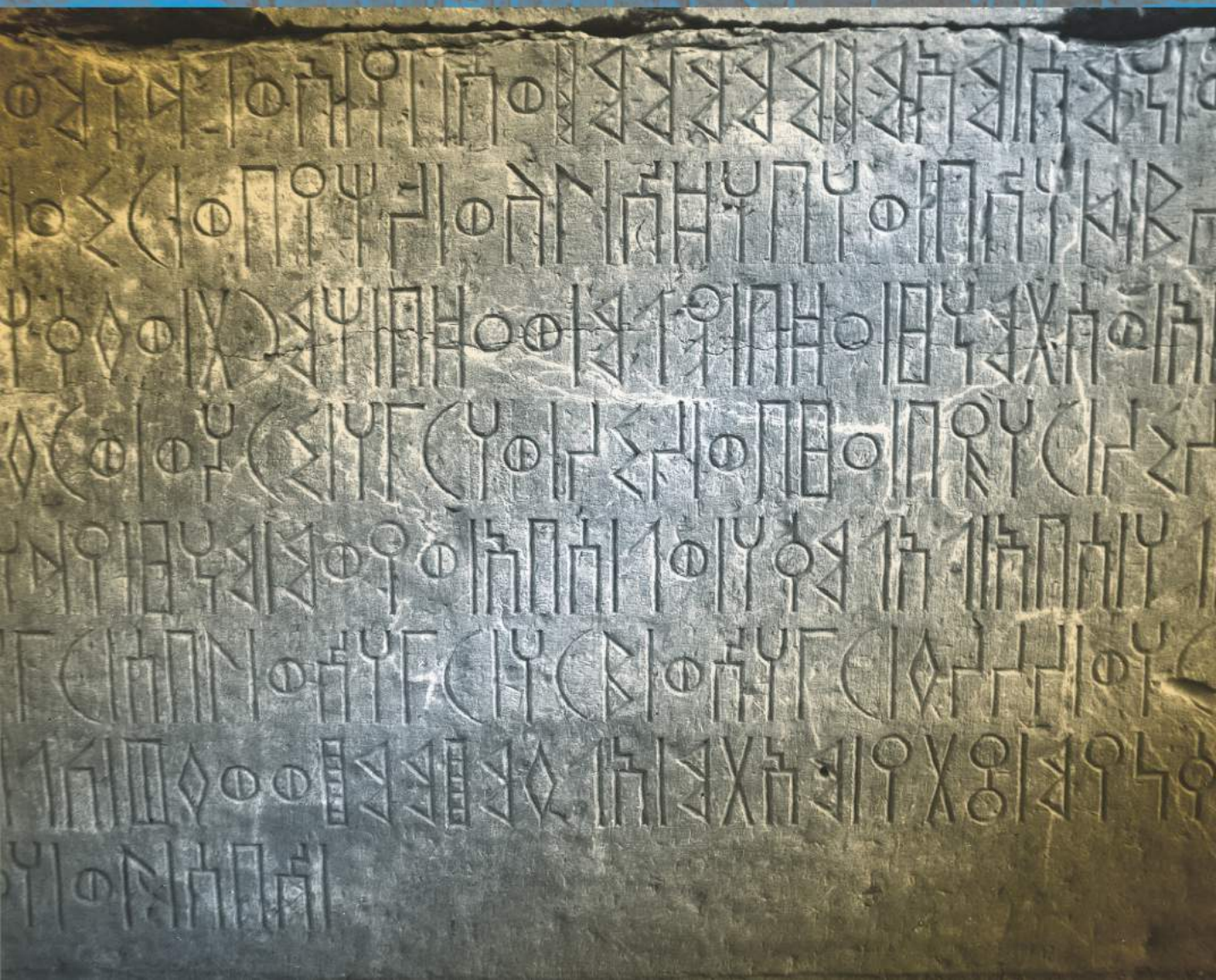


THE CAMBRIDGE  
ENCYCLOPEDIA OF THE  
WORLD'S ANCIENT  
LANGUAGES

EDITED BY ROGER D. WOODARD





THE CAMBRIDGE  
ENCYCLOPEDIA OF THE  
**World's Ancient  
Languages**

Edited by

**ROGER D. WOODARD**

*Andrew V. V. Raymond Professor of the Classics  
Professor of Linguistics  
University of Buffalo (The State University of New York)*



**CAMBRIDGE**  
UNIVERSITY PRESS





Dedicated to the memory of Henry Hoenigswald  
Scholar, gentleman, friend

# CAMBRIDGE

## UNIVERSITY PRESS

University Printing House, Cambridge CB2 8BS, United Kingdom

Cambridge University Press is part of the University of Cambridge.

It furthers the University's mission by disseminating knowledge in the pursuit of education, learning and research at the highest international levels of excellence.

[www.cambridge.org](http://www.cambridge.org)

Information on this title: [www.cambridge.org/9780521562560](http://www.cambridge.org/9780521562560)

© Cambridge University Press 2004

This publication is in copyright. Subject to statutory exception and to the provisions of relevant collective licensing agreements, no reproduction of any part may take place without the written permission of Cambridge University Press.

First published 2004

8th printing 2015

Printed and bound in the United Kingdom by Berforts Information Press Ltd.

*A catalogue record for this publication is available from the British Library*

ISBN 978-0-521-56256-0 Hardback

Cambridge University Press has no responsibility for the persistence or accuracy of URLs for external or third-party internet websites referred to in this publication, and does not guarantee that any content on such websites is, or will remain, accurate or appropriate.

# Contents

<i>List of figures</i>		<i>page</i> vii
<i>List of tables</i>		xi
<i>List of maps</i>		xiv
<i>List of contributors</i>		xv
<i>Preface</i>		xvii
<i>List of abbreviations</i>		xviii
1 Introduction	ROGER D. WOODARD	1
2 Sumerian	PIOTR MICHALOWSKI	19
3 Elamite	MATTHEW W. STOLPER	60
4 Hurrian	GERNOT WILHELM	95
5 Uartian	GERNOT WILHELM	119
6 Afro-Asiatic	JOHN HUEHNERGARD	138
7 Ancient Egyptian and Coptic	ANTONIO LOPRIENO	160
8 Akkadian and Eblaite	JOHN HUEHNERGARD and CHRISTOPHER WOODS	218
9 Ugaritic	DENNIS PARDEE	288
10 Hebrew	P. KYLE MCCARTER, JR.	319
11 Phoenician and Punic	JO ANN HACKETT	365
12 Canaanite dialects	DENNIS PARDEE	386
13 Aramaic	STUART CREASON	391
14 Ge'ez (Aksum)	GENE GRAGG	427
15 Ancient South Arabian	NORBERT NEBES and PETER STEIN	454
16 Ancient North Arabian	M. C. A. MACDONALD	488
17 Indo-European	HENRY M. HOENIGSWALD, ROGER D. WOODARD, and JAMES P. T. CLACKSON	534
18 Hittite	CALVERT WATKINS	551

---

19	Luvian	H. CRAIG MELCHERT	576
20	Palaic	H. CRAIG MELCHERT	585
21	Lycian	H. CRAIG MELCHERT	591
22	Lydian	H. CRAIG MELCHERT	601
23	Carian	H. CRAIG MELCHERT	609
24	Attic Greek	ROGER D. WOODARD	614
25	Greek dialects	ROGER D. WOODARD	650
26	Sanskrit	STEPHANIE W. JAMISON	673
27	Middle Indic	STEPHANIE W. JAMISON	700
28	Old Persian	RÜDIGER SCHMITT	717
29	Avestan	MARK HALE	742
30	Pahlavi	MARK HALE	764
31	Phrygian	CLAUDE BRIXHE	777
32	Latin	JAMES P. T. CLACKSON	789
33	Sabellian languages	REX E. WALLACE	812
34	Venetic	REX E. WALLACE	840
35	Continental Celtic	JOSEPH F. ESKA	857
36	Gothic	JAY H. JASANOFF	881
37	Ancient Nordic	JAN TERJE FAARLUND	907
38	Classical Armenian	JAMES P. T. CLACKSON	922
39	Etruscan	HELMUT RIX	943
40	Early Georgian	KEVIN TUIITE	967
41	Ancient Chinese	ALAIN PEYRAUBE	988
42	Old Tamil	SANFORD B. STEEVER	1015
43	Mayan	VICTORIA R. BRICKER	1041
44	Epi-Olmec	TERRENCE KAUFMAN and JOHN JUSTESON	1071
45	Reconstructed ancient languages	DON RINGE	1112
	<i>Index</i>		1129

# Figures

1.1	Proto-Elamite tablet. Courtesy of Matthew W. Stolper	3
1.2	Indus Valley inscriptions. From Parpola 1994	4
1.3	Cretan Hieroglyphic inscription and portrait stamped on a sealing. From Evans 1909:272	5
1.4	The Phaistos Disk (side A). From Evans 1909:plate XII	5
1.5	The Caslir Situla. From Conway, Whatmough, and Johnson 1933:II, 28	10
7.1	From the cube statue of Senenmut, Berlin Museum	165
7.2	Photograph of Coptic text. All rights reserved, The Metropolitan Museum of Art, Rogers Fund, 1912. (12.180.237)	168
8.1	Sample of Old Babylonian lapidary signs (After Labat 1988)	221
8.2	Comparison of Sumerian, Babylonian, and Assyrian cursive cuneiform signs (After Labat 1988)	222
8.3	Law (§150) from the Code of Hammurabi (turned 90°) with Transliteration, Transcription, and Translation. (Autographed text after Harper, 1904)	225
10.1	The seven full vowels of Tiberian Hebrew	326
11.1	Vowel phonemes of Standard Phoenician	370
14.1	The Ethio-Semitic subfamily	428
14.2	The Semitic language family	429
14.3	Vowels of Ge'ez	437
15.1	The great inscription of Karib'il Watar (c. 685 BC) R 3945: section of lines 11–20	457
16.1	Pre-Islamic Arabia	489
16.2	Examples of the Ancient North Arabian scripts	491
16.3	Letter-forms in the Ancient North Arabian scripts	496
18.1	The seal of Suppiluliumas I	554
24.1	The vowel phonemes of Classical Attic Greek	617
27.1	Part of a scroll manuscript of the Anuvatapta-gatha or “Songs of Lake Anavatapta” in Gandhari Prakrit, c. first century AD. Reproduced by kind permission of the Board of the British Library	701
28.1	Old Persian vowels	724

29.1	Avestan vowels	746
30.1	Pahlavi vowels	767
33.1	South Picene inscription. From A. Marinetti, <i>Le Iscrizioni Sudpicene</i> , p. 219	814
33.2	Oscan inscription. From Yves Duhoux 1988, "A propos des Inscriptions Osques dites Iuvilas et du texte Vetter 94," in Yoël Arbeitman (ed.), <i>A Linguistic Happening in Memory of Ben Schwartz</i> , by permission of Peeters Publishers, Louvain-la-Neuve, 1988	815
34.1	Venetic votive inscriptions. From G. B. Pellegrini and A. L. Prosdocimi, <i>La Lingua Venetica</i> , 1967, pp. 163 (1a), 488 (1b)	841
34.2	Venetic epitaphs. From G. B. Pellegrini and A. L. Prosdocimi, <i>La Lingua Venetica</i> , 1967, pp. 56 (2a), 330 (2b), 195 (2c)	842
34.3	Latino-Venetic inscriptions. From G. B. Pellegrini and A. L. Prosdocimi, <i>La Lingua Venetica</i> , 1967, p. 222	846
36.1	The Germanic languages	881
38.1	The vowel phonemes of Classical Armenian	926
43.1	Classification of Mayan languages. Dotted lines represent less secure groupings. After Campbell and Kaufman (1985:fig. 1)	1042
43.2	Approximate locations of Mayan languages in AD 1550	1043
43.3	Tikal, stela 31, back. After Jones and Satterthwaite (1982:fig. 52b)	1045
43.4	The Leyden Plaque (left = back; right = front). Drawing by Linda Schele	1047
43.5	Phonetic complementation ( <i>a</i> , Leyden Plaque B9. <i>b</i> , Leyden Plaque B10. <i>c</i> , Balakbal, Stela 5 [Justeson and Mathews 1990:fig. 12]. <i>d</i> , Palenque, Foliated Cross Tablet, D17 [Maudslay 1889–1902:IV, plate 82]. <i>e</i> , Tikal, Stela 31, B6 [Jones and Satterthwaite 1982:fig. 52b]. <i>f</i> , Balakbal, Stela 5 [Justeson and Mathews 1990:fig. 12]. <i>g</i> , Tikal, Temple IV, Lintel 3, E3b [Jones and Satterthwaite 1982:fig. 74]. <i>h</i> , Tikal, Stela 31, C13 [Jones and Satterthwaite 1982:fig. 52b]. <i>i</i> , Pomona Panel, L1 [Schele and Miller 1986:fig. III.12]. <i>j</i> , Yaxchilan, Lintel 46, G3 [Graham 1979:101]. <i>k</i> , Yaxchilan, Lintel 31, I4 [Graham 1979:71]. <i>l</i> , Tila A, B5. <i>a</i> and <i>b</i> from a drawing by Linda Schele. <i>l</i> from Thompson 1971, fig. 16 30)	1048
43.6	Phonetization ( <i>a</i> , Tikal, Stela 31, G12–H12. <i>b</i> , Balakbal, Stela 5. <i>c</i> , Caracol, Stela 1, A3–B3. After Justeson and Mathews [1990:fig. 12])	1048
43.7	Syllabic and logosyllabic spellings ( <i>a</i> , Rio Azul, pot, D [Stuart 1988:fig. 2]. <i>b</i> , Piedras Negras Lintel 3, P2. <i>c</i> , Rio Azul, pot, L [Stuart 1988:fig. 2]. <i>d</i> , Tikal, Stela 31, M3 [Jones and Satterthwaite 1982:fig. 52a]. <i>e</i> , Caracol, Stela 16, B18 [Beetz and Satterthwaite 1981:fig. 15]. <i>f</i> , Naranjo, vase, B' [Coe 1973:103 (drawing by Michael D. Coe)]. <i>g</i> , Tikal, Stela 31, L1 [Jones and Satterthwaite 1982:fig. 51a]. <i>h</i> , Tikal, Stela 31, P2 [Jones and Satterthwaite 1982:fig. 52a]. <i>i</i> , Rio Azul, Tomb 12, C [Graham and Mobley 1986:456]. <i>j</i> , Rio Azul, pot, B [Stuart 1988:fig. 2]. <i>k</i> , Yaxchilan, Lintel 2, Q [Graham and von Euw 1977:15]. <i>l</i> , Yaxchilan, Lintel 43, D4 [Graham 1979:95]. <i>m</i> , Kabah, Structure 1 [Proskouriakoff and Thompson 1947:fig. 1g]. <i>n</i> , Copan, Altar S, glyph I [Maudslay 1889–1902:I, plate 94]. <i>o</i> , Copan, Stela A, F4 [Maudslay 1889–1902:I, plate 30]. <i>p</i> , Chichen Itza, Monjas, Lintel 2A, B1 [Graham 1977:269]. <i>b</i> after a drawing by John Montgomery)	1049
43.8	Early Classic syllables	1050



- 43.9 Polyvalence (*a*, Piedras Negras, Lintel 3, D'6. *b*, Yaxchilan, Lintel 37, C6a [Graham 1979:83]. *c*, Houston Lintel, E3. *d*, Piedras Negras, Lintel 12, E1. *e*, Piedras Negras, Throne 1, L. *f*, Yula, Lintel 1, B3. *g*, Piedras Negras, Altar 2 support, G3. *h*, Yaxchilan, Lintel 42, G3 [Graham 1979:93]. *a*, *d*, and *g* after drawings by John Montgomery. *c* after a field sketch by Ian Graham. *e–f* from Thompson 1971, figs. 4 22 and 33 31) 1051
- 43.10 Alternative spellings of *ʔahaw* (*a*, Uaxactun, fresco, glyph 60. *b*, Tikal, Stela 31, D14 [Jones and Satterthwaite 1982:fig. 52b]. *c*, Copan, Stela C, A2b. *d*, Quirigua, Stela D, D14. *e*, Copan, Stela D, A4b. *f*, geometric variant. *g*, Yaxchilan, Lintel 23, O5b [Graham 1982:136]. *h*, Yaxchilan, Hieroglyphic Stairway 3, step IV, B3a [Graham 1982:170]. *i*, Yaxchilan, Lintel 3, J1 [Graham and von Euw 1977:17]. *j*, Piedras Negras, Stela 3, F5a [Marcus 1976:fig. 12]. *k*, Piedras Negras, Throne 1, H'3 [Morley 1937–1938:fig. 111]. *l*, Piedras Negras, Lintel 3, V4. *a*, *c–e* from Thompson 1971, figs. 10 47, 11 24, 33, 34. *l* after a drawing by John Montgomery) 1052
- 43.11 Alternative spellings of *ʔuniw* (*a*, Naranjo, Hieroglyphic Stairway, V1b. *b*, Seibal, Hieroglyphic Stairway, K1a. *c*, Dos Pilas, Stela 8, I13. *d*, Yaxchilan, Hieroglyphic Stairway 3, step I, D1a [Graham 1982:166]. *a* and *b* from Thompson 1971, figs. 18 23, 27. *c* after a drawing by Ian Graham) 1053
- 43.12 Numbers (*a*, Monte Alban, Stela 12 [Marcus 1976:fig. 3]. *b*, La Mojarra, Stela 1, M7. *c*, Palenque, Foliated Cross Tablet, D17 [Maudslay 1889–1902:IV, plate 82]. *d*, Piedras Negras, Stela 3, A7. *e*, Balakbal, Stela 5, D10. *f*, Palenque, Sun Tablet, A6. *g*, Quirigua, F, C8b. *h*, Palenque, Cross Tablet, A5. *i*, Copan, Temple 11, north door, west panel. *j*, Piedras Negras, Stela 3, F4. *k*, Piedras Negras, Lintel 3, G1. *l*, Palenque, Inscriptions Temple, east panel, S9 [Maudslay 1889–1902:IV, plate 60]. *e–i* and *k* from Thompson 1971, figs. 4 16, 24 12, 60, 25 11, 33 21, 29. *b* after a drawing by George Stuart in Winfield Capitaine 1988:15. *d* and *j* after a drawing by John Montgomery) 1054
- 43.13 Landa's "alphabet." After Tozzer (1941:170) 1054
- 43.14 Contraction and cluster reduction (*a*, Palenque, Cross Tablet, S16 [Maudslay 1889–1902:IV, plate 77]. *b*, Tikal, Temple I, Lintel 3, D3–C4 [Maudslay 1889–1902:III, plate 74]. *c*, Copan, Altar Q, A3 [Maudslay 1889–1902:I, plate 93]. *d*, Palenque, Sun Tablet, O9 [Maudslay 1889–1902:IV, plate 89]. *e*, Machaquila, Structure 4, V3 [Graham 1967:fig. 39]. *f*, Uaxactun, Stela 13, A5b [Graham 1986:163]. *g*, Copan, Temple 11, step, glyph 22 [Maudslay 1889–1902:I, plate 8]. *h*, Yaxchilan, Hieroglyphic Stairway 5, glyph 84 [Graham 1982:179]. *i*, Machaquila, Structure 4, V2–V3 [Graham 1967:fig. 39]. *j*, Tikal, Temple IV, Lintel 3, B4 [Jones and Satterthwaite 1982:fig. 74]) 1057
- 43.15 Verbal inflection (*a*, Palenque, Inscriptions Temple, middle panel, H2 [Stuart 1987:fig. 36d]. *b*, Piedras Negras, Lintel 3, J1. *c*, Piedras Negras, Altar support, A2 [Bricker 1986:fig. 148b]. *d*, Palenque, Inscriptions Temple, middle panel, C8 [Robertson 1983:fig. 96]. *e*, Piedras Negras, Throne 1, A'1. *f*, Yaxchilan, Hieroglyphic Stairway 3, step I, tread, D1b [Graham 1982:166]. *g*, Yaxchilan, Hieroglyphic Stairway 3, step I, tread, A2 [Graham 1982:166]. *h*, Palenque, Temple 18, jambs, D18a [Sáenz 1956:fig. 5]. *i*, Piedras Negras, Stela 12, A16a. *j*, Yaxchilan, Lintel 31, I3b [Graham 1979:71]. *k*, unprovenienced ceramic vessel [after a drawing by David Stuart]. *l*, unprovenienced ceramic vessel [Grube 1991:figs. 5i and 9a]. *m*, Deletaille panel, C2 [Ringle 1985:fig. 2].

---

<i>n</i> , Palenque, Foliated Cross Tablet, N7 [Maudslay 1889–1902:IV, plate 82]. <i>o</i> , Dos Pilas, Stela 8, F14. <i>p</i> , Copan, Altar U, K2 [Maudslay 1889–1902:I, plate 98]. <i>b</i> , <i>e</i> , and <i>i</i> after drawings by John Montgomery. <i>o</i> after drawing by Ian Graham)	1061
43.16 Inflection of <i>b'ah</i> ( <i>a</i> , Yaxchilan, Lintel 2, F–G [Graham and von Euw 1977:15]. <i>b</i> , Copan, Stela 10, H3–G4a [Maudslay 1889–1902:I, plate 111]. <i>c</i> , Copan, Stela 6, D8–D9 [Maudslay 1889–1902:I, plate 107])	1063
43.17 Word order ( <i>a</i> , Dos Pilas 8, I5–I6. <i>b</i> , Quirigua, Stela A, B16–B17 [Maudslay 1889–1902:II, plate 7]. <i>c</i> , Quirigua, Structure 1, G–I [Morley 1937–1938:V, plate 175]. <i>d</i> , Yaxchilan, Lintel 46, F3–G3 [Graham 1979:101]. <i>e</i> , Machaquila, Structure 4, F1–F2 [Graham 1967:fig. 39]. <i>a</i> after a drawing by Ian Graham)	1064
44.1 The Mije-Sokean language family	1072
44.2 La Mojarra Stela	1073

# Tables

1.1	A partial inventory of Cypro-Minoan characters	<i>page</i> 6
1.2	Characters of the Meroitic script	7
1.3	Irish Ogham (Craobh-Ruadh); font courtesy of Michael Everson	8
3.1	Middle Elamite and Early Neo-Elamite (before c. 650 BC) syllabic signs: V, CV, VC	68
3.2	Late Neo-Elamite (after c. 650) and Achaemenid Elamite syllabic signs: V, CV, VC (values in parentheses are not attested in Achaemenid)	68
6.1	The consonantal phonemes of Common Semitic	142
6.2	Proto-Semitic personal pronouns	150
7.1	Monoconsonantal hieroglyphic signs	164
7.2	The Coptic alphabet	167
7.3	Personal pronouns in Earlier Egyptian	176
7.4	Earlier Egyptian numerals and their Sahidic Coptic outcome	182
8.1	The consonantal phonemes of Akkadian (post-Old Akkadian period)	230
9.1	The Ugaritic cuneiform consonantal script	291
9.2	The Ugaritic noun: absolute state	296
9.3	The Ugaritic verb: G-stem	306
10.1	The Hebrew alphabet	322
10.2	The Tiberian representation of the principal Hebrew vowels	323
10.3	The consonantal phonemes of Hebrew	325
10.4	The vowel classes of Hebrew verbs	351
11.1	The Phoenician consonantal script	368
11.2	The consonantal phonemes of Standard Phoenician	369
11.3	The enclitic personal pronouns of Standard Phoenician	375
13.1	Aramaic consonantal scripts	393
13.2	Aramaic vowel diacritics	394
13.3	Old Aramaic consonantal phonemes	396
14.1	The Ge'ez syllabary	432
14.2	The labiovelar symbols	432
14.3	The Ethiopic numerals	433

14.4	Semitic consonantal correspondence series	434
14.5	Consonants of Ge'ez	436
14.6	Strong verb-stem paradigms	446
15.1	The Ancient South Arabian consonantal script	456
15.2	The consonantal phonemes of Ancient South Arabian	458
16.1	The consonantal phonemes of Ancient North Arabian	497
16.2	The cardinal numerals of Dadanitic	522
17.1	Proto-Indo-European nominal endings	543
18.1	Old Hittite noun inflection	560
18.2	Neo-Hittite noun inflection	561
18.3	First- and second-person pronouns	562
19.1	Examples of Hieroglyphic Luvian logograms; courtesy of Roland D. Woodard	578
21.1	The Lycian alphabet; courtesy of Roland D. Woodard	593
22.1	The Lydian alphabet; courtesy of Roland D. Woodard	602
23.1	A subset of characters of the Carian alphabet; courtesy of Roland D. Woodard	610
24.1	The Greek alphabet	615
24.2	The consonantal phonemes of Classical Attic Greek	616
24.3	Ablauting noun patterns of PIE	627
25.1	The Linear B script (from O. Dickinson, <i>The Bronze Age</i> , 1994, p. 196)	654
25.2	The Cypriot syllabary	655
25.3	Epichoric Greek alphabets	656
26.1	The Devanāgarī script	675
26.2	The consonantal phonemes of Sanskrit	676
28.1	The Old Persian cuneiform script	719
29.1	The Avestan writing system (from Hoffmann and Forssman 1996; p. 41)	744
29.2	Consonantal sounds of Avestan	746
29.3	First- and second-person personal pronouns of Avestan	755
29.4	Demonstrative pronouns of Avestan	756
30.1	The Pahlavi script	765
30.2	Consonantal sounds of Pahlavi	767
30.3	Pahlavi verb inflection	772
32.1	The Archaic Latin alphabet	791
32.2	The consonantal phonemes of Latin	792
32.3	Latin nominal paradigms	797
33.1	National Oscan alphabet, c. 250 BC	816
33.2	Umbrian alphabet, Iguvine Tablets I–Vb7, c. 250 BC	817
33.3	The consonantal phonemes of Oscan	818
33.4	The consonantal phonemes of Umbrian	820
33.5	Sabellian noun stems	823
33.6	Sabellian personal endings	827

34.1	Venetic <i>alphabet princeps</i> (c. 550 BC). From G. Fogolari and A. L. Prosdocimi, <i>I Veneti Antichi</i> , 1988, p. 333	844
34.2	Spelling of Venetic dental stops	845
34.3	The consonantal phonemes of Venetic	846
35.1	The Celtic adaptation of the Iberian script; courtesy of Roland D. Woodard	858
35.2	The Lugano script; courtesy of Roland D. Woodard	860
35.3	Hispano-Celtic nominal inflection	865
35.4	Lepontic nominal inflection	866
35.5	Gaulish nominal inflection	867
35.6	Present endings of Continental Celtic	870
35.7	Preterite endings of Continental Celtic	870
36.1	Wulfila's alphabet	883
36.2	Gothic nominal stems	894
36.3	Gothic strong and weak verb paradigms	900
37.1	The Northwest Germanic futhark	909
37.2	Ancient Nordic nominal stems	913
38.1	The Armenian alphabet	924
38.2	The consonantal phonemes of Classical Armenian	925
39.1	The Etruscan alphabet of archaic inscriptions	945
40.1	The Early Georgian Asomtavruli script with numerical values	969
40.2	The Early Georgian consonants	970
40.3	Declension of <i>mojγw[a]r</i> – “leader”	971
40.4	Early Georgian verb classes (“conjugations”)	976
40.5	Case shift	977
40.6	Set S (morphological subject) suffix groups	981
40.7	Early Georgian verb paradigms	982
41.1	Development of the Chinese script. From Norman, <i>Chinese</i> , 1988, p. 66	991
42.1	The Tamil syllabary	1017
43.1	The consonantal phonemes of Mayan	1055
44.1	Sources and characteristics of Epi-Olmec and related texts	1074
44.2	Epi-Olmec syllabary	1076
44.3	Comparison of Sokean linguistic features	1104

# Maps

1	The Ancient languages of Northeastern Africa and Arabia	<i>between pages</i> 453–454
2	The Ancient languages of Anatolia and surrounding regions (languages which are first attested in the first millennium BC or later are marked in italics)	575–576
3	The Greek dialects of the first millennium BC and neighboring languages	649–650
4	The Ancient languages of Italy and surrounding regions (for the Greek dialects of Italy and Sicily, see Map 3)	839–840
5	The distribution of attested Ancient languages within Asia (for the languages of Asia Minor, see Map 2)	987–988



# Contributors

VICTORIA R. BRICKER	Tulane University
CLAUDE BRIXHE	Université de Nancy
JAMES P. T. CLACKSON	Cambridge University
STUART CREASON	University of Chicago
JOSEPH F. ESKA	Virginia Polytechnic Institute and State University
JAN TERJE FAARLUND	Universitetet i Oslo
GENE GRAGG	University of Chicago
JO ANN HACKETT	Harvard University
MARK HALE	Concordia University, Montreal
HENRY M. HOENIGSWALD	University of Pennsylvania
JOHN HUEHNERGARD	Harvard University
STEPHANIE W. JAMISON	Harvard University
JAY H. JASANOFF	Harvard University
JOHN JUSTESON	State University of New York at Albany
TERRENCE KAUFMAN	University of Pittsburgh
ANTONIO LOPRIENO	Universitaet Basel
M. C. A. MACDONALD	Oxford University
P. KYLE MCCARTER, JR.	The Johns Hopkins University
H. CRAIG MELCHERT	University of North Carolina
PIOTR MICHALOWSKI	University of Michigan
NORBERT NEBES	Friedrich-Schiller-Universität Jena
DENNIS PARDEE	University of Chicago
ALAIN PEYRAUBE	Université de Paris
DON RINGE	University of Pennsylvania

HELMUT RIX	Albert-Ludwigs-Universität Freiburg
RÜDIGER SCHMITT	Universität des Saarlandes
SANFORD B. STEEVER	New Canaan, Connecticut
PETER STEIN	Friedrich-Schiller-Universität Jena
MATTHEW W. STOLPER	University of Chicago
KEVIN TUIITE	Université de Montréal
REX E. WALLACE	University of Massachusetts at Amherst
CALVERT WATKINS	Harvard University
GERNOT WILHELM	Julius-Maximilians-Universität
ROGER D. WOODARD	University of Buffalo (The State University of New York)
CHRISTOPHER WOODS	Harvard University

## Preface

In the following pages, the reader will discover what is, in effect, a linguistic description of all known ancient languages. Never before in the history of language study has such a collection appeared within the covers of a single work. This volume brings to student and to scholar convenient, systematic presentations of grammars which, in the best of cases, were heretofore accessible only by consulting multiple sources, and which in all too many instances could only be retrieved from scattered, out-of-the-way, disparate treatments. For some languages, the only existing comprehensive grammatical description is to be found herein.

This work has come to fruition through the efforts and encouragement of many, to all of whom the editor wishes to express his heartfelt gratitude. To attempt to list all – colleagues, students, friends – would, however, certainly result in the unintentional and unhappy neglect of some, and so only a much more modest attempt at acknowledgments will be made. Among those to whom special thanks are due are first and foremost the contributors to this volume, scholars who have devoted their lives to the study of the languages of ancient humanity, without whose expertise and dedication this work would still be only a *desideratum*. Very special thanks also go to Dr. Kate Brett of Cambridge University Press for her professionalism, her wise and expert guidance, and her unending patience, also to her predecessor, Judith Ayling, for permitting me to persuade her of the project's importance. I cannot neglect mentioning my former colleague, Professor Bernard Comrie, now of the Max Planck Institute, for his unflagging friendship and support. Kudos to those who masterfully translated the chapters that were written in languages other than English: Karine Megardoomian for Phrygian, Dr. Margaret Whatmough for Etruscan, Professor John Huehnergard for Ancient South Arabian. Last of all, but not least of all, I wish to thank Katherine and Paul – my inspiration, my joy.

Roger D. Woodard  
Christmas Eve 2002

# Abbreviations

Any abbreviation that deviates from the form given below is noted within the text of the individual chapter or within a chapter-specific list.

## Linguistic terms

abl.	ablative	dir.	directive
abs.	absolutive	dir. obj.	direct object
acc.	accusative	disj.	disjunctive
act.	active	du.	dual
adj.	adjective	dur.	durative
adv.	adverb (adverbial)	emph.-pcl.	emphatic particle
all.	allative	encl.	enclitic
anim.	animate	eq.	equative
aor.	aorist	erg.	ergative
art.	article	ext.	extended
asp.	aspirated	fem.	feminine
aux.	auxiliary (verb)	final-pcl.	final-particle
caus.	causative	fut.	future
cl.	clause	gdve.	gerundive
coll.	collective	gen.	genitive
com.	common	ger.	gerund
comp.	comparative	impf.	imperfect
comt.	comitative	impftv.	imperfective
conj.	conjunction	impv.	imperative
conjv.	conjunctive	inan.	inanimate
conn.	connective	inc.	inclusive
cons.	consonant	indef. art.	indefinite article
constr.	construct (state)	indet.	indeterminate
cont.	continuant	indic.	indicative
cop.	copula	inf.	infinitive
dat.	dative	instr.	instrumental
def. art.	definite article	interr.	interrogative
dem.	demonstrative	intr.	intransitive
det.	determinate	iter.	iterative
detv.	determinative	juss.	jussive
dial.	dialect	loc.	locative
		mediopass.	mediopassive
		mid.	middle

N.	noun
neg.	negative
neut.	neuter
nom.	nominative
NP	noun phrase
num.	number
obj.	object
obl.	oblique
opt.	optative
part.	participle
pass.	passive
pcl.	particle
per.	person
perf.	perfect
perfv.	perfective
perfvz.	perfectivizer
pert.	pertinentive
pl.	plural
pluperf.	pluperfect
poss. suff.	possessive suffix
postp.	postposition
PP	prepositional phrase
prec.	precative
preC.	preconsonantal
pref.	prefix
prep.	preposition
pres.	present
pret.	preterite
preV.	prevocalic
pro.	pronoun
prosp.	prospective
quot.	quotative particle
refl.	reflexive
rel. pro.	relative (pronoun)
rel./connec.	relative/connective
sg.	singular
soc.	sociative case
SOV	Subject–Object–Verb (word order)
spec.	specifier
stat.	stative
subj.	subject
subjunc.	subjunctive
subord.	subordinate/subordinator/ subordination marker
subord.-pcl.	subordinating particle
suff.	suffix
splv.	superlative
s.v.	<i>sub voce</i>

top.	topicalizer
tr.	transitive
V.	verb
var.	variant
vent.	ventive
voc.	vocative
vow.	vowel
VP	verbal phrase

### Languages

Akk.	Akkadian
Ar.	Arabic
Ass.	Assyrian
Av.	Avestan
Bab.	Babylonian
Cis. Gaul.	Cisalpine Gaulish
Eg.	Egyptian (Old, Late, Earlier)
Eng.	English
Etr.	Etruscan
Gk.	Greek
Gmc.	Germanic
Go.	Gothic
Hispr.-Celt.	Hispano-Celtic
Hitt.	Hittite
IE	Indo-European
Lat.	Latin
Lep.	Lepontic
Luv.	Luvian
Lyc.	Lycian
MA	Middle Assyrian
MB	Middle Babylonian
NA	Neo-Assyrian
NB	Neo-Babylonian
OA	Old Assyrian
O. Akk.	Old Akkadian
O. Av.	Old Avestan
OB	Old Babylonian
OHG	Old High German
OP	Old Persian
PG	Proto-Greek
PGmc.	Proto-Germanic
PIE	Proto-Indo-European
PIIr.	Proto-Indo-Iranian
PIr.	Proto-Iranian
PMS	Proto-Mije-Sokean
PS	Proto-Semitic
PSo.	Proto-Sokean
SB	Standard Babylonian

Skt. Sanskrit  
 Sum. Sumerian  
 Y. Av. Young Avestan

**Other**

abbr. abbreviation

dict. dictionary  
 intro. introduction  
 lit. literally  
 NA not applicable  
 NS new series  
 trad. traditional  
 translit. transliteration



# Introduction

ROGER D. WOODARD

## 1.1 Preliminary remarks

What makes a language ancient? The term conjures up images, often romantic, of archeologists feverishly copying hieroglyphs by torchlight in a freshly discovered burial chamber; of philologists dangling over a precipice in some remote corner of the earth, taking impressions of an inscription carved in a cliff-face; of a solitary scholar working far into the night, puzzling out some ancient secret, long forgotten by humankind, from a brittle-leaved manuscript or patina-encrusted tablet. The allure is undeniable, and the literary and film worlds have made full use of it.

An ancient language is indeed a thing of wonder – but so is every other language, all remarkable systems of conveying thoughts and ideas across time and space. And ancient languages, as far back as the very earliest attested, operate just like those to which the linguist has more immediate access, all with the same familiar elements – phonological, morphological, syntactic – and no perceptible vestiges of Neanderthal oddities. If there was a time when human language was characterized by features and strategies fundamentally unlike those we presently know, it was a time prior to the development of any attested or reconstructed language of antiquity. Perhaps, then, what makes an ancient language different is our awareness that it has outlived those for whom it was an intimate element of the psyche, not so unlike those rays of light now reaching our eyes that were emitted by their long-extinguished source when dinosaurs still roamed across the earth (or earlier) – both phantasms of energy flying to our senses from distant sources, long gone out.

That being said, and rightly enough, we must return to the question of what counts as an ancient language. As *ancient* the editor chose the upward delimitation of the fifth century AD. This *terminus ante quem* is one which is admittedly “traditional”; the fifth is the century of the fall of the western Roman Empire (AD 476), a benchmark which has been commonly (though certainly not unanimously) identified as marking the end of the historical period of *antiquity*. Any such chronological demarcation is of necessity arbitrary – far too arbitrary – as linguists accustomed to making such diachronic distinctions as *Old English*, *Middle English*, *Modern English* or *Old Hittite*, *Middle Hittite*, *Neo-Hittite* are keenly aware. Linguistic divisions of this sort are commonly based upon significant political events and clearly perceptible cultural shifts rather than upon language phenomena (though they are surely not without linguistic import as every historical linguist knows). The choice of the boundary in the present concern – the ancient-language boundary – is, likewise (as has already been confessed), not mandated by linguistic features and characteristics of the languages concerned.

However, this arbitrary choice, establishing a *terminus ante quem* of the fifth century, is somewhat buttressed by quite pragmatic linguistic considerations (themselves consequent

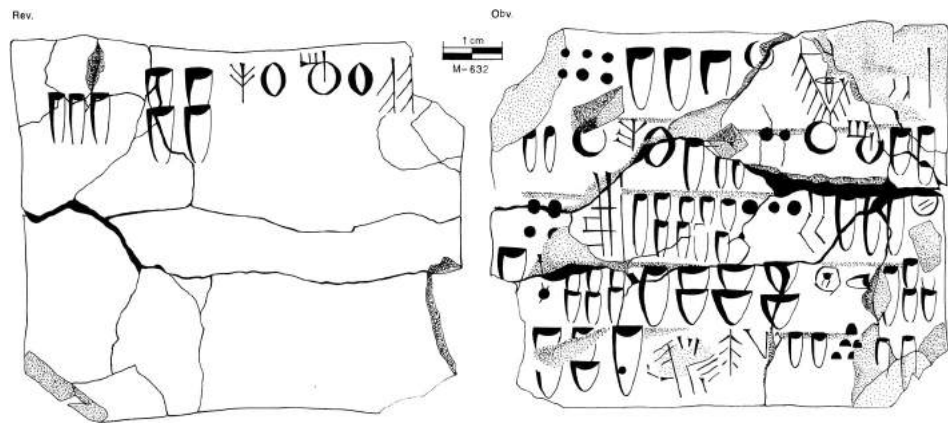
to the whim of historical accident), namely the co-occurrence of a watershed in language documentation. Several early languages first make a significant appearance in the historical record in the fourth/fifth century: thus, Gothic (fourth century; see Ch. 36), Ge'ez (fourth/fifth century; see Ch. 14, §1.3.1), Classical Armenian (fifth century; see Ch. 38), Early Old Georgian (fifth century; see Ch. 40). What newly comes into clear light in the sixth century is a bit more meager – Tocharian and perhaps the very earliest Old Kannada and Old Telegu from the end of the century. Moreover, the dating of these languages to the sixth century cannot be made precisely (not to suggest this is an especially unusual state of affairs) and it is equally possible that the earliest attestation of all three should be dated to the seventh century. Beginning with the seventh century the pace of language attestation begins to accelerate, with languages documented such as Old English, Old Khmer, and Classical Arabic (though a few earlier inscriptions preserving a “transitional” form of Arabic are known; see Ch. 16, §1.1.1). The ensuing centuries bring an avalanche of medieval European languages and their Asian contemporaries into view. Aside from the matter of a culturally dependent analytic scheme of historical periodization, there are thus considerations of language history that motivate the upper boundary of the fifth century.

On the other hand, identifying a *terminus post quem* for the inclusion of a language in the present volume was a completely straightforward and noncontroversial procedure. The low boundary is determined by the appearance of writing in human society, a graphic means for recording human speech. A system of writing appears to have been first developed by the Sumerians of southern Mesopotamia in the late fourth millennium BC (see Ch. 2, §§1.2; 2). Not much later (beginning in about 3100 BC), a people of ancient Iran began to record their still undeciphered language of Proto-Elamite on clay tablets (see Ch. 3, §2.1). From roughly the same period, the Egyptian hieroglyphic writing system emerges in the historical record (see Ch. 7, §2). Hence, Sumerian and Egyptian are the earliest attested, understood languages and, *ipso facto*, the earliest languages treated in this volume (on the problem of undeciphered languages like Proto-Elamite, see §1.2).

It is conjectured that humans have been speaking and understanding language for at least 100,000 years. If in the great gulf of time which separates the advent of language and the appearance of Sumerian, Proto-Elamite, and Egyptian societies, there were any people giving written expression to their spoken language, all evidence of such records and the language or languages they record has fallen victim to the decay of time. Or the evidence has at least eluded the archeologists.

Though no human language is documented prior to the late fourth millennium BC, it is still possible in certain instances to recover the linguistic system of a deeply archaic, preliterate people, using the remarkable methodology commonly dubbed the *comparative method of historical linguistics* (see Ch. 45). The development of the comparative method and the discovery of the linguistic principles which make it possible was one of the greatest, if perhaps one of the less recognized, scientific achievements of the nineteenth century. Following upon the pioneering efforts of philologists such as the English jurist, Sir William Jones, scholars like Rasmus Rask, Franz Bopp, August Schleicher, Jacob Grimm, Karl Verner, Karl Brugmann, and Hermann Osthoff, among still others (on whom, see, *inter alia*, Mallory 1989:9–23, Lehmann 1967), developed the comparative method and applied it in the reconstruction of Proto-Indo-European, the parent language of the Indo-European language family. Though spoken between the fifth and third millennia BC and nowhere attested by written record, the grammar and lexicon of the language are well known through their reconstruction (see Ch. 17).

The comparative method has likewise been used to reconstruct the parent of the Semitic languages, Proto-Semitic (see Ch. 6, §§2–3), spoken sometime prior to the third



**Figure 1.1** Proto-Elamite tablet

millennium BC. The reconstruction of Proto-Semitic's own parent language, Proto-Afro-Asiatic, from which are also descended Egyptian, Berber, Cushitic, Chadic, and perhaps still other languages is ongoing (see Ch. 6, §1).

With the exception of only a few languages of Mesoamerica, the native American languages of antiquity are known solely through reconstruction. Those exceptions are provided by the Mayan languages (see Ch. 43, §1) and Epi-Olmec (see Ch. 44, §§1–2); though the Zapotec language is perhaps attested as early as about the sixth/fifth century BC, the inscriptional evidence is very meager and the language is poorly understood at present (see below, §1.3.8).

Not every language which is attested in the period that extends from the beginning of written records to the fifth century AD is treated in this volume. There are generally two reasons for this exclusion: (i) the written remains of the language can be read (to a greater or lesser extent), but the evidence of the language which is provided by these records is sufficiently meager to limit significantly a knowledge of the language and, consequently, to proscribe any attempt to offer a meaningful grammatical description of it; and (ii) the written remains of the language cannot be understood – in other words, the recorded language has not yet been deciphered. Languages falling into the latter category will be addressed first, and then those of the former.

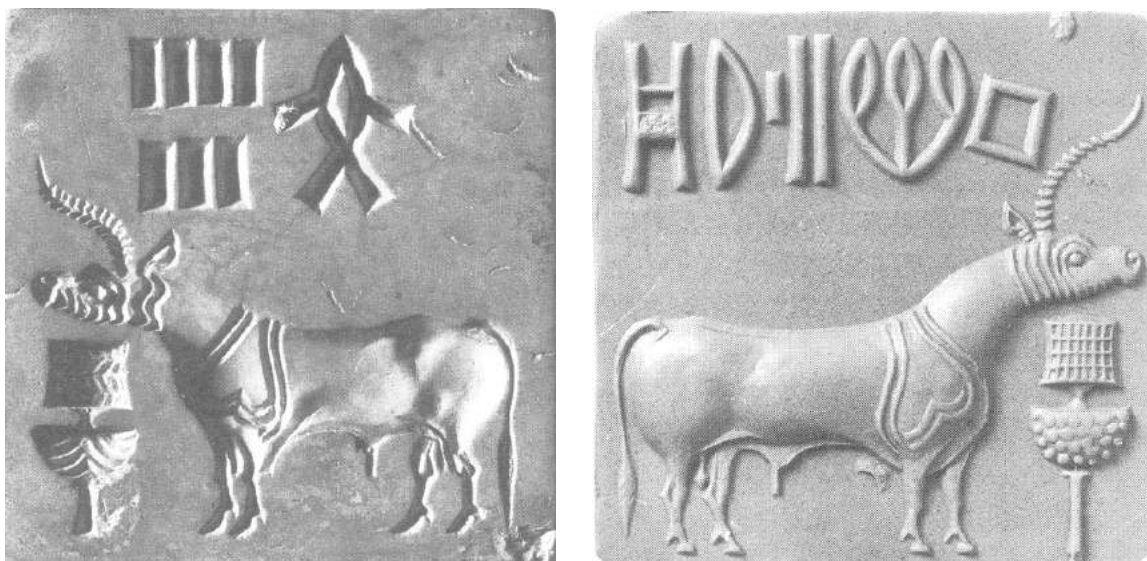
## 1.2 Undeciphered languages

### 1.2.1 Undeciphered Elamite

The earliest attested undeciphered script (late fourth millennium BC) is the one called *Proto-Elamite*. Tablets bearing this script have been recovered in large numbers from the same Iranian region in which ancient Elamite is attested, hence the name Proto-Elamite (see Fig. 1.1). Far fewer in number are the inscriptions from the same general area which are written in the script called *Linear Elamite*. For discussion of each of these, see Chapter 3, §§2.1–2.2 (see also Englund 1996).

### 1.2.2 The Indus Valley language

In the middle of the third millennium BC, writing emerges in the archeological record of the Harappan culture of the Indus Valley. The characters of this Indus Valley script (see Fig. 1.2) are of a well-developed, somewhat conventionalized pictographic nature at the earliest



**Figure 1.2** Indus Valley inscriptions

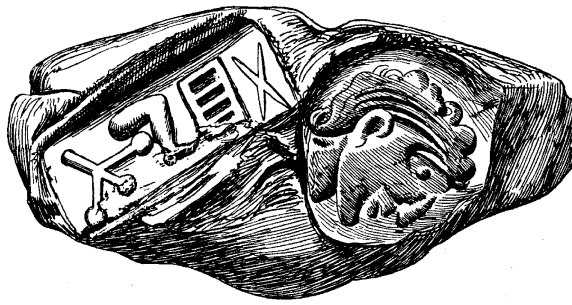
phase of the script's attestation (possibly suggesting some earlier unattested developmental stage). The number of characters identified likely reveals that the script operates with both logograms (symbols representing entire words) and syllabograms (phonetic symbols having the value of a syllable). Lying behind the Indus Valley script may well be a Dravidian language (see Ch. 42, §1) or possibly an early form of Indo-Aryan (see Ch. 26, §1). On the Indus Valley script and its attempted decipherment, see Parpola 1996 and 1994.

### 1.2.3 Cretan languages

Examples of three distinct, undeciphered scripts have survived in the remains of the Minoan civilization of ancient Crete. The oldest of these is called *Cretan Hieroglyphic* or *Cretan Pictographic* (see Fig. 1.3) and its use is dated to the period 2000–1600 BC, seal stones providing the bulk of examples. The pictographic symbols making up the script probably have a syllabic value.

The second of the undeciphered Cretan scripts is known from only a single document, the *Phaistos Disk* (dated to about 1700 BC; see Fig. 1.4). The disk has been the object of repeated attempts at decipherment since its discovery in the early twentieth century. While success has often been claimed, none of the proposed decipherments carries conviction.

*Linear A*, the third of the Minoan scripts, is the best represented of the three. Dating from about the mid-nineteenth to mid-fifteenth centuries BC, Linear A documents partially overlap chronologically with those written in Cretan Hieroglyphic, though in terms of historical development, the former may trace its origins to the latter. Linear A, in turn, appears to be the source of the Mycenaean Greek script, Linear B (see Ch. 25, §§1.1; 1.2; 2.1), though a simple direct linear descent is not probable. Of the three Minoan scripts, Linear A holds the greatest hope for decipherment. Recent work by Brown (1990) and Finkelberg (1990–1991) has taken up a notion proposed by Palmer in the middle of the twentieth century (e.g., Palmer 1968) which would identify the Linear A language as a member of the Anatolian subfamily of Indo-European. On the Cretan scripts see, *inter alia*, Chadwick 1990; Palaima 1988; Woodard 1997.



**Figure 1.3** Cretan  
Hieroglyphic inscription  
and portrait stamped on  
a sealing



**Figure 1.4** The Phaistos  
Disk (side A)

Mention should also be made of the undeciphered language called *Eteo-Cretan*. Much later than the three Bronze Age Minoan scripts, Eteo-Cretan is preserved in inscriptions written in the Greek alphabet. On Eteo-Cretan, see Duhoux 1982.

#### 1.2.4 Cypriot languages

Prior to the emergence of Greek writing on Cyprus, attested by about the middle of the eleventh century BC (and the somewhat later appearance of Phoenician; see Ch. 11, §1.2; Ch. 24, §2), the island was inhabited by a people, or by groups of people, who were recording their speech in the undeciphered set of scripts called *Cypro-Minoan* (see Table 1.1). As the name suggests, these Cypriot writing systems appear to have their origin in a writing system of Minoan Crete, Linear A being the likely candidate. *Archaic Cypro-Minoan* is the name given to the script found on only a single inscription, dated to about 1500 BC.

**Table 1.1 A partial inventory of Cypro-Minoan characters**

I	𐀀	𐀁	𐀂	𐀃	𐀄
𐀅	𐀆	𐀇	𐀈	𐀉	𐀊
𐀋	𐀌	𐀍	𐀎	𐀏	
𐀐	𐀑	𐀒	𐀓	𐀔	
𐀕	𐀖	𐀗	𐀘	𐀙	
𐀚	𐀛	𐀜	𐀝	𐀞	
𐀟	𐀠	𐀡	𐀢	𐀣	
𐀤	𐀥	𐀦	𐀧	𐀨	
𐀩	𐀪	𐀫	𐀬	𐀭	
𐀮	𐀯	𐀰	𐀱	𐀲	
𐀳	𐀴	𐀵	𐀶	𐀷	

This script has been analyzed as the likely ancestor of the more widely attested *Cypro-Minoan 1*, found in use between approximately the late sixteenth and twelfth centuries BC. A distinct script, *Cypro-Minoan 2*, has been found on thirteenth-century documents from the site of Enkomi. Yet a third, *Cypro-Minoan 3*, dating also to the thirteenth century BC, has turned up not on Cyprus but in the remains of the ancient Syrian city of Ugarit (see Ch. 9, §1; on the Cypro-Minoan scripts, see especially Masson 1974, 1977; Palaima 1989).

Cypro-Minoan 1 appears to have provided the graphic model for the Greek syllabary of Cyprus (see Ch. 25, §2.2). This Greek syllabic script was in turn not only used for writing Greek but also adopted for some other language of Cyprus, as yet undeciphered, dubbed *Eteo-Cypriot*. The Eteo-Cypriot inscriptions are commonly regarded as the documentary remains of an indigenous people of Cyprus who had withstood assimilation to the communities of Greek and Phoenician settlers. After Greek and Phoenician settlement of Cyprus, Eteo-Cypriots appear to have concentrated particularly in the area of Amathus (on the Eteo-Cypriot inscriptions, see Masson 1983:85–87).

### 1.2.5 Byblic language

As well as the various Canaanite and Aramaic scripts and languages preserved in the archaeological remains of Syria-Palestine that are treated in this volume (see Chs. 6, 9–13), there is an additional script, attested by a small number of inscriptions, which is commonly (though not universally) regarded as undeciphered. In the course of his excavations at the site of the ancient city of Byblos (Biblical Gebal) on the coast of the modern state of Lebanon, the French archeologist Maurice Dunand unearthed inscriptions, on bronze and stone, executed in a previously unknown script. Many of the symbols are of a hieroglyphic nature, some apparently descended from or inspired by characters of the Egyptian hieroglyphic script; the Byblian script thus bears the tag *Pseudo-Hieroglyphic*, or, less commonly, *Proto-Byblic*. The script, judging by the number of identified symbols (114 by Dunand's analysis), is likely syllabic. As early as 1946 (a year after Dunand's publication of the inscriptions), the decipherment of Byblian Pseudo-Hieroglyphic was announced by a distinguished French philologist, Edouard Dhorme, who read the language of the script as Phoenician. Dhorme's proposed decipherment and others which have followed (see Daniels 1996:29–30 for discussion of subsequent attempts) have not been received with confidence and the script and its language still reside in the undeciphered column.



**Table 1.2 Characters of the Meroitic script**

Character			Character		
Hieroglyphic	Cursive	Transcription	Hieroglyphic	Cursive	Transcription
	ḡz	a		ḡ	l
	ḡ	e		ḡ	h
	ḡ	i		ḡ	s
	/	o		ḡ	se
	///	y		ḡ	k
	ḡ	w		ḡ	q
	v	b		ḡ	t
	ḡ	p		ḡ	te
	ḡ	m		ḡ	to
	ḡ	n		ḡ	d
	ḡ	ne		ḡ	word-divider
	ḡ	r			

### 1.2.6 North African languages

In the ancient Nubian city of Meroë (in the north of modern Sudan – the great city which Herodotus calls the μητρόπολις τῶν ἄλλων Αἰθιοπῶν, “the capital of all Ethiopia”), the Egyptian scripts must have been long known and utilized. By the third century BC, however, with the rise of the Meroitic kingdom, a native writing system appeared and continued in use for recording Meroitic language until the fourth century AD. Two varieties of the script are known: a hieroglyphic script based on Egyptian Hieroglyphic, and a cursive form based on Egyptian Demotic (see Ch. 7, §2.1).

The phonetic values of the symbols of the Meroitic writing system, unlike those of many of the undeciphered languages discussed thus far, have been purportedly identified. The majority of symbols have been assigned the value of a single consonant or vowel sound (i.e., the script is analyzed as fundamentally alphabetic), with a small set of syllabic CV (consonant + vowel) symbols filling out the inventory of characters (compare Ugaritic’s consonantal script, supplemented by three CV characters; see Ch. 9, §2.2); see Table 1.2. While Meroitic texts can thus be given a phonetic reading, the language uttered in such a reading cannot be understood with the exception of a very few words, chiefly proper nouns. On the Meroitic script and language, see Wenig 1982, Griffith 1911, 1912.

For the *Ancient Lybian* or *Numidian* script, or scripts, which have been read as recording archaic Berber, but which some would regard as undeciphered, see Chapter 6, §1.1.3 with references. On the partially deciphered *Proto-Sinaitic* inscriptions, see Chapter 12, §2.2.

### 1.2.7 European languages

From Portugal and Spain come ancient inscriptions recorded in those scripts called *Iberian*, broadly divided into two groups, Northeast and South Iberian. The latter group includes the variety of the script called *Turdetan*, after the ancient Turdetanians, of whom the Greek geographer Strabo wrote: “These are counted the wisest people among the Iberians; they write with an alphabet and possess prose works and poetry of ancient heritage, and laws

**Table 1.3 Irish Ogham (Craobh-Ruadh); font courtesy of Michael Everson**

Symbol	Transcription	Name	Symbol	Transcription	Name
┐	b	beithe	┌	h	úath
└	l	luis	┘	d	dair
┘	f	fern	┐	t	tinne
┘	s	sail	┘	c	coll
┘	n	nin	┘	q	ceirt
┐	m	muin	┐	a	ailm
┐	g	gort	┐	o	onn
┐	ng	gétal	┐	u	úr
┐	z	straif	┐	e	edad
┐	r	ruis	┐	i	idad
✕	ea	ébad	◊	oi	ór
✕	ia	iphín	◊	ui	uilen
⌘	ae	emancholl			

composed in meter, six thousand years old, so they say” (*Geography* 3.1.6). One form of the Northeast Iberian writing system was adopted by speakers of Celtic for recording their own language (*Hispano-Celtic* or *Celtiberian*; see Ch. 35, especially §2.1), and these Celtic documents are interpretable (for the language, see Ch. 35, especially §§3.1; 3.4; 4.2.1.1; 4.3.6; 5.1). However, the Iberian scripts were used principally for a language or languages which are not understood, in spite of the fact that there also occur Iberian-language (*Old Hispanic*) inscriptions written with the Greek and Roman alphabets, and even bilingual texts. On the Iberian scripts and language(s) see, *inter alia*, Untermann 1975, 1980, 1990, 1997; Swiggers 1996; Diringer 1968:193–195.

While the South Picene language of eastern coastal Italy appears to be demonstrably Indo-European (belonging to the Sabellian branch of Italic; see Ch. 33), the genetic affiliation of its meagerly attested northern neighbor, North Picene, remains uncertain (though the two were formerly lumped together under the name *East Italic* or *Old Sabellian*). Though completely readable (being written in an Etruscan-based alphabet), North Picene remains largely impenetrable, in spite of the fact that a Latin-North Picene bilingual exists (a brief inscription, the identity of the non-Latin portion of which has been disputed). For an examination toward a tentative translation of the long North Picene inscription, the *Novilara Stele*, see Poultny 1979 (providing a summary of earlier attempts at interpretation).

The documentation of Insular Celtic – the Celtic languages of Ireland and Britain – (as opposed to Continental Celtic; see Ch. 35) which has survived from antiquity is very meager indeed, and is limited to Irish. The script used in recording this early Irish is the unusual alphabetic system called Ogham (see Table 1.3); most of its characters consist of slashing lines, longer and shorter (notches being used at times for vowel characters), giving the impression that it was originally designed to be “written” by means of an ax or some similar sharp instrument, with wood serving as a medium. The Ogham inscriptions, which date as early as the fourth century AD (and perhaps as early as the second century), can be read (owing to our knowledge of later Irish) but consist largely of personal names and provide little data on which can be constructed a linguistic description of Ogham Irish. For such descriptions of Insular Celtic, the linguist must await the appearance of Old Irish and Old

Welsh manuscripts in about the eighth century AD (and hence Ogham Irish is not treated in the present volume).

There is, however, a second ancient language of Britain which is written with a variety of Ogham, the language of *Pictish*. The Picts, who receive their name from Latin *Picti* “painted ones” (presumably referring to the practice of tattooing, though other etymologies have been proposed), inhabited portions of modern Scotland, along with the Scots, a Celtic people of Irish origin. A much broader, earlier distribution of the Picts has also been claimed. The Picts are known for their production of stone monuments on which are engraved intriguing images of animals and other designs, at times accompanied by Ogham inscriptions. The language of the Pictish Ogham inscriptions is not understood; it is not Celtic and probably not Indo-European. On the Pictish language, see Jackson 1980; for Ogham generally, see McMannus 1991.

### 1.3 Insufficiently attested languages

The differences between the languages of this group and the preceding are in some cases only a matter of degree (rather than one of kind), and not sharply one of intelligible versus unintelligible.

Among those recorded languages of antiquity which can be read and understood to an appreciable degree but which were judged too meagerly attested to be included in the present volume of grammatical descriptions, several are languages which were spoken on the Italian peninsula and the neighboring island of Sicily.

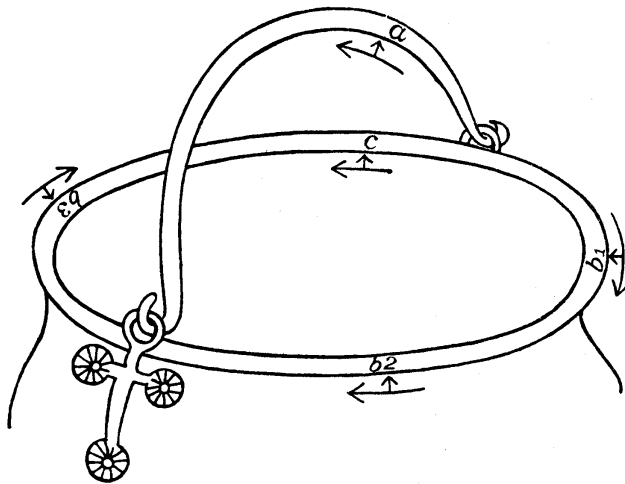
#### 1.3.1 Sikel

From Sicily come several inscriptions written in a language which appears to be Indo-European; a number of glosses are claimed as well (see Conway, Whatmough, and Johnson 1933 II:449–458; on Sikel generally, see Pulgram 1978:71–73 with references). The name assigned to the language, Sikel or Siculan, is that given by Greek colonists to the native peoples of Sicily whom they there encountered in the eighth century BC. Little is known about the ethnicity of these Siceli. The form *esti* occurs in Sikel, seemingly the archetypal Indo-European “(s)he is.” Interpretations of other inscriptional forms show considerable variation. Tradition held that the Siceli had migrated to Sicily from the Italian peninsula: thus, Varro (*On the Latin Language* 5.101) writes that they came from Rome; Diodorus Siculus (*Library of History* 5.6.3–4) records that the Siceli had come from Italy and settled in the region of Sicily formerly occupied by a people called the Sicani. On the basis of the available linguistic evidence, however, Sikel cannot be demonstrated to be a member of the Italic subfamily of Indo-European (see Ch. 32, §1).

On the inscriptional fragments from western Sicily identified as *Elymian*, see Cowgill and Mayrhofer 1986:58 with references.

#### 1.3.2 Raetic and Lemnian

From the eastern Alps, homeland of the tribes called Raeti by the Romans, come a very few inscriptions in a language which has been claimed to bear certain Indo-European characteristics. For example, from an inscription carved on a bronze pot (the Caslir Situla; see Fig. 1.5) comes the Raetic form *-talina* which has been compared to Latin *tollo* “I raise” (see Pulgram 1978:40 with additional references). However, similarities to Etruscan have



**a lavisešeli**

**b 1, 2 velxanu | lup·nu pitiave |**

**3 kusenkustrinaxe**

**c felna vinutalina.**

**Figure 1.5** The Caslir Situla

also been identified and the two are perhaps to be placed in a single language family, along with a language attested on the island of Lemnos in the north of the Aegean Sea. Lemnian is known principally from a single inscribed stele bearing the engraved image of a warrior, dated to the sixth century BC. On these connections, see Chapter 39, §1.

Of the Raeti, the Roman historian Livy (*History* 5.33.11) writes, following upon his discussion of the Etruscans: “Undoubtedly the Alpine tribes also have the same origin, particularly the Raeti, who have been made wild by the very place where they live, preserving nothing of their ancient ways except their language – and not even it without corruptions.”

### 1.3.3 Ligurian

The Ligurians were an ancient people of northwestern Italy. Writing in the second century BC, the Greek historian Polybius (*Histories* 2.16.1–2) situates the Ligurians on the slopes of the Apennines, extending from the Alpine junction above Marseilles around to Pisa on the seaward slopes and to Arezzo on the inland side. Another Greek, Diodorus Siculus (*Library of History* 5.39.1–8), writes of the Ligurians eking out a life of hardship in their heavily forested, rock-strewn, snow-covered homeland and of the extraordinary stamina and strength which this lifestyle engendered in both men and women.

The Ligurian language appears to be attested in certain place names and glosses, some of which have been assigned Indo-European etymologies. For example, Pliny the Elder, a Roman author of the first century AD, in describing the grain called *secale* in Latin, noted

that its Ligurian name (the name among the Taurini) is *asia* (*Natural History* 18.141). If the Ligurian form was once *sasia* (see Conway, Whatmough, and Johnson 1933 II:158), then, it has been proposed, the word may find relatives in Celtic – Welsh *haid* and Breton *heiz* “barley.” The location of its speakers, abutting Celtic areas (and Strabo writes of Celtoligurians; *Geography* 4.6.3), might itself be taken to suggest an affiliation with the Indo-European family, but such a relationship cannot be confirmed by the available linguistic evidence.

### 1.3.4 Illyrian

The historical peoples called Illyrian occupied a broad area of the northwest Balkans. Evidence for an Indo-European intrusion into the region can be identified by the late third millennium BC; an identifiable “Illyrian” culture appears only in the Iron Age (see, *inter alia*, Wilkes 1992:28–66). By the first century AD, the Greek geographer Strabo, in describing that part of Europe south of the Ister (the Danube) can identify as Illyrian those people inhabiting the region bounded on the east by the meandering Ister, on the west by the Adriatic Sea, and lying above ancient Epirus (*Geography* 7.5.1). For the Romans, the province of *Illyricum* denotes a rather larger administrative area. The term “Illyrian” can, however, be used by classical authors to designate a variety of peoples in and beyond the Balkans (see the discussion in Katičić 1976:156–163).

Within the northwestern Balkan region itself there was considerable cultural diversity, with not only the so-called Illyrian tribes being present, but Celts as well, by at least the third century BC. Strabo writes of the Iapodes dwelling near Mount Odra (close to the border of modern Slovenia and Croatia) whom he calls a mixed Celtic and Illyrian tribe (*Geography* 4.6.10) and who, he adds, use Celtic armor but are tattooed like the Illyrians and Thracians (*Geography* 7.5.4; on the Thracians see §1.3.5). In his account of the wars which various Illyrian tribes waged against one another and against the Romans, the Greek historian and Roman citizen, Appian of Alexandria, writing in the second century AD, preserves a tradition in which one hears echoes of such Balkan ethnic diversity. Appian (*Roman History* 10.2) records that the Illyrians received their name from Illyrius, a son of Polyphemus (the cyclops of Homer’s *Odyssey*) and the nymph Galatea, and that Illyrius has two brothers, Celtus and Galas, namesakes of the Celts and the Galatae (the latter commonly being synonymous with “Celt” and perhaps used here to invoke descent from Galatea).

The Illyrian language presents an unusual case. While the Illyrians are a well-documented people of antiquity, not a single verifiable inscription has survived written in the Illyrian language (on two proposed Illyrian inscriptions, one demonstrably Byzantine Greek, see Katičić 1976:169–170). Even so, much linguistic attention (perhaps a disproportionately large amount) has been paid to the language of the Illyrians. Chiefly on the basis of Illyrian place and personal names, the language is commonly identified as Indo-European. To provide but two examples, the frequently attested name *Vesclaves* has been etymologized as a reflex of Proto-Indo-European *\*wesu-klēwes* (“good fame”), with Sanskrit *Vasuśravas* being drawn into the analysis; the place name *Birziminium*, interpreted as meaning “hillock,” has been traced to the Proto-Indo-European root *\*b<sup>h</sup>erǵh-*, source of, *inter alia*, Germanic forms such as Old English *beorg* “hill” (see Katičić 1976:172–176 for discussion). This onomastic evidence is supplemented by the survival of just a very few glosses of Illyrian words; for example, the Illyrian word for “mist” is cited as *rhinos* (ῥινός) in one of the scholia on Homer; see Katičić 1976:170–171, who compares Albanian *re*, earlier *ren*, “cloud.” Extensive study of Illyrian was undertaken by Hans Krahe in the middle decades of the twentieth

century, who, along with other scholars, argued for a broad distribution of Illyrian peoples considerably beyond the Balkans (see, for example, Krahe 1940); though in his later work, Krahe curbed his view of the extent of Illyrian settlement (see, for example, Krahe 1955). Radoslav Katičić (1976:179–180) has argued, on the basis of a careful study of the onomastic evidence, that the core onomastic area of Illyrian proper is to be located in the south-east of that Balkan region traditionally associated with the Illyrians (centered in modern Albania).

The modern Albanian language, it has been conjectured, is descended directly from ancient Illyrian. Albanian is not attested until the fifteenth century AD and in its historical development has been influenced heavily by Latin, Greek, Turkish, and Slavic languages, so much so that it was quite late in being identified as an Indo-European language. Its possible affiliation with the scantily attested Illyrian, though not unreasonable on historical and linguistic grounds, can be considered little more than conjecture barring the discovery of additional Illyrian evidence.

### 1.3.5 Thracian

At the northern end of the Aegean Sea, stretching upward to the Danube, lived in antiquity people speaking the Indo-European language of Thracian. The ancestors of the Iron Age Thracians had probably arrived in the Balkans as a part of the movement which brought the forebears of the Illyrians. For the Greeks, Thrace was a place wild and uncultivated, home to both savage Ares and Dionysus, god of wine who inspired frenzy and brutality in his worshippers. Herodotus (*Histories* 5.3; 9.119) writes of the Thracian practices of human sacrifice and widow immolation, and of the enormous population of the Thracians (second only to the Indians) and their lack of political unity. Were they unified, surmises the historian, they would be the most powerful people on the face of the earth.

Though the Thracian language is not well preserved, its attestation, unlike that of Illyrian, is sufficient to place its membership in the Indo-European family practically beyond doubt. A few short Thracian inscriptions survive (see Brixhe and Panayotou 1994a:185–188), but more valuable are the numerous glosses (e.g., *bólinthos* “European bison,” cf. Old Norse *boli* “bull”; *brûtos* “beer,” cf. Old English *breowan* “to brew”) coupled with the evidence of place and personal names. For a summary of the evidence see Katičić 1976:138–142; Brixhe and Panayotou 1994a:188–189; see also Cowgill and Mayrhofer 1986:54–55, with references. Onomastic evidence may suggest the occurrence of a language boundary within the Thracian area, demarcated by Mount Haemus. South of this boundary the language evidenced has been distinguished as Thracian, while that to the north has been called Daco-Mysian.

According to Greek tradition, the Phrygians of Anatolia had migrated from the Balkans (see Herodotus, *Histories* 7.73, who writes that the Phrygians were formerly called the Briges and had been neighbors of the Macedonians; on the Macedonians see §1.3.6), a view with which modern scholarship is generally in agreement. The Phrygian language does show certain similarities to Thracian, and some linguists have argued for linking the two in a single linguistic unit (Thraco-Phrygian). The appropriateness of the subgrouping is, however, uncertain; see Chapter 31, §1.5.

### 1.3.6 Macedonian

North of the Greeks, bracketed by Illyrians and Thracians, lived the Macedonians. Much uncertainty surrounds the linguistic status of the Macedonian peoples. Though, under the patronage of Macedonian kings, Philip the Second and his son Alexander the Great, Greek

culture would be spread across the Mediterranean and Near Eastern world and the Greek language would become a lingua franca (the Attic-based Koine dialect; see Ch. 24, §1) spoken from Italy to India, it remains unclear if Greek was the native language of the Macedonians (see Brixhe and Panayotou 1994b:206–207 for a synopsis of ideas about the identity of Macedonian).

To be sure, the Greek orator Demosthenes, in the fourth century BC, can revile and lambaste Philip as one of the *barbaroi* (“barbarians,” those who do not speak Greek, i.e., those who *babble*; *Orations* 3.17) and rehearse how in the old days the Macedonian king had been rightly subject to the Greeks, as *barbaroi* should be (*Orations* 3.24). He can skewer Philip with the charge that, not only is he not a Greek and unrelated to the Greeks, he is not even a *barbaros* from some worthwhile place, but he is a plague out of Macedonia – a place from which you cannot even acquire a good slave (*Orations* 9.31). A century earlier, Herodotus had told the story of an ancestor of Philip, Alexander the First (a contemporary of Herodotus), who had been allowed to compete in games at Olympia – though *barbaroi* were excluded from the competition – because he was able to demonstrate satisfactorily that he himself was descended from a Greek banished from Argos (*Histories* 5.22; 8.137–139).

Explicit references to “Macedonian speech” exist. Plutarch, the Greek savant of the first and second centuries AD, when writing of Cleopatra (*Life of Antony* 27.3–4), the last of the Ptolemies (the Macedonian kings of Egypt), lauds her linguistic abilities, reporting that she could speak the languages of the Ethiopians, Troglodytes, Hebrews, Arabs, Syrians, Medes, and Parthians. In contrast, her male predecessors had not even learned Egyptian and some had even “ceased to speak Macedonian” (μακεδονίζειν ἐκλιπόντων). Presumably they had continued to speak Greek (i.e., had not taken a vow of silence). Athenaeus, a Greek writer of the later second century AD, in his account of a “Learned Banquet” (*Deipnosophistae* 3.121f–122a), places on the lips of one of the guests, the cynic Cynulcus, a Latin word *decocta* (a kind of drink made by boiling and then rapidly cooling a liquid); in turn, Athenaeus has another guest, Ulpian (an “Atticist,” promoting the use of untainted Attic Greek), rebuke Cynulcus for uttering a barbarism (!). Cynulcus fires back, retorting that even in the best old Greek one finds Persian loanwords and that he knows many Attic Greeks “using Macedonian speech” (μακεδονίζοντας; a participle from Plutarch’s verb). Elsewhere, Plutarch uses an adverb *makedonistí* (μακεδονιστί) having the same sense. For example, in his *Life of Alexander* (51.4), Plutarch recounts how the Macedonian conqueror, in a fit of rage, refusing to be quieted by his body guards, shouted out for the *hypaspistai* (Macedonian infantry troops, one contingent of the army of Alexander), “calling *in Macedonian* – and this was a sign of a great disturbance.” The precise sense of “speaking Macedonian” in these and other passages can be and has been debated; yet when these references to Macedonian speech are considered in their context, it is not difficult for one to conclude that what is being reported is the use of a distinct, non-Greek (“barbarian”) Macedonian language.

In contrast, however, other classical authors explicitly identify the Macedonians as a Greek people. Polybius, the Greek historian of the second century BC, for example, describes Macedonians and Greeks as being *homophylos* (ὁμόφυλος) “of the same race” or “akin” (*Histories* 9.37.7). For references to other, similar texts, see Katičić 1976:107–108.

An interesting case is provided by an instance in which Macedonians identify themselves as Greeks and speakers of Greek. The Roman historian Livy (first centuries BC and AD), writing of events in the war waged by Philip the Fifth of Macedon and his Arcarnanian Greek allies against Athens, with Rome as its own ally, records a meeting of the council of the Aetolian Confederacy, at which representatives from Philip, from Athens and from Rome address the council, each seeking Aetolian assistance in the war (200 BC). In his speech to the council,



the Macedonian ambassador refers to the Romans as “a foreign people set apart more by *language* and customs and laws than by the space of sea and land” (31.29.12). In contrast, “Aetolians, Acarnanians and Macedonians [are] people of the *same language* . . . [and] with foreigners, with barbarians *all Greeks* are, and will be, at eternal war” (31.29.15). The dialect of the Aetolian Confederacy, a league of the Aetolians of northwest Greece, was the Northwest Greek Koine, a “common” dialect used throughout regions controlled by the Confederacy (see Ch. 25, §1.1.5). Is it this lingua franca to which Livy has his Macedonian diplomat self-servingly refer? One could well imagine that it would be the Macedonian’s *langue de choix* on such an occasion. The Acarnanians also inhabited northwest Greece, though Acarnanian inscriptions from this period are written in the Doric Koine, only slightly different from the Aetolian dialect.

Surviving Macedonian texts have not proved helpful in identifying the native language of the Macedonians. Most of the Macedonian inscriptions are written in Attic Greek, the dialect broadly disseminated by Philip and Alexander. A fourth-century BC inscription found recently in the remains of the great Macedonian city of Pella appears to be written in a variety of Northwest Greek and has led to conjectures that this may be the previously unattested Macedonian language (see the comments of Brixhe and Panayotou 1994b:209 along with the mention of other finds in n.19).

The evidence provided by Macedonian glosses is conveniently summarized by Katičić (1976:108–112), who analyzes these as belonging to three different classes. One class consists of words that are quite close to known Greek lexemes, some, though probably not all, of which appear likely to be loanwords directly from Greek: for example, *kommárai*; compare Greek *kámmaroi* (κάμμοροι), a type of lobster (pl.). A second set is made up of Macedonian words which have no Greek counterparts, such as *aliē* “boar.” The third group is similar to the first to the extent that it consists of Macedonian words apparently having Greek counterparts; it differs from the first class, however, in that these Macedonian words are perhaps to be analyzed as cognates of the Greek lexemes, rather than borrowings. In other words, by such an analysis, the related Macedonian and Greek forms have evolved historically from words occurring in a common parent language, either Proto-Indo-European or, alternatively, some later, intermediate Balkan Indo-European language. Compare, for example, Macedonian *adē* “sky” and Greek *ait<sup>h</sup>ēr* (αἰθήρ); Macedonian *kebalá* “head” (cf. *gabalá* which the Greek lexicographer Hesychius also glosses as “head,” without identifying the linguistic source of the word) and Greek *kep<sup>h</sup>alē* (κεφαλή). If such sets are rightly analyzed as cognates, the Macedonian language departs conspicuously from Greek in showing voiced unaspirated rather than voiceless aspirated reflexes of the earlier Indo-European voiced aspirated stops (on the Greek development, see Ch. 24, §3.7.1).

### 1.3.7 Messapic

The Messapii were a people of southeast Italy, inhabiting ancient Calabria (the Sallentine peninsula, the “heel” of the Italian “boot”). Strabo, the Greek geographer, records (*Geography* 6.3.1) that the Greeks give the name *Messapia* to that region, also called *Iapygia*, but adds that the locals of the area make a distinction between the Salentini (in the south) and the Calabri. Northward lies the country of the Peucetii and of the Daunuii (Apulia). For Polybius (*Histories* 3.88.4), however, Iapygia is the region inhabited by the Daunuii, Peucetii, and Messapii (though elsewhere he writes of “Iapyges and Messapii”; see *Histories* 2.24.11).

Messapic survives in a large number of inscriptions, recording chiefly proper names, dating from about the sixth to the first century BC (the most abundantly attested ancient language not to receive individual treatment in this encyclopedia), including many recent



finds from a grotto in Lecce (see Santoro 1983–1984). This language of ancient Italy is Indo-European, but not Italic; that is, it is not a member of the subfamily to which belong Latin and Sabellian (see Chs. 32 and 33). No close genetic affiliation with any other known Indo-European language can be definitively demonstrated, though a close connection to Illyrian has been alleged. Indeed, the Messapic materials provided a major component of the evidence adduced by Krahe and others for the study of Illyrian. There do exist ancient traditions about the settling of southeast Italy by Illyrian peoples. For example, Pliny (*Natural History* 3.102) makes cursory reference to the story that the “Paedicali” of Apulia were descended from nine young men and nine young women of Illyria. A linking of the two languages, Illyrian and Messapic, must, however, remain a linguistically unverifiable hypothesis until such time as Illyrian is better attested.

### 1.3.8 Zapotec

Far away from Italy and the Balkans, in Mesoamerica, yet another language of antiquity is attested. Zapotec is one of several documented early Mesoamerican languages, others being Mayan (Ch. 43), Epi-Olmec (Ch. 44), Mixtec and Aztec; both of the last-named are attested by about AD 1100 but are best known from sixteenth-century AD manuscripts (*inter alia*, for Mixtec, see Marcus 1992:57–67 and Jansen 1992:20–33; for Aztec, see Marcus 1992 and Prem 1992:53–69; on the pictographic records of the Tlapanecs, see Vega Sosa 1992:34–52). Zapotec inscriptions, carved in stone like those of the Mayans and Epi-Olmecs, may date as early as 500 or 600 BC (though the earliest uncontroversial dates are between 400 and 200 BC) and are last attested in about AD 900 (as with Mixtec and Aztec, Zapotec manuscripts also occur in the sixteenth century, though the corpus is small). Several dozen short inscriptions exist, as well as a large number of calendrical citations, providing perhaps one to three hundred distinct glyphic components.

Owing to the difficulty in obtaining information on this language, a brief grammatical sketch of Zapotec, based on our present, limited understanding of the language, has been included as an appendix to Chapter 44.

## 1.4 Format and conventions

Each chapter, with only the occasional exception, adheres to a common format. The chapter begins with an overview of the history (including prehistory) of the language, at least up to the latest stage of the language treated in the chapter, and of those peoples who spoke the language (§1, HISTORICAL AND CULTURAL CONTEXTS). Then follows a discussion of the development and use of the script(s) in which the language is recorded (§2, WRITING SYSTEMS); note that the complex Mesopotamian cuneiform script, which is utilized for several languages of the ancient Near East – Sumerian (Ch. 2), Elamite (Ch. 3), Hurrian (Ch. 4), Urartian (Ch. 5), Akkadian and Eblaite (Ch. 8), Hittite (Ch. 18), Luvian (Ch. 19) – and which provides the inspiration and graphic raw materials for others – Ugaritic (Ch. 9) and Old Persian (Ch. 28) – is treated in most detail in Chapter 8, §2. The next section presents a discussion of phonological elements of the language (§3, PHONOLOGY) identifying consonant and vowel phonemes, and treating matters such as allophonic and morpho-phonemic variation, syllable structure and phonotaxis, segmental length, accent (pitch and stress), and synchronic and diachronic phonological processes. Following next is discussion of morphological phenomena (§4, MORPHOLOGY), focusing on topics such as word structure, nominal and pronominal categories and systems, the categories and systems of finite verbs and other verbal elements (for explanation of the system of classifying Semitic verb

stems – G stem, etc. – see Ch. 6, §3.3.5.2), compounds, diachronic morphology, and the system of numerals. Treatment of syntactic matters then follows (§5, SYNTAX), presenting discussion of word order and coordinate and subordinate clause structure, and phenomena such as agreement, cliticism and various other syntactic processes, both synchronic and diachronic. The description of the grammar closes with a consideration of the lexical component (§6, LEXICON); and the chapter comes to an end with a list of references cited in the chapter and of other pertinent works (BIBLIOGRAPHY).

To a great extent, the linguistic presentations in the ensuing chapters have remained faithful to the grammatical conventions of the various language disciplines. From discipline to discipline, the most obvious variation lies in the methods of transcribing sounds. Thus, for example, the symbols *ś*, *ṣ*, and *ṭ* in the traditional orthography of Indic language scholarship represent, respectively, a voiceless palatal (palato-alveolar) fricative, a voiceless retroflex fricative, and a voiceless retroflex stop. In Semitic studies, however, the same symbols are used to denote very different phonetic realities: *ś* represents a voiceless lateral fricative while *ṣ* and *ṭ* transcribe two of the so-called emphatic consonants; the latter a voiceless stop produced with a secondary articulation (velarization, pharyngealization, or glottalization), the former either a voiceless fricative or affricate, also with a secondary articulation. Such conventional symbols are employed herein, but for any given language, the reader can readily determine phonetic values of these symbols by consulting the discussion of consonant and vowel sounds in the relevant phonology section.

Broad phonetic transcription is accomplished by means of a slightly modified form of the International Phonetic Alphabet (IPA). Most notably, the IPA symbols for the palato-alveolar fricatives and affricates, voiceless [ɧ] and [tɕ] and voiced [ʒ] and [dʒ], have been replaced by the more familiar [š], [č], [ž], and [j] respectively. Similarly, [ɣ] is used for the palatal glide rather than [j]. Long vowels are marked either by a macron or a colon.

In the phonology sections, phonemic transcription, in keeping with standard phonological practice, is placed within slashes (e.g., /p/) and phonetic transcription within square brackets (e.g., [p]); note that square brackets are also used to fill out the meaning of a gloss and are employed as an element of the transcription and transliteration conventions for certain languages, such as Elamite [Ch. 3] and Pahlavi [Ch. 30]). The general treatment adopted in phonological discussions has been to present transcriptions as phonetic rather than phonemic, except in those instances in which explicit reference is made to the phonemic level. Outside of the phonological sections, transcriptions are usually presented using the conventional orthography of the pertinent language discipline. When potential for confusion would seem to exist, transcriptions are enclosed within angled brackets (e.g., <p>) to make clear to the reader that what is being specified is the *spelling* of a word and not its *pronunciation*.

## Bibliography

- Bader, F. (ed.). 1994. *Langues indo-européennes*. Paris: CNRS Editions.
- Bricker, V. 1992. *Epigraphy*. Supplement to the Handbook of Middle American Indians, vol. V. Austin: University of Texas Press.
- Brixhe, C. and A. Panayotou. 1994a. "Le thrace." In Bader 1994, pp. 179–203.
- . 1994b. "Le macédonien." In Bader 1994, pp. 205–220.
- Brown, E. 1990. "Traces of Luwian dialect in Cretan text and toponym." *Studi micenei ed egeo-anatolici*, pp. 225–237.
- Chadwick, J. 1990. "Linear B and related scripts." In *Reading the Past*, pp. 137–195. Introduction by J. Hooker. Los Angeles/Berkeley: University of California Press; London: British Museum.

- Conway, R., J. Whatmough, and S. Johnson. 1933. *The Prae-Italic Dialects of Italy* (3 vols.). Cambridge, MA: Harvard University Press.
- Cowgill, W. and M. Mayrhofer. 1986. *Indogermanische Grammatik*, vol. I. Heidelberg: Carl Winter.
- Daniels, P. 1996. "The first civilizations." In Daniels and Bright 1996, pp. 21–32.
- Daniels, P. and W. Bright (eds.). 1996. *The World's Writing Systems*. Oxford: Oxford University Press.
- Dhorme, E. 1946. "Déchiffrement des inscriptions pseudo-hiéroglyphiques de Byblos." *Syria* 25:1–35.
- Diringer, D. 1968. *The Alphabet: A Key to the History of Mankind*, vol. I. New York: Funk and Wagnalls.
- Duhoux, E. 1982. *L'Étéocrétois*. Amsterdam: Gieben.
- Dunand, M. 1945. *Byblia grammata: Documents et recherches sur le développement de l'écriture en Phénicie*. Beirut: Ministry of Education of Lebanon.
- Englund, R. 1996. "The Proto-Elamite script." In Daniels and Bright 1996, pp. 160–164.
- Finkelberg, M. 1990–1991. "Minoan inscriptions on libation vessels." *Minos* 25–26:43–85.
- Griffith, F. 1911. *Karanog: The Meroitic Inscriptions of Shablul and Karanog*. Philadelphia: University Museum.
- . 1912. *Meroitic Inscriptions II*. London: Egypt Exploration Fund.
- Jackson, K. 1980. "The Pictish language." In F. Wainwright (ed.), *The Problem of the Picts*, pp. 129–160. Perth: Melven.
- Jansen, M. 1992. "Mixtec pictography: conventions and contents." In Bricker 1992, pp. 20–33.
- Katičić, R. 1976. *Ancient Languages of the Balkans*, part 1. The Hague: Mouton.
- Krahe, H. 1940. "Der Anteil der Illyrier an der Indogermanisierung Europas." *Die Welt als Geschichte* 6:54–73.
- . 1955. *Die Sprache der Illyrier* (2 vols.). Wiesbaden: Otto Harrassowitz.
- Lehmann, W. 1967. *A Reader in Nineteenth-Century Historical Indo-European Linguistics*. Bloomington: Indiana University Press.
- Mallory, J. 1989. *In Search of the Indo-Europeans*. London: Thames and Hudson.
- Marcus, J. 1992. *Mesoamerican Writing Systems*. Princeton: Princeton University Press.
- Masson, E. 1974. *Cyprominoica*. Göteborg: Paul Åströms Forlag.
- . 1977. "Présence éventuelle de la langue hourrite sur les tablettes chypro-minoennes d'Enkomi." *Revue Roumaine de Linguistique* 22:483–488.
- Masson, O. 1983. *Les inscriptions chypriotes syllabiques*. Paris: Édition E. de Boccard.
- McMannus, D. 1991. *A Guide to Ogam*. Maynooth: An Sagart.
- Palaima, T. 1988. "The development of the Mycenaean writing system." In J. Olivier and T. Palaima (eds.), *Texts, Tablets and Scribes*, pp. 269–342. Supplement to *Minos* 10.
- . 1989. "Cypro-Minoan scripts: Problems of historical context." In Y. Duhoux, T. G. Palaima, and J. Bennet (eds.), *Problems in Decipherment*, pp. 121–187. Louvain-la-Neuve: Peeters.
- Palmer, L. 1968. "Linear A and the Anatolian languages." In *Atti e memorie del 1° congresso internazionale di micenologia*, vol. I, pp. 339–354. Rome: Edizioni dell'Ateneo.
- Parpola, A. 1994. *Deciphering the Indus Script*. Cambridge: Cambridge University Press.
- . 1996. "The Indus Script." In Daniels and Bright 1996, pp. 165–171.
- Poultney, J. 1979. "The language of the Northern Picene inscriptions." *Journal of Indo-European Studies* 7:49–64.
- Prem, H. 1992. "Aztec writing." In Bricker 1992, pp. 53–69.
- Pulgram, E. 1978. *Italic, Latin, Italian*. Heidelberg: Carl Winter.
- Santoro, C. 1983–1984. *Nuovi studi messapaci*. Galatina, Italy: Congedo.
- Swiggers, P. 1996. "The Iberian Scripts." In Daniels and Bright 1996, pp. 108–112.
- Untermann, J. 1975. *Monumenta Linguarum Hispanicarum i, Die Münzlegenden*. Wiesbaden: Dr. Ludwig Reichert.
- . 1980. *Monumenta Linguarum Hispanicarum ii, Die Inschriften iberischer Schrift aus Südfrankreich*. Wiesbaden: Dr. Ludwig Reichert.
- . 1990. *Monumenta Linguarum Hispanicarum iii, Die iberischen Inschriften aus Spanien*. Wiesbaden: Dr. Ludwig Reichert.
- . 1997. *Monumenta Linguarum Hispanicarum iv, Die tartessischen, keltiberischen und lusitanischen Inschriften*. Wiesbaden: Dr. Ludwig Reichert.

- Vega Sosa, C. 1992. "The annals of the Tlapanecs." In Bricker 1992, pp. 34–52.
- Wenig, S. 1982. "Meroe, Schrift und Sprache." In W. Helck and W. Westendorf (eds.), *Lexikon der Ägyptologie*, vol. 4, pp. 104–107. Wiesbaden: Otto Harrassowitz.
- Whittaker, G. 1992. "The Zapotec writing system." In Bricker 1992, pp. 5–19.
- Wilkes, J. 1992. *The Illyrians*. Oxford: Blackwell.
- Woodard, R. 1997. "Linguistic connections between Greeks and non-Greeks." In J. Coleman and C. Walz (eds.), *Greeks and Barbarians*, pp. 29–60. Bethesda, MD: CDL Press.

# Sumerian

PIOTR MICHALOWSKI

## 1. HISTORICAL AND CULTURAL CONTEXTS

### 1.1 Introduction

Of all the extinct languages of the ancient world, Sumerian has the longest literary tradition, extending over roughly three thousand years. The time span and geographical spread of the spoken language is not known and is the subject of much speculation. Presumably it was once the major vernacular in the southern part of Mesopotamia, but it is impossible to establish if it was ever spoken outside of this enclave. In modern terms this would be the area of Iraq south of Baghdad. Estimates on the time of the demise of spoken Sumerian range from the third to the middle of the second millennium BC (see Michalowski, 2002 [2005], 2006). It seems that even in early times Sumerian speakers came into contact with Semitic languages, as evidenced by numerous loanwords from early Semitic. Some have hypothesized additional Mesopotamian substrate languages, but the evidence for this is lacking (Rubio 1999b).

The native designations for the “land of Sumer” are *kiĝir* (written *ki-en-gi*) in Sumerian and *māt šumerim* in Akkadian. Related to this are the respective language labels *eme-gir*<sub>15</sub> and *šumeru*, which have been the subjects of much etymological speculation. If *gir*<sub>15</sub> means “native,” then the Sumerian terms would mean “native land” (*ki.ĝir*) and “native language” (Steinkeller 1993:112–113). The origins and meaning of the Akkadian *šumeru* – the source of modern renditions such as *Sumer* – remain unknown. Equally opaque are the native geographical concepts. We know that beginning in the middle of the third millennium BC, southern Mesopotamia was thought of as divided between the “Land of Sumer” in the south and the “Land of Akkad” to the north, but it is difficult to establish any native border between the two. A broken passage in a hymn to the main temple of the city of Nippur seems to place that city at the dividing point, but the implications of the line are unclear.

### 1.2 Textual evidence

The oldest Sumerian texts – perhaps even the oldest written texts known to us – are the approximately five thousand clay tablets found discarded in debris in the ceremonial center of the city of Uruk, written in an early form of the *cuneiform* script (see §2). These tablets, which are dated around 3200 BC, have been seriated, on the basis of script, format, and content, into two general groups corresponding, in theory, to archeological levels from the site: Uruk IV and III, although they were not actually found in those levels.

Close to 90 percent of these early tablets are administrative records, but there are also word lists that were used in the teaching of the writing system (about 670 of the total known

5,820 archaic texts). One composition among these has been considered by some to be a narrative literary composition; others think it is a word list. In light of later usage of such compositions in the educational system, the difference between the two categories may be less than it appears to be. While the general transactions can be understood, the texts cannot all be precisely read; even the actual number of discrete signs is disputed, with estimates ranging from just over 700 to almost 2,000. Some have argued that the system was not linked to any language or was meant to represent an unknown, pre-Sumerian tongue. The existence of phonetic glosses within certain signs, however, strongly suggests that the administrative language was indeed Sumerian. Thus the sign AMA, which is used later for Sumerian *ama*, “mother,” contains within it the phonetic indicator, or gloss, *am*<sub>6</sub>, to help distinguish it from similar signs and to prompt the proper identification. The latter phase of archaic cuneiform, Uruk III, is attested not only in Uruk and possibly at Larsa in the south, but also farther north at Jemdet Nasr, Uqair, and Tell Asmar, demonstrating the relatively rapid spread of the new invention.

We do not know how long this particular phase of cuneiform lasted, nor do we have any evidence for the changes that must have taken place early in the third millennium. We have to wait for about four hundred years for our next archaic texts from Ur, dated approximately 2800 BC (Wright 1969). The 375 tablets from this city are primarily administrative documents; additionally, as at Uruk, one also finds pedagogic word lists, and one possible literary mythological fragment. Although these laconic tablets are difficult to translate, the notation of a few morphological elements and phonetic glosses provides convincing evidence that the language of the texts is indeed Sumerian.

The next larger groups of texts from Sumer are the Early Dynastic III texts from Fara (ancient Shuruppak), Abu Salabikh, Nippur, and Adab from around 2500 BC. The majority of tablets found at the first two of these sites are literary, and now for the first time we have evidence for an extensive written poetic tradition. This literature was widely distributed wherever cuneiform was taught; some of the same compositions have been discovered, in slightly later copies, far to the west, during excavations of the Syrian city of Ebla. Syrian scribes used cuneiform to write a Semitic language that we call Eblaite (see Ch. 8), but they also copied Sumerian and Akkadian literary texts, including word lists, that they inherited from Sumer and from northern Babylonia. Many cities in northern Babylonia and in Syria used writing, as is documented by the roughly contemporary tablets from Ebla, Mari, Tell Beydar and Tell Brak. There are small differences in the manner in which cuneiform was used in these places, but these are only variations within a common tradition. Moreover, sometime before the middle of the third millennium, cuneiform had already been fully adapted to write Semitic languages, including Eblaite and Akkadian.

One of the characteristic peculiarities of Early Dynastic literature is the existence of a separate manner of writing that has been termed UD.GAL.NUN (UGN), from a sequence of graphemes commonly found in these texts. With a few exceptions, the signs used are the same as in “normal” Sumerian, but the values (or “readings”) of these signs are clearly different. Only a small number of these have been deciphered, among them the sequence that originally gave this system its name: UD corresponds to the classifier (see §2) *dingir* “god, divine name,” GAL to *en*, and NUN to *lil*<sub>2</sub>. These three signs therefore spell out the name of the chief god of Sumer, Enlil, or Ellil, normally written as *en-lil*<sub>2</sub>. This was not a local tradition, since texts of this type have been found at Nippur and Abu Salabikh as well as at Fara; its purpose and origins are simply unknown to us. This manner of writing disappears forever after this period, and remains but a reminder of the complex route that writing took from its origins, with many experiments and dead ends that have not been documented to date.

### 1.3 Akkadian and Sumerian

With the rise of Akkad around the year 2350 BC, the Semitic Akkadian (see Ch. 8) becomes one of the official languages of Sumer and joins the older language as a vehicle of administration and communication. Semitic had been written in the north, but was only sporadically attested in Sumer. Now certain communities limited themselves exclusively to Akkadian for written communication; others retained Sumerian for local accounts but used the other language to communicate with the central government. Very little literature has survived from this period, leaving us in the dark concerning schooling and scribal education.

Soon after the collapse of the Akkadian state, Sumer and Akkad were once again dominated by one royal house, this time centered at the old city of Ur. The Third Dynasty of Ur (c. 2112–2004 BC) ruled for almost exactly a century and left behind an unprecedented number of bureaucratic records. There are approximately forty thousand published administrative texts from this time, and countless more remain in museums and private collections. This documentation is almost exclusively Sumerian, but small numbers of Akkadian texts from northern sites suggest that our large sample is skewed by chance of discovery and that Sumerian was not the sole official language of the time. The documents from Puzrish-Dagan, Ur, Umma, Girsu, Eshnunna, and Nippur do indicate that the central bureaucracy preferred Sumerian as a written language, but small archives from northern places such as Ishan Mizyad indicate that Akkadian was used as well. The Ur III kings oversaw writing reforms and a drastic change in the school tradition. Most of the Early Dynastic literary legacy was discarded and new texts, many of them honoring contemporary rulers, were composed. Most of these are known only in later copies, but a sizable group of Ur III Sumerian literary tablets from Nippur awaits publication.

After the collapse of the Ur III state, Sumerian retained its status as an official language in the south, while in the north, Akkadian dialects began to take over in writing. The last Sumerian archival letter dates from the time of Lipit-Eshtar of Isin (c. 1873–1865 BC), and by the middle of the nineteenth century BC Sumerian was no longer used for administrative and accounting purposes. Letters, wills, and other everyday texts were written in Akkadian; Sumerian stock phrases were often employed in legal and administrative documents, but they were undoubtedly read aloud in the Semitic vernacular. Schooling, however, remained primarily in the old tongue. Indeed, this is the period that has left us the largest quantity of Sumerian literary compositions. We have a good knowledge of educational practices in southern cities such as Nippur, Isin, Uruk, and Ur. The curriculum consisted of the study of lexical lists, proverbs, and a few easy royal hymns in the early stages, after which the student graduated to the copying of a broad range of compositions, including royal and divine hymns, epics, laments, epistolary texts, as well as idealized debates, and a small number of legal, historical, and historiographic texts. Liturgical and magical texts are more common in northern and peripheral cities.

### 1.4 The status of Sumerian in antiquity

For inestimable years Sumerian was a living language in southern Mesopotamia. It was the first language in Western Asia that was committed to writing and this, if nothing else, assured its prestige status for millennia to come. By the Old Babylonian period it was limited to schools and temples, and until the end of the use of cuneiform it remained a high prestige liturgical language that was studied, with various levels of success, throughout the Near East.



## 1.5 External affiliation

Sumerian is an isolate, like Ainu, Etruscan, Basque, or Burushaski. Over the years various unsuccessful attempts have been made to link it with a variety of languages or language families, among them Chinese, Tibetan, Hungarian, Turkish, and Indo-European. These attempts have sometimes been flavored with nationalist fervor. More recently some scholars have tried to include Sumerian within the hypothetical Nostratic proto-language of Eurasia, while others have excluded it from such reconstructions.

## 1.6 General characteristics

The isolate Sumerian is an agglutinating language. The word order of simple declarative sentences is strongly SOV, although this impression may be skewed by the highly formal nature and limited rhetorical scope of much of the sample. Heads and dependents are marked, nominal cases are marked with postpositions, genitives succeed the nouns they modify, adjectives follow nouns, and subordinate clauses usually, but not always, precede main ones.

Sumerian is generally characterized as an ergative language because the main participants of an action are marked according to a system that formally recognizes agents of transitive clauses as different from transitive patients and intransitive subjects. The former are marked by the ergative case, the latter by the absolutive. Few languages are fully ergative. Sumerian, like many other languages, shows various splits: while nominal marking is fully ergative, independent personal pronouns, verbal imperatives and cohortatives, as well as certain participial constructions, are nominative-accusative. Verbal concord works on a split determined by aspect: the perfective is ergative, and the imperfective is nominative-accusative. Sumerian is not alone in this respect and aspectual splits of this type are found in various unrelated Asian languages, including Georgian, Burushaski (Tibet), the Iranian Pashto, as well as in certain Indo-European languages of India. This has led some (Nichols 1993, also implied in Anderson 1985:182) to suggest that this may be an areal phenomenon.

It is usually remarked that ergativity is a strictly morphological phenomenon in Sumerian and there is no evidence that it triggers any syntactic operations (Michalowski 1980, Zólyomi 1996a), but this is a matter that requires further investigation.

## 1.7 The later use of Sumerian

Little is known at present about the use of Sumerian in the centuries immediately following the fall of the Old Babylonian state around 1595 BC. Akkadian was now widely used for written communication throughout the Near East, from Iran to Anatolia, the Levant, and even Egypt. Some selected Sumerian texts were transmitted to these areas and were used in the study of cuneiform, but most of the Old Babylonian compositions were discarded, and were never read again until modern times. The same holds true for subsequent Babylonian and Assyrian periods: Akkadian was the major language, and Sumerian was studied in school and used in liturgical contexts, although the old language was sometimes used in Babylonian building inscriptions in the late second and early first millennia. Sumerian prayers, laments, and incantations remained in use in rituals, indeed they were studied, edited, and reedited and new texts continued to be composed even after the conquest of Babylonia by Alexander of Macedon in 331 BC. Even as late as the third century numerous Sumerian liturgical texts were redacted and written anew in cities such as Uruk and Babylon, including large numbers of prayers and



incantations. There are even a handful of tablets with Sumerian or Akkadian exercises on one side and Greek transcriptions on the other. It is difficult to date these texts, but some would claim that they might be as late as the second century AD (Geller 1997).

## 1.8 Sumerian dialects

Because of the official nature of written Sumerian, the study of possible dialectal distinctions is somewhat problematic. There are synchronic and diachronic variations and these have sometimes been ascribed to dialectal differences. For example, in particular places during the third millennium, a verbal prefix *i-* is written *e-* in certain contexts; in other places there is a prefix *a-* that rarely occurs elsewhere. Are such isolated isoglosses sufficient to speak of dialects? Only recently Krispijn (2002 [2005]) has attempted to define a specific Lagash-area dialect on the basis of a number of phonological and morphological features. In a literary depiction of an idealized and perhaps satirized school examination, a teacher asks a student if he knows the languages of priests, metalworkers, shepherds, and so forth (Sjöberg 1975:166). This document has been interpreted as providing information on “dialects” or, better, sociolects, but most probably it only refers to knowledge of technical terms connected with these professions that were included in word lists that were memorized and copied as part of scribal training.

### 1.8.1 The “main dialect” and the “women’s tongue”

The main dialect distinction in Sumerian, as reflected in native terminology, is between *eme-ĝir*<sub>17</sub> (EG) and *eme-sal* (ES). The former seems to be the native term for what we could call Standard Literary Sumerian. The latter is restricted to ritual texts – primarily those used by lamentation priests (*gala*) – and to the direct speech of certain goddesses and their messengers in literary texts, although these same goddesses speak fluent “Standard Sumerian” in other compositions. On the basis of false etymology, and misunderstandings of the distribution of Emesal, it has been often called a “woman’s tongue,” leading some to invoke unnecessary ethnographic analogies. Likewise, it has been claimed that the *gala* priest and the divine messengers were eunuchs (e.g., Boisson 1992:434), although there is no evidence for castration, human or divine, in ancient Sumer.

The sign SAL has three basic readings, *mi*<sub>2</sub>, *munus*, and *sal*. The first represents only the phonological sequence /mi/ (with very limited distribution), the second means “woman,” and the third means “thin.” Thus the term *eme-sal* – and the reading is assured because of the Akkadian loan *emešallu* – refers to some sort of pronunciation, but its origins and use in living speech cannot be determined. Emesal is not attested before the Old Babylonian period. At that time Emesal texts are primarily, although not exclusively, attested in northern Babylonian cities such as Kish and Sippar, but are much less common in the school texts from Nippur, Ur, and other cities in the south. This may be attributed to differences in school curricula. It is also possible that cult texts were transmitted mainly orally in southern Babylonia, but written down in the northern area. This may have been one of the consequences of massive social and political upheavals during the last quarter of the eighteenth century BC that led to the abandonment of many southern settlements and the emigration of much of the population northwards. By the first millennium BC, the majority of Sumerian texts were liturgical Emesal compositions, aside from incantations, which continued to be copied and recited in the main dialect. Thus, most literate priests used Emesal more than the old Standard Sumerian.

Various attempts have been made to explain the origins and “dialectal” status of Emesal. Alster (1982) thought that it might be related to the UD.GAL.NUN texts of the Early Dynastic period. Others have sought its origins in regional dialects. Bobrova (1989) suggested that it was the dialect of a cultic center of the goddess Inanna, since this goddess speaks in ES in literary texts. There is no evidence at present to support this claim. Bauer (1998:436) has noted that some of the sound changes that are characteristic of ES can be sometimes found in third-millennium texts from the Lagash area; this led him to propose that Emesal was related to, if not based on, the local version of Sumerian, which was hidden from our view by the scribes who wrote in the standard version of the language.

The main distinctions between the two forms of Sumerian are phonological. Thus, EG *ḡ* corresponds to ES *m* (EG *ḡar* ~ ES *mar* “to place”), or EG *d* corresponds to ES *z* (EG *udu* ~ ES *eze* “sheep”). A small number of basic terms have unexplainable lexical alternates: EG *ereš* ~ ES *gašan* “queen, mistress,” or EG *nitadam* ~ ES *mudna* “betrothed man.” A full list of correspondences as well as a listing of the known ES words can be found in Schretter 1990.

## 1.9 The study of Sumerian

Because of a lack of known cognate tongues, and because Sumerian died out thousands of years ago, it is extremely difficult to establish a reliable grammar or lexicon of the language. Despite much progress over the years, there is still much disagreement about basic grammatical facts, and it is impossible to do justice to all the debates on the matter in a short survey. Many complex issues have had to be simplified or presented in an abbreviated fashion; because of a lack of any proper study, issues of syntax have suffered disproportionately. The following remarks represent an attempt to present the author's present opinions, tempered by a selective representation of other points of view. One should also note that despite the large number of surviving cuneiform tablets, there are severe limitations on what can be recovered. Not only was Sumerian written for thousands of years after it was no longer the vernacular, but what was written has preserved only a part of the language. The surviving texts consist primarily of highly conventionalized administrative documents, academic word lists, and poetic compositions; there is very little literary prose. As a result, one must always keep in mind that we are dealing with highly formalized forms of verbal art far removed from any putative language of the streets, constrained by certain conventions with restricted rhetorical scope.

## 2. WRITING SYSTEM

### 2.1 Cuneiform writing

Sumerian is written with a script known as *cuneiform* – impressed onto moist clay tablets, although there are also monumental texts inscribed on stone and other hard surfaces. Once dry, clay is extremely durable and therefore tens of thousands of such tablets have survived to the present day. It is impossible to quantify the available Sumerian language remains or to estimate what lies buried in museums and in the unexcavated mounds of the Near East.

Although some popular theories propose evolutionary precursors to this writing system, it seems much more probable that it was invented as a system, with all of its characteristic features intentionally bound into a comprehensive notational structure. The signs on the earliest tablets were drawn with a reed stylus (see Ch. 8, Fig. 8.2). Very soon the technique

changed, and the end of the stylus was used to impress wedges to make up a grapheme, and this manner of writing persisted from that time on. The wedge-like look of the script gave us the modern name cuneiform, from Latin *cuneus* “wedge.” In one Sumerian poem the signs are described as *gag*, “nail(s).” The earliest writing system, which has been variously designated as *archaic cuneiform* or *proto-cuneiform*, was designed for recording transactions, and thus the texts consist almost entirely of word and number signs.

The early history of cuneiform might be characterized as one of an uneasy adaptation of an autonomous communication system to accommodate natural language. By the middle of the third millennium the new system was capable of representing full utterances, but it was still something of a mnemonic device to the extent that no attempt was made to represent with precision all aspects of language. Only kernel elements were noted, and these were not inscribed in the order in which they were read. Thus a verb, which in later writing might have numerous affixes, would only carry one or two prefixes. The reader was expected to provide the missing elements and to unscramble the signs into their proper sequence. The graphic elements needed for fairly accurate phonological representation of Sumerian language were all in place, as was the case in contemporary Egyptian, but that was not the goal of the recording system.

## 2.2 Signs and conventions

The sign repertoire consists of three different types of signs: (i) *semantic classifiers* – Assyriologists refer to them as *determinatives*; (ii) *syllabograms* (also called *phonograms*) or phonetic signs; and (iii) *logograms*, or word signs. Signs have multiple values, and some can even function in all three capacities. Thus, the wedge sequence  $\rightarrow\uparrow$  can be read, depending on the context, (i) as the classifier for a divine name; (ii) as the syllabogram *an*; and (iii) as the noun *an* “heavens” or as *diĝir* “god.”

Certain conventions are used in the transliteration of sign sequences into the Roman alphabet. Sequences of cuneiform signs that represent roots and affixes are linked in transliteration by dashes, while morphemes are separated by periods. Similar or homophonous readings have been numbered, and modern scholars represent these indices with accents and/or with subscripted numbers. For example, the Sumerian word for “house, temple” can be transliterated either as *é* or as *e<sub>2</sub>* (the actual phonological shape was closer to /ha/). The unpronounced classifiers (determinatives) are transliterated with raised letters; for example the classifier for a divinity (*diĝir*) is abbreviated to <sup>d</sup>: for example, <sup>d</sup>*en-lil<sub>2</sub>* “(god) Enlil”; <sup>giš</sup>*tukul*, “(wooden) weapon”; *ur<sub>15</sub><sup>ki</sup>* “Ur<sup>(city)</sup>.” Sign names and signs with uncertain readings are represented in capital letters. The transliteration conventions are modern, but historic, and do not represent the current state of our knowledge about semiotics, morphology, or phonology. They are relicts of the decipherment of cuneiform, which has a long history going back almost two hundred years (Bottéro 1992).

## 2.3 Logographic writing

The early writing system is primarily logographic. Syllabograms were originally used to represent minimal grammatical information, and to assist in reading word signs by providing pronunciation glosses. Later sign usage and modern conventions of transliteration sometimes obscure this principle. For example, the Sumerian word for “ear” or “wisdom” is written with three signs and is commonly transliterated as *ĝeštug<sub>2</sub>*. Originally, the middle sign alone had the value *ĝeštug* and the first and third signs, *ĝeš* and *tug<sub>2</sub>* respectively, were phonetic complements. A more accurate transliteration would thus be *ĝeš<sup>2</sup>ĝeštug<sup>tug<sub>2</sub></sup>*. One

could argue that this rendered the middle sign redundant, but such instances only demonstrate the consistent use of word signs and the avoidance of syllabic spellings for roots. In principle the syllabic writing of roots was reserved, from the middle of the third millennium on, for loanwords. For example, the Sumerian word for “road,” *kaskal*, was written logographically with a single sign, but its synonym, borrowed from Akkadian, was written syllabically as *har-ra-an*.

The elementary indications of grammatical morphemes in Early Dynastic writing were in a sense also logographic, that is they did not always accurately represent phonological shapes but only a conventional form of a morpheme. Thus, to cite a classic example, the modal prefix *he-* (or *hV-*) is written as follows (see Civil and Biggs 1966:14):

(1)	2500 BC	2400–2000 BC	1800 BC
	he <sub>2</sub> - (/_e,i,a,u/)	he <sub>2</sub> - (/_e,i/)	he <sub>2</sub> - (/_e,i/)
		ha- (/_a,u/)	ha- (/_a/)
			hu- (/_u/)

Around 2400 BC the signs began to be written in the order in which they were to be read, and by 2000 BC most, if not all, grammatical elements were represented in writing. The general nature of the signs remained the same, but the structure of the system changed. Logograms and syllabograms were combined according to certain principles, but this does not mean that cuneiform writing moved towards a precise phonological representation of Sumerian.

## 2.4 The evolution of syllabic spelling

The complex move towards the implementation of a full syllabic repertoire was probably driven by multiple motivations. The application of cuneiform to represent Semitic languages such as Akkadian and Eblaite required the development of such a syllabary, as did the need to represent Semitic personal names in Sumerian texts. Such a full syllabary is known for Eblaite as early as 2500 BC, but the first adaptation of cuneiform to Semitic must have taken place somewhat earlier. Because of the word structure of Semitic, which requires the representation of changes that take place within roots, one could not simply use the same combinatorial principles that one used to write Sumerian. The distinct structure of discontinuous Afro-Asiatic roots favored a full syllabary rather than a logographic writing system, and therefore someone applied cuneiform to these languages by exploiting the CV and VC signs of the Sumerian script.

Certain conventions helped in interpreting the written segments, such as the use of the sequence CV–VC to express the sequence CVC. Although in Syrian Semitic writings the signs were written in the proper linguistic order, the texts from Sumer still exhibit a fairly free order of signs within a case of writing.

Eventually, these syllabic practices were partially applied back to Sumerian, and in the Early Dynastic texts we find an incipient use of syllabograms for loanwords, and for limited marking of bound morphemes. Loans and other syllabically spelled words are subject to certain conventions, such as the use of CV signs for the sequence CVC, as in *li<sub>2</sub>-ga* for *lidga* (a measure of capacity). Nominal case endings and possessive pronouns are sometimes written, sometimes omitted. Only one or two verbal affixes are provided to the reader.

The full syllabary would eventually be applied to Sumerian as well, but not in the same manner as in Semitic. Because Sumerian roots are often monosyllabic and do not take infixes, roots continued to be written with logograms. Syllabograms are used for morphological elements, but because of the nature of a syllabary, sign usage follows certain conventions and

does not render linguistic units precisely. A series of graphemes that we would transliterate as *he<sub>2</sub>-en-ĝar* could be transcribed as *he.i.n.ĝar* or as *he.n.ĝar.Ø*, depending on one's view of grammar, but not as *he.en.ĝar*.

## 2.5 Comparison of earlier and later systems

The differences between the nuclear early system and the fully developed second-millennium version of cuneiform can be illustrated by examples from a passage that is preserved in both versions. Here is a line from a third-millennium literary composition, followed by the manner in which the clichéd formula was written in Standard second-millennium Sumerian, a glossed version of the latter, and a translation (see Civil and Biggs 1966:12):

(2)	<i>Third millennium</i>	<sup>d</sup> en-ki	isimud	gu <sub>3</sub>	de <sub>2</sub>
	<i>Second millennium</i>	<sup>d</sup> en-ki-ke <sub>4</sub>	isimud-ra	gu <sub>3</sub>	mu-un-na-de <sub>2</sub> -e
	<i>Transcription</i>	Enkik.e	isimud.ra	gu.Ø	mu.na.de.e
		Enkik-ERG.	Isimud-DAT.	voice-ABS.	PREF.-DAT.-pour-NOM.
		“The god Enkik says to [his vizier] Isimud”			

An unusual writing in one such early text reveals that prefixes usually not expressed in writing could occasionally surface (Civil and Biggs 1966:3):

(3)	<i>Third millennium</i>	dur <sub>3</sub>	gu <sub>3</sub> -di	nab-sa <sub>10</sub> -sa <sub>10</sub>
	<i>Second millennium</i>	dur <sub>3</sub>	gu <sub>3</sub> -di	na-ab-ta-sa <sub>10</sub> -sa <sub>10</sub>
	<i>Transcription</i>	dur	gudi.Ø	na.b.ta.sa.sa
		ass	braying-ABS.	PREF.-PRO.-ABL.-buy
		“You should not buy a braying ass”		

By the beginning of the second millennium BC, the Standard Sumerian orthography had been established that would be used, with only minor adjustments, down to the very end of cuneiform writing.

In addition to the word- and morpheme-centered manner of writing, there exists a less stable and less formalized way of writing the language syllabically. Texts of this type, which first appear in northern Babylonia and peripheral areas in Old Babylonian times, write out free morphemes by means of syllabograms rather than by means of logograms. Thus for example, the Standard Sumerian sequence *sipa* <sup>d</sup>ur-<sup>d</sup>namma-ke<sub>4</sub> mu-na-an-šum<sub>2</sub> “he gave to the shepherd [king] Ur-Namma” is rendered as *si-pa ur-an-na-ma-ke mu-na-an-šu* in the so-called syllabic orthography. The five hundred or so texts of this type are mainly, but not exclusively, ritualistic.

## 3. PHONOLOGY

The phonology of the language is not well understood, and it is fair to say that it will never be fully recovered. There are many reasons for this; chief among them are the manner in which the language was encoded in writing, as well as modern misconceptions as to the nature of the script. Cuneiform was deciphered backwards, that is, it was first read in its latest incarnation, thousands of years after its origins. The Semitic Akkadian language was recovered first, and when Sumerian was discovered, it was read by means of sign values established for Akkadian. As a result, certain Sumerian phonemes that were not used in Akkadian were not initially identified. The repertoire of Sumerian phonemes currently

recognized still looks suspiciously close to the Akkadian repertoire; this may be due to chance, to our inability to recognize certain sounds, or to convergence of the two systems.

### 3.1 Consonants

The following chart presents the conservative current view of the Sumerian consonantal inventory.

#### (4) Sumerian consonantal phonemes

b	d	g
p	t	k
	s	š
	z	
		h
m	n	ḡ
	l	r
		ī

#### 3.1.1 Stops

Ambiguities in the use of the cuneiform script to write Sumerian and Akkadian have led to many debates about the nature of Sumerian stops. Observing the behavior of certain loans from Sumerian into Akkadian, Gelb (1961:33) argued against voiced stops in Sumerian and suggested that the distinction was between voiceless aspirated stops ( /p<sup>h</sup>/, /t<sup>h</sup>/, and /k<sup>h</sup>/ ) and voiceless unaspirated stops ( /p/, /t/, and /k/ ). Some have followed his hypothesis; Jacobsen (1957:92, n. 1) proposed that the opposition was between rounded and unrounded stops. There are serious flaws in these reconstructions, as noted by Rubio (1999a:141). For the present it seems most sensible to follow the traditional view and to argue for a voiced versus voiceless distinction. Civil (1973a:34) has observed that voiceless stops become voiced when they occur before an ending that begins with a vowel (*kalak*/*kalaga* “mighty”), although he also notes that the rule may have to be reversed.

The occurrence of a phonemic glottal stop /ʔ/ is uncertain. Spellings such as *sa-a* “cat” are commonly transcribed as *sa’a* (as if /saʔa/), but this is presently best seen as a Sumerological convention rather than a phonological claim.

#### 3.1.2 Sonorants

Sumerian has both nasal and liquid phonemes. The evidence for phonemic glides is less straightforward.

##### 3.1.2.1 Nasals

The writing system makes a clear distinction between /m/ and /n/. There is some uncertainty about their behavior in word-final position. Certain words ending in a nasal have a different consonant when followed by vocalic ending; thus *ezen* “festival” but *ezem-ma*. This variation may be interpreted as a change either of /n/ to /m/ before a vowel, or of /m/ to /n/ in word-final position.

The nasal /n/ also regularly becomes /l/ before /b/. This is commonly encountered in the verbal prefix chain when the prefix *nu-* is followed by *ba*/i- (written *la-ba-* or *li-bi<sub>2</sub>-*), but also within words as in the ES *la-bar* (EG *nağar*) “carpenter.” An unusual change of /l/ to /n/ before /g/ is found in early syllabic writings for the word *lugal* “king” (*nu-gal*). This,

however, may have to be interpreted as hypercorrection based on analogy with composites formed with *nu-* such as *nu-kiri*<sub>6</sub> “gardener” and so forth.

The identity of the phoneme commonly written *ḡ* is somewhat problematic (see Krecher 1978). As Civil (1973a:61) has noted, it is regularly only found before the vowels /a/, /i/, and /e/; it has variously been described as a velar nasal, a labiovelar nasal or as a nasalised labiovelar, and has been represented phonetically by notations such as /ŋ/, /ŋm/, or /ŋg/ (Black 1990:107–108). One should not exclude the possibility that Sumerian at one point had more than one such nasal – retroflex, palatal, as well as labial – as is the case, for example, in certain Dravidian languages.

### 3.1.2.2 Liquids

Because of certain writing conventions, Diakonoff (1967:49) proposed a phonemic distinction between the lateral liquids /l/ and velar /ɭ/. This has not gained wide acceptance. The phonological status of /l/ and /r/ is difficult to determine, and there are examples of an interchange of these phonemes in final and medial position (Civil 1973c: 174).

### 3.1.2.3 Glides

Standard transliterations of Sumerian do not recognize the existence of glides. Third-millennium texts from Syria, however, provide spellings that suggest the existence of a labial /w/, a palatal /y/ (and possibly one or two other sonorants; see Civil 1984:80).

## 3.1.3 Other consonants

Because of certain writing conventions, alterations, loans, and syllabic spellings, other phonemes have been suggested over the years. Civil (1973a) has drawn attention to the alternation of [g] and [b] in certain words, concluding that these spellings represent a distinct phoneme, either the labiovelar /g<sup>w</sup>/ or /gb/. The most widely debated extra phoneme of Sumerian has been variously notated as /dr/, /d<sup>r</sup>/, /ṛ/, and, most recently as [ts<sup>h</sup>] (Jagersma, 2002 [2005]). If the last-named is correct, it was an affricate that had disappeared early on from the language, but which in certain cases was reflected in historical spellings.

## 3.1.4 Apocope

It is generally assumed that word-final consonants are dropped, but it is unclear if this applies in all situations. Hence most CVC signs also have a CV transliteration: for example, the sign read as *šag*<sub>4</sub> “heart” by some, is read as *ša*<sub>3</sub> by others.

## 3.2 Vowels

The vowels of Sumerian correspond to those found in Akkadian:

### (5) Sumerian vowel phonemes

/i/	/u/
/e/	
/a/	

In Sumerian, however, unlike Akkadian, vowel length is not phonemic. Some have argued for the existence of a mid-back vowel /o/ (Lieberman 1979), but this has not found wide support. There is no evidence for the existence of diphthongs. In third-millennium texts



from the Syrian city of Ebla, certain words are unexpectedly written with final *-n*; this may be Semiticization or an indication of nasalization of final vowels in early Sumerian (Civil 1984:79).

### 3.2.1 Vowel harmony

Sumerian words show a very strong tendency towards vowel harmony, both within roots and morphophonologically, but the issue has never been analyzed in detail. Thus, many bisyllabic native words in the language repeat the same vowel: *kalam* “land,” *pirig* “lion,” or *murub<sub>4</sub>* “center.” Loans sometimes do conform to this tendency (e.g., *ugula* “captain, foreman” from Akkadian *waklu*), and sometimes do not (e.g., *akkil* “cry” from Akkadian *ikkilu*). Diakonoff (1983:87) thought that Sumerian had total vowel harmony, but as Boisson (1997:41) notes, no other language shows such a degree of harmony. It is probably safer to state that the language has a strong tendency towards harmony, but that the degree of the phenomenon may be masked by our transliteration system. There are many bisyllabic words with two different vowels, especially /a/ and /i/: for example, *agrig* “provider,” *gisal* “oar,” or *apin* “plow.” There are also bisyllabic words with other vowel sequences: for example, *dedal* “ashes,” *bugin* “bucket,” or *gizbun* “banquet.” Vowel harmony seems to operate strongly, but not totally, within the verbal prefix chain, but does not affect the stems, nor does it operate on nominal prefixes. Individual elements in compounds also retain their original vowels, as in *a<sub>2</sub>-tuku* “benefit, profit.”

## 3.3 Accent and intonation

Over the years there have been suggestions that Sumerian was a tonal language. The underlying assumption was that because the language had so many homophones, some additional distinctions were necessary, hence the tonal hypothesis. Many, but not all, Sumerian homophones are an illusion based on the system of transliteration (Parpola 1975). The only clearly identifiable prosodic feature is typologically predictable: rising phrase intonation to mark questions is sometimes expressed through the writing of additional vowels at the end of a clause.

## 4. MORPHOLOGY

### 4.1 Word formation

Sumerian distinguishes between nominal and verbal bases. The controversial category of adjectives will be discussed below; here it is assumed that most adjectives are verbs. The only recent discussions of Sumerian word formation are those of Diakonoff (1967:51–54), Kienast (1975), Schretter (1993), and Attinger (1993:155–158). This is a modified version of their analysis. One should bear in mind that the form of Sumerian words is sometimes obscured by inconsistent transliteration (on the CVC ~ CV transliteration variation, see §3.1.4).

#### 4.1.1 Basic Word Structure

Basic words were built on the following phonotactic patterns: (i) V (e.g., a “water”). There are few such roots. Most words transliterated as simple vowels are actually CV, such as *e<sub>2</sub>*



“house, temple, estate”, (/ha/) or a, “father”, /aya/ or /yaya/); (ii) CV (e.g., *ki*, “earth”); (iii) VC (e.g., *ud* “day”); (iv) VCV (e.g., *ama* “mother”); (v) VCVC (e.g., *amar* “calf”); (vi) VC<sub>1</sub>C<sub>1</sub>VC (e.g., *addir* “river crossing, wage”); (vii) CVC (e.g., *dub* “tablet”); CVCV (e.g., *gaba* “breast”); (ix) CVCVC (*munus* “woman”).

While the syllabic cuneiform script does not represent consonant clusters directly, heterogeneous clusters undoubtedly existed. In medial position one can recognize the following patterns: (i) CV<sub>1</sub>C<sub>1</sub>C<sub>2</sub>V<sub>1</sub>C (e.g., *kiskil*, “young woman,” written ki-sikil); (ii) (C)V<sub>1</sub>C<sub>1</sub>C<sub>2</sub>V<sub>2</sub>C (e.g., *ḡeštug* “ear, wisdom”); (iii) V<sub>1</sub>C<sub>1</sub>CV<sub>2</sub>C (e.g., *irkab*, “bat,” *adkin* “salted meat”). Initial and final clusters cannot be directly spelled out in cuneiform, but there are patterns of the type (i) C<sub>1</sub>C<sub>2</sub>VC (e.g., *lgud* “thick”) or (ii) CV<sub>1</sub>CV<sub>1</sub>C<sub>1</sub>C<sub>2</sub> (e.g., *kurušt* (*kurušda*) “ox fattener”).

#### 4.1.2 Compound forms

In addition to primary nouns and verbs, Sumerian has a rich repertoire of composite forms. For compound verbs see below §4.6. The least productive is a concatenation of two nouns. A form N<sub>2</sub> N<sub>1</sub> replaces the normal order of N<sub>1</sub> N<sub>2</sub>+gen. These are found only in poetry and are archaic or archaizing: for example. *an-ša(g)* “heavens + center” for “center-of-the-heavens.” Two nouns may also occur in normal order without genitive marker, as in *ereš-dingir*, “lady + god” for “priestess.”

Compound nouns are also formed from a noun a verbal/adjectival root such as *dub-sar*, “tablet + write” for “scribe.” In addition, nouns may be created from compound verbs without any affixes: *sa<sub>2</sub>-dug<sub>4</sub>*, “delivery.” Finally, nouns may be formed from frozen verbal forms: *u<sub>3</sub>-na-(a)-dug<sub>4</sub>* “letter,” literally “when you speak to him/her”; *ga-an-tuš* “tenant,” literally “I want to sit”; *ba-an-ḡi<sub>4</sub>* “answer,” literally “he/she answered.”

#### 4.1.3 Apophony

Apophony (or ablaut) may have played a limited role in word formation, but requires further study. At present it can be recognized in a small number of basic adjectives: for example, *gal/gul* “large/larger” (Civil 1982:12).

#### 4.1.4 Reduplication

Reduplication plays a highly restricted role in word formation. It appears that basic color terms share reduplicated stems: for example, *babbar* < *bar<sub>6</sub>-bar<sub>6</sub>* “white”; *kukku* < *ku<sub>10</sub>-ku<sub>10</sub>* “black”; and possibly *sig<sub>17</sub>-(sig<sub>17</sub>)* “blue/green” (Civil 1987:155). There is also a small class of echo words, nouns created by duplication with a vowel alternation (CV<sub>1</sub>C-CV<sub>2</sub>V), all restricted to the semantic class of noise: for example, *dum-dam...za* “to clamor”; *suh<sub>3</sub>-sah<sub>4</sub>...za* and so forth (Civil 1966). There are also isolated examples such as *nunuz* (< \**nuz-nuz*) “eggs” or of onomatopoetic words such as *zi...pa-an-pa-an* “to breathe.” The morpheme *-didli*, which means “one by one,” was originally *dil-dil* “one-one.” Reduplicated nouns and adjectives mark plurality (see §4.2.3), while reduplicated verb-stems can mark imperfect aspect and plurality of absolutes (see §4.6.3).

### 4.2 Nominal morphology

Sumerian nominal forms consist of a base and a series of affixes, primarily suffixes. The one prefix position is occupied by derivational morphemes; all other affixes come after the stem. Nouns are marked for gender (animate and inanimate), number, and case.

Although these affixes are ordered in a strict sequence when there is only a single noun, the matter is more complex when more than one is involved. In possessive constructions only the dependent noun takes a genitive marker: for example,

- (6) dumu lugal.ak  
son king-GEN.  
“The king’s son”

When two genitives are involved, the suffixes are added cumulatively (i.e., displaced) after the last noun. For example,

- (7) sa-a dumu lugal-la-ka  
sa’a dumu lugal.ak.ak  
cat son king-GEN.-GEN.  
“The cat of the son of the king”

In more complex sequences the affixes come at the end of a noun phrase; as a result, nouns that are within the phrase receive no marking at all. Sumerian is therefore a language with case displacement and globally final NP-marking, to use Aristar’s terminology (1995: 432, 445).

In schematic positional terms, the *noun chain* could be represented as follows (where PRO represents “possessive pronouns”):

(8) Sumerian noun chain

1	2	3	4	5	6	7
DERIVATIONAL	N <sub>1</sub>	N <sub>2</sub>	GEN.	PRO.	PL.	CASE
MORPHEMES						

#### 4.2.1 Derivational morphemes (position 1)

There are two derivational prefixes. The first, *nam-*, forms abstracts (e.g., *lugal* “king,” *nam-lugal* “kingship”); the second, *niĝ<sub>2</sub>-*, forms nouns out of verbs (e.g., *ba* “to bestow,” *niĝ<sub>2</sub>-ba* “gift”). The former presents few problems; the latter is more complicated.

Originally *niĝ<sub>2</sub>* was the inanimate relative pronoun. Many Sumerologists write that *niĝ<sub>2</sub>* is a noun meaning “thing,” but there is little to substantiate this claim. The prefix is used in ways that are not always clear to us and may have been lexicalized to some extent. It can be prefixed to certain adjectives such as *daĝal* “broad, wide,” but the difference between *daĝal* and *niĝ<sub>2</sub>-daĝal(a)* eludes us at present. One possibility is that this forms a superlative; if this is indeed the case, it was not generalized for all adjectives. More probable is that the forms with *niĝ<sub>2</sub>-* are no longer adjectives but are nouns, and therefore stand in possessive relationship with other nouns. Thus, the royal epithet *sipa gin.a* (*sipa gi-na*) means “just/true shepherd,” but *sipa niĝ.gin.ak.e* (*sipa niĝ<sub>2</sub>-gi-na-ke<sub>4</sub>*) means “shepherd of justice.” One should also note that there are a large number of *niĝ<sub>2</sub>*-compounds in Sumerian in which the element has no apparent semantic role.

Attinger (1993:155) does not consider the preceding to be derivational morphemes, arguing that only the prefix *nu-* serves this role. He follows the standard opinion, based primarily on etymological grounds, that *nam-* is a substantive derived from *me* “to be” and that *niĝ<sub>2</sub>* is a noun meaning “thing” that forms “concrete nouns.”

It is not clear if *nu-* should be viewed as a derivational morpheme or simply as a nominal formant. It is found in a small group of nouns denoting professions such as *nu-banda<sub>3</sub>* “captain” or *nu-kiri<sub>6</sub>* “gardener” (Edzard 1967). It is possible that the formant is related

to *lu*<sub>2</sub> “person, man.” The pronunciation with /n/ is indicated by loans into Akkadian such as *nukaribbu* and *laputtu* (with change of *n* > *l* / *\_\_* *b*; see §3.1.2.1). Early texts, however, indicate that *lu*<sub>2</sub> may have been pronounced as /nu/, as evidenced by such syllabic spellings as *nu-gal* for *lugal* “king” (etymologically, or folk etymologically, from *lu*<sub>2</sub> *gal* “great man”).

The formant *nam-* is also found in compound verbs (e.g., *nam...tar* “to decide fate”). Difficult to analyze are words such as *til* “life, to live, give health” which can function as verbs as well as nouns. These also create forms with the abstract prefix and it is difficult to distinguish the differences between *nam-til* and *til*.

#### 4.2.2 Possession (position 4)

A noun can be followed by an adjective (*lugal gal* “great king”), or by another noun in possessive relationship (Zólyomi 1996b). In that instance the second, possessed, noun, is marked by the suffix *-ak*. Thus, *lugal kalam.ak* “king of the land.” This is written as *lugal kalam-ma* in obedience to two rules: that in order to add a vocalic ending to a consonant-final root one use a CV sign, and the loss of final consonants. In rare instances there can be two or even three genitives, but no more than that. Note that the genitive *-ak* occupies a different position than the other case affixes.

There is another possessive construction in Sumerian that topicalizes the possessed noun. In the Sumerological literature this is called an *anticipatory genitive*; it is limited to literary texts and often results in tortured modern translations such as “the land – its king was.” The possessed noun is fronted and carries the genitive suffix; the possessor follows and is marked with a third-person possessive pronoun. Thus, with *lugal kalam.ak* “king of the land,” compare *kalam.ak lugal.bi* “the land’s king.”

#### 4.2.3 Number (position 6)

Singular is unmarked, but plurality can be expressed in a number of ways. Animate plural nouns take a suffix *-ene*, but there is no equivalent plural morpheme for inanimates. Hence an unmarked inanimate noun may be plural and the number is only marked by means of plural verbal agreement. The same holds true for collective nouns, such as *eren*<sub>2</sub> “troops” which take no plural marker but can trigger plural or collective verbal agreement.

If an animate or inanimate plural noun is followed by an adjective, the latter is reduplicated (e.g., *lugal / na*<sub>4</sub> *gal gal* “great kings/stones”); this can, in some animate cases, be combined with the plural suffix as in *lugal gal gal.ene* “great kings.” Plurality can also be expressed by reduplication of the stem, as in *lugal lugal* “kings.” It is commonly accepted that this signifies totality (i.e., “all kings”), but this remains to be fully documented. In addition, one encounters reduplicated nominals with the ending *-ene*, as in *lugal lugal.ene* “kings,” but the nuances of this formation elude us at present.

Two additional markers of plurality are usually cited: *-meš* and *-hi-a*. The ending *-meš* is the third-person plural copula, that is a form of the verb “to be”; *hi-a*, however, is not a plural marker at all, but an adjective meaning “mixed, of various sorts.” Thus, *udu hi-a* means not “sheep” (pl.) but rather “various types of smaller cattle.” Both have limited distribution, although the exact limits have not been studied. Since there is no formal morphological marker for inanimate plurals, the marker *-meš* may have developed from the copula to supplement the paradigm (as a sort of pseudo-morphological marker for paradigm leveling) and minimize ambiguity. It is commonly found in administrative lists and as a marker of plurality of Sumerograms in Akkadian texts, but is much less common in Sumerian narratives.

Since Akkadian used only morphological means of marking plurality, paradigm leveling may also account for the new composite plural morpheme *-bi.ene* that begins to appear in Old Babylonian literary texts. Thus, *iri.bi.ene* does not mean that the city was considered somehow metaphysically personified; it is simply a new way of expressing “cities.”

#### 4.2.4 Case (position 7)

Sumerian has two direct and seven oblique cases. With the exception of the equative, all of these are also marked on the verb, albeit the direct cases occupy different ranks from the obliques.

(9)	<i>Ergative</i>	-e	
	<i>Absolutive</i>	-Ø	
	<i>Dative</i>	-ra	<-ar/-ir/-ur>
	<i>Comitative</i>	-da	<-ta/-da <sub>5</sub> >
	<i>Ablative/Instrumental</i>	-ta	<-da>
	<i>Allative</i>	-(e)še	<-še <sub>3</sub> /-e <sub>3</sub> /-aš/-eš/-eš <sub>2</sub> /uš>
	<i>Equative</i>	-gin	<-gin <sub>7</sub> >
	<i>Locative 1</i>	-a	
	<i>Locative 2 (terminative)</i>	-e	

The ergative case marks the most agent-like argument of transitive clauses (corresponding to the transitive subject in English).

The absolutive case marks the patient of transitives (corresponding to English direct objects), as well as the single core argument of intransitives (corresponding to English intransitive subjects). The absolutive is also the citation form for nouns:

- (10) A. *lugal.e*    *iri.Ø*    *mu.n.hul.Ø*  
king-ERG.    city-ABS.    PREF.-erg-destroy-ABS.  
“The king destroyed the city”
- B. *Lugal.Ø*    *i.gin.Ø*  
king-ABS.    PREF.-go-ABS.  
“The king went”

The dative marks the beneficiary of an action (*lugal.ra* “for the king”) but also functions as a locative with animates (“upon the king”), in concert with the observations of Kuryłowicz (1964) and Aristar (1996) about the typological associations of datives with animates and locatives with inanimates. It also marks the secondary agent of causative constructions.

The comitative (or propriative) indicates accompaniment (*lugal.da* “with the king”).

The ablative case is also used in an instrumental manner (*tukul.ta* “by means of a weapon”) and with numbers it is used in a distributive sense (*min.ta* “two each”). The allative (usually called terminative in the literature) and the ablative denote movement towards (*iri.(e)še* “to/towards the city”) and away from a goal (*iri.ta* “from the city”), respectively.

The equative denotes comparison (*tukul.gin* “like a weapon”).

The locative 1 marks the inanimate place where an action takes place (*iri.a* “in the city”); while the locative 2, called locative-terminative by Sumerologists, marks propinquity (*iri.e* “next to the city”). The locative cases also mark the syntactic object of compound verbs (see §4.6.1); together with the allative they can also be used to mark the goal or object of certain verbs of affection and cognition.

There are some examples of idiomatic or verb-specific uses of certain cases with idiosyncratic meanings. In later Sumerian one sometimes encounters a redistribution of case functions under the influence of Akkadian. For example, the Akkadian preposition *ina* is both locative and instrumental, and under its influence Sumerian *-ta*, originally ablative and instrumental, acquires a locative meaning.

As is to be expected, low animacy nouns do not take ergative or dative; and high animacy nouns cannot take ablative/instrumental, allative, or locative suffixes.

In addition, Sumerian contains a set of discontinuous morphemes built by means of an initial word – often a body part – an optional bridging genitive morpheme, and a locative or directional case ending (*-a*, *-e*, *(e)še*, *-ta*). These can bracket nouns or nominalized clauses. Thus, for example, *bar eg-ba-ka* means “because of that ditch”:

- (11) *bar*                      *eg.bi.ak.a*  
       because of      ditch-PRO.-GEN.-LOC.

Body parts are *bar* “exterior” (“because of”); *da* “side” (“next to”); *igi* “eye” (“before”); *eĝer* “back” (“behind”); *murub*<sub>4</sub> “waist, middle” (“in the midst”); *šag*<sub>4</sub> “heart” (“inside”); *ugu* “forehead” (“before”); and *zag* “side” (“outside of”). A few other morphemes may also play this role, including *en-na*, of unknown origin (“until”); *ki* “earth” (“in, from”); *mu* “name” (“for”); and the abstract prefix *nam-* (“for the sake of”). These discontinuous morphemes allow for the spatial determination of animates, which as a rule cannot take the simple locative and allative case suffixes.

Diakonoff (1967: 56) lists *-ak.eš* as a case (he calls it causative); no other grammar does so. It is built by adding the allative to a bridging morpheme, which is the genitive. This properly belongs with the complex morphemes discussed above, as it is an abbreviation of *mu...-ak.(e)še* “because.”

## 4.2.5 Gender

Sumerian had two genders, animate and inanimate. The animate class covers humans and divinities, everything else is inanimate; perhaps one should use the terms “personal” and “impersonal.” Gender is not marked directly on the noun, but only surfaces in cross-reference, in pronouns, which are dominated by animates, and verbal concord.

## 4.3 Pronouns

As is to be expected in a head-marking language, the principal participants in an action are marked by affixes on Sumerian verbs, and therefore personal pronouns do not normally appear in sentences (Rhodes 1997). They are only used for emphasis, topicalization, and topic shift. Given the limited rhetorical range of Sumerian poetry, and the predominance of third-person narrative, it is not surprising that independent pronouns are relatively rare in the preserved texts, especially first- and second-person plural forms.

### 4.3.1 Personal pronouns

Unlike nouns, which show ergative case marking, independent personal pronouns can only be used as transitive and intransitive subjects, and thus have to be interpreted as nominative, albeit without any corresponding accusative form. The nominative marker is *-e*; it is possible that this is a deictic element (see Woods 2000 [2005]). In addition to nominative forms, personal pronouns have dative, terminative, comitative, and equative forms; as animates

they do not take local cases. Nothing is known about the inanimate third person, although it is possible that this function was fulfilled by *ur<sub>5</sub>* (or *ur<sub>5</sub>-bi*). As already noted, not all forms are attested. In addition to the normal forms encountered in texts, lexical texts (see §6) list compounds of singular and plural forms such as *za-e-me-en-ze<sub>2</sub>-en* for the second person. Such forms may simply be speculative grammatical constructions, or they may indicate that Sumerian originally had an inclusive/exclusive distinction that was incomprehensible to speakers of Akkadian. The personal pronouns are presented in (12) (OB = Old Babylonian):

(12)		<i>Singular</i>	<i>Plural</i>
	<i>Nominative</i>		
	1st	ġa <sub>2</sub> -e	me-(en)-de <sub>3</sub> -(en)
	2nd	za-e	me-en-ze <sub>2</sub> -en
	3rd	e-ne (pre-OB a-ne)	e-ne-ne
	<i>Dative</i>		
	1st	ġa <sub>2</sub> -a-ra/ar	
	2nd	za-a-ra/ar	
	3rd	e-ne-ra	e-ne-ne-ra
	<i>Comitative</i>		
	1st	(a/e)-da	
	2nd	za-(a/e)-da	
	3rd	e-ne-da	e-ne-ne-da
	<i>Terminative</i>		
	1st	ġa <sub>2</sub> -(a/e)-še <sub>3</sub>	
	2nd	za-(a/e)-še <sub>3</sub>	
	3rd	e-ne-še <sub>3</sub>	e-ne-ne-še <sub>3</sub>
	<i>Equative</i>		
	1st	ġa <sub>2</sub> -(a/e)-gin <sub>7</sub>	
	2nd	za-(a/e)-gin <sub>7</sub>	
	3rd	e-ne-gin <sub>7</sub>	e-ne-ne-gin <sub>7</sub>

#### 4.3.2 Possessive pronouns

Possessive pronouns affixed to nouns are etymologically related to the independent pronouns.

(13)		<i>Singular</i>	<i>Plural</i>
	<i>First</i>	-ġu <sub>10</sub>	-me
	<i>Second</i>	-zu	-zu-(e)-ne-(ne)
	<i>Third animate</i>	-a-ni	-a-ne-ne
	<i>Third inanimate</i>	-bi	-bi-(e-ne)

#### 4.3.3 Reflexive pronouns

Reflexive pronouns are not well attested. There is no ergative form. The base is *ni<sub>2</sub>-*, to which can be added possessive pronouns and case endings such as the locative. The absolutive paradigm is as follows:

(14)		<i>Singular</i>	<i>Plural</i>
	<i>First</i>	ni <sub>2</sub> -ġu <sub>10</sub>	
	<i>Second</i>	ni <sub>2</sub> -zu	
	<i>Third animate</i>	ni <sub>2</sub> -(te-a-ni)	ni <sub>2</sub> -te-a-ne-ne
	<i>Third inanimate</i>	ni <sub>2</sub> -bi	ni <sub>2</sub> -ba/bi-a

#### 4.3.4 Interrogative pronouns

Unlike personal pronouns, interrogatives work on the ergative pattern (for a different view see Huber 1996:186). In these pronouns the normal marking of animate with *n* and inanimate with *b* is reversed:

- (15) *Ergative*    a-ba-(a) “who?”  
           *Absolute*   a-ba “who?” a-na “what?”

Both pronouns can occur with suffixes. The animate form takes only the enclitic copula and personal pronouns. The inanimate form can be combined with certain postpositions, the copula, as well as possessive pronouns.

#### 4.3.5 Relative pronouns

Sumerian uses two substantives in the function of relative pronouns. Both are related to the derivational morphemes discussed in §4.2.1. The animate pronoun is *lu*<sub>2</sub>, literally “man, human,” as in *lu*<sub>2</sub> *e*<sub>2</sub> *du*<sub>3</sub>-a “who built the temple.” The inanimate equivalent, which is often translated as “thing,” although the etymology may be questioned is *niĝ*<sub>2</sub>: *niĝ*<sub>2</sub>-*du*<sub>11</sub>-*ga*-*ni* (*niĝ.dug.ani*) “what he/she said.”

### 4.4 Adjectives

No proper study of adjectives exists; recent grammars contain limited information on this category (Thomsen 1984:53–65; Attinger 1993:167–168). The only preliminary study is Black (2002 [2005]). It is generally agreed that Sumerian had only a limited number of “true” adjectives and that most are uninflected verbs with the nominalizer *-a* (there is a complex debate on this issue; see, most recently, Krecher 1993, Schretter 1996). There are only a handful of adjectives that are not attested as verbal roots, and, for lack of a better analysis, one should maintain that all Sumerian adjectives are in fact verbs (Gragg 1968). In form, adjectives are bare uninflected verbal roots followed by  $\emptyset$  or by *-a*. This suggests that at a certain level they are simply reduced predicates. The distribution of these two forms is not clear. Most adjectives appear in one or the other, but some are attested in both forms.

Certain adjectival constructions are unclear at present. A small group of adjectives carries the derivational prefixes *niĝ*<sub>2</sub>- and *nam*- (see §4.2.1). We do not know what the difference is between *daĝal(a)* “wide, teeming” and *niĝ*<sub>2</sub>-*daĝal(a)*, or between *kas dug*<sub>3</sub> “sweet beer” and *kas niĝ*<sub>2</sub>-*dug*<sub>3</sub>. Since *niĝ*<sub>2</sub>- usually makes nouns out of verbs, this may be construed as a nominal construction. It is also conceivable that *niĝ*<sub>2</sub>- is here the inanimate relative pronoun and that this is a calque from Akkadian.

### 4.5 Adverbs

Sumerian adverbs are formed from nominal and verbal bases. Most commonly they are formed with a suffix *-bi* (originally probably an inanimate deictic) which can only be added to verbal (“adjectival”) roots, either directly or following the nominalizing suffix *-a*: for example, *gal-bi* “greatly,” *dug*<sub>3</sub>-*bi* “tenderly,” *gibil-bi* “anew,” or *ul*<sub>4</sub>-*la-bi* “rapidly.” A different suffix *-(e)še*, homonymous with the allative case, created manner adverbs from nouns as well as adjectives: thus, *u*<sub>4</sub>-*de-eš*<sub>(2)</sub> “as the day,” *gal-le-eš* “grandly.” In Old Babylonian



texts one begins to encounter the cumulative use of both suffixes as in *gibil-bi-eš<sub>3</sub>* “anew.” In some cases, adjectives can be used as adverbs without any suffix, such as *gal* “great” but also “greatly” (Krecher 1987:74). A postulated class of adverbs in *-a* has been questioned (Attinger 1993:170).

#### 4.5.1 Modal and temporal adverbs

The most common modal adverbs are the following: *i<sub>3</sub>-gi<sub>4</sub>-in-zu* “moreover, what’s more”; *i<sub>3</sub>-ge<sub>4</sub>-en* “truly, in fact”; *a-na-aš-am<sub>3</sub>*, *a<sub>2</sub>-še<sub>3</sub>* “how is it (that).” Temporal adverbs are as follows: *a-da-lam* (*a-da-al*, *i-da-al*) “(but) now”; and *i<sub>3</sub>-ne-eš<sub>2</sub>* “now.”

#### 4.5.2 Interrogative adverbs

These consist of a stem *me(n)*, complemented by directional suffixes or the enclitic copula. The most common forms are these: *me-a* “where?” *me-še<sub>3</sub>* “where to?” and *me-na-am<sub>3</sub>* “when?”

### 4.6 Verbal morphology

The analysis of verbal structure is the most controversial part of modern Sumerian grammatical study. It was also of concern to Akkadian-speaking ancients, who compiled comparative paradigms of Sumerian and Akkadian verbal forms and attempted to isolate morphological elements that they considered equivalent to ones found in their own language (Black 1984). It would be impossible to give an adequate accounting of all competing visions of the Sumerian verb in the present context; what follows is my own relatively simple analysis with selective references to competing theories. For fuller bibliographical information see Thomsen (1984), Attinger (1993), and Römer (1999).

Sumerian verbs consist of a verbal root and morphological affixes that mark certain verbal categories. The affixes mark categories such as mood, concord, and aspect. Verbs are either simple or compound. In certain verbs the base may be reduplicated to mark the imperfective, iterative action, or plurality of patient.

Compound verbs are construed with a noun and an inflected verbal base (Karahashi 2000). The noun is inanimate, indefinite, and generic; it is the semantic patient of the verb but it does not constitute a core argument of a clause, hence it is not marked by a direct case ending. The direct object of the clause is marked as oblique, usually with the locative 2 *-e*, less often with locative 1 *-a*, and with dative *-ra* on a small group of verbs, most of them verbs of emotion, and with still other cases. A good example is the verb *in-(še<sub>3</sub>) . . . dub<sub>2</sub>* “to insult” which takes the dative, although the verb takes the locative rather than the dative prefix:

- (16) 

ud-bi-a	gi	gīš-ra	in-še <sub>3</sub>	mu-ni-in-dub <sub>2</sub>
ud.bi.a	gi.(e)	gīš.ra	in.še	mu.ni.n.dub
day-PRO.-LOC.	reed-(ERG.)	tree-DAT.	N.-ALL.	PREF.-LOC2-ERG.-insult

“Then (lit. ‘on that day’) Reed insulted Tree”

Many compound verbs have transparent etymologies, such as *ki* “earth” + *tag* “strike, touch” = “to lay a foundation, to spread.” The incorporated noun is sometimes a body part, *šu* “hand” or *ka* “mouth.” Others consist of a noun and an auxiliary verbal root such as *dug<sub>4</sub>* “to speak” or *ak* “to make,” verbs which otherwise appear independently. Some verbs of this type may be doubly compounded with an auxiliary and it is unclear if this has any semantic



consequences; thus *šu . . . bal* and *šu bal . . . ak* both mean “to overturn.” A substantial group of compound verbs has no apparent etymological transparency, such as *ki . . . aḡ<sub>2</sub>* “to love” (lit. “place” + “to measure out”). Small subsets allow for expansion of the noun by an adjective (e.g., *šu zi . . . ḡar* “hand” + “true” . . . “place” = “to bestow, grant”). One has the impression that by the time we actually observe the language, noun–verb compounding was no longer productive. A frozen set had entered the lexicon, but new verbs were not being created.

Attinger (1993) has suggested that compound verbs are an example of noun incorporation, a phenomenon attested in many languages of the Americas, Southeast Asia, and elsewhere (Mithun 1984, 1985). Some have denied this, arguing that in Sumerian this is a syntactic and not a morphological issue (Zólyomi 1996a), but this is a theoretical question that covers all of noun incorporation. Huber (1996) likewise comes out against incorporation in this language, but once again it is a definitional question. The Sumerian data suggest either what has been termed loose incorporation (Mithun 2000) or, more probably, what Miner (1986) calls “noun stripping.” In such constructions the nouns are “stripped” of their affixes but remain as separate phonological entities; the nouns are backgrounded but remain as independent words.

#### 4.6.1 Transitivity

Most Sumerian verbs are strictly transitive or intransitive. There exists a small class of labile, or ambitransitive, verbs that can be either transitive or intransitive. Examples are *gu<sub>7</sub>* “to eat ~ to feed”; *naḡ* “to drink, give to drink, water”; *uṣ<sub>2</sub>* “to kill ~ to die”; *tuṣ* “to sit ~ to seat”; *kud<sup>r</sup>* “to enter, bring in”; *us<sub>2</sub>* “to follow, reach, let reach.” Two such verbs are semantically similar, but differ in the animacy of the subject/patient: *til* “to live, dwell, be healthy ~ to settle, give life/health” used when people are involved; and *lug* “to pasture, settle” which is used for animals. One should note, however, that *til* can be used of inanimates with the meaning “to be/make healthy.”

#### 4.6.2 Valence

Matters of valence in Sumerian have been disputed, but no consensus has been reached. It is clear that simply deleting the agent can form impersonal passives; as a consequence, this often results in a change of verbal prefixes, but there is no specific passive marker as such. The existence of other forms of valence change mechanisms, be it antipassive or causative, is difficult to ascertain at present (see, most recently, Attinger 1993:195–199, though most of his examples are actually labile verbs).

#### 4.6.3 Aspect/Tense

Opinion is divided on whether the two forms of the Sumerian verb differ in tense or in aspect, although in recent years most scholars have come to speak of the latter rather than the former. Certain verbs utilize stem reduplication to create one of the forms, and therefore typologically it is unlikely that tense is involved (see Anderson 1985:170). For the sake of the present discussion we shall use the terms *perfective* and *imperfective* to designate these two forms; one could also designate them as *completive* and *incompletive* since the only thing that most scholars agree on is that one denotes a complete and the second an incomplete action.

Ancient lexical and grammatical texts provide us with the Akkadian names of the two basic verbal forms: *hamṭu* and *marû*. There has been much discussion of the exact meaning of these words as well as of whether these technical terms describe the Sumerian verbal forms

or their Akkadian translations. Uncertainties aside, the terms have often been used in the modern literature in order to avoid labeling the specific aspectual or temporal qualities of the Sumerian verb. It now appears fairly certain that these Akkadian grammatical terms means simply “short” and “long” (Civil 2002: 69–100) and that the perfective (i.e., “short”) form was considered the unmarked citation category. At the present time the full significance of the two forms is open to debate and the use of “perfective” and “imperfective” here is purely conventional.

#### 4.6.3.1 Marking of aspect

Verbs mark these distinctions in three separate ways: through (i) agreement, (ii) stem reduplication, and (iii) suppletion. The perfective is the unmarked aspect and the perfective stem is the citation form. Reduplication and suppletion also serve to mark the plural of absolutes, that is, plural intransitive subjects and transitive objects; Sumerologists refer to this as *free reduplication*. Most verbs achieve this by means of stem reduplication, but a small class of verbs has suppletive plural forms. On rare occasions imperfective verbal roots can be tripled or even quadrupled to mark plurality of absolutes; with perfects this marks both intense action and plurality of absolutes.

There has been some disagreement concerning the marking of the two aspects. Yoshikawa (1968) in a pioneering study proposed three classes of verbs: those that formed the imperfective by affixation (-e); by reduplication; and by alternation of roots. It seems fairly certain, however, that there is no affixation group, and that the suffix -e belongs to the agreement-markers (Thomsen 1984:116).

More than half of Sumerian verbs have no overt aspectual morphology; the distinctions are expressed by means of different agreement patterns for the two aspects (e.g., *šum* “to give,” *dal* “to fly”). A much smaller group of verbs utilizes partial or full stem reduplication to form the imperfective. The writing system makes it difficult to discern when a root is fully or partially reduplicated, but as a rule CV and VC roots are fully copied (e.g., *si* ~ *si-si* “to fill,” *ur* ~ *ur.ur* “to drag”), while CVC roots are reduced to CV (*gar* ~ *ga.ga* “to place,” but cf. *gar* ~ *gar-gar*, “to pile up”), although the final consonant may resurface before a vocalic ending. Often this is not written, but forms such as *ga<sub>2</sub>-ga<sub>2</sub>-am<sub>3</sub>* illustrate the principle well. A very small class of verbs displays root suppletion for aspect as well as number (Steinkeller 1979). As a result, one can say that there were two “regular” ways of distinguishing aspect in Sumerian: through agreement and by stem reduplication.

#### 4.6.3.2 Regular verbs

Regular verbs may be represented as follows, utilizing the *šum* (written <šum<sub>2</sub>>) “to give,” *gi* (written <gi<sub>4</sub>>) “to return,” and *gar* “to place”:

(17)	<i>Perfective</i>		<i>Imperfective</i>	
	šum	<šum <sub>2</sub> >	šum	<šum <sub>2</sub> >
	gi	<gi <sub>4</sub> >	gi.gi	<gi <sub>4</sub> -gi <sub>4</sub> >

Superficially, it would seem that there was also a reduced reduplication group:

(18)	<i>Perfective</i>		<i>Imperfective</i>
	ḡar	<ḡar>	ḡa.ḡa    <ḡa <sub>2</sub> -ḡa <sub>2</sub> >

Although it remains to be fully demonstrated, it is most probable that all CVC verbs copied only CV in reduplication, although this is often obscured by the writing system. Thus, the reduplication of *gar* is written as <ga<sub>2</sub>-ga<sub>2</sub>>, but the reduplication of *kin* “to seek” is written as <kin-kin>, which must be read as *ki<sub>3</sub>-ki<sub>3</sub>*.

### 4.6.3.3 Suppletive verbs

The suppletive verbs are similar in meaning to such verbs found in unrelated languages, including many North American tongues. Most of them are intransitive or labile. The complex paradigms of these verbs began to conform to the regular verbs already at the end of the third millennium, when singular roots began to replace the plural forms. For comparative purposes it is necessary to list these Sumerian verbs in full.

(19) Verb	[Perfective]		[Imperfective]	
	Singular	Plural	Singular	Plural
“to bring”	<i>de</i> <sub>6</sub>	<i>lah</i> <sub>4</sub>	<i>tum</i> <sub>2/3</sub>	<i>lah</i> <sub>4</sub>
“to go”	<i>gin</i>	<i>er</i>	<i>du</i>	<i>su</i> <sub>8</sub> -( <i>b</i> )
“to stand”	<i>gub</i>	<i>su</i> <sub>8</sub> -( <i>g</i> )	<i>gub</i>	<i>su</i> <sub>8</sub> -( <i>g</i> )
“to sit”	<i>tuš</i>	<i>durun</i>	<i>dur</i> <sub>2</sub>	<i>durun</i>
“to speak”	<i>dug</i> <sub>4</sub>	<i>e</i>	<i>e</i>	<i>e</i>
“to kill/die”	<i>uš</i> <sub>2</sub>	<i>ug</i> <sub>5/7</sub>	<i>ug</i> <sub>5/7</sub>	<i>ug</i> <sub>5/7</sub>
“to live, be healthy, dwell” (animate)	<i>til</i>	<i>še</i> <sub>x</sub> ( <i>SIG</i> <sub>7</sub> )	<i>til</i>	<i>še</i> <sub>x</sub> ( <i>SIG</i> <sub>7</sub> )
“to live, dwell, pasture” (inanimate)	<i>lug</i>	<i>še</i> <sub>x</sub> ( <i>SIG</i> <sub>7</sub> )	<i>lug</i>	<i>še</i> <sub>x</sub> ( <i>SIG</i> <sub>7</sub> )
“to enter, bring in”	<i>ku</i> <sub>4</sub>	<i>sun</i> <sub>5</sub>	<i>ku</i> <sub>4</sub> - <i>ku</i> <sub>4</sub>	<i>sun</i> <sub>5</sub>

Three other verbs have a limited form of suppletion that consists of adding a final consonant in the imperfect: *e*<sub>3</sub> ~ *e*<sub>3</sub> [*d*] “to go out”; *ri* ~ *rig* “to pour out”; *ti/e* ~ *teḡ* “to approach.” This set of three is commonly referred to as an *alternating class*, but the limited number of verbs obviates the creation of a separate fundamental category.

In the simplest terms, the Sumerian verb may be represented in the following manner:

### (20) Sumerian verbal chain

1	2	3	4	5
MOOD	CONJUNCTION	FOCUS	INDIRECT OBJECT	DIMENSIONAL PREFIXES
6	7	8	9	10
AGREEMENT	ROOT	ED	AGREEMENT	NOMINALIZATION

### 4.6.4 Mood (position 1)

The traditional description of modes distinguishes between pairs of homophonous prefixes that differ in meaning depending on the mood. Thus *he*- is “precativ” with the imperfective, but “assertative” with the perfective. As a result, translations of texts are replete with “let him/her” and “verily he/she . . .” There are reasons to reject this interpretation; certain modal prefixes are indeed usually associated with one aspect or the other, but this results from the semantics of the mode and not from any formal constraints. The following reinterpretation of the modes results in part from the author’s own observations, but mainly from the work of Civil (2002 [2005]) which obviates much earlier research on the subject.

Unlike previous writers, Civil makes reference to *deontic* and *epistemic* notions of modality (Palmer 1986). To cite Chung and Timberlake (1985:246): “The epistemic mode deals with alternative worlds with respect to a given world at a given time point; the alternative worlds are those that could exist instead of a given world. The deontic mode also deals with a given world and with alternative worlds, but the alternative worlds are those that could develop out of the given worlds.” In Sumerian, deontic functions are distributed over four

forms, the deontic subjunctive-optative, both negative and positive, the cohortative, as well as the imperative. A variety of epistemic functions are encoded by the positive and negative epistemic subjunctive-optative markers.

#### 4.6.4.1 Indicative

The normal indicative has no prefix in this position; the negative carries the prefix *nu-*. Thus, *lugal-e iri mu-un-gul* “the king destroyed the city” but *lugal-e iri nu-mu-un-gul* “the king did not destroy the city.” There are also rare cases of *nu* as a predicate, as in *lu<sub>2</sub>-še<sub>3</sub> lugal-ġu<sub>10</sub> in-nu* “that man yonder is not my king.”

#### 4.6.4.2 Deontic subjunctive-optative

This prefix is used to make commands, give advice, or exhort someone to do the speaker's bidding, or to express the desires and wishes of the speaker. This results in phrases with counterparts to English “should,” “please,” or “may.” The positive prefix is *he-*, written with the sign *he<sub>2</sub>*, although from Old Babylonian times on the writing shows vowel harmony with what follows (written *ha* or *hu*), and the negative is *na-*.

#### 4.6.4.3 Epistemic subjunctive-optative

This function expresses conditions dependent on actions from another clause or phrase, often resulting in dependent clauses or conditionals. The positive prefix is *he-*; the negative is *bara-* (written as *ba-ra-*).

The subjunctive-optative modals are treated somewhat differently in traditional grammars, which correlate four different prefixes with the two aspects of the Sumerian verb, here marked as p(erfective) and i(mperfective). Thomsen (1984:193–199) is representative. In this system *he-* is affirmative (p) or precative (i); the negative counterpart is *bara-*, which is negative affirmative (p), or vetitive (i) for first person, otherwise it is prohibitive *na-*, also with the imperfect. The prefix *na-* (see §4.6.4.6) with perfect aspect is affirmative.

#### 4.6.4.4 Cohortative

The prefix *ga-* renders the intent or willful pronouncement of the speaker: for example, *ga-na-ab-dug<sub>4</sub>* (*ga.na.b.dug.Ø*) “I have decided to tell it (=b) to him myself.” Such forms almost always use the perfect aspect, but agreement (see §4.6.10) is nominative-accusative, rather than ergative (Michalowski 1980:97). The prefix marks the accusative rather than the ergative, as is usual in the perfective. During the Old Babylonian period a first-person plural form appears, with imperfective aspect and the first-person plural ending *-enden* marking the nominative: *ga-mu-na-dur<sub>2</sub>-ru-ne-en-de<sub>3</sub>-en* (*ga.mu.na.durun.enden*) “We want to prostrate ourselves before him!”

#### 4.6.4.5 Prefix of anteriority

The prefix *u-*, often written with the sign *u<sub>3</sub>*, marks an action that precedes another action in a sequence. Such forms are usually translated as temporal clauses “when . . .”; in bilingual texts they are often rendered by imperatives. Traditionally such constructions are labeled prospective.

#### 4.6.4.6 Other modal prefixes

Civil calls the prefix *na-* a marker of reported speech. In earlier treatments it is regarded as an affirmative (volitive) marker. Although it seems to be a homonym of the negative

subjunctive-optative (see §4.6.4.2), it may in fact have originally had a different phonological shape. Unlike the negative prefix, this *na-* is usually combined with the perfective aspect. It is often found in contexts where traditional or mythological lore is reported, or in formulaic introductions to narratives and speeches. It is best illustrated by the standard opening formula of Sumerian letters of the late third millennium: *PN<sub>1</sub>-ra u<sub>3</sub>-na-(a)-dug<sub>4</sub> PN<sub>2</sub> na-(ab)-be<sub>2</sub>-a* “When you address PN<sub>1</sub>, this is what PN<sub>2</sub> says to him.”

There is some evidence, however, that this prefix had other functions before the Ur III period. In third-millennium literary texts *na-* is one of the few prefixes that are regularly written before the verbal root, often with the sign *nam<sub>2</sub>*, and it is used much more commonly than in later periods. This grapheme goes out of use in the second millennium, when it is merged, together with some other similar signs, into *še<sub>3</sub>*. One could speculate that originally *na-* had a narrative foregrounding function that was lost in later Sumerian. The fact that it is apparently homophonous with the negative subjunctive-optative raises additional questions. It may be that this is a historical accident, but it is also possible that the consonants of the two prefixes were different.

Another uncertain modal prefix is *ša-* (Jacobsen 1965:73 called it “contrapunctive”). It is documented only in literary texts. As Civil notes, the distribution of this prefix is somewhat puzzling, as a third of occurrences in the middle second-millennium school curriculum are limited to four compositions. It is not perhaps accidental that one of these, *The Instructions of Shuruppak*, is attested already in Early Dynastic copies, and the second, *The Collection of Temple Hymns*, is ascribed to a princess who lived c. 2300 BC. It is possible that two different processes resulted in two different written forms of the same grammatical element, or even in the split of one into two: a change in meaning of *na-* and the misreading of the sign *nam<sub>2</sub>* as *še<sub>3</sub>*.

A rare modal prefix, found only in literary texts, is *nu-uš-*, charmingly named “frustrative” by Jacobsen (1965:82), and apparently means “if only, would that.” Civil considers it a rhetorical interrogative particle, meaning “why not?”

#### 4.6.4.7 Imperative

The morphology of the imperative in Sumerian is completely different from that of other moods, and is not marked by any characteristic affix. Copying the root to the front of the verbal form, which is always the perfective singular root, creates imperatives: thus *mu lugal mu-ni-in-pad<sub>3</sub> (mu.ni.n.pad)* “He/she wore by the name of the king”; but *mu lugal pad<sub>3</sub>-mu-ni-ib<sub>2</sub> (pad.mu.ni.b)* “take the oath by the name of the king!” The agreement prefix *b-*, now moved after the root, in the imperative always marks the accusative, that is, the transitive object; in the corresponding indicative sentence *n-* marked the agent.

The unmarked singular second-person referent of the imperative is always nominative, that is, either transitive or intransitive subject. In early texts, this is always deleted; in the Old Babylonian times an overt plural form was created by analogy with the cohortative, resulting in forms such as *du<sub>11</sub>-ga-na-ab-ze<sub>2</sub>-en (dug.a.ba.b.enzen)* “you all say it!”

The nominative/accusative agreement pattern of the imperative is not surprising; this is a pragmatic universal (Michalowski 1980:97; Payne 1982:90). Note that the last form cited above has the vowel *a* after the root. This can be interpreted either as an insertion to avoid a cluster or confusion with infinitives, or as an allomorph of the conjugation prefix *i-*. The latter otherwise never occurs in imperatives. The few attested forms of the type *gar-i<sub>3</sub>* are probably to be interpreted as *gar.(a)ni* “when he/she placed” and are not imperatives at all (Attinger 1993:299). Other examples of imperatives are *dug<sub>4</sub>-ga-na-ab* “say it to him/her!”, and *tuš-a* “sit!”

#### 4.6.5 Conjunction (position 2)

The second rank is occupied by the conjunction prefix *inga-*, which means “as well, also, too.” The rank of the prefix has been the subject of some debate; it comes after the modals, but is rarely followed by conjugation prefixes. Writings such as *nam-ga-* are probably to be analyzed as *na.(i)nga*.

#### 4.6.6 Focus (conjugation) prefixes (position 3)

The prefixes that fall in this position constitute the most controversial part of Sumerian grammar. No two Sumerologists appear to agree fully on their form, meaning, etymology, and identity; the number of ranks that they occupy is equally disputed. It would be impossible to do justice in this short survey to the various opinions that have been expressed. I have therefore chosen to present my own working hypotheses on the subject and only mention selected previous opinions on the matter. For the numerous interpretations of these prefixes see the references offered by Thomsen (1984:182–185), with important newer discussions by Black (1986:77), Wilcke (1988), Attinger (1993:261–288), Jagersma (1993), as well as a study by Vanstiphout (1985) on foregrounding and backgrounding strategies in Sumerian.

Rather than split these prefixes into three, four, or even five separate ranks, I prefer a minimalist position according to which there are only four distinct “conjugation” prefixes: *mu-*, *ba-*, *i-* (or *V-*), and *imma-*. Gragg (1973a:93) and Civil (in Karahashi, 2002 [2005]) apparently take similar positions. I do not break these down into smaller components, as do many others. Most Sumerologists consider this position obligatory, and restore a hypothetical *i-* even in cases when it is not written. In my opinion the neutral *i-* is not marked after a modal prefix. Rather than consider the position obligatory, one should simply state that a finite verbal form cannot begin with any of the final three positions before the root.

The prefix *imma-* is most commonly considered as a compound, often etymologized as containing both *i-* and *b(a)-* as well as a locative element *a*. According to the analysis followed here, the first two are mutually exclusive and the third element does not exist. Rather than view *imma-* as a “compound” I would suggest that it represents a form of reduplication of *mu-*, in which the initial consonant is copied and the cluster is reinforced by an initial vowel.

The meanings of these prefixes are as contested as their ranking. The prefix *mu-* appears to mark focus on control over an action that is within the control and propinquity of the agent. When such control is loosened, absent – and this includes the absence of an agent in a clause – the prefix *ba-* is used. When the focus is intensified, as with verbs denoting movement towards the agent, or the agent manipulates an object, such as a tool, the prefix *imma-* is often used. When focus is not specified, the prefix is *i-*. There is a rare prefix *a-*; in Old Babylonian literary texts it is probably an allomorph of *i-*, but in earlier texts it seems to be used, in Nippur at least, to mark verbs without agents. Yoshikawa (1992) considers *ba-* to mark reduced valency, which may fit well into this scheme.

I must reiterate the contested nature of these issues. The reader should be aware that there are many graphemic and morphophonological matters that remain unresolved. For example, a sequence such as *im-ROOT* or *i<sub>3</sub>-im-ROOT* has been interpreted as *i* followed by a “ventive” prefix that signifies “hither.” I much prefer to view the *m* as a reflex of *n* (the animate third-person pronoun); it is also possible that there are other morphophonemic or even prosodic processes at play here that are represented by the extra vowel, but this is a complex issue that cannot be debated in the present work. One should also note that

the writing conventions as well as the forms of these prefixes show much synchronic and diachronic variation.

#### 4.6.7 The prefix *al-*

There exists another verbal prefix of undetermined rank, namely *al-*. The rank cannot be specified because, with rare exceptions, this morpheme cannot coexist with any other verbal affix, although such forms can be nominalized. The forms with *al-* are intransitive, and appear to correspond to Akkadian inflected verbal adjectives (“statives”).

Attinger (1993: 269) and Edzard (2003: 111) deny the existence of a separate prefix and consider *al-* to be an allomorph of the vocalic prefix *a-*.

#### 4.6.8 Indirect object (position 4)

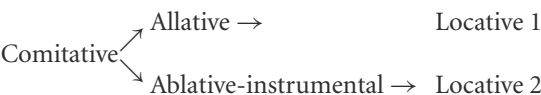
The dative prefixes are normally classed together with the dimensional elements of the next position. For structural reasons they are set apart here in their own rank. The dimensional prefixes, when they do refer to arguments of a clause, mark adjuncts; this position, however, cross-references the beneficiary, that is, a core or core extension argument. Unlike the markers that correspond to the oblique cases of nouns (dimensional prefixes), datives have different forms for different persons:

- (21) *First*      *a*      *me*  
           *Second*   *ra*    ?  
           *Third*    *na*   *ne*

The first person always follows the prefix *mu-* and together they are realized in writing as *ma-*, as in *ma-an-šum<sub>2</sub>* (*mu.a.n.šum.Ø*) “he/she gave [it] to me”; or *ma-an-dug<sub>4</sub>* (*mu.a.n.dug.Ø*) “he/she said [it] to me.” The second person is also found after *mu-*; in early texts this sequence is subject to vowel harmony, and is usually, but not always, written as *ma-ra*.

#### 4.6.9 Dimensional prefixes (position 5)

The forms and meanings of the prefixes that occupy this rank are fairly well established, due to a great extent to the work of Gragg (1973a). These prefixes are coreferential with the oblique case marking of the noun: dative, comitative, ablative-instrumental, allative, and the two locative cases. Most of them are phonologically similar or identical to those of the noun, and are presumably of the same etymological origin. As Gragg has shown, these prefixes are often connected to certain roots and are lexicalized to a degree. The dative and locatives differ in certain respects from the other prefixes and may have a different common origin. The dimensional prefixes follow one another in a set order:

- (22) 

##### 4.6.9.1 *Comitative*

The comitative prefix is usually written as *da* (Old Sumerian *da<sub>5</sub>*); sometimes as *di*, *de<sub>3</sub>*, and *de<sub>4</sub>* when followed by the locative 2 (terminative). The affix can be preceded by a pronominal element: first person is either *Ø* or *e*, second is *e*, and third person is *n* or *b*, for animates and



inanimates respectively. In the plural, only the third person is attested; this element, written as PI, with unknown reading, is only found in Old Sumerian documents.

Although homophonous with the equivalent nominal suffix, in most instances the verbal prefix does not copy a corresponding marker appearing on a noun. Sometimes, especially after the prefix *ba-*, *da-* must be interpreted as a writing for the ablative *ta-* (see §4.6.9.3). The comitative occurs with verbs that include the semantic notion of accompaniment, such as “to speak with” (*dug*<sub>4</sub>) “to compare with” (*sa*<sub>2</sub>), or “to counsel with” (*ad...ḡi*<sub>4</sub>), as well as with verbs of emotion. Still other verbs take this prefix, including those meaning to “flee,” “to escape” (*zah*<sub>3</sub>). An important function of this prefix is the marking of potential – referred to as abilitative in the literature.

#### 4.6.9.2 Allative (*terminative, directive*)

The allative prefix was originally written as *še*<sub>3</sub>-, but beginning with the Ur III period it was expressed by means of the sign *ši*-. Unlike the comitative, it is closely related semantically to the nominal suffix, and denotes movement towards a goal. It is therefore frequently found on verbs of motion and often marks nuances of meaning that are connected with its basic function. It is also often found with compound verbs that denote attention; these all include body parts as the “stripped” noun (see §4.7.1). Examples are *igi...ḡar/du*<sub>3</sub>/*kar*<sub>2</sub>, which all denote different ways of seeing and include the noun *igi* “eye”; and *ḡeštug*<sub>2</sub>...*gub/ḡar* “to pay attention to/listen” incorporating *ḡeštug*<sub>2</sub> “ear.”

#### 4.6.9.3 Ablative-instrumental

The primary meaning of this prefix, *ta-*, is ablative, although there are rare cases of instrumental usage. There is another prefix *ra-* that obviously also has ablative meaning and has been the subject of some debate. In Ur III documents one finds the expression *ud-ta ud x ba-zal* ~ *ba-ra-zal* ~ *ba-ta-zal* ~ *ba-ra-ta-zal*, which means “at the end of the xth day,” literally “from the month the xth day having passed.” This has created much confusion about the possible existence of two such prefixes, but Civil (1973b:27) ingeniously suggested that these writings all express the realization of /ta/ as /dʳ/ in intervocalic position.

#### 4.6.9.4 Locative 1

The locative prefix *ni-* corresponds to the nominal locative case ending *-a*. Unlike the already discussed dimensional prefixes, it primarily resumes locatives in the clause. This includes true locative adverbials as well as the logical direct objects of compound verbs and third participants of causative constructions. It is sometimes written as *in-* immediately before the root, and is thus confused with the third-person ergative marker (see §4.6.10; Attinger 1993:234).

#### 4.6.9.5 Locative 2 (*locative-terminative*)

The second locative, which corresponds to the nominal suffix *-e*, is somewhat more difficult to isolate, and its identity as well as morphophonemic shape are disputed. Civil (1976:90) has proposed that it is a vocalic element or glide; his theory has been fully investigated, with reference to the many other theories on the subject, by Karahashi (2002 [2005]). The morphophonemic realizations in writing, following other morphemes, obscure its prototypical shape. After conjugation prefixes the main writings are *bi*<sub>2</sub>- (*ba.i*), *imma-* (*imma.i*), *mu-NI-* (*mu.i*); after indirect object prefixes, *mu-e* (*mu.i*), *ri-* (*ra.i*), *ni* (*na.i*); after the comitative, *di* or *de*<sub>3</sub> (*da.i*); after “ablative” *ra-* it is *ri-*.

The function of this prefix is similar to that of locative 1: concord with locatives, and with logical direct objects of compound verbs.



#### 4.6.10 Agreement prefixes and suffixes (positions 6 and 9)

The final position before the root is occupied by agreement prefixes (position 6), although in the plural these prefixes work cumulatively with the second suffix position (position 9). These affixes cross-reference the core arguments of the clause – ergative, absolutive, nominative, and accusative. As already noted, *perfective* verbs have *ergative* agreement, and *imperfective* verbs have *nominative-accusative* agreement. The reconstruction of the forms is somewhat complicated by morphophonemic changes. The prototypical paradigms presented here do not apply in all cases, as the agreement markers may be used for different functions with different verbs (see Yoshikawa 1977). Ergative agreement is marked by prefixes in combination with suffixes, absolutive by suffixes:

##### (23) Agreement affixes – perfective aspect

<i>Ergative agreement affixes</i>		
	<i>Singular</i>	<i>Plural</i>
<i>First</i>	Ø/e-	(-enden)
<i>Second</i>	e-	(-enzen)
<i>Third animate</i>	n-	(-eš)
<i>Third inanimate</i>	b-	

<i>Absolutive agreement suffixes</i>		
	<i>Singular</i>	<i>Plural</i>
<i>First</i>	-en	-enden
<i>Second</i>	-en	-enzen
<i>Third</i>	-Ø	-eš

In the imperfective aspect, the suffixes mark the nominative subject (i.e., both transitive and intransitive subjects) in the first and second person; the use of these suffixes is not obligatory with transitive verbs. However, there is a three-way split in the third person, with separate suffixed-marking of transitive and intransitive subjects, as well as distinct prefixed-marking of transitive objects (see Woods 1999).

##### (24) Agreement affixes – imperfective aspect

<i>Nominative agreement suffixes</i>		
	<i>Singular</i>	<i>Plural</i>
<i>First</i>	-en	-enden
<i>Second</i>	-en	-enzen
<i>Third (transitive subject)</i>	-e	-ene
<i>Third (intransitive subject)</i>	-(e)	-eš

<i>Accusative agreement prefixes</i>		
	<i>Singular</i>	<i>Plural</i>
<i>First</i>	(0/e-/ /a-)	?
<i>Second</i>	(e-)	?
<i>Third animate</i>	(n-)	?
<i>Third inanimate</i>	(b-)	?

#### 4.6.11 The morpheme *-ed* (position 8)

The first rank after the root is occupied by the suffix *-ed* (Edzard 1967, Steiner 1981). The form and function of this element have been much debated. Certain theories recognized a marker of the imperfective *-e*, and as a result it was unclear if the suffix was defined as *-ed*, *-d*, or *-de*. With the elimination of this imperfective marker it seems relatively certain that the form of this suffix is *-ed*, although in cuneiform writing, the consonant is dropped in final position. The meaning is less clear. This morpheme seems to refer to the future and to purpose, especially used in subordinate clauses with nonfinite verbal forms such as *iri daḡal-e-de<sub>3</sub>* (*daḡal.ed*) “in order to widen the city”; or *iri daḡal-la-da* (*daḡal.ed.a*) “the city that is/has to be widened” (Civil 1999–2000). In finite forms – which are less frequently attested – it seems to have a future function combined with a prospective obligatory modal nuance. The latter results in the incompatibility of *-ed* with modal prefixes with the exception of the indicative, although a few late literary examples of usage are attested.

More commonly this suffix is added to the bare verbal root and is followed by a vowel /e/ in final position to preserve its final consonant. This vowel is not subject to harmony; thus we have *nam tar-e-de<sub>3</sub>* “in order to/able to determine destinies”; but *šum<sub>2</sub>-mu-de<sub>3</sub>* “in order to/to be able to give.” The verb is always imperfective; this is only overtly apparent in those that have a distinct imperfective form, such as *ḡa<sub>2</sub>-ḡa<sub>2</sub>-de<sub>3</sub>* “to place.”

#### 4.6.12 Nominalization (position 10)

The final position of the verbal chain, after *-ed* and the pronominal suffixes, is occupied by *-a*, which creates nouns out of verbs and turns main clauses into dependent ones (Krecher 1993). Once this happens, the nominalized entity can take suffixes as if it were a noun: pronouns, as well as simple and compound and discontinuous case morphemes.

The nominalizer *-a* can be attached to both finite and nonfinite verbal forms – that is to verbs with a full set of prefixes, or to the bare root alone. With finite verbs this creates subordinate clauses dependent on another verb, on a noun, or on a relative pronoun, as in *lu<sub>2</sub> e<sub>2</sub> in-du<sub>3</sub>-a* “the one who built the temple.” The morpheme can also be attached to verbs in indirect speech clauses dependent on a limited set of verbs of speaking. Because of the restricted rhetorical range of written Sumerian, this usage is relatively rare in the preserved corpus.

With nonfinite verbal forms the suffix *-a* creates participles that are usually equivalent to English active and past participles: *kur a-ta il<sub>2</sub>-la* “mountain rising from the waters”; *ḡeštug šum<sub>2</sub>-ma* “given wisdom.” For other uses of the nominalizer see §5.4.

#### 4.6.13 Other suffixes

There are two other morphemes that have traditionally been assigned the same final rank as the nominalizer *-a*. Both are rarely used and both are unattested before the Old Babylonian period. Although it is difficult to prove, one might question if these are really bound morphemes or if they are independent particles. The first of these is the marker of direct speech, *-eše* and the second, *-ḡišen*, marks irrealis and seems to be equivalent to “if only” or “were it that.” As Civil (2002 [2005]) has observed, *eše* is not an affix but a frozen verbal form meaning “they said.”

#### 4.6.14 The enclitic copula

The Sumerian verb *me* “to be” can be used independently, but is most commonly attested as a copula (Gragg 1968). It is intransitive, occurs only in the perfective, and takes only the pronominal endings:

(25)	<i>Singular</i>	<i>Plural</i>
<i>First</i>	i <sub>3</sub> -me-en	-me-en-de <sub>3</sub> -en
<i>Second</i>	i <sub>3</sub> -me-en	-me-en-ze <sub>2</sub> -en
<i>Third</i>	-me	-me-eš

In its use as a copula it is morphologically identical to the forms of (25) except that the third-person singular is *-am*, originally written *-am<sub>6</sub>*, but later as *-am<sub>3</sub>*. The final consonant of the copula is dropped in early texts; from Ur III on, this happens regularly only in the third-person singular. The copula can even be added to the conjugated form of “to be” as in the following: *pi-lu<sub>5</sub>-da u<sub>4</sub>-bi-ta e-me-a* (*e.me.am*) “these were the conventions of earlier times.”

The functions of the copula are multifold. With nouns it often takes the place of an independent pronoun as a predicate: *sipa-me-en e<sub>2</sub> mu-du<sub>3</sub>* “I (lit. I am) [the king,] the shepherd, have built the temple.” It can function as a simple predicate, sometimes following a bridging genitive morpheme: *an-ta-sur-ra ġa<sub>2</sub>-a-kam* (*ġa.ak.am*) “the Antasura [shrine] is mine!” It is often used pausally or emphatically after a complete sentence, appended to a finite verbal form.

The copula is also used in comparisons, much like the nominal equative ending *-gin<sub>7</sub>*: thus, *e<sub>2</sub> kur gal-am<sub>3</sub>* “the temple is akin to a large mountain.”

### 4.7 Numerals

#### 4.7.1 Cardinals

As a rule, Sumerian number words are written with number signs and are not spelled out syllabically; hence there is some uncertainty about the forms of the words, and not all numbers are attested. The following tentative list is based on word lists, as reconstructed by Diakonoff (1984) and Civil (1982:6–7). According to this analysis, there are five primary words that were originally compounded to create the numbers six through nine. Small numbers were counted decimally, large numbers in multiples of sixty.

(26)	1	diš, dili, aš	9	(y)ilimmu (ya + limmu)
	2	min	10	hu(wu)
	3	eš (written eš <sub>5</sub> )	20	niš
	4	limmu	30	ušu
	5	ya (written ia <sub>2</sub> )	40	nimin
	6	aš (written aš <sub>3</sub> ; ya + aš)	50	ninnu
	7	imin (ya + min)	60	ġi/eš
	8	ussu (ya + eš)	360	šar

#### 4.7.2 Ordinals

Ordinal numbers are formed with the cardinal number word, followed by the bridging genitive morpheme and the enclitic copula *-am*: thus, *min-(a)-kam* (*min.ak.am*) “second.”

## 5. SYNTAX

The syntax of Sumerian is perhaps the most neglected part of the grammar, and its complexities can only be hinted at in the limited space available here. The language is head-final; subordinate and relative clauses appear to the left of the main clause. Although Sumerian morphology is primarily ergative, it seems that ergativity plays little or no role in interclausal syntax; indeed it may very well be that the language is one of those that have no syntactic pivot (Zólyomi 1996a:106), although this is a matter that requires full investigation. Sentences are either simple or complex. The rich verbal morphology of Sumerian encodes much syntactic information, but the morphological and syntactic relationships between clauses and sentences have not been extensively studied.

### 5.1 Simple sentence word order

Simple sentences as a rule follow SOV order, although the object can be moved right and complements moved left for pragmatic purposes. Sentences with all three components in proper order are primarily third person:

- (27) lugal-e      e<sub>2</sub>              mu-un-du<sub>3</sub>  
       lugal.e      e.Ø              mu.n.du.Ø  
       king-ERG.   temple-ABS.   PREF.-ERG.-build-ABS.  
       “The king built the temple”

Unmarked first-person agents are only expressed by verbal agreement markers; pronouns are used only for emphasis or topicalization:

- (28) ḡa<sub>2</sub>-e              uri<sup>ki</sup><sub>5</sub>-ma      ga-na-aḡ<sub>2</sub>  
       ḡa.e              urim.a              ga.na.aḡ  
       PRO.-ERG.    Urim-LOC.    PREF.-DAT.-pay  
       “I want to pay him back in [the city of] Ur myself”

Agents are not obligatory; clauses without overt agents correspond to Akkadian or English impersonal passives:

- (29) e<sub>2</sub>              ba-du<sub>3</sub>  
       e.Ø              ba.du  
       temple-ABS.    PREF.-build  
       “The temple was built”

Because Sumerian has such a complex verbal morphology, a finite verbal form can by itself constitute a well-formed sentence:

- (30) bi<sub>2</sub>-in-dug<sub>4</sub>  
       ba.i.n.dug.Ø  
       PREF.-LOC2-ERG.-speak-ABS.  
       “He/she said it”

The copula can function as the predicate:

- (31) dumu    uri<sup>ki</sup><sub>5</sub>-ma      me-en  
       dumu    urim.ak      me.en  
       son      Urim-GEN.    COP.  
       “I am a citizen of [the city of] Ur”

A nominalized verb can be turned into a full predicate by addition of the copula:

- (32)  $bi_2$ -in-dug<sub>4</sub>-ga-gin<sub>7</sub>-nam  
 ba.i.n.dug.a.gin.am  
 PREF.-LOC.-ERG.-speak-NOMINALIZER-EQUATIVE-COP.  
 “It was just as he had said”

## 5.2 Coordination

Two nouns can be seriated together to express conjunction as in *an ki* “heavens and the earth.” The compound morpheme *-bi.da* (possessive pronoun and comitative case-marker) is also used to conjoin two nouns, as in *an ki-bi-da* “heavens and the earth.” From the latter half of the third millennium one encounters the sporadic use of loanword  $u_3$ , presumably borrowed from Semitic *u* (originally \**wa*), which is attested in both Eblaite and Old Akkadian. Simple sentences can be seriated with conjunctive, resultative, or disjunctive meaning. Again, beginning in the latter half of the third millennium, one finds the occasional use of the conjunction  $u_3$ .

- (33) A.  $ni_2$       ba-da-te      su      ba-da-zi  
 ni.Ø      ba.da.te      su.Ø      ba.da.zi  
 N.-ABS.    PREF.-COMT.-fear    N.-ABS.    PREF.-COMT.-be terrified  
 “I was afraid, I was terrified”
- B.  $giš$       ba-gur<sub>4</sub>      kuš-bi      nu-da-dar  
 $giš$ .Ø      ba.gur      kuš.bi      nu.da.dar  
 tree-ABS.    PREF.-grow    thick bark-PRO.    NEG.-COMT.-split  
 “The tree grew thick, [but] its bark did not split”
- C. dub-sar    me-en    na-ru<sub>2</sub>-a    ab-sar-re-en  
 dubsar    me.en    narua.Ø    a.b.sar.en  
 scribe    COP.    stela-ABS.    PREF.-ACC.-write-NOM.  
 “I am a scribe, [therefore] I can write stele”

Seriated clauses can even have temporal or implicational relationships that are not marked by any particle or morphological marker:

- (34)  $a_2$ -a $g_2$ - $g_a_2$     lugal- $g_a_2$ -ke<sub>4</sub>       $i_3$ -gub-be<sub>2</sub>-en      nu-dur<sub>2</sub>- $u_3$ -de<sub>3</sub>-en     $bi_2$ -dug<sub>4</sub>  
 a'a $g_a$       lugal. $g_a$ .ak.e      i.gub.en      nu.dur.ed.en      bi.dug  
 orders    king-PRO.-GEN.-LOC2.    PREF.-stand-NOM.    NEG.-sit-SUFF.-NOM.    PREF.-speak  
 “I said: ‘When acting on His Majesty’s (lit. “my king’s”) orders, I stand, I do not sit!’”

## 5.3 Subordination

Other complex sentences consist of a main clause preceded by a relative or other subordinate clause. The predicate of the subordinate clause is always nominalized; it may be a full verbal form (S) or, as is most often the case, a nominal form of the verb, with nominalization, but without the normal affixes (S'). Once nominalized, nominal markers such as pronouns and case endings may follow the verbs.

Full verbal forms, once nominalized, can be treated as nouns and can be bracketed by various discontinuous morphemes to create relative clauses. The first elements in these sequences include nouns (*ud* “day,” *e $g_2$ er* “back”) as well as particles such as *en-na* (“until”) and *mu* (“because”), which do not carry meaning by themselves, but only as part of specific

constructions. The nominalized verb is then followed by a locative or directional case ending, sometimes combined with the bridging genitive morpheme.

Subordinate clauses can also be introduced by conjunctions such as *tukum-bi* “if, in the event that,” or by nouns such as *ud* “day.” Thus *ud-da* (*ud.a* “on the day”) means “when” and *ud-ba* (*ud.bi.a* “on that day”) means “then.”

Subordination can also be marked on the predicate with the modal prefix *he-* in its epistemic function (see §4.6.4.3). As explained by Civil (2002 [2005]), when the subordinate clause comes first, it is conditional:

- (35) *u<sub>2</sub>-gu*    *he<sub>2</sub>-ni-ib-de<sub>2</sub>*                      *ki-bi*                      *ga-mu-na-ab-ḡi<sub>4</sub>*  
*ugu.Ø*    *he.ni.b.de*                              *ki.bi*                      *ga.mu.na.b.ḡi*  
 N.-ABS.    PREF.-LOC2.-ERG.-lose    place-PRO.    PREF.-PREF.-DAT.-ACC.-return  
 “Should it be lost, I will replace it for him”

If the *he-* clause is in final position, it marks a situation that is made possible by the main clause:

- (36) A. *u<sub>2</sub>-lal-e*                      *mu-un-du<sub>3</sub>*                      *amar-e*                      *ha-ma-an-gu<sub>7</sub>-e*  
           *ulal.e*                              *mu.ni.du*                      *amar.e*                      *ha.ma.ni.gu.e*  
           sweet-plant-LOC2    PREF.-LOC2.-plant    calf-LOC2    PREF.-DAT.-LOC2-eat-NOM.  
           “He planted the sweet-plant, so that the calf(calves) could eat them”  
 B. *uri<sup>ki</sup><sub>5</sub>-ma*    *gi*    *zi-bi*    *lal<sub>3</sub>-am<sub>3</sub>*    *ku<sub>6</sub>*    *ha-ma-gu<sub>7</sub>-e*  
           *urim.ak*    *gi*    *zi.bi*    *lal.am*    *ku.Ø*    *ha.ma.gu.e*  
           Ur-GEN.    reed    zi-PRO.    sweet-COP.    fish-ABS.    PREF.-DAT.-eat-NOM.  
           “The *zi* reeds of the [the city of] Ur are sweet, and so the fish eat them”

Commonly a nominalized full verbal predicate can be followed by the ablative suffix *-ta* to create temporal clauses that mark both temporal sequence and a form of contemporaneity:

- (37) *ba-tu-ud-en-na-ta*    *nitah*    *kala-ga*                      *me-en*  
           *ba.tud.en.a.ta*    *nitah*    *kalaga.Ø*                      *men*  
           PREF.-born-NOM.-NOMINALIZER-ABL.    male    mighty-ABS.    COP.  
           “Ever since my birth I have been a mighty male”

More complex and varied are subordinate clauses that are construed with reduced verbal forms (*S'*) (see Gragg 1973b:90–91; 1973c; Michalowski 1978:117; Civil 1999–2000). The simplest such predicates consist of (i) the root and the nominalizer *-a*; (ii) the root followed by the morpheme *-ed*; or (iii) the root and *-ed* + copula (obligation). The first creates simple relative clauses:

- (38) *e<sub>2</sub>*                      (*lugal-e*)                      *du<sub>3</sub>-a*  
           *e.Ø*                      (*lugal.e*)                      *du.a*  
           temple-ABS.    (king-ERG.)    build-NOMINALIZER  
           “The temple that the king built”

The second and third constructions differ in meaning:

- (39) A.  $e_2$ -a-ni       $du_3$ -de<sub>3</sub>      ma-an-dug<sub>4</sub>  
           e.ani      du.ed      ma.n.dug  
           temple-PRO. build-SUFF. PREF.-ERG.-speak  
           “He told me to build his temple”
- B.  $e_2$ -a-ni       $du_3$ -da      ma-an-dug<sub>4</sub>  
           e.ani      du.ed.am      ma.n.dug  
           temple-PRO. build-SUFF.-COP. PREF.-ERG.-speak  
           “He told me that I had to build his temple”

Temporal clauses are often created on the patterns based on S'-a (i.e., the type of [38]), followed by pronouns and other endings. These create a complex paradigm with nominative/accusative rather than ergative agreement. One exception aside, only singular forms are encountered in texts. Once again the third person distinguishes between animate and inanimate forms:

(40)	Singular	Plural
<i>First</i>	S'-a-ġu-ne	
<i>Second</i>	S'-a-zu-ne	S'-ed-a-enzen
<i>Third animate perfect</i>	S'-(a)-ani	
<i>Third animate imperfect</i>	S'-ed-ani	
<i>Third inanimate</i>	S'-a-bi	

Examples of temporal clauses follow:

- (41) A.  $ka_2$   $e_2$ -gal-la-še<sub>3</sub>      gub-a-ġu<sub>10</sub>-ne  
           ka      egal.ak.še      gub.a.ġu.ne  
           gate      palace-GEN.-ALL.      go-NOMINALIZER-PRO.-SUFF.  
           “When I arrived at the gate of the palace”
- B.  $ku_4$ - $ku_4$ -da-ġu<sub>10</sub>-ne  
           ku.kud.a.ġu.ne  
           enter-NOMINALIZER-PRO.-SUFF.  
           “When I entered”
- C. ud-bi-a      ġiš-e       $e_2$ -gal-la       $ku_4$ - $ku_4$ -da-ni  
           ud.bi.a      ġiš.e      egal.a      ku.kud.ani  
           day-PRO.-LOC.      tree-ERG.      palace-LOC.      enter-PRO.  
           “Then, as Tree entered the palace”

The third-person forms can also occur with an addition of the ablative *-ta*: *gur-re-da-ni* “when he returns,” but *gur-ru-da-ni-ta* “upon his return.” According to Gragg (1973c:128), the latter indicates a time subsequent to an action, while the former relates a time at which something happened.

## 6. LEXICON

Although Sumerian died out millennia ago, the countless surviving cuneiform tablets preserve a rich and varied lexicon. In addition to words in narrative contexts, we also have access to an extensive native Mesopotamian lexicographical tradition in the form of ancient

monolingual and multilingual lexical lists (Civil 1975). These lexical texts, which were designed for use in teaching the art of writing, have a long tradition, from the very beginnings to the very end of cuneiform use. The early versions are monolingual, but by the second millennium the entries were all provided with Akkadian translations; outside of Mesopotamia other languages were added. The longest composition of this type contains almost ten thousand entries. These lists are arranged by various criteria, graphic, semantic, etc. Some have compound words or sign combinations, and some late bilingual lists are arranged according to the Akkadian translations. The complex nature of these texts should not be underestimated. They include many terms that had long gone out of use, or were no longer properly understood. Some words were simply invented by scribes who were not native Sumerian speakers. Most important, one has to respect the organizational structure of a specific list type to properly understand the semantics of an entry.

The lexical texts contain many words that are not otherwise documented in any Sumerian texts. But the lexicon of the literary and administrative tablets must also be treated with caution. The language of written texts is often conservative and resistant to the changes taking place in the vernacular. Many if not most extant Sumerian texts were written and composed after the language was no longer spoken in the streets, and therefore one has to view diachronic developments differently than one would if this were a living tongue.

It would seem that much of the lexicon is native Sumerian, but this is difficult to gauge correctly in view of the lack of a modern dictionary, and of related languages, and because of some of the ambiguities of the script. Over the millennia, Sumerian came into direct and indirect contact with many other languages and borrowed lexical items from various donors. The majority of such loanwords come from Semitic, mostly from forms of Akkadian. Loans are often, but not always, written out syllabically. Thus, the Semitic loan *dam-ha-ra* "battle" is written with three signs, but *ugula* "overseer," likewise a Semitic loan, is written logographically with only one grapheme (*PA*). The writing of some changed over time. In early texts the Sumerian word for copper – originally a culture word that came into the language through some Semitic intermediary – is written syllabically as *a-ru<sub>12</sub>-da*, then as *urudu<sup>a-ru<sub>12</sub>-da</sup>* or as *a-ru<sub>12</sub>-da<sup>a</sup>urudu*, and finally later on simply as *urudu*.

In the past, scholars have claimed that certain basic culture words were borrowed from one or more hypothetical substrate languages, sometimes referred to as Proto-Tigridian and Proto-Euphratic, and that one of them may even have been Indo-European. More recently Rubio (1999b) has shown that these lexemes are either native Sumerian, Semitic loans, or culture words that show up in various languages; and while one cannot discount the possibility of some substrates at some time, the current linguistic evidence does not support this in any way.

Semitic loans have a long history in Sumerian. The earliest such borrowings exclude any Semitic endings (*har-ra-an* "road" from Akkadian *harranum*); Old Akkadian period loans end in *-a* (*ugula* from Akkadian *waklum*); and second-millennium loans retain the Akkadian nominative ending and mimation (*pu-uh<sub>2</sub>-ru-um* "assembly" from Akkadian *puhrum*), although there are exceptions to these rules. There are also rare borrowings from Hurrian, for example, *tibira* "metal worker," although this may be a culture word, and from unknown sources, as is the case with *lams(a)r* "brewing vat." There are other "wandering words" that appear in many different languages: Sumerian *ugu<sub>4</sub>-bi* "monkey" belongs together with Akkadian *pagû*, as well as reflexes in Hebrew and Egyptian; *za<sub>3</sub>-gin<sub>2</sub>* "lapis lazuli" compounded with *za* "stone" is of the same unknown origin as Akkadian *uqnu* or Hittite *ku(wa)nna(š)*. More complex is the matter of Sumerian (*h*)*urin*, *er<sub>in</sub>*, Akkadian *a/erû* "eagle." Civil (1983:3) seeks the origins of these words in a form *\*haran*, and points to Hittite *haran-* "eagle." The root appears commonly in Indo-European, but the ultimate origin is



unknown. It has been proposed that a number of Sumerian agricultural terms belong to this category, but this requires further investigation (Rubio 1999b). Borrowed words usually replace native ones, as exemplified by the Semitic loan *iri* “city,” for which the original Sumerian word is not preserved; but sometimes they were used alongside the native term, as *unken* “assembly” coexisted with *pu-uh<sub>2</sub>-ru-um*. There are even poetic examples of the rhetorical use of synonymic word pairs, with the native term preceded by the borrowed one, as in *har-ra-an kaskal* “road (road).” This seems to be the order encountered in most languages that have such pairs (Boeder 1991).

## 7. READING LIST

The most convenient place to read about Sumerian grammar is Thomsen 1984, supplemented by the important remarks of Attinger 1993, and by more recent studies listed in Römer 1999. The two classic highly influential older grammars are Poebel 1923 and Falkenstein 1959, to which one has to add the idiosyncratic but often brilliant insights of Jacobsen 1965 and 1988. Readers of Russian should not ignore the important but often overlooked contributions of Diakonoff (1967 = 1979) and of his student Kaneva (1996). Interesting insights are also found in the older sketches of Jestin (1951) and Lambert (5 fascicles, 1972–1978). The latest grammar is that of Edzard (2003). No reliable introductory primer is currently available, but there is an excellent reader that contains a selection of texts in cuneiform with a sign list, bibliography, and glossary (Volk 1997).

There is no complete modern dictionary of the language; the first volumes of the monumental *Pennsylvania Sumerian Dictionary* (Sjöberg *et al.* 1984–) are now available, but they currently only cover words beginning with the letters A and B. The web-based version of the full Dictionary, now edited by S. Tinney, is available on line at: <http://psd.museum.upenn.edu/epsd>.

It is impossible to list here all the published Sumerian sources; for the important word lists see Civil 1975, and for a survey of literary compositions see Michalowski 1995 with bibliography. For all types of texts and the secondary literature consult Römer 1999.

## Bibliography

- Alster, B. 1982. “Emesal in Early Dynastic Sumerian? What is the UD.GAL.NUN-orthography?” *Acta Sumerologica* 4:1–6.
- Anderson, S. 1985. “Inflectional morphology.” In T. Shopen (ed.), *Language Typology and Syntactic Description III: Grammatical Categories and the Lexicon*, pp. 150–201. Cambridge: Cambridge University Press.
- Aristar, A. 1995. “Binder-anaphors and the diachrony of case displacement.” In F. Plank (ed.), *Double Case: Agreement by Suffixaufnahme*, pp. 431–447. Oxford: Oxford University Press.
- . 1996. “The relationship between dative and locative: Kurytowicz’s argument from a typological perspective.” *Diachronica* 13:207–224.
- . 1997. “Marking and hierarchy: types and the grammaticalization of case-markers.” *Studies in Language* 21:313–368.
- Attinger, P. 1993. *Éléments de linguistique sumérienne. La construction de du<sub>11</sub>/e/di «dire»* (Orbis Biblicus et Orientalis, Sonderband). Fribourg/Göttingen: Editions Universitaires/Vandenhoeck and Ruprecht.
- Bauer, J. 1975. “Zum /dr-/Phonem des Sumerischen.” *Welt des Orients* 8:1–9.
- . 1998. “Der vorsargonische Abschnitt der mesopotamischen Geschichte.” In J. Bauer, R. K. Englund, and M. Krebernik (eds.), *Mesopotamien. Späturuk-Zeit und Frühdynastische Zeit*.

- pp. 431–585. *Orbis Biblicus et Orientalis*, 160/1. Fribourg/Göttingen: Editions Universitaires/Vandenhoeck and Ruprecht.
- Black, J. 1984. *Sumerian Grammar in Babylonian Theory*. Studia Pohl, Series Maior 12. Rome: Biblical Institute Press.
- . 1986. Review of Thomsen 1984. *Archiv für Orientforschung* 33:77–83.
- . 1990. “The alleged ‘extra’ phonemes of Sumerian.” *Revue d’Assyriologie* 84:107–118.
- . 2000 [2005] “Some Sumerian adjectives.” *Acta Sumerologica* 22:3–27.
- Bobrova, L. and A. Militarëv. 1989. “Towards the reconstruction of Sumerian phonology.” In *Lingvistieska rekonstrukcia i drevneiaia istoriia Vostoka*, part 1, pp. 96–105. Moscow: Nauka.
- Boeder, W. 1991. “A note on synonymic parallelism and bilingualism.” *Studia Linguistica* 45:97–126.
- Boisson, C. 1992. “The Sumerian pronominal system in Nostratic perspective.” In V. Shevoroshkin (ed.), *Nostratic, Dene-Caucasian, Austric and Amerind: Materials from the First International Interdisciplinary Symposium on Language and Prehistory, Ann Arbor, 8–12 November, 1988*, pp. 433–461. Bochum: Universitätsverlag Dr. Norbert Brockmeyer.
- . 1997. “The phonotactics of Sumerian.” In I. Hegeds, P. A. Michalove, and A. Manaster Ramer (eds.), *Indo-European, Nostratic, and Beyond: Festschrift for Vitalij V. Shevoroshkin*, pp. 30–50. Journal of Indo-European Studies Monograph 12. Washington: Institute for the Study of Man.
- Bottéro, J. 1992. “The ‘avalanche’ of decipherments in the ancient Near East between 1800 and 1930.” In J. Bottéro, *Mesopotamia: Writing, Reasoning and the Gods*, pp. 55–66. Chicago: Chicago University Press.
- Chung, S. and A. Timberlake. 1985. “Tense, aspect, and mood.” In T. Shopen (ed.), *Language Typology and Syntactic Description III: Grammatical Categories and the Lexicon*, pp. 202–258. Cambridge: Cambridge University Press.
- Civil, M. 1966. “Notes on Sumerian lexicography, I.” *Journal of Cuneiform Studies* 20:119–124.
- . 1973a. “From Enki’s headaches to phonology.” *Journal of Near Eastern Studies* 32:57–61.
- . 1973b. “The Sumerian writing system: Some problems.” *Orientalia* NS 42:21–34.
- . 1973c. “Notes on Sumerian lexicography, II.” *Journal of Cuneiform Studies* 25:171–177.
- . 1975. “Lexicography.” In S. Lieberman (ed.), *Sumerological Studies in Honor of Thorkild Jacobsen*, pp. 123–157. Assyriological Studies, 20. Chicago: University of Chicago Press.
- . 1976. “The song of the plowing oxen.” In B. Eichler et al. (eds.), *Kramer Anniversary Volume: Cuneiform Studies in Honor of Samuel Noah Kramer*, pp. 83–95. Alter Orient und Altes Testament, 25. Neukirchen-Vluyn: Neukirchener Verlag.
- . 1982. “Studies in Early Dynastic lexicography, 1.” *Oriens Antiquus* 21:1–26.
- . 1983. “Early Dynastic spellings.” *Oriens Antiquus* 22:1–5.
- . 1984. “Bilingualism in logographically written languages: Sumerian in Ebla.” In L. Cagni (ed.), *Il bilinguismo a Ebla*, pp. 75–97. Naples: Istituto Universitario Orientale.
- . 1987. “The early history of HAR-ra: The Ebla link.” In L. Cagni (ed.), *Ebla 1975–1985. Dieci anni di studi linguistici e filologici*, pp. 131–158. Naples: Istituto Universitario Orientale.
- . 1999–2000. “Reading Gilgameš: history and stories.” *Aula Orientalis* 17–18:179–189.
- . 2002. “The Forerunners of Marû and Hamtu in Old Babylonian.” In T. Abusch (ed.), *Riches Hidden in Secret Place: Ancient Near Eastern Studies in Memory of Thorkild Jacobsen*, pp. 63–72. Winona Lake: Eisenbrauns.
- . 2000 [2005]. “Modal prefixes.” *Acta Sumerologica* 22:29–42.
- Civil, M. and R. Biggs. 1966. “Notes sur les textes sumériens archaïques.” *Revue d’Assyriologie* 60:1–16.
- Cooper, J. 1996. “Mesopotamian cuneiform: Sumerian and Akkadian.” In P. Daniels and W. Bright (eds.), *The World’s Writing Systems*, pp. 37–57. Oxford: Oxford University Press.
- Diakonoff, I. 1967. “Shumerskij iazyk.” In I. Diakonoff, *Iazyki drevnei Perednei Azii*, pp. 35–84. Moscow: Nauka.
- . 1979. “Shumerskij iazyk.” In I. Diakonoff, I. M. Dunaevskaia, Iu. D. Desherier, et al. (eds.), *Iazyki drevnei Perednei Azii (nesemitskie): iberiisko-kavkazskie iazyki: paleoaziatskie iazyki*, pp. 7–36. Iazyki Azii i Afriki 3. Moscow: Nauka.
- . 1983. “Some reflections on numerals in Sumerian: towards a history of mathematical speculation.” *Journal of the American Oriental Society* 103:83–93.

- Edzard, D. 1963. "Sumerische Komposita mit dem 'Nominalprefix' nu-." *Zeitschrift für Assyriologie* 55:91–112.
- . 1967. "Das sumerische Verbmorphem /ed/ in den alt- und neusumerischen Texten." *Heidelberger Studien zum Alten Orient* 1:29–62.
- Edzard, D. O. 2003. *Sumerian Grammar*. Handbuch der Orientalistik, 1.71. Leiden: Brill.
- Falkenstein, A. 1959. *Das Sumerische*. Handbuch der Orientalistik. Leiden: Brill.
- Gelb, I. 1961. *Old Akkadian Writing and Grammar* (2nd edition). Materials for the Assyrian Dictionary, 2. Chicago: University of Chicago Press.
- Geller, M. 1997. "The first wedge." *Zeitschrift für Assyriologie* 87:43–95.
- Gragg, G. 1968. "The syntax of the copula in Sumerian." In J. Verhaar (ed.), *The Verb "Be" and its Synonyms, 3: Philosophical and Grammatical Studies*, pp. 86–109. Foundations of Language, Supplementary Series, 8. Dordrecht: Reidel.
- . 1973a. *Sumerian Dimensional Infixes*. Alter Orient und Altes Testament – Sonderreihe 5. Neukirchen-Vluyn: Neukirchener Verlag.
- . 1973b. "Linguistics, method, and extinct languages: the case of Sumerian." *Orientalia NS* 42:78–96.
- . 1973c. "A class of 'when' clauses in Sumerian." *Journal of Near Eastern Studies* 32:124–134.
- Huber, C. 1996. "Some notes on transitivity, verb types, and case with pronouns in Sumerian." *Wiener Zeitschrift für die Kunde des Morgenlandes* 86:177–189.
- Jacobsen, T. 1957. "Early political development in Mesopotamia." *Zeitschrift für Assyriologie* 52:91–140.
- . 1965. "About the Sumerian verb." In *Studies in Honor of Benno Landsberger on his Seventy-fifth Birthday, April 21, 1965*, pp. 71–102. Assyriological Studies 16. Chicago: Chicago University Press.
- . 1988. "The Sumerian verbal core." *Zeitschrift für Assyriologie* 78:161–220.
- Jagersma, B. 1993. Review of J. Hayes, *A Manual of Sumerian Grammar and Texts*. Malibu: Undena, 1990. *Bibliotheca Orientalis* 50:422–425.
- . 2000 [2005]. "Sound change in Sumerian: the so-called /dr/ phoneme." *Acta Sumerologica* 22:81–87.
- Jestin, R. 1951. *Abrégé de grammaire sumérienne*. Paris: Paul Geuthner.
- Kaneva, I. 1996. *Shumerskii iazyk* (The Sumerian language). St. Petersburg: Tcentr "Peterburgskoe Vostokovedenie."
- Karahashi, F. 2000. "Sumerian Compound Verbs with Body-Parts." Doctoral dissertation, University of Chicago.
- . 2000 [2005]. "The locative-terminative verbal infix in Sumerian." *Acta Sumerologica* 22:113–133.
- Kienast, B. 1975. "Zur Wortbildung des Sumerischen." *Zeitschrift für Assyriologie* 65:1–27.
- Krecher, J. 1978. "Das sumerische Phonem ġ." In B. Hruška and G. Komoroczy (eds.), *Festschrift Lubor Matouš II*, pp. 7–73. *Assyriologia*, 5. Budapest: Eötvös Loránd Tudományegyetem, Ökori Történeti Tanszék.
- . 1987. "Morphemeless syntax in Sumerian as seen on the background of word-composition in Chukchee." *Acta Sumerologica* 9:67–88.
- . 1993. "The suffix of determination -/al/." *Acta Sumerologica* 15:81–98.
- Krispijn, Th. 2000 [2005]. "The change of Official Sumerian in the city state of Lagash." *Acta Sumerologica* 22:153–175.
- Kurylowicz, J. 1964. *The Inflectional Categories of Indo-European*. Heidelberg: Carl Winter.
- Lambert, M. 1972–1978. *Grammaire sumérienne rédigée à l'intention des élèves de l'École du Louvre*. Five fascicles. Paris: École du Louvre.
- Lieberman, S. 1979. "The phoneme /o/ in Sumerian." In M. Powell, Jr. and R. Sack (eds.), *Studies in Honor of Tom B. Jones*, pp. 21–28. *Altes Orient und Altes Testament* 203. Neukirchen-Vluyn: Neukirchener Verlag.
- Michalowski, P. 1978. "Two Sumerian literary letters." *Journal of Cuneiform Studies* 30:114–120.
- . 1980. "Sumerian as an ergative language, part 1." *Journal of Cuneiform Studies* 32:86–103.
- . 1995. "Sumerian literary traditions: an overview." In J. Sasson, J. Baines, G. Beckman, et al. (eds.), *Civilizations of the Ancient Near East*, pp. 2277–2289. New York: Scribners.

- . 2000 [2005]. "The life and death of the Sumerian language in comparative perspective." *Acta Sumerologica* 22:177–202.
- . 2006. "The Lives of the Sumerian Language." In S. Sanders (ed.), *Margins of Writing, Origins of Culture: New Approaches to Writing and Reading in the Ancient Near East. Papers from the Symposium held February 25–26, 2005*, pp. 157–82. Chicago: The Oriental Institute.
- Miner, K. 1986. "Noun stripping and loose incorporation in Zuni." *International Journal of American Linguistics* 52:242–254.
- Mithun, M. 1984. "The evolution of noun incorporation." *Language* 60:847–894.
- . 1985. "Diachronic morphologization: The circumstances surrounding the birth, growth, and decline of noun incorporation." In J. Fisiak (ed.), *Papers from the 6th Conference on Historical Linguistics*, pp. 365–394. Amsterdam: John Benjamins; Poznań: Adam Mickiewicz University Press.
- . 2000. "Incorporation." In G. Booij, C. Lehmann, J. Mugdan, et al. (eds.), *Morphology: An International Handbook on Inflection and Word Formation*, pp. 916–928. Berlin: Walter de Gruyter.
- Nichols, J. 1993. "Ergativity and linguistic geography." *Australian Journal of Linguistics* 13:39–89.
- Palmer, F. 1986. *Mood and Modality*. Cambridge Textbooks in Linguistics. Cambridge: Cambridge University Press.
- Parpola, S. 1975. "Transliteration of Sumerian: problems and prospects." *Studia Orientalia* 46:239–257.
- Payne, T. 1982. "Role and reference related subject properties and ergativity in Yup'ik Eskimo and Tagalog." *Studies in Language* 6:75–106.
- Poebel, A. 1923. *Gründzüge der sumerischen Grammatik*. Rostock: Selbstverlag.
- Rhodes, R. 1997. "On pronominal systems." In I. Hegedus et al. (eds.), *Indo-European, Nostratic, and Beyond: Festschrift for Vitalij V. Shevoroshkin*, pp. 293–319. Journal of Indo-European Studies Monograph 12. Washington: Institute for the Study of Man.
- Römer, W. 1999. *Die Sumerologie. Einführung in die Forschung und Bibliographie in Auswahl*. Alter Orient und Altes Testament 262. Neukirchen-Vluyn: Neukirchener Verlag.
- Rubio, G. 1999a. Review of A. S. Kaye (ed.), *Phonologies of Asia and Africa*. *Language* 75:138–142.
- . 1999b. "On the alleged 'Pre-Sumerian substratum.'" *Journal of Cuneiform Studies* 51:1–16.
- Schretter, M. 1990. *Emesal-Studien: Sprach- und literaturgeschichtliche Untersuchungen zur sogenannten Frauensprache des Sumerischen*. Innsbrucker Beiträge zur Kulturwissenschaft. Sonderheft 69. Innsbruck: Institut für Sprachwissenschaft der Universität Innsbruck.
- . 1993. "Sumerische Phonologie: Zu Konsonantenverbindungen und Silbenstruktur." *Acta Orientalia* 54:7–30.
- . 1996. "Überlegungen zu den Wortarten des Sumerischen." *Wiener Zeitschrift für die Kunde des Morgenlandes* 86:399–412.
- Sjöberg, Å. 1975. "The Old Babylonian eduba." In S. Lieberman (ed.), *Sumerological Studies in Honor of Thorkild Jacobsen*, pp. 159–179. Assyriological Studies, 20. Chicago: University of Chicago Press.
- Sjöberg, Å. et al. (eds.). 1984–. *The Sumerian Dictionary of the University Museum of the University of Pennsylvania*. Philadelphia: Babylonian Section, University Museum.
- Steiner, G. 1981. "The vocalization of the Sumerian verbal morpheme /-ED/ and its significance." *Journal of Near Eastern Studies* 40:21–41.
- Steinkeller, P. 1979. "Notes on Sumerian plural verbs." *Orientalia* NS 48:54–67.
- . 1993. "Early political development in Mesopotamia and the origins of the Sargonic Empire." In M. Liverani (ed.), *Akkad: The First World Empire*, pp. 107–129. Padua: Sargon.
- Thomsen, M.-L. 1984. *The Sumerian Language: An Introduction to its History and Grammatical Structure*. Mesopotamia, 10. Copenhagen: Akademisk Forlag.
- Vanstiphout, H. 1985. "On the verbal prefix /i/ in Standard Sumerian." *Revue d'Assyriologie* 79:1–15.
- Volk, K. 1997. *A Sumerian Reader*. Studia Pohl, Series Maior 18. Rome: Pontifical Biblical Institute.
- Wilcke, C. 1988. "Anmerkungen zum 'Kunjugationspräfix' /i/- und zur These vom «silbischen Charakter der sumerischen Morpheme» anhand neumerischen Verbalformen beginnend mit i-ib-, i-im- und i-in-." *Zeitschrift für Assyriologie* 78:1–49.
- Woods, C. 2000 [2005]. "Deixis, person, and case in Sumerian." *Acta Sumerologica* 22:303–334.

- 
- Wright, H. 1969. *The Administration of Rural Production in an Early Mesopotamian Town*. Anthropological Papers, 38. Ann Arbor: University of Michigan.
- Yoshikawa, M. 1968. "The *marû* and *ḥamṭu* aspects in the Sumerian verbal system." *Orientalia* NS 37:401–416.
- . 1977. "Some remarks on the Sumerian verbal infixes –n-/-b- in the preradical position." *Journal of Cuneiform Studies* 29:78–96.
- . 1992. "The valency-change system in the Sumerian verbal prefixes." *Acta Sumerologica* 14:395–402.
- Zólyomi, G. 1996a. Review of Attinger 1993. *Bibliotheca Orientalis* 53:95–107.
- . 1996b. "Genitive constructions in Sumerian." *Journal of Cuneiform Studies* 48:31–47.
- . 1999. "Directive infix and oblique object in Sumerian: an account of the history of their relationship." *Orientalia* NS 68:215–253.

# Elamite

MATTHEW W. STOLPER

## 1. HISTORICAL AND CULTURAL CONTEXTS

### 1.1 Sources

Texts in Elamite come from the modern provinces of Khuzestān and Fārs, in southwestern Iran. Most are from ancient Susa and the plains of Khuzestān around it, from ancient Persepolis and Anshan (modern Tall-i Malyān) in the high valleys of Fārs, from sites on the way between Susiana and the Persepolis–Anshan area, or from the coast of Fārs. Achaemenid multilingual rock inscriptions of c. 520–450 BC with Elamite versions are also found in central western Iran, near Hamadān, and in eastern Turkey, near Van. Elamite texts on clay tablets from c. 600–550 BC have been found at the Assyrian city of Nineveh, in northeastern Iraq, at the Urartian fortress at modern Armavir Blur in Armenia, and at Old Kandahar in modern Afghanistan. The oldest dated texts are from about 2300 BC, the latest from about 350 BC. The first to come to modern attention were the inscriptions of the Achaemenid kings (c. 522–330 BC), whose Old Persian texts were often accompanied by Elamite and Akkadian versions, all deciphered in the 1840s. Other Elamite texts include royal display or dedicatory inscriptions written on bricks, glazed tiles or other architectural elements, or on stone or metal objects; administrative texts written on clay tablets; engravings on cylinder seals naming the owners of the seals; and a few legal texts, letters, and literary or scholarly texts on clay tablets.

### 1.2 History of the language and its speakers

The indigenous name for the country of Elam, *Hatamti*, is reflected in Sumerian *Elama*, Akkadian *Elamtu*, Hebrew *‘Elām*, and other forms. The indigenous name of the language is not attested. The usual modern name *Elamite* (used as early as Sayce 1874:467) corresponds to Sumerian and Akkadian usage (e.g., Sumerian *eme Elama*, “language of Elam”). Other modern names once given to the language are *Scythian* and *Median*, on the supposition that the languages of the Achaemenid royal inscriptions were those of dominant populations in the Achaemenid empire; *Susian*, in recognition of the fact that the language used beside Old Persian and Babylonian in the Achaemenid royal inscriptions was related to the language found in older texts from Susa; and *Anzanite*, on the view that the language found on texts from Susa was not original there, but was introduced by rulers from Anshan, whose location was a matter of conjecture.

Because there is some disagreement about the historical geography of ancient Iran, there is also uncertainty about the area in which Elamite was actually spoken. On the maximal



view, the political and cultural area of Elam, where the Elamite language must have been commonly used, extended in the late third and early second millennia over the entire highland territory of Iran, as far northwest as Azerbaijan and as far southeast as Baluchistān. Early Elamite states conquered and held Khuzestān and promoted the use of Elamite there in a population that also spoke and wrote in Akkadian, wrote in Sumerian, and perhaps also included some speakers of Amorite and Hurrian. By the middle of the first millennium, however, after the immigration of Iranian speakers and the rise of the Achaemenid Persian state, the territory called Elam was confined to Khuzestān and the adjoining mountains of Lurestān and northwestern Fārs. The Persian rulers who made the Elamite Anshan into Persia proper continued to write inscriptions and administrative records in Elamite in much of highland Iran (e.g., Vallat 1993, 1998). Critics of this view consider the original Elamite political and language area to be much smaller, but to include Fārs, Khuzestān, and extensions of uncertain distance to the northwest and southeast. Most modern appraisals agree in considering Khuzestān, where most early Elamite texts originate, to be at the edge, not at the center, of the Elamite area, a region where Elamite language coexisted or competed with Sumerian and Akkadian.

The unsettled question of the eastern extent of the Elamite language area is connected with the hypothesis that Elamite is related to the Dravidian languages, considered in various forms since the 1850s. A comprehensive proposal of phonological, lexical, and morphological correspondences and developments, with an inference that Proto-Dravidian and an ancestor of Elamite separated from a common Proto-Elamo-Dravidian before 3000 BC, and probably in the fifth millennium BC (McAlpin 1981), has been embraced by some students of Elamite (e.g., Khačikjan 1998:3 following Diakonoff) and ignored by others. It has not been systematically criticized, and it has not yet had practical consequences for the study of Elamite grammar or lexicon (Zadok 1995:243).

The framework of Elamite history is built chiefly on texts from Mesopotamia. Sumerian, Babylonian, and Assyrian states had intermittent, sometimes intense diplomatic, political, military, and commercial connections with the intermontane valleys of Elamite Iran, rich in timber, semi-precious stones and metals, and sometimes in population. The same Mesopotamian states sometimes fought Elamite states for control of Susiana. The chronological phases into which the Elamite language proper is divided are primarily political phases. (The earliest texts from Elamite territories, however, are in undeciphered scripts called Proto-Elamite and Linear Elamite; see below, §2.1.)

### 1.2.1 Old Elamite (c. 2600–1500 BC)

Early Sumerian rulers recorded skirmishing with Elamites in southern Mesopotamia as early as c. 2650 BC. The Old Akkadian rulers of southern Mesopotamia (c. 2300–2100) recorded battles with Elamite rulers and campaigns against Elamite highland regions; they took control of Susa. When Old Akkadian power broke down, Susa fell under the control of a ruler from the interior highland, Puzur-Inšušinak of Awan, who also claimed to control other highland territories. Any political integration that lay behind this claim was short-lived, as Sumerian rulers of the Third Dynasty of Ur (c. 2000–1900) reestablished control over the whole of lower Mesopotamia and over Susa, and pushed into the highland districts surrounding Khuzestān with punitive military campaigns, tribute-taking, the creation of occupied provinces in nearer valley-systems, and the maintenance of active diplomacy with more distant territories. In reaction, Elamite states of the interior coalesced in an alliance that sacked Ur, destroyed the Mesopotamian state and its empire, and took its king to die in captivity in Anshan. By about 1750 BC, this alliance had reached the zenith of Elamite

power, becoming the largest regional state of the time, exercising sway over smaller competing alliances in Mesopotamia and northern Syria, and sending expeditionary armies to promote its interests. A defeat at the hands of Hammurabi of Babylon removed the Elamites from Mesopotamian affairs, but the Elamite monarchy remained in place until c. 1500 BC.

#### 1.2.1.1 *Old Elamite texts*

Elamite texts from this long interval are scarce. They include three tablets of uncertain literary or scholarly character (at least one of them excavated in southern Mesopotamia), a treaty with an Old Akkadian king, and four royal inscriptions from about 1800–1700 BC, only one of them nearly complete (Steve 1992:19; Vallat 1990). Additional evidence comes from Elamite names and words that occur in Sumerian and Akkadian texts from Elamite territories, above all in several hundred legal and administrative texts from Susa (Lackenbacher 1998; Zadok 1995:244). There are also five passages in Sumerian and Old Babylonian texts that are perhaps incantations in Elamite (van Dijk *et al.* 1985:4 and 9ff. Nos. 4, 5, and 18; Hinz and Koch 1987:1322 s.vv. Inc. 70 E–H).

### 1.2.2 **Middle Elamite (c. 1500–1000 BC)**

After about 1450 BC, scattered texts from sites in Khuzestān mention a series of “kings of Susa and Anshan,” and after c. 1400 numerous inscriptions, most of them in Elamite, attest the reigns of two dynasties of “kings of Anshan and Susa” who controlled Susa and nearby sites and eventually resumed warfare with contemporary Assyria and Babylonia. These wars culminated c. 1150 in Elamite raids on the cities of Babylonia, from which the Elamites took trophies that include some of the ancient Mesopotamian monuments that are most celebrated in modern times, including the Victory Stele of Naram-Sin of Akkad and the stele with the Laws of Hammurabi. The wars continued with a Babylonian attack on Elam in c. 1120 BC. Thereafter, sources for Elamite political history fade away.

#### 1.2.2.1 *Middle Elamite texts*

Texts from this period, usually considered the classical period of Elamite language and culture, include about 175 royal inscriptions on bricks, steles, reliefs, statues, and large and small votive objects. Most of them are from Susa or Choghā Zanbīl, a few from other sites in Khuzestān, a site in the valleys on the road to Fārs, a site on the Fārs coast of the Persian Gulf, and one from Anshan (Steve 1992:19–21; add van Soldt 1982:44–48; de Maaijer 1996:70–72). Among them is a single Elamite–Akkadian bilingual building inscription. Elamite administrative tablets from Anshan are attributed either to the end of this period or to the earliest phase of Neo-Elamite, as are two fragmentary legal and administrative texts from Susa (Stolper 1984:5–10; Steve 1992:21). Elamite words and titles also appear in Akkadian administrative texts from Haft Tepe, near Susa, written at the beginning of the period (Lackenbacher 1998:343; Zadok 1995:241).

### 1.2.3 **Neo-Elamite (c. 1000–550 BC)**

By c. 750 BC, when Mesopotamian sources on Elam reappear, much of central and western Iran had been populated by speakers of Iranian languages who lived among, pushed aside, or amalgamated with other ethnic and linguistic groups. The Mesopotamian texts reflect episodic conflict between the Neo-Assyrian empire, then reaching the height of its power, and Neo-Elamite kings who controlled Khuzestān. The theaters of conflict were the central Zagros



valleys, where the Assyrians tried to protect the fringes of a new province, and Babylonia, where the Assyrians tried to stabilize political control against incessant resistance, while the Elamites tried to support buffers against the Assyrians in both places. In the mid-640s these encounters led to an Assyrian sack of Susa and a tour of looting and destruction around the adjoining plains of Khuzestān.

After the fall of the Assyrian Empire, 612–10 BC, successor states arose on Elamite territory, one based at Susa, another probably in the highland valleys to the north of Khuzestān, others in the valleys to the southeast, between Khuzestān and Fārs, and another in central Fārs. The rulers in Fārs were Persians who assumed the Elamite title “king of Anshan.” Their descendant was Cyrus the Great (550–530 BC), who conquered Iran, Anatolia, and Mesopotamia to lay the foundations of the Achaemenid Persian Empire.

### 1.2.3.1 *Neo-Elamite texts*

Elamite texts from the first phase of this period are very scarce. Texts from after c. 750 BC include about thirty royal inscriptions, most on bricks and stele pieces from Susa, but also including rock inscriptions of a local ruler in eastern Khuzestān, and inscriptions of post-Assyrian local rulers on portable objects (Steve 1992:21–23, partially redated by Vallat 1996a; and add Caubet 1995, Donbaz 1996, and Vallat 1996b, Baššāš-e Kanzaq 1997:19–22; Bleibtreu 1999:21, 54; Henkelman forthcoming). An omen text and a hemerological text in Elamite are assigned to the period before 650 BC. From the period after 650 come a small group of legal texts from Susa, an archive of about 300 administrative texts, also from Susa, letters from Susa, Nineveh, and Armavir Blur in Armenia, and some unprovenienced letters and administrative texts, and Elamite inscriptions on cylinder seals from Susa and heirloom seals used at Persepolis (Steve 1992:22–23, and add Vallat 1997b and Jones in Garrison and Root 2001).

## 1.2.4 **Achaemenid Elamite (550–330 BC)**

Under the Achaemenids the region administered from Susa became the province of Elam (Old Persian *Huja*, corresponding in multilingual inscriptions to Elamite *Hatamti*, ~ *Haltamti*), and Fārs became Persia proper (Old Persian *Pārsa*, corresponding to Elamite *Parsa* ~ *Paršan* ~ *Paršaš*). Darius I (522–486 BC) and his successors built palace complexes at Susa, which became the main political center of the imperial court, and at Persepolis, not far from the old Elamite center at Anshan. They used Elamite for display and recording, but did not give the Elamite history from which they had emerged any other prominence.

### 1.2.4.1 *Achaemenid Elamite texts*

Elamite was the first language used by the Achaemenids for formal inscriptions. The Elamite version of the great inscription of Darius I at Bisitūn (Behistān), near Kermānshāh, was the first and for a short time the only version on the rock face. In later royal inscriptions, however, the Elamite always accompanies an Old Persian text to which it usually corresponds very closely. The inscriptions are on prepared rock faces, on architectural elements, reliefs and sculpture from royal residences, on a small number of portable objects and cylinder-seals. Most Achaemenid administrative texts belong to two archives excavated at Persepolis, from about 500–450 BC, but the contemporary pieces from Susa and Old Kandahar imply wider use of Elamite recording (Steve 1992:23–24; add Garrison 1996 [Achaemenid administrative text from Susa], Scheil 1939, No. 468 [administrative text from Susa, probably Achaemenid], Helms 1982:13, 1997:101 [Elamite administrative fragments from Old Kandahār]).

### 1.2.5 Later Elamite

Under Hellenistic and Parthian rule, Elam continued to be a geographical and cultural entity, mostly called “Elymais” in Greek sources, but without leaving a continuing record of the Elamite language. In the tenth century AD the geographer Iṣṭahrī mentioned an unaffiliated language spoken in Khuzestān, called Khūzī, and Muqaddasī added that Khūzīs spoke an incomprehensible language, said by Muḥammad to be devilish, but whether a survival of Elamite lies behind these remarks is doubtful (Cameron 1948:18, n. 115; Khačikjan 1998:1).

## 1.3 Status of Elamite in antiquity

The Sumerian king Shulgi of Ur (c. 2000) claimed that he knew Elamite well enough to answer Elamite messengers in their own tongue (Civil 1985:73), but Hammurabi of Babylon (c. 1750) listed Elam among distant mountain lands which had languages that were “twisted” (Gadd *et al.* 1928:44–45, No. 146), a perception of outlandishness also reflected in the Old Babylonian Elamite incantations. Later Mesopotamian scholarly texts characterized plants, tools, or wagons as Elamite, correlated an Elamite calendar with the standard Sumero-Babylonian calendar, and glossed a few Elamite words, but apparently gave little attention to Elamite language.

Since the earliest Elamite texts include probable literary or scholarly pieces, Elamite may have been used more widely as a language of learning than the known sample suggests. Moreover, the writing of Elamite for display and recording may have been more widespread at an earlier date among Elamites of highland Iran than the known sample, dominated by texts from the Mesopotamian border, would indicate. In the known sample, Elamite became the preeminent language for the display inscriptions of Elamite rulers after about 1400 and for administrative and legal recording after about 1100, and by about 600 it was also used for scholarship and for international correspondence.

The hypothesis of Achaemenid “alloglottography” (Gershevitch 1979) holds that Achaemenid Elamite was a medium for transmitting texts that were conceived and dictated in Old Iranian languages, to be read out and understood as Old Iranian texts, and hence that the use of Iranian words and congruence between Elamite and Iranian morphology or syntax are not matters of borrowing or interference, but explicit notations of the underlying text. This hypothesis (which has been neither widely embraced nor rebutted) implies a literate bilingual or multilingual population who knew a living version of Elamite.

## 1.4 Elamite dialects

Dialects of Elamite have been postulated to account for variations in syntax (Grillot-Susini and Roche 1987:11; Grillot-Susini 1994:1; Khačikjan 1998:47 n. 129), but no dialects have been identified or described. Of the main chronological periods, most descriptive attention is given to Middle Elamite and Achaemenid Elamite. Neo-Elamite has not been systematically analyzed, although it is represented by the largest variety of text types and might allow discrimination between chronological development and dialect differences.

The frequent characterization of Elamite as “poorly understood” means in practice that sharp differences in the translation of individual Elamite texts reflect disagreements about grammar and lexicon. Behind these disagreements lies a nearly complete consensus on the identification of morphemes and paradigmatic sets of forms, as well as a general agreement that knowledge of Elamite phonology is seriously limited. The main areas of disagreement are on the meaning of particular morphemes, especially the verbal auxiliary *ma-*, the

verb- and clause-suffixes *-t(a)* and *-a*; on the construction of pronouns and pronoun clusters with verbs and directional elements; on the understanding of morphological or syntactic differences between Middle and Achaemenid Elamite; and on the meanings of words.

## 2. WRITING SYSTEMS

### 2.1 Proto-Elamite

The earliest texts from the area where Elamite was spoken and written appear in scripts called *Proto-Elamite* and *Linear Elamite*. Neither script has been deciphered. It is plausible but not provable that both scripts rendered versions of Elamite language.

Proto-Elamite writing was so named at a time when “Elamite” was mostly used as a geographical term, not as the name of a language, so the name “Proto-Elamite” originally described texts without ambiguity – the first texts from Elamite territory, but not necessarily in the language that came to be called Elamite. Proto-Elamite writing was impressed or incised on clay tablets. About 1,600 texts are known, most of them from Susa, others from sites across southern and eastern Iran, as far south as Kermān and as far east as Seistān. The tablets are from archeological contexts dated c. 3100–2900 BC. Most of the tablets, perhaps all of them, are administrative records, having clear entries with groups of signs followed by groups of numerals, sometimes with a corresponding total on the reverse. They use sexagesimal and bisexagesimal systems that are identical with approximately contemporary Proto-Cuneiform texts from Mesopotamia. They also use a decimal system that is without a parallel in archaic Mesopotamian texts. About 5,000 attested forms of nonnumerical characters (a few clearly pictographic, most abstract patterns) probably represent about 1,000 signs or less, with paleographic variations (Brice 1962–1963; Meriggi 1971–1974; Friberg 1978–1979; Vallat 1986; Damerow and Englund 1989; Englund 1996, 1998).

### 2.2 Linear Elamite

This script is known from eighteen inscriptions carved on stone objects and incised on clay objects, and one inscription punched on a silver vase. Most are from Susa, one from Fārs and one from southeastern Iran. One occurs with a counterpart text in Old Akkadian (perhaps not a close translation) in the name of Puzur-Inšušinak, c. 2100 BC (see §1.2.1). Most or all of the texts are probably dedicatory inscriptions. Only 103 sign forms are attested, 40 of them attested only once (Hinz 1969:11–44, 1975a; Meriggi 1971–1974: I 184–220; André and Salvini 1989; Salvini 1998).

### 2.3 Elamite cuneiform

Readable Elamite texts are written in versions of the same cuneiform script that was developed in Mesopotamia to write Sumerian and Akkadian from the early third millennium BC on, and that was also adapted to write Eblaite, Hittite, Hurrian, and Urartian. The first progress of the nineteenth-century decipherers of cuneiform scripts came from work on inscriptions of the Achaemenid Persian kings in Old Persian, Akkadian, and Elamite. The decipherers recognized that the Akkadian and Elamite versions were written in two varieties of a single script. Hence, when the readings of the Akkadian texts were confirmed, they were also applied to Elamite cuneiform. Evidence from the Elamite versions themselves, however, did not contribute to the decipherment.

Like other versions of Mesopotamian cuneiform script, Elamite cuneiform includes several types of characters: those with syllabic values (syllabograms), those indicating words (logograms), unpronounced characters indicating semantic categories (determinatives), and numerals. Some symbols belong to more than one of these categories; some have more than one syllabic value; and some syllabic values are represented by more than one sign or sequence of signs. Regarding the last two points: in all periods, *polyphony* of signs (that is, single signs with two or more syllabic values) and *homophony* of signs (that is, two or more signs with the same syllabic value) are less common than in Mesopotamian scripts, and more often limited to writings of particular words or sequences. In Achaemenid Elamite, homophony and polyphony are almost (but not entirely) eliminated (Steve 1992).

### 2.3.1 Syllabograms

Syllabic symbols occur having the values V (vowel), VC (vowel + consonant), CV, and CVC (including C<sub>1</sub>VC<sub>1</sub> and C<sub>1</sub>VC<sub>2</sub>). Almost all syllabic values of Elamite signs are the common values of the same signs in Mesopotamian cuneiform; a few are uncommon in Mesopotamia and specialized in Elamite; and a few are unique to late Elamite writing. Mesopotamian VC and CVC signs do not distinguish between voiced, voiceless, and emphatic final stops, and some CV signs do not distinguish between voiced and voiceless initial stops; the counterpart Elamite signs also do not represent a corresponding distinction between stops.

### 2.3.2 Logograms

As in Mesopotamian cuneiform, almost all logograms are Sumerograms, that is, historical writings of Sumerian words used to indicate words with the same meaning in Akkadian or in Elamite. The Elamite words written with Sumerograms are sometimes unknown (e.g., Sumerian DUMU, “son,” Akkadian *māru*, Elamite *šak*; but Sumerian ŠE.BAR, “barley,” Akkadian *uṭṭētu*, Elamite uncertain). Akkadian loanwords appear in Elamite, but Elamite cuneiform lacks Akkadograms of the kind found in Hittite cuneiform (see Ch. 18, §2).

### 2.3.3 Determinatives

Most determinatives precede the words they qualify. The postpositive determinatives found in Mesopotamian cuneiform (for example, marking the preceding words as names of birds or plants) do not occur in Elamite. Some determinatives have the same value as the counterpart signs in Mesopotamian cuneiform: for example, signs that mark the following words as divine names, as personal names, as feminine personal names or words describing women, or as wooden things. Others are Mesopotamian signs used with determinative values specific to Elamite cuneiform: for example, a horizontal wedge to mark a following place name or location (commonly indicated in Mesopotamian cuneiform with different sign, postposed). The only postpositive determinative is the sign that in Mesopotamian cuneiform has the value MEŠ and marks the preceding word as a plural, but in Elamite cuneiform marks the preceding word as a logogram (this usage is also found with lower frequency in some so-called “peripheral” cuneiform writing – that is, cuneiform orthography for Akkadian in non-Akkadian speaking environments, e.g., Ugarit and Nuzi – and in Neo-Assyrian; van Soldt 1991:428–429). Postposed MEŠ also marks some pseudologograms (that is, historical spellings of Elamite words, e.g., Achaemenid Elamite *puhu* “boy,” *ulhi* “house,” both with nonphonemic *-h-*), and MEŠ sometimes appears after apparently ordinary syllabic spellings

(e.g., *tar-mu*<sup>MEŠ</sup> [a grain]). Conversely, not all logograms are followed by MEŠ (e.g., EŠŠANA = *sunki*-, “king,” DUMU = *šak*-, “son” are never followed by MEŠ [Vallat 1987a]).

### 2.3.4 Direction and division

As in Mesopotamian cuneiform and other adaptations of it, writing runs left to right, top to bottom. Word division is not ordinarily marked. Determinatives do not double as word-dividers, since most of the preposed determinatives also have common syllabic values (e.g., GIŠ [determinative for wooden objects, fruits, etc.] is used syllabically with the value *iz*), and postpositive MEŠ may be followed by signs indicating grammatical morphemes (e.g., LÚ (= *ruh*) “man” in <sup>DIS</sup>LÚ<sup>MEŠ</sup>-*ip*, “men”). In most Elamite texts, lines of writing are not divided at word boundaries, as they are in Mesopotamian cuneiform.

### 2.3.5 Graphemic inventories and spelling practices

The inventory of Mesopotamian cuneiform signs and the uses of the signs were adapted for writing Elamite. Most of the adaptations were motivated by economy, few if any by specific properties of the Elamite language. In all periods, Elamite used a smaller inventory of cuneiform signs than Mesopotamian scripts; a little more than 200 signs are attested overall. For any period, only 100–140 signs are attested.

The forms of cuneiform characters found in Old Elamite, Middle Elamite, and early Neo-Elamite texts are similar in composition and general appearance to forms in contemporary Mesopotamian scripts, with very few idiosyncrasies. Forms of many signs in Neo-Elamite texts after about 650 BC and in Achaemenid Elamite inscriptions and tablets are sharply and systematically distinct from forms in contemporary Mesopotamian scripts. To a modern eye, the difference is perhaps as great as the difference between standard and *Fraktur* forms of the Roman alphabet.

Royal inscriptions, which dominate the corpus of Old and Middle Elamite texts, use few logograms. Administrative texts, numerous only in Neo-Elamite and Achaemenid Elamite, use many. Conversely, Middle Elamite and Neo-Elamite inscriptions use more syllabic signs, with more syllabic values, than Achaemenid Elamite texts. Logograms are not used to write verbs, rarely used to write adjectives (other than “big” and “small” or “male” and “female”), and never used to indicate grammatical categories (such as plurality or noun derivation).

Loss of some CV symbols made it impossible to mark a consistent distinction between two kinds of labial, palatal, and dental stops consistently (utilizing the signs that distinguished voiced from voiceless in Akkadian cuneiform). Furthermore, loss of some VC values, mostly for sonorants and fricatives (*up*, *us*, *uš*, *al*, *ar*), made it impossible to write certain CVC sequences with the unambiguous combination CV<sub>1</sub>-V<sub>1</sub>C. These sequences were commonly represented with “broken writings” of the type CV<sub>1</sub>-V<sub>2</sub>C, in which V<sub>2</sub> is always *i* or *u*: for example, late Neo-Elamite, Achaemenid Elamite *du-iš* versus Middle Elamite *du-uš* for *duš*, “he received.” Similar broken writings were even used when not required by the inventory of syllabic signs: for example, singular *šá-lu-ur* and *šá-lu-ir* (not required); plural *šá-lu-ip* (required) “gentleman/men”; singular *li-ba-ir* (required), plural *li-ba-ap* and *li-ba-ip* (not required), “servant(s).” Some word-final variations between required broken spellings and “harmonic” spellings with different vowels, however, may represent loss of vowel distinction or presence of consonant clusters at ends of words: for example, *du-nu-iš* (required), *du-na-iš* (not required), *du-na-áš* (harmonic), “he gave” (Justeson and Stephens 1994).

**Table 3.1 Middle Elamite and Early Neo-Elamite (before c. 650 BC) syllabic signs: V, CV, VC**

V Symbols: a e i u, ú

CV Symbols				VC Symbols		
ba	be	bi		ap	ip	up
pa		pi	pu			
ga		gi	gu?	ak	ik	uk
ka, ka <sub>4</sub>		ki	ku			
da		di	du	at	it	ut
ta	te	ti	tu, tu <sub>8</sub> ?			
sa		si	su	as	is	us
za		zi, zí	zu			
ša, šá	še	ši	šu	aš, áš	iš	uš
ma	me	mi	mu	am	im	um
na	ne	ni	nu	an	en	in
la		li	lu	al	el	il
ra		ri	ru	ar	ir	ur
ha		hi	hu	Vh, V?		

**Table 3.2 Late Neo-Elamite (after c. 650) and Achaemenid Elamite syllabic signs: V, CV, VC (values in parentheses are not attested in Achaemenid)**

V Symbols: a e i u, ú

CV Symbols				VC Symbols		
ba	be					
pa		pi	pu	ap	ip, íp	
		gi				
ka <sub>4</sub>		ki	ku	ak	ik	uk
	te	ti	tu, tu <sub>4</sub>	at	it	ut
da			du			
sa		si	su	as	is	(us)
za		zí				
šá, šà	še	ši	šu	áš	iš	
ma	me	mi	mu	am	im	um
na		ni	nu	an	en	in
la		li	lu		el?	ul
ra		ri	ru	(ar)	ir	ur
ha		hi	hu	Vh		

In Achaemenid Elamite, as in late Mesopotamian cuneiform scripts, CVC signs may be ambiguous as to vowel color (e.g., *tup-pi-ra*, *tup-pi-ip* ~ *ti-pi-ra*, *ti-pi-ip*, “scribe(s)”; *šá-tin* ~ *šá-tan*, “priest”). In Achaemenid Elamite, and sometimes earlier, as in Mesopotamian cuneiform, CVC sequences are sometimes made unambiguous by plene writings of the types CVC-VC- (e.g., *tan-an-* beside *tan-* and *da-an-*), CV-CVC (e.g., *-ri-ráš-*) or CVC-CV- (e.g., *gal-li-*, *gal-lu-* beside *gal-*; *hal-la-tam<sub>5</sub>-ti* beside *hal-tam<sub>5</sub>-ti*, “Elam”; see Vallat 1989).

Late Neo-Elamite and Achaemenid Elamite introduced some syllabic values not found in Mesopotamian (e.g., *mak<sub>x</sub>* [KUR], *tam<sub>x</sub>* [GIM]), as well as one syllabic character not found in Mesopotamian cuneiform (*ra<sub>x</sub>* [from SAL+BAR]), and two determinatives: (i) the horizontal wedge (equivalent to AŠ) to mark place names, words indicating locations, and certain other words (e.g., “month” and “day”); and (ii) the signs BE and HAL, graphic variants of each other, to mark personal names and words indicating persons. Neo-Elamite and Achaemenid Elamite variants of some logograms betray misunderstanding of their Mesopotamian graphic etymologies: for example, Neo-Elamite E.GAL for historically correct É.GAL, “palace”; Achaemenid Elamite SI.KAK (once, probably erroneously) beside historically correct ŠI.KAK, “spear”; and the Sumerograms ANŠE “equid,” GEŠTIN “wine,” and NUMUN “seed” treated as combinations of two signs (PA+x, DIN+KAK, NU+MAN, respectively), sometimes separated by line divisions.

### 2.3.6 Transliteration and transcription

Elamite forms are represented below in sign-by-sign transliteration, morphological transcription, or conventional transcription. In *transliteration*, hyphens mark off syllables, logograms are in capitals, and determinatives are superscript (e.g., *hu-ud-da-an-ti*, <sup>d</sup>IT<sup>MEŠ</sup>). In *morphological transcription*, placed within square brackets herein (not within slanting brackets, as often, in order to avoid confusion with phonemic representation), hyphens mark off morphemes, and parentheses sometimes indicate vowels or doubled consonants that are inherent in syllabic writings but are apparently not morphemic (e.g., [hut<sup>a</sup>-n-t(i)]). *Conventional transcriptions* are commonly used representations that reflect underlying transliterations but do not consistently reflect inferred phonology or morphology (e.g., singular *hupirri*, plural *hupibe*, written *hu-pír-ri*, *hu-pi-be*, probably to be analyzed [hupi-r(i)], [hupi-p(e)]). Unattested or reconstructed forms are marked with \*.

The following abbreviations are used: DN (divine name); GN (geographical name); PN (personal name); RN (royal name).

## 3. PHONOLOGY

The use of Mesopotamian cuneiform script presents obstacles to recognizing Elamite phonology. In ordinary use for writing Akkadian, the script distinguishes only three vowels consistently (*a*, *i*, *u*) and a fourth in some sequences (*e*); it does not render initial or final consonant clusters or medial clusters of more than two consonants unambiguously; it does not distinguish voicing of syllable-final stops. The simplification of the script for use with Elamite further narrowed the possibilities of expressing distinctions. Changes in Elamite phonology were not necessarily accompanied by corresponding changes in writing; thus, although *h* was probably no longer phonemic in Achaemenid Elamite, Achaemenid Elamite writing retained a complete set of hV signs, a Vh sign, and some hVC signs, and *h* is written frequently, in some words regularly. Writing conventions for expressing phonological features peculiar to Elamite are not easily recognized or interpreted. The greatest obstacle to understanding Elamite phonology and its phonetic realization, however, is the lack of a securely identified close cognate language with a well-known phonology.

Resources for the study of Elamite phonology include transcriptions of words and names from other languages in Elamite texts (Iranian words and names in Achaemenid Elamite have been much studied (see, e.g., Hinz 1975b; Mayrhofer 1973; Tavernier 2002), but Akkadian and West Semitic words and names in Achaemenid and earlier Elamite have not);



transcriptions of Elamite words and names into other languages and scripts (words and names in Sumerian and Akkadian texts from Mesopotamia and Iran contemporary with Old Elamite and Middle Elamite have not been studied comprehensively as sources for phonology); and spelling variations within Elamite texts.

Much of this evidence, especially spelling variation, is ambiguous in that it may support either inferences about phonology or inferences about writing conventions. Conversely, where the writing does express phonemic distinctions that do not have counterparts in Mesopotamian languages, it cannot make their phonetic realizations plain.

### 3.1 Consonants

The consonantal inventory of Elamite is summarized in (1), though this summary is qualified below:

#### (1) Elamite consonants

p	t	k
b	d	g
	s	š
	z	
	v/f(?)	h
m	n	
	l	r

#### 3.1.1 Stops

There are two series of stops, ordinarily indicated in transliteration and transcription by *p*, *k*, *t* versus *b*, *g*, *d*. Elamite syllabaries do not allow consistent distinction of all pairs in all positions. Regular geminate spellings of medial stops in some words (e.g., *hutta-* (not \**hu-ta-*) “do,” *-ikki* “to” versus *igi* “brother”) and regular choices of initial signs in others (e.g., *pari-* (not \**ba-ri-*) “go”) indicate that a phonemic distinction was made. However, spelling variations within Elamite (e.g., *dumanpi*, *dumanba* but not \**dumanpa*) and Elamite transcriptions of foreign words and names (e.g., Middle Elamite *pi-it* for Akkadian *bīt(u)*, “house,” Achaemenid Elamite *Ba-ir-šá* (never \**Pa-*) for *Pārsa* “Persepolis”) indicate that the two series were not distinguished by voicing. A contrast between tense (rendered with *p*, *k*, *t*) and lax (rendered with *b*, *g*, *d*) stops, as in Dravidian, is sometimes suggested (e.g., Reiner 1969:115; Khačikjan 1995).

##### 3.1.1.1 Allophonic variation

Spelling variations like *Šu-šu-ga* ~ *Šu-šu-un-ka*, “Susa (+ marker of grammatical concord),” *šu-ul-lu-me-ka* ~ *šu-ul-lu-me-en-ka* (a verbal form of uncertain meaning), *hi-nu-ka* ~ *hi-nu-un-ka* “(which) we (will) have,” and perhaps *su-un-ki-ir* ~ *su-gir* “king” (all Middle Elamite) suggest nasal allophones of the velar series. Late Neo-Elamite royal inscriptions from Khuzestān that spell a final first-person morpheme *-k* (below) with signs containing *h* suggest spirantization of the velar (Khačikjan 1995:109; Vallat 1996a:387). An affricated pronunciation of dentals may lie behind an Achaemenid Elamite spelling *zī-da-el* ~ *zī-za-el* (Hinz and Koch 1987:1288), and Akkadian writings of Elamite names of the eighteenth–seventeenth centuries BC (hence contemporary with Old Elamite) with such variations as *tempti-* ~ *šimti-* “lord,” *kutir-* ~ *kušir* ~ *kusir-* “carrier” (Zadok 1984:3; Vallat 1996c:315).



### 3.1.2 Fricatives

At least three fricatives (sibilants), transcribed as *s*, *š*, and *z*, are expressed with signs which have common Mesopotamian values including *s*, *š* and *z* (or *š*).

Variations between spellings with *s* and *š* (e.g., *Insušnak* ~ *Inšušinak* [a divine name], *mušika* ~ *muššika* ~ *musika*, “it is counted”) suggest that *s* can represent an affricate. Moreover, variations between spellings with *š* and *z* (e.g., *Anšan* ~ *Anzan* [a place name], both in Elamite and in Akkadian), along with the use of signs with *z* to transcribe Iranian /č/, and the Achaemenid Elamite spellings *ku-ti-iš* and *ku-iz*, for [kutš], “he carried,” suggest the existence of an Elamite phoneme /č/. However, the spellings *ku-iz-iš-da* and even *ku-iz-da-ti-iš-da* may suggest that the writers perceived a cluster /-tšt-/ to be clarified with the same graphic convention used otherwise for CVC signs (§2.3.5).

In Old and Middle Elamite, syllabic symbols with Akkadian values including *ḥ* consistently represent a phoneme transcribed as *h*. Its phonetic value is uncertain, but it was not a velar fricative like Akkadian /ḥ/. Spellings cease to be consistent when this /h/ ceases to be phonemic in late Neo-Elamite and Achaemenid Elamite, though many historical spellings with *h* and sometimes with *-hh-* occur.

A labial fricative such as /f/ or /v/, not represented unambiguously or consistently, is suggested by the spelling variations *ligawe* ~ *likame*, *suhterwe* ~ *suhterme*, and Akkadian *Ši/Ši-we-*, *Še-ep-* ~ Elamite *Si-me-* (in the royal name *Sim/we-palar-huhpak*); see Khačikjan 1995:107, 1998:8.

### 3.1.3 Sonorants

Elamite possesses nasal and liquid phonemes; the phonemic status of glides is less clear.

#### 3.1.3.1 Nasals

Both /n/ and /m/ are unambiguously represented in Elamite spelling. Some words are regularly spelled with geminate *m* or *n*, but a phonemic distinction is uncertain.

From at least Middle Elamite on, /n/ was assimilated to following /l/ (e.g., /ullina/ < [un lina]) and perhaps to following palatal and dental consonants. In late Neo-Elamite and Achaemenid Elamite, /n/ was labialized before a bilabial stop and written as *m* (e.g., *tahhampa* < \**tahhanpa*; *sitmamba* ~ *sitmap*; *dumamba* and even *du-ma-ma* ~ *dumanba* (all plural forms on verbal stems), but also the exceptional *tah-ha-ma-am-ri*, perhaps back-formed from the plural; also in pronoun-verb phrases like *ú-um beša* “he (who) created me,” *um parimanka* “I will (not) be coming there”; see Paper 1955:62; Vallat 1996a:387–388). Achaemenid Elamite spellings *hu-ut-tan-ti* ~ *hu-ut-tam<sub>5</sub>-ti* do not indicate dissimilation, but reflect a graphic convention also found in late Mesopotamian cuneiform: *CVm* ~ *CVn* ~ *CV-Vn*.

#### 3.1.3.2 Liquids

The liquids /l/ and /r/ are written unambiguously. The writing of Akkadian *La-gamāl* as Elamite *Lagamar* and of Elamite *Ruhuratir* as Akkadian *Lahuratil* (both divine names), as well as Achaemenid Elamite *ka-ri-ri* ~ *kar-li* “lamb,” suggest a non-trilled [r] (Khačikjan 1995: 107f., 1998:8f.). The spelling variations *pi-ri-ip* ~ *pa-ri-ip* “they went to, reached,” *pa-ri-iš* ~ *pa-iš* “he/they set out, went,” and perhaps *mar-ri-ia* ~ *ma-ú-ri-ia* “I seized” (all Achaemenid Elamite) suggest a vocalic [ɾ].

### 3.1.3.3 Glides

The phonemic status of [y] and [w] is unclear. The intervocalic use of the syllabograms *-i-* or *-ia-* represents a glide [y]. In contrast, word-initial *ia-*, rare except in proper nouns, represents juncture between syllables or words: thus, *ia-ak* for *a-ak* “and” in the sequence *intikka yak*; *ia-áš-pu*, a *Kulturwort* corresponding to Akkadian *ašpu* (a semi-precious stone), in the sequence *rišakki i yašpu*. Word-initial *a-a* represents two syllables separated by a glide or juncture (*a-a-ni* ~ *a-hi-in*, *a-ah-in* “family(?)”, *A-a-pír* [a place-name]). In Achaemenid Elamite, the sequence *(-)ú-uC* also represents glide, syllable-boundary, or word-initial juncture (*hu-ut-ti-ú-ut* ~ *hu-ud-da-hu-ut* ~ *hu-ud-du-ud-da* “we made”; *hu-ut-ti-ip* ~ *ú-ut-ti-ip* “makers”). But the unique Achaemenid Elamite spelling *a-áš-šá-ir-ki-* for *Manšarki* (a month name) seems to suggest some phoneme with allophones [y], [w] and perhaps [ʔ].

## 3.2 Vowels

The vowels /a/, /i/, and /u/ are expressed unambiguously. The vocalic phoneme /e/ is confirmed by minimal pairs (e.g., *tetin* “beam(?)” vs. *titen-* “lying”) and supported by transcriptions of foreign words (e.g., *alumelu* from Akkadian *ālum ēlum*), but it is not often expressed unambiguously by the writing system. Final [-e] and [-i] were probably not distinguished phonemically. Many spellings with final *Ci* probably indicate final [-C], especially in clusters (e.g., *hu-ud-da-an-ti* for [huttant] “you do”). Disagreement prevails concerning the existence of phonemic /o/, sometimes postulated on the basis of distinctive uses of the signs *u* and *ú* (Paper 1955:17; Khačikjan 1998:6).

Contemporary variation in spellings using signs with *u* and signs with *i* in some words (e.g., *tu<sub>4</sub>-ru-iš* ~ *ti-ri-iš*, *mu-ši-in* ~ *mi-ši-na*, all Achaemenid Elamite) may reveal a common reduced allophone shared by /i/ and /u/. Variation in spellings of vowels in the final syllables (e.g., *dunuš* ~ *dunaš* “he gave”) may indicate a reduced vowel or a final cluster with sonorant.

Spelling variations like *Hu-ban* ~ *Hu-um-ban* (a divine name), *te-em-ti* ~ *te-ep-ti* “lord,” *na-ra-an-da*, *na-ra-an-te* ~ *na-ra-da*, *na-ra-te* “daily” suggest the existence of nasalized vowels.

Vowel length is not phonemic. Most long writings of vowels are susceptible to graphic explanations: for example, avoidance of one-sign spellings of open monosyllables (*a-ak* vs. *a-gi* for /ak/ “and”), or marking of final vowel versus final consonant cluster (*te-la-ak-ni-e* vs. *te-la-ak-ni* for /telakni/ not /<sup>\*</sup>telakn/).

Diphthongs do not occur. In Achaemenid Elamite, some spellings with *-a-uC* appear to reflect the pronunciation of following sonorants: *mauriya* ~ *marriya* (perhaps with vocalic [ɾ]) “I seized”; *zaumip* ~ *zammip* ~ *zamip* (perhaps with labial continuant; see §3.1.2) “laborers.”

### 3.2.1 Vowel contraction

Monosyllabic pronouns in clusters, and pronouns in constructions with directional elements were often susceptible to contraction and written without word-division. The sequence *i u* does not normally occur; *u i* contracts to *u*, and *i i* to *i*: thus, [li-n-a ap u in] written *li-na-pu-un*; [pat-r ir u-r] written *pa-at-ru-ur* (Reiner 1969:99, Grillot 1983:210, Grillot-Susini and Roche 1987:9).

### 3.3 Accent

Neutralization of some final vowels and elision of some medial vowels suggests that stress was nonfinal, probably initial (Grillot-Susini and Roche 1987:11, 1994:15; Khačikjan 1998:10).

## 4. MORPHOLOGY

### 4.1 Word formation

Elamite is an agglutinative language. Most roots are of one or two syllables, of the types CV (*da-* “place”), VC (*ki* “one”), CVC (*nap* “god,” *ruh* “man”), VCV (*igi* “brother”), CVCV (*zana* “lady”), and perhaps CVCCV (*sunki-* “king,” *tingi-* “take away” [or: CVCV *su<sup>n</sup>ki-*, *ti<sup>n</sup>gi-* ?]). Some roots produce only nominal forms, others both nominal and verbal forms. All inflection is marked with suffixes attached to a root or to a base derived from a root with the addition of a thematic vowel, a derivational suffix, by reduplication, or by compounding. Most trisyllabic bases can be identified as composites or loanwords (Grillot-Susini 1994:1–8).

### 4.2 Nominal morphology

Nominal inflection affects substantives, attributes of substantives (including clauses), demonstratives and pronouns, numerals, the negative particle, and some verbal forms (derived from the bare verb-stem (gerunds or participles), and from the “nominal conjugations” formed on the verb-stem with suffixed *-k-* (Conjugation II) and suffixed *-n-* (Conjugation III)).

#### 4.2.1 Gender, person, and number

Nominal inflection distinguishes two genders, animate and inanimate. Inflection of animates distinguishes three personal classes, corresponding to the three persons of verbal inflection. The first-person (I-class) form is sometimes called *locutive*; the second-person (you-class) *allocutive*; and the third-person (he-it-class) *delocutive*. Inflection of third-person animates distinguishes singular and plural. These suffixes mark agreement (i) between subject and verb, and (ii) between parts of possessive and attributive constructions and subordinate clauses (see below and §5.2); the gender/person/number suffixes are as follows:

(2)	<i>Animate</i>		
	<i>Singular 1st</i>	-k	([sunki-k] “I, king”)
	<i>2nd</i>	-t	([hutta-n-t] “you, doing” [katu-k-t] “you, living”)
	<i>3rd</i>	-Ø	([nap] “he, god,” [zana] “she, lady”)
		-r	([nap-ir] “he, god,” [sunki-r] “he, king”)
	<i>Plural 3rd</i>	-p	([nap-ip] “they, gods,” [sunki-p] “they, kings”)
	<i>Inanimate</i>		
	<i>3rd</i>	-Ø	([hal] “town, land,” [mur] “place”)
		-me	([sunki-me] “kingdom, kingship”)
		-n	([siya-n] “temple,” [muru-n] “earth”)
		-t	([hala-t] “clay, mud brick”)

Third-person suffixes are derivational. The animates indicate agent nouns (*huttira* “maker, doer”), members of a class, or persons (*Babilira*, *Babilip* “Babylonian(s)”; *libar*, *libap* “servant(s)”). The inanimate *-me* indicates abstracts (*takkime* “life”). In Achaemenid Elamite, *-ta* ~ *-te* indicates generality (*marrita* “everything”). Doublets are common: thus, Achaemenid Elamite [muši-n] ~ [muši-me] “account.”

In possessive and attributive constructions, the suffixes appropriate to the possessor or the determined substantive are added to the possessed or attribute. Consider the following Middle Elamite examples:

- (3) A. [u PN šak PN<sub>2</sub>-k(i) sunki-k GN-GN<sub>2</sub>-k(a)]  
 “I, PN, son of PN<sub>2</sub>, king of GN (and) GN<sub>2</sub>”  
 with first-person suffixes throughout  
 B. [PN meni-r GN ak GN<sub>2</sub>-r(i) šak hanik PN<sub>2</sub>-r(i) ak PN<sub>3</sub>-r(i)]  
 “he, PN, ruler(?) of GN and GN<sub>2</sub>, beloved son of PN<sub>2</sub> and PN<sub>3</sub>”  
 with third-person suffixes throughout

In Neo-Elamite a postposition *-na* (derived from the neutral inanimate *-n* with final “relative” *-a*), sometimes expresses possession, and in Achaemenid Elamite most possession and some attributive relationships are expressed with *-na*: Neo-Elamite [zalmu PN-na] “image of PN”; Achaemenid Elamite [halmi PN-na] “seal[ed document] of PN.”

#### 4.2.2 Case

Only personal pronouns (see §4.3.2) are marked for nominal case, distinguishing between an object-case and a subject/indirect object-case. Other spatial relationships and relationships between nouns and verbs are expressed with resumptive pronouns and with postpositions attached to nouns, to noun phrases, or to clauses.

#### 4.2.3 Indeclinable nominals

Kinship terms in which possessive or attributive relationships are inherent (*šak* “son,” *puhu* “child,” *igi* “brother,” *šutu* “sister,” *amma* “mother,” *rutu* ~ *riti* ~ *irti* “wife,” *ruhušak* “sister’s son”) are indeclinable; that is, they do not have markers of gender and person where other nouns have such markers (Reiner 1969:88). As the possessed noun in some possessive constructions, they are marked with nominal suffixes that refer to the possessor: Neo-Elamite, Achaemenid Elamite [PN šak-r(i)], [PN riti-r(i)] (Grillot-Susini and Roche 1987:23).

#### 4.2.4 Adjectives

Elamite adjectives do not constitute a distinct morphological class. They are marked with the personal suffixes and postpositions that express attributive and possessive constructions, including the personal marker of the modified substantive ([*temti riša-r*] “great lord”; [*upat lansiti-p(a)*] “brickwork (anim. pl.!) of gold,” i.e., gilded or enameled?); and the possessive postposition *-na* ([*sunki-na*] “of the king,” i.e., “royal”), productive in Achaemenid Elamite: e.g., GURUŠ-*na* “male”; MUNUS-*na* “female” (the Elamite words underlying the logograms are unknown); *punna*, *berna*, etc. (qualifying animals). There are no comparative or superlative forms. Superlatives are expressed with a possessive construction: Achaemenid Elamite [*akka irša-r-a napi-p(e)-na*] ~ [*akka irša-r-a nap-b(e)-r(a)*] (corresponding to Old Persian *haya maθišta bagānām*) “[Ahuramazda] the greatest of the gods”; Middle Elamite [*riša-r napi-p(i)-r(a)*] “[Inšušinak], greatest of the gods.”

### 4.3 Pronouns

Elamite has demonstrative, personal, possessive, relative, indefinite, and resumptive pronouns.

#### 4.3.1 Demonstrative pronouns

The Middle Elamite demonstrative pronouns are *hi* ~ *i* (animate singular and inanimate) and *ap* ~ *api* (animate plural). Achaemenid Elamite distinguishes between near-deictic *hi* ~ *i* and *ap* “this, these,” and far-deictic *hube* (inanimate), *hupirri* (animate singular), *hupibe* (animate plural) “that, those.” The demonstrative pronouns also serve as third-person personal pronouns.

#### 4.3.2 Personal pronouns

The personal pronouns distinguish an “unmarked” nominative/dative form for subjects or indirect objects, and a “marked” accusative form for direct objects:

(4)	<i>Singular</i>		<i>Plural</i>	
	<i>Nominative</i>	<i>Accusative</i>	<i>Nominative</i>	<i>Accusative</i>
<i>1st</i>	u	un	nika ~ nuku	nukun
<i>2nd</i>	ni ~ nu	nun	num ~ numi	numun
<i>3rd</i>	i ~ hi	ir ~ in	ap ~ ap(p)i	ap(p)in
<i>Inan.</i>	i ~ in	i ~ in		

In Achaemenid Elamite, first-person singular accusative pronouns written *unan*, *unahan*, *unanku* ~ *uhanaunku* also occur. Analysis of them is a matter of disagreement (Paper 1955:95 and Khačikjan 1998:22). Also in Achaemenid Elamite, *ha-ap* appears once as a variant spelling of *ap*.

#### 4.3.3 Possessive pronouns

Possessives of the personal pronouns are formed like other possessive constructions, by adding the suffixes appropriate to the possessor (see §4.2.1) or by adding the possessive postposition *-na* ~ *-ni*: Middle Elamite [napir-u-r(i)] “my god,” [sunkip urip-u-p(e)] “kings, my predecessors,” [takkime puhu nika-me-na ~ nika-me-me] “the live(s) of our children”; Achaemenid Elamite [ulhi nuka-me] “our house,” [libar-u-r(i)] “my servant,” [libar-e-r(i)] “his servant,” [sunkime appi-ni] “their kingship (= rule over them),” but first-person singular with an enlarged base [libame u-ni-na] “my servitude (= servitude to me)” and first-person plural without animate/inanimate distinction [kir akkayaš nuka-me] “one colleague of ours.”

In addition, there is a third-person animate singular possessive suffix *-e* that may derive from the pronoun *hi* ~ *i*, without suffix: Middle Elamite [PN ak puhu-e] “PN and her children”; Middle Elamite, Achaemenid Elamite [hiš-e] “his name.” A corresponding third-person plural animate possessive is formed by adding *-e* to the demonstrative/personal pronoun: Middle Elamite [hiš(-)api-e] “their name”; Achaemenid Elamite [puhu appi-e] “their boys.” Hinz and Koch 1987 diverge from Hallock 1969 and others in interpreting final *-še* in Achaemenid Elamite writings of substantives of Iranian origin as representing the Old Persian possessive *-šay*, rather than as including Elamite *-e*.

Achaemenid Elamite also has a first-person possessive suffix *-ta* (only in the construction [u atta-ta] “my father”) and a second-person singular possessive suffix *-ni* (NUMUN-*ni* “your lineage,” [širi-ni] “your s”).

#### 4.3.4 Relative pronouns

The Elamite relatives are animate *akka* “who” and inanimate *appa* “which, what.” A corresponding animate plural *akkap(e)* also appears in Neo-Elamite and Achaemenid Elamite. In Achaemenid Elamite the inanimate form doubles as the accusative of the animate: *appi* 9 *sunkip appa u... mauriya* “these are the nine kings whom I captured.” In Achaemenid Elamite, the relative pronouns appear as calques on the Old Persian relative pronouns and articles, *haya/hayā/taya*, connecting substantive and attribute or possessor and possessed pronouns; such calques are frequent in multilingual royal inscriptions (PN *akka Makuš* “PN the Magian,” *taššup appa* PN-*na* “the troops of PN,” *taššup appa unina* “my troops”); the usage also occurs in administrative texts (PN *akka* GN-*ma kurdabattiš* “PN the chief of workers at GN”). Occasional uses of the relative pronouns in expressing dates, however ([<sup>d</sup>ITI<sup>MEŠ</sup> *appa* NN-*na-ma*] “in the month of NN,” *bel appa* 24-*ummemana* “the 24th year”), do not have Old Persian parallels.

The inanimate substantive *mur(u)*, unmarked and undeclined, serves as the locative relative “where”: Middle Elamite [muru huma-hš(i)-ta in-me durna-h] “where they took (it) I do not know”; Achaemenid Elamite [mur halmarraš hi kuši-k-a] “where this fortress is built.”

#### 4.3.5 Indefinite pronouns

An animate indefinite pronoun, “anyone,” is formed from the relative *akka* with personal suffix *-r*; it occurs in negated clauses: for example, Middle Elamite [sunki-p uri-p-u-pi akka-r(a)... in-r(i) hutta-n-r(a)] “(what) former kings, any (of them) did not do,” i.e., “what no former king did”; Achaemenid Elamite [appa-n-lakki-me akka-r(i) inni hutta] “I did not commit a trespass against anyone.”

The inanimate indefinite *aški* “anything,” also found in negated clauses, is perhaps formed with the numeral *ki* “one” (Hinz and Koch 1987:88; Khačikjan 1998:29; otherwise Hallock 1969:670).

#### 4.3.6 Resumptive pronouns

Nominal constituents of a clause are frequently “resumed” by one or more pronouns placed immediately before the verb at the end of the clause. In Middle Elamite these resumptives are in clusters: [ap u in (written *a-pu-un*) duni-h] “to them [the gods] I gave it [the temple].” In contrast, Achaemenid Elamite normally allows only a single resumptive to precede the verb: *u* DN *un nušgišni* “I, may Ahuramazda protect me”; *u* PN *ir halpi* “I, PN, I killed him”).

The element *aha* (Middle Elamite, Neo-Elamite) ~ *ah* (Neo-Elamite, Achaemenid Elamite) ~ *ha* (Achaemenid Elamite) also appears before the verb at the end of a clause, replacing or, less often, preceded by resumptive personal pronouns. In Achaemenid Elamite it is commonly transcribed as a proclitic. In Middle Elamite it sometimes takes nominal suffixes *-r*, *-n*, or *-t* to mark concord. Characterizations of this formant disagree. On a narrow interpretation, it is a locative and only a locative, indicating “here,” “there,” or even both “here” and “there” contrasted in a single phrase. Some contexts are susceptible only to translation with locatives: Middle Elamite [ir aha-r murta-h] “I placed him [the image

of the god] in it [the temple],” expressed elsewhere [sian-r(a) ir murta-h] “his temple, I placed him” (see Grilhot-Susini and Roche 1987:20–21, but cf. Grilhot 1970:235 n. 40; Giovinazzo 1989:13–14). On a broad interpretation *aha* ~ *ah* ~ *ha* is a general oblique resumptive pronoun, referring to substantives of any gender and number, and indicating not only “in, at it” but also “to, for, with it” (see Hallock 1969:9, 1973:148 n. 4; Stolper 1984:25; Malbran-Labat 1995:80; cf. Khačikjan 1998:25). Some contexts are susceptible only to translation with nonlocatives: thus, Middle Elamite [upat... tepu-h ulhi i aha kuši-h] “I fashioned bricks, with them I built this house.” The comparison among Achaemenid Elamite *hupimer* “then, after that,” *hamer* “then,” and *hami* “there” favors identifying *ha* as demonstrative and pronominal. An agnostic view identifies Achaemenid Elamite *ha-* as a prefix or particle of uncertain function and meaning (Grilhot-Susini, Herrenschildt, and Malbran-Labat 1993:51; Tucker 1998:175).

In Achaemenid Elamite administrative texts *kaš* sometimes replaces *hi* as an oblique singular resumptive pronoun (Hallock 1969:9). Vallat (1987b), accounting for this non-paradigmatic form as a ghost word arising from the misreading of an archaic form of the sign *hi*, is not supported by collation.

### 4.3.7 Reflexive pronouns

The reflexive *du(h)-*, perhaps related to the verb *du-*, “take, receive,” occurs with possessive suffix *-e* in Middle and Neo-Elamite ([hiš duh-e] “his own name”) and in Achaemenid Elamite ([halpi duh-e-ma] “by his own death” (i.e., a natural death). In Achaemenid Elamite, it also forms an animate plural (also in possessive constructions, e.g., [GUD<sup>MES</sup> du-p-e-ma ~ du-p(i)-ni-ma] “for their own cattle” vs. pleonastic [GUD<sup>MES</sup> du appi-ni-ma]), and an animate singular object-case, like the personal pronouns (e.g., [du-n nušgiš] “protect yourself”).

In Achaemenid Elamite, the element *hisu* indicates emphasis of the subject of an action: *hisu x makiš* “he himself consumed x [grain].” It also appears with a “generalizing” inanimate suffix *-t* ([PN hisu-t(a) x du-ma-k-a] “x [grain] was received by PN himself”), but it is not marked for case or number.

### 4.3.8 Other pronouns

“Each, every” is expressed in Achaemenid Elamite with *unra* (referring to persons: 90 *kurtas unra* 20-irmaki *dušda* “90 workers received a twentieth [measure of wine] each”) and *lurika* (referring to animals and inanimates: UDU.NITA<sup>MES</sup> *lurika* x ŠE.BAR<sup>MES</sup> *ha-lika* “for each sheep x barley was delivered”). The form *unra* varies with *unra-na*, with the adjectival *-na* suffix.

“All” is expressed in Achaemenid Elamite by *marrida*, with the “generalizing” *-t* (*hupe marrida*... *hutta* “I did all that”), also *marribepda* ~ *marbepda*, with animate plural marker ([taššup marri-p(e)-p-t(a) ~ mar-p(e)-p-t(a)] “all the people,” but elsewhere *taššup marrida* (otherwise Hinz and Koch 1987, segmenting a word *marri*, plural *mar(ri)bep* from *da* “also”).

## 4.4 Nominalized negative particle

In Middle Elamite and Neo-Elamite, and exceptionally in Achaemenid Elamite, the negative particle *in-* takes nominal suffixes (first-person singular *in-ki*, third-person *in-ri*, \**in-pi*, inanimate *in-ni*, *im-me* (< \**inme*)) indicating concord with the logical subject (either the



subject of the verb or the subject of attention). In Achaemenid Elamite, the inanimate form *inni* is general: [taššup appa unina in-ni tiriman-p(i)] “people who do not call themselves mine.”

## 4.5 Verbal morphology

Verb bases are simple (*ta-* “put,” *dunu-* “give”), compound (*mur-ta-* “put in place”), or reduplicated. Reduplicated bases are mostly of the type  $C_1V_1C_1C_2V_2-$  (*beti-* > *bepti-* “rebel”), rarely of the form  $C_1V_1C_1V_1-$  (*li-* > *lili-* “give, deliver”) or the form  $C_1V_1C_1V_1C_2V_2-$  (*tallu-* > *tatallu* (earlier \**taltallu*) “write”). The change of meaning that reduplication conveys is not established; Steiner (1990:152–153) proposes plurality of action or patient.

### 4.5.1 Verb conjugations

Verbs produce three primary sets of forms labeled “conjugations”: one “verbal conjugation” (*Conjugation I*) and two “nominal conjugations” (most often called *Conjugation II* and *III*, also called *participles*, *paraverbal forms*, or *appellatives*). Particular verbs do not belong to a single conjugation; most verbs produce forms in more than one conjugation. All three conjugations distinguish three persons and two numbers. The *nominal conjugations* are formed by adding the suffixes that mark person, gender, and number in nouns (see §4.2.1). The *verbal conjugation* is formed by adding suffixes that are specific to verbs.

#### 4.5.1.1 Middle Elamite verbs

Conjugations I–III of Middle Elamite are presented in (5)–(7), utilizing *kulla-* “pray”; *hap(i)-* “hear”; *hutta-* “do”; *turu-* “say”; and *tahha-* “help(?)”:

#### (5) Conjugation I (verbal conjugation) – Middle Elamite

	<i>Singular</i>	<i>Plural</i>
<i>1st</i>	[kulla-h]	[kulla-hu]
<i>2nd</i>	[hap-t]	[hutta-h-t]
<i>3rd</i>	[hutta-š]	[hutta-h-š]

#### (6) Conjugation II (base + -k-) – Middle Elamite

	<i>Singular</i>	<i>Plural</i>
<i>1st</i>	[-k-k]	
<i>2nd</i>	[-k-t]	
<i>3rd animate</i>	[hutta-k-r]	[hutta-k-p]

#### (7) Conjugation III (base + -n-) – Middle Elamite

	<i>Singular</i>	<i>Plural</i>
<i>1st</i>	[hutta-n-k]	
<i>2nd</i>	[hutta-n-t]	
<i>3rd animate</i>	[hutta-n-r]	[tahha-n-p]

Since the personal suffixes on nouns include no first-person plural, no first-person plural form is expected in (6)–(7). Two clear first-person plurals with a suffix *-nunk* (*turununki* “we say,” *hinunka* “we get [children]”) may correspond to Conjugation II first-person singulars (*hinka*, Neo-Elamite *turunka*). There is, however, disagreement on the analysis of these



forms, and of a counterpart Achaemenid Elamite first-person plural on a base enlarged with auxiliary (-)ma-, *tiri(-)ma-nun* “we call ourselves” (summarized by Khačikjan 1998:36; Tucker 1998:188 n. 41).

#### 4.5.1.2 Achaemenid Elamite verbs

Conjugations I–III of Achaemenid Elamite are presented in (8)–(10), illustrated with *marri-* “hold”; *hutta-* “do”; *šinnu-* “come”; *kātu-* “live”; *na-* “say”:

#### (8) Conjugation I (verbal conjugation) – Achaemenid Elamite

	<i>Singular</i>	<i>Plural</i>
<i>1st</i>	[marri-Ø (~ -y, -ʔ)]	[hutta-Ø-ut] (written <i>-hu-ut</i> and <i>-ú-ut</i> )
<i>2nd</i>	[* -t]	[* -t]
<i>3rd</i>	[hutta-š]	[hutta-h-š]

As a result of the loss of phonemic /h/ and inconsistency in the writing of historical *h*, singular and plural were not distinguished in the third person – at least not distinguished in writing. A juncture or syllable boundary was still pronounced at the end of first-person singular forms, however, reflected in writings of forms with suffixed *-a* as *marriya*, *pariya*, *beliya*, *tengiyya*. The first-person plural form, marked with an enclitic *-ut* that also appears on nominal forms ([*sunkip-ut*] “we are kings”), was productive (Hallock 1973:151).

#### (9) Conjugation II (base + -k-) – Achaemenid Elamite

	<i>Singular</i>	<i>Plural</i>
<i>1st</i>	[šinnu-(k)-k-ut]	
<i>2nd</i>	[kātu-k-t]	
<i>3rd animate</i>	[hutta-k-Ø]	[šinnu-Ø-p]

The third-person forms expected from the Middle Elamite paradigm occur as nouns or attributive adjectives (inanimate *katuka*, animate singular *katukra*, animate plural *katukpe*) but not clearly as predicates (Tucker 1998:171–173). The ending of the first-person singular, always written *-gi-ut*, apparently contains the same particle *-ut* found in the Conjugation I first-person plural, and on nominal forms and phrases (*sunkir appi-ni-gi-ut* “I am king of them,” *titu-kur-ra-gi-ut* “I am (not) a liar”), where *-gi-ut* corresponds to Old Persian *āham* “I am” and is parallel to *ha-um*, an Elamite transcription of Old Persian *āham*.

#### (10) Conjugation III (base + -n-) – Achaemenid Elamite

	<i>Singular</i>	<i>Plural</i>
<i>1st</i>	[na-n-k]	
<i>2nd</i>	[na-n-t]	
<i>3rd animate</i>	[na-n-r]	[na-n-p]

For the first-person plural *tirimanun*, “we call ourselves,” see the Achaemenid Elamite forms noted in §4.5.1.1. A similar form, *hutti-nun* has been treated as a first-person plural of *hutta-* “do,” although it occurs in the phrase *hutti-nun-(h)uba*, corresponding to an Old Persian infinitive meaning “[in order] to do [battle]”; analysis of the form is disputed (summarized in Khačikjan 1998:37).

#### 4.5.1.3 Auxiliary and suffixed (-)ma-

Verb phrases occur in Middle and Neo-Elamite in which an auxiliary *-ma-*, with endings of Conjugations I, II, or III, follows either (i) a bare verb base ([*miši-ma-n-*] “becoming dilapidated”), or (ii) Conjugation II or III stems (Neo-Elamite *pali-k-ma-n-k*, *pera-n-ma-n-k*), or (iii) verbal nouns with animate marker *-r* (*pepsi-r-ma-h* “I renovated”). In Achaemenid Elamite, the element *-ma-* only follows the bare verbal-stem and precedes the personal suffixes, producing secondary sets of forms that are usually called Conjugations Im, IIm, and IIIm. Attested Achaemenid forms are presented in (11):

#### (11) Achaemenid Elamite secondary conjugations

		<i>Singular</i>	<i>Plural</i>
Conjugation Im	<i>1st</i>	<i>-ma-Ø</i>	
	<i>3rd</i>	<i>-ma-š</i>	
Conjugation IIm	<i>3rd</i>	<i>-ma-k</i>	<i>-ma-p</i>
Conjugation IIIm	<i>1st</i>	<i>-ma-n-k</i>	
	<i>3rd</i>	<i>-ma-n-ra</i>	<i>-ma-n-p</i>

Conjugation Im forms are rare, except for the verb *du-ma-* “receive.” Conjugation IIm plural forms are also rare.

#### 4.5.1.4 Conjugation functions

There is broad agreement on the distinctions of meaning among the conjugations, but authorities differ in emphasis on aspect, transitivity, and/or voice (perfective/imperfective, active/passive, taking one, two, or three arguments). Conjugation I is mostly active, transitive, sometimes intransitive (including verbs of motion and verbs of speaking), having neutral or absolute aspect, mostly of past tense. Conjugation II is mostly intransitive or passive, perfective in aspect hence often past. Conjugation III is transitive or intransitive, imperfective, non-past (see, among others, Hallock 1959; Grillot 1970:216–218; McAlpin 1981:71 and 80; Khačikjan 1998:33–36; see also Malbran-Labat 1990, distinguishing verbs with a single argument, with no Conjugation I, from verbs with two or three arguments in Conjugation I but fewer arguments in Conjugations II and III).

There is only partial consensus on the meaning of auxiliary (-)ma- (Malbran-Labat 1986): durative (Labat 1951:36); intensive or emphatic, iterative and durative (Hallock 1959:18); indicating will, intent, decision, or declaration (Grillot and Vallat 1975, Grillot-Susini and Roche 1987:36); uncertain, indicating change of state (Khačikjan 1998:36).

When Achaemenid Elamite reflects translation of an underlying Old Persian text or simply contact with Old Iranian speakers, historically original distinctions are affected by calquing on Old Iranian. Old Persian subjunctives with future meaning are regularly translated with Conjugation III forms, and Old Persian presents usually with Conjugation IIIm forms (McAlpin 1981:71; Tucker 1998:181–182).

### 4.5.2 Verb moods

Several modal uses of various conjugation forms can be identified.

#### 4.5.2.1 Precative or optative

Forms of Conjugations I and II with the suffix *-ni* ~ *-na* are precative or optative: Middle Elamite [*tela-k-ni*] “may it be dedicated(?)”; Neo-Elamite [*hutta-hš-ni*] “may they do”; Achaemenid Elamite [*kata-k-t(i)-ni*] “may you live”; [*dunu-š-ni*] “may he give.” Achaemenid

Elamite forms in *-ni* sometimes correspond to Old Persian optatives: thus, [sura-k nima-k-ni], and [šura-k-ni], both rendering Old Persian *miḍa kariyaiš* “would do harm.” The particle *-ni* may also be asseverative (Middle Elamite [hutta-h-ni] “I indeed made,” [šatu-h-ni] “I will truly š.”; see Grilhot 1978:29 n. 65) and perhaps concessive (Middle Elamite [kuši-k-ni] “although(?) it was built [formerly of unbaked brick, I rebuilt it of baked brick]”).

#### 4.5.2.2 Imperative

In Middle Elamite, the second person of Conjugation I serves as the imperative (*kullak-ume hap-t(i)* “hear my prayer”). In Achaemenid Elamite, the third person of Conjugation I serves as an imperative: [mite-š . . . halpi-š] “go, defeat.” In a parallel phrase the first of two imperatives, an intransitive, is rendered with the bare stem: [mite ~ mida . . . halpi-š]. See also Vallat 1994:266, arguing for *iddu* < \**in du* “he is to receive it,” a bare stem used as third-person imperative or optative.

#### 4.5.2.3 Prohibitive

Prohibitives are Conjugation III (imperfective, non-past) forms preceded by the particle *anu* ~ *ani*: for example, Middle Elamite [par ani kutu-n] “may he not be assured of(?) progeny”; Neo-Elamite [anu i-n kuti-n-k(i)] “I must surely not support(?) him”; Achaemenid Elamite [hupe anu hutta-n-t(i)] (written *huttamti*) “do not do that”; [anu u ir turna-n-p(i)] (written *turnampi*) “lest they know me,” corresponding to Old Persian *mā taya-* with a subjunctive.

### 4.5.3 Nonfinite verbals

The bare verbal stem used as a substantive is usually termed an “infinitive”: for example, Achaemenid Elamite GN-*mar* GN<sub>2</sub> *laki* “a crossing from GN to GN<sub>2</sub>,” occurring at the end of the text, in a statement otherwise construed with a finite form [pari-š] “they went.” The form is labeled a Conjugation I infinitive in Hallock 1965; a Participle I in Khačikjan 1998:41. Stems with animate personal markers are agent nouns: Achaemenid Elamite [lipte kuti-r-a] “bow carrier”; called Conjugation I participle in Hallock 1965. Stems with suffixed *-k* and *-n*, that is, the bases of Conjugations II and III, are passive-intransitive perfective (sometimes past) participles and active imperfective (non-past) participles, respectively. Participles in *-k* also form substantives or adjectives: [katu-k-r-a] “living”; [halpi-k-r-a] “dead”; [hutta-k hali-k] “(what is) made with effort(?)”. The stem with *-n* or *-na* is also a non-past or imperfective infinitive: for example, Middle Elamite *kukkunum pittena* “[the god commanded me] to make an enclosure of (?) the *k*”; Achaemenid [tuppi talli-ma-n-a] “[I ordered] an inscription to be written.” Such constructions are termed Conjugation III infinitive in Hallock 1965; verbal noun or supine in Khačikjan 1998:42. Compare, however, Achaemenid [šaparakumme hutta-ma-n-r-a] “[he came] to do battle,” with a Conjugation III third-person form translating an Old Persian infinitive.

### 4.5.4 Other verbal morphemes

Additional suffixes can be appended to verbal forms.

#### 4.5.4.1 The suffix *-a*

This suffix attaches to verbal forms of all conjugations in all periods. It is usually the final morpheme of the form (but note Achaemenid Elamite [kuši-š-t-a-p(e)] “women who have given birth” and similar forms; see §5.6). In Middle Elamite it also attaches to some nominal

forms, including nominalized clauses, either replacing or following markers of gender and person: [DN GN-r-a] “DN [the god] of GN”; [siyan . . . in-me (written *imme*) kuši-hš(i)-me-a (written *kušihšima*)] “the temple which they did not build.” Divergent characterizations of the function of *-a* include the following:

1. Suffixed *-a* is determinative and subordinating. It first marked determining attributes of nouns and nominal predicates of subordinate clauses, then also marked verbal predicates of subordinate clauses. In Achaemenid Elamite *-a* appears mostly on subordinate verbs. In all periods, clauses introduced with relative pronouns or conjunctions may also omit *-a* (Grillot 1970, 1973; Grillot-Susini and Roche 1987:25, 40; Steiner 1990:144, 153). In an extreme form of this interpretation, Achaemenid Elamite forms in *-a* are subordinate and only subordinate, usually with temporal implication, but also with causal and other nuances: [hutta-k-a] “which is done,” hence “which has [previously] been done” (see, among others, Giovinazzo 1989; Vallat 1994:272).
2. Alternatively, *-a* is connective. It does not express subordination but coordination: thus, Achaemenid Elamite [marri-š-a (written *maurišša*) appin halpi-š] “he seized and killed them”; [marri-k-a u-ikki tengik] “he was seized and brought to me” – both corresponding to Old Persian main clauses; Middle Elamite [pepši-h-a kuši-h] “I restored and built” (see Hallock 1959:5–6, 11–12, 1973:150–151; and cf. Steiner 1990:144, comparing Elamite relative *-a* to the use of the Akkadian enclitic conjunction *-ma* in paratactic syntax to express subordination).
3. With less precision, *-a* is a semantic auxiliary expressing “non-finiteness and semantic connection . . . primarily looking forward to the finite verb.” See McAlpin 1981:80 (cf. 71); in general, Khačikjan 1998:50–51.

It is probable that *-a* is determinative-relative through Middle Elamite and probably later. That *-a* is always subordinating and only subordinating in Achaemenid is less well grounded. Counterexamples for all proposals occur, notably many Achaemenid Elamite administrative texts in which all verbs are marked with *-a* (see also Tucker 1998:165, n. 2, noting Achaemenid Elamite leveling in the distribution of *-a*).

#### 4.5.4.2 The suffix *-ti* ~ *-ta*

Disagreement also prevails over the characterization of a suffix *-ti* (and *-t(i) + -a > -ta*) found on verbs of all periods. It appears mostly on third-person forms of Conjugation I (Middle Elamite [kuši-š-t-a], Achaemenid Elamite [hutta-š-t-a]), rarely on other forms (Achaemenid Elamite Conjugation II second person [huttu-k-t-a]). In Middle Elamite, forms with *-ti* ~ *-ta* often occur in subordinate clauses; in Achaemenid royal inscriptions, they occur only in subordinate clauses; in Achaemenid administrative texts they often occur at the ends of texts.

The suffix *-ti* ~ *-ta* is characterized by some as marking finality or completeness (Hallock 1959:6–7; McAlpin 1981:71); by others as marking past time, translatable with perfect or pluperfect tenses (Hinz and Koch 1987 *passim*), most often distant past time, anteriority with respect to another verb (Grillot-Susini and Roche 1987:33; Vallat 1994:272). Most passages can be plausibly translated with past tenses that indicate anteriority: Middle Elamite [akka kukši-š-t-a imma durna-h] “I do not know [the former kings] who had built it [the temple]”; Achaemenid Elamite [akka Makuš šari-š-t-a] “[I rebuilt the temples] which the Magian had destroyed.” Khačikjan 1998:53 suggests historical development in the function of *-ti* ~ *-ta* from a nominalizing clitic (after Labat 1951:38 and Paper 1955:49), made obsolete as the system of marking nouns for gender and person became less articulated, to a completive and/or pluperfect marker.

## 4.6 Adverbs

Some Achaemenid phrasal adverbs are formed with nouns and postpositions (see §5.1.1): *da'e* “other” > [daʔe-ikki] “differently”; *irša-* “big” > [irše-ikki] “much, many”; *šit-* “night,” *na-* “day” > [šit-ma-na na-ma-na] “by night, by day,” /nan-na/ “daily”; [hi-ma] “here,” [hupe-ma] “there,” [hupe-ma-mar] “from there.” Dimensional elements provide the heads of other derived adverbs: Neo-Elamite [ukku-mi-na] “above”; [pat-mi-na] “below”; Achaemenid Elamite [me-ni] ~ [me-mi] “then”; [me-ša, mešši-n, me-šamerašae] “afterward.” Others with a derivational suffix *-ta* have doublets without *-ta*: Achaemenid Elamite [ha-me-r ~ ha-me-r-ta] “then, after that”; [hupi-me-r ~ hupi-me-r-ta] “then, after that”; [am ~ am-ta] “now”; [šašša ~ šašša-ta] “formerly.” Others are derived from nouns with various formants (Middle Elamite [šut-ki-me šat-ki-me] “by night, by day”; Achaemenid Elamite [na-zirna, na-randa] “daily”), or from participles (Achaemenid Elamite [kappa-k-a] “together”; [zilla-k-a] or [šilla-k-a] “greatly, much”). Others are simply bare stems: Achaemenid Elamite *yani* “afterwards”; *zila* “thus,” but usually phrasal *hi zila* “thus.”

Achaemenid Elamite distributive constructions are formed with nouns or numerals, usually paired, marked with the postposition or derivational suffix *-na*: [10 ruhip-na ak 10 ruhip-na] “[1 sheep to be received] by each group of ten”; <sup>d</sup>ITI<sup>MEŠ</sup>-na <sup>d</sup>ITI<sup>MEŠ</sup>-na “[one unit of wine to be received] monthly,” compare [kurtaš hupipe-na unra-na <sup>d</sup>ITI<sup>MEŠ</sup>-na x duš-t-a] “136 of their workers received x [barley] each per month.” The suffix is usual but optional: *ruh-ra ruh-ra* <sup>d</sup>ITI<sup>MEŠ</sup>-na <sup>d</sup>ITI<sup>MEŠ</sup>-na “each man, per month”; *5-ip ak 5-ip... 5-ip ak 5-ip-na... 5-ip-na ak 5-ip-na... 5-ip-na* (all in a single text).

## 4.7 Interjections

A vocative interjection *e* appears in pre-Achaemenid Elamite: for example, *e DN* “o, DN!.” In Achaemenid inscriptions, Old Persian vocative cases have no corresponding formant in the Elamite version: *ruhirra*, corresponding to the Old Persian vocative *martiyā* “o, man,” though perhaps *malla e*, corresponding to the Old Persian vocative *marikā* “o, subject.”

## 4.8 Compounds

Compound nouns are of several constructions: (i) noun plus noun (*kik-murun* “sky-earth” > “world”); (ii) participle plus participle (*huttak-halik* “done-perfected” > “handiwork, accomplishment”); (iii) infinitive plus agent noun (*paha-hutip* “protect-doers” > “protective gods”); (iv) infinitive plus infinitive (*hutta-hali* “handiwork, accomplishment”). Compound verbs consist of a noun plus verb: *mur-ta-* “place-put” > “establish”; *kur-ma-* “hand-intend(?)” > “entrust” (see Grillot 1984:190 n. 25).

## 4.9 Numerals

Cardinal numerals may take nominal suffixes: [ki-r] “one,” 1-*ir*, 2-*ip*, 3-*ip*, and so forth; [bel ki-ma] “in one year”; [ki-r x duš] “one (man) received x (grain)”; [1-ir šalu-r] “one gentleman,” but *samidakurra ki* “one *samida*-maker” (all examples from Achaemenid Elamite).

In Achaemenid Elamite, ordinal numbers are usually followed by *-ummema* ~ *-ummena* ~ *-ummemana*, probably to be analyzed as including the nominal suffix *-me* and the postpositions *-ma* and *-na* (Hallock 1969:76). Less frequent variant forms are *-umme*, *-mema*, *-mena*, and *-memana*.

In Achaemenid Elamite, fractions are formed with a suffix *-irmaki* ~ *-kurmaki* (Cameron 1948:38f; Hallock 1969:73).

## 5. SYNTAX

### 5.1 Word order and typology

The subject of attention usually occurs in sentence-initial position. In Middle and Neo-Elamite, the verbal predicate is normally at the end, indirect objects precede direct objects, attributes and clauses follow the nouns they modify, resumptive pronouns and adverbs precede verbs, so the common sentence order is:

- (12) Subject (+ modifier) – Indirect object (+ modifier) – Direct object (+ modifier)  
– Resumptive pronoun(s) – Adverb – Verb

As partially illustrated in the following example: [sian DN-me sunki-p uri-p-u-p(e) GN in-me kuši-hš(i)-me-a u GN kuši-h] “the temple of DN which kings who were before me did not build in Susa I built (at) the acropolis.”

In Achaemenid Elamite the verb is often but not always final. Free and irregular word order does not always reflect translation from Old Persian: thus, [meni sunki-me hupi-r(ri) GN-(i)p-na hutta-š] “then the kingship he of the Elamites exercised” corresponds to Old Persian *haw xšāya θiya abava Ūjai* “he became king in Elam”; [ap dunu-k-a <sup>SAL</sup>MUNUS<sup>MEŠ</sup> appa GN hami-ma-n-p(i) gal-ma] “[grain] to them was given, women who in GN were grinding(?), as rations.”

Khačikjan 1993, 1998:63–66 reviews the discussion of ergativity in Elamite (Kammenhuber 1974:204; Steiner 1979, 1990: 151, 159; Wilhelm 1978, 1982; Diakonoff 1981), concluding that Elamite was “an early nominative language [i.e., based on a fundamental opposition of subject vs. object] that had retained some features typical of ergative [i.e., based on a fundamental opposition of agent vs. patient] languages.”

#### 5.1.1 Postpositions

Elamite is chiefly postpositional, though prepositions occur as well. In Achaemenid Elamite, spatial and temporal relationships are expressed with postpositions, either enclitic (*-ma* “in, on”; *-ikki* “to”; *-mar* ~ *-ikki-mar* “from”; *-lakka* “across”) or separable ([hi da-k-a] > *idaka* “with”; [hat-i-ma] > *hatima* ~ *hatuma* “within, throughout”; *tubaka* “concerning”; *tibba* “before(?)”). A preposition *kuš* “to(ward), until” occurs both in Middle Elamite and in Achaemenid Elamite: [kuš Purattu ir pari-h] “I went toward the Euphrates.” (see §5.5.)

In pre-Achaemenid Elamite, postpositions per se are less numerous and less frequent. Locative *-ma* “in” and possessive *-na* “of” are common in Middle Elamite. Other postpositions are occasional in Neo-Elamite: *-ikki* “to” and perhaps *-tibba* “before” (perhaps adverbial; see Grillett-Susini and Roche 1987:29). Most spatial and temporal relationships in pre-Achaemenid Elamite are indicated by “directional words” combined with pronouns in postpositional constructions. The directional words originate either as nouns (*ukku* “head” > “on”; *pat* “foot, base” > “under”; *si* “face(?)” > “before”; *me* “(?)” > “after”), or as verbs (*li*- “give” > *lina* “for”; *tuk*- “(?)” > *tikka*- “for the sake of”). Two types of postpositional constructions occur, subject to different interpretations.

One analysis distinguishes postpositional constructions as governing internally and governing externally, or as long and short constructions. The long construction, governing internally, consists of (i) the *governing* noun or an anaphoric pronoun referring to the governing noun; (ii) the directional element with a nominal suffix (see §4.2.1) referring to the governing noun; (iii) a pronoun referring to the *governed* noun plus a nominal suffix again referring to the *governing* noun:

- (13) A. [i-r                      pat-r                      u-r                      ta-t-ni]  
           him-ANIM. SG.    under-ANIM. SG.    me-ANIM. SG.    place-2ND PER.-OPT.  
           “May you place him under me”  
       B. [RN    ukku-r                      i-r                      murta-n]  
           RN    over-ANIM. SG.    it-ANIM. SG.    put in place-IMPERF.  
           “Establishing RN over it”

The short construction, governing externally, consists of (i) the governed noun, (ii) an anaphoric pronoun referring to the *governing* noun with a nominal suffix marking concord with the *governed* noun, and (iii) the dimensional element with a nominal suffix again referring to the *governing* noun, and usually with determinative or subordinating *-a* (see Grillot 1983; Grillot-Susini and Roche 1987:27–28):

- (14) [DN    i-r                      šara-r-a                      ani                      uzzu-n]  
           DN    he-ANIM. SG.    beneath-ANIM. SG.-SUBORD.    NEG. WISH    go about-IMPERF.  
           “May he not go about(?) beneath the Sun God”

Another analysis distinguishes constructions in which the governed word is a *substantive* from constructions in which the governed word is a *pronoun*. In the first (corresponding to the short, external construction), (i) the relationship between the governing element and the governed substantive is unmarked, and (ii) the governed noun (*napi-r*) is followed by a resumptive pronoun referring to the governing element (*i* = *zalmu*) and by (iii) a dimensional element with nominal suffix referring to the governing element:

- (15) [zalmu... DN    napi-r                      u-r(i)                      i    sima-Ø                      ta-h]  
           statue            DN    god-ANIM. SG.    me-ANIM. SG.    it    before-INAN.    place-1ST PER.  
           “The statue, I placed it before my god, DN”

In the second (corresponding to the long, internal construction), (i) the governing noun or an anaphoric pronoun referring to it is followed by (ii) the directional element with a nominal suffix referring to the governing noun and (iii) a personal pronoun indicating the governed noun (Khačikjan 1998:45–47):

- (16) [peti-p                      pat-p                      u                      p-rabba-k-na]  
           be hostile-ANIM. PL.    under-ANIM. PL.    me    ANIM. PL.-bind-PERFV.-OPT.  
           “May enemies be bound beneath me”



These same constructions sometimes appear in Achaemenid Elamite:

- (17) A. [sunki-r                  murun    hi    ukku-r(i)]  
           king-ANIM. SG.   earth    this    on-ANIM. SG.  
           “King on this earth”  
       B. [PN...    me-r(i)                  i-r                  ta-k-a                  sa-k]  
           PN            after-ANIM. SG.    he-ANIM. SG.    put-PERF.-REL./CONNEC.    go-PERF.  
           “He got under way(?) after PN”

Note the phrasal adverb [i-n tukki-me], a long construction corresponding to the Middle Elamite short construction [i-n-tikka], both “for the sake of it, therefore.”

## 5.2 Agreement

A distinctive feature of Elamite syntax is “bracketing” (Bork 1933–1934), in which nominal suffixes that identify gender and person mark the constituents of possessive and attributive constructions and subordinate clauses (see §4.2.1).

In possessive and attributive constructions, the suffixes appropriate to the possessor or the determined substantive are added to the possessed or attribute; consider the following Middle Elamite examples:

- (18) A. [u PN šak PN<sub>2</sub>-k(i) sunki-k GN-GN<sub>2</sub>-k-a]  
           “I, PN, son of PN<sub>2</sub>, king of GN (and) GN<sub>2</sub>”  
           with first-person suffixes throughout  
       B. [PN meni-r GN-r ak GN<sub>2</sub>-r(i) šak-Ø hanik-Ø PN<sub>2</sub>-r(i) ak PN<sub>3</sub>-r(i)]  
           “He, PN, ruler(?) of GN and GN<sub>2</sub> beloved son of PN<sub>2</sub> and PN<sub>3</sub>”  
           with third-person suffixes throughout

The last noun in a sequence is always marked, but not all elements in the series are necessarily marked (in [18B] [hani-k], Conjugation II participle, not \*[hani-k-r]). The suffix on the final element is sometimes doubled, without apparent change of meaning: [u PN šak PN<sub>2</sub>-**ki-k** liba-k hanik-Ø DN-**ki-k**]; see Grilhot 1978:6, suggesting that the final *-k* marks the end of the clause, Grilhot-Susini and Roche 1987:24, suggesting that the first suffix marks agreement and the second marks determinacy.

A single noun may govern more than one possessor: thus, [puhu kuši-k u-p(e) ak PN-p(e)] “children born of me and PN.”

### 5.2.1 Other possessive and attributive constructions

For kinship terms in possessive and attributive constructions see §4.2.3. In Neo-Elamite and Achaemenid Elamite, kinship expressions sometimes invert the word order that is usual in Middle Elamite inscriptions: [<sup>f</sup>PN PN<sub>2</sub> riti-r(i)], “<sup>f</sup>PN, PN<sub>2</sub>’s wife”; [PN PN<sub>2</sub> šak-r(i)], “PN, PN<sub>2</sub>’s son.” Since the inverted construction is already occasional in Middle Elamite ([lika-me riša-r(i)] “enlarger of the realm”), its later use is probably not a calque on Old Persian. The construction may reflect the syncopation of a resumptive pronoun: [šak (i)-r], [riti (i)-r] (Hallock 1962:54, Grilhot-Susini and Roche 1987:23).

In Neo-Elamite, descent is also expressed PN šak PN<sub>2</sub>-*na*. The postposition *-na* (probably to be analyzed as the neutral inanimate *-n* + *-a*), sometimes expresses possession or other qualification in Middle Elamite: *erentum-na* ~ *erentum-ma* ~ *erentum-ia* “[made] of baked brick.” In Achaemenid Elamite most possession and some attributive relationships



are expressed by the (so-called) genitive *-na*: Neo-Elamite *zalmu PN-na* “image of PN”; Achaemenid Elamite *halmi PN-na* “seal[ed document] of PN.” Occasional inversion of the word order in Achaemenid Elamite is probably a calque on Old Iranian: *PN-na miyatukka* “viaticum of (= issued by) PN”; *[hupirri-na gal-ma]* “as his rations.”

### 5.3 Resumptive pronoun-verb constructions

Verbs of Conjugation I are often preceded by one, two, or three resumptive pronouns that refer to the arguments of the verb. In Middle Elamite, pronouns that refer to logical indirect object, subject, and/or direct object of the clause regularly appear in that order; they may be contracted in writing, and some or all pronouns may be omitted: *[ap u in (written a-pu-un) duni-h]* “to them I gave it,” with variant *[ap u (written a-pu ú) duni-h]* “to them I gave” (see Grillett 1978:31; Grillett-Susini and Roche 1987:18, 39). In Achaemenid Elamite, pairs or groups of resumptive pronouns do not occur before Conjugation I verbs. Single resumptive pronouns refer to subjects or objects: *PN... sunkime hupirri marriš* “PN, he seized the kingship”; *u PN... ir halpi* “I, PN, him I killed”; *u DN un nuškišni* “I—may DN protect me.”

Verb forms of Conjugation II and Conjugation III are often but not always preceded by resumptive pronouns: *(V)n* for the first and second persons, *(V)r* and *(V)p* for third-person animates. The same pronominal forms that mark the objects of transitive Conjugation I verbs thus mark the agents of Conjugation II and III forms (in a typically ergative fashion): Neo-Elamite *[anu i n (written in) kuti-n-k(i)]* “I will truly not support(?) him”; Middle Elamite *[nu u n (written un) tahha-n-t-a]* “[O DN] you command(?) me”; *[u r (written ur) tahha-n-r-a]* “he [DN] commands(?) me”; Neo-Elamite *[u ip tahha-n-p-a]* “they [DN and DN<sub>2</sub>] command(?) me”; Achaemenid Elamite *[GN-ikki ir pari-k]* “he arrived at GN,” but *[anu u ir (not \*ip) turna-n-p(i)]* “lest they(!) know me”; *[hi zila ap (i)r titu-k-a]* “thus he lied to them,” but *[hi zila titu-k-a]* “thus he lied” (see Khačikjan 1998:35 and 65, Grillett-Susini and Roche 1987:35; cf. Malbran-Labat 1990 and Grillett 1978:19, 25. Grillett (1978:20–21), however, demurs, taking *(V)r-* and *(V)p-* as vestigial elements referring to the agent, but *(V)n* as marking the logical object).

In Achaemenid Elamite, indirect objects of verbs of all conjugations are regularly expressed with resumptive pronouns (Haddock 1969:9).

### 5.4 Coordination

The conjunction *ak* (usually spelled *a-ak*, sometimes *a-gi*, *ia-ak*), meaning both “and” and “or,” connects (i) words or (ii) clauses. Consider the following Middle Elamite examples:

- (19) A. *[siyan DN ak DN<sub>2</sub>-me]*  
       “Temple of DN and DN<sub>2</sub>”  
       B. *[sunki-r peti-r ak tari-r akka melka-n-r-a... ak lansit-e du-n-r-a ak hiš RN sukuš a-ak (written su-ku-ša-ak) i-m-e-ni aha-r ta-n-r-a]*  
       “A king, enemy or ally, who destroys [the temple] or takes its gold or erases the name of RN and puts his own there”

In Achaemenid Elamite inscriptions, it also introduces a new paragraph: *[ak RN sunki-r na-n-r]* “and RN the king declares,” where the Old Persian and Akkadian versions have no conjunction.

Another conjunction, *kudda* “and” occurs in Achaemenid Elamite, sometimes coupled with *ak*: *kudda Paršip ak kudda Madabe ak kudda dayauš appa dae* “Persia and Media and the other countries.” A possible Neo-Elamite occurrence of *kudda* raises doubt about the

suggestion that *kudda* is not an Elamite word but a graphic device meant to indicate that “and,” written in Elamite as *ak*, was to be read out in Old Persian as *utā* (Gershevitch 1979:132; Zadok 1995:243). In one inscription Elamite *utta* transcribes Old Persian *utā* “and.”

Coordinate clauses are thus connected with *ak*; in Achaemenid Elamite also with *kudda*; or asyndetically. In pairs of closely coordinated verbs, suffixes on the second verb may apply to both verbs (without a conjunction): thus, Middle Elamite [e DN hutta-t u-n duni-t-ni] “O, DN, may you do [and] give me”; and so with participles, [hutta-k hali-k-u-me] “what I made and finished(?)”.

## 5.5 Subordination

Achaemenid Elamite uses subordinating conjunctions, including (i) simple conjunctions (*anka* “if, when”; *kuš* “until” (also prepositional “to(ward)”; *sap* “as, when”); (ii) phrasal conjunctions (*sap innu* “as long as”; *meni sap anka* “after”); and (iii) phrasal conjunctions with the relative *appa*, perhaps calques on Old Persian conjunctions compounded with relative *taya* (*appa anka* “as”; *sap appa* “when”). In pre-Achaemenid Elamite, *anka* appears once at the head of a clause ([*anka ruri-n-a ak miši-ma-n-a*] “if [the temple] . . . -s and becomes dilapidated”), and *kuš* appears only as a preposition.

Most subordinate clauses precede the verb of the main clause. In Achaemenid Elamite, purpose clauses governed by *šera*, “order,” are formed with infinitives of Conjugation III with auxiliary *-ma-* and follow the governing verb: *meni u šera DUB<sup>MEŠ</sup> tallimana* “then I ordered an inscription to be written” (Grillot 1973:155–162; Grillot and Vallat 1975:215; Grillot-Susini and Roche 1987:41).

Since Achaemenid Elamite verb forms marked with *-ta* are often final, a correlate of the view that *-ta* marks anteriority is the supposition that temporal clauses referring to anterior action often follow the clauses that refer to prior action: [*du-š-a . . . hutta-š-t-a*] “(barley) which he received, because he had previously done . . .” (see Vallat 1994:272–273).

## 5.6 Relative clauses

Elamite relative clauses may be introduced by the relative pronouns *akka* “who” or *appa* “which.” The clause may follow its antecedent (e.g., Middle Elamite [*sunki-r akka ta-š-t-a*] [lit. “king-ANIM. SG. who he-has put”] “the king who set up [the stele]”), or the relative clause may occur without an expressed antecedent (e.g., Middle Elamite [*akka ulhi i melka-n-r-a*] [lit. “who house this he-destroys”] “he who destroys this house”).

There is another way, predominant in Middle Elamite, in which Elamite forms relative clauses. Attributive relative clauses may also be marked like other attributes, by adding nominal suffixes to the verb at the end of the clause. In Middle Elamite *-a* is often attached to the nominal suffix ([*sian . . . in-me kuši-hš(i)-me-a*] “the temple that they had not built”), but the presence of *-a* is optional ([*lika-me i-r hani-š-r(i)*] “whose realm DN loves,” Grillot 1978:11). In Neo- and Achaemenid Elamite examples, *-a* attaches to the verb form before the nominal suffix: Neo-Elamite [6-(i)p ANŠE.KUR.RA<sup>MEŠ</sup> tukka-š-t-a-p(e)] “six people who fed(?) horses”; Achaemenid Elamite [6 MUNUS<sup>MEŠ</sup>-na kuši-š-t-a-p(e)] “six [women] who gave birth to girls”). In Middle Elamite, a relative pronoun can optionally (pleonastically) occur at the head of such clauses ([*sian appa* (variant omits *appa*) kuši-h-me-a] “the temple that I built”); no Neo-Elamite or Achaemenid Elamite examples combine this construction with a relative pronoun. See Grillot 1978:8–15; Grillot-Susini and Roche 1987:24, 41; Khačikjan 1998:59–60; Hallock 1969:37, 1978:115, 1973:149 (the last-mentioned demurs on Middle Elamite examples).

The occurrence of these two types of relative constructions varies over time. In Achaemenid Elamite, use of the relative pronoun is regular, but it is uncommon in Middle Elamite. Conversely, the nominal construction of relative clauses (with *-a*) is common in Middle Elamite but rare in Achaemenid (and Neo-) Elamite.

## 5.7 Direct discourse

The close of a quoted statement is indicated by a form of *ma-*, probably identical with the verbal auxiliary *ma-*, with suffixes in agreement with the speaker: *manka* (Conjugation III first-person singular), *mara* ~ *mar* and *mapa* (animate singular and plural agent nouns) and *maka* (passive participle, impersonal).

In Achaemenid Elamite, a verb that indicates speaking (*turu-* ~ *tiri-* “tell, speak”; *na-* “say”) usually introduces the quoted statement:

- (20) A. [hi zila ap tiriya mite-š . . . halpi-š ma-n-k-a]  
 “I told them thus, ‘Go, defeat (the enemy)’”  
 B. [na-n-ri PN šera-š ma-r-a]  
 “He said ‘PN gave the order’”

In Neo-Elamite, verbs of speaking sometimes follow the quoted statement plus *ma-*:

- (21) A. [ir unsa-h-a mara tiri-n-r-a]  
 “[PN] who says ‘I made an exchange(?) with him’”  
 B. [akka zalmu . . . in-k(i) in-dunu-n-k(u) mar turu-n-r-a]  
 “He who says ‘I will not give the statue’”

The verb of speaking in such constructions may, however, be omitted.

Neo-Elamite and Achaemenid Elamite letters begin with verbs of speaking, introducing the body of the letter as quoted matter to be spoken to the addressee:

- (22) [PN turu-š PN<sub>2</sub> na-n turu-š]  
 “Tell PN [the addressee], PN<sub>2</sub> [the sender] speaks, saying”

*Mara* and *mapa* are not added at the end of the letter, where the end of direct speech is self-identifying (but cf. *u nun turriya nanki . . . hupirri mušin huttanra manka* “I spoke to you, saying ‘ . . . He will do the accounting,’” apparently quoting from a previous letter).

## 6. LEXICON

Without a body of bilingual texts, an indigenous scholarly tradition, or a well-known language that is closely related to Elamite, few pre-Achaemenid Elamite words can be translated with precision and many can be translated only with guesses. The geographical and chronological distribution of the lexicon has not yet been analyzed. A comprehensive collection of parsed forms, useful for problems in Elamite grammar, has not been made (Zadok 1995:243).

Elamite words in Akkadian texts from southwestern Iran, where Elamite was also spoken and where it was often the language of the rulers, include titles of officials, names of professions, and words for realia (in legal and administrative texts), and architectural terms and titles or kinship terms (in dedicatory inscriptions). Elamite words in Akkadian texts from Mesopotamia include titles and terms describing people in or from Elam and a small number of common nouns that may be actual loanwords. A few other Elamite nouns are identified and glossed in Mesopotamian lexical texts (Zadok 1995:244–245; Vallat 1998; Stolper 1978). Elamite words appear in personal names, often of people identified as Elamites, in Sumerian

and Akkadian texts of many periods (Zadok 1984, 1991). *Kam/bnaskires*, the name or sobriquet of rulers in Parthian Elymais, is probably the survival of an Elamite title, *kapnuškir*, “treasurer” (Alram 1986:139–153, Stolper 2000:287).

Akkadian words in pre-Achaemenid Elamite building inscriptions are mostly proper nouns, including names of places and buildings (*alumelu* ~ *alimeli* “acropolis,” *abul mišari* “gate of justice”), epithets of gods and rulers (*melki ilani* “king(s) of the gods”), and names of votive objects (*nur kibрати* “light of the world”). Possible Sumerian or Akkadian words for materials or objects in administrative texts may be *Kulturwörter* or Akkadograms (written as Akkadian but read as Elamite): Middle or Neo-Elamite *zabar* “copper or bronze,” *anaku* “tin,” *kušuru* “beam,” Achaemenid Elamite *paru* “mule,” *basbas* “duck” (Stolper 1984:21–22).

Achaemenid Elamite inscriptions contain transcriptions of Old Iranian words, not always representing forms identical to those used in the corresponding Old Persian texts: for example, Elamite *miššadanašpena*, transcribing Old Persian \**visadanānām*, where the Old Persian text has the non-Persian form *vispazanānām*, “of all kinds” (genitive plural). Transcribed Iranian words include terms with specific cultural nuance (*irdama* corresponding to Old Persian *artāvā* “blessed [in death]”), and occasional common words and particles (enclitic *-aham*, *-me* corresponding to Old Persian *āham* “I was,” *-mai* “my”). Achaemenid Elamite administrative texts include transcriptions of hundreds of Iranian words, many unattested in Old Iranian (e.g., *miyatukka* < Iranian \**viyātika* “authorization, viaticum”), some also found as loanwords in Achaemenid texts in other languages (*kanzabarra* < Iranian \**ganzabara* “treasurer,” Akkadian *ganzabaru*, Aramaic *gzbr* and *gnzbr*, etc.; see Hinz 1975b). The Elamite transcriptions represent both Persian and non-Persian Iranian forms (*misapušša*, *miššaputra* corresponding to Persian \**viṭapuša-*, non-Persian \**visapuθra-*, “prince”). For those who hold that Achaemenid Elamite texts are not translations, but Elamographic transcriptions of texts that are dictated in Iranian and read out in Iranian, these forms are not foreign words or loanwords but explicit writings of the underlying text (Gershevitch 1979).

## 7. READING LIST

Hinz and Koch 1987:133–168 offers comprehensive bibliography of works on Elamite texts, language, and history published between 1711 and 1986, arranged chronologically. Later items are listed in the journals *Abstracta Iranica*, *Archiv für Orientforschung*, and *Orientalia*.

Potts 1999 surveys current knowledge of the archeology and history of Elam from prehistory to the Islamic conquest. A short current survey of Elamite history is Vallat 1997a (but many contemporary historians of the ancient Near East will hesitate over the geographical framework). An encyclopedic survey of Susa in Elamite and Iranian history is Steve *et al.* 2002. Longer surveys, including Cameron 1936, Hinz 1965, 1972–1973, and Carter and Stolper 1984, are out of date.

A current survey of Elamite grammar is Khačikjan 1998. Among earlier surveys, Labat 1951, Reiner 1969, and Grillo-Susini and Roche 1987 represent successive generations of a school that gives greatest prominence to Middle Elamite evidence; Paper 1955 and Hallock 1959, 1965, 1969:8–10 are explicitly confined to Achaemenid Elamite; McAlpin 1981:63–83 includes separate parallel treatments of Achaemenid and Middle Elamite as a basis for systematic comparison with Proto-Dravidian.

The lexicon Hinz and Koch 1987 covers texts of all periods, including proper names and Elamite words in non-Elamite texts, and includes a survey of published treatments of many entries. Hallock 1969:664–776 is a nearly complete glossary of Achaemenid Elamite. Elamite personal names are collected in Zadok 1984, Elamite place names in Vallat 1993.

Useful transcriptions and editions of most pre-Achaemenid Elamite royal inscriptions are in König 1965. Other collections of Elamite inscriptions are Steve 1967 (Middle Elamite texts from Chogha Zanbil), 1987 (pre-Achaemenid and Achaemenid inscriptions from Susa), and Malbran-Labat 1995 (pre-Achaemenid building inscriptions from Susa).

The synoptic edition of Achaemenid Elamite multilingual texts of Weissbach 1911 is dated but not replaced. A recent edition of the longest Achaemenid Elamite royal inscription, at the monument of Darius I at Bisitun, is Grillot-Susini, Herrenschmidt, and Malbran-Labat 1993. A compendium of the Elamite versions of the Achaemenid inscriptions is to appear in the *Corpus Inscriptionum Iranicarum*. Editions of Achaemenid Elamite administrative texts from Persepolis are Cameron 1948 and Hallock 1969, 1978. All translations of Elamite texts merit reading with some reservation.

## Bibliography

- Afram, M. 1986. *Nomina Propria Iranica in Nummis*. Iranisches Personennamenbuch, IV. Vienna: Österreichische Akademie der Wissenschaften.
- André, B. and M. Salvini. 1989. "Réflexions sur Puzur-Šušinak." *Iranica Antiqua* 24:53–72.
- Baššāš-e Kanzaq, R. 1997. "Šarḥ-e čahār mōred az katibehā-ye manqūr bar ašiyā-ye ġār-e Kalmākare." In *Yādnāme-ye gerdehamā'ēye bāstān-šenāsī-Šūš*, vol. I, pp. 13–22. Tehran: Sāzmān-e Mirāš-e Farhangī-ye Kešvar (National Organization of Cultural Heritage).
- Bleibtreu, E. 1999. "Ein vergoldeter Silberbecher Assurbanipals (668–627)." In W. Seipel (ed.), *Schätze des Orients. Meisterwerke aus dem Miho Museum. Kunsthistorisches Museum Wien, 22 Juni–31 Okt. 1999*, pp. 21–30, 54–56. Vienna(?): SKIRA.
- Bork, F. 1925. "Elam, B. Sprache." In M. Ebert (ed.), *Reallexikon der Vorgeschichte*, vol. III, pp. 70–83. Berlin: Walter de Gruyter.
- . 1933–1934. "Die elamische Klammer." *Archiv für Orientforschung* 9:292–300.
- Brice, W. 1962–1963. "The writing system of the Proto-Elamite account tablets of Susa." *Bulletin of the John Rylands Library* 45:15–39.
- Cameron, G. 1936. *History of Early Iran*. Chicago: University of Chicago Press.
- . 1948. *Persepolis Treasury Tablets*. Oriental Institute Publications 65. Chicago: University of Chicago Press.
- Carter, E. and M. Stolper. 1984. *Elam: Surveys of Political History and Archaeology*. Near Eastern Studies 25. Berkeley/Los Angeles: University of California Press.
- Caubet, A. 1995. "Acquisitions: Moyen orient ancien, Musée du Louvre, Département des Antiquités Orientales, 1: Gobelet." *Revue du Louvre* 4:81.
- Civil, M. 1985. "Sur les 'livres d'écolier' à l'époque paléo-babylonienne." In J.-M. Durand and J.-R. Kupper (eds.), *Miscellanea babylonica, mélanges offerts à Maurice Birot*, pp. 67–78. Paris: Editions Recherche sur les Civilisations.
- Damerow, P. and R. Englund. 1989. *The Proto-Elamite Texts from Tepe Yahya*. American School of Prehistoric Research, Bulletin 39. Cambridge, MA: Peabody Museum of Archaeology and Ethnology.
- De Maaijer, R. 1996. "Recently acquired cuneiform texts." *Oudheidkundige Mededelingen uit het Rijksmuseum van Oudheden te Leiden* 76:69–84.
- Diakonoff, I. 1973. "Bemerkungen zu einer neuen Darstellung altkleinasiatischer Sprachen." *Orientalistische Literaturzeitung* 68:14–16.
- . 1981. "Ist das Elamische eine Ergativsprache?" *Archäologische Mitteilungen aus Iran* NS 14:7–8.
- Dijk, J. van, A. Goetze, and M. Hussey. 1985. *Early Mesopotamian Incantations and Rituals*. Yale Oriental Series, Babylonian Texts 11. New Haven/London: Yale University Press.
- Donbaz, V. 1996. "A median (?) votive inscription on silver vessel." *Nouvelles Assyriologiques Brèves et Utilitaires* 1996, no. 43.
- Englund, R. 1996. "The Proto-Elamite script." In P. Daniels and W. Bright (eds.), *The World's Writing Systems*, pp. 160–164. Oxford and New York: Oxford University Press.
- . 1998. "Elam, iii. Proto-Elamite." In *Encyclopaedia Iranica*, vol. VIII, fascicle 3, pp. 325–330. Costa Mesa, CA: Mazda.

- Friberg, J. 1978–1979. *The Third Millennium Roots of Babylonian Mathematics, I. A Method for the Decipherment, through Mathematical and Metrological Analysis, of Proto-Sumerian and Proto-Elamite Semi-Pictographic Inscriptions*. Göteborg, Sweden: Chalmers University of Technology and University of Göteborg.
- Gadd, C., L. Legrain, S. Smith, et al. 1928. *Royal Inscriptions*. Ur Excavations, Texts, 1. Publications of the Joint Expedition of the British Museum and of the Museum of the University of Pennsylvania to Mesopotamia. London: British Museum.
- Garrison, M. 1996. “A Persepolis Fortification seal on the tablet MDP 11 308 (Louvre Sb 13078).” *Journal of Near Eastern Studies* 55:15–35.
- Garrison, M. and M. Root. 2001. *Seals on the Persepolis Fortification Tablets, I: Images of Heroic Encounter*. Oriental Institute Publications 117. Chicago: Oriental Institute.
- Gershevitch, I. 1979. “The alloglottography of Old Persian.” *Transactions of the Philological Society of London*, pp. 114–190.
- Giovinazzo, G. 1989. “L’expression ‘*ha duš ha duka*’ dans les textes de Persépolis.” *Akkadica* 63:12–26.
- Grillot, F. 1970. “A propos de la notion de subordination dans la syntaxe élamite.” *Journal Asiatique* 258:213–241.
- . 1973. “La postposition génitive *-na* en élamite.” *Cahiers de la Délégation Archéologique Française en Iran* 3:113–169.
- . 1974. “A propos du verbe moyen élamite.” *Journal Asiatique* 262:31–35.
- . 1978. “Les affixes nominaux et les pronoms indépendants de la langue élamite.” *Journal Asiatique* 266:1–35.
- . 1983. “Le mécanisme des groupes nominaux en élamite.” *Journal Asiatique* 271:207–218.
- . 1984. “Trinôme de la royauté en Élam.” *Studia Iranica* 13:185–191.
- Grillot, F. and F. Vallat. 1975. “Le semi-auxiliaire *ma-* en élamite.” *Journal Asiatique* 263:211–217.
- Grillot-Susini (= Grillot), F. 1994. “Une nouvelle approche de la morphologie élamite: racines, bases et familles de mots.” *Journal Asiatique* 282:1–18.
- . 1998. “Elam, v. Elamite language.” In *Encyclopaedia Iranica*, vol. VIII, fascicle 3, pp. 332–335. Costa Mesa, CA: Mazda.
- Grillot-Susini, F. and C. Roche. 1987. *Éléments de grammaire élamite*. Paris: Editions Recherche sur les Civilisations.
- Grillot-Susini, F., C. Herrenschimidt, and F. Malbran-Labat. 1993. “La version élamite de la trilingue de Behistun: une nouvelle lecture.” *Journal Asiatique* 281:19–59.
- Hallock, R. 1958. “Notes on Achaemenid Elamite.” *Journal of Near Eastern Studies* 17:256–262.
- . 1959. “The finite verb in Achaemenid Elamite.” *Journal of Near Eastern Studies* 18:1–19.
- . 1962. “The pronominal suffixes in Achaemenid Elamite.” *Journal of Near Eastern Studies* 21:53–56.
- . 1965. “The verbal nouns in Achaemenid Elamite.” In *Studies in Honor of Benno Landsberger on his Seventy-Fifth Birthday, April 21, 1965*, pp. 121–125. Assyriological Studies 16. Chicago: University of Chicago Press.
- . 1969. *Persepolis Fortification Tablets*. Oriental Institute Publications 92. Chicago: University of Chicago Press.
- . 1973. “On the Middle Elamite verb.” *Journal of Near Eastern Studies* 32:148–151.
- . 1978. “Selected fortification texts.” *Cahiers de la Délégation Archéologique Française en Iran* 8:109–136.
- Helms, S. 1982. “Excavations at ‘The city and the famous fortress of Kandahar, the foremost place in all of Asia’.” *Afghan Studies* 3–4:1–24.
- . 1997. *Excavations at Old Kandahar in Afghanistan 1976–1978*. Society for South Asian Studies Monograph 2. BAR International Series 686. Oxford: Archaeopress.
- Henkelman, W. Forthcoming. “Persians, Medes and Elamites, Acculturation in the Neo-Elamite Period.” In G. Lanfranchi (ed.), *Continuity of Empire: Assyria, Media, Persia*.
- Hinz, W. 1965. *Das Reich Elam*. Stuttgart: Kohlhammer.
- . 1969. *Altiranische Funde und Forschungen*. Berlin: Walter de Gruyter.
- . 1972–1973. *The Lost World of Elam*. Translated by Jennifer Barnes. London: Sidgwick and Jackson, 1972; New York: New York University Press, 1973 (= translation of Hinz 1965).
- . 1975a. “Problems of Linear Elamite.” *Journal of the Royal Asiatic Society* 2:105–115.
- . 1975b. *Altiranisches Sprachgut der Nebenüberlieferungen*. Göttinger Orientforschung, III. Reihe (Iranica) 3. Wiesbaden: Otto Harrassowitz.



- Hinz, W. and H. Koch. 1987. *Elamisches Wörterbuch*. Archäologische Mitteilungen aus Iran, Ergänzungsband 17. Berlin: Dietrich Reimer.
- Justeson, J. and L. Stephens. 1994. "Variation and change in symbol systems: case studies in Elamite cuneiform." In C. Renfrew and E. Zubrow (eds.), *The Ancient Mind, Elements of Cognitive Archaeology*, pp. 167–175. Cambridge: Cambridge University Press.
- Kammenhuber, A. 1974. "Historisch-geographische Nachrichten aus der althurrischen Überlieferung, dem Altelamischen und den Inschriften der Königen von Akkad für die Zeit vor dem Einfall der Gutäer (ca. 2200/2136)." *Acta Antiqua Academiae Scientiarum Hungaricae* 22:157–248.
- Khačikjan, M. 1993. "To the typological characteristics of the Elamite language." In J. Zablocka and S. Zawadski (eds.), *Šulmu IV: Everyday Life in the Ancient Near East*, pp. 143–150. Adam Mickiewicz University, Seria Historia 182. Poznań.
- . 1995. "Notes on the Elamite phonology." *Studi micenei ed egeo-anatolici* 35:105–109.
- . 1998. *The Elamite Language*. Documenta Asiana 4. Rome: Istituto per gli studi micenei ed egeo-anatolici.
- König, F. 1965. *Die elamischen Königsinschriften*. Archiv für Orientforschung Beiheft 16. Graz.
- Labat, R. 1951. "Structure de la langue élamite (état présent de la question)." *Conférences de l'Institut de Linguistique de Paris* 9:23–42.
- Lackenbacher, S. 1998. "Elam, vii. Non-Elamite Texts in Elam." In *Encyclopaedia Iranica*, vol. VIII, fascicle 4, pp. 342–344. Costa Mesa, CA: Mazda.
- Malbran-Labat, F. 1986. "Le «semi-auxiliaire» élamite -ma. Une nouvelle approche: essai de définition d'un champ sémantique." *Bulletin de la Société de Linguistique de Paris* 81:253–273.
- . 1990. "Système pronominal et système verbal en élamite achéménide." *Bulletin de la Société de Linguistique de Paris* 85:61–90.
- . 1995. *Les inscriptions royales de Suse*. Paris: Editions de la Réunion des musées nationaux.
- Mayrhofer, M. 1973. *Onomastica Persepolitana. Das altiranische Namengut der Persepolis-Täfelchen*. Vienna: Österreichische Akademie der Wissenschaften.
- McAlpin, D. 1981. *Proto-Elamo-Dravidian: The Evidence and its Implications* (= *Transactions of the American Philosophical Society* 71, part 3).
- Meriggi, P. 1971–1974. *La scrittura proto-elamica* (I: *La scrittura e il contenuto dei testi*, 1971. II: *Catalogo dei segni*, 1974. III: *Testi*, 1974). Rome: Accademia Nazionale dei Lincei.
- Paper, H. 1955. *The Phonology and Morphology of Royal Achaemenid Elamite*. Ann Arbor, MI: University of Michigan Press.
- Potts, D. 1999. *The Archaeology of Elam: Formation and Transformation of an Ancient Iranian State*. Cambridge World Archaeology. Cambridge: Cambridge University Press.
- Reiner, E. 1960. "Calques sur le vieux-perse en élamite." *Bulletin de la Société de Linguistique de Paris* 55:222–227.
- . 1969. "The Elamite language." In B. Spuler (ed.), *Altkleinasiatische Sprache*, pp. 54–118. Handbuch der Orientalistik, Leiden/Köln: E. J. Brill.
- . 1992. "Elamite." In W. Bright (ed.), *International Encyclopedia of Linguistics*, pp. 406–409. Oxford: Oxford University Press.
- Salvini, M. 1998. "Elam, iv. Linear Elamite." In *Encyclopaedia Iranica*, vol. VIII, fascicle 3, pp. 330–332. Costa Mesa, CA: Mazda.
- Sayce, A. 1874. "The languages of the cuneiform inscriptions of Elam and Media." *Transactions of the Society of Biblical Archaeology* 3:465–485.
- Scheil, V. 1939. *Mélanges épigraphiques. Mémoires de la mission archéologique de Perse*, 28. Paris: Ernest Leroux.
- Steiner, G. 1979. "The intransitive-passival conception of the verb in the languages of the ancient Near East." In F. Plank (ed.), *Ergativity: Towards a Theory of Grammatical Relations*, pp. 185–216. London: Academic Press.
- . 1990. "Sumerisch und Elamisch: typologische parallelen." *Acta Sumerologica* 12:143–176.
- Steve, M.-J. 1967. *Textes élamites et accadiens de Tchoga Zanbil*. Tchoga Zanbil, III. Mémoires de la Délégation Archéologique en Iran 41. Paris: Paul Geuthner.
- . 1987. *Nouveaux mélanges épigraphiques, inscriptions royales de Suse et de la Susiane*. Ville Royale de Suse 7. Mémoires de la Délégation Archéologique en Iran 53. Nice: Editions Serre.
- . 1992. *Syllabaire élamite: histoire et paléographie*. Civilisations du Proche-Orient, Série II: Philologie, 1. Paris/Neuchâtel: Recherches et Publications.

- Steve, M.-J. F. Vallat, and H. Gasche. 2002. "Suse." *Supplément au Dictionnaire de la Bible*. Paris: Letouzey & Are. Fasc. 78: 359–511.
- Stolper, M. 1978. "Šarnuppu." *Zeitschrift für Assyriologie* 68: 261–269.
- . 1984. *Texts from Tall-i Malyan 1: Elamite Administrative Texts (1972–1974)*. Occasional Publications of the Babylonian Fund 6. Philadelphia: University Museum.
- . 2000. "Ganzabara." *Encyclopaedia Iranica* vol X, fasc. 3, pp. 286–288. Costa Mesa, CA: Mazda.
- Tavernier, J. 2002. "Iranica in de Achaemenidische Periode." Ph.D. dissertation, Katholieke Universiteit Leuven, Dept. Oosterse en Stavische Studies.
- Tucker, E. 1998. "The 'nominal conjugations' in Achaemenid Elamite." In M. Brosius and A. Kuhrt (eds.), *Studies in Persian History: Essays in Memory of David M. Lewis*, pp. 165–194. Achaemenid History 11. Leiden: Nederlands Instituut voor het Nabije Oosten.
- Vallat, F. 1986. "The most ancient scripts of Iran: the current situation." *World Archaeology* 17:335–347.
- . 1987a. "L'expression *da-ma da-ak* en élamite." *Nouvelles Assyriologiques Brèves et Utilitaires* 1987, no. 6.
- . 1987b. "Le pseudo pronom personnel \**kaš* en élamite achéménide." *Nouvelles Assyriologiques Brèves et Utilitaires* 1987, no. 114.
- . 1989. "Les compléments phonétiques ou graphiques en élamite achéménide." *Annali* (Istituto Universitario Orientale di Napoli) 49:219–222.
- . 1990. "Deux inscriptions royales en élamite de l'époque des Epartides (sukkalmah)." *Nouvelles Assyriologiques Brèves et Utilitaires* 1990, no. 137.
- . 1993. *Les noms géographiques des sources suso-élamites*. Répertoire Géographique des Textes Cunéiformes, 11. Beiheft zum Tübinger Atlas des Vorderen Orients, Reihe B, Nr. 7/11. Wiesbaden: Dr. Ludwig Reichert.
- . 1994. "Deux tablettes élamites de l'Université de Fribourg." *Journal of Near Eastern Studies* 53:263–274.
- . 1996a. "Nouvelle analyse des inscriptions néo-élamites." In H. Gasche and B. Hrouda (eds.), *Collectanea orientalia: histoire, arts de l'espace et industrie de la terre, études offertes en hommage à Agnès Spycket*, pp. 385–395.
- . 1996b. "Le royaume élamite de SAMATI." *Nouvelles Assyriologiques Brèves et Utilitaires* 1996, no. 31.
- . 1996c. "L'Élam à l'époque paléo-babylonienne et ses relations avec la Mésopotamie." In *Amurru, 1: Mari, Ebla et les Hourrites*, pp. 297–319. Paris: Editions Recherche sur les Civilisations.
- . 1997a. "Elam, I. The History of Elam." In *Encyclopaedia Iranica*, vol. VIII, fascicle 3, pp. 301–312. Costa Mesa, CA: Mazda.
- . 1997b. "La lettre élamite d'Arménie." *Zeitschrift für Assyriologie* 87:258–270.
- . 1998. "*hašša* et *kiparu*, deux termes élamites dans les textes accadiens de Suse." *Nouvelles Assyriologiques Brèves et Utilitaires* 1998, no. 127.
- Van Soldt, W. 1982. "The cuneiform texts in the Rijksmuseum van Oudheden, Leiden (I)." *Ouheidkundige Mededelingen uit het Rijksmuseum van Oudheden te Leiden* 63:43–59.
- . 1991. *Studies in the Akkadian of Ugarit: Dating and Grammar*. Alter Orient und Altes Testament 40. Kevelaer: Butzon and Bercker; Neukirchen-Vluyn: Neukirchener Verlag.
- Weissbach, F. 1911. *Die Keilinschriften der Achämeniden*. Vorederasiatische Bibliothek 3. Leipzig: Hinrichs.
- Wilhelm, G. 1978. "Ist das Elamische eine Ergativsprache?" *Archäologische Mitteilungen aus Iran* NS 11:7–12.
- . 1982. "Noch einmal zur behaupteten Ergativität des Elamischen." *Archäologische Mitteilungen aus Iran* NS 15: 7–8.
- Zadok, R. 1984. *The Elamite Onomasticon* = Suppl. 40 to *Annali* (Istituto Universitario Orientale di Napoli) 44.
- . 1991. "Elamite onomastics." *Studi epigrafici e linguistici sul vicino oriente antico* 8:225–237.
- . 1995. "On the current state of Elamite lexicography." In *The Lexicography of the Ancient Near Eastern Languages* = *Studi epigrafici e linguistici sul vicino oriente antico* 12:241–252.



# Hurrian

GERNOT WILHELM

## 1. HISTORICAL AND CULTURAL CONTEXTS

### 1.1 History of the language and its speakers

Hurrian is an ancient Near Eastern language widely spoken in the northern parts of the Fertile Crescent (present-day northern Iraq, northern Syria, southeast Turkey) from at least the last quarter of the third millennium BC on until the end of the second millennium BC. It survived for another half millennium in small pockets in the mountainous areas north of ancient Assyria.

A cognate language of Hurrian is Uartian (see Ch. 5) which is attested in texts from the late ninth to the late seventh century BC. Apart from Uartian, Hurrian is an isolated language without a genetic relation to any other known ancient Near Eastern language. A genetic relation between (reconstructed) Proto-Urarto-Hurrian and (reconstructed) Northeast Caucasian has been argued for, but it is not generally accepted. If the connection could be demonstrated, it would be a rather distant one.

Hurrian is first attested in a few words and personal or place names mentioned in Akkadian texts of the Akkade period (twenty-third to twenty-second centuries BC). The term *Old Hurrian* (herein abbreviated OH) has been coined for the language of a royal inscription most likely to be dated to the Ur III period (twenty-first to twentieth centuries BC), but it is also used for the more archaic dialect(s) of the second millennium.

During the first half of the second millennium BC (Middle Bronze Age) there are many hundreds of Hurrian personal names attested from the northern parts of the Fertile Crescent (from the Zagros Mountains in the east to the Mediterranean coast), but only little more than a dozen Hurrian texts, still unintelligible for the most part.

By far the majority of Hurrian texts comes from the second half of the second millennium BC (Late Bronze Age). Hurrian disappeared as a result of political and ethnic shifts occurring from the late fourteenth century BC onwards. Except perhaps in remote mountainous areas east of the upper Tigris, Hurrian became extinct during the Dark Ages, beginning in the twelfth century BC.

The modern name of the language (English “Hurrian,” French “hourrite,” German “hurritisch”) is based on the geographical term *Ḫurri* which is not very well defined (presumably denoting most of Upper Mesopotamia). On the basis of this name, Hittite forms an adjective *ḫurlili* “Hurrian” (adjective in *-li* formed from *ḫur-la-* “inhabitant of the land of Ḫurri”) which qualifies Hurrian language incantations used in Hittite rituals. In the so-called “Mittani letter” (fourteenth century BC, see §1.2), the Hurrian adjective *ḫurroge* (variant

*hurvoḡe*) “Hurrian” refers to the country; it is unknown whether it could also designate the language.

Other terms for the language are obsolete – *Mitanni* (based on the name of a country in Upper Mesopotamia); *Subarian* (based on the geographical term *Subir*, *Subartu*). The earliest Hurrian attestations and the linguistic relationship with Urartian point to an origin in the most northeastern parts of the Fertile Crescent and in the mountainous areas beyond (most northeastern Syria, most northern and northeastern Iraq, southeastern Turkey). A connection with the flourishing Transcaucasian Early Bronze culture is possible, but cannot be demonstrated.

The earliest city-states with Hurrian rulers, and presumably a population which at least in part spoke Hurrian, were under strong southern (Akkadian, Sumerian) cultural influence and military pressure. Already about 1800 BC there was a solid Hurrian element in the populations that lived between the Mediterranean and the Euphrates, most likely as a result of movements at the end of the third millennium BC.

At the end of the sixteenth century BC, the kings of Mittani (conventionally also “Mitanni”) in Upper Mesopotamia united most of the Hurrian-speaking countries under their control. The dynasty preserved some archaic Indo-Aryan traditions of unknown origins (dynastic names, some gods known from Vedic sources, hippological terms). During the fifteenth century Mittani struggled with Egypt for the control of Syria west of the Euphrates. A balance was reached shortly after 1400 by a peace treaty and the beginning of a series of dynastic marriages. It is in this context that King Tušratta addressed the so-called Mittani Letter to Pharaoh Amenophis III. In the middle of the fourteenth century BC, Mittanian power declined rapidly as a consequence of dynastic turmoil and the rise of Hittite and Assyrian power; the Hittites conquered Mittani’s vassal states west of the Euphrates, whereas the east was annexed to Assyria, though the Mittanian dynasty was able to keep control of a part of its former empire for four more generations. Eventually, however, it disappeared in the course of Assyrian military expansion. The Assyrians removed whole population groups from former Mittani and settled Assyrians in their place in order to gain better control of the region. This policy undoubtedly accelerated the disappearance of the Hurrian language.

By about 1400 BC, the Hittite dynasty had already adopted cultural traditions from the Hurrian-speaking parts of southern Anatolia (Kizzuwatna). Consequently, Hittite kings supported Hurrian cults and introduced them into their capital of Hattuša and in several north Anatolian provincial centers (Sapinuwa, Samukha). Between 1400 and the Hittite collapse, Hurrian thus became a language of cult and learning far removed from Hurrian-speaking areas.

## 1.2 Sources

The oldest Hurrian text is the royal inscription of Tišatal, *endan* of Urkeš. The texts from c. 2000–1500 BC are mainly short incantations from places outside Hurrian-speaking areas – Babylonia (Larsam?) and the Middle Euphrates (Mari, Tuttul) – but there also exist a few texts of other, not yet identified genres (from Mari, Tigananu).

The most important source for the study of Hurrian up to the present time has been the Mittani Letter written about 1355 BC. This diplomatic document of nearly five hundred lines was discovered in the Egyptian capital of Amarna in 1887. All the other known messages sent to Egypt by the royal court of Mittani are written in Akkadian, which at that time served as the diplomatic vernacular throughout the ancient Near East. Archeology has not yet been very successful in uncovering cuneiform tablets in Mittani proper; though recent finds from Tell Brak on the Khabur river show that Akkadian was widely used in the area, especially for

deeds (there is also a small fragment of a Hurrian letter). The thousands of texts found at Nuzi (close to Kirkuk, northern Iraq) and Alalakh (close to Antioch on the Orontes) are all written in Akkadian, but often display Hurrian influence on the levels of lexicon, grammar, and personal names (anthroponymy).

The trade center Ugarit on the Mediterranean coast has yielded a small but important and diverse number of Hurrian texts. A bilingual Sumero-Hurrian lexical list displays unusual Hurrian forms. Several other tri- or quadrilingual (Sumerian, Akkadian, Hurrian, and, optionally, Ugaritic) lists of words or divine names (theonyms) as well as a short bilingual (Akkadian, Hurrian) wisdom text have made important contributions to our understanding of the Hurrian lexicon. Some Hurrian texts are written in the Ugaritic alphabetic script. A group of Hurrian cult songs combine a (mostly unintelligible) text with musical terms (based on Akkadian ones). One single Hurrian letter suggests the use of the language also outside the sphere of cult and learning.

At Emar on the middle Euphrates omen texts and a trilingual (Sumerian, Akkadian, Hurrian) god list were found in the 1970s, but by 1999 they remained unpublished.

By far the majority of Hurrian texts come from Hittite libraries. The most important belong to a series of bilingual (Hurrian, Hittite) literary texts, including a myth, a historico-religious tale, and wisdom texts. Among other texts, Hittite rituals with Hurrian incantations or offering-lists stand out, but there are also epics, myths, prayers, and omen texts. Many Hurrian texts are reported to have been found at Sapinuwa (Ortaköy south of Çorum) beginning in 1991; they are also still unpublished.

### 1.3 Dialects

Despite its vast geographical distribution and its attested history of about a millennium, Hurrian is remarkably homogeneous. The two main dialects are that of the Mittani Letter and the dialect (or presumably a group of closely related dialects) called *Old Hurrian* (being much earlier attested than Mittani Hurrian). Old Hurrian is closer to Urartian, which seems to have separated from Proto-Urarto-Hurrian not later than the early second millennium BC. It is also the dialect on which the study of Hurrian proper names rests. The main features of Old Hurrian have become clearer only since 1983 when the above-mentioned Hurro-Hittite bilingual was discovered. The chief differences between the two dialects lie in the verbal system and in syntax, though the much more complicated syntax of the Mittani letter (virtually our only source for Mittani Hurrian) and its wider use of enclitics may be due to the demands of diplomacy. A few dialectal differences within Old Hurrian are discernible.

## 2. WRITING SYSTEMS

### 2.1 Syllabic cuneiform

Hurrian was mainly written in the syllabic cuneiform script of Akkadian. Departing from common Akkadian spelling practice (see Ch. 8, §2), only a few logograms (word signs originally used to write Sumerian, hence also called *Sumerograms*) were used in writing Hurrian texts.

The scribe of the Mittani Letter used a restricted inventory of syllabic symbols (41 CV signs, 31 VC signs, and 26 CVC signs, some of which had two values – *ḫar/ḫur*, *kal/tan*). The Akkadian script adopted for writing Hurrian distinguished only partially between /e/ and /i/;

the Mittani Letter makes full use of the given oppositions (*te/ti*, *še/ši*, *me/mi*, *en/in*, *el/il*). Long vowels are rendered by *plene-spelling* (e.g., *še-e-*, *ta-a-*). The two vowel signs *u* and *ú* are strictly distinguished in the Mittani Letter and in some texts from Hattuša, indicating a phonemic distinction of /o/ versus /u/.

The syllabary of the Mittani Letter does not distinguish (at least in some cases) between voiced and voiceless stops, but utilizes one Akkadian sign out of a pair – *pa*, not *ba*; *ta*, not *da*; *du*, not *tu*; and so forth. In two instances, the script of the Mittani Letter redefines a pair of Akkadian signs: *gi* and *ki* encode a difference no longer in consonantal voicing, but in vowel quality – *gi* is used for /Ke/ and *ki* for /Ki/. Correspondingly, *gu* represents /Ku/ and *ku* /Ko/.

The sign *wa* can be used for a labiodental fricative plus any vowel; in texts from Hattuša a small vowel sign is added in order to facilitate the correct reading.

Going a step beyond Akkadian practice, Hurrian scribes repeated a vowel sign in word-initial position before a single consonant in order to represent a long vowel: for example, *u-u-mi-i-ni* for *ōmīni*.

## 2.2 Alphabetic cuneiform

Some of the texts from Ugarit are written in the Ugaritic cuneiform consonantal (so-called alphabetic) script, presumably by Ugaritic-speaking scribes. These yield important evidence for the phonology and phonetics of Hurrian, as the consonantal script encodes differences in consonants which are obscured by the syllabic script.

## 3. PHONOLOGY

### 3.1 Consonants

Since Hurrian was written with scripts which were designed for other languages, it is difficult – to a degree even impossible – to establish the phonemic inventory of Hurrian. In writing Hurrian words and names, non-Hurrian scribes in Babylonia and Ugarit distinguish between voiced and voiceless consonants in keeping with their own native phonologies. However, the distribution of voiced and voiceless consonants in Hurrian follows a strictly positional pattern – in other words, is allophonic. Obstruents are always perceived as voiceless (i) in word-initial position; (ii) in intervocalic position when long (doubled); and (iii) in contact with another consonant except the sonorants /m/, /n/, /l/, and /r/. Conversely, obstruents are voiced in all other positions: (i) when word-final; (ii) in intervocalic position when short (single); and (iii) in contact with /m/, /n/, /l/, and /r/. The resulting consonantal inventory would then appear to be as follows, with capital letters used noncommittally (“archiphonemically”) to transcribe the obstruents displaying allophonic voicing:

#### (1) Hurrian consonantal phonemes

P		T		K
	F	Š	S	Ḫ
m		n		
		l	r	
w			y	

The affricate /tʃ/ (transcribed as *c*) is uncertain. Though voicing is not phonemic, it is by convention (following E. A. Speiser and I. M. Diakonoff) marked in (bound) transcriptions in order to facilitate research on loanwords into and from Hurrian (*p:b, t:d, k:g, f:v, s:z, š:ž, ḥ:ğ*). Note that also according to convention, the so-called broad transcription (transliteration) of syllabic cuneiform uses single bars (-) to separate syllabic symbols, whereas the morphemic transcription uses double bars (=).

### 3.1.1 Obstruents

Since the inventory of (1) is certainly too small, we have to assume that there were two or more sets of obstruents with different phonemic manners of articulation which remain unknown.

The fricative /F/ appears to be a labiodental, as the Mittani Letter distinguishes /F/ and bilabial /w/, the first one written with the sign *wa*, the second one with *ú*. The phonetic realization of /S/ and /Š/ is unknown; the latter is rendered as an interdental fricative – [θ] or [ð], depending on position – by Ugaritic scribes.

The texts from Hattuša often replace /P/ by a fricative, apparently in all positions. Whether this fricative is identical with the /F/ of the Mittani Letter or different (bilabial) is unknown.

### 3.1.2 Sonorants

The bilabial glide /w/ appears in word-internal and final position; in word-initial position it seems to be restricted to loanwords and foreign names. There is also a glide /y/ which, however, is rendered as *i*, *ī* in modern transcription, as the writing system in many cases is ambiguous.

## 3.2 Vowels

The Mittani Letter distinguishes five phonemic vowels – /a/, /e/, /i/, /o/, /u/ – with two quantities each. Only the Mittani Letter carefully marks vowel length. The texts from Hattuša show an instability of the opposition /e/ : /i/.

The existence of diphthongs is uncertain. At least some sequences of vowels (e.g., the suffix *-ae*, see §4.4.9, instrumental) can be shown as bisyllabic, and may be even divided by a glottal stop.

## 3.3 Phonotaxis

The practice of syllabic cuneiform orthography prohibits the unambiguous representation of biconsonantal clusters in word-initial or final position, and of triconsonantal clusters in word-internal position. There are no hints that such clusters actually exist in Hurrian; moreover, the appearance of anaptyctic vowels suggests that in this respect the script conforms to the language.

The liquids /l/ and /r/ do not appear in word-initial position.

In the language of the Mittani Letter, strict constraints govern final position: vowels or /n/ occur in most cases; the consonants /Š/, /w/, and /F/ or /P/ are limited to one suffix each. In Old Hurrian – especially in the case of divine and place names – /T/, /K/, /Ḫ/, /l/, /m/, and /r/ also appear in final position.

### 3.4 Accent

Hurrian seems to have a stress accent which falls on the penultimate syllable of words (including their suffixes), enclitics not counted. In some cases, stress causes a vocalic change (lengthening and lowering): for example, *túri* “low” versus *turé=na* (this is the morphemic transcription; the transliteration of the cuneiform spelling is *du-ú-re-e-na*) “the low ones.”

### 3.5 Phonological processes

Several Hurrian phonological processes, synchronic and/or diachronic, can be identified.

#### 3.5.1 Anaptyxis

Vowels are inserted under two conditions:

1. Presumably with the shift of stress caused by addition of a suffix: for example, (i) *évri* “lord”: *evérni* “king”; (ii) *talógli* “servant”: *talógl=la* (pl.); (iii) *ḫavúrni* “heaven”: *ḫavurún= nē=ž* (erg.); (iv) *am=om=i=nni* “administrator”: *am=om=i=nín=n(a)= až=i?=na* (gen. pl., double stress?).
2. With the *-n* affix of the jussive (see §4.5.12.2) and ablative (see §4.4.9) before enclitic personal pronouns, except that of the third-person singular: (i) *ḫaž=i=en* “may he listen”: *ḫaž=i=en=i=ll(a)=ān* “may he listen to them”; (ii) *ed(i)=ī=dan* “because of”: *ed(i)=ī=dan=i=lla=man* “by himself . . . them” (but *ed(i)=ī=da(n)=nna=man* “he . . . by himself”).

#### 3.5.2 Segment loss

The regular disappearance of sounds is seen in three contexts:

1. In the morphologically conditioned contact of two vowels, the first one is elided: for example, (i) *šēna* “brother”: *šēniffu-* “my brother”; (ii) *fīradi* “nobleman”: *fīradardi* “nobility” (in morphemic transcription the elided vowel is given in brackets: *šēn(a)=iffu-*, *fīr=ad(i)=ardi*). For an exception see §3.5.3.
2. The vowels /a/ and /i/ are syncopated between (simple) /n/, /r/, /l/ and (archaic?) dental stops: *\*kul=i=l=e* → *kulle* “I should like to say”; *\*ēni=na* → *ēnna* “the gods”; *\*kud=id=e(n)* → *kut=t=e(n)* “may they fell.”
3. The consonantal segment of the genitive suffix *-ve* and the dative *-va* is lost after the plural suffix *-až* (see §4.4.9).

#### 3.5.3 Vowel contraction

The contact of the final (short) /a/ of the enclitic pronouns (see §4.4.10.1) and the initial (short) /a/ of the enclitic connective *-an* results in a long vowel: *-tt(a)=ān* spelled *-Vt-ta-a-an*.

#### 3.5.4 Assimilation

Hurrian shows both consonant and vowel assimilation, progressive and regressive:

1. When in the case mentioned in §3.5.2, 2, the two consonants are different, the second one is assimilated to the first one: *\*avari=ne-* → *avarre* “field”; *\*tād=ugār=i=l=eva* → *tadugarreva* (see §4.5.12.6).

2. The consonantal segment of the genitive suffix *-ve* and dative *-va* is assimilated to a preceding /P/, /T/, or /Š/: *Tēššob* (a god), gen. *Tēššop=pe*; *Ḫebat* (a goddess), gen. *Ḫebat=te* (see, however, §3.5.5). It is partially assimilated to preceding /u/: *šēn(a)=iffu=we* “of my brother.”
3. Personal names composed of a verbal form and a divine name display various assimilations at the junction: for example, *Ag=i=p-Tēššob* → *Ag=i=t-Tēššob*; *Ḫud=i=p-Šimīga* → *Ḫud=i=š-Šimīga*.
4. The vowel of the two suffixes *-Všt-* (see §4.5.2, 2) and *-kkV* (see §4.5.7) assimilate to the preceding vowel (“vowel harmony”).

### 3.5.5 Metathesis

Consonantal stems joined with a suffix exhibit metathesis: for example, *\*kik=ši* → *kiški* “third”; *Kužalḫ=fe* → *Kužap/fḫe* “of the Moongod”; *Ḫebat=fi* → *Ḫebap/fti* “of (the goddess) Ḫebat.”

The dialect of the texts from Nuzi often (in some cases regularly) inverts the sequence *consonant + liquid*, especially when the initial consonant is a fricative: for example, *fağri* : *fargi*; *eğli* : *ełgi*; *evri* : *ervi*; *šadna* : *šanda*.

## 4. MORPHOLOGY

### 4.1 Word classes

Hurrian grammars distinguish the following word classes: nouns, adjectives (mostly derived from nouns), pronouns, numbers, verbs, and particles (including enclitics). Nouns, numbers, and verbs may easily change their word class: for example, *eman* “ten,” *eman=am=ož=aw* “I made tenfold,” *eman=di* “group of ten,” *eman=d=o=g=li* “decurio”; *ḫan=i* “child,” *ḫan=ašt=i=kki* “she will not give birth,” *ḫan=ir(i)=ra* “those who have given birth,” *ḫan=o/u=mb=a=z=ḫe* “fertility”(?).

### 4.2 Roots

Hurrian words are composed of (i) roots, (ii) optional root-complements, and (iii) monofunctional nominal or verbal suffixes in a strictly sequential order. The root is always in initial position. Most roots are monosyllabic, but a few are reduplicated. The morphology of Hurrian is fundamentally of the agglutinating type.

### 4.3 Root-complements

A root can be semantically modified by one or two (possibly three) root-complements. In many cases, the semantic value of the root-complements has not yet been established. Root-complements in most cases are monosyllabic; those which traditionally have been defined as bisyllabic might well be composed of two root-complements. Root-complements are listed in (2), though the list is not exhaustive.

#### (2) Hurrian root-complements

- A. *-ağ-* (*-aḫḫ-*?), *-iğ-*, *-oğ-*, *-uğ-*, meaning unknown: *tapš=āğ=i* “cupbearer,” *pūz=iğ-* “dip into,” *irn=ōğ-* “make equivalent,” *šab=ūğ-* “?”

- B. -*al*-, meaning unknown: *samm=al*- “tear off,” *ḫež=āl*- “be naked”
- C. -*am*-, factitive: *eman=ām*- “make tenfold,” *šin=am*- “double,” *nikkass=am*-, *nissakk=am*- “account” (Akkadian *nikkassu* “account”)
- D. -*an*-, -*ann*-, causative: *keb=ān*- “send,” *ar=ann*- “let give,” *an=an*- “please”
- E. -*an*-, meaning unknown: *ḫab=an*- “go,” *kil=ān*- “?”
- F. -*and*-, meaning unknown (bimorphemic?): *pic=and*- “rejoice”
- G. -*ang*-, meaning unknown (bimorphemic?): *pir=ang*- “flee”; *pūd=ang*- “report (to the authorities)”
- H. -*apš*-, -*epš*-, meaning unknown: *šin=apš*- “change” (*šin* “two”), *kig=apš*- “change repeatedly” (*kig(a)* “three”), *par=apš=i* a qualification of a field, *pur=apš=i* a priest, *tag=apš=i* “horse blanket,” *eğ=epš*- “constrict”
- I. -*ar*-, iterative-frequentative: *am=ar*- “treat badly,” *an=an=ar=ešk=i* “joy”(??), *ḫāž=ar*- “anoint,” *ḫāž=ar=i* “oil,” *pašš=ar*- “send (regularly),” *šid=ar*- “curse constantly,” *šid=ar=ni* “curse” (noun), *tād=ar*- “love constantly,” *fand=ar=i=nni* “cook,” *urb=ar=i=nni* “butcher”
- J. -*až*-, intensive(?): *ḫaž=až*- “do listen”(?)
- K. -*om*-, meaning unknown: *am=om*- “supervise” (*am*- “look at something”), *am=om=i=nni* “chief administrator”
- L. -*ugar*-, reciprocal (bimorphemic?, cf. -*ar*-): *ag=ugar*- “dispatch,” *ašt=ugar=i* “equivalent,” *ḫub=ušt=ugar=a* name of a divine vizier, *tād=uḡār*- “love one another”
- M. -*up(p)*-, meaning unknown: *kad=upp*- “?” (*kad*- “say”), *tān=upp*- “?” (*tān*- “do, make”)
- N. -*ur*-, meaning unknown: *ag=ūr=ni* “chiseling,” *kul=ūr*- “spell over something” (*kul*- “say, speak”)

## 4.4 Nominal morphology

The Hurrian noun (and a small number of nonderived adjectives) consists of (i) a root, which may be semantically modified by (ii) a root-complement or two, (iii) an optional noun-formation suffix, and in most cases (iv) a thematic vowel. In addition, by attachment of (v) derivational suffixes a noun may form a derived noun or adjective. There are two numbers (singular and plural), but no grammatical genders. To a noun (derived or primary) relational, possessive, number, case, and congruence suffixes – in a strictly sequential order – may be added, which may be followed by enclitics.

### 4.4.1 Nominalization of the root

A root may become a noun by addition of the thematic vowel -*i* (presumably -*ə* in final position, -*ē*- before a suffix, also -*e*?, see §4.4.3): thus, *eḡl=i* “salvation” (*eḡl*- “save”), *fur=i* “view; eye” (*fur*- “see”), *ḫan=i* “child” (*ḫan*- “give birth”), *ḫalv=i* “enclosure” (*ḫalv*- “enclose [by wall or fence]”), *ḫezm=i* “girdle” (*ḫezm*- “gird”), *ḫāž=ar=i* “oil” (*ḫāž=ar*- “to anoint repeatedly”), *mad=i* “wisdom” (*mad*- “to be, to prove oneself wise”), *sull=i* “fetter” (*sull*- “bind”).

### 4.4.2 Noun-formation suffixes

Each of the following suffixes can be identified:

1. -(*a*)*d=i*, basic meaning unknown, in some cases collective: the allomorph -*dī* appears when the root ends in a vowel or single post-vocalic *l*, *m*, or *n*, otherwise -*adī*:



*amm=adi* “grandfather, ancestor, elder,” *šaḡadn=adi* “halfshekel” (*še/aḡt-* “half”), *šigl=ade* “shekel” (Akkadian *šiglu*), *pariss=ade* a measure of capacity (Akkadian *parisu*), *kel=di* “luck, well-being,” *\*ḡel=di* “sublime,” *kum=di* “tower” (*kum-* “erect?”), *fīr=adi* “nobleman” (*\*fīr-* “remove, untie”), *nakk=adi* a form of real estate (*nakk-* “release”), *eman=di* “group of ten” (*eman* “ten”), *tumn=adi* “with four spokes” (*tumni* “four”).

2. *-arb*, adjectives denoting age of animals (see §4.7.2, 6).

3. *-ardi*, collectives: *att=ardi* “forefathers” (*attai* “father”), *ēl=ardi* “female relatives” (*ēla* “sister”), *fīr=ad=ardi* “nobility,” *irīn=n(i)=ardi* “class of equals,” *mariyā=nn(i)=ardi* “class of chariot fighters,” *pura=m(e)=ardi* “domestic staff,” *šāl=ardi* “group of daughters.”

4. *-aure*, patient-oriented participle: *ḡuž=aure* “someone who is bound” (i.e., “a prisoner”; see also *-iri*).

5. *-bade*, meaning unknown: *ḡir=i=bade* “fixed by a peg” (*ḡiri* “wood?”), *tīd=i=bade* “counting”; compare *-o/ubade*, negative adjectives (morphology unclear, with negative suffix *-ōv-* or derivational *-o-* [see 4.4.5]?): *nir=o/ubade* “bad” (*nir-* “be good”), *faḡr=o/ubade* “ugly, bad,” *kul=o/ubade* “unnamed,” *naḡḡ=o/ubade* “uninhabited.”

6. *-danni*, *-denni* (*-da/e+nni* ?), OH *-dan*, terms of profession: *abul=dann-* “gate-keeper” (Akkadian *abullu* “city gate”), *ḡāž=ar=denn-* “perfume maker” (*ḡāž=ar=i* “oil”), *šellin=dann-* “administrator,” *en=dan* title of a ruler (Sumerian *en*).

7. *-i(=)di*, meaning unknown: *šug=idi* “one horse carriage”(?) (from *šugi* “one?”), *tar=idi* “pot” (*tari* “fire”), *ḡub=idi* “young male calf” (from *ḡub-* “smash, break?”), presumably *\*pašš=idi* as base for *paššitḡe* “ambassador” (*pašš-* “send”).

8. *-i(=)ri*, agent-oriented participle (cf. also *-aure*): *tab=iri* “someone who has cast (metal),” *pa=iri* “someone who has erected (a building).”

9. *-ki*, meaning unknown: *fut=ki* “son” (*p/fud-* “beget”), *\*katki* “utterance”(?) (*kad-* “speak”), *id=ar=gi* place for deposition of magically negative substances, *it=ki* “mortar” (*id-* “crush”), *\*ar=gi* “gift”(?) (*ar-* “give”). Presumably a group of nominals in *-a/e/i/oški* also contains this suffix *-ki*: *tād=ar=ašk(i)=ae* “affectionately” (?; a nominal used adverbially), *an=an=ar=eški* “joy”(?), *tād=ir=eški* “love”(?), *er=ōški* an object.

10. *-k(k)a*, meaning unknown: *Šav=oš=ka* name of goddess (older form *Ša(v)oža*, cf. *šav=ož=i* “great”), LUGAL-*ka-* “king,” LÚ-*ka-* “man” (both based on Sumerograms), *aštaḡa* “woman” (*ašti* “woman”), *taḡḡag/kka* “man” (*taḡḡe* “man”); personal name *ḡanakka* (*ḡan=i* “child”).

11. *-li*, nouns of profession (cf. also §4.4.6 (4A)): *kēb=li* “hunter” (*kēb-* “put,” presumably traps), *tab=li* “smith” (*tav/b-* “to cast metal”).

12. *-m(m)e*, meaning unknown: *pura=m(m)i/e* “slave”; perhaps also in *ulme/i* “female slave,” *elami* “oath,” *ḡalme* “singing,” *ōlmi/e* “weapon,” *tažme* “gift”(??).

13. *-ni*, individualizing, basic meaning unknown: *everni* “king” (*evre* “lord”), *irīn=ni* “equal” (from *\*irn=i* “?”, this from *irn-* “be equivalent”), *šukka=ni* “single” (*šukki* “one”), *Māžriā=ni* “Egyptian,” *Mitta=ni* (from *Maitta=ni*, toponym based on personal name *Maitta*). The suffix often appears as a form, that is parallel to *a*-stems (see §4.4.3.2): *p/fabni*, *p/faba* “mountain” (note also *faban=ni* “mountain range”), *muž=ni*, *muž=a* “good order,” *tiž=ni*, *tiž=a* “heart.” The suffix is attached to kinship terms and then used in the formation of personal names: *šen=ni* (*šen=a* “brother”), *el=li* (*ēl=a* “sister”), *men=ni* (*men=a* female relative, twin sister[?]), *atta=ni* (*atta=i* “father,” for *-i* see §4.4.4). Often *-ni* is suffixed to bi- or trisyllabic roots, which in some cases can be analyzed as *root plus root-complement*; the basic function remains unclear: *šid=ar=ni* “curse” (*šid-* “curse,” with iterative *-ar-*), *kapp=ar=ni* a vessel (*kapp-* “fill”), *ḡavur=ni* “heaven,” *šugur=ni*, *šegur=ni* “life,” *taržuwa=ni* “man.” Compare also *-ni/-nni* as a derivational suffix (§4.4.6, 2).

14. *-šari* (*-zari* after *n*), collectives: *en=zari* “gods” (*eni* “god”), *furul=z/šari* “temple complex”(?) (*furli* “temples”), *ḥanizari* “children”(?) (*ḥani* “child”), *mariya=n=zari* “corps of chariot owners”(?) (*mariyanni* “chariot owner”), *tip=šari* “matter” (*tivi* “word, matter”).

15. *-(a)=šše, -ži, -zi*, abstracts, but in some cases concrete nouns, especially words for buildings; also used for forming ordinal numbers (see §4.7.1). The allomorphs *-ži* and *-zi* appear after single postvocalic *m, n, l* and *r*. The abstracts in *-a=šše* are exclusively derived from words for high-ranking men or for gods, which often are stems ending in *-a* (see §4.4.3.2).

15A. *Abstract nouns*: *all=a=šše* “queenship” (*alla=i* “queen”); *\*att=a=šše* “position of a father” (*atta=i* “father”), *puram=ži* “slavery” (*purame* “slave”), *šarr=a=šše* “kingship” (*šarri* “king”), *taḥḥ=a=šše* “manliness” (*taḥḥe* “male,” < *\*taḥḥai*??), *taržuwan=zi* “mankind,” *ušt=a=šše* “heroism” (*ušta=i* “hero”), also *ušt=a=n(i)=zi* (*uštani* “hero”), *tamga/ir=a=šše* “gain” (from Akkadian *tamkāru* “merchant”), *itk=a=l=zi* “purity” (*itki* “pure, clean”); *tal=aḡ=o=l=zi* “attraction” (in a ritual of evocation).

15B. *Concrete nouns*: *salam=ži* “statue” (from Akkadian *šalmu* “statue”), *pidar=ži* “stable for cattle” (*pidari* “bull”), *tibiš=ši* “strawstack” (*tibni*, from Akkadian *tibnu* “straw”), *lippur=ži* a building (from *\*nippuri*?).

16. *-umme/i*, infinitives: *itt=ummi* “go,” *faḡr=umme* “be in good relation,” *udr=ummi* “protect”; directive in *-e*: *kur=ušt=umme=n(e)=e* “in order to dig.”

### 4.4.3 Thematic stems

Hurrian nouns are classified as *thematic* or *athematic* according to the presence or absence of a thematic (stem) vowel. While stems in ancient Indo-European languages are similarly distinguished (see Ch. 17, §3.4), the Hurrian and Indo-European processes are quite distinct and should not be confused.

#### 4.4.3.1 *i*-stems

Most nouns have the thematic vowel *-i*. Apparently it has no specific function except to nominalize the root. In many cases it can be shown that *-i* before a suffix is lengthened and lowered to *-ē-* (see §3.4). It is not clear to what extent there exist *e*-stems distinct from the (i) *i*-stems (see e.g. *aš-ḥé* “animal skin,” *šiye* “water,” *ku-un-kal-le-e* “broad-tailed sheep,” and the personal name *Še-iš-we-e* [šēšfe “kid”]), and (ii) the word formation or derivational suffixes *-me, -šše, and -ḡe/-ḥḥe*.

#### 4.4.3.2 *a*-stems

The thematic vowel *-a* marks kinship terms, some divine names, and a few other words. For most *a*-stems there is a form in which *-ni* replaces *-a* (see §4.4.2, 14). Examples are *šēna* “brother,” *ēla* “sister,” *šāla* “daughter,” *nēra* “mother,” *mēna* “twin sister(?)” (see also §4.4.5 for *a*-stem kinship terms with honorific *-i*), *tiža* “heart,” *f/paba* “mountain,” *muža* “good order(?)” Divine names (some attested as elements of personal names only) include: *Šimīga* (beside *Šimīge*), *Išḥara*, *Tamgina* (*Damkina*), *Tilla*, *Naja*, *Ḥamanna*, *Ḥurra*, *Nuza*. For words with the suffix *-kka* see §4.4.2, 11.

#### 4.4.3.3 *o/u*-stems

These stems mainly appear as names of non-Hurrian – in few cases also Hurrian – origin in the texts from Nuzi: *Marduku*, *Šelwulḥu*, *Kelžu*, *Kungu*, *Niru*, *Pendu*, *Šindu* (personal names), *Nullu* (country), *Nuzu* (city), *Šayu* (goddess[?], element in female personal names).

#### 4.4.4 Athematic stems

Stems formed without a thematic vowel seem to occur more frequently in the earliest phase of the Hurrian language (mostly late third millennium BC). Some of the athematic stems later become thematic *i*- or *a*-stems: *šen* “brother” (cf. *šēna* and *-šenni*), *mad* “wisdom” (cf. *maḍi*), *adal* “strong” (also second millennium; seldom *adli*), *muž* divine name (cf. *muža*, *mužni*), *Kažiar* name of the mountain *Ṭūr<sup>c</sup> Abdīn* (cf. later *Kažiari*). Several names of gods, heroes, persons, and places are athematic: *Tēššob* (cf. *Teššoba/i*), *Ḫebat* (cf. *Ḫeba*), *Kužug* (cf. *Kužuğa*, *Kuža*), *Nubadig*, *Tažmiž*, *Šaluz*, *Šeriž* (cf. *Šeri*), *Ḫurriž* (cf. *Ḫurri*); *Gilgamiž*.

#### 4.4.5 Honorifics

Some *a*-stems which denote human beings held in respect add a suffix *-i*: *alla=i* “lady, queen,” *atta=i* “father”; *\*umma=i* “mother”? (attested only as personal name), *ušta=i* “hero.” The name of the sun-god *Šimīge* seems to be a contraction of *Šimīga=i* (cf. *Šimiga*); perhaps also *taḫḫel/taḫe* “man” from *\*taḫḫa=i* (cf. *taḫa* in personal names).

#### 4.4.6 Derivational suffixes

These suffixes, which form either nouns or adjectives, follow the thematic vowel (and in rare cases also the possessive suffixes, for which see §4.4.8). Some of them (*-ni*, *-šše*) are identical in form with the noun-formation suffixes, but their position in the sequence of suffixes is different. In the case of thematic stems in *-i* the “derivational vowel” *-o*- or *-u*- replaces the thematic vowel, whereas stems in *-a* keep it. The derivational pattern has a parallel in the pronominal system which often shows an opposition between an absolutive ending in *-i* and oblique cases with *-o/u*- occurring before the case ending (see §§4.4.10.2 and 4.4.10.4). In very rare cases – apparently in old forms – the thematic vowel is not replaced by the derivational vowel. Moreover, some derivational suffixes follow a different pattern and do not replace *-i* by *-o/u*-.

The derivational suffixes are as follows:

1. *-ge*, *-ḫḫe*, adjectives of appurtenance: the form *-ge*, with the voiced initial consonant and used chiefly with geographical names, is treated in 1A–1E, *-ḫḫe* in 1F; both in 1G–1H.

1A. *i*-stems (*-ge*): *ḫurr/ḫurv=ō=ge* “Hurrian” (*Ḫurri*, *\*Ḫurvi*), *ḫatt=o=ge* “Hittite” (*Ḫatti*), *lupt=o=ge* “Luptian” (Lupti [a town]). When the word ends in *-ni*, *-li*, or *-ri*, the adjectives in *-ge* are commonly formed without the derivational vowel: *kibir=ge=n(a)=až=a* (dat. pl., *Kibri* [a town]); *ḫamar=ge* “belonging to the *ḫamri*-sanctuary; *pabil=g(e)=a* “in Babylonian” (from Akkadian *Bābili*); *bidin=ge* local form of a goddess (*Bidin* [a town]). The derivational vowel may, however, remain: *Ḫiri=ge* “wooded”(?) (name of a country, *ḫiri* “wood”), *atta=šši=ge* “paternal property” (*attai* “father”), *ess=o=šši=ge* “?” (a kind of field).

1B. *a*-stems: *ankuwa=ge*, *ḫattarina=ge*, *šabinuwa=ge*, *tameninga=ge* (all based on names of cities), *alžyḡ* (Ugaritic consonantal spelling) = *\*alažiya=ge* “Cyprian” (*Alažiya* “Cyprus”).

1C. *Athematic nouns*: *tugriš=ḫe*; *mardaman=ge*, *igingalliš=ḫe*, *ažuḡiš=ḫe* (all based on names of cities), *mugiš=ḫe* (*Mug/kiš*, name of a country).

1D. *A special group of words*: based on roots which are all attested in verbal use and which preserve *-i* (cf. §4.4.6, 4B): *pašš=i=ge* “consignment” (*pašš-* “send”), *pa=i=ge* “ready for building (a house)” (*pa-* “build”), *kunz=i=ge* “reverence”(?) (*kunz-* “bow”), *un=i=ge* “offering”(?) (*un-* “bring”), *na=i=ge* “pasture” (*nav-* “graze”).

**1E. Multiplicative numbers:** see §4.7.2, 4.

**1F. *i*-stems (-*h<sub>2</sub>he*):** the form of the suffix with the initial doubled consonant, -*h<sub>2</sub>he*, is used principally with *i*-stems: *h<sub>2</sub>iyar*=*o*=*h<sub>2</sub>he* “gold, golden” (*h<sub>2</sub>iyari* “?”), *šiniber*=*o*=*h<sub>2</sub>he* “of ivory” (\**šinibēri* “ivory” < Akkadian *šinnipīri*), *ašt*=*o*=*h<sub>2</sub>he* “female” (*ašti* “woman, wife”), *tur*=*o*=*h<sub>2</sub>he* “male,” *tiž*=*n*=*o*=*h<sub>2</sub>he* “heart-shaped” (*tiž*=*ni* “heart”), *hažman*=*o*=*h<sub>2</sub>he* “colored like the *hašmānu*-stone,” *šimig*=*o*=*h<sub>2</sub>he* “belonging to the sun-god” (name of a gate).

**1G. -*ge/h<sub>2</sub>he*-complexes:** several suffix complexes seem to contain the suffix -*ge/h<sub>2</sub>he*, such as the following: -*aš<sub>2</sub>h<sub>2</sub>*- (adjectives based on abstracts: *ašt*=*aš<sub>2</sub>h<sub>2</sub>* “female attributes,” *aštašše* “womanliness,” *ašti* “woman”); -*iš<sub>2</sub>h<sub>2</sub>*- (*turiš<sub>2</sub>h<sub>2</sub>* “west,” *turi* “low”); -*uš<sub>2</sub>h<sub>2</sub>*- (utensils: *ağr*=*o*=*š<sub>2</sub>h<sub>2</sub>* “incense bowl,” *ağri* “incense”); -*at<sub>2</sub>h<sub>2</sub>*- (mostly terms for household utensils: *kaz*=*o*/*ul*=*at<sub>2</sub>h<sub>2</sub>* a large bronze pot, from *kazi* “jar?”); -*it<sub>2</sub>h<sub>2</sub>*- (*pašš*=*it<sub>2</sub>h<sub>2</sub>*/*e* “envoy,” *pašš*-, “send”); *nir*=*an*=*it<sub>2</sub>h<sub>2</sub>* a kind of wood); -*o/ut<sub>2</sub>h<sub>2</sub>*- (*na<sub>2</sub>h<sub>2</sub>h<sub>2</sub>*=*o/ut<sub>2</sub>h<sub>2</sub>* a seat).

**1H. Nouns of profession:** such nouns can be derived from adjectives of appurtenance by addition of the suffixes -*li* (see §4.4.2, 11) and -*ri*. For -*li* there are three patterns, presented here from least to most commonly occurring: the first (rare) preserves the suffix -*ge* unchanged: *šina*=*ge*/*i*=*l*- “crown prince; second quality” (*šina* “two”); the second shows the derivational vowel -*o/u*- before -*li*: *mardad*=*o*=*g*=*o*=*li* “carpet weaver” (from Akkadian *mardatu* “carpet”); and the third lacks the derivational vowel: *halz*=*o*=*g(e)*=*li* “district governor” (*halzi* “district”). The suffix -*ri* is seen, for example, in *am*=*om*=*i*=*h<sub>2</sub>h<sub>2</sub>*(*e*)=*o/u*=*ri* “administrator” (*am*- “see”).

Derivational suffixes other than -*ge/h<sub>2</sub>he*- are:

2. -*ni*, -*nni*, adjectives and nouns: the suffix -*ni* is found, for example, in *te*(*yi*)=*ō*=*n*=*ae* “widely” (\**teyi*(?) “much”), *fağr*=*o*=*n(i)*=*ne*=*n* “beautifully” (*fağr*- “be beautiful”), *pic*=*o*=*n(i)*=*ne*=*n* “happily” (*pic*- “please”). Examples of -*nni* are: *mād(i)*=*o*=*nni* “wise” (*madi* “wisdom”), *attan(i)*=*o*=*nni* “father” (*attani* “father”), the personal names *Šenn*=*o*=*nni* (*šen*=*ni* “brother”) and *Men*=*o*=*nni* (*men*=*ni* female relative), *haž*=*i*=*kk*=*o*=*nni* “deaf person” (*haž*- “hear,” -*kk*- is a negative), and terms of profession like *urb*=*ar*=*i*=*nni* “butcher” (*urb*- “slaughter”), *fur*=*o*/*ull*=*i*=*nni* “diviner” (*fur*- “see”).

3. -*ssi*, adjectives and nouns of suitability: *šen(a)*=*iffu*=*ssi* “suitable to my brother,” *ašt*=*o*/*u*=*ss*- a garment (*ašti* “woman”), *pağ*=*o*/*u*=*ss*- a headgear (*pāgi*/*e* “head”).

4. -*ži*/-*šše*:

4A. -*ži*, adjectives: *nīr(i)*=*o*/*u*=*ž(i)*=*ae* “well” (adverb in -*ae*, *nīr*- “be good”), *talāv(i)*=*o*=*ži* “great,” *fağr(i)*=*ō*=*ži* “good,” *šav(i)*=*ō*=*ži* “great.”

4B. -*šše*/*i*, nouns: *itt*=*o*/*u*=*šš*- “garment” (*itt*- “clothe”), *suğr*=*o*/*u*=*šše* “meadow” (*suğri* “grass”), *nakk*=*o*/*u*=*šše* a military class (*nakk*- “release”). A special group of words in -*šše* are based on roots which are all attested in verbal use and which preserve the -*i*: *šar*=*i*=*šše* “desire” (*šar*- “wish, demand”), the personal name *Pašš*=*i*=*šše* “sending” (*pašš*- “send”) (cf. §4.4.6, 1D).

5. -*bur*, negative: *mānn*=*ō*=*bur* “is not” (*manni* “is; he,” see also §4.4.10.2 and §4.5.11); compare also *kuld*=*o*/*ubur* “?”.

6. -*o/ubade*: For this morphologically unclear formant, see §4.4.2, 5.

#### 4.4.7 Relational suffixes

The suffixes -*ne* (sg.) and -*na* (pl.) are anaphoric suffixes which are positioned between the noun and its case endings. They are incompatible with the possessive suffixes (except

perhaps in very rare, but still dubious cases) and they do not occur with names (except in a few cases of appellatives used as names like, e.g., *Kešše* “the one who sets (traps),” i.e., “hunter”). Singular *-ne* never occurs in the absolutive case, but *-na*, a plural marker, does. The two suffixes also precede most of the case endings which mark agreement of genitive modifiers with their head noun (*Suffixaufnahme*, see §5.2). Examples follow: *ērbī=ne=ž* “a dog” (ergative), *ōmin(i)=ne=ve allai* “the lady of the country,” *tažē=nē=va ed(i)=ī=da* “concerning the gift (dative);” *paššitḫe=na* “the envoys,” *evren=n(a)=až=už* “the kings” (ergative), *ōmīn(i)=n(a)=až=a* “in the countries” (essive).

#### 4.4.8 Possessive suffixes

These suffixes take the position after the noun-formation suffixes. They very seldom occur together with derivational suffixes; though in a few attested cases, they precede them. The pronominal element is clearly separate from the number suffix.

The possessive suffixes of the Mittani Letter – first, second and third person – are presented in (3):

(3)	<i>Singular</i>	<i>Plural</i>
1st	-iffə, -iffē-, -iffu-	-iff=až
2nd	-v/b/p	*-v=až (?)
3rd	-i (Hattuša: -ia-/ -iə)	-i=až

In a text from Hattuša, the second plural is attested twice: *ōlmi=šši* “your weapons”; *ede=ž=uda* “towards your body.”

#### 4.4.9 Case and number suffixes

Hurrian is an ergative language. The agent of an action with explicit patient is marked as an ergative, and the patient as an absolutive. If the patient is not explicitly mentioned, the agent is encoded as an absolutive, as is the subject of an action or a state without implication of a patient:

- (4) A. *šēn(a)=iffu=šš(a)=ān ašti šār=ōž=a*  
 “My brother (*šēna*, erg., with encl. pronoun 3rd per. sg.) has asked for a wife (*ašti*, abs.)”  
 B. *šēn(a)=iffə pašš=ōž=i*  
 “My brother (*šēna*, abs.) has sent” (patient possible, but not mentioned)  
 C. *tažē=n itt=ōš=t=a*  
 “The gift (*taže*, abs.) has departed”  
 D. *ēl(a)=iffə mănē=mmaman tupp=e*  
 “My sister (*ēla*, abs.), she herself, is present”

In contexts not yet defined, a different pattern may replace the ergative one. In this instance the agent is encoded as an absolutive and the patient as an essive:

- (5) A. *fandarin(n)i=nā=ma ag=i=b neğern(i)=a*  
 “And the cooks (*fandarinni*, abs.) took up breast meat (*neğerni*, ess.)”  
 B. *el(i)=a fağr=o=ž(i)=a tan=d=i=b*  
 “She gave (lit. ‘made’) a beautiful banquet (both essive)”

In total, nine or, in a wider definition, fourteen (see [7] below) cases have been identified so far. The plural is marked by three suffixes: (i) the relator plural *-na(-)* (see §4.4.7) for most noun cases (including the absolutive, conditioned, however, by the absence of a possessive pronoun): for example, *en(i)=na* “the gods,” *en(i)=n(a)=až=e* “of the gods”; (ii) the plural suffix *-až-* (not in the absolutive after *-na*): *ēn(i)=iff=až=už* “our gods” (ergative), *en(i)=n(a)=až=už* “the gods” (ergative); (iii) the enclitic personal pronoun *-lla* (third-person plural; see §4.4.10.1), only in the absolutive: *en(i)=iffa=lla* “my gods.”

The nine case suffixes of Hurrian are presented in (6). The ergative suffix *-ž* is absent before the enclitic personal pronouns except that of the third-person singular:

(6)	<i>Singular</i>	<i>Plural</i>
<i>Absolutive</i>	—	-na
	—	-lla
<i>Ergative</i>	-ž	-(na=)až=už
<i>Genitive</i>	-ve	-(na=)až=e (- (na=)aš=fe)
<i>Dative</i>	-va	-(na=)až=a (- (na=)aš=fa)
<i>Directive</i>	-da	-(na=)aš=ta
<i>Comitative</i>	-ra	-(na=)až=u=ra
<i>Ablative-instrumental</i>	-n(i)	?
<i>Ablative</i>	-dan(i)	-(na=)aš=tan
<i>Directive</i>	-ē	?

Conventionally, certain additional case suffixes have been identified. The absence of a plural in most instances and syntactic differences show their separate status.

(7)	<i>Singular</i>	<i>Plural</i>
<i>Essive</i>	-a	-až=a
<i>Instrumental</i>	-ae	—
<i>Aequative</i>	-ož	—
<i>Associative</i>	-nni	—
<i>Associative-essive</i>	-nn(i)=a	-až=o=nn(i)=a

#### 4.4.10 Pronouns

In addition to the possessive suffixes of §4.4.8, Hurrian has personal and deictic pronoun suffixes:

##### 4.4.10.1 Enclitic personal pronouns

These pronoun suffixes are restricted to the absolutive. They appear in two variants the distribution of which is not yet clear: a long form ending in *-a* (more frequent in the Mittani Letter) and a short form (more frequent in the texts from Hattuša and elsewhere):

(8)	<i>Singular</i>		<i>Plural</i>	
	<i>long form</i>	<i>short form</i>	<i>long form</i>	<i>short form</i>
<i>1st</i>	-tta	-d	-dilla	-dil
<i>2nd</i>	-mma	-m	-ffa	?
<i>3rd</i>	-nna	-n	-lla	-l

Only in the position after certain particles (see §4.6.1), the pronominal enclitic *-ma/e* is

used for the third-person singular. In the same position, *-lla* has an optional variant *-lle*.

#### 4.4.10.2 Independent personal pronouns

Both these pronouns (except for the second person) and the deictic pronouns have an absolutive stem in *-e/i* and an oblique stem in *-o/u-*:

(9)	SINGULAR	<i>First</i>	<i>Second</i>	<i>Third</i>
	<i>Absolutive</i>	ište(=n)	fe	man(n)i
	<i>Ergative</i>	iž=až	fe=ž	manu=ž
	<i>Genitive</i>	šo=ve	fe=ve	—
	<i>Dative</i>	šo=va	fe=va	—
	<i>Directive</i>	šu=da	fe=u=da	—
	<i>Comitative</i>	šu=ra	—	manu=ra
	<i>Ablative</i>	—	—	manu=dan
	<i>Associative</i>	šo=nn(i)=a	—	—
	PLURAL	<i>First</i>	<i>Second</i>	<i>Third</i>
	<i>Absolutive</i>	šatti=(lla)	fe=lla	mane=l(la)
	<i>Ergative</i>	šiye=ž	fe=ž=už(?)	man=ž=ož
	<i>Dative</i>	—	fe=ž=a	man=z=a
	<i>Comitative</i>	—	—	man=ž=o/u=ra

In the plural, the genitive, directive, ablative, associative, and instrumental (also singular) are unattested.

#### 4.4.10.3 Deictic pronouns

The system of deictic pronouns distinguishes between spatial and anaphoric deixis. There is a special *alternative* pronoun (the one-“the other”); only the anaphoric and the alternative pronouns make reference to the distinction “proximity versus distance”:

(10)			<i>Singular</i>	<i>Plural</i>
	<i>Demonstrative</i>	<i>absolutive</i>	anni	anni=l(la)
		<i>ablative</i>	annu=dan	
	<i>Anaphoric, proximity</i>	<i>absolutive</i>	andi	andi=lla
		<i>genitive</i>	andu=we	
		<i>dative</i>	andu=wa	
		<i>directive</i>	anduw=ē (?)	
	<i>Anaphoric, distance</i>	<i>absolutive</i>		ane=na/ani=lla
		<i>dative</i>	anu=wa	
		<i>ablative</i>	anū=dan	
	<i>Alternative, proximity</i>	<i>absolutive</i>	akki	akki=lla
		<i>ergative</i>	akku=ž	
		<i>ablative</i>	akku=dan	
	<i>Alternative, distance</i>	<i>absolutive</i>	agi	
		<i>dative</i>	agu=wa	
		<i>directive</i>	agu=da	

In addition, Old Hurrian shows an anaphoric resumptive pronoun *ʾalli*.



#### 4.4.10.4 Interrogative and relative pronouns

This pronoun takes the form *ave-* “who.” Attested is an ergative *ave=\**ž=lla “who . . . us?” (see §4.4.9).

### 4.5 Verbal morphology

Verbs seem to be marked for modes of action; some of the pertinent suffixes are only attested on verbal forms, whereas others modify the meaning of the root prior to the distinction of nominal or verbal inflection (see §4.3). The valence of a verb (transitive or intransitive; see §4.5.1) is indicated by the so-called class-markers. Valence may be modified either by changing the class-marker or by using a suffix which indicates intransitivity.

The verb in the Mittani Letter distinguishes three tenses (present, preterite, and future). Old Hurrian appears to distinguish aspect instead – it is not clear whether aspect is a category of the grammar of the Mittani Letter.

In ergative verb forms, three persons (first, second, third) and two numbers (singular, plural) are distinguished. The subject of nonergative forms in the Mittani Letter is not expressed by the verb form morphology, but only by a noun or by an enclitic pronoun (see §4.4.10.1) following the verb or any other constituent of the clause. For the subject suffixes of nonergative forms in Old Hurrian, see §4.5.9. Two negative suffixes, which are distinguished according to ergativity and nonergativity, are incorporated into the verbal form.

#### 4.5.1 Valence

Valence (the number of noun phrases governed by the verb) is indicated by the vowels treated in §4.5.6. Some roots are attested in both transitive and intransitive use (*un-* “come”/“bring,” *faž-* “enter”; *naḫḫ-* “sit down”/“set, place,” *teğ-* “grow up/raise,” *an-* “be pleased/please”). Normally, however, the root is attested in *either* transitive *or* intransitive usage. A change of valence appears to be marked by the suffix *-ol-*: *ḫic=ūḡ=i=vā=en* “may he not hurt [my heart]” *ḫic=ūḡ=ol=(a)=l=ē=tt(a)=ān* “I will grieve.”

#### 4.5.2 Modes of action

Several suffixes which immediately follow the root-complements (see §4.3) seem to mark modes of action:

1. The suffix *-il(l)-* marks the inchoative: *šid=ar=ill=ō=m* “he began to curse.”
2. The function of the two suffixes *-ol-* and *-Všt-* (see §3.5.4, 4) is not yet clear; perhaps the first one marks duration and the second one result. In rare instances they may appear together: for example, *muž=ōl=ō=m* “he shaped [the goblet]”; *pa=ašt=o=m* “he erected [a temple]”; *tav=ašt=ō=m* “he cast [a goblet]”; *teğ=ešt=a=b* “he grew up”; *ḫub=ušt=aw* “I shall break to pieces”; *ti=ol=ōšt=aw* “I shall crush underfoot.”
3. The rare verbal forms in *-uva* (*taž=ol=uva* “he made it shining”) may also define a mode of action.

#### 4.5.3 Undefined verbal suffixes

There are some more verbal suffixes occupying a position close to the root, the meaning of which has not yet been established: for example, *ešḫ-*, *-imd-*, *-upt-*, *-o/ušk-*, *-o/už-*, *upp-*).



#### 4.5.4 Tenses

The tense suffixes are *-ož-* (*-ōž-* before *-t-*) for the preterite, and *-ēd-* (also *-ed-*) for the future. These suffixes have been explained as aspectual (perfective and imperfective respectively), but *-ož-* is never used for a complete action of the future, nor is *-ēd-* for a noncomplete action of the past. The present tense is morphologically unmarked.

#### 4.5.5 The marker of a kind of direction(?) *-t-*

There is no agreement thus far concerning the function of the suffix *-t-* which follows the tense markers in intransitive verbs of movement, but seldom in ergative verbs.

#### 4.5.6 “Class-markers” (suffixes of valence)

Old Hurrian distinguishes between three so-called “class-markers”: (i) *-a-*: one valence, intransitive, apparently only with verbs of motion; (ii) *-i-*: virtually two valences, but only one valence filled (see, however, the construction of §4.4.9 [5]), transitive-nonergative; (iii) *-o-*: two valences, ergative. In the Mittani Letter the forms in *-o-* are (nearly?) completely absent and have been replaced by forms with *-i-*. In ergative forms, *-i-* is not compatible with the tense suffixes *-ōž-* and *-ēd-*. Except in forms with the negative suffix *-ma* (see §4.5.7), it is also absent in the present tense before the personal suffixes of the first singular and plural.

#### 4.5.7 Negation

Nonergative verb forms take the negative suffix *-kkV*, which follows the class-marker. The vowel agrees with the preceding vowel except before enclitic personal pronouns, where it changes to *-a-*. Intransitive negative forms replace the class-marker *-a-* by *-o-*, which, however, remains *-a-* before the enclitic personal pronouns. Ergative forms are negated by the suffix *-va-* (also found in both ergative and nonergative jussives; see §4.5.12.2) or *-ma-* and, especially in the dialect of the tablets from Hattuša, *-ud-* (also lexicalized: *sul=ud-*, *hemz=ud-* “loosen,” literally “untie”).

#### 4.5.8 Ergative person suffixes

The following forms are attested in the Mittani Letter (and partially elsewhere):

(11)	<i>Singular</i>	<i>Plural</i>
1st	<i>-aw</i>	<i>-aw=ž</i>
2nd	<i>-o</i>	—
3rd	<i>-a</i>	—

At Hattuša, the suffix of the second-person plural displays the ending *-āššo/ō* (written *-°a-aš-šu(-u)*), which seems to invert that order of person and plural suffixes displayed in the first and third plural. A form of the third plural is attested at Hattuša: *-a=ž*.

#### 4.5.9 Old Hurrian person suffixes

Old Hurrian has a suffix *-b* which seems to mark the third person (singular and plural) of intransitive and transitive-nonergative verbs. For ergative verbs, the suffix *-m* appears to mark the third-person singular of both agent and patient.

#### 4.5.10 Indicative paradigms

Sample paradigms of verbs in the indicative mood are presented in (12)–(14):

##### (12) Intransitive verbs

- A. *Mittani Letter*: ūn=a “he/she comes”      ūn=a=lla “they come”  
    ūn=ōš=t=a “he/she came”      ūn=o=kka=l “they do not come”  
    ūn=ēt=t=a “he/she will come”  
 B. *Old Hurrian*: naḥḥ=a=b “he sat down”

##### (13) Transitive-nonergative verbs

- A. *Mittani Letter*: ḥill=i “he/she says”  
    ḥill=ož=i=kka=tta “I did not say”  
    ḥill=ož=i “he/she said”  
 B. *Old Hurrian*: ḥill=i=b “he/she said”

##### (14) Ergative verbs

- A. *Mittani Letter*:  
    tād=aw “I love [him/her]”      tād=i=a “he loves [him/her]”  
    ar=ōž=aw “I gave [it]”      ar=ēd=a “he will give [it]”  
    kad=ēd=aw “I shall say [it]”      kul=i=ā=ma “he does not say [it]”  
    ūr=aw=ž “we want [it]”      tāt=ōž=a “he made [it]”  
    ūr=i=uffu=nnā “I do not want him”      irn=ōğ=ož=i=â=ma “he has  
    not made it equivalent”  
    ḥic=ūğ=ož=i=uffu “I have not hurt [him]”  
    koc=ōž=o “you retained [him]”  
 B. *Old Hurrian*:  
    šid=ar=ill=ō=m “he started to curse him”      nakk=i=uffu=ž “we will not  
    release [them]”  
    fur=ud=o=m “he did not care for it”      tūn=id=o “they forced him”  
    naḥḥ=ōžo “she placed them”

#### 4.5.11 Stative verbs

A small group of verbs expresses state: *tupp*- “be, exist”; *mann*- “be”; *irn*- “be equivalent”; *ur*- “exist.” The class-marker is *-i*- (like the class-marker of transitive verbs) or *-e*-, but the negative form is the same one as that of intransitive verbs (of movement) with the class-marker *-a*-, becoming *-o*- (see §4.5.7): *tupp*=*e*, *tupp*=*o*=*kko*; *mann*=*i*, *mann*=*o*=*kko* (at Hattuša also *mann*=*o*/*ubur*).

#### 4.5.12 Nonindicative moods

There is a rich system of nonindicative moods, which is not yet fully understood; these forms do not contain person suffixes, but markers of mood and plurality only. There is no agreement concerning the terminology of nonindicative moods; that terminology used here in most cases follows Bush (1964) and Diakonoff (1971). Only the best-established patterns are mentioned below.

#### 4.5.12.1 Imperative and cohortative

The imperative and the cohortative are formed by the root and the class-marker, optionally followed by an enclitic personal pronoun; the plural is marked by -ž. Both second- and third-person imperative forms occur, as well as a first plural cohortative:

(15)	<i>Singular</i>	<i>Plural</i>
<i>1st</i>		dilla . . . tād=ugār=i=ž “we wish to love each other” fağr=o=š=till(a)=ān “we wish to be friendly to each other”
<i>2nd</i>	un=a, un=a=mma “come!” ar=i ‘give!’ ḥaž=i=mma “listen!” kel=o, kel=o=m “be satisfied!”	sull=ud=i=ž “unbind!”
<i>3rd</i>	kud=o “let it be felled!”  nakk=o=n(na) “let him/her be released!”	itk=o=ž “let [the temples] be purified!”

#### 4.5.12.2 Jussive

The jussive expresses a request in the third person. Its suffix is *-en* (*-in*, in Hattuša also *-an* [personal correspondence from M. Giorgieri]), which in transitive forms follows the class-marker *-i-* or, in Old Hurrian, *-o-*. The final *-n* of the suffix could be a pronominal element (see §4.4.10.1), but it appears in forms of both the third singular and plural. The plural is marked by *-id-*. The negation of the jussive (“vetitive”) is *-va-* (OH *-v(e)-*) after a vowel, *-ov-* after a consonant: *pašš=i=en* “he may send”; *tašp=o=in* “he shall destroy”; Hattuša *ar=i=an* “may he give”; *ḥa=i=en=i=lla* “may he take them”; *tād=ašt=id=en* “may they love us”; *itt=id=en* “may they go”; *ḥaž=āž=i=vā=en* “may he not listen [to them]”; *ur=ōv=en* “may he not exist”; *nakk=id=ōv=en* “may they not let/send.”

#### 4.5.12.3 Modal -l-

The modal suffix *-l-* (under undefined conditions apparently *-ll-*) combines with various mood forms and modifies their meaning in a way which, however, cannot always be well established.

#### 4.5.12.4 Optative

The optative seems to be formed by the modal suffix *-l-* plus the jussive suffix (without its final *-n*). Third-person forms in texts from Hattuša often have the suffix *-ž* in a nonplural usage, presumably in an intensifying function. Forms from the Mittani Letter include: *ḥaž=i=l=e* “I wish to hear”; *kul=(i=l=e)* “I wish to say”; *ḥaž=āž=i=va=ll=i=lla* “I do not wish to hear them.” From Hattuša come: *kad=i=l=e=ž* “may it speak,” *kir=o=l=e=ž* “may it be relieved,” *tag=o=l=e=ž* “may it be shining.”

#### 4.5.12.5 Potential

The potential is formed by the root and the suffix *-eva*: *ai...faž=ēva* “if [the enemy] invades”; *...tta pic=ošt=ēva* “I would rejoice.”

#### 4.5.12.6 Conditional

The conditional is the potential modified by the modal suffix *-l-*; it is also used to express the contingency of an action: *kad=i=l=ēva* “[a word which somebody] might say”; *ḫill=o=l=eva* “he might say”; *ai=n ur=d=o=l=eva* “if it happened”; *ar=(\*i=) r(<l)=eva=ž* “we are ready to give” (on the assimilation of *-l-*, see §3.5.4, 1)

#### 4.5.12.7 Desiderative

The desiderative is a strong wish which may be modified (intensified?) by the modal suffix *-l-*: *itk=id=anni* “may they purify him/her”; *id=i=l=anni* “may he beat him.”

#### 4.5.12.8 Other possible modalities

In both the Mittani dialect and the dialect(s) of the Hattuša tablets, there occur roots with the suffix *-ai*, which Hittite scribes translated by subordinate clauses. Consider the following final (purpose) clauses: *faž=ai=n* “so that he may enter”; *itt=ai=ž=a=lla* “so that they may go”; *ḫaž=āž=ill=āi=n=i=lla* “so that he may (be ready to (? – inchoative)) hear them.” Forms in *-ai* occur with a preceding *-l-* or *-m-*: *šid=i=l=āi* “so that he cursed [him]”; *naḫḫ=i=l=āi* “[he is someone whom his lord] appointed”; *fur=ī=m=ai=n* “[whenever] he sees him”; *faž=o/u=m=ai* “when he entered”; *kunz=i=m=ai* “while he bows.”

### 4.5.13 Verbal nouns

There is evidence of various Hurrian verbal nouns.

#### 4.5.13.1 Infinitives

For the Hurrian infinitive, see §4.4.2, 16.

#### 4.5.13.2 Nominalized verb forms

Finite verbs may be nominalized by the suffix *-šše* and treated like other nouns: *am=om=i=a=šše* “a dignitary” (ergative third singular); *ūr=i=ā=šše=na* “those which he desires”; *ar=ōž=aw=šše=nē=ve* “of that which I gave.”

Several nominalized verb forms in the Mittani Letter contain an element *-mbū-* which has not yet been well defined (a state achieved as the result of an action?): *ur=i=mbū=šš(e)=o=ḫḫ(e)=a=mān* “and in the manner desired.”

## 4.6 Particles

“Particle” will here be defined as a word which cannot take nominal or verbal suffixes but only enclitic particles (see §4.6.4) and pronouns.

### 4.6.1 Introductory particles

Certain particles introduce clauses: *adi* “so”; *ai* “if”; *alaže-* “whether”; *inna-* “when, as soon as”; *i/unu-* “as”; *panu-* “?”; *ia/e-* (a relative).

### 4.6.2 Adverbs

The following adverbial particles are identified: *anam(mi)* “in this manner”; *ḫenni* “now”; *kuro/u* “again, on the other hand”; *šukko* “once”(?); *tiššan* “very”; *undo* “now.”

### 4.6.3 Interjections

The interjectory particles are *oia* “no!” and *au* “behold!”(?).

### 4.6.4 Enclitic particles

In the Mittani Letter, the enclitic particles are as follows: *-an*, *-mān* (connective for words and clauses); *-man* (emphatic, restrictive: “only”); *-mmaman* (emphatic(?)); *-nīn* (function unknown). Old Hurrian shows *-ma* (connective).

## 4.7 Numerals

Only the numerals 1 to 10, 13 or 30, 14(?), 17 or 70, 18 or 80, 10,000 and 30,000 are known.

### 4.7.1 Cardinals and ordinals

Ordinals are formed from cardinals by the suffix *-šše*, *-ži* (see §4.4.2, 15).

(16)	Cardinal	Ordinal
1	šukki, šuga(?)	?
2	šin(a)	šinzi
3	kig(a)	kiški (<*kik=ši)
4	tumni	tumušše, tumunzi
5	nariy(a)	narišše
6	šeže	?
7	šindi	šendešši
8	kira/i	?
9	tamri/a	?
10	eman	emanzi, emassi(?)
13/30	kigman(i)	
14(?)	šinašinda	
17/70	šindeman(i)	
18/80	kir(e)man	kirmanze
10,000	nubi	
30,000	kiga nubi	

### 4.7.2 Other numerals

Various other numeral formations are attested:

1. *Fractions*: \*ša/eḫt- “one-half”; *tumunzalli* “one-quarter of a shekel.”
2. *Collectives*: *tumn=adi* “four-spoked”; *šež=adi* “six-spoked”; *eman=di* “group of ten people.”
3. *Distributives* (with instrumental suffix *-ae*): *kig=ad(i)=ae* “three each” (see §4.4.2, 1).
4. *Multiplicatives*: *šukki* “once” (see §4.4.6, 1E).
5. *Adverbs* (number with factitive, adjective, and essive suffixes): *šug=am=ḡ(e)=a* “simple”; *šin(a)=am=ḡ(e)=a* “twofold”; *tamr=am=ḡ(e)=a* “ninefold,” *eman=am=ḡ(e)=a* “tenfold”; *šinz=o=ḡ(e)=a* “in the second place.”
6. *Expressions of age* (only attested with Akkadian case ending): *šin=arbu* “two years old”; *kig=arbu* “three years old.”

## 5. SYNTAX

### 5.1 Word order

In ergative clauses (see §4.4.9) the agent usually takes the initial position, followed by the patient and the verb (SOV):

- (17) pašš=iṭh(i)=iffu!=ž tive andi kul=ôž=a  
 “My messenger (*paššīṭhe*, erg.) said this word (*tive*, abs.)”

Word order may be changed by topicalization, as seen in both the Mittani Letter (18A–B) and Old Hurrian (18C):

- (18) A. keb=ān=ož=āw=šše=na fur=ēd=ā=ll(a)=ān šēn(a)=iffu=ž  
 “[The things] which I have sent (*keb=ān-*) my brother will see (*fur-*)”  
 B. un=ā=l=an šēn(a)=iffu=wa  
 “They do come (*un-*) to my brother”  
 C. ḥa=i=en id(i)=ia=n nir=o/ubadi erāde=ne=ž  
 “May the bird (*erade*) take (*ḥa-*) the evil from his body (*idi*)”

Participants in the dative or directive may follow the verb, otherwise they are positioned between the ergative subject and the absolutive object. A modifier (including a genitive) may precede or follow its head.

Hurrian has postpositions, which may govern a preceding dative or genitive. The following are found with a preceding dative: *ed=i=da* (directive of *edi* “person, body,” with a third-person singular possessive pronoun) “with reference to, concerning”; *e/ig=i=da* “within”; *fur=i=da* (*furi* “eye”) “with regard to”; *ā(i)=i=da* “in the presence of” (at Hattuša occurs also the preposition *ābi*). Governing a preceding genitive are: *āi=ē* (directive) “in front of”; *ed=i=ē* “because of, about.”

### 5.2 Agreement

A modifier agrees with its head. The case endings copied from the head are preceded by *-ne-* or *-na-* (see §4.4.7). This also applies to genitive modifiers (*Suffixaufnahme*):

- (19) A. šēn(a)=iffu=we=nē=ž ašt(i)=i=ž  
 “My brother’s (*-we* gen.) wife (*ašti*)”  
 B. šēn(a)=iffu=we=nē=va torub(i)=i=va  
 “To my brother’s enemy (*torubi* ‘enemy,’ *-va* dat.)”  
 C. en(i)=n(a)=āž=(v)e=ne=da šarri=ne=da  
 “To the king (*šarri*, *-da* dir.) of the gods (*en(i)=na*)”

Likewise, nominalized ergative verbs are constructed as modifiers which agree with their head. In this case, the head is always the patient of the nominalized verb, regardless of its case form:

- (20) A. šēn(a)=iffu=ll(a)=ān ūr=i=ā=šše=na tivē=na  
 “The things (*tive*) which my brother desires (*ūr-*)”  
 B. tuppe niḡār(i)=rē=ve ar=ōž=aw=šše=nē=ve  
 “The tablet (*tuppe*) of the dowry which I have given (*ar-*)”

### 5.3 Coordinate and subordinate clauses

There seem to be no special verbal forms for subordinate clauses except the verbal nouns mentioned in §4.5.13.1. Particles occurring in initial position of temporal, conditional, comparative, and other clauses have been cited in §4.6.1.

In relative clauses introduced by the particle *īa-*, *īe-*, the head of the relative clause is incorporated within the clause; the verb is nominalized and stands in agreement with the head. The main clause refers to the head of the relative clause by an anaphoric pronoun:

- (21) [[*īa=llā=nīn šēn(a)=iffu=ž...tivē=na tān=ōž=ā=šše=na*]<sub>REL.CL.</sub> ... and *i=ll(a)=ān šimīge=ne=ž ar=ēd=a šēn(a)=iffu=wa*]  
 “The things which my brother has done, those the Sun-god will give to my brother”

In rare cases the verb of the main clause may be incorporated:

- (22) [*īa=mē=nīn ed(i)=iffə pal=āw [šēn(a)=iffu=ž...tād=i=ā=šš(e)=a*]<sub>REL. CL.</sub>]  
 “I know (*pal-*, erg.) that [my brother loves (*tād-*, erg., nominalized, essive) it], my person (*ed(i)=iffə*)”

## 6. LEXICON

Hurrian is still only very incompletely known, especially as far as the lexicon is concerned. Since Hurrians had been in contact with the peoples of the northeastern parts of the Fertile Crescent since at least the last quarter of the third millennium BC (and presumably much earlier), one should expect a considerable stock of Sumerian as well as Akkadian and other Semitic loanwords. There are indeed some words borrowed from Sumerian in the third millennium, like *en=dan* “ruler” (from *en*); other possible Sumerian loans are disputed (*nathī* “bed,” Sumerian *ná* “bed”). Akkadian loanwords are numerous, especially in texts from the Late Bronze Age. They reflect an extended usage of Akkadian as a second language, or at least as the vernacular of written communication and documentation; examples include: *šarri* “(divine) king” (from *šarru* “king”); *šukkalli* “vizier” (from *sukkallu*); *tupšarri* “scribe” (from *tupšarru*); *tamgarašše* “profit” (from *tamkāru* “merchant”); *salamži* “statue” (from *šalmu*); *ḫassissi* “ear” (from *ḫasīsu*); *arni* “guilt” (from *arnu*).

The tribes who established the Mittani dynasty spoke an archaic form of Indo-Aryan, which left some traces in Hurrian: *mariyanni* “chariot owner” (cf. Sanskrit *mārya-* “young man”); *pabro/unni*, *paridanni*, *pingaranni* colors of horses (cf. *babhrú-* “brown,” *palitá-* “grey,” *piṅgalá-* “reddish”).

Though Hurrian played an important role in Anatolia as a language of learning and ritual, it appears not to have borrowed from the Anatolian Indo-European languages (Hittite, Luwian, etc.) – an appearance perhaps connected with the fact that the Hurrian tablets chiefly preserve texts of Hurrian traditions. There is little doubt, however, that Hurrian was not only a literary language, but was spoken in a court and temple milieu at least in the fourteenth century BC. This is revealed by the occurrence of adjectives based on Anatolian place names: *ḫattoḡe* (“belonging to Hatti”), *šabinuvaḡe* (“belonging to Šapinuwa”), and so forth.

## Bibliography

There is no Hurrian grammar or dictionary which presents the present state of our knowledge of the language. The following monographs have to be corrected and supplemented in the light of the research literature of the last decades:

Bush, F. W. 1964. *A Grammar of the Hurrian Language*. Dissertation, Brandeis University. Ann Arbor: University Microfilms.

Diakonoff, I. M. 1971. *Hurrisch und Urartäisch*. Munich: Kitzinger.

Khačikjan, M. L. 1985. *Churritskij i urartskij jazyki*. Jerevan: AN Armanskoj SSR.

Laroche, E. 1980. *Glossaire de la langue hourrite*. Paris: Klincksieck.

Speiser, E. A. 1941. *Introduction to Hurrian*. New Haven: American Schools of Oriental Research.

Wilhelm, G. 1989. *The Hurrians*. Warminster: Aris and Phillips.

After the completion of this grammatical outline (spring 1999), two grammars of Hurrian have appeared:

Giogieri, M. 2000. "Schizzo grammaticale della lingua hurrica", *La parola del passato* 55: 71–277.

Wegner, I. 2000. *Hurritisch. Eine Einführung*, Wiesbaden: Harrassowitz.



# Uartian

GERNOT WILHELM

## 1. HISTORICAL AND CULTURAL CONTEXTS

### 1.1 History of the language and its speakers

From the late ninth to the late seventh century BC, Uartian was written in the empire of the Uartian kings, stretching from present-day Armenia, Azerbaijan, Iranian Azerbaijan, and northeastern Iraq to the Euphrates. Neither its geographical origin can be conclusively determined, nor the area where Uartian was spoken by a majority of the population. It was probably dominant in the mountainous areas along the upper Zab Valley and around Lake Van. The center of Uartu is the region surrounding Lake Van with its capital of ʾṬušpa (citadel of Van). We do not know when the language became extinct, but it is likely that the collapse of what had survived of the empire until the end of the seventh or the beginning of the sixth century BC caused the language to disappear.

Uartian is closely related to Hurrian (see Ch. 4, §1.1), especially to the dialect conventionally called *Old Hurrian* (see Ch. 4, §1.3). Presumably Uartian branched off from Hurrian not much later than approximately 2000 BC.

The earliest inscriptions of the Uartian kings are written in the Neo-Assyrian script and language (a dialect of Akkadian; see Ch. 8), though after a single generation the Uartian language, for most purposes, replaced the use of Assyrian. The eighth century BC witnessed the climax of Uartian power and the greatest production of Uartian royal inscriptions (the Annals of Arġiṣti I and Sardure II).

The term *Uartian* is based on the geographical name Uarṭu, which was used not only by the Assyrians, but by the Uartians themselves when writing in Assyrian. The Uartian equivalent is the name *Bia=i=ne=lə* “the [people] of [the land of] Bia.” The Uartians’ name for their own language is unknown. The terms *Vannic* or German *chaldisch* which can be found in older literature are outdated.

### 1.2 Sources

Nearly all of the Uartian texts occur as commemorative stone inscriptions on walls, column and pillar bases, steles, and rocks. There are a few clay tablets, which display an experienced hand, suggesting that most likely the scarcity of this medium is due to archeological accident rather than lack of use. Metal objects from the treasuries of Uartian kings are often inscribed with short proprietorial notes. Sealed clay bullae suggest that at least in the seventh century BC perishable materials like leather were also used for writing.

The earliest Urartian inscriptions can be dated to approximately 820 BC, and the latest were written during the last decades of the seventh century BC. Though there are more than five hundred inscriptions, their linguistic value is limited because of their extreme repetitiveness. There are basically two genres of inscriptions: (i) those which commemorate the building activities of the kings and (ii) those referring to their military campaigns. Apart from these there are a few texts recording cultic prescriptions, especially one long and complete text from the early period of the kingdom (Meher kapısı). Important landmarks lying close to Assyrian territories were made prominent by the erection of bilingual (Urartian and Assyrian) steles, the most famous of which is the Kelišin (“blue stone”) stele. It marked a Zagros pass leading to Muşasir, which was of utmost importance to the Urartians as the cult center of their supreme god Haldi.

### 1.3 Dialects

No dialects of Urartian have been identified, though see §4.3.5.7.

## 2. WRITING SYSTEMS

### 2.1 Cuneiform script

Urartian was written in the Akkadian cuneiform script. Like its model, the Urartian writing system uses syllabograms and logograms (or *Sumerograms*, transcribed with capitals). The sign shapes of the Urartian royal inscriptions are basically those of Neo-Assyrian royal inscriptions, with one innovation: from c. 810 BC onwards, Urartian inscriptions on stone and metal avoid intersecting wedges. There is no convincing argument that some sign forms prove the influence of older traditions.

The syllabary is extremely restricted, with CV signs (57) prevailing over VC signs (19). Some of the VC values most common in the Assyrian cuneiform script are not used at all (*aK*, *iK*, *uK*, *iḫ*, *uḫ*, *im*, *um*, *en*, *in*, *un*, *uT*). Identifying the reason for this simplification of the Assyrian sign inventory is difficult. There are numerous hints that in some cases CV signs actually represent consonants only (in some of the oldest inscriptions the ergative suffix is written with the sign *-iš*, whereas the normal spelling is *-še*; the Hurrian equivalent ends consonantly, i.e., *-ž*). Moreover, in some cases it can be shown that CV signs are likely to represent [VC] syllables: the word *kure=l(ə)* (written *ku-re-e-li*) “feet” can be linked to Hurrian *ugri* “foot,” also *ure=l* “feet” (the sign *uK* does not belong to the Urartian inventory); for inverted readings of CV signs see also §3.5.1. CVC signs are used, though much less frequently than in Assyrian royal inscriptions.

One reason for the abandonment of several VC signs may have been that the Akkadian syllabary neutralized the opposition of voiced and voiceless consonants. The use of CV signs as C signs may thus have been regarded as a means of representing this opposition at the end of syllables. Especially relevant might be the Assyrian model: the inscriptions of Assurnasirpal II (887–858 BC) quite often replace a final closed syllable by an open syllable (mostly with <Ci>).

The signs *mi* and *ne* do not occur in the Urartian script with syllabic values. Both of the phonetic sequences [mi] and [me] are represented by the sign *me* (transliterated as *me* or *mi*); correspondingly, the sign *ni* is used for both [ne] and [ni] (transliterated as *né* or *ní*). The sign *’a* is used as a variant of *wa* in a number of cases. A few signs are used with values not adopted from Assyrian traditions: for example, *ga*, *gi*, and *gu* used for a voiced fricative (conventionally transliterated as *ya<sub>x</sub>*, *yi<sub>x</sub>* and *yu<sub>x</sub>*).

Graphemes representing homorganic consonants are at times used interchangeably: for example, *'a-al-tú-bi*: *'a-al-du-bi*; *su-du-qu-ú-bi*: *su-tu-qu-bi*.

CVC signs are used, though much less frequently than in Assyrian. Signs for five vowels are attested: *a*, *e*, *i*, *u*, *ú*. Variant spellings suggest that *u* and *ú* render one vocalic phoneme only, whereas – despite some interchangeability – *e* and *i* refer to different vowel phonemes. Only a few homophonous signs are used, namely *tu* and *tú*, *ar* and *ár*; there are sufficient variant spellings to show that these do not represent different vowel phonemes.

Plene-spelling of vowels is common, though the function of such full representation is not straightforward. There are three possibilities, and each probably actually occurs: (i) a plene-spelling may mark a long vowel; (ii) it may define the quality of the vowel of the preceding CV sign; and (iii) it may simply serve aesthetic purposes in filling a line. In addition, there seems to be ambivalent plene-spelling of vowels. Thus, the final vowel of an *i*-stem is reduced to *ə* in word-final position; its graphemic representation is the vowel inherent in a *Ce/Ci*-sign, to which the vowel sign *e* may be added: for example, *pi-LI* (*LI* has the values [le] and [li]) or *pi-LI-e* (both transcribed as *pilə*, “canal,” absolutive). The same grapheme sequence *-Ce/i-e*, however, can also be used as a variant of the normal spelling *-Ce/i-i-e*: for example, *Hal-di-i-e* or *Hal-di-e* (both transcribed as *Haldi=ə*, “to (the god) Haldi,” dative).

In letters, a word-divider is used, though not always and not systematically.

## 2.2 Hieroglyphic script

There are few short inscriptions written in pictographic symbols which have not yet been deciphered. Only two “hieroglyphs” often carved into the neck or body of large storage vessels have been identified, as units of capacity.

## 3. PHONOLOGY

### 3.1 Consonants

The cuneiform script distinguishes the following consonants, though not in all positions, and there is uncertainty regarding the value of some (see below). The evidence for the glides *w* and *y* is indirect (suggested by spellings such as *-ni-i-e*, *a-i-u-*, *a-ú-i*):

#### (1) Urtian consonants

p	t	k	?
b	d	g	
	ʈ	q	
	s	š	ḥ
	z		
	ṣ		
m	n		
	l	r	
w		y	

It is unclear to what extent consonantal phonemes may exist which are not distinguished by the script; nor is there agreement concerning the phonetic interpretation of some of the graphic renderings of phonemes.

It is safe to assume a tripartite phonemic opposition between voiced, voiceless, and some third set of stops and postdental fricatives. The third set is represented by the cuneiform signs for the so-called emphatic consonants of Akkadian (*t, q, ʒ*). In Urartian these perhaps represent voiceless glottalized or aspirated consonants. It is also possible that the labial stops form such a triad with a consonant /*p̥*/ graphemically not distinguished from /*p*/ and /*b*/. It cannot be determined whether *g* can represent a voiced fricative in every position or only intervocalically. The *h* syllabograms might represent both a voiced and a voiceless phoneme. The comparison of place names written in Urartian cuneiform and in Greek or Armenian script does not yield unambiguous results, as, especially in the case of Greek, it can hardly be ascertained through what intermediate phonemic systems these names passed.

The “sibilant” system is particularly difficult to reconstruct since even in Akkadian, and in particular Neo-Assyrian, the phonetic value of the cuneiform characters is uncertain. On the basis of Greek and Armenian renderings of Urartian place names I. M. Diakonoff has suggested interpreting *š, s, z, ʒ* as /*s*/, /*š*/ or /*č*/, /*dz*/ and /*tʃ*/ or /*tʃʹ*/ respectively.

Consonants are (with very few exceptions) not geminated, even when the syllabary allows that possibility. It has been suggested that Urartian lost its geminate consonants (which do exist in Hurrian) before it reached the state of the language preserved in the inscriptions.

Transliterations (marked by single bars) and transcriptions (marked by double bars) in this chapter use the conventional values for the transliteration of the cuneiform signs.

### 3.2 Vowels

The script seems to distinguish four vowel qualities: /*a*/, /*e*/, /*i*/, /*u*/. It is uncertain whether the interchangeable signs *u* and *ú* represent not only /*u*/ but also /*o*/. Vowel length seems to be indicated by *scriptio plena* (see §2.1), and presumably it was phonemic (see also §3.5.3; in the following morphemic transcriptions vowel length is not represented because of the high degree of graphemic variation). The opposition between /*e*/ and /*i*/ seems to be neutralized in final position (realized as [ə]), as can be seen (among other places) among variants using *be* and *bi* indiscriminately (*nu-na-bi*, *nu-na-be* “he came,” in morphemic transcription both rendered as *nun=a=bə*). Schwa may be represented by a plene-spelling of the vowel *e*, e.g., *pi-LI-e* for [pilə].

Spellings like *-ka-i*, *ba-ú-še*, *Te-i-še-ba*, *e-ú-ri*, *ú-ru-li-ia-ni*, *qi-ú-ra-i-e-di*, *ha-ú-li-i-e*, *si-lu-a-di*, *a-ú-e-ra-*, *šu-e*, *Iš-pu-i-ni* suggest the existence of the diphthongs /*ai*/, /*au*/, /*ei*/, /*eu*/, /*ia*/, /*ie*/, /*iə*/, /*ua*/, /*ue*/, /*uə*/, /*ui*/. It is not always clear, however, whether two adjoining graphemic vowels represent a monosyllabic diphthong; in some cases it can be shown by variant spellings that they do not: thus, *te-ra-a-i* with variant *te-ra-y[i<sub>x</sub>]*. Some of the diphthongs seem to be historical spellings, because there are variants with monophthongs: for example, *qi-(i)-ú-ra-*: *qi-ra-* “earth”; *al-su-i-ši-*: *al-su-ši-* “greatness”; *ka-i-ú-ke*: *ka-ú-ke* “before me.”

### 3.3 Phonotaxis

The writing system hides many consonant clusters. Any occurring in initial and final position could not be represented orthographically; it is likely, however, that they did not in fact exist. In medial position most consonantal clusters contain a non-stop as the first consonant: [-*ld*-], [-*lg*-], [-*lh*-], [-*lm*-], [-*ls*-], [-*lʃ/z*-]; [-*rb*-], [-*rd*-], [-*rg*-], [-*rḫ*-], [-*rm*-], [-*rn*-], [-*rq*-], [-*rʃ*-], [-*rš*-], [-*rt*-], [-*rz*-]; [-*mn*-]; [-*šd*-], [-*šg*-], [-*šḫ*-], [-*šm*-], [-*šp*-], [-*št*-], [-*šz*-]. There are also clusters with initial dental and bilabial stops: for example, [-*Tḫ*-], [-*Tg*-], [-*Tq*-]; [-*Ph*-], [-*Pq*-], [-*Pr*-], [-*Pš*-], [-*Pt*-]. There are no Urartian words with initial [r-]; the royal name *Rusa* probably was pronounced *Ursa*.

### 3.4 Accent

Uartian seems to have a stress accent on the penultimate syllable – at least in certain cases, defined by unknown conditions. It is evident that in many cases the final syllable is not stressed, as can be seen from the distribution of *Ce* (interpreted as /Cə/) symbols in word-final position versus *Ci(-i)* before suffixes: for example, *gu-nu-še* : *gu-nu-ši(-i)-ni(-)*; *pi-šu-(ú-)še* : *pi-šu-ši(-i)-ni(-)*; *šu-e* : *šu-i-ni-*.

### 3.5 Phonological processes

Several phonological processes can be identified for Uartian.

#### 3.5.1 Anaptyxis

Some of the attested cases of anaptyxis are most likely graphemic only (see §2.1): *ši-di-iš-tú-ni* : *ši-di-ši-tú-ni* (*šid=išt=u=nə*); *ta-ra-ma-na* : *ta-ar-ma-ni-li* (root: *tarm-*). Other cases appear to be genuinely linguistic; though in the absence of etymologies, they could be explained either by anaptyxis or by syncope: *ni-ir-bi* : *ni-ri-bi*; *zi-il-be*, *zi-il-bi(-i)* : *zi-li-bi(-i)*; *uldə* versus *ul-ú-de-e* “vineyard.”

#### 3.5.2 Syncope

1. The vowel of the plural suffix *-it-* and that of the root-complement *-id-* are lost after [r]: *ar=t=u=me* “they gave me”; *par=t=u* “they took away”; *ter=t=u* “they put up” (compare *kug/y=it=u=nə* “they dedicated”); *šer=d=u=l=(e)yə* (also *šer=id=u=l=(e)yə*) “who hides [it]”; *ar=d=i=l=anə* “he shall give”; *ter=d=i=l=anə* “he shall put up.” Generally the vowel is preserved elsewhere (though see below): *ab=il=id=u-*; *batq=id=u-*; *erš=id=u-*; *nips=id=u-*; *su=id=ul=u-*.
2. Stems ending in [d] followed by the plural suffix *-it-* drop the sequence [d=i]: for example, *za-tú-me* “they built me [a path],” from *zad-* “build” (if – contrary to §3.1 – Uartian had double consonants, the process would have to be described as syncope with assimilation). It is doubtful whether the same process occurs with stem-final [t]; the form cited in favor of this, *\*šid=išt=it=u=lə* → *šidištulə*, could be first-person singular *šid=išt=u=lə*.
3. After the sequence of {[IV], [rV] or [nV]} followed by the suffix *-ne-* or *-na-* (see §4.2.4), the vowel (V) is syncopated and the resulting consonant cluster undergoes progressive assimilation (in the case of [liquid + nasal]) and degemination. Thus, as Hurrian reveals, the diachronic process is as follows: *\*ebani=ne=lə* → *\*eban=ne=lə* → *ebanelə* “countries”; *\*ereli=ne=lə* → *\*erel=le=lə* → *erelelə* “kings”; *\*šeri=na=šə* → *\*šer=ra=šə* → *šerašə* “the other/previous (kings).” Note, however, that when the genitive-suffix /i/ intervenes, the changes do not occur: *ebani=i=na=we* dingir “to the gods of the country.”
4. A vowel is lost when occurring between the final [r] or [l] of a root and the ensuing modal suffix [l] (see §4.3.5.3), with assimilation and degemination as in 3 above: *\*tur=u=l=(e)yə* → *\*tul=l=(e)yə* → *tul(e)yə* “[who] might destroy.”

#### 3.5.3 Vowel contraction

Contraction is difficult to determine because of the ambiguities of the writing system and because of uncertainties surrounding the phonemic system. If the spelling variants with

diphthongs and monophthongs are correctly interpreted as revealing “historical spelling,” rather than accurate synchronic representations (see §3.2), then Proto-Urartian had a tendency to contract diphthongs. Such contraction even occurred across morpheme borders; for example, in the dative of *i*-stems. The well-attested traditional form is *LEXEME-i* = *ə* (graphemic *Ci-(i-)e*), as in *Hal-di-i-e* (dative: “to [the god] Haldi”); but there are also forms ending in *-Ci-i* which have been interpreted as a contraction of [i + ə]: for example, *e-si-i* (dative, “to the place,” from \**esi* = *ə*). The plene-writing of the vowel points to the product of contraction being a long vowel.

### 3.5.4 Assimilation

In addition to those cases of consonant assimilation noted in §3.5.2, **3** and **4**, assimilation of vowels occurs in two contexts: (i) the vowel of the verbal suffix *-Všt-* (see §4.3.2, **6**) assimilates to the preceding vowel (so-called vowel harmony); and (ii) the vowel of the plural suffix *-it-* (see §4.3.4.3) in some cases (e.g., after the root-complement *-id-*) assimilates to a following vowel [u]: \**še(i)r=id=it=u* → *še-i-ri-du-tú*; \**su=id=it=u* → *su-ú-i-du-tú*.

### 3.5.5 Metathesis

Metathesis is seen in *uldu* versus *udul-* “vineyard”; see also §4.2.2, **6**.

## 4. MORPHOLOGY

### 4.1 Word structure

The basic structural characteristics of Urartian seem to be in agreement with those of Hurrian (see Ch. 4, §§ 4.1–3); however, the available data in many cases are insufficient for a functional analysis of the sort possible for Hurrian. A set of root-complements would be expected to modify the semantics of the root, regardless of whether by suffixation, the root forms a noun or verb. Among the less than one hundred semantically more or less well-defined Urartian nouns, however, there seems to be no single one which can be shown to contain the same root-complement as a verb.

### 4.2 Nominal morphology

The noun (and a small number of nonderived adjectives) consists of (i) a root, (ii) a thematic vowel, and (iii) optionally a derivational suffix. Root-complements (see Ch. 4, §4.3) are seldom attested (for a possible exception see below, §4.2.1); and only few word-formation suffixes (see Ch. 4, §4.4.2) can be found within the limited body of material. As in Hurrian, there are two numbers (singular and plural), but no grammatical genders. To a noun (derived or not) relational, possessive, number, case, and congruence suffixes in a strictly sequential order may be added.

#### 4.2.1 Nominalization of the root

Roots may perhaps be nominalized by the suffixes *-i* and *-u*: *kapi* (a measure of capacity) is likely to be connected with the Hurrian root *kapp-* “fill” (cf. Hurrian *kapp=ar=ni* Ch. 4, §4.4.2, **13**). The forms *ḥa=ə* and *ašḥ=ašt=ə* (two terms for offerings, if correctly segmented) are based on *ḥa-* “take” and *ašḥ-* “sacrifice,” and *urb=u* “meat offering” on *urb-* “slaughter.”

## 4.2.2 Noun-formation and derivational suffixes

In Hurrian grammar two types of derivational formations have been distinguished: one utilizes suffixes (*word-formation suffixes*) which directly follow the root (and root-complements), and the other utilizes suffixes (*derivational suffixes*) which follow the so-called thematic vowel. In the latter case, the thematic vowel *-i* is replaced by the “derivational vowel” *-o-*. In this paragraph the two sets of suffixes are treated together in alphabetical order, as the Urartian patterns have not yet been fully examined and understood.

1. *-(a)d=ə* (see Ch. 4, §4.4.2, 1): *ḫu-ra-(a-)de* “warrior” (also Hurrian) might be a loan-word from Hurrian, perhaps via Assyrian.
2. *-ayə*, adjectives and nouns: *ṣi-ra-ba-e* “unirrigated”(?), *du-ru-ba-i-e* “hostile,” *tú-a-i-e*, *tú-a-y[i<sub>x</sub>]* “pure,” *tar=ayə* “mighty,” *al=ayə* “decision” (*al-* “speak,” see §4.3.4.4).
3. *-aurə*, patient-oriented participle (see Ch. 4, §4.4.2, 4): *ag=aurə* “something (a canal) which is conducted,” *ṣid=aurə* “something which is built”; (see also *-u(=)rə*).
4. *-(i)bə*: *atibə* “10,000,” *nir(i)bə* “wild sheep,” *ṭeribə* “?,” *zilibə* “seed,” “offspring.”
5. *-ḫə* (see Ch. 4, §4.4.6, 1): This suffix forms adjectives of appurtenance used with geographical or tribal names (*nisbe*): *Abiliane=ḫə ebanə* “the country of Abiliani” (tribal/personal name), *Diaue=ḫə* “the Diauean [king].” Without parallel in Hurrian is its usage in patronyms: *Argište=ḫə* “the son of Argišti,” *Iṣpuine=ḫə*, *Minua=ḫə*, *Rusa=ḫə*, *Sardure=ḫə*. It forms adjectives and nouns (i) after *u*: *egur=u=ḫə/ḫu* “clean, pure” (in a cultic sense), *tar-a-i-ú-ḫe* “?” (cf. *tarayə* “strong”); (ii) after *i* ( $\rightarrow e$ ): *qar-me-ḫe* “?,” *ter=i=ḫə* “plantation” (*ter-* “plant,” “establish”); and (iii) after *a*: *babanaḫə* (*babanə* “mountainous region”).

The ending *-ṣḫə* is presumably a suffix complex containing the abstract suffix *-ṣə* as in Hurrian (see Ch. 4, §4.4.6, 1G): *ḫuri=ṣḫə* “water supply”(?), *tui=ṣḫə* “clean place”(?), *uri=ṣḫə* (“weapon,” “piece of equipment”).

6. *-ḫalə*, *-lḫə* (metathesized variants), ethnic terms: *mišta=ḫal[ə]* “[the land] belonging to Mišta,” *meliṭi(y)alḫə* “the Meliṭian [king],” *puinialḫə* “the Puinian [king],” *iṣqugulḫə* “the Iṣqugulian,” *puluadiulḫə* “the Puluadian [king].”
7. *-i(=)ptə*, meaning unknown: *mer=i(=)ptə* “?,” *ušt=i=ptə* “campaign” (*ušt-* “go on a campaign”).
8. *-ka*, meaning unknown (see Ch. 4, §4.4.2, 10): *urb=i=ka=nə* “sacrificer”(?) (*urb-* “slaughter”).
9. *-lə*, nouns of profession (see Ch. 4, §4.4.2, 11): *erelə* “king,” *A.NIN=lə* “prince”(?).
10. *-nə*, basic meaning unknown (see Ch. 4, §4.4.2, 13): *ti=nə* “name” (*ti-* “speak”); additional nouns which have roots not attested in other usage, however, also end in *-nə*: *ebanə* “country,” *iaranə* (a sanctuary), *qarqaranə* “armor,” *sirḫanə* (a building). A functionally different suffix *-nə*, which perhaps is to be distinguished etymologically (see Ch. 4, §4.4.6, 2), seems to form adjectives: *quldi=nə* “uninhabited”(?), “vacant”(?). Several forms which have been claimed as adjectives, however, presumably are instrumentals: *piṣuši=nə* “pleasurable” *piṣuṣə* “pleasure,” *gunuši=nə* “by fight.”
11. *-ṣə*, abstract nouns (see Ch. 4, §4.4.2, 15A): *uṣma=ṣə* “might,” *ardi=ṣə* “order,” *arniu=ṣə* “deed,” “exploit,” *bau=ṣə* “order,” *gunu=ṣə* “fight,” *piṣu=ṣə* “joy,” *ṭelzu=ṣə* “(sacrificial) instruction,” *ulgu=ṣə* “life,” *alsui=ṣə* “greatness,” *iṣpui=ṣə* (positive abstract).
12. *-umə*, infinitive? (see Ch. 4, §4.4.2, 16): absolutive(?) *áṣ-ḫu-me* “offering”(?), directive *su-du-me-né-e-de* “?”.
13. *-tuḫə*: LUGÁL-*tú-ḫe* “kingship,” LÚ-(ú)-*tú-ḫe* “human beings,” *ir-nu-tú-ḫe-e* “?”.



14. *-u(=)rə*, “subject-oriented participle” with intransitive verbs (see Ch. 4, §4.4.2, 8): *ušt=u(=)rə* “someone who went out for a campaign,” *man=u(=)rə* “something which existed” (see also *-aurə*).
15. *-usə* (see Ch. 4, §4.4.6, 3): The Hurrian equivalent suggests identifying *u* as a suffix of derivation. In only few cases can it be shown that the suffix forms nouns of suitability as in Hurrian: *urišh=usə* “arsenal, treasury” (*urišhə* “weapon, piece of equipment”), *aših=usə* “building for cereals” (cf. Hurr. *až=o=ge* “meal”), *al=usə* “ruler,” *bad=usə* “perfection(?)”, *pul=usə* “stela,” *te/ir=usə* (measure of capacity).

#### 4.2.3 Thematic stems

(See Ch. 4, §4.4.3) All nouns end in a vowel. The most frequent vowel is *-i* or *-e*, but there is a good number of nouns in *-a* and in *-u*. No noun ends in a consonant, at least in writing (for the restrictions of the writing system see §2.1).

#### 4.2.4 Relational suffixes *-ne-* and *-na-*

(See Ch. 4, §4.4.7) Urartian *-ne-* (sg.) and *-na-* (pl.) are anaphoric suffixes. They precede case endings which mark agreement of genitive modifiers or modifiers in *-hə* or *-usə* (see §4.2.2, 5, 15) with their head noun (*Suffixaufnahme*, see §5.2): *Minua=i=ne=i sila=i* “of the daughter of Minua”; *Ḫaldi=i=n(e)=ə patari=ə* “for the city of Ḫaldi”; *Ḫaldi=i=ne=nəušmaši=nə* “by the might of Ḫaldi.” In addition to marking agreement with the head noun, *-na-* also functions generally as a plural marker, except in the absolutive case (*-na-* never occurs in the absolutive; see §4.2.6): *\*ereli=na=we* → *erel=la=we* “of kings”; *ḫuradi=na=we* “to the warriors”; *arniuši=na=ni* “by the deeds.” In the plural, the suffix of the absolutive plural is *-ne=lə*: *ḫuradi=ne=lə* “the warriors.”

#### 4.2.5 Possessive suffixes

Only two possessive suffixes are well attested. They take the position after the thematic vowels:

1. *First-person singular -ukə, -uka-* (without parallel in Hurrian): *e-ú-ri-u-ke* “to my lord”; *e-ba-ni-ú-ke-e-de* “into my country”; *e-ba-ni-ú-ka-né* “from my country.” The suffix also occurs with preposition: *ka-a-ú-ke* “in front of me”; for the suffix see also §4.2.7.4.
2. *Third-person singular -iye, -i(-), iya-* (as in Hurrian): *e-ba-ni-i-e* “his country”; *ti-i-né... ar-mu-zi-i... zi-il-bi-i* “his name, his family”(?), “his seed” (cf. Akkadian NUMUN-šū); *e-ú-ri-i-e* “to his lord” (cf. Akkadian *ana... EN-šū*); *ulguši=ya=nə edinə* “for his life.”

#### 4.2.6 Case and number suffixes

Urartian, as far as can be determined, is a strictly ergative language. The agent (subject) of a transitive verb appears in the ergative case; while the patient (object) of a transitive verb and the agent of an intransitive take the absolutive case. There are no hints that there exists a pattern as in Hurrian which encodes agent and patient as absolutive and essive respectively (see Ch. 4, §4.4.9). A special pattern for an action with a virtual but not explicit patient may exist, but cannot be proven.

Nine cases have been identified thus far. The principal differences vis-à-vis Hurrian (see Ch. 4, §4.4.9) are as follows: (i) the absolutive plural utilizes the suffix *-ne-* (see §4.2.4)



which in Hurrian is confined to the singular; (ii) the genitive and dative suffixes have a labial continuant only in the plural; (iii) the comitative is marked by a complex suffix (Hurrian *-ra*); (iv) the dative also has the function of the Hurrian directive in *-e*; (v) the Uartian directive is perhaps a complex suffix formed from *-e-* (cf. the Hurrian directive) plus the directive suffix *-də*, the product of Proto-Urarto-Hurrian *\*-da*, preserved in Hurrian and, as archaism, in Uartian; (vi) the Hurrian plural marker *-až-* is unknown to Uartian except for a few archaic forms of the directive and ablative plural.

(2)	<i>Singular</i>	<i>Plural</i>
<i>Absolutive</i>	—	-ne=lə
<i>Ergative</i>	-š, -šə	-na=šə
<i>Genitive</i>	-i	-na=wə
<i>Dative</i>	-ə(ø for a-stems)	-na=wə
<i>Directive</i>	-edə	-na=(e/i)də
<i>archaic</i>	-da	-na=aš=tə
<i>Comitative</i>	-ranə	-na=ranə
<i>Ablative-instrumental</i>	-nə	-na=nə
<i>Ablative</i>	-danə	-na=aš=tanə
<i>Locative</i>	-a	-n(a)=a

#### 4.2.7 Pronouns

Uartian is characterized by each of the following pronominal forms.

##### 4.2.7.1 Possessive pronouns

For the possessive suffixes within the suffix chain of the noun see §4.2.5.

##### 4.2.7.2 Pronominal suffixes

The enclitic personal pronouns of Hurrian (see Ch. 4, §4.4.10.1) are only partially attested in Uartian, and they differ in function and distribution. As in Hurrian, they only refer to the participant in the absolutive case. Contrary to Hurrian, in Uartian they cannot be repeated several times within a clause, they cannot be suffixed freely to various words within the clause, and their position is strictly determined.

The suffix *-də* of the first-person singular corresponds to Hurrian *-tta/-d*, but it only occurs as a suffix of intransitive verbal forms of the first singular (see §4.3.4.1). The suffix of the third singular, *-nə*, corresponds to Hurrian *-nna/-n*. In most cases, it is associated with an absolutive singular serving as the subject of an intransitive verb, but sometimes also with the object of a transitive verb. It can also follow an interrogative/relative pronoun (*alu=š=nə*, see §4.2.7.5). The suffix of the third plural, *-lə*, corresponds to Hurrian *-lla/-l*, but its occurrence is restricted to nouns in the absolutive plural. For *-nə* and *-lə* as verbal suffixes see §4.3.4.1 and §4.3.4.3.

A personal suffix without a morphological or functional equivalent in Hurrian is the dative suffix of the first-person singular, *-mə*: *haš=i=a=l=mə* “they granted to me” (intr.); *ar=u=mə* “he gave me”; *Ḫaldi=š=mə* “(god) Ḫaldi (erg.) . . . me”; *a-šú-me* “when . . . me”; *alu=š=mə* “who . . . me.”

##### 4.2.7.3 Independent personal pronouns

Only the following forms have been identified:

1. *First-person singular*: (i) absolutive subject of an intransitive verb: *ištidə*; (ii) absolutive object of a transitive verb: *šukə*; (iii) ergative: *iešə*; (iv) possessive adjective: *šusə* “my” (with suffix *-usə*, see §4.2.2, 15); (v) locative plural: *šusi=na=a*.
2. *Third-person singular*: (i) absolutive: *manə*; with pronominal suffix *-nə* (see §4.2.7.2): *mani=nə*; (ii) possessive adjective: *masə*; plural *masi=ne=lə* “his.”

#### 4.2.7.4 Deictic pronouns

The two most important deictic pronouns are as follows: (i) the demonstrative pronoun *i(nə)-* refers to the object which bears the inscription or which is close to it. In an often attested curse formula it is used anaphorically for actions mentioned immediately before; (ii) *ina-* seems to be restricted to an anaphoric function, possibly including a sense of distance.

		Singular	Plural
Demonstrative	absolutive	inə	i=ne=lə
	ablative-instrumental		i=na=nə (i-na-(a-)né)
	locative		i=na=a (i-na-a)
Anaphoric	absolutive	ina=nə	ina=ne=lə
	ablative-instrumental		ina=na=nə (i-na-na-né)

In addition, *ina-* serves as the base for other pronouns: (iii) *ina=hə* “such” (dative pl.: *ina=hə=na=wə* “for such / the aforementioned [towns]”); (iv) *in=ukə* “exactly this” (emphatic, identifying; absolutive sg.: *in=ukə* (i-nu-ke(-e)); ablative-instrumental: *in=uka=nə* (i-nu-ka-(a-)né); for the suffix cf. §4.2.5, 1); (v) *in=uka=hə* (a hapax legomenon, genitive adjective: *i-nu-ka-hə-né-e*); (vi) *in=usə* “the said,” “the aforementioned” (absolutive sg.: *in=usə*; ablative-instrumental(?): *in=usi=nə* (i-nu-si-i-né); for the suffix cf. §4.2.2, 15); (vii) *ik=ukə* “the same” (attested only in the ablative-instrumental: *ik=uka=nə* “the same [year/day/road]”; for the suffix cf. §4.2.5, 1; the root might be connected with Hurrian postposition *egi* “in”; see Ch. 4, §5.1).

#### 4.2.7.5 Relative pronoun

The relative pronoun is attested in the forms *alə* (absolutive sg.), *\*ali=ne=lə* → *alelə* (absolutive pl.), and *alu=šə* (ergative). For the vowel shift *i/u* compare the Hurrian pronominal pattern (Ch. 4, §4.4.10.2–3.). The indefinite pronoun *ali=kə* “some” seems to be based on the relative pronoun *alə*; it is, however, indeclinable.

### 4.3 Verbal morphology

Our knowledge of the Uartian verb is particularly limited. In the stereotypical royal inscriptions, the indicative verbs (with one exception) only describe past actions in the first-person singular, and in the third-person singular and plural.

By way of a summary comparison with Hurrian verbal morphology, the following remarks are offered (see below for specific discussion of the Uartian verb). As in Hurrian, verbs may be marked for modes of action; but in comparison with Hurrian, the preserved verbal forms show far fewer root-complements (regardless of what their function may be). Most of the Hurrian root-complements which modify the meaning of the root prior to the distinction of nominal or verbal inflection are not attested at all in Uartian. Also, as in Hurrian, the valence of a verb (see §4.3.1) is indicated by the so-called class-markers (see §4.3.3); and valence may be changed by changing the class-marker. It has not yet been convincingly shown that Uartian morphologically distinguishes verbal aspects (see §4.3.2, 3) or tenses. A verb is not negated by a suffix as in Hurrian, but by a particle which precedes it (see

§4.4.3). The subject of an intransitive verb is marked by enclitic personal pronouns which are – unlike the Hurrian condition – a constituent part of the verbal form. The transitive-ergative verb has suffixes which mark the patient. There are, however, two different markers of the third-person singular patient, which are distributed according to agents. Apart from the pronominal suffixes which are etymologically identical with the Hurrian enclitic personal pronouns of the absolutive (see Ch. 4, §4.4.10.1), no person suffixes have been observed. It is, however, possible that the vowel *a* in the form  $R=u=a=l\partial$  (see §4.3.4.3) is a marker of the third singular agent. The Hurrian plural suffix *-až* has no counterpart in Urartian verbal inflection (as far as it is known). As in Hurrian, a wide variety of nonindicative moods occur.

### 4.3.1 Valence

Valence (the number of noun phrases governed by the verb) is indicated by the vowels treated in §4.3.3. As in Hurrian (see Ch. 4, §4.5.1), some roots are attested in both transitive and intransitive use: *aš-* “enter” (intr.) / “put in”; *kuṭe-* “advance” / *kuṭ-* “send, forward, extend, conquer”; *naḥ-* “sit down” / “carry away”; *ši-* “come” / “bring”; *ušt-* “go on a campaign” / “offer,” “present”. Normally, however, the root is attested in either transitive or intransitive usage. As in Hurrian, a change of valence can be marked by the suffix *-ul-*: *aš=u=bə* “I put in [a garrison],” *aš=ul=a=bə* “[when the country] was occupied,” *aš=ul=a=l[ə]* “[the palaces] were occupied.”

### 4.3.2 Undefined verbal suffixes

There are several verbal suffixes immediately following the root, which are morphologically identical with Hurrian suffixes. The scarcity of varying contexts, however, makes it impossible to prove functional identity.

1. *-an-* (cf. Ch. 4, §4.3 [2D]): *ašt=an=ul-* “?”, *ked=an-* “send,” *uṣḥ=an-* “grant.”
2. *-ar-* (cf. Ch. 4, §4.3 [2I]): *qapq=ar=ul-* “besiege”; compare also *ub=ar(=)d=ud-*, *ṭub=ar(=)d-*.
3. *-id-*, *-ud-*, *-d-* (on the background of Hurrian, see Ch. 4, §4.5.4), *-id-* has been interpreted as a marker of aspect, but there is little Urartian evidence for this or any other interpretation): *ab=il=id-* “rank among,” *batq=id-* “restore”(?), *e/irṣ=id-* “settle,” *iz=id-* “order,” *ne/ik=id-* “?”, *nips=id-* “sacrifice (an animal in a specific way),” *su=id-* “force back” (written with *i*, never *e*), *šer=(i)d-* “?”, *wel=id-* “gather”; *kul=ud-*, *lak=ud-* “?” (both a damaging action); *ṭub=ar(=)d-* “?”, *ub=ar(=)d=ud-* “order.” A form *-ad-* occurs in *atq=an=ad-* “celebrate.”
4. *-il-* (cf. Ch. 4, §4.5.2, 1): *ab=il=id-* “rank among.”
5. *-ul-* (cf. Ch. 4, §4.5.2, 2): *ašt=an=ul-* “?”, *qapq=ar=ul-* “besiege,” *su=id=ul-* “defeat.”
6. *-Všt-* (cf. Ch. 4, §4.5.2, 2): *am=ašt-* “burn down” (tr.), *aṣḥ=ašt-* “offer [an offering],” *a(y)=išt-* “jump,” *šid=išt-* “build,” “erect,” *nul=ušt-* “devastate”(?), *sul=ušt-* “prostrate,” *ul=ušt-* “march (ahead)”(?).

### 4.3.3 “Class-markers” (suffixes of valence)

As in Hurrian (see Ch. 4, §4.5.6), the position following the root and the optional root-complements is occupied by a vowel which is called the “class-marker.” In most cases it is either *-a-* or *-u-*: *-a-* indicates single valence and intransitivity, *-u-* (the equivalent to Hurrian *-o-*) two valences and ergativity.

Some intransitive verbs have a vowel *-e-* or *-i-* of unknown function before the class-marker: *bid=i=a-* “return” (intr.), *ḫut=i=a-* “pray,” *kuṭ=e=a-* “advance” (but tr. *kuṭ-* “send,” etc.).

A few intransitive verbs have a class-marker *-i-*: *sul=ušt=i=bə* “I prostrated myself,” *a(y)=išt=i=bə* “[the horse] jumped.”

Some deviating forms (all with a consonant cluster before the class-marker) are not yet well understood: *uḫ=u=də* “I ordered” (intr., hapax), *ul=ušt=ai=bə* “he marched [ahead]”(?) (besides regular *ul=ušt=a=bə*).

#### 4.3.4 Person suffixes

The person suffixes of the verb follow the class-marker. Only the first singular, and the third singular and plural are well attested. Other forms are either not attested or questionable (e.g., *šid=išt=u=šə* “we built it”[?]).

##### 4.3.4.1 Intransitive verbs

The intransitive verb is conjugated by means of pronominal suffixes (see §4.2.7.2) which correspond to the absolutive enclitic personal pronouns first singular and third singular of Hurrian (see Ch. 4, §4.4.10.1). The third singular is formed by a suffix *-bə* equivalent to the suffix *-b* of Old Hurrian. Only the following forms are attested:

	Singular	Plural
(4)		
1st	nun=a=də “I came”	
3rd	nun=a=bə “he came”	nun=a=lə “they came”

##### 4.3.4.2 Stative verbs

The stative verb *man-* shows a class-marker *-u-* which formally is identical with the transitive class-marker (for the parallel in Hurrian see Ch. 4, §4.5.11). Different from the intransitive verb (and in agreement with Hurrian) the third-person singular is not marked by a verbal person suffix: *man=u=də* “I stayed,” *man=u* “it was” (often with adjective in *-ayə* [see §4.2.2, 2]), *man=u=lə* “they were,” *ali... man=u=lə* [-L1-e] “who may exist,” *ali=lə... man=u=lə* “who (pl.) may be there,” *man=u=l=i=nə* “may it exist!” See also §4.2.2, 14 and §4.3.5.1.

##### 4.3.4.3 Transitive verbs

Of the person suffixes used with transitive verbs, only a subset is attested; agent (ergative) and patient (absolutive) suffixes are shown in (5). The etymology of the suffix of the first-person singular is still controversial.

	Abs. 3rd sg.	Abs. 3rd pl.
(5)		
Erg. 1st sg.	R=u=bə “I... him”	R=u=bə / R=u=lə “I... them”
Erg. 3rd sg.	R=u=nə “he... him”	R=u=a=lə “he... them”
Erg. 3rd pl.	R=it=u=nə “they... him”	R=it=u=lə “they... them”

The first-person singular dative suffix *-mə* (see §4.2.7.2) displaces the absolutive suffix *-nə*: *ar=u=mə* “he gave [it] to me,” *\*zad=it=u=mə* → *za=t=u=mə* “they built me [the road].”

The comparison with the endings of the intransitive verb shows that both paradigms make use of the same pronominal suffixes: *-bə*, *-nə*, and *-lə*. There are no special suffixes for the person of the agent except the suffix *-it-* which marks plurality of the agent (for a possible exception see §4.3). The difference between the first- and the third-person agent is encoded by the use of two different suffixes for the patient: the first person of the agent is marked

by the suffix *-bə* and the third by *-nə*. Both suffixes refer to the third person of the patient, but *-bə* – exactly as with Hurrian *-b* (see Ch. 4, §4.5.9) – is not restricted to the singular. Referring to a plural patient, *-bə* may be replaced by the pluralic *-lə*. The occurrence of the suffixes *-bə*, *-nə*, and *-lə* in both the transitive and the intransitive paradigms is related to the ergative structure of Urartian: both the patient of the transitive verb and the subject of the intransitive verb are encoded as absolutes.

#### 4.3.4.4 The verb *al-*

This verb occurs in one form only: *alə* (*a-LI*, *a-LI-e*), always with a noun in the ergative, but without an absolute. It introduces direct speech in royal inscriptions and letters (e.g., LUGAL=šəalə “[thus] says the king”). It has been suggested that the direct speech as a whole is the patient. The verbal status of Urartian *alə* has long been disputed, but it is confirmed by the Hurrian form *a-lu-i-ib* “he said.”

### 4.3.5 Nonindicative moods

As in Hurrian (see Ch. 4, §4.5.12), there is a considerable variety of nonindicative modal forms. Disagreement still exists over terminology, morphology, and relationship with Hurrian modal suffixes. Only the best-established patterns are mentioned here.

#### 4.3.5.1 Imperative

The second singular imperative, both intransitive and transitive, is formed by the root plus the suffix *-ə* (seldom *-i*): *ul=i* (*ú-LI-i*) “go!”; *ar=ə* (*a-ri*) “give!”; *šat=ə* (*šá-te-e*) “take!”; *ti=ə* (*ti-(i)-e*) “speak!”; *tur=ə* (*tú-RI(-e)*) “destroy!” Presumably the form *ma-nu-ni* “may he be/exist!” is to be interpreted as a third-person singular imperative (for the verb *man-* see §4.3.4.2). Perhaps the vowel *u* conveys a concept of state like the Hurrian imperative in *-o(=mma)* (see Ch. 4, §4.5.12.1); for the form *ma-ni-né* see §4.3.5.2.

#### 4.3.5.2 Jussive

The jussive is a request in the third person, marked by a suffix *-in*. This suffix corresponds to the jussive suffix of Hurrian *-en*, which in the earliest Hurrian has also the form *-in* (see Ch. 4, §4.5.12.2). Transitive verbs add a pronominal suffix after an anaptyctic vowel *i*. The plurality of the agent is marked by *-it-*. Examples follow: (i) *tur=in=i=nə* (in most cases spelled *tú-ri-ni-né*, once *-ni-i-né*) “they (once ‘he’) may destroy him” (usually the plural is not marked; the form *[tu]r=ut=in=e=nə*, which is attested in the same context, might be the expected plural form with *it* → *ut*); (ii) *ar=in=i=nə* (*SUM-ni-né*) “he shall give [a cow]”; (iii) plural: *ha=it=in(ə)* (*ha-i-ti-né*) “they shall take”; (iv) *ašh=ašt=it=in(ə)* (*áš-ha-áš-ti-ti-né*) “they shall give as an offering.” The form *ma-ni-né* “may he be/exist” is presumably an intransitive jussive of the verb *man-*: *man=in(ə)*; see also §4.3.5.1.

#### 4.3.5.3 Modal *-l-*

The modal suffix *-l-* combines with various forms of mood and modifies their meaning in a way which, however, cannot always be well established; for the exact equivalent in Hurrian see Ch. 4, §4.5.12.3.

#### 4.3.5.4 Optative

The optative expresses a wish or a demand. It is formed with the modal suffix *-l-* and a suffix *-ə*: These suffixes are usually spelled as *-LI* or, seldom, as *-LI-e*, but never as *-LI-i-e* (as is sometimes the spelling with the conditional; see §4.3.5.5). Most forms are third person,

but there is also at least one form of the first-person singular. The optative may have a pronominal suffix referring to the patient:

1. *First-person singular*:  $qapq=ar=u=l=i=nə$  ( $qa-ap-qa-ru-li-né$ ) “I wanted to besiege [the city].”
2. *Third-person singular*:  $urp=u=l=ə$  /  $urp=u=l=i=nə$  ( $ur-pu-ú-li-i-né$ ) “he shall slaughter [them]”;  $nips=id=u=l=ə$  /  $nips=id=u=l=i=nə$  “he shall sacrifice [in a specific way]”;  $aš-hu-li-né$  “they shall give as an offering”;  $me-ši-i-ú-li-né$  “they shall gather(?) [the grapes].” It is not yet clear whether the first of two verbs in the standard curse formula belongs here:  $ú-ru-LI-a-né$  . . .  $ú-lu-LI-e$  “may they . . . him [and] destroy [him].” The Urartian optative is presumed to be etymologically connected with the Hurrian form (see Ch. 4, §4.5.12.4).

The optative regularly appears in clauses introduced by  $ašə$ . This word is conventionally translated as the temporal conjunction “when(ever)”; it may, however, contain the ergative suffix  $-šə$  and thus fill the position of the agent in an agent–patient construction, which otherwise would be vacant. If this interpretation is correct,  $ašə$  would encode the concept of an indefinite agent (“when one/they”). Otherwise, the forms in  $ašə$ -clauses would have to be explained as being intransitivized by the suffix  $-ul-$  (see §4.3.1), which, however, is not likely:  $ašə$  . . .  $ašh=ašt=u=l=ə$  ( $áš-ha-áš-tú-LI$ ) “when they make an offering” (in a strict sense: “when they are to make an offering?”);  $ašə$  . . .  $nek=id=u=l=ə$  ( $ni-ki-du-LI$ ) “when they . . . (the canal)”;  $ašə$  . . .  $teš=u=l=ə$  ( $te-šú-LI-e$ ) “when they harvest the vineyard.”

#### 4.3.5.5 Conditional

The conditional is a form with the modal suffix  $-l-$ , which regularly appears in relative clauses of the curse formula introduced by  $alu=šə$ . The verbal form ends in  $-LI-e$  or  $-LI-i-e$  ( $LI$  represents [li] and [le]), which we normalize here as  $-l-(e)yə$ :  $alu=šə inə$  DUB- $te tu=l=(e)yə$  (\* $tur=u=l=(e)yə$ ) “who(ever) might destroy this inscription”;  $a-lu-šə pi-i-tú-LI-i-e$  “who(ever) might smash [it to pieces].”

A good morphological comparison is the Hurrian conditional in  $-eva$  (see Ch. 4, §4.5.12.6).

#### 4.3.5.6 Desiderative

The desiderative expresses a strong wish. In the context of the Urartian annals this may be the wish of the royal author (“he shall . . .”) or the reported wish of an enemy (“I heard that he intended to . . .”). The desiderative is formed by the modal  $-l$  suffix preceded by a “class-marker”  $i$  (cf. Ch. 4, §4.5.6) and a suffix  $-anə$ . The same formation is attested in Hurrian (see Ch. 4, §4.5.12.7). Examples follow:  $ar=d=i=l=anə$  ( $ar-di-la-né$ ) “he shall give [tribute]”;  $ha-i-la-a-né$  “[I heard that the country . . .] intended to conquer [the city . . .]”;  $ir-bi-la-[né]$  “[I heard that the country . . .] intended to raid [the . . .].”

#### 4.3.5.7 Additional moods

There are more modal forms which, however, are either poorly attested or semantically difficult:

1. Formed with a complex suffix: (i)  $-alanə$ :  $ha-ia-la-a-né$  “[never had kings] conquered” (with a glide  $y$  at the morpheme border),  $pa-a-ra-la-né$  “[to which never a king] had brought”; (ii)  $-ulanə$ :  $qu-du-la-a-né$  “?” (cf.  $ú-ru-LI-a-né$  at §4.3.5.4).
2. The following (dialectal?) forms are only attested in one religious text from the early period (meher kapısı):  $ni-ip-si-di-a-a-le$  “they shall sacrifice them [in a special way]”;  $qa-ap-qa-ri-li-né$  “he shall carry around(?)”,  $urp=u=ə$  ( $ur-pu-ú-e$ ) “he shall(?) slaughter.”

3. A letter from Bastam has forms which seem to be jussives (first- or third-person singular?): *a-li-le* (*al-* “speak”), *ar-di-le* (*ar-* “give”).
4. The form *a-ri-a-ni* has been interpreted as nonindicative, though it could be an indicative (*ar=i=a=nə* “he does not give her back”[?]) with a perfect parallel in Hurrian (cf. Ch. 4, §4.5.8.10).
5. *mi...kul=i=u=nə* “they shall not let him [exist]” (cf. Hurrian *ko<sup>2</sup>l-* “let”) may be a vetitive formed by the negative particle *mi* and the indicative.

## 4.4 Particles

“Particle” will here be defined as a word which cannot take nominal or verbal suffixes.

### 4.4.1 Conjunctions

The following conjunctions are identified: *ašə* “when(ever)” see §4.3.5.4; *iu* “when” (temporal clause referring to past); *e’ə* (also written *e-ú-e*, *e-a-i*) “and (also),” *e’ə...e’ə* “as well as”; *mi* “but”; *mi...mi* “neither...nor”; *unə* “or.”

### 4.4.2 Adverbs

The following adverbial particles are identified: *ainey* “anyone,” *gey* “anything,” *henə* “now,” *ištinə* “there,” *ištini=nə* “from there.”

### 4.4.3 Negative particles

Negation is accomplished by the particles *ui* “not” and *mi*, *mi=kui* (prohibitive).

## 4.5 Numerals

The numerals are almost exclusively written with numerical symbols, rather than being spelled out phonetically. In an annalistic text, the expression meaning “in one year” alternates between *šusini* MU and 1 MU. There is, however, a plural *šusina* MU<sup>mes</sup> which is translated into Assyrian by *ina libbi šanāteya* “in my years” (cf. §4.2.7.3). The cardinal 10,000 is *atibi*.

## 5. SYNTAX

### 5.1 Word order

As in Hurrian, the agent in ergative clauses (see §4.2.6) usually takes the initial position, followed by the patient and the verb (SOV), but the sequence *absolute–ergative–verb* (OSV) also occurs. The dative or instrumental of a god’s name regularly precedes the ergative in the first clause of a text or a paragraph:

- (6) A. *Ḫaldi=ə...M.=šə...ini pulusə kuḡ/y=u=nə*  
*Ḫaldi=ə...ini pulusə M.=šə...kuḡ/y=u=nə*  
 “To Ḫaldi M. set up this stela”
- B. *Ḫaldi=i=ne=nə ušmaši=nə M.=šə I.=šə inilə tarma=ni=lə*  
*aṭṭ=u=a=lə*  
 “By the might of Ḫaldi M. [and] I. dug this well”



Otherwise the dative may follow the verb:

- (7) É.GAL šid=išt=u=nə badusi=yə  
“He built a palace up to its perfection(?)”

The verb may be placed in initial position when it is topicalized (particularly in the more vivid inscriptions about military campaigns, often forming chiasms with regular clauses):

- (8) ušt=a=də Mana=idə ebanə at=u=bə  
“Forth I marched towards Mana, and I consumed the land”

In nominal clauses of two absolutes the predicate noun takes the final position:

- (9) Minua=nə . . . LUGÁL tarayə  
“M. is the great king”

In a nominal clause indicating a possessive relation with a genitive, the latter takes the initial position:

- (10) M.=i=ne=i sila=yə Tariria=i inə uldə  
“This vineyard belongs to T., the wife of M.”

Generally, the genitive may precede or follow its head; in names it regularly takes the initial position: for example, *Minua=i patarə* “Minua-town”; *Rusa=i šuə* “Rusa-reservoir.” In titles the genitive follows its head, perhaps following the Akkadian model: LUGÁL *Bia=i=na=wə* “king of the [people] of Bia.”

Other modifiers usually follow their head, whereas the deictic pronouns precede it: LUGÁL *tarayə* “great king”; *ina=ne=lə arniušī=ne=lə* “these exploits.”

Uartian has a considerable number of postpositions, which are partially based on the same forms as the postpositions of Hurrian (especially *edi* “person, body,” see Ch. 4, §5.1). Most Uartian postpositions have a suffix *-nə* which is likely to be the ablative-instrumental suffix; there is, however, one postposition which is of locative origin (*ed(i)=i=a*). In one instance (*ištini=yə*) the spelling suggests the presence of a third-person singular possessive suffix, as with most Hurrian postpositions. It is quite possible in fact that the majority of Uartian postpositions were formed with this suffix (hence the transcription *-(i)=i*). The noun governed by the postposition always takes a case ending (*apt(i)=i=nə* “on the side of,” with abl. *-danə* or abl.-instr. *-nə*; *bed(i)=i=nə* “from the side of, on the part of,” with abl.-instr. or loc.; *ed(i)=i=nə* “for,” with abl.-instr.; *ed(i)=i=a* “to(wards),” with archaic gen./dat.; *ištini=yə* (spelled *-ni-e*, *-ni-i-e*), *ištin(i)=i=nə* “for,” with loc.; *(-)kai*, seldom *(-)ka* “before, in front of” – with dat. (persons), loc. or abl.-instr. (places, objects), *kai* can take the possessive suffix *-ukə* (*kai=ukə* “before me”); *(-)kai=nə* “from (before),” with abl.-instr. or dat.(?); *(-)pei* “under,” with dat. or abl.-instr.; *(-)pe(i)=nə* “from under,” with dat.(?); *(-)šə* “(with)in,” “in the middle of,” with loc.).

Whether or not a postposition is enclitic cannot be determined in most cases, as the inscriptions do not separate words, and the evidence of the letters which utilize a word-dividing sign is insufficient in most instances. The letters seem to confirm, however, that *-ka(i)* is enclitic, and this may be true for *(-)pei(=nə)* and *(-)šə* as well. Even so, the enclitic postpositions clearly have not evolved into true case endings since they are not subject to *Suffixaufnahme* (see §5.2).

A single preposition, *parə* (“to(wards),” “unto,” usually with dat., sometimes with loc.) has been identified thus far.



## 5.2 Agreement

As in Hurrian (see Ch. 4, §5.2), a modifier (genitive modifiers and modifiers in *-hə* or *-usə*; see §4.2.2, 5, 15) agrees with its head. The case endings copied from the head are preceded by the relational suffix *-ne-* or *-na-* (see §4.2.4):

- (11) A. Minua=i=ne=i urišhusi=ne=i  
       “Of the storehouse of Minua” (object)  
       B. Ḫaldi=i=na=wə šešti=na=wə  
       “For the gates of Ḫaldi”  
       C. <sup>lú</sup>AD=si=n(e)=i esi=i  
       “On the paternal throne”

## 5.3 Coordinate and subordinate clauses

Coordinate clauses without a connective form the majority of Urartian texts. There is no pattern of nominalized verbs or verbal nouns as in Hurrian (see Ch. 4, §5.3). Subordinate clauses express a relational or a temporal connection with the main clause. The temporal clause introduced by *ašə* (see §4.3.5.4) in all attested cases displays a special modal form (optative), which, however, seems to express a wish or intention, not a special form of subordination. The relative clause may use the conditional (see §4.3.5.5); this is always the case in curse formulae which express a potential action. When the action is considered a fact, the relative clause uses the indicative:

- (12) alə ab=a=də ḫaš=i(y)=a=l=mə DINGIR<sup>meš</sup>  
       “What I requested, the gods granted to me”

Temporal clauses referring to the past are introduced by *iu* “when”; they always take the indicative:

- (13) iu Ḫaldi=š=mə LUGÁL-tuḫə ar=u=nə naḫ=a=də <sup>lú</sup>AD=si=n(e)=i esi=i  
       LUGÁL-tuḫe=i=ne=i  
       “When Ḫaldi gave me kingship, I sat down on the paternal throne of kingship”

## 6. LEXICON

The Urartian lexicon is even less well known than that of Hurrian. The meaning of less than three hundred words has been established, with varying degrees of exactness. For the less than one hundred roots used in verbal forms, approximately 20 percent are also known in Hurrian. This figure obscures the actual close proximity of the two languages: a considerable part of the Urartian corpus is formed by accounts of military campaigns, a genre absent in Hurrian literature; whereas the majority of Hurrian linguistic material is either related to religious ritual or to diplomacy, which are only poorly reflected in the Urartian corpus.

Aside from the basic phonological differences between Urartian and Hurrian (lack of double consonants in Urartian, lack of phonemic voicing in Hurrian), and the open questions concerning vowel length and opposition of /o/ and /u/ in Urartian, the following roots of Urartian and Hurrian (with one exception: *nun-*) are in total phonological agreement. This is basically true also for the nominal isoglosses below (note, however, differences under

*iština*, *šalə*, *šurə*). Putative isoglosses which show greater diversity have been demonstrated to be wrong or remain doubtful.

1. *Roots used in verbal forms*: *ag-* “guide” (H. “take up”); *al-* (H. *ale-*) “speak”; *am-* “burn”; *ar-* “give”; *ašh-* “make an offering, sacrifice”; *durb-* “become hostile” (H. only in noun *torbi*, *torubi* “enemy”); *ha-* “take”; *haš-* “hear” (H. *haž-*); *hut=i(y)-* “pray” (H. *hud-* “raise”); *kul-* “let” (H. *ko<sup>2</sup>l-*); *man-* “be” (H. *mann-*); *naḥ-* (H. *naḥh-*) “sit down”; *nun-* “come” (H. *un-*); *piš-* \* “rejoice” (only in the noun *pišušə* “joy”; H. *pic-* (written with sV symbols)); *šat-* “take” (H. *šatt-*); *tan-* “do,” “make”; *ti-* “speak” (H. *tive*, *tia*, *tieni* “word”); *urb-* “slaughter”; *ušt-* “to leave for a campaign” (H. *ušt=a=nni* “warrior,” “hero”).
2. *Nouns*: *ate-* “father” (H. *atta=i*); *babanə* “mountainous region” (H. *p/faba*, *p/fabni*, *p/fabanni* “mountain”); *edi-* (see §5.1; H. *edi* “person, body”; see Ch. 4, §5.1); *eurə* “lord” (H. *evri*); *evrišə* “lordship” (H. *evrišše* “lordship”); *harə* “road” (H. *hari*); *huradə* “warrior” (H. *huradi* “[a kind of] warrior”); *išanə* “opposite bank, lakeside” (H. *e/išave* “opposite bank”); *iština* (see §5.1; H. *ištani* “inside, middle”); *kurə* (ukrə?) “foot” (H. *ugri* “foot”); *pilə* “canal” (H. *pilli/a*); *p/bura* “slave” (H. *pura=me*); *qarqaranə* “coat of mail” (H. *kargarni*, a piece of military equipment); *šali* “year” (H. *šawala*); *šəhirə* “living” (H. *še/ugurni* “life”; presumably identical with the archaic element of a personal name *šəgirnī*); *šuḥə* “new” (H. *šuge* “new”); *šurə* “weapon” (H. *šauri*); *tarmanilə* (pl.) “spring,” “well” (H. *tarmani* “spring,” “well”; *tarm-* “drink”); *taršuanī* “man” (H. *taržu(w)ani*); <sup>uzu</sup> *tišnu* (also *tišni*?) a part of the body (H. *tižni*, *tiža* “heart”); *ulə* “another” (H. *oli*).

It cannot be demonstrated that all of these isoglosses were inherited from the proto-language ancestral to Hurrian and Urartian. It is possible that some words (especially military vocabulary) were borrowed from Hurrian into Proto-Urartian in the middle of the second millennium BC.

In several cases, it can be shown that Urartian and Hurrian use different lexemes which apparently only exist in one of the two languages: thus, for “build,” “erect [a building]” Urartian consistently uses *šid=išt-*, whereas Hurrian uses *pa-*.

One word has been claimed as a loan from Akkadian – *kubušə* “helmet.” Even this, however, is questionable, since Akkadian *kubšu* is not a piece of military equipment but a headdress or cap, often made of wool and used by gods, kings, and high officials.

There are no secure examples of borrowings from other languages.

## Bibliography

There is no Urartian grammar which presents the present state of knowledge of the language. The following monographs are in need of correction and supplementation in the light of the research literature of the last decades:

- Benedict, W. 1958. “Urartian phonology and morphology.” Dissertation, University of Michigan.  
 Diakonoff, I. 1971. *Hurrisch und Urartäisch*. Munich: Kitzinger.  
 Friedrich, J. 1933. *Einführung ins Urartäische*. Leipzig: Hinrichs.  
 ———. 1969. “Urartäisch.” In *Altkeinasatische Sprachen*, pp. 31–53. Handbuch der Orientalistik. Leiden/Cologne: Brill.  
 Khačikjan, M. 1985. *Churritskij i urartskij jazyki*. Jerevan: AN Armanokj SSR.  
 Melikišvili, G. 1971. *Die urartäische Sprache*. Rome: Biblical Institute Press.  
 Meščaninov, I. 1958, 1962. *Grammatičeskij stroj urartskogo jazyka*, vols. I–II. Moscow/Leningrad: Academy of Sciences of the USSR.

---

The indices of the two corpora of text editions may serve as a dictionary:

König, F. 1955–1957. *Handbuch der chaldischen Inschriften*. Graz: Selbstverlag Ernst Weidners.

Melikišvili, G. 1960. *Urartskie klinoobraznye nadpisi*. Moscow: Academy of Sciences of the USSR

(Supplement: UKN II, in *Vestnik Drevnej Istorii* 1971/3, pp. 229–255, 1971/4, pp. 267–294).

For history, culture and comprehensive bibliographies see:

Salvini, M. 1995. *Geschichte und Kultur der Urartäer*. Darmstadt: Wissenschaftliche Buchgesellschaft.

Zimansky, P. 1998. *Ancient Ararat. A Handbook of Urartian Studies*. Delmar, NY: Caravan Books.

# Afro-Asiatic

JOHN HUEHNERGARD

## 1. THE AFRO-ASIATIC FAMILY

### 1.1 Introduction

In the following paragraphs only a brief overview of the Afro-Asiatic family can be given, with some of the shared features that have prompted recognition of the family. Work on Afro-Asiatic is still in its infancy, and work on the reconstruction of Proto-Afro-Asiatic has barely begun. The remainder of this chapter will be concerned with one of the two well-known ancient branches of Afro-Asiatic, the Semitic branch (the other ancient branch is Egyptian, for which see Ch. 7; for a probable ancient form of Berber, see below, §1.1.3).

The original homeland of Afro-Asiatic has been the subject of some discussion. Most scholars would place it somewhere in the vicinity of the center of the family's current geographical range, or rather further to the east, in far southern Egypt or northern Sudan. A few scholars, however, have argued for an original location in southwest Asia (see Militarev 1994; Diakonoff 1998).

Older names of the Afro-Asiatic family, still used by some scholars, include *Hamito-Semitic* and *Semito-Hamitic*, names that have generally been abandoned because they imply a subgroup of “Hamitic” languages (i.e., of all languages in the family apart from the Semitic languages) that is not indicated by any isoglosses.

The Afro-Asiatic family comprises at least five and as many as eight branches.

#### 1.1.1 Egyptian

See Chapter 7.

#### 1.1.2 Semitic

See §§2–3 below.

#### 1.1.3 Berber

Berber was formerly spoken across much of Africa north of the Sahara, but with the spread of Islam, it has been reduced to a series of linguistic islands in a sea of Arabic. Even so, Berber languages are still spoken by 10–15 million people. Berber languages in Morocco include Tashelhit in the High Atlas mountains, Tamazight in the Middle Atlas mountains, and Tarifit in the Rif mountains. In Algeria the main Berber language is Kabyle, though several other

forms of Berber also occur. Tuareg is also spoken in Algeria, as well as in Mali and Niger. Smaller Berber dialects are spoken in Tunisia, Libya, Egypt (oasis of Siwa), and Mauritania. The Guanche language of the Canary Islands, extinct since the sixteenth century AD, was probably also a Berber language.

A Berber dialect (or dialects) is probably represented in the corpus of over a thousand Numidian (or Lybian, or Lybico-Berber) inscriptions in a consonantal alphabet that have been found in Tunisia, Algeria, and Morocco. Most of the inscriptions date to the second century BC (only one is actually dated, however, a Numidian–Punic bilingual from 139 BC). The texts are difficult to interpret, and thus of limited use for the earlier history of Berber. The script resembles the Tifinigh (or Tifinagh, “Punic [letters]”) alphabet that is now used among the Tuareg (see O’Connor 1996).

#### 1.1.4 Cushitic

Some forty Cushitic languages are spoken by about 15–20 million people in Ethiopia, Somalia, and surrounding countries. The earliest records of Cushitic languages date to the eighteenth century AD. Cushitic is divided into four branches:

1. *North Cushitic*: the Beja language (claimed by some scholars to be a separate branch of Afro-Asiatic; see below, §1.1.6).
2. *Central Cushitic or Agaw*: formerly the major Cushitic language in Ethiopia, which had significant influence on the later Semitic languages there, today represented by a number of languages with small numbers of speakers (Awngi, Bilin, Kemant, Xamir).
3. *East Cushitic*: numerically by far the largest branch of Cushitic, and itself further subdivided into *Lowland East Cushitic* (including Oromo, formerly called Galla, a pejorative term, with 8–10 million speakers in central Ethiopia; Afar-Saho, along the Red Sea coast of Eritrea; and Somali, the official language of Somalia), *Highland East Cushitic* (or Rift Valley Cushitic, including especially the Sidamo language), and smaller subbranches.
4. *South Cushitic*: includes languages spoken in Kenya and Tanzania (such as Alagwa, Burunge, and Iraqw).

#### 1.1.5 Chadic

Chadic is a very large family of some 140 languages spoken by perhaps 30–40 million people in Cameroon, Central African Republic, Chad, Niger, and Nigeria. One Chadic language, Hausa, also serves as a lingua franca in much of western Africa. There are no records of Chadic languages before the modern period. The Chadic languages are divided into three large branches, each of which is further subdivided:

1. *West Chadic*: including Hausa, the Ron languages, and the Bauchi subbranch.
2. *Central Chadic*: languages such as Bura, Margi, Kotoko-Logone, Masa.
3. *Eastern Chadic*: languages such as Kera, Migama, Mubi.

#### 1.1.6 Other possible branches

The Omotic languages, about forty in number, are spoken by about 3 million people, mostly along the Omo River in southwestern Ethiopia. The most prominent language is Wolaytta, with about 2 million speakers. Omotic was formerly considered a western branch of Cushitic,

but is now considered by many scholars to be an independent branch of Afro-Asiatic, although there continues to be discussion about its status.

The Beja language, spoken by about a million people along the Red Sea coast of Sudan and southeastern Egypt, is usually considered to be North Cushitic, but it has sometimes been proposed as a separate branch of Afro-Asiatic (Hetzron 1980). It has a number of intriguing archaic features.

A language spoken by fewer than twenty individuals along the Woito River in southwestern Ethiopia, called by themselves Ongota and by their neighbors Birale or Birelle, has recently been described (Fleming *et al.* 1992) and claimed to be the remnant of another distinct branch of Afro-Asiatic (Fleming 1999).

## 1.2 Subgrouping of Afro-Asiatic

A number of morphological features indicate that Berber, Egyptian, and Semitic may constitute a *North Afro-Asiatic* subgroup. A connection between Berber and Chadic has also been suggested. Various other, more comprehensive subgroupings of the Afro-Asiatic branches have been proposed, but none has gained a consensus.

Macro-comparisons of Afro-Asiatic with other language phyla, such as Indo-European (the so-called Nostratic hypothesis), have not met with general acceptance.

## 1.3 Features of Afro-Asiatic

Most of the features enumerated here are attested in several, but usually not all, of the branches of the family.

### 1.3.1 Phonology

Phonological commonalities include the pharyngeal fricatives [ʕ] and [ħ], and a third series of consonants (in addition to the usual voiced and the voiceless), often called “emphatic,” which in most of the branches have a glottalized realization (but are pharyngealized in Berber and in Arabic).

### 1.3.2 Morphology

In the morphology, the personal pronouns exhibit a number of common features across Afro-Asiatic. Most branches, for example, have both independent and suffixed forms, the latter used both for objects when attached to verbs and for possession when attached to nouns. Common forms include *\*ʔan(V)* and *\*ʔana(:)k(V)* for the first-person singular, and *\*k* as marker of the second person. Demonstrative pronouns also show a number of common elements across the branches.

The “root and pattern” system of noun and verb bases that is well known among the Semitic languages (see §3.3.1) seems to be a common Afro-Asiatic feature, as does the existence of a preponderance of triconsonantal roots (but also a significant number of biconsonantal roots).

Among inflectional features of the noun may be noted (i) the presence of *\*-t* as marker of feminine; (ii) a case system similar to that of Proto-Semitic (Sasse 1984); (iii) pluralization by means of the insertion of *\*a* before the final root consonant (Greenberg 1955a), as well as other “broken” plurals (see below, §3.3.2.4); (iv) a prefix *\*ma-* to form nouns of place, instrument, and agent; and (v) a denominative adjectival ending *\*-i:(y)*.

In verbal morphology, it is likely that the following may be reconstructed for Proto-Afro-Asiatic: (i) a prefix-conjugation, which marked person much as in Semitic, with

\*ʔ- for first-person singular, \*t- for second-person (and perhaps for third-person fem.), and \*y- for third-person (masc.); (ii) the presence of \*a to mark the imperfective form of the verb; and (iii) a set of derivational consonant affixes, \*s for causative, \*t for reflexive or middle, and \*n for passive.

On the cusp bridging nominal and verbal morphology is the predicate adjective or suffix-conjugation, a predication composed of a verbal adjective and an enclitic subject pronoun (found in Semitic [see §3.3.2.1], Egyptian, and, probably, Cushitic).

## 1.4 Afro-Asiatic vocabulary

One of the greatest hindrances to the reconstruction of Proto-Afro-Asiatic has been the difficulty of establishing clear cognate sets across the vocabularies of the several branches (this has also, of course, impeded efforts to establish sound correspondences across the branches and to reconstruct Proto-Afro-Asiatic phonology). Essentially, this must await the working out of reconstructed proto-vocabularies for the individual branches, which is still in its beginning stages, except for Semitic. Nevertheless, a few lexical items common to at least several of the branches may be mentioned, such as \*lis “tongue,” \*m-w-t/mut “to die,” \*s(i)m “name,” and \*sin(n) “tooth.”

# 2. THE SEMITIC LANGUAGES

## 2.1 Introduction

Semitic is a close-knit family of languages first attested in Akkadian names and loanwords occurring in Sumerian cuneiform texts of the first half of the third millennium BC. Akkadian texts proper begin to appear about 2500 BC, and Eblaite shortly thereafter. Many Semitic languages continue to be spoken to this day, including (i) Arabic in many countries of Asia and Africa; (ii) Amharic, Tigrinya, and other related languages in Eritrea and Ethiopia; (iii) Hebrew in Israel; (iv) South Arabian languages such as Mehri, Jibbāli, and Soqotri in Yemen and Oman; and (v) many varieties of Aramaic, now scattered around the globe.

## 2.2 The prehistory of Semitic

It is not known when Semitic hived off from the common Afro-Asiatic stock, other than that the separation must antedate the third millennium BC; nor can anything be said with confidence about the original homeland or early movements of the ancestral Semitic speakers beyond what has been observed above in §1.1. As noted in §1.2, the closest relatives of Semitic within Afro-Asiatic seem to be Egyptian and Berber.

## 2.3 The subgrouping of the Semitic languages

The earliest partition within the Semitic family separated *West Semitic* from Akkadian and Eblaite (see Ch. 8), which together are termed *East Semitic*. West Semitic languages are characterized by an innovative perfective form of the verb, a suffix-conjugation, exemplified by Arabic *katabtu* “I wrote.” The West Semitic group in turn is comprised of three branches:

1. *Central Semitic*: includes (i) the Northwest Semitic languages Ugaritic; Hebrew, Phoenician, and other Canaanite dialects; and Aramaic (see Chs. 9–13); (ii) the Ṣayhadic (Old or Epigraphic South Arabian) languages (see Ch. 15); and (iii) the various forms of Arabic (see Ch. 16).

2. *Ethiopian Semitic*: attested in the ancient period in classical Ethiopic, or Gəʿəz (see Ch. 14).
3. *Mahrrian Semitic or Modern South Arabian*: not attested until the modern period (unless the Old South Arabian language Ḥaḍramitic reflects an ancient member of this group).

### 3. DESCRIPTION OF PROTO-SEMITIC

#### 3.1 Introduction

What follows summarizes some of the reconstructable features of Proto-Semitic as a linguistic system. It is based, of course, on the work of many scholars, not all of whose studies could be mentioned in such a brief overview; nor has it always been possible to allot space for a detailed defense of some of the reconstructions offered here.

#### 3.2 Phonology

##### 3.2.1 Consonants

Common Semitic is uncontroversially reconstructed with twenty-nine consonantal phonemes. The original pronunciation of the consonants is disputed, but a likely set of phonetic values is given in Table 6.1:

**Table 6.1 The consonantal phonemes of Common Semitic**

Manner of articulation	Place of articulation						
	Bilabial	Inter-dental	Dental/Alveolar	Palatal	Velar	Pharyngeal	Glottal
<i>Stop</i>							
Voiceless	p		t		k		ʔ
Emphatic			tʰ		kʰ		
Voiced	b		d		g		
<i>Affricate</i>							
Voiceless			tʰs				
Emphatic			tʰsʰ				
Voiced			dʒ				
<i>Fricative</i>							
Voiceless		θ	s		x	ħ	h
Emphatic		θʰ					
Voiced		ð			ɣ	ʕ	
<i>Lateral continuant</i>							
Voiceless			ɬ				
Emphatic			ɬʰ				
Voiced			l				
<i>Tap/Trill</i>			r				
<i>Nasal</i>	m		n				
<i>Glide</i>	w			y			



As Table 6.1 indicates, Proto-Semitic was characterized by a number of consonant triads consisting of a voiceless, an ejective (i.e., a so-called emphatic), and a voiced member. For two of the obstruent sets that are lacking an ejective member, namely the bilabial stop and the velar fricative sets, that member has been posited for Proto-Afro-Asiatic – \**p*’ and \**x*’. The consonantal repertoire of Proto-Afro-Asiatic is in general assumed to have been significantly larger than that of Proto-Semitic.

All of the consonants could be geminated.

In the traditional Semitological literature, the consonants here characterized as ejective are normally transcribed with an underdot, for example, *ṭ* for /t’/, *θ̣* for /θ’/, and so forth, as is the voiceless pharyngeal fricative, in other words, *ħ̣* for /ħ/. The velar fricatives /x/ and /ɣ/ are usually written by Semitists as *ħ* and either *g* or *ǧ*, respectively, while the voiceless and emphatic lateral fricatives /ɬ/ and /ɬ’/ are usually written as *ś* and either *š* or *ṣ*, respectively. Further, the consonants here characterized as alveolar affricates are traditionally represented as simple fricatives, *s*, *s*’ or *ś*, and *z*, while the sibilant given above as /s/ is traditionally represented as *š*.

At least one assimilation process may be ascribed to Common Semitic, namely, the assimilation of *w* to a following dental or alveolar, as in Akkadian *ittarad* < \**yawtarad* “he has descended”; Arabic *yattaḥidu* < \**yawtaḥidu* “it will be united”; Hebrew *yis’s’or* < \**yawṭs’ur* “he fashions.” Attested in only part of the Semitic area, perhaps reflecting an areal development, is the assimilation of \**n* to a following consonant, which occurs regularly in Akkadian and the Northwest Semitic languages; compare Common Semitic \**yanθ’ur* “he guarded” > Arabic *yanz’ur*, but Akkadian *is’s’ur*, Hebrew *yis’s’or*, Aramaic *yit’t’ar*. In Šayhadic inscriptions the same assimilation is frequently, but not consistently, reflected.

The consonants \**w* and \**y* were regularly lost in the environment C\_Ũ, with compensatory lengthening of the following vowel, as in \**maka:n* < \**makwan* “place”; \**madi:nat* < \**madyinat* “administrative region.”

The existence of syllabic allophones of the sonorants \**l*, \**m*, and \**n* in certain environments has been suggested to account for a number of phenomena attested in the descendant languages (Testen 1998). Examples include the ancient substantives \**bṇ*- “son” and \**sm*- “name,” the pronominal forms \**sn* “they (masc.)” and \**sn* “they (fem.),” and the proclitic asseverative particle \**l*-.

### 3.2.1.1 Major developments in the descendant languages

In most of the West Semitic languages, the common Semitic alveolar fricative \**s* underwent a change to \**h* when prevocalic (i.e., *s* > *h* / \_V) as in Common Semitic \**suḏa* > West Semitic \**huḏa* “he”; Common Semitic \**yusaḥbir* > West Semitic \**yuḥaḥbir* “he sent across”; Common Semitic \**baytisa* > West Semitic \**baytiha* (eventually to Hebrew *bayṯah*) “to the house.” In an interesting development resulting from the morphological patterning of Semitic (see §3.3.1), this sound change was blocked in most nominal and verbal forms because the conditioning environment – namely, the following vowel – did not appear in all forms; for instance, although \**sarik*’ “stolen” would have developed into \*\**harik*’ – by the sound rule, no change would have occurred in the verbal form \**yasrik*’ “he stole,” where \**s* was not followed by a vowel; a principle of root integrity (essentially an overriding avoidance of root allomorphy) then blocked the change \**sarik*’ > \*\**harik*’ – as well. Thus, \**s* generally remains in West Semitic nominal and verbal roots, but is otherwise missing. In a number of languages, including Aramaic, Hebrew, Jibbāli (Mahrian branch), and the Babylonian form of Akkadian, \**s* became a palato-alveolar *š*; note, for example, Arabic and Ethiopic *sala:m* “well-being,” but Babylonian *šala:mum*, Aramaic *šla:m*, Hebrew *šlom*.

The early dental/alveolar affricates  $*t_s$ ,  $*t_s'$ ,  $*d_z$  were deaffricated in most of the attested languages, becoming  $*s$ ,  $*s'$ , and  $*z$ , respectively. In Arabic and Ethiopic, the new voiceless  $*s < *t_s$  merged with Common Semitic  $*s$ . In the Assyrian form of Akkadian, however,  $*t_s$  became  $\dot{s}$ .

The Common Semitic voiceless and ejective lateral fricatives,  $*t$  and  $*t'$ , underwent changes in most of the attested languages, although the nonejective  $*t$  is still preserved as such in the Mahrian languages, and was probably also pronounced as such in early Akkadian, Hebrew, and Arabic. The ejective  $*t'$  merged with  $*t_s'$  in Hebrew, in Akkadian, and in Ethiopian Semitic (although it is preserved as a distinct phoneme in the earliest classical Ethiopic, pronunciation unknown). In Aramaic, however, it became first  $\gamma$  and finally  $\text{ʔ}$ ; compare Hebrew  $\text{ʔeʔeʔ}$ , Akkadian *ers'etum*, but Aramaic  $\text{ʔarʕa}$ , from Common Semitic  $*ʔarʔ$  - "earth."

In Arabic and perhaps in some other Central Semitic languages, most of the common Semitic ejective or glottalic consonants became pharyngealized, for example,  $*t' > t^h$ ,  $*t_s' > *s' > s^h$ . The velar ejective  $*k'$ , however, became a nonejective uvular stop  $q$ . The Arabic reflexes of  $*\theta'$  and  $*t'$  vary according to dialect, but for the classical language are usually said to be a voiced interdental or dental/alveolar fricative,  $\delta^h$  or  $z^h$ , and a voiced dental/alveolar stop,  $d^h$ , respectively.

As the result of an areal spread, the bilabial stop  $p$  became a labiodental fricative  $f$  in several branches of Semitic, namely, Mahrian, Ethiopian, and the Ṣayhadic and Arabic subbranches of Central Semitic.

A characteristic of the Northwest Semitic languages is the change of initial  $*w$  to  $*y$ , as in Hebrew  $\text{yélēḏ}$ , Aramaic *yalda*: "child" < Common and Central Semitic  $*wald$ .

### 3.2.2 Vowels

Proto-Semitic (and probably Proto-Afro-Asiatic) may be reconstructed with three vowels, high front  $*i$ , high back  $*u$ , and low central  $*a$ , each of which could occur short or long.

On diphthongs and triphthongs see §3.2.3.

The presence of  $*i(:)$  in the base of a Proto-Semitic word seems to have precluded the presence of another high vowel elsewhere in the base. In other words, bases with the vowel melodies  $i \dots i$ ,  $i \dots u$ , and  $u \dots i$  do not seem to have occurred, though bases with two  $u$  vowels,  $CuC(C)u(:)C-$ , can be reconstructed.

Internal reconstruction indicates the existence of a Proto-Semitic rule of vowel syncope:  $a > \phi / aC_1 \text{ \_\_\_ } C_1 V$ , as in  $*k'alalum > *k'allum$  "light, small."

### 3.2.3 Syllable structure

It is likely that only three syllable shapes are to be reconstructed for Proto-Semitic – two open, CV and CV:, and one closed, CVC. These syllable-types may also be classified quantitatively, as either light, CV, or heavy, CV: and CVC. Thus, all syllables contain a single vowel, begin with a single consonant, and end either in a single consonant or in a vowel. The following conditions are not permitted: (i) sequences of two or more consonants word-finally; (ii) sequences of three or more consonants within words; (iii) sequences of two or more vowels; (iv) long vowels in closed syllables.

Since only one vowel quality was permitted in each syllable, true *phonemic* diphthongs did not occur in Proto-Semitic (nor are they attested in most of the descendant languages). Semitists, however, often speak of the phonetic sequences [V + glide] (i.e.,  $Vw$  and  $Vy$ ) as "diphthongs," even though Semitic syllable structure dictates that the glide functions as a

consonant in such cases. The sequences *\*aw* and *\*ay* were common in Proto-Semitic, as in *\*mawt*- “death” and *\*bayt*- “house.” These were frequently monophthongized, usually to [o:] and [e:], respectively, in many of the languages.

Already in Proto-Semitic the sequences *\*iy* and *\*uw* were realized as [i:] and [u:], respectively; thus the noun *\*di:n*- “judgment,” from the root *d-y-n*, may be said to be equivalent to *\*diyyn*- (i.e., a noun of the pattern *CiCC*), and so comparable in form to the noun *\*ḏibḥ*- “sacrifice” from the root *ḏ-b-ḥ*. Similarly, *\*k'u:m*- “height,” from the root *k'-w-m*, is equivalent to *\*k'uwm-* (i.e., of the pattern *CuCC*) and comparable to *\*ḥumk'*- “depth,” from the root *ḥ-m-k'*.

The sequences *VwV* and *VyV*, sometimes called “triphthongs” in Semitic studies, tended to be unstable and to be reduced to “diphthongs” or to simple vowels, as in *\*mawit-/mayt-* (< *mait-*)/*mi:t-/mit-* “dead.” For the sequences *CwV* and *CyV* see §3.2.1.

The implications of reconstructing a set of syllabic allophones of certain consonants (see §3.2.1) require further investigation. Clearly, however, the generalizations just enunciated would need to be modified if forms such as *\*bn-* + case ending – that is, *\*[bṇum]* “son” (CCVC?, CVVC?) – are to be considered valid in Proto-Semitic.

### 3.2.4 Stress

The evidence suggests that Proto-Semitic word stress was not phonemic, but assigned automatically (i) to the rightmost nonfinal heavy syllable (*CV:* or *CVC*), or (ii) in words having only nonfinal light syllables, to the initial syllable: *\*salima*, *\*salimu:*, *\*salimta(:)*, *\*yislam*, *\*yislamu:*, *\*yislām:mu:na*. This is essentially the pattern assumed to operate in both classical Arabic and Akkadian, which are widely separated within the Semitic family.

There are instances in which stress is phonemic in some of the descendant languages, but these are undoubtedly the result of internal developments: for example, classical Ethiopic *'səḥtat* “she erred,” but *səḥ'tat* “error”; *ra'kaba*: “they (fem.) found,” but *rak'aba*: “he found her”; Hebrew *k'ōmō* “she stood” versus *k'ō'mō* “standing” (fem. sg.), *rōs'u* “they ran” versus *rō's'u* “they were pleased.”

## 3.3 Morphology

### 3.3.1 Morphological type and word structure

Common Semitic, like its descendants, may be characterized as a fusional language.

Certain pronouns and a small but important number of isolated substantives, that is, substantives not associated with a verbal root, may be reconstructed for Proto-Semitic as discrete, and complete, lexical items with no formal restrictions other than those imposed by the constraints of syllable structure: for example, *\*ʔanti(:)* “you (fem. sg.),” *\*suʔa* “he,” *\*yadum* “hand,” *\*ʕit'um* “tree,” *\*kalbum* “dog,” *\*ʔu<sup>d</sup>znum* “ear,” *\*ḥima:rum* “(male) donkey,” *\*ʔarnabum* “hare” (see Fox 1998).

But a remarkable characteristic of Semitic morphology is that the majority of words – all verbal forms and most nouns – reflect the interdigitation of a root, consisting of an invariable sequence of consonantal radicals (usually three in number), and a pattern of vowels and other features, which include gemination of one of the root consonants (other than the first) and affixation of a small subset of the consonantal repertoire (especially *ʔ*, *m*, *n*, *s*, *t*, *y*; these also appear commonly in the pronominal systems). As examples the following forms of the root *\*s-l-m* “(to be) whole, sound, well” may be cited, with *R<sub>1</sub>* and so forth representing the root consonants:

- (1) pattern  $R_1 a R_2 i R_3$  (a common adjectival form): \**salim*- “whole, sound, well”;  
 pattern  $R_1 a R_2 a R_3$  (a common verbal noun form): \**sala:m*- “wholeness, well-being”;  
 pattern  $mu R_1 a R_2 R_2 i R_3$  (participle of a derived verbal form): \**musallim*- “(one) who makes whole.”

It is possible that verbal roots consisting of fewer than three consonantal radicals were not unusual in Common Afro-Asiatic. By the Proto-Semitic period, however, the triradical root was the norm, roots that earlier may have had fewer radicals having been conformed to that norm by various analogical developments. Original biradical bases may perhaps be detected in some roots with first radical *w*, such as \**w-θ-b* “dwell,” that lack the *w* in certain forms, such as the verbal noun \**θib-t*- “dwelling,” across the descendant languages and even occasionally in cognates in Egyptian; and in some biform root pairs of the type  $R_1-R_2-R_2 \sim R_1-R_2-w/R_1-R_2-y$ , that must also be reconstructed to the proto-language. Common Semitic also probably had a small number of quadriradical roots, most of them with a sonorant in second position.

Certain constraints on the composition of the Semitic verbal root have been noted (Greenberg 1950): roots with identical first and second radicals are unattested, and roots with identical first and third radicals are extremely rare. In addition, homorganic consonants tend to be avoided within a root, except for the common root type known as the geminate, in which the second and third radicals are identical.

### 3.3.2 Nominal morphology

Reconstruction indicates that Proto-Semitic nouns occurred in two states, bound and free (adjectives also in a third, predicative); two genders, masculine and feminine; three cases, nominative, genitive, and accusative (and perhaps a fourth, directive); and three numbers, singular, dual, and plural. Proto-Semitic did not have a definite or an indefinite article. A definite article first evolved in the Central Semitic branch, while an indefinite article failed to develop in most of the descendant languages (apart from the occasional use of the numeral “one” for “a certain”).

#### 3.3.2.1 State

Proto-Semitic nouns occurred in two syntactic states, either (i) bound to a following qualifier or (ii) not thus bound, in other words, free. Free forms were marked with an ending that exhibited two allomorphs, \**-m* after short vowels, \**-na* after long vowels and diphthongs: for example, nominative singular \**wa:θibum* “inhabitant,” plural \**wa:θibu:na* “inhabitants” (see further below). Bound forms (also called *construct forms*), which lacked this ending, governed an immediately following constituent, which was either a noun in the genitive case (2A–B), a (genitive) pronominal suffix (2C–D), or a nominalized (relative) clause (2E–F):

- (2) A. \**wa:θibu baytim* “inhabitant of the house”  
 B. \**wa:θibu: baytim* “inhabitants of the house”  
 C. \**wa:θibu-su(:)* “its inhabitant”  
 D. \**wa:θibu: su(:)* “its inhabitants”  
 E. \**wa:θibu yamu:tu* “the inhabitant who died”  
 F. \**wa:θibu: yamu:tu:na* “the inhabitants who died”

Nothing was permitted to intervene between a bound form and the constituent governed by it; an attributive adjective (in the free form), for example, followed the construction:

\**wa:θibu baytim salimum* “the sound inhabitant of the house” (vs. \**wa:θibu baytim salimim* “the inhabitant of the sound house”).

In addition to occurring in the bound and free forms, adjectives of verbal roots, when functioning as the predicate of their clause, entered into a special morphological construction which was comprised of the simple base of the adjective (unmarked for case, gender, or number) followed by an enclitic subject pronoun, as in \**salim-ti(:)* “you (fem. sg.) are well”; \**salim-at baʕlatum* “the lady (she) is well.”

The comparative degree was expressed syntactically rather than morphologically; that is, there was no special comparative form of the adjective, a comparison such as “their army is larger than our army” being expressed as “their army is large from/against our army.” For the superlative, Akkadian and Arabic attest a form of the adjective augmented by a prefix resembling the causative marker of the verbal system, but it is unclear whether this reflects a Proto-Semitic feature (Speiser 1952). It is likely that the superlative could be expressed by a bound-form adjective governing a plural noun, as in “the great one of the gods” = “the greatest god.”

### 3.3.2.2 Gender

The evidence of the descendant languages suggests that in Common Semitic any given substantive was construed either as masculine or as feminine. Of the two genders, the masculine was generally unmarked formally, whereas most feminine nouns were marked with an ending. Each of the languages, however, attests a number of unmarked words that are construed as grammatically feminine, including: (i) the words for “mother” (\**ʔimm-*), “ewe” (\**laxir-*), “female donkey” (\**ʔata:n-*); (ii) words for the parts of the body that occur in pairs – for example, \**ʕayn-* “eye,” \**birk-* “knee,” a curious phenomenon that undoubtedly arose because the ending of the dual on nouns (nominative) and verbs and the ending of the feminine plural on some verbal and adjectival forms were formally identical, namely, \*-a; and (iii) a semantically disparate group of other words for inanimate objects that varies from language to language and is thus difficult to reconstruct in the proto-language with any certainty. A few unmarked nouns in each language – again the set varies – are construed as both masculine and feminine.

The marker of the feminine is \*-t or \*-at, which appears after the base but before a case ending; examples are \**baʕl-* “lord,” \**baʕl-at-* “lady”; \**wa:θib-* “inhabitant (masc.),” \**wa:θib-t-* (fem.). The original distribution of \*-t versus \*-at is difficult to recover with certainty. In all of the languages, for reasons of syllable structure, the ending \*-at appears after bases ending in two consonants (a sequence of three consonants being prohibited), as in \**baʕl-at-*. In some of the descendant languages, such as Akkadian and Aramaic, \*-at appears only on such bases, \*-t occurring on all other forms. In Arabic, \*-at has been generalized (with a few exceptions, such as *bin-t-* “daughter”). Classical Ethiopic patterns for the most part like Akkadian and Aramaic, in other words, with \*-t unless \*-at is phonologically necessary; but there are a number of exceptions, such as *ʕəlat* “day,” *xatʕat* “sin.” In Hebrew, \*-at (> Hebrew -ָ, bound-form -אָת) predominates on verbal adjectives (as in *kəbeðā* < \**kabidat-* “heavy”); but otherwise the occurrence of the two endings suggests a certain free variation at an earlier period: for example, *dēlθ* < \**dal-t-* “door,” versus *ʕəmə* < \**ʕam-at-* “female slave.”

The endings \*-at/-t have a number of semantic functions: (i) to mark the feminine singular of adjectives; (ii) to denote the female member of various pairs of words, such as \**baʕl(-at)-*, “lord/lady” and \**kalb(-at)-* “dog/bitch”; (iii) to denote the single member of the class represented by a collective noun (termed in traditional Semitic grammar the *nomen unitatis*),

as in *\*bak'ar-* “cattle,” *\*bak'ar-at-* “a cow” and *\*ḥiṣar-* “hair,” *\*ḥiṣar-at-* “a hair”; and finally (iv) as a suffix on many substantives with no obvious feminine or other common semantic connotations.

The various descendant languages preserve vestiges (less rare in Arabic) of other markers of the feminine that must be reconstructed to Proto-Semitic, including *\*-ay* and *\*-a:ʔ*.

### 3.3.2.3 Case

Traditional Semitic grammar recognizes three cases of the noun, each of which is marked, in singular forms, by one of the short vowels. These cases are given labels borrowed from the classical Indo-European languages: nominative, marked by *\*-u*; genitive, marked by *\*-i*; and accusative, marked by *\*-a*.

The *nominative* is used for the subject of a clause, for the predicate of a verbless equational clause (as in “my brother is the king”), and as a citation form and for extraposition (“as for the king” = nom.). The ending *-u* also functions in a locative sense (*\*libbum* “in the heart”) in Akkadian and vestigially in other languages; it is unclear whether the nominative and locative functions are to be considered reflexes of a single case at an earlier stage.

The *genitive* is an adnominal case, used after all bound forms and all prepositions (many of which originate as bound-form nouns). The ending that marks the genitive, *\*-i*, is undoubtedly connected to the morpheme *\*-i:y* that is suffixed to substantives to form denominative adjectives (see §3.3.2.6 below).

The so-called *accusative* is indeed used to mark the object, usually the direct object, of the verb, but also in a host of other adverbial functions, such as to indicate manner, means, location, and “time when.” If, as has been suggested, Proto-Semitic at an early stage had an ergative verbal system, *\*-a* may have marked the absolutive case (see, e.g., Diakonoff 1988:59,101).

In dual and plural forms, the genitive and accusative are invariably marked by a common set of endings, and the two cases are sometimes jointly termed the *oblique*.

Another common Semitic noun ending that may perhaps be considered a case marker is *\*-isa*, the reflexes of which, in Akkadian, Ugaritic, and Hebrew, have a directional nuance, as in *\*baytisa* “houseward.” In Akkadian, further, and more commonly, it is attached to adjectives to create adverbs: *t'a:bum* “pleasant (nom. sg.),” *t'a:biš* “pleasantly.” This ending occurs only on singular forms.

### 3.3.2.4 Number

The dual was marked by a set of endings attached to the singular base of the noun, following the feminine marker if one was present. The evidence of Old Akkadian, Ugaritic, Šayhadic, and Arabic indicates that the dual was regularly used to indicate “two” of anything. In later Akkadian, in Hebrew, and in early Aramaic the use of the dual came to be restricted to words for naturally occurring pairs of objects and certain time words. In later Aramaic, in Ethiopian, and in some of the Mahrian languages the use of the dual has become obsolescent or has been lost entirely.

The plural in a northern group of the Semitic languages – namely, Akkadian and the Northwest Semitic subbranch – is indicated by a set of endings attached to the singular base of the noun, replacing the case endings of the singular; the feminine ending is altered from *\*(a)t* to *\*-a:t* in the plural. These plural endings may to a certain extent be seen to involve the feature [+length] vis-à-vis their singular counterparts. In the rest of the languages – Ethiopian, Mahrian, Šayhadic, and Arabic – pluralization is normally expressed by means of pattern replacement (called “broken plurals” or “internal plurals”), of the type *\*kalb-* “dog,” plural *\*kila:b-*. Such forms take the same case endings as singular



forms. Since there is evidence for both types of pluralization in both groups of languages, in other words, vestigial use of pattern replacement in the northern group, and the use of external plural endings for certain noun types in the other languages, it is clear that both types are to be reconstructed for Proto-Semitic. It seems plausible that the external plurals were at first restricted to verbal adjectives (the endings are clearly related formally to the endings on predicate adjectives with third-person subjects), and that most other nouns either had plurals formed by pattern replacement or were collectives that had no special plural forms (or, perhaps, were simply unmarked for number). Certain features of the pattern-replacement plurals, such as *a*-insertion between the second and third root radicals, can be traced back to Common Afro-Asiatic (Greenberg 1955a; Ratcliffe 1998).

### 3.3.2.5 Declension

Below is presented a sample Proto-Semitic nominal paradigm, that of the active participle of the root *\*w-θ-b* “to sit, dwell,” including feminine and external plural forms. The elements *-m* and *-na* are present only in the free (unbound) forms of the noun, but missing in bound forms (see §3.3.2.1).

#### (3) The Proto-Semitic nominal paradigm

	<i>Masculine</i>	<i>Feminine</i>
<i>Singular</i>		
<i>Nominative</i>	wa:θibu-m	wa:θib(a)tu-m
<i>Genitive</i>	wa:θibi-m	wa:θib(a)ti-m
<i>Accusative</i>	wa:θiba-m	wa:θib(a)ta-m
<i>Dual</i>		
<i>Nominative</i>	wa:θiba:-na	wa:θib(a)ta:-na
<i>Genitive-accusative</i>	wa:θibay-na	wa:θib(a)tay-na
<i>Plural</i>		
<i>Nominative</i>	wa:θibu:-na	wa:θiba:tu-m
<i>Genitive-accusative</i>	wa:θibi:-na	wa:θiba:ti-m

### 3.3.2.6 Noun derivation

A number of specific nominal patterns, when applied to verbal roots, may be identified with certain semantic classes (see Barth 1894, Fox 2003). Thus, for example, the pattern  $R_1 a R_2 i R_3$  is reconstructable as the active participle of nonstative verbal roots. Nouns of the monosyllabic patterns  $R_1 V R_2 R_3$  are normally substantives rather than adjectives, whereas nouns of the patterns  $R_1 a R_2 V R_3$  tend to be (but need not be) adjectives. Of the monosyllabic patterns just mentioned,  $R_1 i R_2 R_3$  substantives are frequently passive: *\*θik'l-* “weight, what is weighed,” *\*simʕ-* “report, what is heard,” *\*θibħ* “sacrifice, what is sacrificed.” The pattern  $R_1 u R_2 R_3$  is often used for abstracts of stative roots: *\*ʔurk-* “length,” *\*murr-* “bitterness,” *\*t'u:b-* (< *\*t'uyb-*) “goodness.” In general, however, it is only the patterns of such deverbal forms that are reconstructable for the proto-language, not individual lexemes, much reshuffling having occurred in the various branches and individual languages.

Derivational endings include the following: (i) *\*-a:n*, an individualizing morpheme, as in Akkadian *šarra:k'a:num* “the thief in question,” from *šarra:k'um* “thief”; (ii) *\*-i:y*, which forms denominative adjectives (including gentilics), such as *\*sapli:y-* “low,” from

\**sapl*- “bottom, under part”; and (iii) \*-*u:t*, which forms abstracts, as in \**baflu:t*- “lordship,” from \**baʕl*- “lord.”

### 3.3.3 Personal pronouns

The personal pronouns, like nouns, have three numbers, singular, dual, and plural. First-person forms are of common gender, while both the second and the third persons exhibit distinct masculine and feminine forms in the singular and the plural. Duals are of common gender. First-person dual forms are only rarely attested in the descendant languages, and where attested may be later innovations rather than vestiges of Proto-Semitic forms. The enclitic forms of the pronouns distinguish a nominative set, used as the subjects of predicate adjectives, as in \**salim-nu*(:) “we are well” (see §3.3.5.1), and a genitive/accusative set, used as possessive pronouns on nouns, as in \**baytu-ka*(:) “your (masc. sg.) house,” and as objects on verbs, as in \**yanθ’ur-ka*(:) “he guarded you (masc. sg.).” For the first person, distinct genitive and accusative forms existed.

In Table 6.2, the vowels occurring at the ends of many of the forms are marked as optionally long; they are short when word-final, long otherwise. The second- and third-person plural forms must be reconstructed with optional extensions, namely, \*-*u*: on masculine forms (e.g., \**sumu*: in addition to \**sum*), and \*-*na*(:) or \*-*a*: on feminine forms (e.g., \**sinna*(:) or \**sina*: in addition to \**sin*). If Proto-Semitic is to be reconstructed with syllabic sonorants (see §3.2.1), then the second- and third-person dual and plural pronouns may be reconstructed as, for example, second masculine plural \*-*tʔl*/, second feminine plural \*-*tʔl*/, and so forth, rather than with the sequence [homorganic vowel + sonorant] as given in Table 6.2.

Possessive adjectives are attested in several of the Semitic languages, but their divergent construction makes it difficult to reconstruct such forms for the proto-language.

The Semitic languages do not attest a true reflexive pronoun, and it is unlikely that one existed in the proto-language. The reflexive was expressed by a set of derived verbal forms

**Table 6.2 Proto-Semitic personal pronouns**

	Independent	Enclitic	
	Nominative	Nominative	Genitive-accusative
<i>Singular</i>			
1st com.	ʔana(:), ʔana:ku(:)	-ku(:)	-i:/-ya (gen.), -ni: (acc.)
2nd masc.	ʔanta(:)	-ta(:)	-ka(:)
2nd fem.	ʔanti(:)	-ti(:)	-ki(:)
3rd masc.	suʔa	-a	-su(:)
3rd fem.	siʔa	-at	-sa(:)/-si(:)
<i>Dual</i>			
1st com.	ʔ	-nuya:ʔ	-niya:ʔ (gen.), -naya:ʔ (acc.)
2nd com.	ʔantuma:	-tuma:	-kuma:/-kumay
3rd com.	suma:	-a:	-suma:/sumay
<i>Plural</i>			
1st com.	niḥnu(:)	-nu(:)	-ni(:) (gen.), -na(:) (acc.)
2nd masc.	ʔantum	-tum	-kum
2nd fem.	ʔantin	-tin	-kin
3rd masc.	sum	-u:	-sum
3rd fem.	sin	-a:	-sin



(see §3.3.5.2) and by means of a substantive meaning “person” or “body”: for example, *\*yanθ’ur napsa-su(:)* “he guarded his person” = “he guarded himself.”

### 3.3.4 Demonstrative and interrogative pronouns

Proto-Semitic had a determinative-relative pronoun, *\*ðu:* or *\*θu:* (the initial consonant is voiced in West Semitic, unvoiced in Akkadian), declinable for gender, number, and case (e.g., fem. sg. nom. *\*ða:tu*), always used as a bound form, with the meaning “the one of, that of, he/she of,” as in *\*θu: baytim* “the one of the house, he of the house”; *θu: ?anθ’uru* “the one whom I guarded.” It was commonly used in apposition to (and agreeing in case with) an antecedent: *\*baʕlum θu: baytim* “the lord(, the one) of the house,” *\*baytu baʕlim θi: ?anθ’uru* “the house of the lord(, the one) whom I guarded.”

In West Semitic, the determinative-relative pronoun entered into the formation of a set of demonstrative pronouns, such as masculine singular *\*ðin*, feminine singular *\*ða:* “this.” Another demonstrative base was *\*ʔVl(l)*, which appears in the plural of near demonstratives in West Semitic and as a far demonstrative (sg. and pl.) in Akkadian.

The third person pronouns were used as anaphoric or far demonstratives, as in *\*baytum suʔa* “that house,” “the aforementioned house.”

The evidence of the descendant languages for the interrogative pronouns is inconsistent. For “what?,” Akkadian and Ethiopic suggest a form *\*min-*, while Central Semitic has *\*ma:-/mah-*; for “who?” most languages have *\*man-*, while in Ugaritic and Canaanite the form is *\*mi:y-*. A common Semitic interrogative adjective is *\*ʔayy-* “which?”

### 3.3.5 Verbal morphology

Proto-Semitic had two basic indicative forms, which differed primarily in aspect. The forms were conjugated for person, gender, and number by means of prefixes and, in some instances, suffixes. Essentially, a *perfective*, punctive form prefix+R<sub>1</sub>R<sub>2</sub>V<sub>1</sub>R<sub>3</sub> contrasted with an *imperfective* form with gemination of the middle radical, prefix+R<sub>1</sub>aR<sub>2</sub>R<sub>2</sub>V<sub>2</sub>R<sub>3</sub>, as in *\*yanθ’ur* “he guarded” versus *\*yanab’θ’ar* “he guards.”

It seems likely that the bases of these forms were originally verbal adjectives, perfective (and passive) *\*naθ’ur-* “guarded” (note, for example, Akkadian *nas’ir* and Hebrew *nās’ur* with that meaning) and imperfective (and active) *\*naθ’θ’ar-* “guarding” (note the Common Semitic adjectival pattern R<sub>1</sub>aR<sub>2</sub>R<sub>2</sub>V<sub>2</sub>(: )R<sub>3</sub> for nouns expressing durative or habitual activity, as in *\*dayya(:)n-* “judge”). The pattern of the imperfective base, at least, was probably an inheritance from Common Afro-Asiatic (Greenberg 1952).

The perfective paradigm of the root *nθ’r* “to guard” is presented in (4):

(4)	<i>Singular</i>	<i>Dual</i>	<i>Plural</i>
<i>1st com.</i>	<i>?anθ’ur</i>		<i>nanθ’ur</i>
<i>2nd masc.</i>	<i>tanθ’ur</i>		<i>tanθ’uru:</i>
<i>2nd fem.</i>	<i>tanθ’uri:</i>		<i>tanθ’urna(:)</i>
<i>2nd com.</i>		<i>tanθ’ura:</i>	
<i>3rd masc.</i>	<i>yanθ’ur</i>		<i>yanθ’uru:</i>
<i>3rd fem.</i>	<i>tanθ’ur</i>		<i>yanθ’urna(:)</i>
<i>3rd com.</i>		<i>yanθ’ura:</i>	

Akkadian attests a third inflected indicative verbal form, called the *Perfect*, of the structure prefix+R<sub>1</sub>taR<sub>2</sub>V<sub>2</sub>R<sub>3</sub>, as in *\*yantaθ’ar*, which functions as a present perfect, “he has guarded.”

The existence of similar forms in other Afro-Asiatic branches, especially Berber and Cushitic, has been noted, and the suggestion offered that the Akkadian Perfect reflects a Proto-Semitic form that has been lost in West Semitic. But the Akkadian Perfect is formally identical with the perfective form of a Common Semitic – and ultimately also Common Afro-Asiatic – derived, mediopassive verbal class, and it seems likely that the former arose from the latter in an internal Akkadian development, perhaps under Sumerian influence.

In addition to these indicative forms, a number of modal forms may be posited. The imperative was confined to second-person forms, and had the shape of the perfective form without its prefixes, the initial consonant cluster being resolved by either prothesis or anaptyxis:

(5)	<i>Singular</i>	<i>Dual</i>	<i>Plural</i>
<i>2nd masc.</i>	nuθ'ur / ʔunθ'ur		nuθ'uru: / ʔunθ'uru:
<i>2nd fem.</i>	nuθ'uri: / ʔunθ'uri:		nuθ'urna(:) / ʔunθ'urna(:)
<i>2nd com.</i>		nuθ'ura: / ʔunθ'ura:	

By itself or with a prefixed asseverative particle *\*l(a)-*, the perfective form could be used injunctively, as a jussive, “let him guard.” Other modal forms, likewise related to or based on the perfective *\*yanθ'ur*, probably also occurred, but are difficult to reconstruct for Proto-Semitic with certainty, since they appear only in one or two of the branches of the family (e.g., *\*yanθ'ura*, with final *-a*; one or more “energetic” forms, such as *\*yanθ'uran(na)*).

Akkadian verbs in subordinate clauses are obligatorily (and usually redundantly) marked with an ending *-u* or *-ni* (probably < *\*-na*). It is likely that this mark of nominalization is of Proto-Semitic origin. In Central Semitic, the perfective verb with this ending came to be used as an imperfective form, replacing the inherited Proto-Semitic form *\*yanaθ'ar*.

### 3.3.5.1 Verbal nouns

Two verbal adjectives may be reconstructed for Proto-Semitic: (i) an *active participle* of the form  $R_1 a:R_2 iR_3$ , as in *\*na:θ'ir-* “guarding, who guards” (probably only for verbal roots expressing actions); and (ii) a *perfective adjective* of the form  $R_1 aR_2 VR_3$ , the meaning of which depended on the lexical meaning of the root: passive for transitive verbs (6A), resultative for intransitive active verbs (6B), and descriptive for stative verbs (6C):

- (6) A. *\*naθ'ur-* “guarded” (*n-θ'-r* “to guard”)  
 B. *\*waθib-* “having sat, seated” (*w-θ-b* “to sit, dwell”)  
 C. *\*hadaθ-* “new” (*ḥ-d-θ* “to be(come) new”)

The uninflected base of the verbal adjective could be combined with an enclitic nominative form of the person pronouns (see §3.3.3) to create a verbless (and thus tenseless) predication:

- (7) *\*naθ'ur-ta(:)* “you (masc. sg.) are/were guarded”  
*\*waθib-nu(:)* “we are/were seated”  
*\*hadaθ-at* “it (fem.) is/was new”

This construction is also attested in the oldest dialects of ancient Egyptian. In West Semitic the construction evolved in nonstative roots into an active, perfective verb, which began to replace the inherited form *\*yaR<sub>1</sub>R<sub>2</sub>VR<sub>3</sub>*; the development entailed a change of vocalism between the second and third radicals, to *\*a*: *\*naθ'arta(:)* “you (have) guarded,” *\*waθabnu(:)* “we (have) sat.”

It is likely that more than one pattern was used for the *infinitive*, including  $R_1 aR_2 a:R_3$  and  $R_1 iR_2 R_3$ , as in *\*naθ'a:r-* and *\*niθ'r-* “to guard, the guarding.”

### 3.3.5.2 *Derived verbs*

The examples of verbs that have been cited thus far in this chapter reflect the basic stem of the verbal root, which Semitists usually call the *G stem*, after German *Grundstamm*. From this basic stem are derived other stems, each with a fairly predictable semantic range vis-à-vis the basic stem; derivation is by means of one of a set of prefixed consonants or by means of the doubling of the second or third root consonant:

1. *The N stem*: With prefixed *n*, the perfective form of which was based on the basic verbal adjective of the root, turning the latter into a fientic verb; for most roots the semantic result is a passive: for example, G stem perfective *\*yapʔid* “he sought”; adjective *\*pakʔid* “sought,” N stem perfective *\*ya-n-pakʔid* “it became/was sought.”
2. *The C (causative) stem*: With prefixed *s* (originally, in all likelihood, a third-person pronoun serving as an agent), with causative force: *\*yusapʔid* “he caused (someone) to seek”; especially common with verbs of motion: G stem *\*yaʕliy* “he went up,” C stem *\*yusaʕliy* “he caused (something) to go up” = “he sent/took/brought/led up.”
3. *The D (doubled) stem*: Marked by gemination of the second radical, the effect of which was to increase the transitivity of the basic stem (Kouwenberg 1997); for stative verbal roots, the result is a factitive: G stem *\*yiħlal* “it was/became pure,” D stem *\*yuħhallil* “he purified”; for transitive verbal roots, the D stem is most often pluralic.

The G, C, and D stems could all be augmented by a prefixed *t*, associated with the notions of reciprocity, reflexivity, and the mediopassive; perfective forms of these may be illustrated by tG *\*yatpakʔid*; tD *\*yuthallV*; Ct (with *t* following the causative prefix *s*) *\*yustaʕliy*.

4. *The R stem*: With reduplication of the third radical (perfective *\*yVR<sub>1</sub>aR<sub>2</sub>R<sub>3</sub>iR<sub>3</sub>*, imperfective *\*yVR<sub>1</sub>aR<sub>2</sub>aR<sub>3</sub>R<sub>3</sub>aR<sub>3</sub>*, verbal adjective *\*R<sub>1</sub>aR<sub>2</sub>VR<sub>3</sub>R<sub>3</sub>* or *\*R<sub>1</sub>aR<sub>2</sub>R<sub>3</sub>VR<sub>3</sub>*). This stem is likely also to be reconstructed for Proto-Semitic; further investigation is required to elucidate the semantics of the stem, which is only vestigially preserved in most of the languages (apart from Arabic). It tends to involve description of physical qualities or states.

As noted above, most of these derived verbal stems have analogues elsewhere in Afro-Asiatic (Lieberman 1986).

### 3.3.6 *Compounds*

The Semitic languages, and presumably Proto-Semitic as well, exhibit remarkably few instances of compounding in either the nominal or the verbal morphology.

### 3.3.7 *Numerals*

The Proto-Semitic cardinals 1 through 10 were declined like singular nouns, except for 2 which was declined as a dual. They occurred in both masculine and feminine forms. In an unusual – and still unexplained – syntactic phenomenon reflected in nearly all the descendant languages, for the numbers from 3 to 10 the masculine form of the cardinal was used when the counted item was a feminine noun, and the feminine form of the cardinal with masculine nouns. The basic forms of the cardinals were as follows; feminine forms were marked with the addition of *\*(a)t*.

- (8) 1 \*ʔaħad-  
 2 \*θin-/θn-  
 3 \*θala:θ-  
 4 \*ʔarbaʕ  
 5 \*xamis-  
 6 \*sidθ-  
 7 \*sabʕ-  
 8 \*θama:niy-  
 9 \*tisʕ-  
 10 \*ʕaṭr-

The cardinal 20 is the dual of 10, \*ʕaṭra:-; the other tens have the appearance of being duals (plurals in Central Semitic) of the corresponding units: for example, \*θala:θa:- 30. “Hundred” is \*miʔ(a)t-; higher numbers are difficult to reconstruct with certainty (“thousand” is West Semitic \*ʔalp-, East Semitic \*liʔm-; for “10,000; myriad” West Semitic has forms derived from the root *r-b-b* “to be(come) much, many”).

Unlike the cardinals, the ordinals are generally constructed on a single pattern; the pattern, however, varies from language to language (e.g., the pattern  $R_1a:R_2iR_3$  in classical Arabic and classical Ethiopic, as in *θa:liθ-* “third,” *ra:biʕ* “fourth,” *xa:mis-* “fifth”), and so cannot be reconstructed to the proto-language.

### 3.4 Syntax

#### 3.4.1 Word order

Proto-Semitic was probably a VSO (Verb–Subject–Object) language. This is true of the earliest forms of most West Semitic languages. Most dialects of Akkadian were rigidly SOV, but word order in poetic texts is much freer; further, early Akkadian personal names composed of a subject and a verb are frequently VS, as in *Iddin-Si:n* “[the god] Sin has given [a child].” The normal SOV order of Akkadian is undoubtedly due to Sumerian influence.

Modifiers, including adjectives, genitives, and relative clauses, follow their head noun.

#### 3.4.2 Clitics

Semitic is characterized by a number of prefixed monosyllabic relational particles, including the coordinating conjunction \*wa- “and,” the asseverative particle \*l(a)-, and, in West Semitic, the prepositions \*ba- “in,” \*la- “to, for,” and \*ka- “like” (in early Akkadian dialects, too, proclitic forms of certain prepositions are attested: *an-*, *in-*, and *el-* for *ana* “to, for,” *ina* “in,” *eli* “on,” respectively).

Much of the personal pronoun system consists of suffixed morphemes, as in \*la-su(:) “to him” \*baytu-su(:) “his house,” \*ʔanθʕur-su: “I guarded him”; two of these suffixes, denoting indirect and direct objects in sequence, could appear on finite verbs: \*yantinu:-ni:-su(:) “they (masc.) gave me it (masc.)” (Gensler 1998).

The enclitic particle \*-ma(:) served to topicalize the word to which it was attached; in Akkadian and in several modern Ethiopian languages it also developed into a coordinating conjunction. The Proto-Semitic status of other enclitic forms attested in the various languages remains to be investigated, as, for example, \*-mi(:), an emphasizing particle in Northwest Semitic, but a marker of direct speech in Akkadian.

### 3.4.3 Coordination

The essential Proto-Semitic coordinating conjunction was the proclitic particle *\*wa-*, which was used to connect words, clauses (including connecting main clauses to preceding subordinate clauses), and sentences. Unclear as yet are the Proto-Semitic status and functions of the Central Semitic proclitic clause connector *\*pa-*, meaning, *inter alia*, “and then, and so” (for the very common Akkadian enclitic clause connector *-ma*, see §3.4.2).

### 3.4.4 Subordination

Subordinate clauses are less common in Semitic than in some languages, simple coordination usually being preferred. Nevertheless a few subordinating conjunctions may be reconstructed. A general subordinating conjunction was *\*ki:* (also *\*ki(:)ma(:)*), attested in a number of the descendant languages with the meanings “when, because, that.” Several words functioned both as prepositions and as conjunctions: for example *\*ʕad(ay)* “up to, until.” Certain bound-form nouns could also function as the equivalent of conjunctions, as in *\*yawma ʔanθʾuru* “the day (= when) I guarded” (with the accusative of *\*yawm-* “day” used adverbially; also with a preposition: *\*ba-yawmi ʔanθʾuru* “on the day I guarded”).

Subordination was also expressed by means of infinitives, especially with the preposition “in” for circumstance and the preposition “to, for” for purpose and result: *\*ba-naθʾa:ri-su(:)* “in his guarding” = “while he guards/guarded” (or “while guarding him”); *\*la-naθʾa:risu(:)* “for his guarding” = “(in order) that he guard” (or “(in order) to guard him”).

### 3.4.5 Verbless clauses

While a verb “to be, become” can be reconstructed for Proto-Semitic, namely, *\*hawaya*, the notion “to be” was not normally expressed and verbless clauses were a common feature. With a nominal or pronominal subject, the predicate could be (i) adverbial (adverb or prepositional phrase: “he [is] in the house”; “my sister [is] here”); (ii) adjectival, in which case the construction described in §3.3.5.1 was used; or (iii) nominal, with both subject and predicate in the nominative case (*\*ʔimmu-su(:) baʕlatu-nu(:)* “his mother is our mistress”). A third-person pronoun in apposition to the subject could be included in the clause, probably either before or after the predicate (*\*ʔimmu-su(:) siʔa baʕlatu-nu(:)* or *\*ʔimmu-su(:) baʕlatu-nu(:) siʔa* “his mother [she] is our mistress”; the pronoun is traditionally said to function as a copula in such instances).

For existential sentences, the phrase “in it,” *\*ba-su(:)*, with the meaning “there is” may perhaps be reconstructed to Proto-Semitic, since “in it” is so used in classical Ethiopic (*bo* < *\*ba-hu:*, *botu*), various Arabic dialects (*fi:* < *fi:-hi*), and Akkadian (in which by the time of the earliest dialect the construction had developed into a finite verb, *\*basa:ʔum* “to be present, on hand”). A particle *\*yiθ-* “there is/are” can be reconstructed for Central Semitic; it is cognate with an Eblaite infinitive, *i-ša-wu* = /yVθa:wu(m)/, known from a lexical text, where it is equated with Sumerian A/AN.GÁL “be.”

For “to have” Akkadian attests the irregular verb *išûm*, of uncertain etymology (connected by some scholars with *\*yiθ-*, etc., cited in the preceding paragraph, but the few Old Akkadian writings of *išûm* suggest that the middle radical was not *\*θ*). In West Semitic, however, possession is expressed with the dative preposition either as the predicate of a verbless clause or governed by the verb “to be” (e.g., “the lord has a house” by “[a] house [is] to [the] lord”).

### 3.5 Lexicon

Additional research is needed before the percentage of Proto-Semitic vocabulary inherited from common Afro-Asiatic can be estimated.

A few Common Semitic words resemble Indo-European words or roots: \**θawr-* beside PIE \**tauro-* “bull”; \**k’arn-* beside PIE \**kr̥-n-* “horn”; and \**ʕaθtar-* “morning/evening star” beside PIE \**h<sub>2</sub>ste:r-* “star”; the significance of these similarities is unclear. Other words show by their divergent reflexes in the descendant languages, as well as by their unusual patterns, that they were not native to Common Semitic, such as \**b/par<sup>d</sup>zil-* “iron,” \**ʔan(n)a(:)k-* “lead,” \**ʔuk’niy-* “lapis lazuli” (with the last compare Hittite *ku(wa)nna-*, Greek *kūānos*).

## 4. READING LIST

Surveys of the Afro-Asiatic languages and of common Afro-Asiatic features are given in Greenberg 1955b, 1970; Hodge 1971; Sasse *et al.* 1981; Hetzron 1987; D. Cohen 1988; Diakonoff 1988; Petráček 1988; and Hayward 2000. Important works dealing with specific features include Rössler 1950; Greenberg 1952, 1955a; Lieberman 1986; Voigt 1987a; Zaborski 1995. A pioneering treatment of common Afro-Asiatic vocabulary is M. Cohen 1947; the recent dictionary of Orel and Stolbova 1995 has been widely criticized in scholarly reviews.

A recent compendium in which all of the major Semitic languages are covered is Hetzron 1997. The fundamental reference work on Semitic grammar is Brockelmann 1908–1913; other general works on Semitic are Nöldeke 1904, 1910; Bergsträsser 1928; Gray 1934; Kurylowicz 1973; Moscati 1964; Garbini and Durand 1994; Lipiński 1997; Bennett 1998; Stempel 1999; Kienast 2001.

The internal classification or subgrouping of the Semitic languages has been a subject of much discussion, and a consensus has not been reached. The subgrouping presented here is that proposed by Hetzron 1974, 1976; as modified in Huehnergard 1991, 2002; Nebes 1994; and Porkhomovsky 1997.

The current understanding of the consonantal phonology of Proto-Semitic is the result of the work of several scholars, but especially Steiner 1977, 1982; Faber 1984, 1985, 1989; and Voigt 1987b.

The Semitic root and pattern system is discussed recently in McCarthy 1979; Goldenberg 1994; and Fox 2003. The pronominal systems are considered in Barth 1913; Rundgren 1955; Castellino 1962; Pennacchietti 1968; nominal inflection, *inter alia*, in Diem 1975; Voigt 1987a; Ratcliffe 1998. Of the many important studies of the Semitic verbal system only a very small selection may be noted here: Rundgren 1959; Retsö 1989; Tropper 1990.

Works on comparative and historical Semitic syntax continue to be few, but mention should be made of D. Cohen 1984; Khan 1988; and Gensler 1998.

The Common Semitic lexicon was considered in an important series of articles in Fronzaroli 1964–1971. A complete Semitic etymological dictionary does not exist; the fascicles of the *Dictionnaire des racines sémitiques* (Cohen 1970–) that have thus far appeared cover about one-third of the Semitic roots.

### Bibliography

- Barth, J. 1894. *Die Nominalbildung in den semitischen Sprachen* (2nd edition). Leipzig: Hinrichs.  
 ———. 1913. *Die Pronominalbildung in den semitischen Sprachen*. Leipzig: Hinrichs.  
 Bennett, P. 1998. *Comparative Semitic Linguistics: A Manual*. Winona Lake, IN: Eisenbrauns.  
 Bergsträsser, G. 1928. *Einführung in die semitischen Sprachen: Sprachproben und grammatische Skizzen*. Munich: Max Hueber. Translated (with notes, bibliography and an appendix on the



- scripts) by P. Daniels as *Introduction to the Semitic Languages. Text Specimens and Grammatical Sketches*. Winona Lake, IN: Eisenbrauns, 1983.
- Brockelmann, C. 1908–1913. *Grundriss der vergleichenden Grammatik der semitischen Sprachen* (2 vols.). Berlin: von Reuther.
- Castellino, G. 1962. *The Akkadian Personal Pronouns and Verbal System in the Light of Semitic and Hamitic*. Leiden: Brill.
- Cohen, D. 1970–. *Dictionnaire des racines sémitiques ou attestées dans les langues sémitiques*. Paris/The Hague: Mouton; Louvain-la-Neuve: Peeters.
- . 1984. *La phrase nominale et l'évolution du système verbal en sémitique: études de syntaxe historique*. Collection Linguistique 72. Paris: Société de Linguistique de Paris.
- . (ed.). 1988. *Les langues chamito-sémitiques*. In Jean Perrot (general ed.), *Les langues dans le monde ancien et moderne*, part 3. Paris: CNRS Editions.
- Cohen, M. 1947. *Essai comparatif sur le vocabulaire et la phonétique du chamito-sémitique*. Bibliothèque de l'Ecole des Hautes Etudes 248. Paris: Champion.
- Diakonoff, I. 1988. *Afrasian Languages*. Moscow: Nauka.
- . 1998. "The earliest Semitic society: linguistic data." *Journal of Semitic Studies* 43:209–219.
- Diem, W. 1975. "Gedanken zur Frage der Mimation und Nunation in den semitischen Sprachen." *Zeitschrift der Deutschen Morgenländischen Gesellschaft* 125:239–258.
- Faber, A. 1984. "Semitic sibilants in an Afro-Asiatic context." *Journal of Semitic Studies* 29:189–224.
- . 1985. "Akkadian evidence for Proto-Semitic affricates." *Journal of Cuneiform Studies* 37:101–107.
- . 1989. "On the nature of Proto-Semitic \*l." *Journal of the American Oriental Society* 109:33–36.
- Fleming, H. 1999. "Afroasiatic internal taxonomy: new methods and better results." Paper presented at the 26th annual meeting of the North American Conference on Afroasiatic Linguistics, Baltimore, April 1999.
- Fleming, H., A. Yilma, A. Mitiku, et al. 1992. "Ongota (or) Birale: a moribund language of Gemu-Gofa (Ethiopia)." *Journal of Afroasiatic Languages* 3:181–225.
- Fox, J. 1998. "Isolated nouns in the Semitic languages." *Zeitschrift für Althebraistik* 11:1–31.
- . 2003. *Semitic Noun Patterns*. Harvard Semitic Studies. Winona Lake, IN: Eisenbrauns.
- Fronzaroli, P. 1964–1971. "Studie sul lessico commune semitico." *RANL* 19 (1964):155–172 (I), 243–280 (II); 20 (1965):135–150 (III), 246–269 (IV); 23 (1968):267–303 (V); 24 (1969):285–320 (VI); 26 (1971):603–642 (VII).
- Garbini, G. and O. Durand. 1994. *Introduzione alle lingue semitiche*. Brescia: Paedea.
- Gensler, O. 1997. "Reconstructing quadriliteral verb inflection: Ethiopic, Akkadian, Proto-Semitic." *Journal of Semitic Studies* 42:229–257.
- . 1998. "Verbs with two object suffixes: a Semitic archaism in its Afroasiatic context." *Diachronica* 15:231–284.
- Goldenberg, G. 1994. "Principles of Semitic word-structure." In G. Goldenberg and S. Raz (eds.), *Semitic and Cushitic Studies*, pp. 29–64. Wiesbaden: Harrassowitz.
- Gray, L. 1934. *Introduction to Semitic Comparative Linguistics*. New York: Columbia University.
- Greenberg, J. 1950. "The patterning of root morphemes in Semitic." *Word* 6:162–181.
- . 1952. "The Afro-Asiatic (Hamito-Semitic) present." *Journal of the American Oriental Society* 72:1–9.
- . 1955a. "Internal *a*-plurals in Afroasiatic (Hamito-Semitic)." In J. Lukas (ed.), *Afrikanistische Studien*, pp. 198–204. Berlin: Akademie-Verlag.
- . 1955b. *Studies in African Linguistic Classification*. New Haven.
- . 1970. *The Languages of Africa* (3rd edition). Indiana University Research Center for Language Sciences, Publication 25. Bloomington: Indiana University Press.
- Hayward, R. 2000. "Afroasiatic." In B. Heine and D. Nurse (eds.), *African Languages: An Introduction*, pp. 74–98. Cambridge University Press.
- Hetzron, R. 1974. "La division des langues sémitiques." In A. Caquot and D. Cohen (eds.), *Actes du premier congrès international de linguistique sémitique et chamito-sémitique, Paris 16–19 juillet 1969*, pp. 181–194. The Hague/Paris: Mouton.
- . 1976. "Two principles of genetic reconstruction." *Lingua* 38:89–108.
- . 1980. "The limits of Cushitic." *Sprache und Geschichte in Afrika* 2:7–126.
- . 1987. "Afroasiatic languages." In B. Comrie (ed.), *The World's Major Languages*, pp. 645–53. New York: Oxford University Press.

- Hetzron, R. (ed.). 1997. *The Semitic Languages*. London/New York: Routledge.
- Hodge, C. (ed.). 1971. *Afroasiatic: A Survey*. The Hague/Paris: Mouton. Reprinted from T. A. Sebeok (ed.), *Current Trends in Linguistics*. Vol. VI: *Linguistics in South West Asia and North Africa*, pp. 237–661. The Hague/Paris: Mouton, 1970.
- Huehnergard, J. 1991. "Remarks on the classification of the Northwest Semitic languages." In J. Hoftijzer and G. van der Kooij (eds.), *The Balaam Text from Deir 'Alla Re-evaluated. Proceedings of the International Symposium held at Leiden 21–24 August 1989*, pp. 282–293. Leiden: Brill.
- . 2002. "Comparative Semitic linguistics." *Israel Oriental Studies* 20 (S. Izre'el, ed., *Semitic Linguistics: The State of the Art at the Turn of the Twenty-First Century*):213–245.
- Khan, G. 1988. *Studies in Semitic Syntax*. London Oriental Studies 38. Oxford: Oxford University Press.
- Kienast, B. 2001. *Historische semitische Sprachwissenschaft*. Wiesbaden: Harrassowitz.
- Kouwenberg, N. 1997. *Gemination in the Akkadian Verb*. Studia Semitica Neerlandica. Assen: Van Gorcum.
- Kurylowicz, J. 1973. *Studies in Semitic Grammar and Metrics*. London: Curzon Press. (Translation and revision of *L'apophonie en sémitique* 1961.)
- Lieberman, S. 1986. "The Afro-Asiatic background of the Semitic N-stem: toward the origins of the Semitic and Afro-Asiatic verb." *Bibliotheca Orientalis* 43:577–628.
- Lipiński, E. 1997. *Semitic Languages: Outline of a Comparative Grammar*. Orientalia Lovaniensia Analecta 80. Louvain-la-Neuve: Peeters/Departement Oosterse Studies.
- McCarthy, J. 1979. *Formal Problems in Semitic Phonology and Morphology*. Ph.D. dissertation, MIT. (= New York: Garland, 1985.)
- Militarev, A. 1994. "Home for Afrasian: African or Asian? Areal linguistic arguments." In C. Griefenow-Menis and R. Voigt (eds.), *Cushitic and Omotic Languages: Proceedings of the Third International Symposium, Berlin, March 17–19, 1994*, pp. 13–32. Cologne: Rüdiger Köppe.
- Moscatti, S. (ed.). 1964. *An Introduction to the Comparative Grammar of the Semitic Languages: Phonology and Morphology*. Wiesbaden: Harrassowitz.
- Nebes, N. 1994. "Zur Form der Imperfektbasis des unvermehrten Grundstammes im Altsüdarabischen." In W. Heinrichs and G. Schoeler (eds.), *Festschrift Ewald Wagner zum 65. Geburtstag*. Vol. I, *Semitische Studien unter besonderer Berücksichtigung der Südsemitistik*, pp. 59–81. Beirut/Stuttgart: Steiner.
- Nöldeke, T. 1904. *Beiträge zur semitischen Sprachwissenschaft*. Strasburg: Trübner.
- . 1910. *Neue Beiträge zur semitischen Sprachwissenschaft*. Strasburg: Trübner.
- O'Connor, M. 1996. "The Berber Scripts." In P. Daniels and W. Bright (eds.), *The World's Writing Systems*, pp. 112–116. New York/Oxford: Oxford University Press.
- Orel, V. and O. Stolbova. 1995. *Hamito-Semitic Etymological Dictionary: Materials for a Reconstruction*. Leiden: Brill.
- Pennacchietti, F. 1968. *Studi sui pronomi determinativi semitici*. Pubblicazioni del seminario de semitistica, ricerche 4. Naples: Istituto Orientale.
- Petráček, K. 1988. *Altägyptisch, Hamitosemitisch und ihre Beziehungen zu einigen Sprachfamilien in Afrika und Asien*. Praha: Univerzita Karlova.
- Porkhomovsky, V. 1997. "Modern South Arabian languages from a Semitic and Hamito-Semitic perspective." *Proceedings of the Seminar for Arabian Studies* 27:219–223.
- Ratcliffe, R. 1998. *The Broken Plural Problem in Arabic and Comparative Semitic: Allomorphy and Analogy in Non-Concatenative Morphology*. CILT 168. Amsterdam: John Benjamins.
- Retsö, J. 1989. *Diathesis in the Semitic Languages: A Comparative Morphological Study*. Studies in Semitic Languages and Linguistics 14. Leiden: Brill.
- Rössler, O. 1950. "Verbalbau und Verbalflexion in den semitohamitischen Sprachen. Vorstudien zu einer vergleichenden semitohamitischen Grammatik." *Zeitschrift der Deutschen Morgenländischen Gesellschaft* 100:461–514. Translated as "The structure and inflection of the verb in the Semito-Hamitic Languages: preliminary studies for a comparative Semito-Hamitic grammar," in Y. Arbeitman and A. Bomhard (eds.), *Bono Homini Donum: Essays in Historical Linguistics in Memory of J. Alexander Kerns*, pp. 679–748. Amsterdam: John Benjamins, 1981.
- Rundgren, F. 1955. *Über Bildungen mit s/š- und n-t-Demonstrativen im Semitischen*. Uppsala: Almqvist and Wiksell.



- . 1959. *Intensiv und Aspektkorrelation*. Uppsala: Lundequistska; Wiesbaden: Otto Harrassowitz.
- Sasse, H.-J. 1984. "Case in Cushitic, Semitic and Berber." In J. Bynon (ed.), *Current Progress in Afro-Asiatic Linguistics: Papers of the Third International Hamito-Semitic Congress*, pp. 111–126. Amsterdam/Philadelphia: John Benjamins.
- Sasse, H.-J., L. Störk, and E. Wolff. 1981. "Afroasiatisch." In B. Heine, T. Schadeberg, and E. Wolff (eds.), *Die Sprachen Afrikas*, pp. 129–262. Hamburg: Helmut Buske.
- Speiser, E. 1952. "The 'Elative' in West-Semitic and Akkadian." *Journal of Cuneiform Studies* 6:81–92.
- Steiner, R. 1977. *The Case for Fricative-Laterals in Proto-Semitic*. American Oriental Series, 59. New Haven: American Oriental Society.
- . 1982. *Affricated Šade in the Semitic Languages*. The American Academy for Jewish Research Monograph Series 3. New York.
- Stempel, R. 1999. *Abriß einer historischen Grammatik der semitischen Sprachen*. Nordostafrikanisch/Westasiatische Studien 3. Frankfurt am Main: Peter Lang.
- Testen, D. 1998. *Parallels in Semitic Linguistics: The Development of Arabic la- and Related Semitic Particles*. Studies in Semitic Languages and Linguistics 26. Leiden: Brill.
- Tropper, J. 1990. *Der ugaritische Kausativstamm und die Kausativbildungen des Semitischen: Eine morphologisch-semantische Untersuchung zum Š-Stamm und zu den umstrittenen nichtsibilantischen Kausativstämmen des Ugaritischen*. Abhandlungen zur Literatur Alt-Syrien-Palästinas 2. Münster: Ugarit Verlag.
- Voigt, R. 1987a. "Derivatives und flektives *t* im Semitohamitischen." In H. Jungraithmayr and W. Müller (eds.), *Proceedings of the Fourth International Hamito-Semitic Congress, Marburg, 20–22 September, 1983*, pp. 85–107. CILT, 44. Amsterdam/Philadelphia: John Benjamins.
- . 1987b. "Die Personalpronomina der 3. Personen im Semitischen." *Welt des Orients* 18:49–63.
- Zaborski, A. 1995. "Problems of Hamitosemitic pronouns." *Sprawozdania z Posiedzeń Komisji Naukowych* 38:59–62.

# Ancient Egyptian and Coptic

ANTONIO LOPRIENO

## 1. HISTORICAL AND CULTURAL CONTEXTS

### 1.1 Introduction

Ancient Egyptian and its latest historical stage, Coptic, represent a branch of the language family variously called Afro-Asiatic, Hamito-Semitic, or Semito-Hamitic (see Ch. 6; also Diakonoff 1965; Hodge 1971; Zaborski 1992:36–37). The Afro-Asiatic family comprises, from antiquity to the present time, a number of languages spoken and written in the eastern Mediterranean world, in northern Africa, and in western Asia. These languages are characterized by the following linguistic features: (i) a preference for fusional or flecational morphology; (ii) the presence of bi- and triconsonantal lexical roots, capable of being variously inflected; (iii) a consonantal system displaying a series of pharyngealized or glottalized phonemes, called “emphatics,” alongside the voiced and the voiceless series; (iv) a vocalic system originally limited to the three vowels *a*, *i*, *u*; (v) a nominal feminine suffix *-at*; (vi) a case system consisting of no more than two or three cases; (vii) a nominal prefix *m-*; (viii) an adjectival suffix *-ī* (called *nisba*, the Arabic word for “relation”); (ix) an opposition between prefix conjugation for verbal actions and suffix conjugation for verbal states; (x) a conjugation pattern singular first person *’a-*, second person *ta-*, third person masculine *ya-*, feminine *ta-*, first person plural *na-*, with additional suffixes in the other persons.

The individual branches of the Afro-Asiatic family are as follows:

1. *Egyptian*: Within Afro-Asiatic, Egyptian shows the closest relations to Semitic and Berber.
2. *Semitic*: The largest family of languages within Afro-Asiatic (Hetzron 1992:412–417).
3. *Berber*: A group of related languages and dialects currently spoken in competition with Arabic by a few million speakers in northern Africa from the Atlantic coast to the oasis of Siwa and from the Mediterranean Sea to Mali and Niger (Willms 1980). Modern Berber is most probably the descendant of the Libyan languages spoken in the same area in antiquity.
4. *Cushitic*: A group of languages spoken by 15 million people in eastern Africa, from the Egyptian border in northeast Sudan to Ethiopia, Djibouti, Somalia, Kenya, and northern Tanzania (Sasse 1992:326–330).
5. *Chadic*: A group which comprises about 140 languages spoken by more than 30 million speakers in sub-Saharan Africa around Lake Chad (Newman 1992:253–254).
6. *Omototic languages*: Spoken by approximately one million speakers along the Omo River and north of Lake Turkana in southwest Ethiopia, formerly thought to represent the western branch of Cushitic (Fleming 1976:34–53).

The productive history of Egyptian, which spans from 3000 BC to AD 1300, can be divided into two main stages, characterized by a major change from synthetic to analytic patterns in the nominal syntax and the verbal system (Junge 1985). Each of these two stages can be further subdivided into three different phases, which affect mainly the sphere of graphemics (Kammerzell 1995).

## 1.2 Earlier Egyptian

This is the language of all written texts from 3000 to 1300 BC, surviving in formal religious texts until the third century AD. Its main phases are as follows:

1. *Old Egyptian* (Edel 1955–1964): The language of the Old Kingdom and of the First Intermediate Period (3000–2000 BC). The main documents of this stage of the language are royal rituals such as the “Pyramid Texts,” and funerary texts, especially “autobiographies” containing accounts of individual achievements inscribed in the rock tombs of the administrative elite.

2. *Middle Egyptian* (Gardiner 1957): Also termed *Classical Egyptian*, from the Middle Kingdom to the end of Dynasty XVIII (2000–1300 BC). Middle Egyptian is the language of classical Egyptian literature, which comprises ritual texts, for example the “Coffin Texts” inscribed on the sarcophagi of the administrative elite; wisdom texts that convey the educational and professional expectations of contemporary Egyptian society, for example the “Instructions of the Vizier Ptahhotep”; narratives relating adventures of a specific hero and representing the vehicle of individual, as opposed to societal, concerns (the most famous specimen of this genre is the “Tale of Sinuhe”); hymns and poetical texts with religious contents, written in praise of a god or of the king. Besides literary texts, the Middle Egyptian corpus comprises administrative documents, for example the Kahun papyri, and historical records.

3. *Traditional Egyptian*: The language of religious texts (rituals, mythology, hymns) from the New Kingdom to the end of Egyptian civilization. Late Middle Egyptian coexisted with later Egyptian for more than a millennium in a situation of diglossia (Vernus 1996:560–564). From a grammatical point of view, Late Middle Egyptian maintains the linguistic structures of the classical language, but on the graphemic side, especially in the Greco-Roman period, it shows an enormous expansion of the set of hieroglyphic signs.

Earlier Egyptian is characterized by a preference for synthetic grammatical structures: it displays a full set of morphological suffixes indicating gender and number; it exhibits no definite article; it maintains the VSO order in verbal formations:

- (1)  $\underline{s}dm$                        $zh^3w$        $n$        $sb^3.t-j$   
       listen.PROSP.    scribe    to    teaching.FEM.-me  
       “May the scribe listen to my teaching”

## 1.3 Later Egyptian

Later Egyptian is documented from Dynasty XIX down to the Middle Ages (1300 BC–AD 1300). Its main phases are as follows:

1. *Late Egyptian* (1300–700 BC): The language of written records from the second part of the New Kingdom (Černý and Groll 1984; Junge 1996; Neveu 1996). It conveys the rich entertainment literature of Dynasty XIX, consisting of wisdom texts and tales, but also of new literary genres, such as mythology or love poetry. Late Egyptian was also the vehicle of Ramesside bureaucracy as documented by the archives of the Theban necropoleis or by

school texts. Late Egyptian is not a wholly homogeneous linguistic reality; rather, the texts of this phase of the language show various degrees of interference with classical Middle Egyptian, with the tendency of older or more formal texts, such as historical records or literary tales, to display a higher number of borrowings from the classical language, as opposed to later or administrative texts, where Middle Egyptian forms are much rarer (Winand 1992:3–25).

2. *Demotic* (seventh century BC to fifth century AD): The language of administration and literature from the pharaonic Late Period to late antiquity (Johnson 1991). While grammatically close to Late Egyptian, it differs from it radically in its graphic system. Important texts in Demotic are narrative cycles and moral instructions.

3. *Coptic* (fourth to fourteenth century AD): The language of Christian Egypt, written in a variety of Greek alphabet with the addition of six or seven Demotic signs to indicate Egyptian phonemes absent from Greek (Lambdin 1983). As a spoken, and gradually also as a written language, Coptic was superseded by Arabic from the ninth century onward, but it survives to the present time as the liturgical language of the Christian church of Egypt and in a few linguistic traces it left in spoken Egyptian Arabic (Vittmann 1991).

Besides displaying a number of phonological evolutions, later Egyptian develops analytic features: suffixal markers of morphological oppositions are dropped and functionally replaced by prefixal indicators; the demonstrative “this” and the numeral “one” evolve into the definite and the indefinite article; periphrastic patterns in the order SVO supersede older verbal formations (Hintze 1950):

- (2) mare-p-sah      sôtm    e-ta-sbô  
       OPT.-the.scribe    listen    to-the.FEM.my-teaching  
       “May the scribe listen to my teaching”

## 1.4 Dialects

Owing to the centralized nature of the political and cultural models underlying the evolution of Ancient Egyptian society, there is hardly any evidence of dialect differences in pre-Coptic Egyptian (Osing 1975; Lüddeckens 1975). However, while the writing system probably originated in the south of the country, the origins of the linguistic type represented by earlier Egyptian are to be seen in Lower (northern) Egypt, around the city of Memphis, which was the capital of the country during the Old Kingdom. The linguistic origins of Later Egyptian lie in Upper (southern) Egypt, in the region of Thebes, the cultural, religious, and political center of the New Kingdom (Zeidler 1992:208; Schenkel 1993:148).

Coptic is known through a variety of dialects differing mostly in the graphic rendition of Egyptian phonemes, and to a lesser extent also in morphology and lexicon. The most important dialect is *Sahidic* (from Arabic *al-šaʿīd* “Upper Egypt”), the written standard of the Theban area. Sahidic is the first dialect of Coptic literature. *Bohairic* (from Arabic *al-buḥayra* “Lower Egypt”), the dialect of Alexandria, eventually became the language of the liturgy of the Coptic church. Other important dialects of Coptic literature are *Akhmimic* from the city of Akhmim (Greek Panopolis) in Upper Egypt; *Subakhmimic*, also called *Lycopolitan*, spoken in the area of Asyut (Greek Lycopolis) in Middle Egypt; and *Fayyumic*, the variety of Coptic from the oasis of Fayyum, in the upper western corner of the Nile Valley (Kasser 1991b).




## 2. WRITING SYSTEMS

### 2.1 Principles

The basic graphic system of the Egyptian language from about 3000 BC to the first centuries of our era is composed of *hieroglyphs* (Fischer 1977). This term is the Greek counterpart to the Egyptian expression *mdw.w-nṯr* “god’s words.” Hieroglyphs were used primarily for monumental purposes, their main material support being stone or, less frequently, papyrus. For cursive uses the hieroglyphic system developed two handwritten varieties, called *Hieratic*, documented from the Old Kingdom through the third century AD, and *Demotic*, from the seventh century BC to the fifth century AD. Beginning in Hellenistic times, hieroglyphs and their manual varieties were gradually superseded by alphabetic transcriptions of words, and then of whole texts, inspired by the Greek alphabet with the addition of Demotic signs to render Egyptian phonemes unknown to Greek. The final result of this process is the emergence of *Coptic*. Unlike other writing systems of the Ancient Near East, for example Mesopotamian cuneiform, hieroglyphs were never used to write down any language other than Egyptian, except for their later adoption in Nubia for the writing of Meroitic (third century BC to fourth century AD; Wenig 1982). However, the Proto-Sinaitic inscriptions of the second millennium BC (Givón 1982) show that hieratic signs may have inspired the shape of Northwest Semitic consonantal signs. As for Demotic, some of its sign groups were adopted and phonetically reinterpreted in Meroitic.

Because of the formal similarities with Egyptian hieroglyphs, the term “hieroglyph” has also been applied to the writing system of Luwian, an Anatolian language related to cuneiform Hittite, spoken and written during the Late Bronze and Iron Ages (between c. 1500–700 BC) in southern and southwestern Anatolia and northern Syria: hence the misleading designation “Hittite hieroglyphs” by which they are often referred to (Gelb 1963:81–84).

The Egyptian hieroglyphs constitute a variable set of graphemes, ranging from about 1000 in the Old Kingdom (third millennium BC) down to approximately 750 in the classical language (second millennium BC), then increasing to many thousands during the Ptolemaic and Roman rule in Egypt, from the third century BC to the second century AD. They are pictographic signs representing entities and objects, such as gods or categories of people, animals, parts of the human or animal body, plants, astronomical entities, buildings, furniture. But these pictograms are not organized within a purely ideographic system; rather, they represent a combination of phonological and semantic principles (Schenkel 1984). An Egyptian word usually consists of two components:

1. A sequence of phonograms, each of which represents a sequence of one, two, or three consonantal phonemes; hence their label as *monoconsonantal* (such as  /m/), *biconsonantal* (such as  /p-r/), or *triconsonantal* signs (such as  /h-t-p/). Phonograms convey a substantial portion of the phonological structure of the word: normally all the consonants, less regularly the (semiconsonantal) glides *j* and *w*. The vowels remain unexpressed in the writing. Bi- and triconsonantal signs may be accompanied by other phonograms, mostly monoconsonantal, which spell out one or two of their phonemes, allowing in this way a more immediate interpretation of the phonological sequence; these signs are called *phonetic complements*.

Egyptian writing displays a set of twenty-four signs of monoconsonantal value (see Table 7.1). Although these cover almost completely the inventory of consonants and glides – an exception being the liquid /l/, conveyed by the graphemes <n>, <r>, or <n+r> – the

**Table 7.1 Monoconsonantal hieroglyphic signs**

Sign	Entity depicted	Transliteration	Phonological value
	vulture	ʾ (aleph)	earlier /R/ > later /ʔ/
	flowering reed	j (yod)	earlier /y/ > later /ʔ/
	(1) two reed flowers	jj or y	/y/
	(2) two strokes		
	human forearm	ʿ (ayin)	earlier /d/ > later /ʕ/ as in Arabic kaʿba
	quail chick	w (waw)	/w/
	foot	b	/b/
	stool	p	/p/
	horned viper	f	/f/
	owl	m	/m/
	water	n	/n/
	human mouth	r	/r/
	reed shelter	h	/h/ as in English he
	twisted wick	ḥ	/ħ/ as in Arabic aḥmad
	placenta	ḫ	/x/ as in German Buch
	animal's belly	l	/ç/ as in German ich
	bolt	z	/θ/ as in English think
	folded cloth	s	/s/
	pool or lake	š	/ʃ/ as in English she
	hill slope	q	/q/ as in Arabic qurʾān
	basket with handle	k	/k/
	stand for jar	g	/k/
	bread loaf	t	/t/
	tethering rope	č	/č/ as in English choke
	human hand	d	/t/
	snake	ḏ	/č/



hieroglyphic system never became fully phonetic, but always maintained the original combination of logograms and phonograms.




The phonological value of the phonograms is derived from the name of the represented entity by means of the rebus principle, i.e., by applying the same phonological sequence to other entities semantically unrelated to them. For example, from the representation of water \**maw* is derived the phonological value of this sign as /m-w/. In this process of derivation, called the *consonantal principle*, only a segment of the original sequence of phonemes of the represented entity, usually consisting of the strong consonants, is isolated to function as phonogram: thus, the sign for a house \**pa:ruw*, is used for the sequence /p-r/. In later times, the consonantal principle was expanded by the so-called *acrophonic principle*, i.e., the derivation of a phonological value from the first consonantal sound of the represented entity.

2. The sequence of phonograms is usually followed by a *semagram*, called a *determinative*, which classifies a word according to its semantic sphere: for example, a sitting man expresses the lexical realm of “man, mankind”; a sitting man touching his mouth the domain of




**Figure 7.1** From the cube statue of Senenmut, fifteenth century BC

“eating, speaking, thinking, sensing”; a scribe’s equipment  the area of “writing”; a stylized settlement  identifies the word as a toponym.

Many items of the basic vocabulary of Egyptian are expressed by semagrams which indicate their own semantic meaning. They do this (i) iconically, by reproducing the object itself; (ii) indexically, by portraying an entity whose name displays a similar phonological structure; or (iii) symbolically, by depicting an item metaphorically or metonymically associated with the object. These signs are called *logograms* or *ideograms*: for example, the hieroglyph which represents the enclosure of a house  is used to indicate iconically the concept “house” (*prw*); the sign representing a duck  means “son” (*z3*) by virtue of the phonetic similarity between the Egyptian words for “duck” and for “son”; the cloth wound on a pole  a sacred emblem placed on the pylons of Egyptian temples, through symbolic association denotes “god” (*ntr*).

Unlike most other systems of pictographic origin, such as Mesopotamian cuneiform or Chinese logograms, Egyptian hieroglyphs kept their original iconicity throughout their entire history without developing stylized forms. From about 2150 BC, Egyptian developed a subsystem of hieroglyphic orthography to express a sequence of *consonant* + *vowel*. In this subsystem, dubbed “syllabic orthography” (Schneider 1992; Zeidler 1993; Hoch 1994:487–504) and mostly used for the writing of words of foreign origin, three consonantal symbols (*/ʾ/*, */y/*, */w/*) were used to express vowels, in a procedure similar to the use of *matres lectionis* in Northwest Semitic orthography.

The writing system also possessed a set of hieroglyphic signs used to convey logographically the numbers  $10^0$ : $10^6$  and the fractions  $1/2$ ,  $1/3$ , and  $1/4$  (Loprieno 1986). To indicate natural numbers, signs appear repeated and organized sequentially from the highest to the lowest (  $356 = 3 \times 100, 5 \times 10, 6 \times 1$ ).

The basic orientation of the Egyptian writing system, and the only one used in the cursive varieties, is from right to left, with signs facing to the right; in monumental texts, the order may be inverted to left to right for reasons of symmetry or artistic composition.



Hieroglyphic writing conventions could be modified by addressing the figurative content of the sign. First of all, signs could become the vehicle for the expression of a cultural attitude vis-à-vis the entity it represented. For example, signs referring to the divine or royal sphere preceded in the writing any other sign belonging to the same compound noun, regardless of their actual syntactic positions. Conversely, a sign referring to a negatively connotated entity (for example an enemy) could be modified by means of substitution or mutilation of one of its features, in order to neutralize its negative potential. Secondly, the array of functional values of a specific sign could be expanded beyond the limits of the fixed convention: a sign could be given a different phonological value from the traditionally established one, especially by using it to indicate only the first consonantal phoneme of the corresponding word according to the acrophonic principle. This type of connotational expansion of the hieroglyphic system is found sporadically from the Old Kingdom onward, but developed dramatically in Ptolemaic times, leading to a radical change in the laws regulating the use of hieroglyphs.

## 2.2 Historical evolution

While the principles described above basically apply to the entire history of Egyptian writing, their distribution varies somewhat in the course of time. In the *archaic period*, around 3000 BC, the emergence of writing in Egypt is associated with a gradual development of a centralized system of government covering the entire country. In the inscriptions from this period on seals, palettes, and other monuments pertaining to the royal or administrative sphere, phonological and semantic principles are already intertwined, with a high number of signs functioning as logograms (Kahl 1994). In the *Old Kingdom* (Dynasty III–VI, 2750–2150 BC), the quantity and the complexity of written documents expands dramatically. Phonetic complementation may precede or follow the main sign. In the *classical system* of the Middle Kingdom (2050–1750 BC), which remained in use until the end of Dynasty XVIII (c. 1300 BC), a developed school system for the education of the bureaucratic elite fixes Egyptian orthography by reducing the number of graphic renditions allowed for any given word. The conventional orthography of the word usually consists either of a logogram, for the most basic nouns of the lexicon, or of a sequence of phonograms, often complemented, followed by a determinative. The inventory of hieroglyphs now totals about 750 signs (Gardiner 1957:438–548). During Dynasty XIX (1310–1195 BC), major changes affect the writing conventions of hieroglyphs and especially of Hieratic. In monumental texts, the space units within which sequences of hieroglyphs are formally arranged undergo an aesthetic readjustment. Changes are even more significant in manual writing, with a constant interface between traditional historical writing and the evolved phonetic reality.

With the decay of a powerful centralized government in the first millennium BC, centrifugal tendencies affect writing conventions as well. During Dynasty XXVI (seventh century BC), cursive *Demotic* develops at first in the north of the country, where the royal residence was located, and is gradually extended to the southern regions. Unlike Hieratic, which has sign groups that mirror the shape of the original hieroglyphs rather closely, Demotic signs break away from this tradition and adopt a set of stylized, conventional forms, in which the connection to the hieroglyphic counterpart is hardly perceivable, and which are therefore more likely to be used in purely phonetic function. Determinatives have now lost to a large extent their function as lexical classifiers. The development of Demotic marks the beginning of a divorce between monumental and cursive writing which will have a dramatic impact on the evolution of the hieroglyphic system as well. Demotic remained in literary and administrative use until the end of the Roman period.



**Table 7.2 The Coptic alphabet**

Character	Transcription	Character	Transcription
Ⲁ	a	Ⲡ	r
Ⲃ	b	Ⲃ	s
Ⲅ	g	Ⲅ	t
Ⲇ	d	Ⲇ	u
Ⲉ	e	Ⲉ	ph
Ⲋ	z	Ⲋ	kh
Ⲍ	ē	Ⲍ	ps
Ⲏ	th	Ⲏ	ō
Ⲑ	i	Ⲑ	š
Ⲓ	k	Ⲓ	f
Ⲕ	l	Ⲕ	ḥ
Ⲗ	m	Ⲗ	h
Ⲙ	n	Ⲙ	j
Ⲛ	ks	Ⲛ	c
Ⲝ	o	Ⲝ	ti
Ⲟ	p		

In *Ptolemaic* and *Roman* times (fourth century BC to third century AD), an increasing consciousness of the symbolic potential inherent in the relation between hieroglyphic signs and semantic meanings led to the development of previously unknown phonetic values and also of so-called cryptographic solutions. This evolution, which originated in priestly circles and remained until the end the monopoly of a very restricted intellectual community, threatened the accessibility of the system, favoring a dramatic increase in the number of signs, which now reaches many thousands (Daumas 1988–1995) and exploiting the full array of potential meanings of the hieroglyphic sign. And it is exactly this radical change in the nature of the writing system in the Greco-Roman period which is at the origin of the view, held in the Western world from late antiquity to the emergence of modern Egyptology, of the symbolic, rather than phonological, character of the hieroglyphic writing (Fowden 1986:13–74). With few exceptions, the Ptolemaic system was applied only to monumental writing.

## 2.3 Coptic

The first two centuries of our era saw the development of a whole corpus of mostly magical Egyptian texts in Greek letters, with the addition of Demotic signs to supplement it when phonologically required, known in the literature as *Old Coptic*. The pressure to adopt an alphabetic system increased with the Christianization of the country, when religious reasons contributed to the divorce between Egyptian culture and its traditional writing systems. The last dated hieroglyphic inscription is from the year AD 394. Demotic texts substantially decrease in number, Egyptian being replaced by Greek as a written language (Bagnall 1993:235ff.). The last Demotic graffito is dated to AD 452. In the following century, the new convention, which we call *Coptic*, appears completely established: the Egyptian language is now written in a Greek-derived alphabet. By the fifth century, the Egyptian elite had already lost the knowledge of the nature of hieroglyphs: the *Hieroglyphiká* of Horapollo, a hellenized Egyptian, offer a “decipherment” of the hieroglyphs fully echoing the late antique symbolic speculations (Boas and Grafton 1993).



## 2.4 Decipherment

The interest in matters Egyptian remained vivid in the West for the following centuries (Iversen 1961:57–123), but it was only in modern times that the understanding of the writing system was recovered. In the seventeenth century Athanasius Kircher recognized the linguistic derivation of Coptic from the language of the hieroglyphs (which he still took to be a symbolic writing), and in the eighteenth century Jean Barthélemy suggested that the cartouches which surround some hieroglyphic words contain divine and royal names. In 1799, during Napoleon's expedition to Egypt, the discovery of the so-called Rosetta Stone, a trilingual (Hieroglyphic, Demotic, and Greek) document from the Ptolemaic period found in the Egyptian town of Rosetta, in Arabic *Al-rašīd*, provided the possibility of comparing the same text in two unknown writing systems (Demotic and hieroglyphs) and in Greek; this event opened the way to the actual decipherment. First results were reached by the Swede Johan David Åkerblad for the Demotic section and especially by the English physician Thomas Young, who, however, did not progress beyond the royal names. The most decisive contribution to the decipherment was achieved by the French scholar Jean-François Champollion in his *Lettre à M. Dacier* (1822), and especially in the *Précis du système hiéroglyphique* (1824). On the basis of the writing of Greek names in the hieroglyphic text, Champollion was able to establish the presence of a phonetic component in the system, breaking away from the traditional symbolic approach (Iversen 1961:124–145).

## 3. PHONOLOGY

### 3.1 Phonemes and graphemes

The exact phonological value of many Egyptian phonemes is obscured by difficulties in establishing reliable Afro-Asiatic correspondences (Schenkel 1990:24–57). Vocalism and prosody can be partially reconstructed on the basis of (i) Akkadian transcriptions of Egyptian words and phrases from the second millennium BC, (ii) Greek transcriptions from the Late Period (corresponding roughly to spoken Demotic), and (iii) the Coptic evidence of the first millennium AD. In the sketch of Egyptian phonology presented below, Egyptological transliterations are given in *italic*, whereas underlying phonological realities are rendered between slashes and, since they are scholarly reconstructions, always preceded by an asterisk (note that by convention a dot is used to separate the root from morphological affixes; e.g., *sn.t* “sister” < root *sn* + feminine marker *t*). As for Coptic, in spite of a certain number of graphic idiosyncrasies, all dialects share a relatively uniform phonological system. For example, the graphic conventions of Sahidic – as opposed to those of Bohairic – do not distinguish between voiceless and ejective plosives (Sahidic *tôre*, Bohairic *thôri* = /t<sup>h</sup>o:rə/ “willow” ~ Sahidic *tôre*, Bohairic *tôri* = /t̥o:rə/ “hand”); or between velar and glottal fricatives (Sahidic *hrai*, Bohairic *hrai* = /hraj/ “above” ~ Sahidic *hrai*, Bohairic *xrai* = /xraj/ “below”). Yet the presence of the corresponding oppositions in Sahidic can be established on the basis of comparative dialectology and of the different impact of these phonemes on their respective phonetic environment (Loprieno 1995:40–50).

### 3.2 Consonants

#### 3.2.1 Stops and affricates

The stops and affricates of Earlier Egyptian are presented in (3):

## (3) Earlier Egyptian stops and affricates

	<i>Bilabial</i>	<i>Dental</i>	<i>Palatal</i>	<i>Velar</i>	<i>Uvular</i>	<i>Glottal</i>
<i>Voiced</i>	b = /b/	( <sup>c</sup> = /d/)	—	—	—	—
<i>Voiceless</i>	p = /p <sup>h</sup> /	t = /t <sup>h</sup> /	<u>t</u> = /č <sup>h</sup> /	k = /k <sup>h</sup> /	—	—
<i>Ejective</i>	—	d = /t̥/	<u>d</u> = /č̥/	g = /k̥/	q = /q/	ʔ = /ʔ/

In the Egyptian phonological system, the opposition between voiceless and voiced phonemes (Schenkel 1993:138–146) appears limited to bilabial stops (4A), whereas in the other series the articulatory opposition – when present – is between voiceless and ejective stop or affricate (4B–C). The voiceless varieties displayed the optional feature of aspiration in pretonic and high-sonority environments:

- (4) A. bilabial /b/ ~ /p<sup>h</sup>/: Bohairic *bôk* “servant” ~ *p<sup>h</sup>ôk* “yours.MASC.SG.”  
 B. dental /t<sup>h</sup>/ ~ /t̥/: Earlier Eg. *tm* “to complete” ~ *dm* “to sharpen”  
 C. palatal /č<sup>h</sup>/ ~ /č̥/: Earlier Eg. *tr.t* “willow” ~ *dr.t* “hand”

The dental series is typologically complex: while it probably exhibited a tripartite opposition *voiceless–voiced–ejective* in the earliest periods, the voiced stop *\*d/* evolved into a pharyngeal fricative *\*ʕ/* before the emergence of Middle Egyptian (Zeidler 1992:206–210), and then to a glottal stop (and eventually zero) in Coptic (5A). During the second millennium BC, the voiceless dental /t/ shows the tendency to be dropped in final position (5B):

- (5) A. *\*d/* > *\*ʕ/* > /ʔ/ or /ø/:  
 Old Eg. *ʕs* *\*da:š/* > Late Eg. *\*ʕa:š/* > Coptic *ôš* / (ʔ) *o:š/* “to call”  
 B. *t* > ø / \_\_\_#:  
 Old Eg. *sn.t* *\*sa:nat/* > Late Eg. *\*sa:nə(t)/* > Coptic *sône* / *so:nə/* “sister”

During the late second millennium BC, the place of articulation of stop consonants tends to be moved to the frontal region (Osing 1980:946): uvulars and velars are palatalized (6), palatals become dentals and dentals are dropped in final position (7):

## (6) Uvular and velar palatalization

- A. Late Eg. *kʕm* *\*k<sup>h</sup>aʔm/* > Coptic *côm* / *kʕo:m/* “garden”  
 B. Old Eg. *gr* *\*k̥a:r/* > Coptic *cô* / *kʕo:ʔ/* “to cease”  
 C. Old Eg. *qd* *\*qat̥/* > Coptic *cot* / *kʕot/* “form”

## (7) Palatal &gt; dental; dental &gt; ø / \_\_\_#

- A. Old Eg. *dr.t* *\*č̥a:rat /* > Late Eg. *\*ta:rə(t)/* > Coptic *tôre* / *to:rə/* “hand”  
 B. Old Eg. *rm̥t* *\*ra:mač/* > Late Eg. *\*ra:mə(t)/* > Coptic *rôme* / *ro:mə/* “man”

During the first millennium BC, the opposition between uvulars and velars is neutralized: Coptic exhibits a new tripartite opposition “palatalized : voiceless (with optional aspiration) : ejective” in the velar series:

(8) /k<sup>h</sup>/ ~ /k̥/ ~ /q/ > /kʕ/ ~ /k<sup>h</sup>/ ~ /k̥/

- A. *kô* “shrine” (from Eg. *\*k<sup>h</sup>/*) ~ *cô* “to cease” (from Eg. *\*k̥/*)  
 B. *côb* “weak” (from Eg. *\*k̥/*) ~ *kôb* “to double” (from Eg. *\*q/*)  
 C. *cot* “form” ~ *kot* “wheel” (both from Eg. *\*q/*)

## (9) Stops and affricates in Sahidic Coptic

	<i>Bilabial</i>	<i>Dental</i>	<i>Palatal</i>	<i>Velar</i>	<i>Glottal</i>
<i>Palatalized</i>	—	—	—	c=/kʲ/	—
<i>Voiced</i>	b=/b/	(d=/d/)	—	(g=/g/)	—
<i>Voiceless</i>	p=/p <sup>(h)</sup> /	t=/t <sup>(h)</sup> /	j=/ç <sup>(h)</sup> /	k=/k <sup>(h)</sup> /	—
<i>Ejective</i>	—	t=/t̥/	j=/ç̥/	k=/k̥/	/ʔ/

It should be noted that the opposition between voiceless and ejectives is neutralized as voiceless (unmarked) in posttonic position (10A), and that voiced dentals and velars are only found in Greek borrowings or as a result of assimilation of the corresponding voiceless in nasal environments (10B):

- (10) A. *sôtm* /so:təm/ < /so:təm/ “to hear” ~ *sôtp* /so:təp/ < /so:t<sup>(h)</sup>əp/ “to choose”  
 B. *tooun-g* < *tooun-k* “stand up!”

## 3.2.2 Fricatives

In Old Egyptian, all fricative consonants were voiceless; in Middle Egyptian, as we have just seen (in [5] above), a voiced pharyngeal /ʕ/ evolved from earlier /d/ via lateralization.

## (11) Fricatives in Earlier Egyptian

	<i>Labio-dental</i>	<i>Inter-dental</i>	<i>Alveolar</i>	<i>Palato-alveolar</i>	<i>Palatal</i>	<i>Velar</i>	<i>Pharyngeal</i>	<i>Glottal</i>
<i>Voiceless</i>	f=/f/	z=/θ/	s=/s/	š=/š/	ħ=/ç/	ḫ=/x/	ḥ=/ħ/	h=/h/
<i>Voiced</i>	—	—	—	—	—	—	—	—

The interdental *z* merged very early with the alveolar *s* (/θ/ > /s/). In the first millennium BC, the tripartite opposition between fricatives in the palatal region (/š/ ~ /ç/ ~ /x/) was reduced to a bipartite one (/š/ ~ /x/), with a partial redistribution of the original articulation (Osing 1976:401–402; 503):

- (12) A. \*/x/ > /š/:  
 Old Eg. *ḥm* \*/da:xam/ > Late Eg. \*/ʕa:xəm/ > Coptic *ôšm* /ʔo:šəm/ “to extinguish”  
 B. \*/ç/ ~ /š/ > /x/:  
 Old Eg. *zhw* \*/θa:çraw/ ~ /θa:šraw/ > Coptic *sah* /sax/ “scribe”  
 C. \*/ç/ > /x/:  
 Old Eg. *ḥm* \*/da:çamv/ “falcon” > Late Eg. \*/ʕa:çam/ > Coptic *ahôm* /ʔaxo:m/ “eagle”

A similar neutralization affected in the first millennium BC the opposition between pharyngeal /ħ/ and glottal /h/ (Osing 1976:367–368):

- (13) A. Old Eg. *ḥ<sup>3</sup>t* \*/ḥu:rit/ > Late Eg. \*/ḥe:ʔə(t)/ > Coptic *hê* /he:ʔ/ “beginning”  
 B. Old Eg. *h<sup>3</sup>w* \*/haru:w/ > Late Eg. \*/həʔe:ʔ/ > Coptic *hê* /he:ʔ/ “season”

The distribution of fricative phonemes in Sahidic Coptic is thus as follows:

## (14) Fricatives in Sahidic Coptic

	<i>Labiodental</i>	<i>Alveolar</i>	<i>Palato-alveolar</i>	<i>Velar</i>	<i>Glottal</i>
<i>Voiceless</i>	f=/f/	s=/s/	š=/š/	ḫ=/x/	h=/h/
<i>Voiced</i>	—	(z=/z/)	—	—	—

It should be noted that the voiced alveolar fricative *z* is only found in Greek borrowings or as a result of assimilation in nasal environments:

- (15) Coptic *anzêbe* < *ansêbe* “school”

### 3.2.3 Sonorants

Historical evolutions affecting nasals, liquids, and glides during the second millennium BC (Loprieno 1995:38) involved (i) the loss of the uvular vibrant (i.e., nonlateral) /ʀ/ and its lenition to glottal stop /ʔ/ and eventually *ø* (16), and (ii) the loss of final vibrants and glides (17) in the same environments in which a final voiceless dental *t* was dropped (see [7]):

- (16) /ʀ/ > /ʔ/ > *ø*

Old Eg. *kʷmw* \*/k<sup>h</sup>armaw/ > Late Eg. \*/k<sup>h</sup>aʔm/ > Coptic *côm* /k<sup>ʷ</sup>o:m/ “garden”

- (17) /r, ɣ, w/ > *ø* / —#

A. Old Eg. *ḥpr* \*/xa:par/ > Late Eg. \*/xa:pə(r)/ > Coptic *šôpe* /šo:pə/ “to become”

B. Old Eg. *ntṛ.w* \*/nac<sup>h</sup>u:raw/ > Late Eg. \*/nəṯ<sup>h</sup>e:rə/ “gods” > Coptic *ntêr* /ənt<sup>h</sup>e:r/ “idols”

- (18) Sonorants in the Egyptian domain

	Labial	Dental	Palatal	Uvular
Nasal	m = /m/	n = /n/	š = /š/	
Vibrant	—	r = /r/	—	(ʒ = /ʀ/)
Lateral	—	/l/	—	—
Glide	w = /w/	—	j = /y/	—

### 3.3 Vowels

The set of vowels posited for Earlier Egyptian (Osing 1976:10–30) is the same as for most Afro-Asiatic languages in their earliest stage of development (Diakonoff 1965:30–31):

- (19) Vowels in Earlier Egyptian

	Short	Long
Front	/i/	/i:/
Central	/a/	/a:/
Back	/u/	/u:/

This system underwent a certain number of historical changes, only some of which can be discussed here. First and foremost, because of the presence of a strong expiratory stress, Egyptian unstressed vowels gradually lost phonological status, until in Coptic they are generally realized as *schwa*; only the short unstressed /a/ is maintained in pretonic position in specific phonetic environments (Schenkel 1990:91–93):

- (20) A. Old Eg. *rmṯ nj km.t* \*/ramac-ni-k<sup>h</sup>u:mat/ > Coptic *rmnkême* /rəmənkh<sup>h</sup>e:mə/ “Egyptian man”

B. Old Eg. *jnk* \*/janak/ > Coptic *anok* /ʔanok/ “I”



Stressed vowels underwent a global *Lautverschiebung*: during the second millennium BC, long /u:/ turned into /e:/, while short stressed /i/ and /u/ merged into /e/. In the main Coptic dialects and unless followed by glottal stop, this /e/ evolved into /a/:

- (21) A. Old Eg. *rn* \*/rin/ > Late Eg. \*/ren/ > Coptic *ran* /ran/ “name”  
 B. Old Eg. *m³ꜥ.t* \*/murdat/ > Late Eg. \*/meʔʕə(t)/ > Coptic *me* /meʔ/ “truth”  
 C. Old Eg. *km.t* \*/kʰu:mat/ > Late Eg. \*/kʰe:mə(t)/ > Coptic *kême* /kʰe:mə/ “Egypt”

Around 1000 BC, long /a:/ became /o:/ (/u:/ after nasals) and short /a/ became /o/, a change limited to the same portion of the Coptic linguistic domain to which /i/, /u/ > /e/ applies:

- (22) A. Old Eg. *ntr* \*/na:car/ > Coptic *noute* /nu:tə/ “God”  
 B. Old Eg. *sn* \*/san/ > Coptic *son* /son/ “brother”

(23) Vowels in Sahidic Coptic

	Unstressed	Stressed	
		Short	Long
Front			i, ei /i:/
		e /e/	ê /e:/
Central	e /ə/		
	a /a/	a /a/	
Back			ou /u:/
		o /o/	ô /o:/

### 3.4 Stress and syllabic patterns

In Earlier Egyptian, the stress lay on the ultimate (oxytone) or penultimate (paroxytone) syllable of a word (Schenkel 1990:63–86). Closed (CVC) and open (CV) syllables can be found in pretonic, tonic, and posttonic position. The stressed vowel of a penultimate open syllable is always long (CV:); according to some scholars, extrasyllabic additions under oxytone stress could generate syllables of the type CV:(C) or CVC(C) (Loprieno 1995: 36–37):

(24) Earlier Egyptian syllabic structures

	Pretonic	Tonic	Posttonic
Open	\$CV\$	\$'CV:\$	\$CV#
Closed	\$CVC\$	\$'CVC\$	\$CVC#
Doubly closed		\$'CVCC#	
Long		\$'CV:C\$	

These syllabic structures were modified under the influence of the strong expiratory stress which always characterized the Egyptian domain (Fecht 1960) and prompted significant typological changes in morphology and syntax. The gradual loss of short unstressed vowels led to the emergence of complex consonantal clusters in syllable onset (i.e., word-initially) in Coptic (Loprieno 1995:48–50):

## (25) Coptic syllabic structures

	<i>Pretonic</i>	<i>Tonic</i>	<i>Posttonic</i>
<i>Open</i>	\$CV\$ #CCV\$	\$'CV:\$ #'CCV:\$	\$CV#
<i>Closed</i>	\$CVC\$ #CCVC\$	\$'CVC\$ #'CCVC\$	\$CVC#
<i>Doubly-closed</i>		\$'CVCC# #'CCVCC#	
<i>Long</i>		\$'CV:C# #'CCV:C#	

Examples for the evolution of oxytone patterns follow:

## (26) A. CV'CVC &gt; CCVC

Old Eg. *wḏḥ* \*/waḥah/ > Coptic *outah* /wəṭah/ “fruit”

## B. CVC'CVC &gt; CVC'CVC

Old Eg. *nmḥw* \*/numḥiw/ “poor” > Coptic *rmhe* /rəmhe?/ “free”

Examples for the evolution of paroxytone patterns:

## (27) A. 'CVCCVC &gt; CVCC

Old Eg. *ḥmtw* \*/xamtaw/ > Bohairic *šomt* /šomt/ “three”

## B. CV'CVCCVC &gt; CCVCC

Old Eg. *ḥjm.wt* \*/ḥijamwat/ > Coptic *hiome* /hjom?/ “women”

## C. CV'CV:CVC &gt; CCV:C

Old Eg. *psḏw* \*/psi:ḥaw/ > Coptic *psit* /psi:t/ “nine”

## 4. MORPHOLOGY

## 4.1 Word formation

Earlier Egyptian is a language of the flecnal or fusional type, in which morphemes are unsegmentable units combining many grammatical functions. Morphological forms exhibit a number of correspondences with the patterns of word formation in other Afro-Asiatic languages (Schenkel 1990:94–121). In recent years, scholars have also emphasized the importance of prehistoric contact between Egyptian and Indo-European (Ray 1992:124–136; Kammerzell 1994:37–58).

The basic structure of an Egyptian word is a lexical *root*, an abstract phonological entity consisting of a sequence of consonants or semiconsonants which vary in number from one to four, with an overwhelming majority of biconsonantal, triconsonantal, and so-called weak roots, which display a vocalic or semivocalic last radical (i.e., phoneme) or a gemination of the second radical. Superimposed on the root as a separate morphological tier is a vocalic or semivocalic pattern, which together with the root forms the so-called *stem*, the surface form acquired by the root; the stem determines the functional class to which the word belongs. It is transformed into an actual *word* by means of inflectional affixes (in Egyptian for the most part suffixes), which convey deictic markers and other grammatical functions such as gender, number, tense and aspect, and voice (Reintges 1994).

Vocalic skeletons generally determine the structure of nominal patterns and of basic conjugational forms, whereas semivocalic suffixes convey the expression of the plural, of adjectival forms of the verb (participles and relative forms), and of some conjugational



patterns. A *j*- or *w*- prefix can be added to biconsonantal roots to form triradical nominal stems; conversely, a triconsonantal root may lose a semivocalic glide and be reduced to a biradical stem. Examples of consonantal additions prefixed to a root are *s*- for causative stems, *n*- for singulative nouns and reflexive verbs, and *m*- for nouns of instrument, place, or agent. Egyptian stems resulting from the addition of a consonantal phoneme to a root tend to be lexicalized as new autonomous roots rather than treated as grammatical forms of the basic root: Egyptian, therefore, does not possess a full-fledged paradigm of verbal stems conveying semantic nuances of a verbal root similar to the ones known in Semitic.

Common modifications of the root are:

1. The reduplication of the entire root or of a segment thereof. This pattern affects the semantic sphere, creating new lexemes: from *sn* “brother,” *snsn* “to befriend”; from *gmj* “to find,” *ngmgm* “to be gathered” (with the *n*-prefix of reflexivity); from *snb* “to be healthy” *snbb* “to greet.”
2. The gemination of the last radical, which affects the grammatical sphere: *dd* “to say” > *ddd.t* “what has been said”; *mrj* “to love” > *mrr-j* “that I love”; *sdm* “to hear” > *sddmm-f* “he will be heard” (Reintges 1994:230–240).

## 4.2 Nominal morphology

Both nouns and adjectives are included in this category.

### 4.2.1 Nouns

In Earlier Egyptian, nouns are built by adding to the stem a zero- or a non-zero-suffix, depending on whether the stem ends in a consonant, in which case the suffix is zero, or a vowel, in which case a *w*-suffix is added. The feminine marker is a *t*-suffix added to the masculine noun; the plural displays a *w*- or *ww*-suffix; the dual has a *j*-marker added to the stem of the plural in masculine, and to the stem of the singular in feminine, nouns:

#### (28) Nouns in Earlier Egyptian

	<i>Masculine</i>	<i>Feminine</i>
<i>Singular</i>	-Ø, -w	-t
<i>Dual</i>	-w-j	-t-j
<i>Plural</i>	-Ø, -w, -ww	-t, -j-t, -w-t

### 4.2.2 Adjectives

Adjectives are morphosyntactically treated like nouns. In a common derivational pattern, called *nisbation*, a morpheme masculine *\*ij*, feminine *\*it* is added to a stem, which may be different from the stem of the singular or plural noun, to form the corresponding adjective: *nṯr*\*/na:čar/ “god,” *nṯr.w*\*/nač<sup>h</sup>u:ra(w)/ “gods,” *nṯrj*\*/nučriy/, *nṯrj.t*\*/nučrit/ “divine.”

## 4.3 Pronouns

### 4.3.1 Personal pronouns

There are four sets of personal pronouns (Kammerzell 1991), including one reserved for the stative form of the verb (see §4.4.1). Stressed pronouns are used for the topicalized subject

**Table 7.3 Personal pronouns in Earlier Egyptian**

		Stressed	Unstressed	Suffix
SINGULAR	1st com.	jnk	wj	-j
	2nd masc.	ntk, <u>t</u> wt	<u>t</u> w	-k
	2nd fem.	ntt, <u>t</u> mt	<u>t</u> n	- <u>t</u>
	3rd masc.	ntf, swt	sw	-f
	3rd fem.	nts, stt	sj, st	-s
(twt, tmt, swt and stt are archaic forms found mainly in Old Kingdom religious texts)				
DUAL	1st com.		nj	-nj
	2nd com.	nttnj	<u>t</u> nj	- <u>t</u> nj
	3rd com.	ntsnj	snj	-snj
PLURAL	1st com.	jnn	n	-n
	2nd com.	nttn	<u>t</u> n	- <u>t</u> n
	3rd com.	ntsn	sn, st	-sn

of noun clauses in the first and second person (29A), and for the focalized subject of verbal cleft sentences (29B):

- (29) A. jnk            jt-k  
           I.TOPIC    father-you  
           “I am your father”  
       B. nts            s-nḥ                    rn-j  
           she.FOCUS    CAUS.-live.PART.    name-me  
           “She is the one who makes my name live”

Unstressed pronouns are used for the object of verbal phrases (30A), for the subject of adjective clauses (30B) and of adverb clauses (30C):

- (30) A. h<sup>3</sup>b-f            wj  
           send.PERF.-he    me  
           “He sent me”  
       B. nfr            tw    ḥn<sup>c</sup>-j  
           be.good.PART.    you    with-me  
           “You are happy with me”  
       C. mk            wj    m-b<sup>3</sup>ḥ-k  
           behold    me    in-presence-you  
           “Look, I am in front of you”

Suffix pronouns are used as the subject of verb phrases, as possessive marker, and as the object of prepositions:

- (31) dj-k                    r-k                    n-j                    ḥ.t-j  
           give.PROSP.-you    toward-you    to-me    thing.FEM.-me  
           “You shall indeed (lit. ‘toward-you’) give me my possessions”

#### 4.3.2 Demonstrative pronouns

Demonstratives are characterized by a deictic element preceded by the indicator of gender and number: masculine *pn*, *pf*, *pw*; feminine *tn*, *tf*, *tw* – for example, *rmṯ pf* “that man,” *ḥjm.t*

*tn* “this woman.” They follow the noun they refer to. The plurals (originally neuter) *nw*, *nf*, *nn* are also used as pronouns in partitive constructions with the determinative pronoun *nj*; *nn nj srjw.w* “these officials” < \* “this of officials.” This determinative pronoun *nj*, feminine *n.t*, plural *n.w* is used primarily as a genitive marker: *rmṯ.w n.w km.t* “men of Egypt” > “Egyptians.” On Egyptian articles, see §4.6.2.

### 4.3.3 Relative pronouns

The relative pronoun, masculine *ntj*, feminine *nt.t*, plural *ntj.w* “who, which, that,” is morphologically derived from the determinative pronoun. In Earlier Egyptian, these pronouns agree in gender and number with the head noun, which must be semantically specific. Characteristic of Earlier Egyptian is the presence of a relative pronoun masculine *jwṯj*, feminine *jwṯ.t*, plural *jwṯj.w*, which semantically incorporates negation (“who/which/that not”):

- (32) *jwṯj*            *phr-f*            *ḏd.w*            *m*    *ḥ.t-f*  
           who.not    vent.AOR.-he    say.PART.IMPF.PASS.    in    belly.FEM.-him  
           “He who does not vent what is said in his belly”

### 4.3.4 Interrogative pronouns

Interrogative pronouns are *m* “who? what?,” *jḥ* “what?,” *jšst* “what?” They can be combined with prepositions or particles to form complex pronouns: *jn-m* “who?,” *ḥr-m* “why?,” literally “on-what?”

## 4.4 Verbal morphology

### 4.4.1 Finite verb-stems

Earlier Egyptian finite verb phrases display a limited number of stems (three or four) indicating tense, aspect, and voice followed by the pronominal suffix (33A) or nominal subject (33B):

- (33) A. *ḥḥ-s*  
           live.PROSP.-she  
           “She will live”  
       B. *ḥ3b*            *ḥjm.t*            *z3-s*  
           send.PERF.    woman    son-her  
           “The woman sent her son”

Typical Egyptian verb “inflection” (utilizing the suffix pronouns) is illustrated in (34) with the verb-stem *sdm* “hear”:

- |               |           |               |             |
|---------------|-----------|---------------|-------------|
| (34) SINGULAR | 1st com.  | <i>sdm-j</i>  | “I hear”    |
|               | 2nd masc. | <i>sdm-k</i>  | “you hear”  |
|               | 2nd fem.  | <i>sdm-t</i>  | “you hear”  |
|               | 3rd masc. | <i>sdm-f</i>  | “he hears”  |
|               | 3rd fem.  | <i>sdm-s</i>  | “she hears” |
| PLURAL        | 1st com.  | <i>sdm-n</i>  | “we hear”   |
|               | 2nd com.  | <i>sdm-tn</i> | “you hear”  |
|               | 3rd com.  | <i>sdm-sn</i> | “they hear” |

In addition to variations in the stem, a few verbal features are indicated by complementizers inserted between the stem and the subject. The most important of these indicators are *n* for the preterite tense, *t* for nonparadigmatic occurrences of the perfective aspect and for the prospective aspect of a few irregular verbs, *w* for prospective aspect and passive voice (in perfective stems), *tw* for passive (in nonperfective stems).

A particular verbal stem of nominal (probably relative) origin displays the tonic vowel between the second and the third radical, and in weak verbal classes the reduplication of the second radical: *stp-\*/satap-/* (choose.REL.), *mrr-\*/marar-/* (love.REL.). A similar verbal form indicates in Semitic languages the imperfective aspect; in Egyptian, its function is to mark the verb phrase as pragmatic theme of the sentence in which it appears (Polotsky 1976:4–25). In these sentences, the pragmatic rheme is usually a modifier or an adverb clause:

- (35) *jrr*            *ḥm-k*            *r*    *mrj.t-f*  
          do.IMPF.    Majesty-your    to    desire.REL.FEM.-him  
          “Your Majesty acts as he desires”

The imperative has no suffix element in the singular, but sometimes, especially with weak verbs, a semivocalic suffix in the plural.

Egyptian also exhibits a verbal form, called Old Perfective, Stative, or Pseudoparticiple, which indicates the wide semantic range of “perfectivity,” from perfect aspect (with intransitive verbs) to passive voice (with transitive verbs). This form is built with a special set of suffixes that are etymologically linked to the forms of the Semitic suffix conjugation (Schenkel 1990:104–108; Kammerzell 1991:165–199):

- (36) *mk*            *wj*    *jj-kw*  
          behold    me    come.STATIVE-me  
          “Look, I have come,” i.e., “I am here”

#### 4.4.2 Nonfinite verbals

Nonfinite forms of the Egyptian verb are (i) the *participles*, with nominal stems derived from the verbal root (e.g., *sḏm-\*/sa:ḥim/* “hearer”); and (ii) the *infinitives*, which display a suffix *ø* in the regular verbs (*sḏm-\*/sa:ḥam/* “to hear”), *t* in some classes of weak verbs (*mrj.t-\*/mirjit/* “to love”), and *w* after verbs of negative predication, such as *tm* (*tm jr.w-\*/tam-ja:raw/* “not to do,” lit. to complete-to do.NEG.INF.).

### 4.5 Particles

The basic negative particle is *n*, which is used for *contradictory negation* (37A); when combined with the adverb *js* “indeed,” this morpheme expresses *contrariety* (37B; see Loprieno 1991):

- (37) A. *n*            *rd-f*                    *n-j*            *mw*  
          not    give.PERF.-he    to-me    water  
          “He did not give me water”  
       B. *n-js*            *jt-j*                    *rdj*                    *n-j*  
          not-indeed    father-me    give.PART.    to-me  
          “It was not my father who gave [it] to me”

A morphological variant of *n*, conventionally transcribed *nn*, is used in noun clauses to negate the *existence* (37A) and in verb clauses to negate the *prospective* aspect (38B):

- (38) A. nn            m3<sup>c</sup>.tjw  
          not.exist    trust.ADJ.PL.  
          “There are no trustworthy people”  
       B. nn            mwt-k  
          not.exist    die.PROSP.-you  
          “You shall not die”

## 4.6 Morphological evolution

Under the pressure of a strong expiratory stress, which reduced the distinctive function of unstressed vowels, the flectional system underwent a profound crisis in Later Egyptian, requiring a reorganization of the morphological carriers of information.

### 4.6.1 Nominal lexicalization

The general trend was to replace synthetic structures by analytic constructions: for example, nominalized participles (39) or abstract nouns (40) were replaced by lexicalized compounds with nominal classifiers (Till 1970:71–75):

- (39) PARTICIPLE    >    “MAN-WHO”-V  
       t3w                ref-jioue  
       steal.PART.        “MAN-WHO”-steal.INF.  
       “Thief”  
  
 (40) ABSTRACT NOUN    >    “THING-OF”-N  
       r3            nj    km.t            mnt-rm-n-kême  
       mouth    of    Egypt            “THING-OF”-man-of-Egypt  
       “Egyptian language”

### 4.6.2 Articles

Later Egyptian develops two sets of articles. The *indefinite article* comes from the numeral *wj* “one”:

- (41) N.[−SPEC]    >    INDEF.ART.-N.  
       Earlier Eg. *sn.t* “a.sister” > Late Eg. *w(t)-sn(.t)* > Coptic *ou-sône* “a-sister”

whereas the *definite article* (Loprieno 1980) derives from a grammaticalized anaphoric pronoun:

- (42) N.[+SPEC]    >    DEF.ART.-N.  
       Earlier Eg. *rm.t* “the.man” > Late Eg. *p3-rm(t)* > Coptic *p-rôme* “the man”

The definite article also attracts the pronominal affix indicating the possessor, which in Earlier Egyptian followed the head noun (43A). Similarly, deictics now precede the noun they modify (43B):

- (43) A. N.-SUFFIX            >    DEF.ART.-SUFFIX-N.  
       sn-f                        pe-f-son  
       brother-his                the-his-brother  
       “His brother”

B. N.-DEICTIC	>	DEICTIC-N.
hjm.t tn		tei-shîme
woman this.FEM.		this.FEM.-woman
"This woman"		

### 4.6.3 Coptic morphological markers

Thus, because of the described loss of regular flectional patterns, the only device by which Coptic conveys the distinction between different patterns (masculine vs. feminine, nominal vs. verbal) is through the presence of morphological markers preceding the noun (44A–C; a zero-marker in the case of C):

- (44) A. *rmt*: stem \*ramač- + MASC.SG.Ø = \*/ra:mač/ > Coptic *p-rôme* "the man"  
 B. *sn*: stem \*san- + FEM.SG. at = \*/sa:nat/ > Coptic *t-sône* "the sister"  
 C. *hpr*: stem \*xapar- + INF. Ø = \*/xa:par/ > Coptic *šôpe* "to become"

### 4.6.4 Verbal lexicalization

The evolution towards a lexicalization of compound expressions also affected the verbal system (Winand 1992:20). In many instances, an earlier verbal lexeme is replaced in Later Egyptian, particularly in Coptic, by an auxiliary of generic meaning ("to do," "to give," "to take," etc.) followed by the verbal infinitive or by a noun object:

- (45) VERBAL LEXEME > AUXILIARY + NOUN  
 wd<sup>c</sup> r-hap, ti-hap  
 judge.INF. do.INF.-law, give.INF.-law  
 "to judge"

Participles were superseded by analytic constructions with the relative pronouns (46A), while finite VSO forms were replaced by a paradigm of SVO-constructions, called "sentence conjugations" or "clause conjugations" (Polotsky 1960), resulting from the grammaticalization of a form of the verb "to do" followed by the infinitive (46B):

- (46) A. PARTICIPLE > RELATIVE CONSTRUCTION  
 Old Eg. > Late Eg. > Coptic  
 sdm p<sup>3</sup>-ntj (h<sub>r</sub>) sdm p-et-sôtm  
 hear.PART.IMPF. the.one-who-(on-)hear.INF. the.one-who-hear  
 "the hearer"
- B. VSO > SVO  
 Old Eg. > Late Eg. > Coptic  
 sdm.h<sub>r</sub>-f h<sub>r</sub>-j<sub>r</sub>-f-sdm ša-f-sôtm  
 hear-AOR.-he AOR.-do-he-hear.INF. AOR.-he-hear  
 "He usually hears"

### 4.6.5 Coptic verb morphology

In this way, Coptic ultimately maintains only two flectional patterns from most verbal roots: (i) the infinitive for process predicates, and (ii) the so-called "qualitative," derived from the third masculine singular (rarely third feminine singular) form of the Old Perfective, for statives (Polotsky 1990:197–221):

- (47) *f-kôt*                      ~    *f-kêt*  
       he-build.INF.            it-build.STAT.  
       “He builds”            “It is built”

Thus, with the productivity of root and stem variations massively reduced, Later Egyptian gradually moves toward the polysynthetic type which to a large extent characterizes Coptic:

- (48) *Earlier Eg.*  
       *jw*                      *sdm.n-j*            *h̥rw*  
       “SITUATION”        hear-PRET.-I        voice  
                                  > *Late Eg.*  
                                  *jr-j-sdm*                      *w<sup>c</sup>-h̥rw*  
                                  do.PRET.-I-hearing        a-voice  
                                  > *Coptic*  
                                  *ai-setm-ou-h̥roou*  
                                  PRET.-I-hear-a-voice  
                                  “I heard a voice”

Nonfinite forms of the Coptic verb are the *infinitive*, which usually indicates (i) activities (*ei* “to come”), (ii) accomplishments (*ôô* “to conceive”), or (iii) achievements (*cine* “to find”); and the *qualitative*, which conveys states (*eet* “to be pregnant”). Although participial functions, as we saw above, are analytically conveyed by relative constructions, there are still few remnants of Ancient Egyptian synthetic participles (*mai-noute* “lover of god” > “pious”). Finite verbal forms now consist of a marker which conveys aspectual, temporal, or modal features, followed by the nominal or pronominal subject and by the infinitive (for actions) of the verb: *a-prôme sôtm* “the man heard,” *a-i-hmoos* “I sat down.” In the present and imperfect tense, which are treated as adverbial constructions, the infinitive can be replaced by the qualitative (for states): *ti-hkaeit* “I am hungry.” The most important verbal markers are as follows (the double stroke indicates pronominal subjects, the simple stroke nominal subjects):

1. *e=, ere-*: circumstantial present (*e=i-hkaeit* “while I am hungry”)
2. *ša=, šare-*: aorist of habit (*ša=i-ka pa-joi na=i* “I keep my ship for me”)
3. *me=, mere-*: negative aorist (*me=f-sôtm* “he cannot hear”)
4. *e=e, ere-e-*: prospective of wish (*e=s-e-šôpe* “may it happen,” “amen”)
5. *nn(e)=, nne-*: negative prospective (*nne=f-eibe ša-eneh* “may he never be thirsty”)
6. *mar(e)=, mare-*: optative (*mare-pe=k-ran ouop* “hallowed be your name”)
7. *(n)tare=, (n)tare-*: final (*aitei tar=ou-ti nê=tn* “ask, that you may be given”)
8. *šant(e)=, šante-*: completive (*šante-prê hôtp* “until the sun sets down”)
9. *mpat(e)=, mpate-*: negative completive (*mpat=f-ei* “he has not yet come”)
10. *a=, a-*: preterite (*a-ouša šôpe* “a festival took place”)
11. *mp(e)=, mpe-*: negative preterite (*mpi-raše* “I did not rejoice”)
12. *ne=, nere-*: imperfect (*nere-tmaau n-iêsous mmau* “Jesus’ mother was there”)
13. *nter(e)=, ntere-*: temporal (*ntere=f-je nai* “when he said these things”)
14. *n=, nte-*: conjunctive (*e=k-e-nau n=g-eime* “may you see and understand”)

In addition to these so-called sentence (or clause) conjugations, Coptic displays (i) an inflected form of the infinitive (*p-tre=f-sôtm* “the fact that he hears”); (ii) a special suffix conjugation for adjective verbs (*nanou=f* “he is good”); and (iii) a marker for the future of the present and imperfect tense (*ti-na-sôtm* “I shall hear”).

## 4.7 Numerals

Numerals precede the noun to which they refer. The number 5 is etymologically derived from the word for “hand”; 20 is the dual of 10; 50 through 90 represent the plural forms of the respective units 5 through 9. Ordinals are derived from cardinals through the addition of a suffix *.nw* (from 2 to 9: *hmt.nw* “third”), later through the prefixation of the participle *mḥ* “filling” to the cardinal number: *mḥ-20* “twentieth”).

### (49) Egyptian numerals

**Table 7.4 Earlier Egyptian numerals and their Sahidic Coptic outcome**

1 <i>w</i> <sup>ʿ</sup> <i>w</i> <sup>ʿ</sup> <i>wuʿʿuw</i> /	10 <i>mḏw</i> <sup>*</sup> <i>/mu:čaw/</i>	100 <sup>*</sup> <i>š(n).t</i> <sup>*</sup> <i>/š(iny)ut/</i>
> <i>oua</i> <sup>ʿ</sup> <i>/waʿ/</i>	> <i>mēt</i> <sup>ʿ</sup> <i>/me:t/</i>	> <i>še/šeʿ/</i>
2 <i>sn.wj</i> <sup>*</sup> <i>/sinuw waj/</i>	20 <sup>*</sup> <i>ḏwtj</i> <sup>*</sup> <i>/čawa:taj/</i>	200 <sup>*</sup> <i>š(n).tj</i> <sup>*</sup> <i>/š(iny)u:taj/</i>
> <i>snau</i> <sup>ʿ</sup> <i>/snau/</i>	> <i>jouôt</i> <sup>ʿ</sup> <i>/čwo:t/</i>	> <i>šet</i> <sup>ʿ</sup> <i>/še:t/</i>
3 <i>hmtw</i> <sup>*</sup> <i>/xamtaw/</i>	30 <i>m</i> <sup>ʿ</sup> <i>b³</i> <sup>*</sup> <i>/maʿbVR/</i>	300–900 <sup>*</sup> <i>hmtw-š(n.w)t</i> , etc.
> <i>šomṇt</i> <sup>ʿ</sup> <i>/šomṇt/</i>	> <i>maab</i> <sup>ʿ</sup> <i>/ma²b/</i>	
4 <i>jfdw</i> <sup>*</sup> <i>/jift aw/</i>	40 <sup>*</sup> <i>hm.w</i> <sup>*</sup> <i>/ḥVmew/</i>	1,000 <i>ḥ³</i> <sup>*</sup> <i>/xar/</i>
> <i>ftouu</i> <sup>ʿ</sup> <i>/ftou/</i>	> <i>hme</i> <sup>ʿ</sup> <i>/hmeʿ/</i>	> <i>šo</i> <sup>ʿ</sup> <i>/šoʿ/</i>
5 <i>djw</i> <sup>*</sup> <i>/t i:jaw/</i>	50 <sup>*</sup> <i>dj.w</i> <sup>*</sup> <i>/t ijjaw/</i>	10,000 <i>ḏb</i> <sup>ʿ</sup> <i>/čVbaʿ/</i>
> <i>tiou</i> <sup>ʿ</sup> <i>/t i:u/</i>	> <i>taiou</i> <sup>ʿ</sup> <i>/tajjəu/</i>	> <i>tba</i> <sup>ʿ</sup> <i>/tbaʿ/</i>
6 <i>sjsw</i> <sup>*</sup> <i>/saʿsaw/</i>	60 <sup>*</sup> <i>sjs.w</i> <sup>*</sup> <i>/saʿsew/</i>	100,000 <i>ḥfn</i>
> <i>soou</i> <sup>ʿ</sup> <i>/sou/</i>	> <i>se</i> <sup>ʿ</sup> <i>/seʿ/</i>	
7 <i>sf ḥw</i> <sup>*</sup> <i>/safxaw/</i>	70 <sup>*</sup> <i>sf ḥ.w</i> <sup>*</sup> <i>/safxew/</i>	1,000,000 <i>ḥḥ</i> <sup>*</sup> <i>/ḥaḥ/</i>
> <i>sašf</i> <sup>ʿ</sup> <i>/sašf/</i>	> <i>šfe</i> <sup>ʿ</sup> <i>/šfeʿ/</i>	> <i>hah</i> <sup>ʿ</sup> <i>/hah/</i>
8 <i>ḥmnw</i> <sup>*</sup> <i>/xama:naw/</i>	80 <sup>*</sup> <i>ḥmn.w</i> <sup>*</sup> <i>/xamnew/</i>	
> <i>šmoun</i> <sup>ʿ</sup> <i>/šmu:n/</i>	> <i>hmene</i> <sup>ʿ</sup> <i>/xmneʿ/</i>	
9 <i>psḏw</i> <sup>*</sup> <i>/pisi:čaw/</i>	90 <sup>*</sup> <i>psḏ.j.w</i> <sup>*</sup> <i>/pisčijjaw/</i>	
> <i>psit</i> <sup>ʿ</sup> <i>/psi:t/</i>	> <i>pestaiau</i> <sup>ʿ</sup> <i>/pəstajjəw/</i>	

## 5. SYNTAX

### 5.1 Sentence-types and word order

Egyptian syntax knows three types of sentences: the so-called noun clauses, adverb clauses, and verb clauses.

#### 5.1.1 Noun clauses

In noun clauses, the predicate is a noun, whether substantive or adjective (Doret 1989–1992; Loprieno 1995:103–131). In categorical statements or qualifying adjectival sentences, the normal order of constituents is *Predicate–Subject* (50A); a demonstrative *pw* “this” functioning as copula may be inserted between the two phrases (50B):

- (50) A. *nfr* *mṇ-j*  
           be.good.PART. path-me  
           “My path is good”



- B. dmj.t      pw      jmn.t  
 city.FEM.    COP.    West.FEM.  
 “The West is a city”

The syntactic order Predicate(–Copula)–Subject is modified into a pragmatic order *Topic–Comment* in (i) classifying sentences when the subject is a first- or second-person pronoun (51A), (ii) identifying sentences when both the subject and the predicate are semantically determined or specified (51B), and (iii) in cleft sentences, in which the predicate is a participle and the subject is focalized (51C) (Loprieno 1988:41–52):

- (51) A. ntk    jtj      n      nmḥw  
           you    father    for    orphan  
           “You are a father to the orphan”  
       B. zh³w-f      pw      ḥrw  
           scribe-him    COP.    Horus  
           “His scribe is Horus”  
       C. jn            sn.t-j      s-‘nh                    rn-j  
           FOCUS    sister-me    CAUS.-live.PART.    name-me  
           “My sister is the one who makes my name live”

### 5.1.2 Adverb clauses

In adverb clauses, the predicate is an adverbial phrase or a prepositional phrase (Loprieno 1995:144–172). The word order is always Subject–Predicate. In Earlier Egyptian, main adverb clauses are often introduced by particles functioning as discourse markers (52A); in the absence of a discourse marker, the clause is to be understood as syntactically dependent (52B):

- (52) A. jw                    nzw    jr                    p.t  
           “SITUATION”    king    towards    heaven.FEM.  
           “Now the king is [directed] towards heaven”  
       B. ḥr.t-k                    m    pr-k  
           rations.FEM.-you    in    house-you  
           “[Because] your rations are in your house”

### 5.1.3 Verb clauses

In verb clauses, the predicate is a verbal phrase (Loprieno 1995:183–220); the word order is Predicate–Subject:

- (53) jj.n-j                    m      nw.t-j  
           come-RET.-I    from    city.FEM.-me  
           “I came from my city”

As we observed in §4.4.1, a peculiarity of Egyptian syntax is that the predicate of verb clauses may function as the theme of the utterance. In general, Egyptian verbal syntax displays a comparatively high incidence of topicalization and focalization phenomena. The most common topicalization device is the extraposition of the topicalized argument through the particle *jr* “concerning” (54A); used as a conjunction, the same particle introduces the protasis of a hypothetical clause (54B):

- (54) A. jr                      sf                      wsjr      pw  
           concerning yesterday Osiris COP.  
           ‘As for ‘yesterday’, it means ‘Osiris’ ’
- B. jr                      jqr-k                      grg-k                      pr-k  
           concerning be.important.PROSP.-you found.PROSP.-you house-you  
           ‘If you become wealthy, you should found a household’

Unmarked VPs not introduced by discourse markers are less frequent than in related languages, mostly functioning as embedded or modal clauses:

- (55) h‘y-k  
           appear.PROSP.-you  
           ‘May you appear’

## 5.2 Prepositional phrases

The most frequent prepositions are *m* ‘in, with’; *n* ‘to, for’; *r* ‘toward’; *mj* ‘as, like’; *hr* ‘on’; *hr* ‘under’; *hn* ‘with’; *hft* ‘according to’; *hnt* ‘before.’ Prepositional phrases follow the noun or the verb they modify. Particularly noteworthy is the presence of the preposition *hr* ‘near’; its original semantic value ‘beneath’ was applied to any situation in which the two participants A and B belong to different hierarchical levels:

- (56) A. dd-f                      hr                      ms.w-f  
           say.PROSP.-he beneath child.PL.-him  
           ‘He will say to his children’
- B. jm³hy                      hr                      ntr      ʕ  
           honor.PASS.PART. beneath god great  
           ‘Honored by the great god’

## 5.3 Coordination and subordination

The presence or absence of morphemes indicating *paragraph initiality* is an important syntactic feature of adverb and verb clauses in Egyptian. The general rule is as follows: (i) adverbial and verbal patterns introduced by a discourse particle are *initial* main clauses; (ii) whereas bare patterns are *noninitial* clauses – either (a) paratactically juxtaposed to the initial predication as noninitial coordinate main clauses or (b) controlled by it as subordinate clauses. This flexibility in sentence patterns, which can appear as main sentence or as subordinate clause, depending on the syntactic environment, is a common feature of Egyptian syntax, being shared by the majority of patterns, whether nominal, adverbial, or verbal.

The dialectics between the initial main sentence introduced by a particle and the noninitial coordinate bare adverb clause is captured in the following example:

- (57) jw                      hnw                      m sgr                      jb.w                      m gmw                      rw.tj  
           ‘SITUATION’ residence in silence heart.PL. in mourning portal.FEM.DUAL  
           wr.tj                      htm.w  
           great.FEM.DUAL shut.STAT.  
           ‘The Residence was in silence, the hearts in mourning, the Two Great Portals shut’

An example of coordinate verb clause syntax is provided by the following passage, in which a series of noninitial main clauses is paratactically linked to the the initial verb form:

- (58) jrj.t-j      šm.t    m    ḥnt.yt    nj    k3-j                    spr            r    ḥnw    pn  
 make.INF.-I   go.INF.   in   sail south   not   think.PERF.-I   reach.INF.   to   residence   this  
 ḥmt.n-j            ḥpr                    ḥ3.yt                    nj    dd-j            ḥb-j                    r-s3-f  
 think.PRET.-I   happen.PROSP.   turmoil.FEM.   not   say.PERF.-I   live.PROSP.-I   after-it  
 nmj.n-j            m3.tj    m    h3.w    nh.t                    zm3.n-j                    m    jw-snfrw  
 pass.PRET.-I   Maaty   in   area   sycamore   arrive.PRET.-I   in   island-Snefru  
 “I made a journey southward, and I did not plan to reach the residence; I thought  
 that there would be turmoil and I did not expect to survive after it; I crossed the  
 lake Maaty in the Sycamore neighborhood, and I arrived at Snefru Island”

It is important to appreciate the difference between *initiality* as a property of discourse and *independence* versus *subordination* as syntactic features of the clause. In the examples of (57) and (58), there are only main clauses, in the sense that – if taken individually – all clauses represent well-formed Egyptian sentences paratactically organized within a chain of discourse (Collier 1992). In both cases, however, only the first sentence is initial: in the case of (57), it is introduced by an overt particle of initiality (*jw*), which indicates that the corresponding adverbial sentence (*ḥnw m sgr*) opens a new segment of discourse; in the example of (58), the initial verb form, a so-called *narrative infinitive*, provides the temporal and aspectual references for the chain of paratactically linked clauses.

We need, therefore, to draw a distinction between the level of *clause* and the level of *discourse*. Adverbial and verbal sentences introduced by a particle are always main clauses; noninitial patterns may be paratactically linked main clauses or embedded subordinate clauses. The difference between forms with and without introductory particle lies on the discourse level, in that the sentence introduced by an initial proclitic particle opens a segment of text. In this respect, rather than operating with the traditional two levels of clausal linkage (parataxis vs. hypotaxis, or coordination vs. subordination), it seems suitable to analyze Egyptian syntactic phenomena by positing three forms of linkage between sentences:

1. *Parataxis*, i.e., the linkage between main clauses: this linkage usually remains unexpressed in Egyptian syntax, as in the case of bare adverbial, pseudoverbal, or verbal sentences which follow an initial main clause within a chain of discourse. Specimens of paratactic chains were provided in (57)–(58).

2. *Hypotaxis*, i.e., a semantic, rather than syntactic, dependency of a sentence on the discourse nucleus: hypotactically linked clauses are usually introduced by particles such as *jsk*, *jhr* or *js*; their semantic scope and their pragmatic setting can be properly understood only in reference to the message conveyed in the textual nucleus, as in example (59), which in the original text immediately follows the example of (57):

- (59) jst                    r-f    zbj.n                    ḥm-f                    mšc    r    t3-tmhj.w    z3-f  
 meanwhile   to-it   send-PRET.   Majesty-him   army   to   land-Libyans   son-him  
 smsw   m   ḥrj                    jry  
 elder   in   superior   thereof  
 “Meanwhile, His Majesty had sent off to the land of the Libyans an army whose  
 leader was his elder son”

3. *Subordination*, i.e., the syntactic dependency of a clause on a higher node, which itself can be a main or a subordinate clause: subordination is usually signaled by morphological markers such as prepositions (for example *m* “in” > “when”) governing nominalized verbal phrases, conjunctions (such as *hr-ntt* “because”), or particles (*jr* “if”):

- (60) ṛḫ.n-j            qd-k            tw-j            m    zšj            m    wn-k            m    šms.wt  
 know.PRET.-I    character-you    indeed-I    in    nest    in    be.AOR.-you    in    following  
 jt-j  
 father-me  
 “I knew your character while still in the nest, when you were in my father’s  
 following”

In the absence of an overt marker of dependency, subordination can also be determined by syntactic control. In this case, one speaks of “embedding,” as in the case of adverbial or verbal sentences functioning as virtual relative clauses or controlled by a verb of perception:

- (61) gmj.n-j            nb-j            ṛnḫ.w            wdḏ.w            snb.w            ḥntj-f  
 find.PRET.-I    lord-me    live.STAT.    whole.STAT.    healthy.STAT.    sail south.AOR.-him  
 “I found my Lord (may he be alive, prosperous and healthy) travelling southward”

### 5.3.1 Relativization

As an example of the complex interface between overt and embedded subordination, let us consider relativization. Specific antecedents (Loprieno 1995:202–208) are resumed by an *overt marker of relativization*, such as (i) the relative pronoun masculine *ntj*, feminine *nt.t*, plural *ntj.w* “who, which, that” in adverb clauses (62A); or (ii) an agreement-marker inflected in the relative verb form – (a) a participle in the presence of coreferentiality of antecedent and subject of the relative clause (62B); (b) a finite relative form in its absence (62C):

- (62) A. mtr-n            wj            rmt.w            km.t            ntj.w            jm            ḥnʿ-f  
 witness-PRET.    me            man.PL.    Egypt    who.PL.    there    with-him  
 “Egyptians who were there with him bore witness for me”  
 B. ḏj-s            ḫ.t            nb.t            nfr.t            wʿb.t  
 give.PROSP.-she    thing.FEM.    every.FEM.    be.good.PART.FEM.    be.pure.PART.FEM.  
 prr.t            ḫr            wdḥ-s  
 exit.PART.FEM.    on    altar-her  
 “May she give every good and pure thing which goes up on her altar”  
 C. ḥʿs.t            nb.t            rwj.t-n-j            r-s  
 country.FEM.    every.FEM.    advance.REL.FEM.-PRET.-I    against-it  
 “Every country against which I advanced”

Nonspecific antecedents, on the other hand, are modified by relative clauses which lack overt agreement-markers (Collier 1991; Loprieno 1995:158–161). They are syntactically subordinated through embedding into the main clause:

- (63) k.t            n.t            msdr            ḏj-f            mw  
 another.FEM.    that-of.FEM.    ear    give.AOR.-it    water  
 “Another (remedy) for an ear which gives off water”

## 5.4 Syntactic evolution

Syntactic patterns prove rather stable throughout the history of Egyptian. Late Egyptian (Satzinger 1981) and Coptic (Polotsky 1987:9–43) display the same variety of sentence-types as Earlier Egyptian:

1. *Noun clauses*: With an unmarked (syntactic) order Predicate–Subject when the subject is a noun (64A), and with a marked (pragmatic) order Topic–Comment in three environments: (i) when the subject is a pronoun (64B); (ii) when both the subject and the predicate are semantically specific (64C); and (iii) in cleft sentences, in which the predicate is a participle and the subject is focalized (64D):

- (64) A. ou-me      te      te-f-mnt-mntre  
           a-truth    COP.    the-his-thing-witness  
           “His testimony is true”  
       B. anok      ou-šôs  
           TOPIC-I    a-shepherd  
           “I am a shepherd”  
       C. t-arkhê            n-t-sophia                    te      t-mnt-mai-nûte  
           the-beginning    that.of-the-wisdom    COP.    the-thing-lover-god  
           “The beginning of wisdom is piety”  
       D. p-nûte      p-et-sooun  
           the-god      the.one-who-know.INF.  
           “God is the one who knows” (= “Only God knows”)

2. *Adverb clauses* (Polotsky 1990:203–224): in which the predicate is an adverbial or a prepositional phrase; the order is Subject–Predicate:

- (65) ti-hm-pa-eiôt  
       I-in-the-my-father  
       “I am in my father”

3. *Verb clauses* (Polotsky 1990:175–202): in which the predicate is a verbal phrase built according to the SVO-patterns described in §4.6.4; in these patterns, the subject can be extraposed to the right of the predicate and anticipated by a cataphoric pronoun in the regular syntactic slot:

- (66) a-u-rîme                    nci-ne-snêu  
       PRET.-they-weep.INF.    namely-the-brother.PL.  
       “The monks wept”

In Coptic verbal sentences, the tendency to have the verb phrase function as theme or rheme of the utterance reaches its full development: in the former case, the verb phrase is preceded by a relative marker *e-* or *nt-* and is described in Coptological literature as “second tense” (Polotsky 1987:129–140); in the latter, the form is preceded by the circumstantial marker *e-* and is described as “circumstantial” (Polotsky 1990:225–260):

- (67) nt-a-n-jpo-f                    e-f-o                    n-blle  
       REL.-PRET.-we-beget-him    “WHILE”-he-do.STAT.    as-blind  
       “He was born to us blind” (lit. “That we begot him was while he is as blind”)

## 6. LEXICON

Owing to Egypt’s geographically protected location, Ancient Egyptian does not display in its earlier phase (from 3000 BC) detectable influences from other languages, although the neighboring languages certainly did contribute to the lexical development of historical Egyptian. The majority of the lexicon is of Afro-Asiatic origin and displays convergences especially

with the Semitic and Libyan branches of this family (Schenkel 1990:49–57): for example, *sp.t* “lip,” cf. Arabic *šafat-un*, *sflw* “seven,” cf. Arabic *sab-un*, *jnm* “skin,” cf. Berber *a-glim*. There is also, however, some evidence for the impact of an Indo-European adstratum in the area of basic vocabulary (Kammerzell 1994:37–58): for example, Egyptian *jr.t* \*/jla:čat/ “milk,” compare Greek *gala*, *galak-tos*; or *hntj* \*/xant-ij/ “before,” compare Latin *ante*. In some cases, for the same concept, for example “heart,” Egyptian displays the coexistence of an Afro-Asiatic (*jb\*/jib/*, cf. Akkadian *lubb-um*) and of an Indo-European connection (*h<sup>3</sup>tj* \*/h<sup>3</sup>urtiy/, cf. Latin *cor*, *cord-is*), probably rooted in different dialectal areas of the country.

During the New Kingdom (c. 1500–1100 BC), contacts with the western Asiatic world led to the adoption of a considerable number of especially West Semitic loanwords (Hoch 1994), many of which remained confined to the scholarly and administrative sphere: for example *tp<sub>r</sub>* from Northwest Semitic *sôpēr* “scribe”; *mrkbt* (Coptic *berecôout*) from Northwest Semitic *merkābā* “chariot”; *mryn* from Mitanni (Iranian) *maryannu* “chariot-fighter.”

In the Late Period, after the seventh century BC, when the productive written language was Demotic, a limited number of (mostly technical) Greek words entered the Egyptian domain: *gawma* from *kauma* “fever”; *wynn* from *hoi Iōnes* “the Ionians” i.e., “the Greeks.” The impact of Greek vocabulary became more dramatic with the Christianization of the country, Hellenistic Greek being the language in which the Christian Scriptures were transmitted in the Eastern Mediterranean world. The number of Greek loanwords in Coptic is therefore very high (Kasser 1991a) – depending on the nature of the text, up to one-third of the lexical items found in a Coptic text may be of Greek origin. Most of these words stem from the spheres of (i) religious practice and belief (*angelos* “angel,” *diabolos* “devil,” *ekklēsia* “church,” *agios* “saint,” *sōtēr* “savior,” etc.); (ii) administration (*arkhōn* “governor,” *oikonomei* “to administer,” etc.); and (iii) high culture (*anagnōsis* “recitation,” *logikos* “spiritual,” etc.). In some texts translated from Greek, the influence of this language extends to the realm of syntax. A limited number of words from the military context are Latin (*douks* “general”), whereas documents from the end of the first millennium begin to display the adoption of loanwords from Arabic (*alpesour* from *al-bāsūr* “hemorrhoids”). The terms referring to the basic vocabulary, however, usually remained those of Egyptian origin: for example, “man” *rm<sub>t</sub>* > *rōme*; “woman” *hjm.t* > *shime*; “water” *mw* > *mau*; “two” *sn.wj* > *snau*.

## 7. READING LIST

The bibliography contains all the works referred to in this chapter. In the case of particularly important grammatical tools, I have added a short comment on their contents. In addition, some books have been listed which may prove valuable as further reading on a general topic on history or grammar of Ancient Egyptian or Coptic.

### Bibliography

- Allen, J. 1984. *The Inflection of the Verb in the Pyramid Texts*. Bibliotheca Aegyptia 2. Malibu: Undena. (A detailed morphological analysis of the earliest corpus of Egyptian texts.)
- Bagnall, R. 1993. *Egypt in Late Antiquity*. Princeton: Princeton University Press.
- Boas, G. and A. Grafton (eds.). 1993. *The Hieroglyphics of Horapollo*. Bollingen Series 23. Princeton: Princeton University Press.
- Černý, J. and S. Groll. 1984. *A Late Egyptian Grammar* (3rd edition). Studia Pohl, Series Maior 4. Rome: Pontifical Biblical Institute. (A structural grammar of colloquial, i.e., scholastic and administrative, Late Egyptian.)

- Comrie, B. 1981. *Language Universals and Linguistic Typology*. Chicago: University of Chicago Press.
- Crum, W. 1939. *A Coptic Dictionary*. Oxford: Clarendon Press. (The standard dictionary of Coptic, with very detailed philological references.)
- Collier, M. 1991. "The relative clause and the verb in Middle Egyptian." *Journal of Egyptian Archaeology* 77:23–42.
- . 1992. Predication and the circumstantial *sdm(.f)/sdm.n(.f)*. *Lingua Aegyptia* 2:17–65.
- Daumas, F. (ed.). 1988–1995. *Valeurs phonétiques des signes hiéroglyphiques d'époque gréco-romaine* (4 vols.). Montpellier: University of Montpellier.
- Davies, W. 1987. *Egyptian Hieroglyphs*. London: British Museum. (An introductory presentation of the writing system with many examples and references.)
- Diakonoff, I. 1965. *Semito-Hamitic Languages. An Essay in Classification*. Moscow: Academy of Sciences of the USSR.
- Doret, É. 1989–1992. "Phrase nominale, identité et substitution dans les Textes des Sarcophages." *Revue d'Égyptologie* 40:49–63; 41:39–56; 43:49–73.
- Edel, E. 1955–1964. *Altägyptische Grammatik*. *Analecta Orientalia* 34–39. Rome: Pontifical Biblical Institute.
- Erman, A. and H. Grapow. 1926–1953. *Wörterbuch der ägyptischen Sprache*. Berlin: Akademie-Verlag. (The fundamental dictionary for Old, Middle, and Late Egyptian.)
- Faulkner, R. 1962. *A Concise Dictionary of Middle Egyptian*. Oxford: Oxford University Press. (A lexicographical aid for the translation of Middle Kingdom texts.)
- Fecht, G. 1960. *Wortakzent und Silbenstruktur*. Ägyptologische Forschungen 21. Glückstadt: Verlag J. J. Augustin. (The standard analysis of the syllabic patterns of Egyptian.)
- Fischer, H. 1977. "Hieroglyphen." In W. Helck and E. Otto (eds.), *Lexikon der Ägyptologie*, vol. II, pp. 1189–1199. Wiesbaden: Otto Harrassowitz.
- Fleming H. 1976. "Cushitic and Omotic." In M. Bender, J. D. Bowen, R. L. Cooper, et al. (eds.), *Language in Ethiopia*, pp. 34–53. Oxford: Oxford University Press.
- Fowden, G. 1986. *The Egyptian Hermes. A Historical Approach to the Late Pagan Mind*. Princeton: Princeton University Press.
- Gardiner, A. 1957. *Egyptian Grammar, Being an Introduction to the Study of Hieroglyphs* (3rd edition). Oxford: Oxford University Press. (Still the most detailed tool for the philological study of the classical language.)
- Gelb, I. 1963. *A Study of Writing* (revised edition). Chicago: University of Chicago Press.
- Giveon, R. 1982. "Protosinaitische Inschriften." In W. Helck and W. Westendorf (eds.), *Lexikon der Ägyptologie*, vol. IV, pp. 1156–1159. Wiesbaden: Otto Harrassowitz.
- Graefe, E. 1997. *Mittelägyptische Grammatik für Anfänger* (5th edition). Wiesbaden: Otto Harrassowitz. (A user-friendly manual for academic instruction in classical Egyptian.)
- Hetzron, R. 1992. "Semitic languages." In W. Bright (ed.), *International Encyclopedia of Linguistics*, vol. III, pp. 412–417. Oxford: Oxford University Press.
- Hintze, F. 1950. "Konversion und analytische Tendenz in der ägyptischen Sprachentwicklung." *Zeitschrift für Phonetik und Allgemeine Sprachwissenschaft* 4:41–56.
- . 1980. "Zur koptischen Phonologie." *Enchoria* 10:23–91. (A very influential generative analysis of Coptic phonology.)
- Hoch, J. 1994. *Semitic Words in Egyptian Texts of the New Kingdom and Third Intermediate Period*. Princeton: Princeton University Press. (A companion for issues of comparative Egyptian–Semitic phonology and for Semitic vocabulary in Egyptian.)
- Hodge, C. (ed.). 1971. *Afroasiatic. A Survey*. *Janua Linguarum Series Practica* 163. The Hague/Paris: Mouton.
- Iversen, E. 1961. *The Myth of Egypt and its Hieroglyphs in European Tradition*. Princeton: Princeton University Press. (For the history and the cultural milieu of the decipherment of hieroglyphs.)
- Johnson, J. 1991. *Thus Wrote 'Onchsheshonqy. An Introductory Grammar of Demotic* (2nd edition). *Studies in Ancient Oriental Civilization* 45. Chicago: Oriental Institute. (A short but comprehensive introduction to Demotic script and grammar.)
- Junge, F. 1985. "Sprachstufen und Sprachgeschichte." In *Zeitschrift der Deutschen Morgenländischen Gesellschaft*, Supplement 6, pp. 17–34. Stuttgart: Steiner. (The most modern presentation of the history of Egyptian.)
- . 1996. *Einführung in die Grammatik des Neuägyptischen*. Wiesbaden: Otto Harrassowitz.



- Kahl, J. 1994. *Das System der ägyptischen Hieroglyphenschrift in der 0.-3. Dynastie*. Göttinger Orientforschungen IV/29. Wiesbaden: Otto Harrassowitz.
- Kammerzell, F. 1991. *Personalpronomina und Personalendungen im Altägyptischen. Ägypten im afro-orientalischen Kontext. Aufsätze zur Archäologie, Geschichte und Sprache eines unbegrenzten Raumes. Gedenkschrift Peter Behrens. Afrikanistische Arbeitspapiere*, pp. 177–203. Special issue edited by D. Mendel and U. Claudi. Cologne: University of Cologne. (The most thorough treatment of Egyptian pronouns, also important for comparative issues.)
- . 1994. *Panther, Löwe und Sprachentwicklung im Neolithikum*. *Lingua Aegyptia Studia Monographica* 1. Göttingen: Seminar für Ägyptologie und Koptologie.
- . 1995. "Zur Umschreibung und Lautung." In R. Hannig, *Großes Handwörterbuch Ägyptisch-Deutsch (2800-950 v.Chr.)*, xxiii–lix. Mainz: Philipp von Zabern.
- Kasser, R. 1991a. "Vocabulary, Copto-Greek." In Aziz S. Atiya (ed.), *The Coptic Encyclopedia*, vol. VIII, pp. 215–222. New York: Macmillan.
- . 1991b. "Dialects; Dialects, Grouping and Major Groups of." In Aziz S. Atiya (ed.), *The Coptic Encyclopedia*, vol. VIII, pp. 87–88, 97–101. New York: Macmillan.
- Lambdin, T. 1983. *Introduction to Sahidic Coptic*. Macon: Mercer University Press. (The standard academic handbook for teaching classical Coptic.)
- Loprieno, A. 1980. "Osservazioni sullo sviluppo dell'articolo prepositivo in egiziano e nelle lingue semitiche." *Oriens Antiquus* 19:1–27.
- . 1986. "Zahlwort." In W. Helck and W. Westendorf (eds.), *Lexikon der Ägyptologie*, vol. VI, pp. 1306–1319. Wiesbaden: Otto Harrassowitz.
- . 1988. "On the typological order of constituents in Egyptian." *Journal of Afroasiatic Languages* 1:26–57.
- . 1991. *Topics in Egyptian Negations. Ägypten im afro-orientalischen Kontext. Aufsätze zur Archäologie, Geschichte und Sprache eines unbegrenzten Raumes. Gedenkschrift Peter Behrens. Afrikanistische Arbeitspapiere*, pp. 213–235. Special issue edited by D. Mendel and U. Claudi. Cologne: University of Cologne.
- . 1995. *Ancient Egyptian: A Linguistic Introduction*. Cambridge: Cambridge University Press. (A general presentation of linguistic features and historical grammar of Egyptian.)
- Lüddeckens, E. 1975. "Demotisch." In W. Helck and E. Otto (eds.), *Lexikon der Ägyptologie*, vol. I, pp. 1052–1056. Wiesbaden: Otto Harrassowitz.
- Neveu, F. 1996. *La langue des Ramsès. Grammaire du Néo-Égyptien*. Paris: Chéops.
- Newman, P. 1992. "Chadic." In W. Bright (ed.), *International Encyclopedia of Linguistics*, vol. I, pp. 253–254. Oxford: Oxford University Press.
- Osing, J. 1975. "Dialekte." In W. Helck and E. Otto (eds.), *Lexikon der Ägyptologie*, vol. I, pp. 1074–1075. Wiesbaden: Otto Harrassowitz.
- . 1976. *Die Nominalbildung des Ägyptischen* (2 vols.). Mainz: Philipp von Zabern. (The fundamental reference work on the vocalic patterns of the language from Middle Egyptian through Coptic.)
- . 1980. "Lautsystem." In W. Helck and W. Westendorf (eds.), *Lexikon der Ägyptologie*, vol. III, pp. 944–949. Wiesbaden: Otto Harrassowitz.
- Polotsky, H. 1960. "The Coptic conjugation system." *Orientalia* 29:392–422.
- . 1971. *Collected Papers*. Jerusalem: Magnes Press. (For the development of the "Standard theory" of Egyptian syntax.)
- . 1976. "Les transpositions du verbe en égyptien classique." *Israel Oriental Studies* 6:1–50.
- . 1987–1990. *Grundlagen des koptischen Satzbaus*. Decatur, GA: Scholars Press.
- Ray, J. 1992. "Are Egyptian and Hittite related?" In A. Lloyd (ed.), *Studies in Pharaonic Religion and Society in Honour of J. Gwyn Griffiths*, pp. 124–136. London: Egypt Exploration Society.
- Reintges, C. 1994. "Egyptian root-and-pattern morphology." *Lingua Aegyptia* 4:213–244.
- . 1997. *Passive Voice in Older Egyptian. A Morpho-Syntactic Study*. Holland Institute of Generative Linguistics. Dissertations 28. The Hague: Holland Academic Graphics.
- Sasse, H.-J. 1992. "Cushitic languages." In W. Bright (ed.), *International Encyclopedia of Linguistics*, vol. I, pp. 326–330. Oxford: Oxford University Press.
- Satzinger, H. 1981. "Nominalsatz und Cleft Sentence im Neuägyptischen." In D. Young (ed.), *Studies Presented to Hans Jakob Polotsky*, pp. 480–505. Beacon Hill: Pirtle and Polson.



















- Schenkel, W. 1990. *Einführung in die altägyptische Sprachwissenschaft. Orientalistische Einführungen*. Darmstadt: Wissenschaftliche Buchgesellschaft. (An essential companion for the study of the history of linguistic thinking in Egyptology and an indispensable tool for the study of the prehistory of Egyptian phonology and its comparative aspects.)
- . 1993. "Zu den Verschuß- und Reibelauten im Ägyptischen und (Hamito-)Semitischen. Ein Versuch zur Synthese der Lehrmeinungen." *Lingua Aegyptia* 3:137–149.
- . 1984. "Schrift." In W. Helck and W. Westendorf (eds.), *Lexikon der Ägyptologie*, vol. V, pp. 713–735. Wiesbaden: Otto Harrassowitz. (A systematic presentation of the features of the hieroglyphic system.)
- Schneider, T. 1992. *Asiatische Personennamen in ägyptischen Quellen des Neuen Reiches*. Orbis Biblicus et Orientalis 114. Freiburg/Göttingen: Vandenhoeck and Ruprecht.
- Shisha-Halevy, A. 1986. *Coptic Grammatical Categories*. Analecta Orientalia, 53. Rome: Pontifical Biblical Institute. (A complete grammatical study of Sahidic Coptic.)
- Spiegelberg, W. 1925. *Demotische Grammatik*. Heidelberg: Carl Winter.
- Till, W. 1970. *Koptische Grammatik* (2nd edition). Leipzig: Verlag Enzyklopädie.
- Vernus, P. 1990. *Future at Issue. Tense, Mood and Aspect in Middle Egyptian: Studies in Syntax and Semantics*. Yale Egyptological Studies 4. New Haven: Yale Egyptological Seminar. (Especially important for the interface of aspect and mood in early Middle Egyptian.)
- . 1996. "Langue littéraire et diglossie." In A. Loprieno (ed.), *Ancient Egyptian Literature. History and Forms*, pp. 555–564. Probleme der Ägyptologie 10. Leiden: Brill.
- Vittmann, G. 1991. "Zum koptischen Sprachgut im Ägyptisch-Arabisch." *Wiener Zeitschrift für die Kunde des Morgenlandes* 81:197–227.
- Wenig, S. 1982. "Meroe, Schrift und Sprache." In W. Helck and W. Westendorf (eds.), *Lexikon der Ägyptologie*, vol. IV, pp. 104–107. Wiesbaden: Otto Harrassowitz.
- Willms, A. 1980. *Die dialektale Differenzierung des Berberischen*. Afrika und Übersee Berlin: Reimer.
- Winand, J. 1992. *Etudes de néo-égyptien, I. La morphologie verbale*. Aegyptiaca Leodiensia 2. Liège: CIPL. (A complete morphological and syntactic analysis of the verb in Late Egyptian.)
- Zaborski, A. 1992. "Afroasiatic languages." In W. Bright (ed.), *International Encyclopedia of Linguistics*, vol. I, pp. 36–37. Oxford: Oxford University Press.
- Zeidler, J. 1992. "Altägyptisch und Hamitosemitisch. Bemerkungen zu den Vergleichenden Studien von Karel Petrůček." *Lingua Aegyptia* 2:189–222.
- . 1993. "A new approach to the Late Egyptian 'syllabic orthography'." In *Sesto Congresso internazionale di Egittologia*. Atti, vol. II, pp. 579–590. Torino: Italgas.

























# Sign list

















Listed below are the hieroglyphic signs most often found in Middle Egyptian texts, arranged into 27 groups on the basis of what they depict. The selection and order are those most commonly used by Egyptologists, based on the list in Gardiner's *Egyptian Grammar*, with some additional signs.<sup>1</sup> Each sign is identified as to what it depicts (as far as possible) and its uses, whether phonogram, ideogram, or determinative, arranged in order of frequency; words in SMALL CAPITALS indicate the class of words with which a sign is used as determinative. At the end of the sign list is a supplemental list of signs arranged by shape.

## A. Human Beings, Male









1		seated man	Phonogram <i>j</i> (1s suffix pronoun). Determinative MAN; also in 1s pronouns <i>jnk</i> , <i>uj</i> , <i>.kw/kj</i> . Ideogram for <i>zj</i> "man" or <i>ḥw</i> "companion." With B1 and plural strokes, determinative PEOPLE and ideogram for <i>rmꜥ</i> "people."
2		man with hand to mouth	Variants  (A68),  (A84). Determinative SPEAK, THINK, EAT, DRINK, and for emotions such as LOVE and HATE.
3		man sitting on heel	Determinative SIT.
4		man with hands raised	Determinative WORSHIP; also HIDE (for A5).
5		man hiding behind wall	Determinative HIDE.
6		A1 + W54	Variant of D60.
7		fatigued man	Determinative WEARY, WEAK, SOFT.
8		man performing <i>hnw</i>	Determinative in <i>hnw</i> "jubilation."
9		man with basket on head	Variant  (A119) in <i>ḥ3j</i> . Determinative LOAD, CARRY, WORK. Ideogram for <i>ḥtp</i> "load," <i>ḥ3j</i> "carry, lift," <i>k3t</i> "work."
10		man with oar	Determinative SAIL, ROW.
11		man with scepter and crook	Determinative FRIEND.
12		soldier	Determinative SOLDIER. Ideogram with plural strokes for <i>mꜥꜥ</i> "expeditionary force, army."
13		prisoner	Determinative ENEMY.

1 A number of signs that Gardiner placed in category Aa ("Unclassified") have since been identified. The sign R13 is included as a separate entry in G. The supplemental sign R61 is listed under I as well as R, and Y10 under M as well as Y. Additional signs are numbered, where possible, after the list in N. Grimal, J. Hallof, and D. van der Plas, eds., *Hieroglyphica* (Publications interuniversitaires de recherches égyptologiques informatisées, 1: Utrecht, Centre for Computer-aided Egyptological Research, Utrecht University, 1993). Such signs are placed where they belong in each group rather than in their numerical position: thus, for example, A359 after A28.






- 14  wounded man Variant  (A14a). Determinative DIE, ENEMY.
- 15  man falling Variant  (A97). Determinative FALL, DIE. Ideogram for *hr* "fall."
- 16  man bowing Determinative BOW.
- 17  child Variant  (A17a). Determinative CHILD, YOUNG; in hieratic also SIT (for A3), DIGNITARY (for A21). Ideogram for *hrd* "child." Phonogram *nnj* "child" in *nnj-nsut* "Herakleopolis."
- 18  child with Red Crown Determinative CHILD-KING.
- 19  old man with staff Determinative OLD, DISTINGUISHED. Ideogram for *j3w* "old," *smsw* "eldest," *wr* "great, chief." Phonogram *jk* in *jky* "miner" (from *j3k* "age"). In hieratic sometimes for A25.
- 20  old man with forked staff Variant of last. Determinative in *smsw* "elder," also ideogram for same.
- 21  dignitary Determinative DIGNITARY. Ideogram for *stj* "official." Also as variant of A11 and A22. In hieroglyphic not always distinguishable from A19–20.
- 22  statue on base Determinative STATUE. The form often varies.
- 23  king Determinative KING.
- 24  man striking Determinative FORCE, EFFORT. Ideogram for *nht* "victory."
- 25  man striking Determinative in *hwj* "hit," often  (striking the phonogram).
- 59  man threatening Determinative DRIVE OFF.
- 26  man beckoning Variant  (A366). Determinative CALL. Ideogram for *j* "oh!" and *ʕ* "call."
- 27  man running Phonogram *jn* in *jn* "by" (from *jn* "messenger").
- 28  excited man Determinative HIGH, JOY, MOURN, FRUSTRATION.
- 59  man with arms clasped Determinative in *hsj* "freeze." Also rare variant of A1 (man pointing to himself).
- 29  man upside down Determinative INVERT.
- 30  man worshipping Determinative WORSHIP, RESPECT.
- 31  man shunning Determinative TURN AWAY.
- 32  man dancing Determinative DANCE.
- 33  man with stick and bundle Variant  (A166). Determinative in *mnjw* "herdsman," also ideogram for same. Determinative WANDER, STRANGER.
- 34  man pounding Determinative in *hwsj* "pound, construct"
- 35  man building a wall Determinative in *qd* "build," also ideogram for same.
- 37  man in vat Variant  (A36). Determinative in *ʕtj* "brewer," also ideogram for same.
- 38  man with two animals Variant  (A39, with two giraffes). Ideogram for *qjs/qsj* "Qus" (town).














40		seated god	Determinative GOD, KING. Variant of A1 for 1s pronouns when the speaker is a god or the king.
41		seated king	Variant  (A42). Determinative KING. Variant of A1 for 1s pronouns when the speaker is the king.
43		king with White Crown	Variant  (A44). Determinative of <i>nsut</i> "king," also ideogram for same. Determinative of <i>wsjr</i> "Osiris."
45		king with Red Crown	Variant  (A46). Determinative of <i>bjtj</i> "king of Lower Egypt," also ideogram for same.
47		shepherd seated	Determinative in <i>z3w</i> "guard," also ideogram for same. Ideogram for <i>mnjw</i> "herdsman." Sometimes variant of A48.
48		seated man with knife	Phonogram <i>jr</i> in the nisbe <i>jrj</i> "pertaining to."
49		foreigner with stick	Determinative FOREIGNER.
50		noble on chair	Determinative DIGNITARY, DECEASED. Variant of A1 for 1s pronouns when the speaker is deceased. Also variant of A51.
51		noble on chair, with flail	Determinative in <i>špsj/šps</i> "noble," also ideogram for same. Determinative DIGNITARY, DECEASED.
52		seated noble with flail	Determinative DIGNITARY, DECEASED.
53		mummy standing	Determinative MUMMY, STATUE, LIKENESS, FORM. Ideogram for <i>twt</i> "likeness, statue."
54		mummy recumbent	Determinative DEAD.
55		mummy on bed	Determinative LIE, DEAD. Ideogram for <i>sḏr</i> "lie down." The mummy is sometimes replaced by a man when used in/for <i>sḏr</i> "lie down."

## B. Human Beings, Female























1		seated woman	Variant  (B24). Determinative FEMALE. Rarely variant of A1 when the speaker is female.
2		pregnant woman	Determinative PREGNANT.
3		woman giving birth	Variant  (B4). Determinative in <i>msj</i> "give birth," also ideogram for same.
5		woman nursing	Determinative in <i>mn<sup>c</sup>t</i> "nurse."
6		nurse with child	Determinative in <i>mn</i> "rear, foster."
7		seated queen	Determinative in queens' names.

## C. Anthropomorphic Gods
















1		god with sun-disk	Variant  (falcon-headed, C2). Determinative in <i>r<sup>c</sup></i> "Re," also ideogram for same.
3		ibis-headed	Determinative in <i>ḏhwtj</i> "Thoth," also ideogram for same.
4		ram-headed	Variant  (C5). Determinative in <i>ḏnmw</i> "Khnum," also ideogram for same.

- 6  jackal-headed Determinative in *jnpw* "Anubis" and *wꜥ-wꜥwt* "Wepwawet," also ideogram for same.
- 7  Seth-animal-headed Determinative in *sth/stš* "Seth," also ideogram for same.
- 8  Min figure Determinative in *mnw* "Min," also ideogram for same.
- 9  goddess with horned disk Determinative in *hwt-hrw* "Hathor," also ideogram for same.
- 10  goddess with feather Variants  (C10a),  (C175a). Determinative in *mꜣt* "Maat" (as goddess), also ideogram for same.
- 11  *hh*-figure Ideogram for *hh* "million (§ 9.1)" and "Heh" (god supporting the sky).
- 12  Amun figure Determinative in *jnmw* "Amun," also ideogram for same.
- 17  Montu figure Determinative in *mnṯw* "Montu," also ideogram for same.
- 18  Tatjenen figure Determinative in *tꜣ-tꜣnj* "Ta-tjenen," also ideogram for same.
- 19  Ptah figure Variant  (C20). Determinative in *pṯh* "Ptah," also ideogram for same.


















## D. Parts of the Human Body




















- 1  head Ideogram for *tp* and *dꜣdꜣ* "head." Phonogram *tp* in *tpj* "first." Determinative HEAD.
- 2  face Ideogram for *ḥr* "face." Phonogram *ḥr*.
- 3  hair Determinative HAIR, SKIN, COLOR; also words associated with hair: BALD, MOURN, WIDOW. Ideogram for *wš* "missing."
- 4  eye Phonogram *jr*. Determinative for actions associated with the eye. Ideogram for *jrt* "eye."
- 5  eye with paint Variants  (D6) and  (D7a). Determinative for actions associated with the eye.
- 140  two eyes Determinative in *ptr* "see, look," also ideogram for same.
- 7  eye with paint Determinative ADORN. Also determinative in *ʿn* "beautiful" and *ʿnw* "Tura" (quarry near Cairo), from the Semitic root *ʿjn* "eye."
- 8  eye enclosed Variant of D7 as determinative in *ʿn* "beautiful" and *ʿnw* "Tura."
- 9  eye weeping Determinative in *rmj* "weep," also ideogram for same.
- 10  eye with falcon markings Determinative in *wꜣꜣt* "Sound Eye (of Horus)," also ideogram for same.
- 11  part of D10 Ideogram for 1/2 heqat (§ 9.7.3).
- 12  part of D10 Ideogram for 1/4 heqat (§ 9.7.3). Also determinative in *dꜣd* "pupil" and *mꜣꜣ* "see," the latter as variant of D4.
- 13  part of D10 Ideogram for 1/8 heqat (§ 9.7.3). Also determinative EYEBROW.
- 14  part of D10 Ideogram for 1/16 heqat (§ 9.7.3).
- 15  part of D10 Ideogram for 1/32 heqat (§ 9.7.3).
- 16  part of D10 Ideogram for 1/64 heqat (§ 9.7.3).
- 17  D15 + D16 Determinative of *tjt* "image," also ideogram for same.
- 18  ear Determinative in *msḥr* "ear," also ideogram for same.
- 19  face in profile Variant  (D20). Determinative NOSE, FACE, and associated actions. Ideogram for *fnꜣ* "nose." Phonogram *ḥnt*. In hieratic not always distinguishable from U31 or Aa32.

21		mouth	Phonogram <i>r</i> . Ideogram for <i>r</i> “mouth.”
154		mouth plus water	Determinative in <i>j<sup>c</sup>w-r</i> “breakfast,” also ideogram for same.
22		mouth plus 2 strokes	Ideogram for <i>rwj</i> 2/3 (§ 9.6).
23		mouth plus 3 strokes	Ideogram for <i>hmt-rw</i> 1/4 (§ 9.6).
24		lip with teeth	Variant  (D24a). Determinative in <i>spt</i> “lip,” also ideogram for same. Sometimes in error for F42.
25		two lips and teeth	Determinative in <i>sptj</i> “lips,” also ideogram for same.
26		lips and water	Determinative SPIT, SPEW.
27		breast	Variant  (D27a). Determinative BREAST, NURSE. Ideogram for <i>mnd</i> “breast.”
28		two arms	Phonogram <i>k3</i> . Ideogram for <i>k3</i> “ka” (variant  D29).
30		two arms and tail	Determinative in <i>nḥb-k3w</i> “Assigner of Kas” (a god).
32		two arms embracing	Variant  in hieratic. Determinative EMBRACE, OPEN.
31		D32 plus U36	Variant  . Ideogram for <i>hm-k3</i> “ka-servant” (mortuary priest).
33		arms and oar	Phonogram <i>hn</i> (from <i>hnj</i> “row”).
34a		arms with shield and mace	Variant  (D34). Ideogram for “ <i>h3</i> “fight”
35		gesture of negation	Ideogram for <i>nj</i> “not” and phonogram <i>nj</i> or <i>n</i> (§ 8.2.6), especially in <i>nn</i> “not”; <i>jw</i> or <i>jwt</i> in <i>jut</i> “that not” and <i>juwj</i> “which not” (§§ 12.9, 26.29.5). Determinative NEGATION.
36		forearm	Phonogram <i>c</i> . Ideogram for <i>c</i> “arm, hand.” Often variant for D37–44.
36a		forearm	Ideogram for <i>cwj</i> “arms, hands.”
12a		forearm with water	Determinative in <i>j<sup>c</sup>j</i> “wash,” also ideogram for same.
37		forearm with X8	Phonogram <i>dj</i> in forms of <i>rdj</i> “give.” Also variant of D38.
38		forearm with bread	Phonogram <i>mj</i> or <i>m</i> . Determinative in <i>jmj</i> “give!” (§ 16.2.3).
39		forearm with pot	Determinative OFFER. Sometimes variant of D37–38.
18a		O43 + D36	Ideogram for <i>šzp</i> “receive.”
40		forearm with stick	Determinative FORCE, EFFORT. Ideogram for <i>ḥ3j</i> “measure, evaluate.” Rarely variant of D37.
41		forearm with palm down	Determinative ARM and actions associated with the arm or hand. Ideogram <i>rnn</i> “shoulder.” Phonogram <i>nj</i> .
42		forearm with palm down	Determinative in <i>mḥ</i> “cubit” (§ 9.7.1), also ideogram for same.
43		forearm with flail	Phonogram <i>hw</i> .
44		forearm with scepter	Determinative in <i>hwp</i> “manage,” also ideogram for same.
45		forearm with brush	Variant  (D251). Determinative in <i>ḏsr</i> “sacred, clear away, raise the arm,” also ideogram for same.
46		hand	Phonogram <i>d</i> . Ideogram for <i>drt</i> “hand.”
46a		hand with water	Ideogram for <i>jdrt</i> “fragrance.”
47		hand	Determinative of <i>drt</i> “hand” when spelled with phonograms.
48		hand without thumb	Ideogram for <i>šzp</i> “palm” (§ 9.7.1).
49		fist	Determinative GRASP.
50		finger	Ideogram for <i>ḏb<sup>c</sup></i> “finger” and <i>ḏb<sup>c</sup></i> “10,000” (§ 9.1). When doubled, determinative ACCURATE.


- 51  finger  
Determinative for actions associated with the finger: *h3j* "measure," *t3j* "take," *dqr* "press." Determinative in *ni* "fingernail," also ideogram for same. Determinative FRUIT, FLOWER, also ideogram for *dqrw* "fruit," *q3w* "flour."
- 52  penis  
Determinative MALE. Phonogram *mt*. With E1, ideogram for *k3* "bull."
- 53  penis with fluid  
Determinative PENIS and associated actions, also MALE. Determinative of *b3h* in *m b3h* "in the presence of," *dr b3h* "since," *r b3h* "before," also ideogram for same.
- 279  testicles  
Determinative in *hrwj* "testicles," also ideogram for same.
- 280a  pelvis and vulva  
Phonogram *hm*. Ideogram for *jdt* "vulva, cow."
- 54  walking legs  
Determinative MOTION. Phonogram *juw* in forms of the verb *juw* "come." Ideogram for *nmtt* "step."
- 55  legs walking backwards  
Determinative REVERSE.
- 56  leg  
Determinative FOOT and associated actions. Ideogram for *rd* "foot." Phonogram *pd* (from *p3d* "knee"). Ideogram for *w<sup>c</sup>rt* "district" (from *w<sup>c</sup>rt* "shin"), *sbq* "excellent" (from *sbq* "leg"), *ghs* "gazelle."
- 57  leg with knife  
Determinative MUTILATE. Ideogram for *j3tw* "place of execution" and *sj3tj* "cheater" (from *j3t* "short").
- 58  foot  
Phonogram *b*. Ideogram for *buw* "place, thing."
- 59  D36 + D58  
Phonogram *b*.
- 60  D58 + W54  
Ideogram for *w<sup>c</sup>b* "clean, pure."
- 61  stylized toes  
Variants  (D62) and  (D63). Determinative in *s3h* "toe; kick, touch with the foot," also ideogram for same.

## E. Mammals


- 1  bull  
Determinative CATTLE. Ideogram for *k3* "bull, ox" *jhw* "cattle."
- 166  bulls  
Plural of E1.
- 177  two bulls joined  
Determinative in *hns* "go back and forth."
- 176  bull tied for slaughter  
Determinative *rhs* "slaughter," also ideogram for same. Ideogram for *k3* "bull" as offering.
- 2  bull charging  
Determinative in *sm3* "wild bull." Ideogram for *k3* in *k3 nht* "victorious bull" (epithet of the king).
- 3  calf  
Determinative in *bhz* "calf" and *wmdw* "short-horned cattle."
- 4  sacred cow  
Determinative in *hz3t* "sacred cow."
- 5  cow and calf  
Determinative in *3ms* "solicitous."
- 6  horse  
Determinative HORSE. Ideogram for *ssmt* "horse."
- 7  donkey  
Determinative in *3* (originally *j<sup>c</sup>3*) "donkey."
- 8  kid  
Variant  (E8a). Phonogram *jb*. Determinative GOAT.
- 9  newborn bubalis  
Phonogram *juw*.
- 10  ram  
Variant  (E11). Determinative SHEEP. Ideogram for *b3* "ram," *hnmw* "Khnun."
- 12  pig  
Determinative PIG.
- 13  cat  
Determinative in *mjuw/mjt* "cat."


- 14  dog (saluki)  
 15  jackal recumbent  
 17  jackal  
 18  jackal on standard  
 20  Seth animal  
 22  lion  
 23  lion recumbent  
 128  two lions joined  
 24  panther or leopard  
 25  hippopotamus  
 26  elephant  
 27  giraffe  
 28  oryx  
 29  gazelle  
 30  ibex  
 31  goat with collar  
 32  baboon  
 33  monkey  
 34  hare

Determinative DOG.

Variant  (D16). Determinative in *ḥnpw* "Anubis," also ideogram for same. Ideogram (D15) for title *ḥḥ-sḥt3* "master of secrets."

Determinative in *z3b* "jackal; dignitary," also ideogram for same.

Variant  (E19). Determinative in *wṗ-w3wt* "Parter of the Ways (Wepwawet)," also ideogram for same.

Variant  (E21). Ideogram for *stḥ/stḥ* "Seth." Determinative TURMOIL, CHAOS. In hieratic often for E7 and E27.

Determinative in *m3j* "lion," also ideogram for same.

Phonogram *rw* (from *rw* "lion"). In hieratic often for U13.

Determinative in *3kr* "Horizon (god)," also ideogram for same.

Determinative in *3by* "panther, leopard," also ideogram for same.

Determinative HIPPOPOTAMUS.

Determinative in *3bw* "elephant" Ideogram for *3bw* "Elephantine" (in modern Aswan).

Determinative in *sr* "foretell." Determinative in *mmj* "giraffe," also ideogram for same.

Determinative in *m3ḥd* "oryx."

Determinative in *ghs* "gazelle."

Determinative in *nj3w, nr3w, n3w* "ibex."











Determinative in *sḥ* "privilege," also ideogram for same.


Determinative BABOON, MONKEY, FURIOUS.

Determinative in *gjf* "monkey."

Phonogram *wn*.

## F. Parts of Mammals


- 1  head of ox  
 2  head of charging bull  
 3  head of hippopotamus  
 4  forepart of lion  
 5  head of bubalis  
 7  head of ram  
 9  head of leopard  
 11  head and neck of animal  
 12  head and neck of jackal  
 13  horns


Variant  (F63). Ideogram for *k3* "cattle" (in offering formulas).

Determinative in *dnd* "rage."

Determinative in *3t* "power," and *3t* "moment," also ideogram for latter.

Ideogram for *ḥ3t* "front" and related words.


Variant  (F6). Determinative in *ḥ3* "skilled," and related words, also ideogram for same. Determinative in *sḥ3* "prayer" and *bḥnt* "pylon."

Variant  (F8). Determinative in *ḥyt* "worth" (from *ḥt* "ram's head"), also ideogram for same.



































Determinative in *phṯj* "strength," also ideogram for same (often doubled).


Variant  (F10). Determinative NECK, THROAT and related actions.






























Phonogram *wst*.

Phonogram *wṗ*. Ideogram for *wṗt* "brow." For  see O44.






- 14  F13 + M4
- 16  horn
- 17  F16 + W54
- 18  tusk
- 19  jawbone of ox
- 20  tongue
- 21  ear of bovine
- 22  hindquarters of feline
- 23  foreleg of ox
- 25  leg and hoof of ox
- 26  goatskin
- 27  cowskin
- 28  cowskin
- 29  cowskin with arrow
- 30  water-skin
- 31  three fox-skins
- 32  animal's belly and udder
- 33  tail
- 34  heart
- 35  heart and windpipe
- 36  lung and windpipe
- 37  spine and ribs
- 39  spine and spinal cord
- 40  spine and spinal cord
- 41  vertebrae
- 42  rib
- 43  ribs
- 44  joint of meat
- 45  cow uterus
- 46  intestine
- 50  S29 + F46
- Variant  (F15). Ideogram for *wpt-rmpt* "Opening of the Year" (New Year's Day).
- Phonogram *b*. Determinative HORN, also ideogram for same.
- Determinative in *bw* "purification," also ideogram for same.
- Determinative TOOTH and associated actions. Phonograms *bh* and *hw*. Determinative in words with root *hj3*.
- Determinative in *rt* "jaw."
- Phonogram *ns*. Determinative for actions associated with the tongue. Ideogram for *ns* "tongue" and *jmj-r* "overseer" (§ 8.9). Sometimes for Z6.
- Phonograms *sdm* and *jdn*. Determinative EAR and associated actions. Ideogram for *msdr* "ear" and *drd* "leaf"
- Phonogram *ph*. Determinative END, BOTTOM. Ideogram for *phwj* "end" and *kf3* "discreet" (from *kf3* "bottom").
- Variant  (F24). Determinative in *hps* "strong arm; foreleg," also ideogram for same. Determinative in *mshijw* "Foreleg" (Ursa Major).
- Phonogram *whm*. Ideogram for *whm/whmt* "hoof"
- Phonogram *hn*. Ideogram for *hnt* "hide, skin."
- Determinative HIDE, MAMMAL. Sometimes for N2.
- Phonogram *s3b* in *s3b* "dappled." Sometimes for U23.
- Determinative of *stj* "shoot," also ideogram for same. Phonogram *st*.
- Phonogram *3d*.
- Phonogram *ms*.
- Phonogram *h*. Ideogram in *ht* "belly, body."
- Determinative in *sd* "tail," also ideogram for same.
- Ideogram for *jb* "heart." Determinative in *h3tj* "heart."
- Phonogram *nfr*.
- Phonogram *zm3*.
- Variants  (F38),  (F37b). Determinative BACK. Ideogram for *j3t* "back." Sometimes for M21.
- Determinative in *jm3h* "honor" (Essay 21), also ideogram for same. Determinative in *jm3h* "spinal cord," also ideogram for same. Occasionally for F37 as determinative.
- Phonogram *3w*.
- Variant of Y10. Determinative in *psd* "back."
- Phonogram *spr*. Determinative in *spr* "rib," also ideogram for same.
- Determinative in *sphr* "ribs."
- Determinative in *jw<sup>c</sup>* "inherit" and related words, also ideogram for same. Phonogram *jsw*. Determinative in *jw<sup>c</sup>* "femur," *swt* "tibia."
- Determinative in *jdt* "vulva, cow," also ideogram for same.
- Variants  (F47),  (F48),  (F49). Determinative MIDST, TURN, INTESTINE. Determinative in *wdb* "shore" (from *wdb* "turn").
- Phonogram *sphr*.

51		piece of meat	Also  ,  ,  . Determinative FLESH. Ideogram for <i>kns</i> "vagina" and (tripled) <i>h<sup>c</sup>w</i> "body." Phonogram <i>js</i> in <i>jst</i> "Isis" and <i>ws</i> in <i>wsjr</i> "Osiris" in some Coffin Texts.
52		excrement	Determinative in <i>hs</i> "excrement."
<b>G. Birds</b>			
1		Egyptian vulture	Phonogram 3. Often distinguishable from G4 only by flatter head.
2		two vultures	Phonogram 33.
3		U1 + G1	Phonogram m3.
4		buzzard	Variant  (G4a). Phonogram <i>tjw</i> . G4 often distinguishable from G1 only by rounder head.
5		falcon	Ideogram for <i>hrw</i> "Horus."
6		falcon with flail	Determinative in <i>bjk</i> "falcon."
7		falcon on standard	Determinative DIVINE. Also variant of A1 when the speaker is a god or the king.
R13		falcon on standard	Ideogram for <i>jmnt</i> "West" (older form of R14).
7b		falcon in boat	Variant  (G7a). Ideogram for <i>nmfj</i> "Nemti" (a god).
8		G5 + S12	Ideogram for <i>bjk nbw</i> "Gold Falcon" (title of the king: Essay 6).
9		falcon with sundisk	Ideogram in <i>r<sup>c</sup>-hrw-(3htj)</i> "Re-Harakhti" (Essays 4, 12, 16).
10		falcon in Sokar bark	Determinative in <i>zkr</i> "Sokar" (a god) and <i>hnw</i> "Sokar-bark."
11		falcon image	Variant  (G12). Determinative in <i>h<sup>m</sup>/h<sup>m</sup>/h<sup>m</sup></i> "idol" and <i>snbt</i> "breast."
13		falcon image with plumes	Determinative in <i>spdw</i> "Sopdu" (a god). Ideogram for <i>hrw nhnj</i> "Horus of Hierakonpolis."
14		vulture	Phonogram <i>mjwt/mjt/mwt/mt</i> , most common in <i>mwt (mjwt)</i> "mother." Determinative in <i>nrt</i> "vulture" and words with root <i>nr</i> .
14a		vulture on basket	Determinative in <i>nhbt</i> "Nekhet" (goddess).
15		vulture with flail	Determinative in <i>mwt (mjwt)</i> "Mut" (goddess), also ideogram for same.
16		G14a + I13	Ideogram for <i>nbty</i> "Two Ladies" (title of the king: Essay 6).
17		owl	Phonogram <i>m</i> .
18		two owls	Phonogram <i>mm</i> .
20		G17 + D36	Variant  (G19 = G17 + D37). Phonogram <i>mj</i> , <i>m</i> .
21		guinea-fowl	Phonogram <i>nh</i> . Ideogram for <i>nh</i> "guinea-fowl." Often with body like G1 or G43, but with "horns" and lappet of G21.
22		hoopoe	Phonogram <i>db/db</i> in <i>dbt/dbt</i> "brick."
23		lapwing	Variant  (G24). Determinative in <i>rhwt/rhyt</i> "subjects," also ideogram for same.
25		crested ibis	Phonogram 3h.

- 26  ibis on standard  
Variant  (G26a). Ideogram for *dhwtj* “Thoth.” Determinative in *hhj* “ibis.”
- 27  flamingo  
Phonogram *dšr* “red.” Determinative in *dšr* “flamingo.”
- 28  black ibis  
Phonogram *gm*.
- 29  jabiru  
Phonogram *b3*.
- 30  three jabirus  
Ideogram for *b3w* “impressiveness.”
- 31  heron  
Determinative HERON.
- 32  heron on a perch  
Determinative in *b<sup>c</sup>hj* “inundate,” also ideogram for same.
- 33  egret  
Determinative in *sd3/sd3d3* “tremble.”
- 34  ostrich  
Determinative in *njw* “ostrich.”
- 35  cormorant  
Phonogram *q*.
- 36  forktailed swallow  
Phonogram *ur*. Determinative in *mmt* “swallow.”
- 37  sparrow  
Determinative SMALL, BAD. Distinguished from G36 by the rounded tail.
- 38  goose  
Phonogram *gb* in *gbb*, *gbw* “Geb.” Determinative BIRD, INSECT. Variant of G39 as phonogram *z3*. Determinative in *uf3* “discuss,” *wzf* “idle,” *wdfj* “delay,” *htm* “perish, destroy.”
- 39  pintail duck  
Phonogram *z3*. Determinative in *zrt/zrt/zjt* “pintail duck.” Often distinguishable from G38 only by more pointed tail.
- 40  pintail duck flying  
Phonogram *p3*. Occasional variant of G41.
- 41  pintail duck landing  
Phonogram *p3*, especially in hieratic. Determinative in *hnj* “land, alight” and other words with *hn*. Determinative in *shwj* “gather” and *qmyt* “gum.” In combination with T14, determinative in *qm3* “throw,” *qm3j* “create,” and words with *tn*/*tn*.
- 42  fattened bird  
Determinative in *u33* “fatten,” also ideogram in same. Determinative in *df3w* “food.”
- 43  quail chick  
Phonogram *w*. Ideogram for *w* “chick.”
- 44  two quail chicks  
Phonogram *uw*.
- 45  G43 + D36  
Phonogram *w<sup>c</sup>*.
- 46  G43 + U1  
Phonogram *m3w*.
- 47  duckling  
Phonogram *t3*. Ideogram *t3* “duckling.”
- 48  ducklings in nest  
Variants  (G48a),  (G49). Determinative in *z3* “nest,” also ideogram for same.
- 50  two plovers  
Ideogram for *rhjtj* “washerman.”
- 51  bird and fish  
Determinative in *h3m/hjm* “catch fish.”
- 52  bird picking up grain  
Determinative in *snm* “feed.”
- 53  human-headed bird  
Ideogram for *b3* “ba.”
- 54  plucked bird  
Phonogram *snd/snd*. Determinative in *w3n* “wring the neck of birds.”

## H. Parts of Birds

- 1  head of duck
- 2  head of a crested bird
- 3  head of spoonbill
- 4  head of vulture (G14)
- 5  wing
- 6  feather
- 7  claw
- 8  egg



Ideogram for *3pd* “bird” (in offering formulas). Determinative in *wšn* “wing the neck of birds.” Variant of H2.

Determinative in *m3c* “temple (of the head),” occasionally also *m3c* “correct, true, real.” Phonograms *p3q* (variant of H3), *wšm*.

Phonogram *p3q*.

For G14 as determinative in *nrt* “vulture” and words with root *nr*. Ideogram for *nm* “people.”

Determinative WING and associated actions.
















Variants  (H6a),  (H6b). Phonogram *šw*. Ideogram for *šwt* “feather.” Determinative in *m3ct* “Maat” (Essay 10), also ideogram for same.

Phonogram *š3* in *š3t* “Shat” (a place). Determinative in *j3ft* “claw.”

Ideogram for *z3* “son” in proper names. Determinative in *swht* “egg.”

Determinative in *pct* “the elite.”


## I. Reptiles, Amphibians, and their Parts

- 1  gecko
- 2  turtle
- 3  crocodile
- 4  crocodile on shrine
- 5  crocodile with curved tail
- 6  crocodile scales
- 7  frog
- 8  tadpole
- 9  horned viper
- 10  cobra
- R61  emblematic cobra
- 11  two cobras
- 12  erect cobra
- 13  cobra on basket
- 14  snake

Phonogram *ʕ3*. Determinative LIZARD.

Determinative in *štyw* “turtle,” also ideogram for same.

Determinative CROCODILE, AGGRESSION. When doubled, ideogram for *jty* “sovereign.”

Variant  (I5a, crocodile image). Determinative in *sbkw* “Sobek,” also ideogram for same.

Determinative in *s3q* “collect,” also ideogram for same.

Phonogram *km*.

Determinative FROG. Ideogram for *wšm* “*nḥ* “repeating life” (epithet of deceased).


Ideogram for *hfm* “100,000” (§ 9.2). Determinative TADPOLE.

Phonogram *f*. Determinative in *jty* “father.”

Phonogram *d*.

Determinative in *tnjw* “desert border,” also ideogram for same.




Phonogram *dd*.

Variant  (I64). Determinative in *ʕrt* “uraeus” and names of goddesses.

Determinative in *w3dt* “Wadjet” (a goddess) and name of goddesses.


Variant  (I15). Determinative SNAKE, WORM.





## K. Fish and Parts of Fish

- 1  bulti
- 2  barbel
- 3  mullet

Phonogram *jn*. Determinative in *jnt* “bulti.”


Determinative in *bwt* “abomination.”

Phonogram *ʕd* in  “*d-mr* “district administrator.” Determinative in *ʕdw* “mullet.”

- 4  oxyrhynchus  
 5  pike  
 6  fish scale  
 7  blowfish








Phonogram *h3*. Ideogram in *h3t* "oxyrhynchus."

Determinative in *bzj* "introduce." Determinative FISH, FISHY.

Variant . Determinative in *nšmt* "fish scale," also ideogram for same.

Determinative in *špt* "angry."

## L. Insects and Invertebrates

- 1  scarab beetle  
 2  bee or wasp  
 3  fly  
 4  locust  
 5  centipede  
 6  shell  
 7  emblematic scorpion

Phonogram *hpr*. Determinative in *hpr* "scarab beetle," also ideogram for same.


Ideogram for *bjt* "bee; honey," and *bjj* "King of Lower Egypt."

Determinative in *ʿff* "fly."














Determinative in *znḥm* "locust"


Ideogram in *sp3* "Sepa" (place near Heliopolis). Determinative in *zp3* "centipede."

Phonogram *h3* in *h3wt* "offering table."

Variant  (L7a). Determinative in *sqṭ* "Selket" (a goddess), also ideogram for same.

## M. Vegetation

- 1  tree  
 2  plant  
 3  stick  
 4  rib of palm branch  
 5  M4 + X1  
 6  M4 + D21  
 7  M4 + Q3  
 8  pool with lilies  
 9  lily (lotus)  
 10  lily (lotus) bud  
 11  flower on stem  
 12  lily (lotus) plant  
 13  papyrus

Variant  (M1a, with M3). Determinative TREE; also in *mʿr* "fortunate." Phonogram *jm3*, often with only G17 *m* as complement = *jm(3)*.

Determinative PLANT. Phonogram *ḥn*. Determinative in *jzj* "light," *jz* "tomb," *js* "old" (from *jzw* "reeds"). Rarely for AT as determinative or in IS pronouns (from *j* "reed"). Occasional variant of T24.

Phonogram *ḥt*. Determinative WOOD. Ideogram for *ḥt* "wood, stick, tree, mast." Also vertically as determinative of *dʿr* "seek."

Ideogram for *rnpt* "year" and *ḥsbt* "regal year" (§ 9.9). Determinative in *rnptj* "young." Determinative TIME in *tr* "time, season." When doubled, ideogram for *snf* "last year."

Determinative TIME in *tr* "time, season," also ideogram for same. Variant of M6.

Determinative TIME in *tr* "time, season," also ideogram for same. Determinative of some roots ending in *tr* and *ṛj*.

Determinative in *rnptj* "young," also ideogram for same.


Phonogram *ḫ3*. Ideogram for *ḫt* "Inundation (season)" (§ 9.8). Ideogram for *ḫ3* "pool, marsh."






































Determinative in *zšḥj* "lily (lotus)," also ideogram for same.

Determinative in *nḥbt* "lily (lotus) bud."

























Determinative in *wḏn* "dedicate, offer," also ideogram for same. Occasional variant of F46 as determinative in *wḏb* "shore."

Phonogram *ḫ3*. Ideogram for *ḫ3* "1,000" (§ 9.1) and "lily (lotus)."



Variant  (M14, with I10). Phonogram *w3ḏ/w3ḏ*, also *wḏ/wḏ*. Ideogram for *w3ḏ* "papyrus column."

15		clump of papyrus with buds	Determinative for <i>mḥw</i> "Delta," also ideogram for same. Determinative PAPYRUS, SWAMP. Phonogram <i>ḥ</i> in <i>ḥt</i> - <i>ḥt</i> "Chemmis" (Delta town).
16		clump of papyrus	Phonogram <i>ḥ</i> 3. Variant of M15 in <i>mḥw</i> "Delta."
17		reed	Phonogram <i>j</i> . When doubled, phonogram <i>y</i> . Occasional variant of A1. Ideogram for <i>j</i> "reed."
18		M17 + D54	Variant  . Phonogram <i>j</i> in forms of <i>jj</i> "come."
19		emblem for offerings	Determinative in <i>ʿḥb</i> "offer," also ideogram for same.
20		field of reeds	Determinative in <i>ḥt</i> "field" and <i>ḥtj</i> "peasant," also ideogram for same. Occasional variant of M21.
21		reeds with root	Determinative in <i>sm</i> "grass" and <i>sm</i> "help."
22		rush	Phonogram <i>nḥb</i> . When doubled, phonogram <i>nn</i> .
23		sedge	Phonogram <i>sw</i> . Ideogram for <i>nswt</i> "king." Ideogram for <i>swt</i> "sedge." Occasional variant of M24 and M26.
163		M23 + Aa1	Ideogram for <i>rh-nswt</i> "king's acquaintance."
24		M23 + D21	Variant  (M25). Ideogram for <i>rsu</i> "south."
26		flowering sedge	Variant  (M27, with D36). Phonogram <i>šmʿ</i> . Ideogram for <i>šmʿw</i> "Nile Valley" (Upper Egypt).
28		M26 + V20	Ideogram in title <i>ur mḏw-šmʿw</i> "chief of the tens of the Nile Valley."
29		pod	Phonogram <i>nḏm</i> "pleasant."
30		root	Determinative in <i>bnr</i> "sweet," also ideogram for same.
31		rhizome	Variant  (M32). Determinative in <i>rd</i> "grow," also in <i>rwḏ</i> "firm."
33		grain	Variants  ,  . Ideogram for <i>jtj</i> "grain." Determinative GRAIN.
34		sheaf of emmer	Ideogram for <i>btj</i> (originally <i>bdt</i> ) "emmer," also determinative for same.
35		heap of grain	Determinative HEAP.
36		bundle of flax	Variant  (M37). Phonogram <i>ḏr</i> . Determinative in <i>dm3</i> "bundle."
38		bundle of flax	Determinative in <i>mḥʿw</i> "flax" and <i>dm3</i> "bundle."
Y10		bundle of stems	Determinative in <i>šʿt</i> "murderousness" (from <i>šʿ</i> "cut").
39		basket of fruit or grain	Determinative VEGETABLES.
40		bundle of reeds	Phonogram <i>jz</i> .
41		piece of wood	Determinative WOOD.
42		rosette	Phonogram <i>wn</i> . In hieratic indistinguishable from Z11.
43		grapes on trellis	Variant  (M43a). Determinative VINE, WINE, GARDENER, FRUIT. Ideogram for <i>jrp</i> "wine" and <i>k3ny</i> "gardener."
43b		wine or olive press	Determinative in <i>šzmuw</i> "Shesmu" (god of the wine or olive press), also ideogram for same.
44		thorn	Determinative in <i>spd</i> "sharp," also ideogram for same. Determinative in <i>srt</i> "thorn." Determinative in <i>t-ḥḏ</i> "white-bread" (as bread of this form).

## N. Sky, Earth, Water


- 1  sky
- 2  sky with scepter
- 4  sky with rain
- 5  sun
- 5a  sun with two strokes
- 6  sun with uraeus
- 7  N5 + T28
- 8  sun with rays
- 9  moon
- 11  crescent moon
- 12  crescent moon
- 64  N11 + N14
- 13  half N11 + N14
- 14  star
- 15  star in circle
- 16  strip of land with sand
- 18  strip of sand
- 19  two strips of sand
- 20  tongue of land
- 21  tongue of land
- 23  irrigation canal
- 24  irrigation canal system
- 25  mountain range
- 76  N25 on standard

Determinative SKY, ABOVE. Ideogram for *hrj* "upper" (§ 8.6.7). Determinative in *rw* "gate" and *h3yt* "ceiling, portal," also ideogram for latter.

Variants  (N3, with oar),  (N46b, with star). Determinative NIGHT. Ideogram for *grh* "night."

Determinative DEW, RAIN. Ideogram for *j3dt* "dew."


Determinative SUN, DAY, TIME. Ideogram for *r<sup>c</sup>* "sun, Re," *hrw* "day," and *sw* "day" (in dates: § 9.8).


Variant  (N5 + N23). Determinative TIME.


Determinative in *r<sup>c</sup>* "Re," also ideogram for same.

Ideogram for *hrt-hrw* "daytime, course of the day."

Determinative SUNLIGHT. Phonogram *wbn* (from *wbn* "rise"). Ideogram for *hnmmt* "human beings."

Variant  (N10). Phonogram *psd* in *psdt* "Ennead" and *psdntjw* "new-moon festival." Variant of X6 in *p3t* "origin."

Variant  as determinative. Determinative in *j<sup>c</sup>h* "moon," also ideogram for same. Ideogram for "month" (*3bd*) in dates (§ 9.8). Occasional variant of F42. Determinative in *w<sup>c</sup>h* "carob bean," also ideogram for same. Determinative in *šzp* "palm" (measure: § 9.7.1), also ideogram for same.



Variant  as determinative. Determinative in *j<sup>c</sup>h* "moon," also ideogram for same. Occasional variant of F42.

Ideogram for *3bd* "month."

Ideogram for *mddjunt* "15th-day festival."


Determinative STAR, TIME. Phonogram *sb3* (from *sb3* "star"). Phonogram *dw3* (from *dw3* "morning"). Ideogram for *wnwt* "hour."

Ideogram for *dw3t* "Duat" (Essay 2).



Variants  (N16d),  (N17). Ideogram for *t3* "land, earth, world." Phonogram *t3*. Determinative in *dt* "estate" and *dt* "eternity."

Ideogram for *jw* "island." Determinative DESERT, FOREIGN LAND. Ideogram for *st3t* "aroura" (§ 9.7.2).

Ideogram for *3ht* "Akhet" (Essay 2) in *hrw-3htj* "Harakhti" (Essay 12).

Variant  (N22). Phonogram *wqb/wdb* in *wqb* "turn." Determinative LAND, especially in *wqb* "shore." Determinative in *h3b-sd* "Sed Festival."

Determinative LAND. Ideogram for *jdb* "bank," when doubled *jdbwj* "Two Banks" (a term for Egypt).

Variants  . Determinative LAND, especially IRRIGATED LAND. Also used in variant of N5a. Ideogram for *gbb/gbw* "Geb."

Determinative of *sp3t* "nome," also ideogram for same. Determinative in names of nomes and divisions of Egypt, also in *h3p* "garden." Ideogram for *dt* "estate, farm."

Ideogram for *h3st* "desert cliffs, foreign land." Determinative DESERT, FOREIGN LAND.

Ideogram for *h3* "Ha" (desert god).










26		mountain	Phonogram <i>ḏw</i> . Ideogram for <i>ḏw</i> “mountain.”
27		sun rising above mountain	Ideogram for <i>3ḥt</i> “Akhet” (Essay 2).
28		sun's rays above hill	Phonogram <i>ḥꜥ</i> , especially in <i>ḥꜥj</i> “appear.”
29		sandy slope	Phonogram <i>q</i> .
30		hill with shrubs	Determinative in <i>j3t</i> “mound,” also ideogram for same.
31		path with shrubs	Variant  (N31e). Determinative for <i>w3t</i> “road,” also ideogram for same. Determinative ROAD, DISTANCE, POSITION. Ideogram for <i>w3j</i> “tend, start” (from <i>w3t</i> “road”). Phonogram <i>ḥr</i> in <i>jn-ḥrt</i> “Onuris” (a god), <i>ḥrw</i> “Horus,” and <i>ḥrw r</i> “except” (from <i>ḥrj</i> “go far away”).
32		lump of clay	Variant of Aa2 and F52.
33		grain of sand	Variants  (N33a),  ,  ,  . Determinative SAND, MINERAL, PELLET. When single, occasional substitute for signs with bad connotations, such as A14 and Z6. When triple, occasional substitute for plural strokes. Determinative in words with <i>qd</i> (from <i>qđj</i> “go around”).
34		ingot of metal	Variant  (N34a). Ideogram for <i>ḥmt</i> “copper, bronze.” Determinative COPPER, BRONZE.
35		ripple of water	Phonogram <i>n</i> .
35a		three ripples of water	Ideogram for <i>mw</i> “water.” Determinative WATER. Phonogram <i>mw</i> .
36		canal	Determinative BODY OF WATER. Phonogram <i>mr</i> and <i>mj</i> . Ideogram for <i>mr</i> “canal.”
37		basin	Variants  (N37a),  (N38),  (N39), etc. Phonogram <i>š</i> . Ideogram for <i>šj</i> “basin, pool, lake.” Determinative of <i>st3t</i> “aroura” (§ 9.7.2), also ideogram for same. Variant of X4 as determinative of <i>zn</i> “open” and <i>znj</i> “pass.” Variant of O36.
40		N37 + D54	Phonogram <i>šm</i> in forms of <i>šmj</i> “go.”
41		well with water	Variants  (N42),  (D280a). Determinative WELL. Determinative in <i>bj3</i> “cauldron, copper” and words with root <i>bj3</i> . Determinative in <i>phwv</i> “outer limits,” also ideogram for same (tripled). Often for D280a.

## O. Structures and Parts of Structures






















1		schematic house plan	Proportions vary. Ideogram for <i>pr</i> “house.” Phonogram <i>pr</i> . Determinative BUILDING, PLACE.
2		O1 + T3	Ideogram for <i>pr-ḥd</i> “treasury.”
3		O1 + P8 + X3 + W22	Ideogram for <i>pṛt-ḥrw</i> “invocation offering.”
4		reed shelter	Phonogram <i>h</i> . Ideogram for <i>h(?)</i> “courtyard.”
5		winding wall	Phonogram <i>nm</i> . Determinative in <i>mrrt</i> “street.” Phonogram <i>mr</i> in <i>mr-wr</i> “Mnevis” (sacred bull of Heliopolis).
6		plan of enclosure	Variant  (O7). Ideogram for <i>ḥwt</i> “enclosure.”
8		O7 + O29	Ideogram for <i>ḥwt-ꜥ3t</i> “Great Enclosure” (temple of Heliopolis).
9		V30 + O6	Ideogram for <i>nbt-ḥwt</i> “Nephthys.”
10		O6 + G5	Ideogram for <i>ḥwt-ḥrw</i> “Hathor.”
11		palace plan with battlements	Variant  (O12). Ideogram for <i>ḥ</i> “palace.”
104		O11 + T3	Ideogram for <i>ḥ-ḥd</i> “White Palace” or “Palace of the Macc” (a shrine).










- 13  enclosure with battlements Variant  (O14). Determinative in *sbḥ* “wall in” and related words.
- 15  enclosure + W10 + X1 Variant  (O15a). Ideogram for *wṣḥt* “broad hall.”
- 16  cornice with cobras Variant  (O17). Determinative for *t3yt* “curtain,” also ideogram for same and *t3ḥtj* “he of the curtain” (title of the vizier). O17 variant of S22 in *t3-wr* “port”
- 18  shrine in profile Determinative in *k3r* “shrine,” also ideogram for same.
- 19  shrine with poles Determinative in *pr-wr* “Great House” (original shrine of Upper Egypt at Hierakonpolis), also in *jtrt ḥmꜥt* “Nile Valley Shrine” (same).
- 20  shrine Determinative SHRINE.
- 21  shrine façade Determinative in *zḥ* “booth,” also ideogram for same.
- 22  booth with pole Determinative in *zḥ* “counsel, advice” and *zḥ* “tent, booth,” also ideogram for latter.
- 23  double platform Determinative in *ḥ3b-sd* “Sed Festival,” also ideogram for same.
- 24  pyramid and enclosure wall Determinative PYRAMID.
- 25  obelisk Determinative in *tḥn* “obelisk,” also ideogram for same.
- 26  stela Determinative STELA, also ideogram for *wḏ* “stela.”
- 27  columned hall Determinative HALL. Determinative of *ḥ3wḥ* “dusk” (from *ḥ3* “office”), also ideogram for same.
- 28  column with tenon Phonogram *j(w)n*. Ideogram for *jwn* “column.”
- 29  wood column Variant . Phonogram ʿ3.
- 30  support Determinative SUPPORT, also ideogram for *zḥnt* “support.”
- 31  door leaf Variant  (O31a). Variant  in ʿ3wḥ “door” (two door leaves). Determinative OPEN. Determinative in ʿ3 “door,” also ideogram for same.
- 32  gateway Determinative DOORWAY, also ideogram for *sb3* “doorway.”
- 33  palace façade Determinative in *srḥ* “serekh” (Essay 6).
- 34  doorbolt Phonogram *z*. Ideogram for *z* “doorbolt” Variant of R22.
- 35  O34 + D54 Phonogram *z* in *zj* “go away, perish,” *zy* “which?” (§ 5.11), *zḥj* “send away, go away,” and *mz* “bring.”
- 36  wall Determinative WALL. Ideogram for *jnb* “wall.”
- 37  wall falling Determinative TOPPLE, TILT.
- 38  corner Determinative CORNER. Ideogram for *qnbt* “council.” Determinative or ideogram for *tm* in the administrative title *ḥrj (n) tm* “chief of the *tm*.”
- 39  stone block or brick Determinative STONE, BRICK.
- 40  stairs Determinative STAIRWAY, TERRACE. Ideogram for *rwd* “stairs” and *ḥtjw* “terrace.”
- 41  double stairs Determinative STAIRWAY, ASCEND.
- 43  fence Variant  (O42). Phonogram *ḥzp*, *sḥp*.
- 44  emblem of Min Variant  (O44a) Determinative in *j3t* “office,” also ideogram for same.

- 45  domed structure Variant  (O46). Determinative in *jp3t* “private quarters,” also ideogram for same.
- 47  enclosed mound Variant  (O48). Ideogram for *nhn* “Hierakonpolis” and *mhnt* “jasper.”
- 49  area with intersection Variant  (O49a). Ideogram for *nwt* “town.” Determinative TOWN, SETTLEMENT.
- 50  threshing floor with grain Phonogram *zp* in *zp* “occasion, event,” *zpj* “be left over,” and related words. Determinative in *zpt* “threshing floor.”
- 51  pile of grain Variant  (O51b). Determinative in *snwt* “granary,” also ideogram for same.







































## P. Ships and Parts of Ships

- 1  boat on water Variants  (P1c),  (P26). Determinative BOAT. Ideogram for “boat” (various readings: *dpt*, *h<sup>c</sup>w*, *jmw*, *q3q3w*).
- 1a  boat capsized Determinative in *pn<sup>c</sup>* “capsize.”
- 2  boat under sail Determinative in *hntj* “sail upstream.”
- 3  sacred boat Variants  (P30),  (P34). Determinative GOD’S BOAT. Ideogram for *wj3* “sacred bark.”
- 3a  P3 + M23 Ideogram for *wj3-nswt* “king’s bark.”
- 4  boat with net Variant  (P4a). Phonogram *w<sup>c</sup>h<sup>c</sup>*.
- 5  mast with sail Determinative WIND, AIR. Ideogram for *t3w* “air” and *nfw* “sailor.”
- 5f  sail Determinative in *ht3w* “sail.”
- 6  mast Variant  (P7). Phonogram *h<sup>c</sup>*.
- 8  oar Variant  in  *m3<sup>c</sup> hrw* “justified” (Essay 8). Phonogram *hrw*. Determinative OAR. Ideogram for *hjp<sup>t</sup>* “oar.”
- 9  P8 + I9 Ideogram for *hr,ff* “says, said” (§ 22.18).
- 10  steering oar with rope Determinative in *hmw* “rudder” and *hmy* “steerer.”
- 11  mooring stake Determinative in *mjnj* “moor, die” and related words. In hieratic often identical with T14.








## Q. Domestic and Funerary Furniture





































- 1  seat Ideogram for *st* “seat, place.” Phonogram *st*. Phonogram *ws* in *wsjr* “Osiris.” Phonogram *htm*.
- 2  portable seat Phonogram *ws* in *wsjr* “Osiris.” Ideogram for *st* “seat.”
- 3  stool Phonogram *p*.
- 4  headrest Determinative in *wsu* “headrest.”
- 5  chest Determinative CHEST, BOX.
- 6  coffin Determinative in *qrs* “bury” and related words, also ideogram for same.
- 7  brazier with flame Determinative FIRE. Ideogram for *sff* “temperature.” When doubled, ideogram for *nsrsr* “flame” in *ju-nsrsr* “Island of Flame” (locality of creation and in the Duat).














## R. Temple Furniture and Sacred Emblems

- 1  table with offerings  
Variants  (R2),  (R36a). Determinative in *h3wt/h3yt* "altar," also ideogram for same.
- 3  low table with offerings  
Determinative in *wḏḥw* "offering table," also ideogram for same.
- 4  bread loaf on mat  
Phonogram *hṭp*. Ideogram for *hṭp* "offering slab."
- 5  censer  
Variant  (R6). Phonogram *k(3)p*. Determinative in *k3p* "fumigate," also ideogram for same.
- 7  bowl with smoke  
Determinative of *snṯr* "incense," also ideogram for same. Variant of W10a/Aa4.
- 8  cloth wound on pole  
Ideogram for *nṯr* "god." Phonogram *nṯr*. Determinative GOD.
- 9  R8 + V33  
Determinative for *bd* "incense," also ideogram for same.
- 10  R8 + T28 + N29  
Variants  (R10e),  (R50). Ideogram for *hṯj-nṯr/hṯt-nṯr* "necropolis."
- 11  reed column  
Phonogram *ḏd*, also doubled with the same value. Ideogram for *ḏd* "djed-column/amulet"
- 12  carrying standard  
Determinative in *j3t* "standard." Usually part of other signs.
- 14  feather on standard  
Variant  (R13). Ideogram for *jmnt* "West" and *wnmj* "right"
- 15  spear emblem  
Variant  (R15b). Ideogram for *j3b* "East, left." Variant of U23.
- 16  scepter with feathers  
Determinative in *wḥ* (emblem of Qus), also ideogram for same.
- 17  wig with feathers on pole  
Variants  (R17b),  (R18). Determinative in *t3-wr* "This" (nome of Abydos), also ideogram for same.
- 19  S40 with feather  
Ideogram for *w3st* "Thebes" (town and nome).
- 20  Seshat emblem  
Variant  (R21). Ideogram for *s3st* "Seshat" (a goddess).
- 22  Min emblem  
Variants  (R22a),  (R23),  (R23a). Ideogram for *mnw* "Min" (a god). Without standard, phonogram *hm* in *hm* "shrine" and *hm* "Letopolis" (town in the Delta).
- 24  Neith emblem  
Variants  (R24a),  (R24b),  (R24c),  (R25),  (R25a),  (R25b). Determinative in *nṯt* (originally *nrt*) "Neith," also ideogram for same.
- 61  emblematic cobra  
Determinative in *tnjw* "desert border," also ideogram for same.

















## S. Regalia and Clothing
























- 1  White Crown  
Variant  (S2). Determinative WHITE CROWN. Ideogram for *hḏt* "White Crown."
- 47a  S1 on standard with flail  
Determinative in *b3b3y* "Babay" (a god), also ideogram for same.
- 3  Red Crown  
Variant  (S4). Determinative RED CROWN. Phonogram *n*. S3 variant of L2 as emblem of King of Lower Egypt.
- 5  Double Crown  
Variant  (S6). Determinative in *shmtj* "Double Crown," also ideogram for same. Determinative CROWN.

7		Blue Crown	Determinative in <i>ḥprš</i> “Blue (War) Crown,” also ideogram for same.
8		Atef Crown	Determinative in <i>štf</i> “Atef Crown,” also ideogram for same.
9		double plumes	Determinative in <i>šwtj</i> “double plumes,” also ideogram for same.
10		headband	Phonogram <i>mdḥ</i> . Determinative in <i>wšḥw</i> “wreath” and <i>mdḥ</i> “headband,” also ideogram for latter.
11		broad collar	Determinative in <i>wšḥ</i> “broad collar,” also ideogram for same. Phonogram <i>wšḥ</i> .
12		bead collar	Variant  (S12a). Ideogram for <i>nbw</i> “gold” and related words. Determinative PRECIOUS METAL.
13		S12 + D58	Phonogram <i>nb</i> .
14		S12 + T3	Ideogram for <i>ḥḏ</i> “silver.”
14a		S12 + S40	Ideogram for <i>ḏ<sup>c</sup>m</i> “electrum.”
15		faience pectoral	Variants  (S16),  (S17),  (S17a). Determinative in <i>ṯḥn</i> “sparkle” and related words, also ideogram for same. Ideogram for <i>šzmt</i> “malachite” and related words.
18		bead necklace	Determinative in <i>mnjt</i> “bead necklace, counterweight,” also ideogram for same.
19		seal on necklace	Ideogram for <i>ḥtm</i> “seal” and related words.
20		seal on necklace	Determinative SEAL. Ideogram for <i>ḥtm</i> “seal” and <i>š(n)<sup>c</sup>tj</i> “ring” (§ 9.7.3). Variant of E31.
21		ring	Determinative RING.
22		shoulder knot	Phonogram <i>s(3)ḫ</i> . Determinative in <i>t3-wr</i> “port (of ship),” also ideogram for same.
23		knotted cloth	Phonogram <i>dmdḏ/dmd</i> . Different from Aa6.
24		knotted belt	Phonogram <i>ṯ3z</i> . Ideogram for <i>ṯ3zt</i> “knot, vertebra.”
25		garment with ties	Ideogram for <i>j<sup>c</sup>3w</i> “guide, dragoman, interpreter.”
26		kilt	Determinative in <i>šndyt</i> (originally <i>šndwt</i> ) “kilt,” also ideogram for same.
130a		strip of cloth	Determinative in <i>d3jw</i> “cloak,” also ideogram for same.
27		cloth with two fringes	Determinative in <i>mnḥt</i> “cloth,” also ideogram for same.
116		cloth with four fringes	Determinative in <i>jfdj</i> “four-ply linen,” also ideogram for same.
118		cloth with six fringes	Determinative in <i>sjsj</i> “six-weave linen,” also ideogram for same.
28		cloth with fringe + S29	Variant  (V48). Determinative CLOTH.
29		folded cloth	Phonogram <i>s</i> . Abbreviation for <i>snb</i> in  “ <i>nḥ.(w)-(w)ḏ3.(w)-s(nb.w)</i> ” (§ 17.20.2).
30		S29 + I9	Phonogram in <i>šf</i> “yesterday.”
31		S29 + U2	Phonogram <i>sm3</i> .
32		cloth with fringe	Phonogram <i>sj3</i> . Ideogram for <i>sj3t</i> “fringed cloth.”
33		sandal	Determinative SANDAL. Ideogram for <i>ṯbt</i> “sandal,” <i>ṯbw</i> “sandalmaker.”
34		sandal strap	Phonogram “ <i>nḥ</i> .” Ideogram for “ <i>nḥ</i> “sandal strap” and “mirror.”




















- 35  sunshade or fan  
Variant  (S36). Ideogram for *šwt* “shadow, shade.” Determinative in *sryt* “fan,” also ideogram for same. Doubled (S36), ideogram for *hjpwj* “Hepwi” (a god).
- 37  fan  
Determinative in *hw* “fan,” also ideogram for same.
- 38  crook  
Phonogram *hq3*. Determinative in *hq3t* “scepter,” also ideogram for same. Variant of S39.
- 39  shepherd's crook  
Phonogram *wt* in *wt* “flock” (from *wt* “crook”).
- 40  animal-headed staff  
Phonogram *w3s*. Ideogram for *w3s* “staff” of this shape. Ideogram for *j3tt* “milk, cream” and “latet” (milk goddess). Doubled, phonogram *w3b* in *w3bwj* “Wabwi” (name of a nome) and *w3bw* “Wabut” (a town). Variant of S41 and R19.
- 40a  S40 on standard  
Variant of S40 as ideogram for *j3tt* “milk, cream” and “latet.”
- 41  animal-headed staff  
Phonogram *dcm* in *dcmw* “fine gold” (from *dcm* “staff” of this shape).
- 42  scepter  
Phonogram *shym*. Determinative in *hpr* “manage,” also ideogram for same, especially in titles. Phonogram *b3*. Ideogram for *b3* “scepter” and “stela.” Ideogram for *shym* “sistrum.”
- 42a  lotus-bud scepter  
Determinative in *nhbt* “lotus-bud scepter,” also ideogram for same.
- 43  staff  
Phonogram *md*. Ideogram for *mdw* “staff.”
- 44  staff with flail  
Determinative for *3ms* “staff,” also ideogram for same.
- 45  flail  
Determinative in *nh3h3w* “flail,” also ideogram for same.





























## T. Warfare, Hunting, and Slaughter

- 1  mace with flat head  
Phonogram *mn*.
- 2  T3 tilted  
Determinative SMITE.
- 3  mace with round head  
Variant  (T4). Phonogram *hd*. Ideogram for *hd* “mace” of this shape.
- 5  T3 + I10  
Phonogram *hd*.
- 6  T3 + I10 + I10  
Phonogram *hdd*.
- 7  axe  
Determinative AXE and related words.
- 7a  axe  
Determinative in *3qhw* “axe” of this shape.
- 8  dagger  
Phonogram *tp*. Determinative in *mtpt* “dagger” of this shape.
- 8a  dagger  
Determinative in *b3gsu* “dagger” of this shape.
- 9  bow  
Variants  (T9a),  (T10). Phonogram *pd/pd*. Determinative in *pd* “bow,” also ideogram for same and words of the same root.
- 11  arrow  
Phonogram *zum*. Determinative ARROW.
- 12  bowstring  
Phonogram *rud/rud*. Determinative in words with *3r* (*3j*, *3jr*, from *3r* “restrain”). Ideogram for *d3r* “subdue.” Determinative for *rud* “bow-string,” also ideogram for same.
- 13  pieces of wood tied  
Phonogram *rs* in *rs* “wake” and related words.











14		throw-stick	Variant  (T15). Determinative with G4T in words with <i>tn/tn</i> . Determinative FOREIGN. Determinative in <i>qm3</i> "throw" and <i>qm3j</i> "create," also ideogram for same. Ideogram for <i>3m</i> "Asiatic," <i>thnw</i> "Libya." Ideogram for <i>hq3t</i> "heqat" (§ 9.7.4). Variant of D50 as determinative ACCURATE; of M3 as determinative in <i>d3r</i> "seek"; of P11 as determinative in <i>mjnj</i> "moor, die"; of S39 as phonogram in <i>wt</i> "flock"; and of T13 and Aa6.
16		scimitar	Determinative in <i>hps</i> "scimitar."
17		chariot	Determinative in <i>wrrt</i> "chariot," also ideogram for same.
18		crook with package attached	Phonogram <i>šms</i> .
19		bone harpoon head	Variant  (T20). Phonogram <i>qs</i> . Determinative BONE, TUBE. Determinative in <i>qrs</i> "bury," <i>twr</i> "pure" (from <i>twr</i> "tube"). Ideogram for <i>gnwt</i> "annals" and <i>gnwtj</i> "sculptor" (often double in the latter).
21		harpoon	Variant  . Phonogram <i>wc</i> in <i>wc</i> "one" and related words.
22		arrowhead	Variant  (T23). Phonogram <i>sn</i> .
24		fishing net	Phonogram <i>h/jh</i> . Determinative NET.
25		reed float	Phonogram <i>db3/db3</i> .
27		bird trap	Variant  (T26). Determinative in <i>šht</i> "trap," also ideogram for same.
28		butcher's block	Phonogram <i>hr</i> .
29		T30 + T28	Determinative in <i>nmt</i> "slaughtering place," also ideogram for same.
30		knife or saw	Determinative KNIFE, SHARP. Ideogram for <i>dmt</i> "knife."
31		knife sharpener	Variants  (T32),  (T33). Phonogram <i>šm</i> in <i>šm</i> "guide" and related words.
35		butcher knife	Variant  (T34). Phonogram <i>nm</i> . Determinative in <i>nm</i> "butcher knife."





## U. Agriculture, Crafts, and Professions

1		sickle	Variant  (U2). Phonogram <i>m3</i> . Determinative REAP, CROOKED.
3		U1 + D4	Phonogram <i>m3</i> in <i>m3</i> "see."
4		U1 + Aa1	Variant  (U5). Phonogram <i>m3c</i> in <i>m3c</i> "true, correct," and related words.
6		hoe	Variants  (U7),  (U6a),  (U7a). Phonogram <i>mr</i> . Determinative HACK. Variant of U8.
8		hoe	Phonogram <i>hn</i> (from <i>hnn</i> "hoe").
9		grain-measure with grain	Determinative GRAIN. Ideogram for <i>hq3t</i> "heqat" and <i>jpt</i> "oipe" (§ 9.7.4).
10		M33 + U9	Ideogram for <i>jtj</i> "barley, grain." Variant of U9 as determinative.
11		S38 + U9	Variant  (U12). Ideogram for <i>hq3t</i> "heqat" (§ 9.7.4).
109		pitchfork	Variant  (U109a). Determinative (U109) in <i>šdb</i> "obstacle," also ideogram for same. Determinative in <i>bj</i> "collect" and <i>bt</i> "pitchfork."
13		plow	Variant  (U14). Phonogram <i>šn</i> . Phonogram <i>hb</i> . Determinative PLOW. Ideogram for <i>prt</i> "seed."
15		sled	Phonogram <i>tm</i> .

- 16  loaded sled with jackal's head Determinative in *bj3* "wonder" and related words, also ideogram for same. Determinative SLED.
- 17  pick and basin Variant  (U18). Phonogram *gry*.
- 19  adze Variant  (U20). Phonogram *nuw*.
- 21  adze and block of wood Phonogram *stp/stp*.
- 22  chisel Determinative in *mnh* "functional." Determinative CARVE.
- 23  chisel Phonograms *3b* and *mr*.
- 25  drill for stone Variant  (U24). Ideogram for *hmwt* "craft" and related words.
- 26  drill for beads Variant  (U27). Ideogram for *wb3* and related words. Occasional variant of U24-25.
- 29  fire-drill Variant  (U28). Phonogram *d3*. Abbreviation for *wd3* in  *nh.(w)-(w)d3.(w)-s(nb.w)* (§ 17.20.2).
- 30  kiln Phonogram *t3*.
- 31  baker's rake Determinative in *hnr* "restrain" and related words, also ideogram for same. Determinative in *rlh/jlh* "restrain." Determinative in *rlhtj* "baker," also ideogram for same. Variant of D19-20.
- 32  pestle and mortar Determinative in *smn* "set, fix" (from *smn* "flatten dough"). Determinative POUND, HEAVY. Determinative in *hzmh* "natron; bronze," also ideogram for same.
- 33  pestle Phonogram *tj/t*.
- 34  spindle Variant  (U35). Phonogram *hsf*. Determinative in *hsf* "spin."
- 36  launderer's club Phonogram *hm*.
- 37  razor Determinative in *h3q* "shave."
- 38  scale Determinative in *mht* "scale," also ideogram for same.
- 39  upright of scale Variants  (U40),  (U40a). Determinative in *wtz* "hold up, carry, wear" and *tzj* "pick up."
- 41  plumb bob Determinative in *th* "plumb bob."

## V. Rope, Baskets, and Cloth






























- 1  coil of rope Variant  (V1a). Determinative ROPE, TIE, COIL. Ideogram for *3t* "100" (§ 9.1). Phonogram *3n* in *3nt* "dispute." Different from Z7.
- 2  VI + O34 Determinative in *st3* "pull" and *33* "hasten." Ideogram for *st3t* "aroura" (§ 9.7.2).
- 3  three VI + O34 Ideogram *st3w* in *r-st3w* "necropolis" (of Giza).
- 4  lasso Phonogram *w3*.
- 5  looped rope Determinative in *sntj* "lay out," also ideogram for same.
- 6  cord with ends up Phonogram *3s* and *3sr*. Ideogram for *3srw/3s* "linen." Variant of V33.
- 7  cord with ends down Variant  (V8). Phonogram *3n*.
- 9  round cartouche Determinative in *3nw* "circuit" (of the sun), also ideogram for same. Determinative in *3nw* "cartouche."

10		cartouche	Surrounding names of kings, queens, and some gods. Determinative in <i>šnw</i> “cartouche” and <i>rn</i> “name.”
11		end of cartouche	Determinative in <i>dhj</i> “dam” and <i>ph3</i> “split.” Ideogram for <i>ph3</i> , a kind of grain. Ideogram for <i>djw</i> / <i>dyt</i> “shriek.”
12		string	Determinative in <i>šh</i> “loosen,” <i>ʿrq</i> “bind,” <i>šfdw</i> “papyrus scroll,” and other words associated with STRING. Determinative in <i>ʿrq</i> “swear” and <i>ʿrqy</i> “last day of the month” (§ 9.8) (from <i>ʿrq</i> “bind”), also ideogram for latter. Ideogram for <i>šh</i> “loosen.” Determinative in <i>fnhw</i> “Phoenicians.”
13		hobble	Variant  (V14). Phonogram <i>t</i> / <i>t</i> .
15		V13 + D54	Phonogram <i>jt</i> in forms of <i>jtj</i> “take possession.”
16		hobble for cattle	Variants  (V16a),  (V17, rolled-up tent),  (V18). Phonogram <i>z3</i> in <i>z3</i> “protection” and related words.
19		hobble for cattle	Determinative SHRINE in <i>k3r</i> “shrine,” <i>qnj</i> “palanquin” (also <i>qnj</i> “sheaf”), <i>štyt</i> “Sokar shrine.” Determinative in <i>tm3</i> “mat” and <i>tm3</i> “cadaster,” also ideogram for latter. Determinative in <i>h3r</i> “sack” (§ 9.7.4), also ideogram for latter. Determinative in <i>mdt</i> “stable, stall,” also ideogram for latter.
20		V19 without horizontal	Ideogram for <i>mdw</i> “10” (§ 9.7.1).
21		V20 + I10	Phonogram <i>md</i> .
23		whip	Variant  (V22). Phonogram <i>mh</i> .
24		cord wound on stick	Variant  (V25). Phonogram <i>wd</i> / <i>ud</i> .
26		spool with thread	Variant  (V25, without thread). Phonogram <i>ʿd</i> / <i>ʿd</i> . Determinative in <i>ʿd</i> “reel,” also ideogram for same.
28		wick	Phonogram <i>h</i> .
29		swab	Phonograms <i>w3h</i> and <i>sk</i> . Determinative in <i>h3r</i> “ward off.” Variant of M1 in <i>mʿr</i> “fortunate.”
30		basket	Phonogram <i>nb</i> .
31		basket with handle	Variant  (V31a) in hieroglyphic transcriptions of hieratic texts, where the handle always faces the front. Phonogram <i>k</i> .
32		wicker satchel	Variant  (V96). Determinative in <i>g3wt</i> “bundle,” hence also in <i>g3w</i> “absence, lack,” hence also in <i>d3rw</i> “need.” Determinative in <i>msnw</i> “harpooner.” Phonogram <i>msn</i> in <i>msn</i> “Mesen” (a Delta town).
33		bag	Variants  (V34),  (V35). Determinative in <i>ʿrf</i> “pack, envelop,” <i>stj</i> “perfume,” and <i>šs(r)</i> “fine linen.” Phonogram <i>g</i> in a few words. Ideogram for <i>s3rw</i> “grain.” Determinative LINEN.
36		receptacle of cloth	Phonogram <i>hn</i> .
37		bandage	Determinative in <i>jdr</i> “herd,” also ideogram for same. Determinative in <i>jdr</i> “bandage.”
38		bandage	Determinative in <i>wt</i> “wrapping.”
39		tie	Ideogram for <i>tjt</i> “Isis-knot” (amulet).




## W. Stone and Ceramic Vessels







1		oil-jar	Determinative OIL. Ideogram for <i>mrht</i> “oil.”
---	---	---------	--













- 2  W1 without ties  
Phonogram *b3s* in *b3stt* "Bastet" (goddess). Determinative in *b3s* "oil jar." Variant of W1.
- 3  alabaster basin  
Variant  (W4). Determinative FEAST. Ideogram for *h3b* "feast."
- 5  T28 + W3  
Ideogram for *h3j-h3bt* "lector priest."
- 6  metal vessel  
Determinative in *wh3t* "cauldron."
- 7  granite bowl  
Variant  (W8). Determinative in *m3t* "granite" and *m3t* "proclaim." Determinative in *3bw* "Elephantine," also ideogram for same. Determinative in *3bt* "family."
- 9  stone jug  
Phonogram *hnm*.
- 10  cup  
Determinative in words with "b. Determinative in *ush* "wide" and related words, also ideogram for same. Phonogram *hnu* in *hnwt* "mistress" (from *hnt* "cup"). Determinative CUP. Variant of N41 in words with *h3j*.
- 10a  pot  
Variant  (Aa4). Phonogram *b3* in conjunction with E10 or G29.
- 12  jar stand  
Variant  (W11). Phonogram *g*. Determinative in *nst* "seat," also ideogram for same. Variant of W13 and O45.
- 13  pot  
Determinative in *d3rt* "red-ware," also ideogram for same.
- 14  water jar  
Phonogram *hz/hs*. Determinative in *hzt* "water jar" and *snbt* "jar," also ideogram for former.
- 15  water jar with water  
Variant  (W16). Determinative in *qbb* "cool" and *qbh* "cool, water," also ideogram for latter.
- 18  water jars in a rack  
Variants  (W17),  (W18a),  (W17a). Phonogram *hnt*. Ideogram for *hntw* "jar-rack."
- 19  milk jug with handle  
Phonogram *mj* (originally *mr*). Determinative in *mhr* "milk jug."
- 20  milk jug with cover  
Variant  (W59). Determinative in *jrt* "milk."
- 21  wine jars  
Determinative in *jrp* "wine."
- 22  beer jug  
Variant  (W23). Determinative POT. Ideogram for *hnqt* "beer" in offering formulas. Ideogram for *wdpw* "waiter."
- 24  pot  
Phonogram *nw*. Phonogram *jn* in *jnk* (1s pronoun). Variant of N33 in words with *qd*. Determinative in *d3d3t* "council" and *nhbt* "Nekhbet" (goddess), for unknown reasons. Often combined with Aa27 as phonogram *nd*. Variant of W22-23 as determinative.
- 24a  W24 + N35a  
Ideogram for *m-hnw* "inside" = *m(w)-h(r)-nw*.
- 25  W24 with legs  
Phonogram *jn* in forms of *jnj* "get, fetch, bring."
- 54  pot pouring water  
Variant of D60 and A6.

## X. Bread



















- 1  flat loaf of bread  
Phonogram *t*. Ideogram for *t* "bread." Often phonogram for *(j)t(j)* "father," alone or in conjunction with I9.
- 2  tall loaf of bread  
Variant  (X3). Determinative BREAD, FOOD. Ideogram for *t* "bread" in offering formulas. Ideogram for *dhwtj* "Thoth." Variant of X1 as phonogram for *(j)t(j)* "father."

- 4  bread roll  
 Variants  (X4a), and  (X5). Determinative BREAD, FOOD. Determinative in words with *zn* (from *znw* "food offerings"). Variant of W3.
- 6  round loaf of bread  
 Determinative in *p3t* "origin" and related words; and in *p3t* "loaf."
- 7  half-loaf of bread  
 Determinative BREAD. Doubled, ideogram for *wnm* "eat."
- 8  bread mold  
 Phonogram *dj/d* (originally *ḏj*) in forms of *rdj* "give," rarely in other words.



































## Y. Writing, Games, and Music

- 1  papyrus scroll  
 Variants  (Y2),  (Y1a). Determinative WRITING, ABSTRACT CONCEPTS. Ideogram for *dmd* "total." Ideogram for *md3t* "scroll" and *md3t* "chisel."
- 3  scribe's kit  
 Variant  (Y4). Ideogram for *zh3* "write" and related words. Determinative in *n<sup>c</sup>* "smooth" and *tns* "ruddy" and related words, also ideogram for same. Determinative in *mnhd* "scribe's kit."
- 5  game board and pieces  
 Phonogram *mn*.
- 6  game piece  
 Determinative in *jb3* "game piece," also ideogram for same. Determinative in *jb3* "dance," also ideogram for same.
- 7  harp  
 Determinative in *bjnt* "harp."
- 8  sistrum  
 Determinative in *z3st* "sistrum." Variant of S42.
- 10  bundle of stems  
 Determinative in *3<sup>c</sup>t* "murderousness" (from *3<sup>c</sup>* "cut").

## Z. Strokes and Figures

- 1  stroke  
 Used as ideogram of signs meant to be read as ideograms rather than phonograms (§ 3.3). Occasionally transferred to phonograms: for example,  *hr* "face" but also preposition *hr* "upon." Determinative in *w<sup>c</sup>* "one," also ideogram for same. Written one to nine times as ideogram for numerals 1 to 9 (§ 9.1). Substitute for A1.
- 5  diagonal stroke  
 Replacement for complex or dangerous signs.
- 4  two strokes  
 Variant  (Z49). Phonogram *j* as ending. Determinative DUAL.
- 2  three strokes  
 Variants  (Z2c),  (Z3a),  (Z3),  and  (Z2a-b),  (N33a). Determinative PLURAL. Also used with words that are plural in meaning, such as collectives, food, and minerals, and with singular words ending in *w* or *wt* ("false plurals"): § 4.6. Determinative in *hmt* "think" (from *hmtw* "three").
- 6  hieratic variant of A13-14  
 Determinative DIE, ENEMY. Sometimes similar to F20.
- 7  from hieratic variant of G43  
 Phonogram *w*. Different from V1.
- 8  oval  
 Determinative ROUND, OVAL.
- 9  crossed sticks  
 Variant  (Z10). Determinative BREAK, CROSS, NUMBER. Phonograms *sw3/zw3* in *sw3j* "pass" and *zw3* "cut off," *sd* in *sd<sup>t</sup>* "flame," *šbn* "mix" and related words, *hbs* in *hbsw* "cultivation," *wp* in *wp-st* "detail, breakdown," and *wr* in a few words.
- 11  crossed planks  
 Phonogram *jm*. Variant of M42.

## Aa. Unclassified

- 1  placenta?
- 2  pustule or gland
- 3  Aa2 with liquid emerging
- 4  pot
- 5  part of a ship
- 6  unknown
- 7  unknown
- 8  irrigation channel?
- 9  unknown
- 10  unknown
- 11  platform
- 13  unknown
- 16  front half of Aa13
- 17  lid
- 19  unknown
- 20  bag for clothing
- 21  unknown
- 24  warp between stakes
- 25  unknown
- 26  unknown
- 27  spindle
- 28  builder's level
- 31  frieze element
- 32  bow
- Variant . Phoneme *h*.
- Determinative SWELLING, UNHEALTHY. Variant of a number of older signs: F52 and N32 as determinative EXCREMENT, CLAY; M41 as determinative in *ʿ*ʿ "cedar"; V32 as determinative in *g3w* "absence, lack" and *g3wt* "bundle"; V38 as determinative in *wt* "bandage" and related words, and *stwh* "treat," also ideogram for former; W6 as determinative in *wh3t* "cauldron," also phonogram *wh3* in same and in *wh3t* "oasis"; W7 as determinative in *m3t* "granite" and *3bw* "Elephantine"; Z10 as determinative in *h3b* "count," also ideogram for same.
- Variant of Aa2 as determinative SWELLING, UNHEALTHY.
- Variant of W10a.
- Variant  (Aa5a). Phonogram *h(j)p*. Ideogram for *h3pt* "oar."
- Determinative in *tm3* "cadaster" and *tm3* "mat." Different from S23.
- Variant . Determinative in *sqr* "smash."
- Phonogram *qn*. Determinative of *sp3t* "estate, farm," also ideogram for same. Determinative of *d3d3t* "council." Variant of N24 as ideogram in *sp3t* "nome"; O34 as phonogram *z* in *zmjt* "desert"; V26 as phonogram *ʿd*.
- Determinative in *hwd* "rich."
- Determinative in *drt* "writing."
- Variants , ,  (Aa12). Phonogram *m3c*. Determinative in *mt3t* "platform."
- Variants ,  (Aa14),  (Aa15, with horizontals parallel). Phonograms *jm* and *m*. Variant of Aa16.
- Ideogram for *gs* "side, half," phonogram *gs*.
- Variant  (Aa18). Phonogram *s3*. Ideogram for *s3* "back."
- Determinative in *hr* "prepare" and *hrj* "terrified" and related words. Determinative in *t3r* "secure."
- Phonogram *pr*.
- Variant  (Aa22). Phonogram *wdc*. Ideogram for *wdc* "judged one" (term used in place of *sth/sts* "Seth").
- Variant  (Aa23). Determinative in *md* "puncture, press, adhere" and related words, also ideogram for same.
- Ideogram in *zm3* "stolist" (priest's title).
- Determinative in *sbj* "rebel."
- Phonogram *nd*. Often used in conjunction with W24.
- Variant  (Aa29). Phonogram *qd*.
- Variant  (Aa30). Determinative in *hkr* "adorn" and related words, also ideogram for same.
- Variant . Phonogram *stj/stj* in *t3-stj* "Nubia" and *stj* "ocher."

From James P. Allen, *Middle Egyptian: An Introduction to the Language and Culture of Hieroglyphs*, Cambridge, 1999, pp. 423–448.

# Akkadian and Eblaite

JOHN HUEHNERGARD AND CHRISTOPHER WOODS

## 1. HISTORICAL AND CULTURAL CONTEXTS

### 1.1 Classification and dialects

Akkadian, the language of the ancient Babylonians and Assyrians, is the oldest known Semitic language and the most widely attested member of the eastern branch of the Semitic family – the other member being the closely related Eblaite, which is considered by many Assyriologists to be a dialect or subbranch of Akkadian. The name Akkadian (*akkadû*), used by the Babylonians and Assyrians themselves for their language, derives from the northern Babylonian city of *Akkad(e)*, the capital city built by Sargon in about 2300 BC. While it is not known when speakers of Akkadian, or of its linguistic predecessor(s), first arrived in Mesopotamia, Akkadian personal names first appear in Sumerian texts dated to about 2600 BC. Connected Akkadian texts appear *c.* 2350 and continue more or less uninterrupted for the next two and a half millennia, with the major text genres attested for most periods. Akkadian probably died as a spoken language in the middle of the first millennium BC when it was gradually replaced by Aramaic. However, Akkadian continued in use as a liturgical and learned language until the beginning of the current era; the latest positively dated Akkadian text comes from the first century AD. To date, nearly one million texts have been excavated, and with ongoing excavations in Iraq, Syria, and Turkey this number steadily increases. The majority of these texts remain unpublished.

As may be expected from the remarkably long life and wide distribution of Akkadian, numerous dialects and geographical variations can be identified. Traditionally, Assyriologists acknowledge eight major subphases or dialects which roughly correspond to the major political periods. However, within these often arbitrary divisions, further geographical and chronological distinctions can be delineated. The earliest such major subphase, *Old Akkadian*, the language spoken by the Sargonic kings, refers collectively to the texts from the earliest attestation of Akkadian (mid-third millennium) to the beginning of the second millennium. Owing to the relatively small size of the Old Akkadian corpus, many grammatical forms are thus far still unattested.

Contemporaneous with the date of Old Akkadian materials, *c.* 2400 BC, are the thousands of texts excavated at Ebla in northwest Syria. Although the majority of these texts were written in Sumerian, many were written in a Semitic language, referred to as *Eblaite*, which has striking similarities to Old Akkadian. Eblaite is attested in bilingual (Sumero-Eblaite) lexical texts, administrative documents, treaties, incantations, and several literary texts, some of which have parallel versions from the Mesopotamian sites of Abū Šalābīkh and Fara (Šuruppak). Eblaite is attested only for a few generations during the

middle of the third millennium; thus, the historical development of the language cannot be traced.

The beginning of the second millennium marks a watershed in the development of Akkadian, after which the language is characterized by two broad geographical dialects, *Babylonian* in southern Mesopotamia and *Assyrian* in the north. The two are distinguished by several phonological differences, by minor morphological variations, and, to a limited extent, by lexicon. Significantly, some of the characteristic features (in the areas of phonology and morphology) of both the Assyrian and Babylonian dialects cannot be derived directly from the attested forms of Old Akkadian. Within the broad geographical dialects of Babylonian and Assyrian, chronological divisions of approximately five hundred years, labeled Old, Middle, and Neo-, are recognized for each:

(1)	Old Babylonian (OB)	2000–1500 BC	Old Assyrian (OA)
	Middle Babylonian (MB)	1500–1000 BC	Middle Assyrian (MA)
	Neo-Babylonian (NB)	1000–600 BC	Neo-Assyrian (NA)
	Late Babylonian (LB)	600 BC–AD 100	

Throughout the parallel development of Babylonian and Assyrian, the latter was always the more restricted dialect, limited primarily to Assyria proper. Even the Old Assyrian materials, which hail for the most part from eastern Anatolia (particularly from the site of Kaneš, modern Kültepe), represent primarily the business transactions of native Assyrian merchants residing in far-flung outposts in Anatolia. Babylonian was the more cosmopolitan of the two, reflecting Babylonia's perennial ascendancy in matters of culture, and it was not uncommon for even the Assyrian kings to adopt the Babylonian dialect when recording their inscriptions and annals.

Middle Assyrian is more sparsely attested than Old Assyrian, although it displays a variety of genres, including royal inscriptions, legal and economic texts, and an important collection of laws, the so-called Middle Assyrian laws. Neo-Assyrian is very well-preserved and was the language of Assyria under the important Sargonid dynasty until its fall in the latter half of the seventh century BC.

Old Babylonian is often considered by modern scholars as the classical phase of Akkadian, not only because of the remarkable uniformity of its grammar, but also because literature and scholarship flourished during this period. Old Babylonian is extremely well-preserved, and nearly all major text genres (discussed below) are attested for it. Indeed, the scribes of subsequent periods, in both Babylonia and Assyria, evidently regarded Old Babylonian as a classical language as well, as witnessed by the rise of *Standard Babylonian* (or *Jungbabylonisch*) – a contrived, nonspoken dialect of the first millennium which was based on archaic Old Babylonian features and used for the composition and transmission of literary works such as *Gilgameš* and *Enūma Eliš* as well as for many Assyrian and Babylonian royal inscriptions.

Middle Babylonian is much more sparsely attested than Old Babylonian and is known primarily from letters, economic texts, and a few royal inscriptions. Neo-Babylonian is well-preserved, especially in letters and economic texts written during the time of the short-lived Chaldean dynasty (625–539 BC) and the subsequent Persian occupation. Late Babylonian was written during the late Persian period and subsequent Seleucid occupation of Babylonia; it is heavily influenced by Aramaic, the spoken language of the time.

During the time of Middle Babylonian (particularly c. 1500–1200), Akkadian was used as a lingua franca throughout the ancient Near East. An archive of some 350 letters unearthed at Tell el-Amarna in Egypt records the diplomatic exchanges of the independent states of the ancient Near East, including Babylon, Assyria, Mittani, Hatti, Cyprus, and Egypt.

Additionally, Akkadian archives have been found at Alalakh, Ugarit, and Emar in Syria, and Hattuša in Anatolia, among other sites. The language of these texts, which were written by scribes who were not native Akkadian speakers, is frequently termed *Peripheral Akkadian* because of the heavy influence of the scribes' native tongues and its variance with normative Akkadian grammar.

The grammatical sketch of Akkadian presented here will be largely based on Old Babylonian, although important dialectal variations and diachronic developments will be noted.

## 1.2 Text genres

Akkadian is represented by an extremely wide variety of genres including both personal and court letters, royal inscriptions, annals, treaties, legal texts, law collections, such as the Old Babylonian Laws of Hammurabi, as well as many administrative and economic dockets such as purchase, loan, and rental agreements, and marriage, divorce, and adoption contracts. Additionally, an array of scholarly works are preserved, including historical and chronographical texts, mathematical and medical texts, literary commentaries, and grammatical and lexical compendia. There is also a large number of magical and divinatory texts, as well as ritual and religious texts, particularly for the later periods. Literary works, such as myths and epics, are largely preserved in the archaizing literary language of Standard Babylonian (about 40,000 lines of text are preserved) and, to a lesser degree, in Old Babylonian.

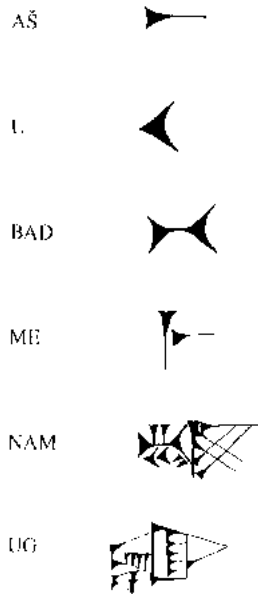
## 1.3 Sumerian and Akkadian

As demonstrated by early Sumerian texts bearing Akkadian personal names, particularly those from the southern Mesopotamian site of Abū Šalābikh, the Sumerian and Akkadian populations commingled and interacted, at least on the border regions between northern and southern Babylonia, from at least the dawn of attested Akkadian, c. 2600 BC, until the death of Sumerian as a spoken language. During the course of this long period of integration, Akkadian was greatly influenced by Sumerian at every level – phonologically, morphologically, syntactically, and lexically. Most of these developments were already underway, if not completed, by the time of the Sargonic kings and the first connected Akkadian texts c. 2350 BC. Thus, while the Akkadian at our disposal may be described as morphologically and syntactically conservative, a form of the language not subject to Sumerian interference has not survived. Significantly, the level of Sumerian influence, especially with regard to lexicon, is markedly greater in Babylonian than in Assyrian.

# 2. WRITING SYSTEM

## 2.1 Description and development

Akkadian and Eblaite used the cuneiform system of writing which the Sumerians devised during the fourth millennium to write their language. The system consists of wedge-shaped graphs (hence *cuneiform*, from Latin *cuneus* “wedge”) which were usually impressed into wet clay with a reed stylus. Other media were also employed, including wax, metal, and particularly stone for the recording of monumental inscriptions.





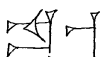
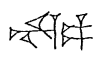
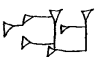




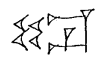


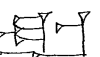


























**Figure 8.1** Sample of Old Babylonian lapidary signs

As the forms in Figure 8.1 reveal, individual signs may consist of one or several wedges. Akkadian script is read from left to right with the notable exception of the Code of Hammurabi, which was read from top to bottom, representing a purposefully archaizing attempt to mimic the Sumerian writing of the early third millennium. Except for the inconsistent use of a word-divider mark in Old Assyrian texts, there are no punctuation marks; often there is not even a dividing space between words.

The earliest Sumerian writing was fundamentally logographic in that a given graph represented a word, or a range of semantically related words. In many cases, but by no means all, the earliest graphs shared a pictographic relationship with their referents, being realistic, albeit stylized and highly conventionalized, depictions of the items represented. But already by the middle of the third millennium BC the graphs were largely stylized and thus unrecognizable as the objects they originally depicted. One of the main motivations for this change was the producing of the signs as a series of impressed wedges, rather than as curvilinear incised lines – a development that greatly facilitated the physical act of writing on wet clay. Very early in the development of the script, *c.* 3000 BC, the logographic signs of early Sumerian began to be assigned purely phonetic values. This was accomplished with the invention of rebus writings whereby a word symbol (logogram; see §2.3.20) such as A (Sumerian words are conventionally transcribed by Roman capitals), the Sumerian word meaning “water,” could be used more generally to represent the syllable [a]. The graph A could thereby be used, for example, to express the Sumerian locative case-marker which also happened to be pronounced [a], solving the problem of how “in-ness” might be expressed pictographically.

Thus, by the Old Akkadian period, when Akkadian is first attested in connected texts, the script was largely phonetic, in part logographic, and graphically stylized in comparison with the earliest signs. As spoken Akkadian evolved along parallel Assyrian and Babylonian courses, the Assyrian and Babylonian scripts, too, developed somewhat independently of

Sign and Meaning	Archaic Sumerian Pictograms c. 3000 BCE	OAKK		OB/OA	MB/MA	NB/NA
KA "mouth"			Bab.			
			Ass.			
KIRI <sub>6</sub> "garden"			Bab.			
			Ass.			
GUD "ox"			Bab.			
			Ass.			
MUŠEN "bird"			Bab.			
			Ass.			
KU <sub>6</sub> "fish"			Bab.			
			Ass.			

**Figure 8.2** Comparison of Sumerian, Babylonian, and Assyrian cursive cuneiform signs

one another. Additionally, there was considerable difference between the shape of the signs used for lapidary inscriptions and those used on clay.

Modern sign lists recognize nearly 600 signs, although there is great variation in the number and even the types of signs attested for various dialects. For the most part, the number of commonly used signs for a given dialect was often considerably less; Old Babylonian, for example, used about 150 frequent signs.

## 2.2 Recovery and modern decipherment

After the last Akkadian texts were written at the beginning of this era, c. AD 100, Akkadian and the cuneiform writing system faded into oblivion. Unlike the situation in Egypt where a form of the language, namely Coptic, lingered on until modern times, even if only as a liturgical language, knowledge of Akkadian and the cuneiform writing system was completely lost. Aside from the isolated reports of travelers during the late Middle Ages,



it was not until the seventeenth century that clay artifacts with “certain unknown characters” were brought back to Europe. The eighteenth century witnessed the first organized missions to collect information and artifacts concerning the ancient Near East. The first and most notable of these was sponsored by the Danish crown in 1761 under the direction of mathematician Karsten Niebuhr. Soon thereafter, philologists began the decipherment of the cuneiform languages. The publication of a long trilingual inscription (Akkadian–Old Persian–Elamite), which Darius engraved high on the rock of Behistun in western Iran, the so-called Rosetta stone of Assyriology, greatly aided attempts at deciphering Akkadian. Once the simpler script of Old Persian was deciphered, it was then possible to begin decoding the Akkadian version. The efforts of three men should be noted, as they contributed most significantly to the decipherment effort – Edward Hincks, Jules Oppert, and Sir Henry Creswicke Rawlinson. In 1857 the Royal Asiatic Society invited these three men, along with the mathematician W. H. Fox Talbot, to prepare independent translations of an unpublished Middle Assyrian text. When the four translations were compared and found to be reasonably close, the decipherment of the Akkadian was officially validated. Since that time a great deal of scholarship has been devoted to the publication of texts, the clarification of the grammar, and the preparation of dictionaries. Akkadian is considered to be well understood at the present time, although it still has a few dark corners.

In 1968 an Italian excavation identified the Syrian site of Tell Mardikh with the ancient city of Ebla, and with the thousands of texts found at that site came the discovery of the Eblaite language. Initially, Eblaite was thought to be an early form of West Semitic because of the location of the site in northwest Syria, well outside Mesopotamia proper. However, at this time there is a general agreement among Assyriologists that Eblaite represents a form of East Semitic and, possibly, even an early dialect of Akkadian. The numerous problems encountered in Eblaite orthography have greatly hampered the decipherment effort, and thus an understanding of the Eblaite language remains very much in its infancy.

## 2.3 Signs

Akkadian is expressed using three types of signs: logograms, phonetic signs, and determinatives. The three types are formally indistinguishable from one another, and certain signs may be used in all three roles in different contexts.

### 2.3.1 Phonetic signs

As noted above, the cuneiform script used by Akkadian is partly syllabic and partly logographic. For all dialects, and for most genres, the use of syllabic writings dominates. Syllables or parts of syllables are expressed by phonetic signs, or *syllabograms*, which may represent a vowel alone (V) or a sequence of a vowel and consonant(s), such as VC, CV, CVC; individual consonants cannot be written. The assignment of VCV or CVCV values to certain signs reflects the application of morphological or morphophonemic rules to the writing system (Reiner 1966:28).

When providing a sign-by-sign rendering of the cuneiform (i.e., a *transliteration*; see §2.4), phonetic signs are given in italics and connected by hyphens: for example, *a-wi-lum* “man”; *e-ka-al-lum* “palac.”

### 2.3.2 Logograms

Word signs, or logograms, are Sumerian words or phrases that must be read with the corresponding Akkadian value. Thus, logograms, often referred to as *Sumerograms*, may

be said to have Sumerian graphic etymologies, but represent Akkadian phonic material (Reiner 1966:26). For example, in the sentence LUGAL *a-na a-lim ik-šu-ud* “the king arrived at the city,” the logogram LUGAL, representing the Sumerian word for “king,” is to be read as *šarrum*, the Akkadian equivalent. Most words may be written either syllabically or logographically; often the scribal conventions of a given dialect or genre dictate the preferred writing, but otherwise the choice is one of scribal whim.

As the above example illustrates, logograms are traditionally transcribed in (nonitalicized) capital letters according to their Sumerian pronunciation. The individual components of a *compound logogram* are separated by a period: for example, DUMU.MUNUS, literally “child.female,” for Akkadian *mārtum* “daughter” (on an additional conventional use of transcription with capitals, see §2.5.2).

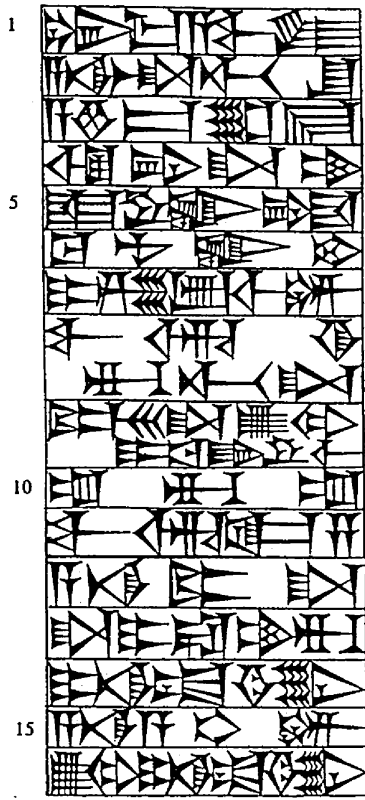
Logograms are often followed by phonetic signs, known as *phonetic complements*, which usually serve to clarify the Akkadian reading of the logogram by specifying the pronunciation of the last part of the word. The use of a phonetic complement may limit the interpretation of a logogram to one Akkadian word from among several possible readings: thus, KUR-*tum* for *mātum* “country,” versus KUR-*ú-um* for *šadûm* “mountain,” where the words for “country” and “mountain” are homonymous in Sumerian. In other cases a phonetic complement may indicate part of the morphological shape of a given Akkadian word, such as (i) the appropriate case ending, as in A.ŠÀ-*lum* (= *eqlum*, nom.) “field,” A.ŠÀ-*lam* (= *eqlam*, acc.), A.ŠÀ-*lim* (= *eqlim*, gen.); or (ii) a possessive pronominal suffix, as in A.ŠÀ-*šu* (= *eqelšu*, nom./acc., or *eqlišu*, gen.) “his field.”

### 2.3.3 Determinatives

Akkadian writing borrowed from Sumerian a subset of logograms used to specify the semantic class to which a given word belongs. Determinatives were a feature only of the writing system and had no phonological value. In transliterations of Akkadian, determinatives are indicated in superscript either before or after the word they modify, according to their placement in the cuneiform text. Although very common with certain words, determinatives were optional and not a mandatory part of the writing of any word. There are roughly nineteen commonly used determinatives; examples include: GIŠ (= *išum*) “wood,” often used before objects made of wood (e.g., <sup>giš</sup>GU.ZA = *kussûm* “throne”); KI (= *eršetum*) “place, land, district,” used after city and country names (e.g., KÁ.DINGIR.RA<sup>ki</sup> = *Bābil* “Babylon”); DINGIR (= *ilum*) “god” (abbreviated<sup>d</sup> in transliteration when used as a determinative), used before god names (e.g., <sup>d</sup>EN.LÍL [= *Enlil* or *Ellil*] “(the god) Enlil”).

## 2.4 Transliteration and transcription

Akkadian is rendered into Latin characters in two distinct forms. A sign-by-sign rendering, or *transliteration*, attempts accurately to reflect the signs expressed in the cuneiform text; as noted above, phonetic signs belonging to a single word are connected with hyphens, while logograms are written in capital letters and are connected with periods. A *transcription*, or *normalization*, attempts to reflect the actual pronunciation, indicating vocalic and consonantal length (i.e., consonantal doubling). The phonetic signs of a given word are connected without their distinguishing diacritic marks (see §2.5.1), and logograms are written with their corresponding Akkadian equivalents. Transliteration and transcription are illustrated in (2) with a portion of Law §150 from the Code of Hammurabi.



**Figure 8.3** Law (§150) from the Code of Hammurabi (turned 90°) with Transliteration, Transcription, and Translation. (Autographed text after Harper, 1904)

(2)	Transliteration	Transcription	Translation
<sup>1</sup>	šum-ma a-wi-lum	šumma awilum	If a man
<sup>2</sup>	a-na aš-ša-ti-šu	ana aššatišu	to his wife
<sup>3</sup>	A.ŠÀ <sup>giš</sup> KIRI <sub>6</sub> É	eqlam kiriam bītam	a field, orchard, house
<sup>4</sup>	ù bi-ša-am	û bišam	or movable property
<sup>5</sup>	iš-ru-uk-šim	išrukšim	gave to her, (and)
<sup>6</sup>	ku-nu-uk-kam	kunukkam	“a sealed document”
<sup>7</sup>	i-zi-ib-ši-im	izibšim	made out for her,
<sup>8</sup>	wa-ar-ki	warki	after
	mu-ti-ša	mutiša	her husband (has died),
<sup>9</sup>	DUMU.MEŠ-ša ú-ul	mārūša ul	her children will not
	i-ba-qá-ru-ši	ibaqqarūši	bring a claim against her;
<sup>10</sup>	um-mu-um	ummum	the mother
<sup>11</sup>	wa-ar-ka-sà	warkassa	her estate
<sup>12</sup>	a-na DUMU-ša	ana mārīša	to her child
<sup>13</sup>	ša i-ra-am-mu	ša irammu	whom she loves
<sup>14</sup>	i-na-ad-di-in	inaddin	will give,
<sup>15</sup>	a-na a-ḫi-im	ana aḫim	to an outsider
<sup>16</sup>	ú-ul u-na-ad-di-in	ul inaddin	she will not give (it).

“If a man gave to his wife a field, orchard, house, or movable property, and made out a sealed document (i.e., contract) for her, after her husband’s death, her children will not bring a claim against her; the mother will give her estate to the child whom she loves – she will not give it to an outsider.”

## 2.5 Characteristics and problems of Akkadian orthography

The cuneiform writing system, in its earliest manifestation, was in essence a type of mnemonic device or *aide-mémoire* that was only loosely tied to the spoken language; it was in this sense an incomplete system in which only the core elements of speech were represented, the decoder or reader of a text having to depend on his knowledge of the language and the context of the message to restore most morphological markers and grammatical elements. Even as Sumerian writing evolved, becoming more closely tied to the spoken language, it was possible to omit certain, more or less predictable, elements from the writing that were presumably present in the utterance. Thus, at no time did the orthography strive to render an exact phonemic representation of the language. This fundamental weakness of the writing system was greatly exacerbated when the script was applied secondarily to express Akkadian; for as poorly as the script represented the phonemes of Sumerian, it expressed to an even lesser degree the phonemic inventory of Akkadian, which was genetically unrelated to Sumerian. The preservation of many Sumerian values along with the addition of new Akkadian values served to complicate the system even more. Consequently, cases of both over-differentiation and under-differentiation of Akkadian phonemes occur within the writing system. Although, during the evolution of the script, scribes attempted to alleviate these problems with the development of new signs and with the secondary differentiations of old ones, these attempts complicated the system still further. Moreover, the structure of Akkadian was less suited to the cuneiform writing system than that of Sumerian, as the agglutinative nature of Sumerian lent itself to the syllabic script more readily than did the inflecting morphology of Akkadian. Some of the more important complications associated with the writing system will be taken up in the following sections.

### 2.5.1 Homophony

Indicative of the over-differentiation of Akkadian phonemes in the writing system is the common existence of two or more signs with the same phonological value. These homophonous signs are distinguished in transliteration by a convention of diacritical marks. The most frequent sign for a given value is unmarked, the second most frequent is marked with an acute accent over the vowel, the third with a grave accent, while the fourth and following signs are marked with subscript index numbers: for example, *gi*, *gí*, *gì*, *gi₄*, *gi₅* . . . *gi₂₇*. Paradoxically, the presence of exceptionally high index numbers such as *gi₂₇* or *še₂₉* does not indicate that homophony played an unduly important role in the Akkadian syllabary. Rather, the system of diacritical marks accounts for the entire time-span of attested Akkadian, including not only the Assyrian and Babylonian dialects, but the peripheral dialects as well. For a given corpus, defined both temporally and geographically, the extent of homophony is much more limited, with at most two or three signs expressing a given value. In many cases the presence of homophonous signs is conditioned by scribal habits; for example, during the Old Babylonian period the Û sign is restricted to writing the conjunctions *u* “and” and *ū* “or,” whereas the Ú sign is used for most other occurrences of /u/ (use of the U sign, “*u*-one,” was uncommon during this period).

### 2.5.2 Polyphony

Many signs may have more than one phonetic value; for example, the logogram UD also has the phonetic values *tam*, *tám*, *pír*, and *par*, among others. The correct value for a given sign must be determined by context. As with homophony, the role of polyphony in Akkadian is limited by dialectal considerations. Note that by Assyriological convention capital letters are

used to indicate a given sign, such as UD, without specifying the phonological value with which it is to be read.

### 2.5.3 Distribution of signs

The failure of the writing system to distinguish many phonemes (i.e., phonemic under-differentiation) manifested itself in the distribution of certain signs. For instance, the writing system did not contrast voiced, voiceless, or emphatic consonants (see §3.1.1) at the end of a syllable, presumably because such distinctions were not significant in Sumerian. In Akkadian, however, the opposition of voiced, voiceless, and emphatic consonants is phonemic. Thus, an Akkadian scribe had to content himself with using, for example, the UD sign to express /ut/ and /uṭ/ as well as /ud/, the AG sign to express /ak/ and /aq/ as well as /ag/, and so on for all VC signs. The same holds true not only for the final consonant of CVC signs, but also for the initial consonant of certain CVC signs: thus, variant readings are possible such as *d/tan*, *d/t/ṭal*, *d/t/ṭim*, *b/pan*, and so forth. For CV signs, voice is often distinguished (although not in Old Akkadian or Old Assyrian); for example, there are different signs for the syllables /ba/ and /pa/, /bi/ and /pi/. However, this distinction is not always made as, for example, there is only one sign for the pair /bu/ and /pu/. Emphatic consonants were almost never distinguished in the early periods: for example, the ZA sign represented both /za/ and /ša/, the DI sign represented /di/ and /ṭi/. Only in the later dialects were emphatics commonly distinguished with certain signs: for example, the KIN sign was used for /qi/ and /qe/ (as opposed to the KI sign for their expression in earlier dialects); KUM for /qu/, and GÍN for /ṭu/. The lack of these distinctions in the writing system clearly increased the degree of polyphony in the syllabary.

Several other features regarding the distribution of signs should be noted. Signs that contain /e/ are often not distinguished from those with /i/. For instance, while there are different signs for the respective syllable pairs /te/, /ti/; /me/, /mi/; /en/, /in/; and /eš/, /iš/, there is only one sign for the syllable pair /ge/ and /gi/, and only one for /ke/ and /ki/. Moreover, certain signs do not distinguish either between /e/ and /i/ or among voiced, voiceless, and emphatic consonants; thus, IG represents the syllables /eg/, /ek/, /eq/ and /ig/, /ik/, /iq/; IB, the syllables /eb/, /ep/ and /ib/, /ip/.

Only for the post-Old Babylonian dialects is there a specific sign for writing the glottal stop ʔ(/ʔ/). In the earlier dialects there were various conventions for expressing this consonant: (i) use of the *h*-signs (ḪA, ḪI, ḪU for syllable-initial /ʔ/, AḪ for syllable-closing /ʔ/, as in *i-na-AḪ-ḪI-id* for *ina* "id" "s/he will heed"); (ii) writing of the appropriate vowel sign (e.g., *le-ú-um* for *leʔum* "to be able"); or (iii) broken writings (see §2.5.4). The sign ʾA that appeared in later periods represents a graphic variant of the AḪ sign.

For each of the two glides, /w/ and /y/, there is only one sign used regardless of the accompanying vowel: the PI sign has the values /wa/, /we/, /wi/, /wu/ (and /aw/, /ew/, /iw/, /uw/); IA, the values /ya/, /yi/, /yu/ (and /ay/).

It has been proposed that the distribution of vowel signs in Old Akkadian distinguishes three degrees of consonantal onset in the representation of the glottal stops (Gelb 1961:24–28). According to this proposal, the Old Akkadian vowel signs, which are purely vocalic in Old Babylonian and later dialects, were used to express consonantal onset when a syllable began or, less often, closed with the reflex of one of the five Proto-Semitic guttural consonants *\*ʔ*(/ʔ/), *\*h*, *\*h*(/ḥ/), *\*ʕ*(/ʕ/), *\*ǵ*(/ǵ/) or the two Proto-Semitic glides *\*w* and *\*y* (referred to as ʔ<sub>1</sub>–ʔ<sub>5</sub> and ʔ<sub>6</sub>–ʔ<sub>7</sub> respectively by Assyriologists). According to this scheme – which, it must be noted, is not strictly adhered to in Old Akkadian orthography – the signs A, E, I, and Ú represent onset (not indicated in the transliteration); the signs Á, Ê, Ì, and Û are used to

express stronger onset; the É sign is used with the value ṛa to express the strongest onset, which is identified with the Proto-Semitic sequences \*ḥa- and \*ha-. Further, it is proposed that the signs BÍ, MÁ, RÍ, LÍ, and LÚ are used either for morphologically long vowels (marked by a macron), or for vowels followed by the reflex of one of the Proto-Semitic gutturals or glides.

### 2.5.4 Spelling conventions and sequences of signs

Akkadian words are always written according to their syllabification (see §3.4); in other words, signs are chosen so that the syllable boundaries are clear. For example, the word *šarrum* “king” may be written syllabically in the following ways: *šar-rum*, *ša-ar-rum*, *šar-ru-um*, *ša-ar-ru-um*, or *ša-rum*, but not as \*\**šar-um* (however, see below for exceptional Old Akkadian writings of this type). The choice of a C<sub>1</sub>V-VC<sub>2</sub> sequence rather than a single C<sub>1</sub>VC<sub>2</sub> sign is largely a matter of scribal preference, although CVC signs do not exist for all CVC combinations: for example, the sequence [paz] must be written *pa-az* since a specific *paz* sign does not exist in the syllabary. Additionally, the historical development of the syllabary may determine the choice, since C<sub>1</sub>VC<sub>2</sub> signs are relatively uncommon (except for CV*m* signs) before the Middle Assyrian and Middle Babylonian periods.

Writings of the type (C)VC-V(C), the so-called “broken writings,” do not conform to the syllabification of the language and thus are generally not tolerated. However, for the early periods, when the glottal stop was not distinguished by a specific sign, broken writings could be used to express it after another consonant. For the Old Akkadian period, such broken spellings additionally indicate the following: (i) a lexical base and a morphological ending (e.g., *i-šar-um* (nom. case ending *-um*) for [išarum] “straight”); (ii) a doubled consonant (e.g., *qar-ad* for [qarrād] (bound form) “strong”); or (iii) a combination of the two (as in *šar-um* for [šarrum] “king”).

Vowel length is, as a general rule, not indicated in Old Akkadian and Old Assyrian, although it is contrastive in the language. Vowel length can only be determined either by the surrounding grammar or by the context. For the Old Babylonian and later dialects, vowel length is similarly left unexpressed in most environments; a notable exception is Middle Assyrian where long vowels are often expressed with an extra vowel sign (so-called *plene writings* of the type CV+V). In the other dialects exceptional indications of vowel length occur when long vowels are derived from roots in which either the second or third consonant was “weak” (see §4.2). With the former, a morphologically long vowel in the middle of the word may occasionally be written with an extra vowel sign: for example, *kīn* “it (masc.) is firm” (< √*kwn*) may be written *ki-in* or *ki-i-in* with no difference in pronunciation. With the latter, a long vowel at the end of a word, resulting from the contraction of two vowels, is often indicated with an extra vowel sign: for example, *ib-nu-ú* for *ibnū* “they (masc.) built” (< *ibni* + *ū*).

Similarly, consonantal length (i.e., consonant gemination), although contrastive, may or may not be indicated in the writing. For Old Akkadian and Old Assyrian, again, length is, as a general rule, not expressed; for the later dialects, the expression of consonantal length is mainly a matter of scribal whim (the exception again being Middle Assyrian, as well as the letters from Hammurabi’s chancellery, which often express double consonants). Thus, for example, the word *inaddišsum* “s/he will give to him” may be written in any of the four following forms in Old Babylonian: *i-na-di-šum*, *i-na-ad-di-šum*, *i-na-di-iš-šum*, *i-na-ad-di-iš-šum*. Similarly, morphologically contrastive word pairs such as *ipparras* “it will be cut” and *iparras* “it will cut” may both be written as *i-pa-ra-as*; only the context can distinguish the two. With the exception of relatively rare writings of the type *i-din-nam* for *iddinam* “he gave to me” (see §2.5.5), expressly written double consonants always reflect long consonants in speech.



### 2.5.5 Morphographemic writings

Certain morphophonemic processes were reflected inconsistently in the orthography. *Morphographemic* writings reveal the constituent morphemes rather than the phonetic character of a given word. For example, /awāssu/ “his word” (representing the lexical base *awāt-* “word” and the 3rd masc. sg. suffix *-šu*) may be written with the partial morphographemic writing *a-wa-at-su* (beside the regular phonetic spelling *a-wa-as-su*). Although morphographemic writings of this type are occasionally encountered in the early periods, they are more frequent in the later dialects. However, Old Akkadian (and frozen Old Babylonian) writings of the type *šar-um* /šarrum/ and *i-din-nam* /yiddinam/, where the writing reflects a lexical base and a suffixed morphological ending, may also be considered morphographemic (see §2.5.4).

## 2.6 Eblaite orthography

The Eblaite orthographic tradition was quite similar to that of Old Akkadian, and, like the latter, Eblaite did not express vocalic or consonantal length. Likewise, the writing system did not distinguish among voiced, voiceless, emphatic, or otherwise similarly articulated consonants. Most syllabograms, therefore, represent a class or group of Proto-Semitic reflexes, some of which may or may not have merged. However, in the case of Old Akkadian, the existence of later dialects with different writing conventions allows, at least in part, for the disentanglement of true phonological mergings from the semblance of mergings suggested by the writing system. Thus, the existence of a single orthographic tradition for Eblaite, as well as Eblaite’s uncertain relationship to Akkadian, does not allow us to comment on possible mergings, or lack thereof, beyond what is suggested by the writing.

Unique to Eblaite are several other orthographic features, essentially archaic in nature, which serve further to obscure the phonology, morphology, and syntax. Foremost among these features is the fact that logograms, which do not reproduce the grammatical forms dictated by context, comprise the vast majority of writings; by some estimates, no less than 90 percent of the Eblaite corpus consists of logograms. Logograms may express any part of speech, including prepositions and verbs. Even pluralization may be expressed by logographic reduplication: for example, *KALAM-tim*.*KALAM-tim* “the lands” (gen. pl.); *UDU.UDU* “all the sheep.” Furthermore, in what may possibly represent a transition between logographic and phonetic spellings, syllabograms may function logographically, masking the expected grammatical form of a word. Specifically, orthographic conventions allow for the nominative singular to be written instead of the expected case or form dictated by context: for example, *il-tum* for expected /ʔilātim/ (fem. gen. pl.) or /ʔilī/ (masc. gen. pl.) “(of) the god(des)es.” Similarly, syllabograms functioning logographically may also be reduplicated to indicate the plural: for example, *NA.SE<sub>11</sub>.NA.SE<sub>11</sub>* for /niši/ or /naši/ “people” (gen. pl.).

Also peculiar to Eblaite are both (i) the frequent omission of certain speech elements, particularly prepositions and possessive suffixes (as in *an-da* ŠEŠ *ù an-na* ŠEŠ “you are [my] brother and I am [your] brother”); and (ii) the combination of Sumerian verbal forms with Eblaite pronominal suffixes (as in *l.NA.SUM-kum* “he has given to you”). When words are expressed syllabically, the lack of certain syllabograms in the Eblaite syllabary often results in the graphic omission of phonemes in certain environments: for example, in the verbal form *a-za-mi-ga* /ʔaš(am)mid-ka/ “I bound you” or “I bind you,” the /d/ phoneme of the syllable /mid/ is not expressed in the writing because of the lack of a VC syllabogram, *id*, in the syllabary; while the fact that *a-za* could represent two open syllables, /ʔaša/, or one closed syllable, /ʔaš/, means that the writing system could not even indicate the intended syllabic structure unambiguously.

**Table 8.1 The consonantal phonemes of Akkadian (post-Old Akkadian period)**

Manner of articulation	Place of articulation						
	Bilabial	Dental/Alveolar	Palato-alveolar	Palatal	Velar	Uvular	Glottal
<i>Stop</i>							
Voiceless	p	t			k		ʔ (/ʔ/)
Voiced	b	d			g		
Emphatic		ṭ (/ṭʰ/)			q (/kʰ/)		
<i>Fricative</i>							
Voiceless			š (/š/ / /s/)		ḫ (/x/)		
Voiced					r (/ɣ/)	~ r (/ʁ/)	
<i>Affricate</i>							
Voiceless		s (/ʰs/)					
Voiced		z (/ʰdʒ/)					
Emphatic		ṣ (/ʰsʰ/)					
<i>Approximant</i>							
Voiced	w			y			
<i>Lateral approximant</i>							
Voiced		l					
<i>Nasal</i>							
Voiced	m	n					

### 3. PHONOLOGY

#### 3.1 Phonemic inventory

##### 3.1.1 Akkadian consonants

The classification of Akkadian phonemes is based largely on the phonemic inventories of other Semitic languages and that postulated for reconstructed Proto-Semitic. To a much lesser degree, it also relies secondarily on the Greek transcriptions of Late Babylonian lexical texts – the so-called *Graeco-Babyloniaca*, as well as on Aramaic transcriptions of Akkadian words and Akkadian glosses of Egyptian. Naturally, the phonetic character of these phonemes is impossible to recover, given that Akkadian is a dead language.

Akkadian distinguishes twenty consonant phonemes – a marked reduction from the original twenty-nine posited for Proto-Semitic. It is commonly assumed that much of this reduction, particularly with regard to the Proto-Semitic *gutturals* (i.e., glottals \*ʔ (/ʔ/)) and \*h; pharyngeals \*ħ (/ħ/) and \*ʕ (/ʕ/), and voiced velar fricative \*ɣ (\* /ɣ/)), is due to extensive and long-term contact with Sumerian, for which these consonants are not attested. As a result, Akkadian underwent a more radical development of its phonological system than any other Semitic language before modern times. In Table 8.1 Akkadian consonantal phonemes are presented using their conventional representations; probable or possible phonetic values follow in parentheses where conventional transcription differs from common phonetic transcription. It must be noted, however, that the places of articulation, and hence the phonetic renderings given here, are approximate – the significance of the chart lies in the phonemic contrasts.

As Table 8.1 indicates, Akkadian possessed four voiceless stops, /p, t, k, ʔ/, three voiced stops, /b, d, g/, and two stops, ṭ and q, known in Semitic linguistics as *emphatics*, which were



probably characterized by glottalic coarticulation (i.e., /t'/ and /k'/ respectively). Similarly, there was a triad of affricates, voiced /<sup>d</sup>z/ (<z>), voiceless /t's/ (<s>), and emphatic /t's'/ (<š>). These became fricatives in later dialects; the voiceless member of this later, fricative set was pronounced [s] in Babylonian, but [š] in Assyrian, while the reflex of Proto-Semitic \*š, which was probably simple [s] originally, continued to be pronounced as such in Assyrian, but as [š] in Babylonian. The only other fricative was the velar /x/ (<ḫ>). There were six sonorants, the glides /w/ and /y/, the nasals /m/ and /n/, the liquid /l/, and, finally, /r/; as the last-mentioned often patterned with ḫ, it was probably realized phonetically as a voiced fricative, either velar ([ɣ]) or uvular ([ʁ]).

### 3.1.2 Eblaite consonants

As noted above, our understanding of the phonemic character of Eblaite is restricted by the limits imposed by the writing system; that is, the syllabic orthography did not distinguish single phonemes, but rather individual syllabograms representing groups of similarly articulated phonemes. It cannot be determined with any certainty which, if any, of these orthographically identical consonants actually merged phonologically. Generally, the phonology of Eblaite appears to be quite similar to that of Old Akkadian, and both languages seem to have maintained reflexes of at least some of the Proto-Semitic gutturals. Eblaite deviates most significantly from Akkadian in the treatment of the liquids /l, r/ (see §3.8.1, 8) and in the distinction between Proto-Semitic \*ḏ and \*z (see §3.8.1, 1).

Only the reflexes of the Proto-Semitic consonants \*ḫ, \*m, and \*n are expressed unambiguously in Eblaite orthography (i.e., with distinctive ḫV- (-Vḫ), mV- (-Vm), and nV- (-Vn) signs, respectively). The reflexes of each of the following Proto-Semitic consonants are expressed ambiguously (i.e., the consonants of each group are represented by a common group of signs): (i) \*b and \*p (with bV- and -Vb signs); (ii) \*d, \*t, and \*ṭ (dV-, -Vd, tV-); (iii) \*g, \*k, and \*q (gV-, -Vg); (iv) \*ś and \*š (śV-, -Vś); (v) \*ḏ and \*θ (šV-, -Vš); (vi) \*z, \*s, \*š, \*ś, and \*ṯ (zV-, -Vś). The reflex of \*l is written with distinctive l-signs (lV-, -Vl), while the reflex of \*r is written with r-signs (rV-, -Vr). Significantly, however, the reflex of \*r may also be written with the set of l-signs (for examples of this phenomenon see §3.8.1, 8). The converse phenomenon, writing the reflex of \*l with r-signs, does not occur.

As in Old Akkadian, the reflexes of the Proto-Semitic gutturals \*, \*ʕ, \*h, \*ḫ, and \*ḡ are evidenced in certain spellings, although the writing system was incapable of representing them properly or of distinguishing separate reflexes fully. However, there appears to have been an attempt, at least in part, to distinguish these reflexes by assigning one group of vowel signs to express \*ʔ and \*ʕ (i.e., a/ʔa<sub>x</sub>(NI), i, ù/u<sub>x</sub>(NI)) and another set to indicate \*h and \*ḫ (i.e., à(É)/a/ʔa<sub>x</sub>(NI), i/i, u<sub>9</sub>(EZENxAN)). The reflex of \*ḡ is attested in a single word, [ḡāribu(m)] “raven,” written with both ga- and ḫa-. As in Akkadian, \*w and \*y were preserved only in certain environments (see §3.8.1, 5, 6). The sequence wV was written with the PI sign (as in Old Akkadian). There was no sign in the syllabary for reflex of \*y; instead, where \*y is expected at the beginning of words and between vowels the following signs were used: a for /(y)a/, i for /(y)i/, u<sub>9</sub>(EZENxAN) for /(y)u/.

### 3.1.3 Akkadian and Eblaite vowels

Akkadian and Eblaite distinguish three primary vowel phonemes, /a/, /i/, and /u/, reflecting the original stock of Proto-Semitic vowels, and a fourth, /e/, which is secondary, derived from either /a/ or /i/:

## (3) Akkadian and Eblaite vowel phonemes

/i(:)/                      /u(:)/  
 /e(:)/  
                               /a(:)/

Vowels may be either long or short; short vowels are transcribed unmarked while two types of long vowels are distinguished in transcription: morphologically long vowels, marked with a macron, *ā*, *ē*, *ī*, *ū*, and long vowels which result from vowel contraction, marked with a circumflex, *â*, *ê*, *î*, and *û*.

## 3.1.3.1 The phoneme /e/

As noted above, the vowel /e/ was a secondary development in Akkadian. The Proto-Semitic gutturals \**ḥ*, \**ʕ*, \**g* had the effect of coloring neighboring \*/a/ vowels to [e] before being lost (see §3.8.1, 4). Only in the Babylonian dialects did the loss of these consonants additionally cause \*/ā/ > [ē]. As these gutturals, or their reflexes, appear to have been retained in Old Akkadian and Eblaite and the change of \*/a/ > [e] does not occur consistently in their presence (see §3.8.1, 4), [e] was simply an allophone of /a/ in these dialects. In other Old Akkadian environments /i/ had an allophone [e], as in the third-person prefix of verbs I-ʕ: for example, *i-mu-ru* [yiʔmurū] beside *e-mu-ru* [yeʔmurū] “they saw.” It appears, then, that for Old Akkadian, and probably Eblaite, *e* was not phonemic. However, with the merging and loss of the Proto-Semitic gutturals, *e* achieved phonemic status in the Old Babylonian period as evidenced by minimal pairs such as *pelūm* “egg” versus *palūm* “reign”; *šērum* “morning” versus *šārum* “wind”; *elīšu* “on it” versus *ilīšu* “his gods” (gen.-acc.); *rēmum* “pity” versus *rīnum* “wild bull,” although the writing system remained unable to reflect the distinction between /e/ and /i/ in many cases.

A vowel [e] also occurred in the post-Old Akkadian dialects as the result of various phonological changes:

1. In Babylonian, [a] and [e] were incompatible in the same word, with the result that long or short [a] was assimilated to [e], a process called *Babylonian vowel harmony*: thus, Bab. *epēšum* versus Ass. *epāšum* “to make”; Bab. *bēlētum* versus Ass. *bēlātum* (nom.) “ladies.” The change is conditioned even by an [e] subsequently lost by vowel contraction: for example *leqūm* < \**leqēm* < \**laqēm* < \**laqēhum* < \**laqāhum* “to take.” Babylonian vowel harmony did not take place across all morpheme boundaries; in addition, a secondary [e] (derived from /i/ or arising by vowel contraction) does not normally condition the change. However, the [a] of the prefixes *a-* and *ta-* often assimilates to [e] in verbs containing that vowel: for example, *ešme* beside *ašme* “I heard”; Bab. *teleqqe* versus Ass. *talaqqe* “you (masc. sg.) take.” The rule of Babylonian vowel harmony was not applicable to Old Akkadian or the Assyrian dialects.
2. The phoneme /a/ became /e/ in words that contained Proto-Semitic \**ʔ* and a Proto-Semitic sonorant \**m*, \**n*, \**r*, or \**l*, as in \**ʔaršatum* > (?)*eršatum* > *eršetum* “earth”; \**šaʔnum* > \**šeʔnum* > *šēnum* “sandal.”
3. The phoneme /i/ had an allophone [e], which occurred immediately before /*ḫ*/ ([x]) or /*r*/: for example, *laberum* “old,” *meḫrum* “copy, reply.”
4. Additionally, the loss of mimation (-*m* occurring in final position) caused [i] of the original word-final sequence [-im] to shift to [e] in Assyrian. Thus, the final vowel of the genitive singular, the accusative-genitive plural of feminine nouns and adjectives, the ventive for the plural, and the second and third feminine singular dative suffixes are pronounced [e] in Middle and Neo-Assyrian: for example, Old Assyrian

*šarritim* > Middle and Neo-Assyrian *šarrete* “queens”; Old Assyrian *illikūnim* > Middle and Neo-Assyrian *illikūne* “they came here”; Old Assyrian *išpuršim* > Middle and Neo-Assyrian *išpurše* “he sent to her.”

5. In Middle Babylonian, Standard Babylonian and Neo-Babylonian the [a] of the second syllable of the preterites, perfects, and precatives of D and Š stems often underwent partial assimilation to the [i] or [e] of the following syllable and appeared as [e]: for example, *ubenni* beside *ubanni* “he built”; *ušeknis* beside *ušaknis* “he subjugated.”

### 3.1.3.2 Akkadian /o/

The phoneme /o/ has been suggested for Akkadian primarily on the basis of the distribution of *u*-signs in certain Old Babylonian lexical texts from the city of Nippur, as well as of Greek transcriptions of Late Babylonian texts (Westenholz 1991). The evidence from the lexical texts suggests that the signs U and U<sub>4</sub> are used to express [ô], while Ú represents [u, ū]. The U and U<sub>4</sub> signs are consistently used to represent the contraction of [ǎ] + [ǔ] (or [ě] (<[ǎ]) + [ǔ]; i.e., where [ē] is derived secondarily from [ā]), as evidenced by such writings as, for example, *na-du-u<sub>4</sub>* “to throw” (*nadā + u*); *tap-pu-u* “companion” (*tappā + u*); *lā e-el-qú-u-ma* “I did not take [oath]” (*elqe + ū + ma*). The Ú sign is used in other environments, as (i) in the reflex of [i] + [u] and [u] + [u] (e.g., *ra-bu-ú* “great” (< *rabi + u*), *zu-uk-ku-ú* “to cleanse” (< *zukku + u*); (ii) in plene writings of [ū] (e.g., *du-lu-ú-tum* “hoisting device” (*dulūtum*)); and (iii) for short [u] (e.g., *im-du-ú* “support” (*imdu*), *ba-a-a-ú* “to walk, pass along” (*bā’u*). These writings are in agreement with later Greek transcriptions such as βιλλoδω[ζ] < *billuda-ū-šu*. The Greek evidence also suggests that Proto-Semitic (PS) \*aw > [ô]: for example, ω “day” = *ōw* < \**yawm-*. It is also proposed that the choice of U-signs in spellings such as *u<sub>4</sub>-ru-ḥu* “hair on the head” (*uruhḥu*) and *u<sub>4</sub>-ḥu-li* “suds” (*uḥūlum*) may indicate /u/ → [o]/\_{/r/, /ḥ/}. Further, the possible minimal pairs *nadôm* “to throw” (< *nadā + um*) and *nadûm* “thrown” (< *nadi + um*) may demonstrate that ô was phonemically distinct from û. However, given the limited evidence for an /o/ phoneme and given that the writing system did not distinguish *u* from *o* outside of the lexical materials, the phonological significance of these data remains uncertain. Moreover, the possibility exists that the scribal differentiation of *u*-signs in these lexical texts reflects the superimposition of the Sumerian phonological system (for which an /o/ phoneme has been similarly proposed), rather than a phonemic distinction within Akkadian (Lieberman 1979).

## 3.2 Possible allophones

In the Assyriological literature other variations in the pronunciation of various vowels and consonants have been pointed out; among these are the so-called vocalic consonants (syllabic allophones of /l, r, n, m/), a palatal nasal [ɲ], a voiced emphatic [g], spirantized variants of the nonemphatic stops (*b̥, p̥, t̥, d̥, k̥, g̥*), and the secondary vowels [ö] and [ü]. Although some of these consonants and vowels may have existed in certain dialects of Akkadian, probably none of them attained phonemic status. In at least several cases, the evidence upon which these alleged allophones rests can be reinterpreted if varying scribal practices and dialectal syllabaries are taken into account.

## 3.3 Length

As noted above, both vocalic and consonantal length are phonemic in Akkadian, although neither is regularly expressed in the script (see §2.5.4). Long vowels are considered to be of two types, morphologically long vowels, marked in transcription with a macron, and long vowels resulting from vowel contraction, marked with a circumflex. However, despite such

apparent minimal pairs as *imlāšu* “he filled it” and *imlāšu* “they (fem.) filled it”; *pānū* “face” and *pānū* “first,” the distinction between the two types of long vowels is not considered to be phonemic by most Assyriologists.

### 3.4 Syllable structure

The syllabification of Akkadian is based on three fundamental rules: (i) every syllable has only one vowel (thus vocalic clusters are divided by a syllable boundary); (ii) no syllable may begin with a vowel (with two exceptions – a word-initial vowel, and the second of two successive vowels); (iii) no syllable may begin or end with two consonants.

### 3.5 Stress

In Akkadian the stress of a given word is predictable in terms of the quantity of its constituent syllables. The fact that the position of word stress is determined by the phonological environment, and the lack of minimal pairs distinguished by stress show that word stress was nonphonemic. The rule for determining stress may be stated as follows: stress falls on the ultimate syllable if it is closed and has a long vowel (either macron or circumflex) or if it is open and has a circumflex vowel: for example, [i'dūk] “s/he killed”; [ib'nū] “they built.” Otherwise, the stress falls on the rightmost nonfinal syllable either closed with a consonant or open with long vowel (macron or circumflex): for example, [i'parras] “s/he will cut”; [mārum] “son.” If neither condition is met, the stress falls on the first syllable: for example, [nadin] “is given”; [ilū] “gods.”

### 3.6 Intonation

Lexical or grammatical tones are not attested for any Semitic language, modern or ancient. However, suprasegmental intonation no doubt conveyed important syntactic information in Akkadian, although its role is scarcely known. Occasionally, a plene writing of the type CV+V is used to express intonation. The rising pitch of a yes-or-no question may be indicated by a plene writing of the final syllable: for example, *in-na-ak-su-ú* ([innaksú]) “Are they cut?”; *ga-me-e-er* ([ga'mér]) “Is it complete?” With penultimate syllables, a plene writing often indicates intonation of emphasis: for example, *te-e-er-ra* [t'erra] “Return (pl.) it!!”; *ne-e-si* ([nèsi]) “It is (indeed) distant!”

### 3.7 Phonotactics

According to some scholars, Akkadian exhibits free variation between [V:C] and [VC:], for example *ḫiṭtu* and *ḫītu* “fault,” except with I- and II-weak verbs where the place of length is functional and serves to distinguish two morphs – thus, *ippuš* “he makes” against *īpuš* “he made”; *ikūnū* “they (masc. pl.) became firm” against *ikunnū* “they (masc. pl.) will become firm” (Reiner 1966:45). However, the occurrence of minimal pairs such as *mārum* “son” versus *marrum* “shovel,” and *šārum* “wind” versus *šarrum* “king” suggests that the distinction between /V:C/ and /VC:/ is in fact contrastive. Variations such as *ḫītu* and *ḫiṭtu* “fault” probably reflect dialectal resolutions of the contact of glottalized (emphatic) consonants and a glottal stop: \*ḫiṭ'um = [xiṭ'-ʔum] > [xiṭ'um] or [xitt'um].

Clusters of three or more consonants are not permissible in Akkadian. Any vowel or consonant except [y] may occur in word-initial position, while consonant clusters (including long consonants of the type C<sub>1</sub>C<sub>1</sub>) and vocalic clusters are not permissible in word-initial

position. Loanwords with initial consonantal clusters are resolved with the insertion of anaptyctic vowels. A glottal stop may have been permissible in word-initial position, but the ambiguous rendering of this consonant prevents certainty in this matter. The glide [w] was only retained in word-initial position in the Old Akkadian, Old Babylonian, and Old Assyrian periods; (see §3.8.1, 5, 6). In word-medial position any vowel, long or short, or cluster of no more than two consonants (of either the  $C_1C_1$  or the  $C_1C_2$  variety) is allowed. The nonoccurrence of certain consonant clusters of the type  $C_1C_2$  limits the possible combinations of this type. In word-final position, any vowel or consonant may occur; however, consonant clusters (both  $C_1C_1$  and  $C_1C_2$ ) and glides are not permitted in word-final position. When consonant clusters arise from morphological processes, they are resolved with anaptyctic vowels or, in the case of long consonants ( $C_1C_1$ ), they may be shortened.

### 3.8 Diachronic developments in Akkadian and Eblaite

#### 3.8.1 Consonantal changes

As noted above, Akkadian exhibits a significant reduction in the number of Proto-Semitic consonants. The following outline illustrates these mergers and losses along with other significant diachronic developments:

##### 3.8.1.1 Mergers

1. Proto-Semitic \* $\delta$  and \* $z$  merged as  $z$ : for example, \* $\delta ak\bar{a}rum$  >  $zak\bar{a}rum$  “to remember”; \* $\dot{u}\delta nun$  >  $uznum$  “ear”; \* $zam\bar{a}rum$  >  $zam\bar{a}rum$  “to make music”; \* $\dot{a}z\bar{a}bum$  >  $ez\bar{e}bum$  “to leave.” However, in Eblaite the syllabary indicates that these phonemes remained distinct (see 3).

2. Proto-Semitic \* $\dot{s}$ , \* $\dot{\zeta}$ , and \* $\theta$  merged as  $\dot{s}$ : for example, \* $\dot{s}ar\bar{a}h\bar{u}m$  >  $\dot{s}ar\bar{a}h\bar{u}m$  “to cry out”; \* $\dot{s}am\bar{a}dum$  >  $\dot{s}am\bar{a}dum$  “to bind”; \* $na\theta\bar{a}rum$  >  $na\dot{s}\bar{a}rum$  “to watch.” These phonemes were expressed with the same set of signs in Eblaite (see 3).

3. Proto-Semitic \* $\dot{\zeta}$ , \* $\dot{s}$  and \* $\theta$  merged as  $\dot{s}$ : for example, \* $\dot{s}apatum$  >  $\dot{s}apatum$  “lip”; \* $\dot{s}ak\bar{a}num$  >  $\dot{s}ak\bar{a}num$  “to place”; \* $\theta al\bar{a}thum$  >  $\dot{s}al\bar{a}\dot{s}um$  “three”. However, during the Old Akkadian period the reflex of \* $\theta$  was still distinct from the merged \* $\dot{\zeta}/\dot{s}$  phoneme; this is clearly demonstrated by the choice of orthographic signs used to express each phoneme. Note that the Old Akkadian phonemes represented in (4) below (and throughout this chapter) as  $\theta$  and  $\dot{s}$  are usually represented as  $\dot{s}$  and  $\dot{s}$  respectively by Assyriologists (according to scholarly convention, though the phonetic particulars of both are actually unknown):

#### (4) The sibilants of Old Akkadian and Old Babylonian

Proto-Semitic		Old Akkadian		Old Babylonian	
		Phoneme	Graphemes	Phoneme	Graphemes
* $\theta$	→	$\theta$	ŠA, ŠI, ŠU	$\dot{s}$	ŠA, ŠI, ŠU
* $\dot{\zeta}$ , * $\dot{s}$	→	$\dot{s}$	SA, SI, SU	$\dot{s}$	ŠA, ŠI, ŠU
* $s$	→	$s$		$s$	SA, SI, SU
* $\delta$ , * $z$	→	$z$	ZA, ZI, ZU	$z$	ZA, ZI, ZU
* $\dot{\zeta}$ , * $\dot{s}$ , * $\theta$	→	$\dot{s}$		$\dot{s}$	ZA, ZI, ZU

Thus, for the Old Akkadian period one encounters, for example, the spelling  $\dot{u}\text{-}\dot{s}a\text{-}ab$  for / $\dot{y}u\theta\theta ab$ / “he dwells,”  $sa\text{-}ap\text{-}ta\text{-}su$  for / $\dot{s}apt\bar{a}\text{-}\dot{s}\bar{u}$ / “his lips (dual),” and  $su\text{-}mu$  from  $\sqrt{*}\dot{s}m$

“name.” The Proto-Semitic sibilants \*s and \*š/z, while remaining distinct in Akkadian, were written with the same signs in the Old Akkadian period (e.g., *zi-ku-ru-um* < \**sukkūrum* “bolt”; *zi-kà(ga)-ar* < \**ḏikar* “male”; *i-za-mar* < \**yazammār* “he sings”), and were only distinguished graphically beginning in the Old Babylonian period.

In Eblaite PS \*ḏ and \*z remained distinct and did not merge as they did in Akkadian. Rather, the reflexes of PS \*ḏ and \*θ were written with the same set of signs; however, as discussed above, this indicates only a possible merger:

#### (5) The sibilants of Eblaite

Proto-Semitic	Eblaite graphemes
*ḏ	ŠA, ŠÈ, ŠU
*θ	ŠA, ŠÈ, ŠU
*ś, *š	SA, SI, SU
*z, *s	ZA, ZI, ZU
*š, *ś, *θ	ZA, ZI, ZU

4. As noted above, the five Proto-Semitic guttural consonants \*, \*h, \*ḥ, \*, \*ğ (ʾ–ʾs) merged as /ʔ/ and were then lost in most environments. Directly before or after consonants, the loss of /ʔ/ resulted in the compensatory lengthening of the preceding vowel: for example, \**milʾum* > *mīlum* “flood”; \**marʾum* > *mārum* “son”; \**nahrūm* > *nārum* “river”. If the lost guttural was at the beginning or end of a word, it was lost with no further change to the word: for example, \**ʾamārum* > *amārum* “to see”; \**halākum* > *alākum* “to go”; \**imlaʾ* > *imla* “he filled.” When ʾ–ʾs originally stood between vowels, its loss left those vowels contiguous and subject to the rules of vowel contraction for that particular dialect (see §3.8.2.4).

Three of the gutturals, \*h, \*, \*ğ (ʾ–ʾs), also colored neighboring *a* vowels to *e* before they were lost (only in Babylonian did the loss of these consonants additionally change *ā* to *ē*): for example, \**aḥpuš* > *ēpuš* “I did”; \**taḥpušā* > *tēpušā* “you (pl.) did”; \**ḥarāṭum* > Ass. *erāšum*, Bab. *erēšum* “to plow”; \**ʾazābum* > Ass. *ezābum*, Bab. *ezēbum* “to leave”; \**ḡaṭāyūm* > *ešūm* “to confuse” (see §3.1.3, 1 for Babylonian vowel harmony). However, the orthographic practices of the Old Akkadian period seem to indicate that the gutturals ʾ–ʾs, or at least the reflex of ʾ–ʾs, were maintained in most environments and the shift *a* > *e* did not occur consistently: for example, Old Akkadian *u-ša-rí-ib* /*yušaʾrib*/ “s/he brought in,” as opposed to Old Babylonian *ú-še-ri-ib* /*ušērib*/; *a-lí-tám* /*ʾalítam*/ “upper” (OB /*elítam*/). Similarly, in Eblaite *a* was usually retained in the vicinity of the preserved reflexes of \*, \*ḥ, and \*ğ: for example, *maʾmadu(m)* “support”; *ḡāribu(m)* “raven.”

#### 3.8.1.2 Segmental loss

5. Proto-Semitic \*w (ʾ) was lost at the end of syllables, unless followed by another \*w. The loss of \*w resulted in the compensatory lengthening of the preceding vowel: for example, \**šuwrid* > *šūrid* “send down (masc. sg.)!”; but *nuwwurum* “to brighten.”

Initial \*w was retained in the early dialects of Akkadian, but was lost by the end of the Old Babylonian period: for example, OB *warḥum* > SB/MB/NB *arḥu* “month.” However, in Middle Assyrian, and subsequently in Neo-Assyrian, initial \*wa became *u*: for example, OA *wardum*, MA/NA *urdu* “slave” (however, in Middle Assyrian/Neo-Assyrian initial *wā* shifts to *ā*: OA *wāšibum* > MA/NA *āšibu* “inhabitant”).

In the Middle Babylonian period, intervocalic *w* was written as <m> (<VmV> = /VwV/): for example, SB/MB/NB *a-me/mi-lu* for /*awīlu*/ “man” (see 18 below). On analogy with <VmV> in verbal forms such as MB *ú-ma-aš-šar* /*uwaššar*/ “he releases,” forms of the verb that originally had initial *w* (i.e., the imperative, the infinitive, and the verbal adjective)



are written with initial <m> and do not show the expected loss of *w* (e.g., MB *mu-uš-šu-ru* /wuššuru/ “to release”). In Middle Assyrian, however, the sequence \*VwV often appears as VV, as in *aṣlu* < *awilum* “man.” In Neo-Assyrian, if an intervocalic <m> is secondary, that is, originating from an etymological *w*, then it often shifts to written <b>: for example, OA *awātum*, NA *abutu* “word” (with Assyrian vowel harmony; see §3.8.2.3).

In Eblaite \**w* was preserved word-initially, as in the early dialects of Akkadian: for example, *wa-ba-lu* /wabālu(m)/ “to bring.” Although \**w* appears also to be preserved intervocalically (e.g., *ma-wu* /mawū/ “water”), some writings may indicate a loss of \**w* in that position: thus, *ga-nu-um* beside *ga-nu-wu* for /ganu(w)u(m)/ “reed.”

6. Proto-Semitic \**y* (ʔ) was also lost at the end of syllables, unless – as was the case with \**w* – followed by another \**y*. Similarly, the loss of \**y* resulted in the compensatory lengthening of the preceding vowel: for example, \**yupahḥar* > *upahḥar* “he gathers”; \**rabiytum* > *rabītum* “great”; but *dayyānum* “judge.”

Initial \**y* was also lost, but only after the change of \**ya-* to \**yi-* (probably preserved in Old Akkadian, in view of plene-writings of the type *i-ik-mi* /yikmi/ “he captured”): for example, \**yaprus* > (OAkk.) *yiprus* > (OB/OA) *iprus* “he cut.” A single intervocalic \**y* is preserved only in the possessive suffix -*ya* (e.g., *bēliya* [gen.] “my lord”), but after long vowels the *y* of the possessive suffix is often written as <ʾ> (e.g., *bēlū-ʾa* “my lords” [nom.]; *šēpē-ʾa* “my feet” [acc.-gen.]).

In Eblaite, word-initial \**ya* often shifted to *y(i)* as in Akkadian: for example, *i-sa-lum* < \**yašarum* “straight.” But note the apparent biforms *i-me-tum* and *a-me-tum* < \**yamintum* “right (side)” (fem. sg.). As the writing system had no direct means of expressing /y/, little else can be said concerning its behavior in other environments; thus, for intervocalic *y* note the spelling *ba-ga-um* for /bakā(y)um/ “to weep.”

7. Beginning with the late Old Babylonian and Old Assyrian periods, *mimation* (the occurrence of -*m*) was lost when word final (retained only when followed by -*ma* and the pronominal suffixes). However, often the spelling does not reflect this change; as a result some CV*m* signs assume CV values; thus, OB *šarrātum* “queens” (nom. pl.), written <šar-*ra-tum*>, becomes MB *šarrātu*, but is still written as in Old Babylonian, as a historical spelling. To reflect the sound change in transliteration, the final sign may be read *tu*<sub>4</sub>, thus *šar-ra-tu*<sub>4</sub>.

8. In Eblaite, *l* and *r* are apparently lost in the spellings of some words: for example, *la-i-mu* and *a-i-mu* for /laḥimu(m)/? < √\**lḥim* “to press together”; *sa-ʾà-lum* and *sa-ʾà-a-um* for /šah(a)rum/ “new moon”?; *ba-a-ḥu-um* for /palāxum/ “to fear.” These spellings possibly suggest that *l* and *r* may be weakened to ʔ, *y*, or perhaps Ø. As the second set of examples indicates, *r* may be written with the set of syllabograms for <l>: for example, *ba-ga-lum* /bakʾ(a)rum/ “cow”; *bu-ga-ru*<sub>12</sub> and *bu-ga-lu* for /buk(a)ru(m)/ “first-born.” However, the converse – writing *l* with <r> signs – is not attested (see §3.1.2).

### 3.8.1.3 Dissimilation

9. By Barth’s Law of Dissimilation, initial \**m* (except for \**#mu-*) dissimilated to *n* in forms containing the other labials, namely, *p*, *b*, *m*: for example, \**markabtum* > *narkabtum* “chariot”; \**mamšarum* > *namšarum* “sword.”

In Eblaite, however, Barth’s Law does not appear to apply: thus, *má-ma-du* /maʕmadu(m)/ = Akk. *nēmedum* “support.”

10. By Geers’ Law of Dissimilation, in roots originally containing two Proto-Semitic emphatic consonants, one of the emphatics dissimilated to its nonemphatic, voiceless counterpart: (i) \**ṭ* became *t* in forms that also contained *q* or *ṣ* (< PS \**ṣ*, \**ṣ̣*, or \**θ*): for example, \**qaṭārum* > *qatārum* “to smoke”; \**ṣabāṭum* > *ṣabātum* “to seize”; (ii) in forms with both *q* and *ṣ*, the first dissimilated,

## (6) {q, ʃ} &gt; {k, s} / – (X) {s, q}

for example, \**qašārum* > *kašārum* “to knot”; \**šayāqum* > *siāqum* “to be narrow.”

11. The first member of a voiced geminate cluster, particularly in the case of dentals, often undergoes nasalization. This phenomenon is characteristic of the post-Old Babylonian periods, but is occasionally attested for some Old Babylonian dialects and sporadically for Old Akkadian. Examples are *ambi* beside *abbi* (< \**anbi*) “I called”; *inandin* and *inamdin* beside *inaddin* “s/he will give”; *inanziq* and *inamziq* beside *inazziq* “he becomes vexed”; *nangāru* and *namgāru* beside *naggāru* “carpenter.” In Middle Babylonian and later, the assimilatory change *dn* > *nn* often took place in forms of the verb *nadānu* “to give.”

## 3.8.1.4 Assimilation

12. Proto-Semitic \**n* assimilated to a following consonant: for example, \**tanθur* > *taššur* “you (masc. sg.) guarded”; \**libintum* > *libittum* “brick.” Exceptions occur when \**n* is the second root consonant: \**anzum* > *enzum* “female goat.”

In Eblaite, \**n* assimilates to the feminine ending: for example, *li-bi/ba-tu/lib{i,a}ttum/* < √\**lbn* “a kind of brick.” In other environments, however, it remains unchanged. Additionally, \**m* may also assimilate to a following consonant: for example, *si-tum* /šittum/ (Akk. *šimtu*) “sign, color”; *ti-da-ḫu-ru<sub>12</sub>* probably /tittaxrū/ < \**timtaḫrū* “they approached.”

13. The dentals *d* and *t* assimilate completely to the feminine ending *-t*: for example, \**paqidtum* > *paqittum* “entrusted”; \**balittum* > *balittum* “alive.” The sibilants *s*, *š*, and *z* become *š* before the feminine ending *-t*: for example, \**paristum* > *parištum* “separated”; \**maruštum* > *maruštum* “sick.” Beginning with the Middle Babylonian and Middle Assyrian periods, this *š* becomes *l*, not only before the feminine ending *-t*, but regularly before dentals *d*, *t*, *ṭ* and sibilants *s*, *š*, *z* as well (i.e., *š* + {*D*, *Z*} > *l* + {*D*, *Z*}): for example, OB *mazzaztum* > *mazzaltu* “place”; OB *aštapar* > *altapar* “I have sent”; OB *uṣziz* > *ulziz* “he caused to stand.” One notable exception to this change is found in the Middle Assyrian preposition *ištu* “from” (cf. Middle Babylonian *ultu*). In Neo-Assyrian the *lt* that originates from *št* often changes to *ss*: for example, MA/MB *altakan* > NA *assakan* “I have placed.” Occasionally, *lt* becomes *ss* even when *lt* was not originally *št*: for example, OA *ilteqe* > NA *isseqe* “he has taken.” Rarely, *d* assimilates to an immediately following *š*, as in *eššu* “new” < \**edšu*.

In Eblaite, as in the later Akkadian dialects, sibilants may shift to *l* before dentals: for example, *dal-da-i-bū* /taltahhibu(m)/ < \*/taštaḥhibu(m)/, from the verb *šaḥābu(m)* (cf. Arabic *saḥaba* “to withdraw, take away”).

14. The infixed *-t-* of the Perfect and of the Gt and Dt stems assimilated completely when immediately before or after the dentals, *d* and *t*, and the sibilants, *s*, *š*, *z*; and when immediately before *š*, but not when after *š* (i.e., {*D*, *Z*} + *t* > {*DD*, *ZZ*}; while *t* + {*D*, *Z*, *š*} > {*DD*, *ZZ*, *šš*): for example, \**idtamiq* > *iddamiq* “it has improved”; \**ḫitdulum* > *ḫiddulum* “to become knotted”; \**istaḫur* > *issaḫur* “he has turned”; \**pitšušum* > *piššušum* “to anoint oneself”; but note *ištakan* “he has placed.” The infixed *-t-* became *d* when immediately after *g*: \**igtamar* > *igdamar* “he has finished.” In the Middle Assyrian and Neo-Assyrian periods, the infix *-t-* became *ṭ* after *q*: for example, *iqṭibi* > *iqṭibi* “he has said.”

15. The sequence of stem-final dental or sibilant (*d*, *t*, *ṭ*, *s*, *š*, *z*, *ṣ*) plus the *š* of the third-person pronominal suffixes yielded *ss* (i.e., {*D*, *Z*, *š*} + *š* > *ss*): for example, \**ikšud-šu* > *ikšussu* “he reached him (masc.)”; \**ikkis-šu* > *ikkissu* “he cut it (masc.) off.” However, there are qualifications to this assimilatory change for the Old Akkadian and Old Assyrian periods regarding stem-final *š* (< PS \**š*, \**ṣ*, \**θ*). In the Old Akkadian period, when the reflexes of PS \**š*/*ṣ* and PS \**θ* were still distinguished, the result of a stem-final *θ* or *š* plus the *š* of the pronominal suffix was *šš*: for example, Old Akkadian *iqiššum* < *iqiš-šum* “he gave to him”;



*erēššunu* < *erēθ-šunu* “their cultivation.” Similarly, in Old Assyrian, after the merger of PS \*š/š and PS \*θ > š, the sequence šš was preserved: for example, Old Assyrian *lubūššunu* < \**lubūš-šunu* “their clothing”; Old Assyrian *epuššum* < \**epuš-šum* “do for him!”

16. In Middle Assyrian and Neo-Assyrian, š before *b* and *p* shifted to *s* (i.e., š > *s* / — [bilabial stop]): for example, OA, OB *wašbat* > MA *usbat* (MB *ašbat*), NA *usbat/uspat* (for *wa* > *u* see 5 above).

17. The bilabial stops *b* and *p* assimilated to the labial nasal *m* at morpheme boundary (i.e., {*p*, *b*} + *m* > *mm*). Aside from relatively rare sandhi writings of the type *ḥiṣimmātim* < *ḥiṣib mātīm* “abundance of the land,” these phonemes are only juxtaposed when the enclitic particle *-ma* follows the morpheme boundary: for example, *irkamma* < *irkab-ma* “he rode and . . .” In Neo-Assyrian, *b* sometimes devoices after *s* (i.e., *b* > *p* / *s* —): for example, *uṣpākūni* beside *uṣbākūni* versus Babylonian *wašbāku* “(where) I dwell.”

18. From Middle Babylonian on (and sporadically already in Old Babylonian), intervocalic *m* was pronounced [w], as in [šawaš] for the god Šamaš, so that etymological *m* and etymological *w* were indistinguishable intervocalically (e.g., OB *awātum* “word” and *amātum* “female slaves” were in MB both pronounced [awātu]); as a result of this development, signs for *m* + vowel came to be used to write both *m* and *w* (see 5 above). In Middle Babylonian, Standard Babylonian, and Neo-Babylonian *m* and the infix *-t-* usually yield the sequence *nd* (less frequently *md*), showing a reciprocal assimilation in voicing and place of articulation: *m* + *t* > *nd*, as in *imtala* > *indala* “he has become full.” Just as OB *wuššurum* is analogously replaced by MB/SB *muššuru* (see 5 above), the Old Babylonian perfect *ūtaššer*, “he has released,” is replaced by *umtaššer*, appearing as MB/SB/NB *undaššer* or *undeššer* (see §3.1.3.1).

From the Middle Babylonian period onwards, *m* that is part of the root may also shift to *n* before the other dentals, *d* and *t*, as well as before š, ṣ, *q*, *k*, as in *anši* < *amši* “I forgot”; *enqu* < *emqu* “wise.” In Neo-Assyrian this secondary *n* completely assimilated to the following consonant: for example, *attaḥar* < \**antaḥar* < *amtaḥar* “I received.” Moreover, in Neo-Assyrian an etymological intervocalic *m* was often written as <’> or deleted from the script entirely: for example, <da’iq>, <deq> for earlier *damiq* “it is good” (< √dmq).

19. The final *m* of the ventive suffix (*-am*, *-nim*, *-m*) and of the locative adverbial *-um* assimilated completely to the consonant of a following pronominal suffix: for example, *ašpurakkum* < \**ašpuram-kum* “I sent to you (masc.)”; *šēpuššu* < \**šēpum-šu* “at his foot.”

20. The assimilation of *l* and *r* to a following consonant is rare but attested: for example, *kilattān* < \**kilaltān* (fem.) “both”; *qarnum/qannum* “horn.” In Neo-Babylonian and Late Babylonian, *r* very often becomes š before *t* and *k*: for example, *lištappud* < *lirtappud* “may he always run”; *šipištu* < *šipirtu* “message.”

21. In the Middle Babylonian and Neo-Babylonian periods the voiceless velar *k* often becomes voiced when immediately following the nasals *m* and *n* (i.e., *k* > *g* / {*m*, *n*} —): for example, *kankum* > *kangu* “sealed document”; *ṭēmka* > *ṭēnga* “your report” (for *m* > *n* see 18 above).

### 3.8.1.5 Metathesis

22. Metathesis of consonants occurs in unprefix verbal forms (i.e., infinitive, imperative, verbal adjective) with an infixed *-t-*, specifically, in forms of the Gt, Gtn, and adjectives of the form *pitrās*, when the first radical is *z*, *s*, ṣ, *d* (and in Old Assyrian š as well); in other words, {*Z*, *D*} + *Vt* > *tV*{*Z*, *D*}: for example, *tišbutum* (Gt inf.) < \**šitbutum* “to grasp one another”; *tizqārum* (*pitrās* adj.) < \**zitqārum* “prominent”; Old Assyrian *tišammeā* alongside *šitammeā* (Gtn pl. imperative) “listen continually!”

### 3.8.2 Vocalic changes

The Proto-Semitic vowels *\*a*, *\*i*, and *\*u* were subject to various developments in Akkadian (for secondary /e/, including Babylonian vowel harmony, see §3.1.3.1; for possible /o/, see §3.1.3.2).

#### 3.8.2.1 Loss of final vowels and resolution of consonant clusters

In Akkadian, short final PS *\*a* and *\*u* were lost; final *\*i* was retained in Old Akkadian, but was also lost thereafter. Exceptions to this change are the prepositions *ana* “to,” *ina* “in,” and the subordination marker *u*. The final consonant clusters created by the loss of these vowels were resolved by the insertion of an anaptyctic vowel. In Babylonian this vowel was of the same quality as the preceding vowel, while in Assyrian an *a* was always inserted, regardless of the preceding vowel: for example, in the bound forms of *rigmum* “noise,” singular nominative *\*rigmu*, accusative *\*rigma* both > *\*rigm* > Ass. *rigam*, Bab. *rigim*; compare the bound form singular genitive *rigmi* (in Old Akkadian) > *\*rigm* (post-Old Akkadian) > Ass. *rigam*, Bab. *rigim* as well.

#### 3.8.2.2 Vowel syncope

In all dialects of Akkadian, the last of two or more non-final short vowels in open syllables was syncope. Before the consonants *l* and *r* vowel syncope was optional: thus, *\*rapašatum* > *rapaštum* (fem. sg.) “wide”; *\*akalum* > *akalum* beside *aklum* “food.” Syncope did not take place at the end of a word (where two successive open syllables were allowed), before vowels, before certain pronominal suffixes, or in some Sumerian loanwords.

#### 3.8.2.3 Assyrian vowel harmony

In the Assyrian dialects a short *a* vowel in an open, unaccented syllable assimilated to the vowel in the following syllable: for example, OA *šarritim* (gen.) versus OB *šarratim* “queen”; Ass. *išbutū* versus Bab. *išbatū* “they (masc.) seized.” Old Assyrian was characterized by an additional rule which stipulated that in the N stem (see §4.2.2), preterite *i* of the second syllable (which resulted from regular Assyrian vowel harmony) remained even though the influencing vowel was syncope: for example, OA *iššiknū* < *\*iššikinū* versus MA *iššaknū* “they were placed.” This rule was not applicable to the Gt stem: for example, OA *ētitīq* “he marched away” beside OA *ētatqū* “they (masc.) marched away.” In Neo-Assyrian, vowel harmony sometimes took place across two consonants (e.g., *idubbub* < *idabbub* “he is speaking”) or when the influencing vowel was lost (e.g., *ittuqtū* < *\*imtaqtū* “they (masc.) have fallen”). For Babylonian vowel harmony see §3.1.3.1, 1.

#### 3.8.2.4 Vowel contraction

In Old Akkadian and during most of the history of Assyrian, vowels that became contiguous with the loss of the Proto-Semitic gutturals *\*ʾ*, *\*h*, *\*ḥ*, *\*ʿ*, *\*ğ* or the glides *\*w* and *\*y* did not contract. In Babylonian, however, these contiguous vowels contracted to an ultralong vowel (marked in transcription by a circumflex) of the quality of the original second vowel. There are two exceptions to this rule: (i) the regular contraction *ā + ī* > *ē* for all Babylonian dialects; and (ii) the sequences *ī + ā* and *ē + ā* do not contract in Babylonian until the end of the Old Babylonian period (except in the northern Old Babylonian dialect exemplified by the many texts from the site of Mari, in which {*i*, *e*} + *ā* > *ē*): for example, Ass. *ibniū* = Bab. *ibnū* “they (masc.) built” (< *\*ibniyū*); Ass. *banāim* = Bab. *banēm* (gen.) “to build” (< *\*banāyim*); Ass. *zakuim* = Bab. *zakīm* “clear” (masc. gen.) (< *\*zakuwim*); Ass. *kalaum* = Bab. *kalūm* “entirety, whole” (< *\*kalaʾum*). Only in the Neo-Assyrian period does Assyrian contract

adjacent vowels:  $a + \{i, e\} > \hat{e}$ ; *ia*, *iu* and *ua* do not contract at the end of a word; however, if one or more syllables follow, with the exception of *ia*, contraction usually takes place (*ia* and *ua* do not contract in verbs II-<sup>3</sup>).

### 3.8.2.5 Contraction of diphthongs

Before a consonant, the Proto-Semitic diphthong *\*aw* contracted to *ū* (for a possible *\*aw* > *ō*, see §3.1.3.2), while *\*ay* contracted to *ī* in Babylonian and *ē* in Assyrian: for example, *\*mawtum* > *mūtum* “death”; *\*baytum* > Bab. *bītum*, Ass. *bētum* “house.” Note that in the case of the sequences *\*aww* and *\*ayy* (in effect, when a geminated glide occurs), contraction did not take place (cf. §3.8.1.2, 5, 6); however, in middle weak verbs (i.e., verbs in which the second radical was originally *\*w*, *\*y*, or one of the five Proto-Semitic guttural consonants *\*ʔ*, *\*h*, *\*ḥ*, *\*ʕ*, *\*ğ*) the following developments may have occurred: *\*yadayyan* > *\*yadīyan* > *idīan* “he judges”; *\*yadawwak* > *\*yadūwak* > Ass. *idū(w)ak*, Bab. *idāk* “he kills.”

In Eblaite, the Proto-Semitic diphthongs *\*aw* and *\*ay* are normally preserved, although the writing system has no means of representing them directly. The diphthong *aw* is expressed with either *-a* or *-a-wa*: for example, *a-mu* and *a-wa-mu* for /yawmū/ “days.” The diphthong *ay* is sometimes written with an extra A sign, possibly in an attempt to express the second element: for example, *ʾa<sub>x</sub>(NI)-a-la-nu* for /ʾaylānu(m)/ a kind of tree; but more often no attempt is made to express the *-y* of the diphthong: for example, *ʾa<sub>x</sub>(NI)-la-nu-um*; *ba-du* for /baytu(m)/ “house”; *ba-nu* for /baynu(m)/ “tamarisk.”

### 3.8.2.6 Simplification of triphthongs

Akkadian simplified the Proto-Semitic triphthongs *\*awi* and *\*ayi* to *ī* in Babylonian and *ē* in Assyrian: for example, *\*kawin* > Bab. *kīn*, Ass. *kēn* “he is true”; *\*šayim* > Bab. *šīm*, Ass. *šēm* “it is established.” Similarly, *\*āya* (and *\*āwa*?) was reduced to *ā*, for example, *\*tāyab* > *tāb* “he is good.”

## 4. MORPHOLOGY

As with the other Semitic languages, Akkadian is characterized by a fusional or inflecting morphology based, with the exception of primitive nouns and the pronominal system, on a consonantal root structure of three consonants or *radicals* (C<sub>1</sub>C<sub>2</sub>C<sub>3</sub> or PRS, after the paradigmatic verb *parāsum* “to cut”). Morphological information is conveyed by vowel patterns, various affixes, and other modifications such as consonant doubling. As noted above, Akkadian morphology is historically quite conservative.

Although Eblaite morphology has not been studied extensively (and the problems encountered in the writing system may prevent such a study), it appears to be remarkably close to that of Akkadian. The only major morphological deviations appear to be the presence of infinitives with prefixed *t-*, which represent an innovation within Eblaite, and third masculine plural verbal forms with *ti-* prefix (see §4.2.1, §4.2.11.3).

### 4.1 Nominal morphology

The Akkadian noun is morphologically marked for case (nominative, accusative, and genitive), gender (masculine and feminine), and number (singular, dual, and plural). There is no dative case for Akkadian nominals; instead indirect objects are expressed by nouns in the genitive governed by prepositions. Certain transitive verbs, however, may govern two accusative nouns or pronouns, one of which represents the indirect object. The accusative case

may also be used to fulfill a wide range of adverbial functions. The dual is only productive for the Old Akkadian period, after which it is generally confined to nouns referring to parts of the body and other natural pairs. Additionally, the Akkadian nominal can assume four possible forms or “states”: (i) the free form or *declined* state; (ii) the *construct* or bound form; (iii) the *absolute* form or state; and (iv) the *predicative* form or the predicative construction (also referred to as the stative in some grammars).

#### 4.1.1 Declined form

Below, the Old Babylonian paradigm for the declined form is given for the noun *ilum* “god” (*iltum* “goddess”):

##### (7) Akkadian noun declension (Old Babylonian)

		<i>Masculine</i>	<i>Feminine</i>
<i>Singular</i>	<i>Nominative</i>	ilum	iltum
	<i>Genitive</i>	ilim	iltim
	<i>Accusative</i>	ilam	iltam
<i>Dual</i>	<i>Nominative</i>	ilān	iltān
	<i>Genitive-accusative</i>	ilīn	iltīn
<i>Plural</i>	<i>Nominative</i>	ilū	ilātum
	<i>Genitive-accusative</i>	ilī	ilātim

The short vowels *u*, *i*, and *a*, which mark the nominative, genitive, and accusative, respectively, are inherited from the Proto-Semitic case system. In the dual and plural, the genitive and accusative forms collapse into a single form marked with *i*, often referred to as the oblique case. The masculine singular and the feminine singular and plural forms exhibit a final *-m*, mimation, which, as noted above (see §3.8.1.2, 7), disappeared after the Old Babylonian and Old Assyrian periods (in Assyrian the loss of mimation resulted in forms with a final *i* changing to *e*; see §3.1.3.1). Dual forms exhibit a final *-n*, *nunation*, which is similarly lost after the Old Babylonian and Old Assyrian periods.

In the later dialects of Akkadian the distinctions in the case system were blurred. The loss of mimation (and nunation) in the Middle Babylonian and Middle Assyrian dialects served to obscure the case and number of many nouns, given that the writing system did not generally distinguish vowel length. Moreover, in the Neo-Babylonian, Neo-Assyrian, and Standard Babylonian dialects, both the nominative and accusative cases are written with *-u* (for example, the nom./acc. sg. *šarru*). And in the Standard Babylonian and Neo-Assyrian dialects, *-ī/ē* is used not only for the oblique plural, but for the nominative as well. In Late Babylonian and to a lesser degree in Neo-Babylonian, the loss of final short vowels often resulted in case endings being dropped completely.

The nominal declension of Eblaite is identical to that of Akkadian. As with the early dialects of Akkadian, Eblaite exhibits mimation although it is not regularly indicated in the script.

There is no definite or indefinite article in Akkadian or Eblaite. Thus, a word such as *šarrum* (Akk.) may represent “king,” “a king,” or “the king” depending on context.

##### 4.1.1.1 Gender

Most unmarked substantives are masculine. However, perhaps 15 percent are either feminine (e.g., *ḥarrānum* “road”) or both masculine and feminine (e.g., *urḫum* “road”). Additionally,

a significant number of substantives (perhaps 30 percent) are masculine in the singular, but are either feminine in the plural (e.g., *eqlum* “field,” pl. *eqlētum*) or have both masculine and feminine plurals (e.g., *kunukkum* “seal,” pl. *kunukkū* or *kunukkātum*). Feminine singular substantives are marked with the allomorphs *-t* and *-at*, the occurrence of which is based solely on phonological grounds: *-at* is employed when the base ends in two consonants, as in *šarratum* (nom.) “queen”; otherwise *-t* is employed, as in *bēltum* “lady” (< *bēlum* “lord”). Feminine plural nouns are universally marked with *-āt*.

#### 4.1.1.2 Number

As the above paradigm indicates, external plurals are standard for Akkadian. However, there may be remnants of internal plurals in the language as exemplified by a relatively small number of nouns: for example, *šuḥārū* “boys, servants” (< \**šuḥarā’u*); plurals with double middle radical, as in *abbū* “fathers” (sg. *abum*); *aḥḥū* “brothers” (sg. *aḥum*); *aḥḥātum* “sisters” (sg. *aḥātum*); *iṣṣū* “trees” (sg. *iṣum*); note also *arrakūtum*, possibly a plural of *arkum* “long.”

#### 4.1.1.3 The terminative and locative endings

In addition to the nominative, genitive, and accusative case endings, nouns may take two other endings. The ending *-iš* has a terminative or locative function, as in *ilīš* “for the god,” *qātiššu* “to his hand” (*qātum* “hand”); *-um* (identical to the nominative case) is used with a locative function: for example, *libbum* “in the heart”; *šaptukki* “on your (fem.) lips” (< *šaptum* + *ki*). Both endings are productive only in Old Akkadian, after which they are restricted to certain nouns and compounds, particularly in the Old Babylonian literary and Standard Babylonian dialects. However, *-iš* and *-um* continue to be used in the formation of adverbs: for example, *rabiš* “greatly”; *apputum* “please” (see §4.3.2). It is not clear whether the terminative and the locative represent vestigial case markers or adverbial formatives. Eblaite also displays the terminative ending *-iš* and the locative ending *-um*: for example, *ga-tum-ma ga-ti-iš* /*k’ātum-ma k’ātiš*/ “from hand to hand.”

#### 4.1.1.4 The particularizing suffix *-ān*

The particularizing suffix *-ān* may appear directly after the root and immediately before the case endings. It serves to identify a particular individual or object from among a general group or class: for example, *nādinānum* “the seller in question, the particular seller”; *šarrāqānum* “the thief in question, that particular thief.” In some instances it may acquire a specific meaning: for example, *rabiānum* “mayor” (lit. “the particular great man”). With plural substantives it indicates a specific group: for example, *ilānū* “(a certain group of) gods”; *šarrānū* “(a certain group of) kings.” In the post-Old Babylonian and Old Assyrian dialects (but attested already in the Old Babylonian and Old Assyrian periods) the words *ālum* “city,” *ilum* “god,” and *šarrum* “king” regularly exhibit the particularizing suffix in their plurals – in such cases, the particularizing meaning of *-ān* has been lost.

#### 4.1.1.5 The abstract suffix *-ūt*

The suffix *-ūt* (Ass. *-utt*) is added to the bases of nouns to form abstracts. The suffix precedes the expected case ending (with feminine forms the marker *-(a)t* is usually dropped): for example, *šarrūtum* “kingship” (< *šarrum* “king”); *ilūtum* “divinity” (< *ilum* “god”); *awilūtum*

“humanity” (< *awīlum* “man”); *dannūtum* “strength” (< *dannum* “strong”); *aššūtum* “wifehood” (< *aššatum* “wife”). Although the Babylonian suffix *-ūt* is identical to the masculine plural ending of adjectives, nouns formed with this suffix are grammatically feminine singular.

#### 4.1.1.6 Noun forms

As in the other Semitic languages, many Akkadian nouns are derived from verbal bases. Nouns that are not associated with verbal bases are referred to as *isolated*. Many such nouns inherited from Proto-Semitic display fixed bases in the singular consisting of two or three consonants with one or two intervening vowels: for example, \**ʿayn*- “eye”; \**kalb*- “dog”; \**šinn*- “tooth”; \**paraš*- “horse”; \**ʿiṣ*- “wood”; \**ʾil*- “god.” However, most native Akkadian nouns are derived from verbal bases, and some – though not most – are classifiable with regard to a predictable pattern and associated meaning. The following list of noun patterns is not in any way exhaustive, but merely illustrates the phenomenon (verbal noun and adjective patterns are not included); the root *p-r-s* (< *parāsum* “to cut”) is used here to represent *C<sub>1</sub>C<sub>2</sub>C<sub>3</sub>*, following the traditional presentation of the Akkadian root.

1. *parrās nouns*. Often designate professions: for example, *dayyālum* “spy” (< *dālum* “to wander”); *errēšum* “farmer” (< *erēšum* “to plow”); *šarrāqum* “thief” (< *šarāqum* “to steal”).
2. *mapras(t) nouns*. Denote place, time, or instrument (such as tools and vehicles): for example, *maškanum* “threshing floor” (< *šakānum* “to place”); *mūšabum* “dwelling” (< *wašābum* “to dwell”); *narkabtum* “chariot” (< *rakābum* “to ride”; for *ma-* > *na-* see §3.8.1.3, 9, Barth’s Law).
3. *purussā’ nouns*. Often denote legal activities: for example, *purussūm* “legal decision” (< *parāsum* “to cut, to decide”); *rugummūm* “legal claim” (< *ragāmum* “to complain”); *uzubbūm* “divorce(-payment)” (< *ezēbum* “to leave”).
4. *pirs (fem. pirist) nouns*. Often associated with a passive nuance: for example, *isiḫtum* “assignment” (< *esēḫum* “to assign”); *šiprum* “message, mission” (< *šapārum* “to send”); *šitrum* “(piece of) writing” (< *šaṭārum* “to write”).

An *ipris* form is attested, a less common variant of *pirs*: for example, *ikribum* “prayer” (< *karābum* “to pray”); *ipṭerū* “ransom” (< *paṭārum* “to loosen, to remove”).

5. *purs (fem. purust) nouns*. Often abstracts of adjectival roots: for example, *šulmum* “well-being” (< *šalāmum* “to be well/whole”); *rupšum* “width” (< *rapāšum* “to be wide”).
6. *taprās nouns*. Associated with the Gt stem (see §4.2.5): for example, *tamḥārum* “battle” (< *mithurum* “to oppose one another”); *tāḥāzum* “battle, combat” (< *athuzum* “to seize one another”).
7. *taprīs and taprist nouns*. Associated with the D stem (see §4.2.3): for example, *talmūdum* “student” (< *lummudum* “to teach”); *tarbītum* “offspring” (< *rubbūm* “to rear”).
8. *Forms with a reduplicated radical*. Especially associated with insect names: for example, *zuqaqīpum* “scorpion”; *kulbābum* “bee”; *adammūmum* “wasp.”

#### 4.1.2 The bound form

Genitival relationships are most commonly expressed by juxtaposing the governing and governed noun. In such constructions, the former occurs in the bound form and the governed noun in the genitive case. The bound form is also used for nouns that syntactically govern



pronominal suffixes and, less often, introduce relative clauses (see §5.12.1): for example, *bīt awīlim* “the house of the man” (< *bītum* “house”); *šarrat mātīm* “queen of the land” (< *šarratum* “queen”); *kalabšu* “his dog” (nom./acc.).

The bound form of a given noun may be described as the shortest form of the noun phonetically possible. As such, the bound form lacks mimation/nunation and case vowels – with the exception of the final *i* of the genitive singular and oblique feminine plural forms which was retained during the Old Akkadian period, but lost thereafter (see §3.8.2.1). However, the bound form in all dialects retained, and probably lengthened, the genitive singular *i* when it occurred before pronominal suffixes: for example, *bēlišu* “his lord (gen.)” versus *bēlšu* (nom.-acc.). Similarly, plural and dual forms retained their case vowels before pronominal suffixes: for example, *mārūka* “your (masc.) sons (nom.)”; *uznāšu* “his (two) ears (nom.)”. The loss of case vowels in all other instances, that is, all forms before nouns, and nominative and accusative singular forms before suffixes, often resulted in an impermissible word-final consonant cluster (see §3.7). Such impermissible clusters were resolved in a variety of ways according to the morphological shape of the base (see §3.8.2.1). One-syllable nouns ending in  $-C_1C_1$  either added a short *-i* (e.g., *libbi* < *libbum* “heart”) or dropped the second consonant (e.g., *ekal* < *ekallum* “palace”). Bases ending in  $-C_1C_2$ , where  $C_2 \neq t$ , insert an anatyptic vowel of the same quality as that of the preceding syllable (in Assyrian, *a* is inserted regardless of the preceding vowel): for example, *puḥur* (Ass. *puḥar*) < *puḥrum* “assembly.” Bases ending in  $-Ct$  either insert an *a*-vowel (e.g., *mārat* < *mārtum* “daughter”) or add an *-i* (e.g., *qīšti* < *qīštum* “gift”). Two-syllable nouns ending in  $-Ct$  exhibit bound forms with an additional *i* vowel: thus, *napišti* < *napištum* “life” (exceptions being feminine participles, which insert *a*: *māḥirat* < *māḥirtum* “rival”). Before suffixes, singular bases ending in  $-C_1C_2$  universally insert *a* before the suffix in the nominative and accusative; as noted above, in the genitive the *i* case vowel is maintained before suffixes: for example, *napištaka* “your (fem.) life” (nom.-acc.) versus *napištika* (gen.). The syntax of the genitive construction is discussed in section 5.4.

The morphological shape of the bound form in Eblaite cannot be determined with any certainty given the ambiguities posed by the writing system. For example, there are writings with and without final vowel (often *-u*, perhaps representing the nominative case): for example, *ḥa-za-nu* GN “mayor of GN”; *ma-lik* GN “king of GN,” where Akkadian would omit the final vowel in both cases. Additionally, in what may represent fixed or logographic writings (see §2.6), there are forms such as *ma-za-lum-sù* where the reading may be either /maθθārūšu/ or /maθθāršu/ “his guard.”

### 4.1.3 The predicative form

Substantives and adjectives may enter into a predicative construction whereby a specific set of pronominal suffixes are added directly to the indeclinable base. The resulting construction constitutes a verbless clause, and as such, predicative forms are not marked for tense. The pronominal suffix serves as the subject, and the substantive or adjective as the predicate: *marṣāku* “I (*-āku*) am sick (*marṣ-*)”; *rabiānu* (post-Old Babylonian *rabānu*) “we (*-ānu*) are great (*rabi-*)”; *bēlū* “they (*-ū*, masc.) are lords (*bēl-*).” The third masculine singular of the predicative form is marked with  $-\emptyset$ : *šar* “He ( $-\emptyset$ ) is king (*šar-*).” The nonfinal *a*-vowels of the other endings are subject to Babylonian vowel harmony whereby *a* shifts to an *e* vowel when there is an *e* elsewhere in the word (see §3.1.3.1, 1). Substantives may appear in the predicative construction only when not followed by any modifiers, including adjectives, genitive nouns, pronominal suffixes, relative clauses, and the emphatic particle *-ma*.

The Old Babylonian paradigm of the predicative construction for the substantive *šarrum* (for example, *šarrāku* “I am/was/will be king”) and the adjective *ezbum* (for example, *ezbēku* “I am/was/will be abandoned”) appears below:

(8) Old Babylonian predicative form paradigm

		Suffix	<i>šarrum</i>	<i>ezbum</i>
Singular	1st com.	-āku	šarrāku	ezbēku
	2nd masc.	-āta	šarrāta	ezbēta
	2nd fem.	-āti	šarrāti	ezbēti
	3rd masc.	-Ø	šar	ezib
	3rd fem.	-at	šarrat	ezbet
Plural	1st com.	-ānu	šarrānu	ezbēnu
	2nd masc.	-ātunu	šarrātunu	ezbētunu
	2nd fem.	-ātina	šarrātina	ezbētina
	3rd masc.	-ū	šarrū	ezbū
	3rd fem.	-ā	šarrā	ezbā

Note the following developments for the enclitic subject pronominal suffixes: (i) in Assyrian the first common plural subject marker is *-āni*; (ii) in Old Assyrian the second masculine singular subject is *-āti*; (iii) in Neo-Assyrian and Neo-Babylonian the first common singular is *-āk(a)* (< *-āku*). Only in Old Akkadian and Old Assyrian is there a third masculine dual form attested, with subject marker *-ā*: for example, *šarrā* “the two are kings”; Old Akkadian also attests a third feminine dual form with the enclitic subject marker *-tā*: for example, *šalimtā* “you two are well.”

The predicative form appears to be ancient, given the appearance of a similar construction in Old Egyptian. In other Semitic languages, however, the form has become an active, perfective verb: for example, Hebrew *ʾāzabū* “they have abandoned” versus Akkadian *ezbū* “they are abandoned.”

The predicative form is also attested for Eblaite, as evidenced by forms such as *na-im* /*našim*/ “he/it is good”; *da-nu-nu* /*dannunū*/ “they (masc.) are strengthened”; *ʾa<sub>5</sub>(NI)-bū-ḥa* /*ʾabbuxā*/ “they (fem.) are girded.” Additionally, both Eblaite and Old Akkadian (as well as Amorite) attest an archaic third masculine singular predicative form ending in *-a*. This form is no longer productive by the time of the first attested texts and is limited primarily to personal names: for example, the Eblaite name *Na-ma-Da-mu* “Damu is pleasant”; and the Old Akkadian name *Šu-be-la* “he is a lord.”

#### 4.1.4 The absolute form

The absolute form of the noun is an indeclinable form, without mimation or case ending, and resembles the third masculine singular of the predicative form (see §4.1.3); in fact, the absolute form may represent in origin an embedded predication. Although the absolute form is not fully understood, several functions are clear: (i) for certain frozen adverbial expressions (e.g., *zikar sinniš* “male (and) female,” *šeher rabi* “small (and) great”); (ii) to express the vocative (e.g., *bēlet* “lady!”, *šar* “king!”); (iii) often for cardinal numbers (e.g., *ištēn* “one,” *šinā* “two”); (iv) in expressions of mass and quantity, where both the number and the unit of measurement appear in the absolute (e.g., *sebe uṭṭet (kaspum)* “seven grains (of silver)”; and (v) in distributive expressions or in expressions in which the substantive is unmistakably singular (e.g., *ina ellat ellat* “with every caravan,” *šanat* “a single year”).



The absolute form is sparsely attested in Eblaite where it is used to express: (i) divine names (e.g., <sup>d</sup>*Ga-mi-iš*, <sup>d</sup>*Ra-sa-ap*); (ii) geographical names (e.g., *A-da-bí-ik<sup>ki</sup>*, *A-da-ti-ik<sup>ki</sup>*); and (iii) month names (e.g., ITI *za-’à-na-at*).

#### 4.1.5 Adjectives

##### 4.1.5.1 Attributive adjectives

In addition to predicative adjectives (see §4.1.3), Akkadian displays attributive adjectives which follow their head nouns and agree with their antecedents in case, gender, and number. For the most part, attributive adjectives are declined like substantives (see [7]), with the exception of the masculine plural endings *-ūtum* (nom.), *-ūtīm* (obl.), which are distinctive. Eblaite likewise exhibits these plural forms: for example, *a(-wa)-mu ’à-mu-tum* /*yawmū hammūtum*/ (= Akk. *ūmū emmūtum*) “hot days.”

Most Akkadian adjectives are derived from verbal roots, and as such, have the form *parVs*, where the second vowel is *i* for active roots (both transitive and intransitive), but unpredictable for adjectival/stative roots (*a*, *i*, or *u*). However, the distinctive vowel of the second syllable is syncopated in all forms except for the attributive feminine singular (see §3.8.2.2) and the predicative form with third masculine singular subject.

Verbal adjectives denote the condition or state resulting from the action of the verb from which it is derived: (i) for transitive verbs, the verbal adjective is passive (e.g., *šabtum* “captured” < *šabātum* “to capture”); (ii) for active-intransitive verbs, the verbal adjective is resultative and perfective (e.g., *maqtum* “fallen, having fallen” < *maqātum* “to fall”); (iii) for stative/adjectival verbs, the verbal adjective is descriptive (e.g., *damqum* “good” < *damāqum* “to be good”). Old Babylonian and Old Akkadian declensions of *šarrum* “king,” *šarratum* “queen,” and the modifying attributive adjective *damqum* “good” are presented in (9):

##### (9) Akkadian adjective declension (Old Babylonian and Old Akkadian)

	<i>Masculine</i>	<i>Feminine</i>
<i>Singular</i>		
<i>Nominative</i>	šarrum damqum	šarratum damiqtum
<i>Genitive</i>	šarrim damqim	šarratim damiqtim
<i>Accusative</i>	šarram damqam	šarratam damiqtam
<i>Dual</i>		
<i>Nominative</i>	šarrān damqūtum	šarratān damqātum
	šarrān damqān (OAKk.)	šarratān damqatān (OAKk.)
<i>Genitive-accusative</i>	šarrīn damqūtīm	šarratīn damqātīm
	šarrīn damqīn (OAKk.)	šarratīn damqatīn (OAKk.)
<i>Plural</i>		
<i>Nominative</i>	šarr(ān)ū damqūtum	šarrātum damqātum
<i>Genitive-accusative</i>	šarr(ān)ī damqūtīm	šarrātīm damqātīm

An adjective modifying more than one noun appears in the plural; compound, mixed gender antecedents are modified by masculine plural adjectives. Dual adjectives are only attested for the Old Akkadian period, after which dual nouns are modified by plural adjectives.

Any adjective may be used substantively: for example, the masculine singular adjective *šabtum* “seized” (< *šabātum* “to seize”) may also be used to express the masculine singular substantive “captive.” Frequently, the feminine singular of an adjective is used substantively to denote an abstract noun: for example, *damiqtum* “good(ness), luck, fame” (cf. masc. sg. *damqum* “good” < *damāqum* “to be good”).

#### 4.1.5.2 Comparatives and superlatives

Akkadian did not possess distinct forms for expressing the comparative and superlative; instead, either the attributive or the predicative adjective was used. The preposition *eli* was used in comparative expressions: for example, *awīlam ša elišu rabū imḥaṣ* “he struck a man who is greater than he.” The superlative is expressed by the bound form of the adjective: for example, *Ištar rabit ilātim* “Ištar is the greatest of the goddesses.” Additionally, the Š stem (see §4.2.4) verbal adjective is often used as a superlative: for example, *šurbūm* “very great, greatest” (< *rabūm* “great”).

#### 4.1.5.3 Denominative adjectives

The addition of the ending *-ī* to the base of a noun, followed by the adjectival case ending, forms an adjective with the meaning “pertaining to, related to (the noun in question).” Common denominative adjectives include *maḥrūm* “former, earlier” (< *maḥrī-um*; cf. *maḥrum* “front”); *elūm* “upper” (< *elī-um*; cf. *elum* “top”); *šaplūm* “lower” (< *šapli-um*; cf. *šaplum* “bottom”). Gentilic adjectives are formed by adding the denominative *-ī* to place names: for example, *Akkadūm* (< *Akkad* + *ī* + *um*) “Akkadian” (masc. nom.).

#### 4.1.5.4 Independent possessive adjectives

Possession may be expressed by a set of possessive adjectives (as well as with pronominal suffixes; see §4.1.6.2), particularly for the Old Assyrian and Old Babylonian periods (although even for these periods, possession is usually expressed by the pronominal suffixes). The possessive adjectives agree with their antecedent nouns in case, gender, and number; they do not agree with the gender of the possessor. These adjectives may be used attributively (e.g., *kaspam yām* “my silver” [acc.]) or, more commonly, as predicates (e.g., *bītum šū yūm* “that house is mine”; *šūrūtum yā’ū* “the black [textiles] are mine” [masc. pl.]). The nominative forms of the attested possessive adjectives are presented in (10); attested Assyrian forms appear in parentheses:

#### (10) Nominative forms of the possessive adjectives

	SINGULAR		PLURAL	
	<i>Masculine</i>	<i>Feminine</i>	<i>Masculine</i>	<i>Feminine</i>
<i>Singular</i>				
1st	yūm (yā’um)	yattu <sup>m</sup> /n	yūt(t)u <sup>m</sup> /n	yāt(t)u <sup>m</sup> /n
2nd	kūm (kuā’um)	kattum (kuātum)	kūt(tu) <sup>m</sup> /n	kāt(tu) <sup>m</sup> /n
3rd	šūm (šūā’um)	šattu <sup>m</sup> /n (šūātum)	šūt(tu) <sup>m</sup> /n	—
<i>Plural</i>				
1st	nūm (niā’um)	niattu <sup>m</sup> /n	nūt(tu) <sup>m</sup> /n	—
2nd	kunūm	(kunūtum)	(ku(w)ā’ūtum)	—
3rd	šunūm	šunūtum	—	—

The unusual nunation (rather than expected mimation) that often accompanies feminine and plural forms of these adjectives, as well as feminine and plural forms of the demonstrative *annūm* “this” (see §4.1.5.5; for example, fem. sg. *annītum*, masc. pl. *annūtum*), is probably the result of partial assimilation to the preceding *t* of such forms.

In the Middle Babylonian period the possessive pronoun *attu-* replaced the possessive adjectives discussed above. It was always combined with the possessive suffixes (see §4.1.6.2): for example, *ḥarrāna attū’a* “my caravan (acc.)”; *šibšu attūšu* “his lease.”

#### 4.1.5.5 Demonstrative adjectives and pronouns

The most common near demonstrative (“this,” “these”) in Akkadian is *annûm* (base *anni-*), feminine *annîtum*; in Neo-Babylonian, *agû*, feminine *agātu* largely replaces *annûm*. Far demonstratives (“that,” “those”) are represented by *ullûm*, feminine *ullitum*, in Babylonian, and *ammîu(m)*, feminine *ammîtu(m)*, in Assyrian. All of the preceding are inflected with adjectival case endings. For the Old Babylonian, Middle Babylonian, and Standard Babylonian periods particularly, the third-person independent personal pronouns (see §4.1.6.1) are used to express the far demonstrative. This pronoun also serves as an anaphoric pronoun, expressing “the aforementioned”: for example, *alpam šuāti išriq* “he stole the aforementioned ox.”

### 4.1.6 Pronouns

#### 4.1.6.1 Independent personal pronouns

Akkadian distinguishes a nominative, dative, and common genitive-accusative case for the independent personal pronoun. As noted above, the third-person nominative and genitive-accusative forms are also used for the anaphoric pronoun. The genitive-accusative forms are marked with *-ti*, the dative with *-ši(m)*. The dative form represents an innovation within Akkadian – and Eblaite, which also attests these pronouns – and is not found in the other Semitic languages; it is most often found after the preposition *ana* “to”: for example, *ana kâšim taklāku* “I trust you” (fem. sg.). Dative forms are attested primarily in Old Babylonian and in the post-Old Babylonian periods; in Old Assyrian the genitive-accusative forms were usually employed for the dative (i.e., after the preposition *ana*). Nominative personal pronouns are often found in verbless clauses, or for emphasis or clarification in verbal clauses, since the pronominal subject is always included in the verb; the genitive-accusative and dative pronouns are used principally for emphasis. Only for the Old Akkadian period is a dual independent personal pronoun attested, third common *šunīti* (gen.-acc.); however, the presence of dual suffixed pronouns (see §4.1.6.2) in both Old Akkadian and Eblaite may suggest that both languages had a full set of yet-unattested, dual independent personal pronouns. The Old Babylonian paradigm for the most common forms of the independent personal pronouns is set out in (11); Old Assyrian forms (when different from the Old Babylonian) are placed in parentheses, and attested Eblaite forms in brackets:

(11)	<i>Nominative</i>	<i>Genitive-accusative</i>	<i>Dative</i>
<i>Singular</i>			
<i>1st com.</i>	anāku [ʾana]	yāti	yāši(m)
<i>2nd masc.</i>	attā [ʾanta]	kāti/a (ku(w)āti) [kuwāti]	kāši(m) [kuwāši(m)]
<i>2nd fem.</i>	attī	kāti (ku(w)āti)	kāši(m)
<i>3rd masc.</i>	šū (šūt) [šuwa]	šuāti/u [šuwāti]	šuāši(m) [šuwāši(m)]
<i>3rd fem.</i>	ši (šit) [šiya]	šiāti	šiāši(m)
<i>Plural</i>			
<i>1st com.</i>	nīnu (nēnu)	niāti	niaši(m)
<i>2nd masc.</i>	attunu [ʾantanu]	kunūti	kunūši(m)
<i>2nd fem.</i>	attina	kināti	kināši(m)
<i>3rd masc.</i>	šunu [šunū]	šunūti	šunūši(m)
<i>3rd fem.</i>	šina	šināti	šināši(m)

#### 4.1.6.2 Pronominal suffixes

The pronominal suffixes are attached to both nouns and verbs: (i) the genitive or possessive suffixes are appended to the bound form of the noun (including nonfinite forms of the

verb), (ii) while the accusative and dative suffixes, which refer to direct and indirect objects, respectively, may only be appended to finite verbs (predicative adjectives may accept dative suffixes as well): for example, *bitni* “our house” (nom., acc.); *išbatniāti* “he seized us”; *alikniāšim* “Come (masc. sg.) to us!” In later dialects the accusative and dative suffixes are no longer carefully distinguished. The ventive morpheme (see §4.2.9) may precede the accusative and dative suffixes; when both an accusative and a dative suffix are attached to a given verb the order is Verb–(Ventive)–Dative–Accusative: for example, *aṭrud(ak)kuššu* (< *aṭrud* + *am*) + *kum* + *šu*) “I sent him to you.” The pronominal suffixes for the Old Babylonian period are given in (12); Old Assyrian forms (when different from the Old Babylonian) are placed in parentheses, and attested Eblaite forms in brackets:

(12)	Possessive (genitive)	Accusative	Dative
<i>Singular</i>			
1st com.	-ī, -(y)a [-ī]	-anni/-nni/-ninni [-ni]	-am/-m/-nim
2nd masc.	-ka [-ka]	-ka	-kum [-kum]
2nd fem.	-ki [-ki]	-ki	-kim
3rd masc.	-šu [-šu]	-šu [-šu]	-šum [-šum]
3rd fem.	-ša [-ša]	-ši	-šim
<i>Plural</i>			
1st com.	-ni [-ni]	-niāti [-ni]	-niāšim (-niāti)
2nd masc.	-kunu	-kunūti (-kunu)	-kunūšim (-kunūti)
2nd fem.	-kina	-kināti (-kina)	-kināšim (-kināti)
3rd masc.	-šunu [-šunu]	-šunūti (-šunu)	-šunūšim (-šunūti)
3rd fem.	-šina [-šina]	-šināti (-šina) [-šināt]	-šināšim (-šināti)

Only for Old Akkadian and Eblaite are dual pronominal suffixes attested: second common dual *-kunī* [-*kumayn*] (gen.), *-kunīšim* [-*kumay(n)*] (dat.), *-kunīt(i)* (acc.); third common dual *-šunī* [-*šumay(n)*] (gen.), *-šunūti* [-*šumay(n)*] (acc.), *-šunīšim* (dat.). As noted above, the third person is expressed with *š* in Old Akkadian, prior to the merger of \**š*/*š* with \**θ*. The first common singular possessive suffix has the allomorphs *-ī*, *-(y)a*, the choice of which depends on the morphological shape of the preceding noun: *-ī* is used after singular nouns in the nominative or accusative (e.g., *mārti* “my daughter,” *epēši* “my doing”); otherwise, *-(y)a* is used, including after the case-vowel of singular nouns and adjectives in the genitive (e.g., *ināya* “my eyes” [dual], *itti abiya* “with my father”); after *-ū* (and sometimes *-ā*), *-(y)a* is often written as *-a* (i.e., written with the A sign; e.g., *mārū’a* “my sons,” *epšētū’a* “my deeds”).

The first common singular dative suffix is identical to the ventive morpheme (see §4.2.9), and the choice of allomorph *-am/-m/-nim* is dependent on morphological considerations (namely, the person and number of the verb on which it appears): *-am* is attached to forms without a vocalic ending (i.e., 3rd com. sg., 2nd masc. sg., 1st com. sg., and 1st com. pl.); *-m* is attached to the form ending in *ī* (i.e., the 2nd fem. sg.); *-nim* is attached to forms ending in *-ū* (i.e., 3rd masc. pl.) and in *-ā* (i.e., 3rd fem. pl. and 2nd com. pl.).

The allomorphs of the accusative suffix, *-anni/-nni/-ninni*, are derived from the ventive/dative suffix + *-ni*, where the *m* assimilates to the following *n*; the distribution of the accusative forms is identical to that of the corresponding ventive/dative suffix. The second- and third-person genitive pronominal suffixes may be apocopated in poetic texts, the singular forms usually following the appropriate case-vowel: for example, *libbuš* “her heart” (nom.; for *libbaša*). In Old Assyrian, after a short *-a*, bisyllabic suffixes are shortened: for example, *tuppašnu* “their tablet” (< *tuppašunu*); *libbaknu* “your (pl.) heart” (< *libbakunu*).

#### 4.1.6.3 Reflexive pronouns

The noun *ramānum* (Ass. *ramunum*) “self, oneself” is used with pronominal suffixes as a reflexive pronoun: for example, *ana ramāniya* “for myself.” The nouns *pagrum* “body, corpse” and *napištum* “life” may also be used as reflexive pronouns when pronominal suffixes are attached: for example, *pagarka ušur* “guard yourself!”

#### 4.1.6.4 Determinative-relative pronoun

The determinative-relative pronoun was fully declined in Old Akkadian and Eblaite. Because of the ambiguity of the writing system, it is not certain whether the initial consonant of the Eblaite form was *θ* as in Old Akkadian or *ð* as in West Semitic. The paradigm of the Old Akkadian determinative-relative pronoun is given in (13); the signs used to represent the attested Eblaite forms appear in brackets:

(13)	SINGULAR		PLURAL	
	<i>Masculine</i>	<i>Feminine</i>	<i>Masculine</i>	<i>Feminine</i>
<i>Nominative</i>	θū [θU]	θât [θU-DU]	θût	θât [θU-DU]
<i>Genitive</i>	θī [θI]	θâti [θU-TI]	θûti [θU-TI]	*θâti [θU-TI]
<i>Accusative</i>	θā [θA]	θât	θût	θât

The determinative-relative pronoun for all subsequent periods, *ša* (< *θā*, the Old Akkadian accusative masculine sing. form), was indeclinable, although declined forms of the pronoun still remained in several frozen expressions: for example, *šūt rēši* “the ones (masc. pl.) of the head,” that is, “courtiers.” As a determinative pronoun, *ša* represents “the one of”: for example, *ša Bābilim* “the one of Babylon”; usually, however, it stands in apposition to a preceding noun, and is translated “of.” Nouns that follow *ša* are always in the genitive case: for example, *šarrum ša mātīm* “the king, the one of the country,” that is, “the king of the country.” As a relative pronoun, *ša* is translated “which, who, etc.” and serves to introduce relative clauses (see §5.12.1).

The relative pronouns in Eblaite are often expressed by the logogram LÚ, Sumerian “man,” which was normally used in Sumerian to introduce relative clauses with male animate antecedents. In Eblaite, however, it is used with animates and inanimates, both male and female.

#### 4.1.6.5 Interrogative/Indefinite pronouns and adjectives

The personal interrogative pronoun in Akkadian and Eblaite is *mannum* “who?,” which is declined for case, but has no special feminine or plural forms. The impersonal interrogative pronoun *mīnum*, or *minûm*, “what?,” is similarly declined for case and likewise is without special feminine or plural forms; note also the phrase *ana mīnim* “for what?” that is, “why?.” Additionally, there is an interrogative adjective *ayyum* “which?,” that agrees with its antecedent in case, number, and gender: for example, *ayyitum iltum* “which goddess?”; *ayyum*, unlike most adjectives, may precede the noun it modifies.

The indefinite pronouns are formed either by the reduplication of the bases of the interrogative pronouns, or by appending the particle *-ma* to the latter. The personal indefinite pronoun is *mamma(n)* (< \**manman*) “anyone, someone,” with the negative adverb *lā*, “not any, no.” The impersonal indefinite pronoun is *mimma* (< *mīnum* “what?”) “anything, something, all,” with negative adverb *lā*, “nothing.” Both *mamma(n)* and *mimma* are indeclinable and both may be used as generalizing relative pronouns, especially in late texts. Additionally, the adjectival *ayyumma* “whichever, any, some,” based on the interrogative adjective, *ayyum* “which?,” agrees with its noun in case, number, and gender.

## 4.2 Verbal morphology

As in other Semitic languages, the verbal morphology of Akkadian is complex. The verbal root usually consists of three consonants; however, Akkadian displays many so-called *weak verbs* – verbs in which one or more of the original root consonants disappeared altogether or were susceptible to certain phonological changes in specific environments. Additionally, roots of four radicals, referred to as *quadriradical* verbs, are attested as well. Akkadian exhibits four finite forms (tenses or aspects): durative, preterite, perfect, and imperative. As in other Semitic languages, Akkadian derives verbs by means of prefixes and modifications of the verbal roots. These set patterns, better known as *stems*, have characteristic meanings and functions; the range of meanings of a verb for a given derived stem can be more or less extrapolated from the basic stem. Each stem has three nonfinite forms: infinitive, participle, and verbal adjective. For reasons of economy, the outline of the Akkadian verbal morphology presented below is based solely on the strong verb, although it should be kept in mind that many of the most basic and frequent verbs in the language are weak.

Knowledge of Eblaite verbal morphology is greatly limited by the paucity of syllabically spelled verbs. However, it is clear that Eblaite exhibits both the preterite and the durative forms. Verbs with infixes *-t-* (see below) may indicate either G perfects, or Gt or Gtn preterites. Other stems attested for Eblaite are the D, Š, Št, N, Ntn, and ŠD (?), although examples for many of these are scarce. The stem-vowel of the various stems and tenses behaves as it does in Akkadian. The West Semitic fientive *qatala* form cannot be confirmed for Eblaite; rather, such forms most likely represent an archaic predicative form (see §4.1.3).

### 4.2.1 G stem

The basic verbal stem is the G stem (*Grundstamm*, sometimes referred to as the B stem for *basic*). Finite forms of the G stem exhibit a *thematic vowel* between the second and third radical which is lexical and unpredictable. In most cases, the vowel is the same for both the durative and the preterite (the perfect shares its theme vowel with the durative, while the imperative shares its theme vowel with the preterite) – *a*, *i*, or *u*. However, the largest class of verbs, the *ablaut class*, displays an *a* in the durative (and perfect) and *u* in the preterite (and imperative). As evidenced throughout Semitic, the various semantic categories may be roughly associated with the different theme-vowel classes: (i) the ablaut class (*a-u*) contains predominantly transitive verbs; (ii) the relatively small *a*-class similarly contains mostly transitive verbs; (iii) the large *i*-class (which represents the confluence of the Proto-Semitic *a-i* and *i-a* classes) is associated with many stative verbs, but active transitive and intransitive verbs appear in this class as well; and (iv) the large *u*-class is associated with many intransitive verbs.

Person, number, and gender are indicated by prefixes and suffixes; the durative, preterite, and perfect of the G stem (and N stem as well, discussed below) share the same set of affixes. Old Babylonian forms are given below:

(14)	Singular		Plural	
3rd com.	i-	3rd masc.	i-	-ū
		3rd fem.	i-	-ā
2nd masc.	ta-	2nd com.	ta-	-ā
2nd fem.	ta- -ī			
1st com.	a-	1st com.	ni-	

For the Assyrian dialects, Old Akkadian, and the Old Babylonian literary dialect, there was a distinct third feminine singular form marked with the prefix *ta-*. In Eblaite the

third feminine singular prefix was regularly *ta-* (with a biform *ti-*): for example, *taqīš*, *tiqīš* “she presented.” Furthermore, Eblaite exhibits a third masculine plural form with prefixed *ti-* (*tīprusū*): for example, *ti-da-ḥa-ru<sub>12</sub>* /tittaxrū/(?) “(the gods) approached”(?) (cf. Akk. *mithurum*); *ti-na-ḥu-uš* /ti?naxū-š/(?) “(the gods) got tired of it”(?) (cf. Akk. *anāḥum*), which is also encountered in Old Akkadian texts from Mari (and in Middle Babylonian-period peripheral texts from Amarna and Ugarit). For Old Akkadian, Old Assyrian, and Eblaite a third common dual form is attested, marked by the prefix *i-* (Ebl. *yi-*) and the suffix *-ā*.

As noted above, the prefix *i-* derives from PS \**ya* (see §3.8.1.2, 6) and/or \**yī*; only in Old Akkadian and Eblaite is the *y* of the prefix preserved. The prefixes *a-* and *ta-* become *e-* and *te-* when used with verbs containing *e* in Babylonian (see §3.1.3.1, 1): for example, Bab. *telqe*, Ass. *talqe* “you (masc. sg.) took”; or with I-*e* verbs in Babylonian and Assyrian: for example, *ēpuš* “I did.” With I-*w* verbs, *u* replaces the *a-* and *i-*vowels of the prefixes: for example, *urid* “I/he descended,” *turid* “you descended.”

#### 4.2.1.1 Durative (*parrVs*)

The durative of the strong verb is characterized by the doubling of the middle radical; the vowel between C<sub>1</sub> and C<sub>2</sub> is invariably *a* (except where *a* > *e* by Babylonian vowel harmony, see §3.1.3.1, 1) and the vowel between C<sub>2</sub> and C<sub>3</sub> is, as noted above, the theme vowel. Note the G durative paradigm for the verb *parāsum* (*a-u* class) “to cut”:

(15)	Singular		Plural
3rd com.	iparras	3rd masc.	iparrasū
		3rd fem.	iparrasā
2nd masc.	taparras	2nd com.	taparrasā
2nd fem.	taparrasī		
1st com.	aparras	1st com.	niparras

The durative describes action that is nonpunctual or imperfective; most often it corresponds to the English present or future. As its name implies, the durative may also be used to describe any durative, nonpunctual action, past, present, or future: for example, *ikan-nak* may also be used with the meanings “he was sealing, he is sealing, he will be sealing.” However, the durative may be used to describe habitual or customary action – *išakkan* “he used to place, he (continually) places, he will place (regularly)” – as well as potential or probable action: for example, *išabbat* “he may/might/could/can/should/would seize.” The precise nuance of the durative can only be determined from the surrounding context. The durative form was inherited from Proto-Semitic, and like forms are found in Ethiopian and modern South Arabian; however, the form was lost in the Central Semitic languages (for example, Arabic, Hebrew, and Aramaic).

#### 4.2.1.2 Perfect (*ptarVs*)

The perfect is characterized by an infixed *-ta-* immediately after the first consonant. With the addition of the plural suffixes (*-ū*, *-ā*, *ī*), subordination marker (*-u*), or ventive morpheme (*-am*), the theme vowel between C<sub>2</sub> and C<sub>3</sub> is lost according to the vowel syncope rule (see §3.8.2.2). The infixed *-t-* assimilates completely to the first consonant when that consonant is a dental or sibilant (except *š*) (see §3.8.1.4, 14). Below is the G perfect paradigm for *parāsum*:



(16)	Singular		Plural
3rd com.	iptaras	3rd masc.	iptarsū
		3rd fem.	iptarsā
2nd masc.	taptaras	2nd com.	taptarsā
2nd fem.	taptarsī		
1st com.	aptaras	1st com.	niptaras

The perfect roughly corresponds to the English present perfect, for example, *nimtaqut* “we have fallen,” and represents actions that have been completed but affect the present. However, the perfect has many other nuances and functions that are dialect- and genre-specific. In Old Akkadian the perfect is only sparsely attested. For the Old Babylonian period, however, the perfect is quite common, especially in letters and in the conditional clauses that comprise the various law codes, including the Code of Hammurabi. In these genres, the perfect has a focusing nuance, denoting the central event in a sequence of events. In the letters of the Old Babylonian period the perfect is used with this nuance in conjunction with the adverbs *inanna* “now” or *anumma* “now, herewith, hereby” to emphasize the immediacy or current relevance of the event, the so-called epistolary or announcement perfect: for example, *inanna wardam ana maḥrika aṭṭardam* “I have now sent the servant to you.” Because of its emphatic character, the perfect rarely occurs in questions or relative clauses.

After the Old Babylonian and Old Assyrian periods, the perfect replaced the preterite as the main form used to express the past, the preterite having become restricted to negative main clauses and subordinate clauses. This distribution of forms is exactly parallel to that of certain pairs of forms in some modern Ethio-Semitic languages (Hetzron 1968); to borrow the terminology used in reference to the marked member of the Ethiopic forms, we may say that the *t* of the Middle Babylonian and Middle Assyrian perfect functions as a main verb marker.

#### 4.2.1.3 Preterite (prVs)

The preterite denotes a punctual, completed action as seen by the speaker or writer: for example, *iddin* “he gave.” It is most often translated by the simple past tense in English. Below is the paradigm of the G preterite for *parāsum* (*a-u*):

(17)	Singular		Plural
3rd com.	iprus	3rd masc.	iprusū
		3rd fem.	taprusā
2nd masc.	taprus	2nd com.	taprusā
2nd fem.	taprusī		
1st com.	aprus	1st com.	niprus

The preterite form is found in secondary uses in other Semitic languages, having been replaced as the past tense by a suffix-pronoun conjugation.

#### 4.2.1.4 Imperative and precative

The imperative occurs only in the second person; in form, the imperative is the preterite without prefix. The resulting initial consonant cluster is generally resolved by inserting the preterite theme-vowel between *C*<sub>1</sub> and *C*<sub>2</sub>: for example, *šabat* “seize! (masc. sg.).” In the feminine singular and common plural forms, the addition of the respective *-ī* and *-ā* suffixes causes the vowel between *C*<sub>2</sub> and *C*<sub>3</sub> to be lost according to the rule of vowel syncope (see §3.8.2.2).



The imperative is complemented by the precative, which is used to express wishes and indirect commands in the first and third person. Like the imperative, the precative is based on the preterite. The Babylonian precative is formed by replacing the first-person singular prefixes with *lu-* and third-person prefixes with *li-*; the first common plural precative consists of an unattached short *i* before the preterite: for example, *lukšud* “may I arrive, let me arrive”; *i niškun* “let us place, may we place.” In Assyrian the precative has the form of the preterite plus a prefixed *l-* (except in the first common plural and third feminine singular, where the particle *lū* appears before the preterite). In Old Akkadian, the precative is identical to the Babylonian form, with the exception of the third feminine singular, which follows the Assyrian model: for example, *lū tamḥur* “may she receive.” Additionally, Old Akkadian attests a third common dual precative: for example, *lilqutā* “may they (dual) gather.” Note the suppletive injunctive (imperative plus precative) paradigm for *parāsum* (*a-u*); Assyrian forms are in parentheses:

(18)	<i>Singular</i>		<i>Plural</i>
3rd com.	liprus	3rd masc.	liprusū
(3rd fem.	lū taprus)	3rd fem.	liprusā
2nd masc.	purus	2nd com.	pursā
2nd fem.	pursi		
1st com.	luprus	1st com.	i niprus
	(laprus)		(lū niprus)

Neither the imperative nor the precative is used with a negative adverb; rather, the prohibitive and the vetitive are used to express negative commands and wishes.

#### 4.2.1.5 Prohibitive and vetitive

The negative counterpart of the imperative is the prohibitive, used to express negative commands and prohibitions. The form consists of the negative adverb *lā* followed by the durative: for example, *lā tašappar* “do not send, you may not send (masc. sg.)”; *lā ipallaḥā* “they (fem.) may not/ shall not fear.” The vetitive, used to express negative wishes, is formed by prefixing *ayy-* to forms of the preterite that have an initial vowel, and *ē-* to forms with an initial consonant (in Assyrian the prefix is *ē-* in all cases): for example, *ē-taškunā* “may you (pl.) not place, you should not place”; *ayy-ašpur* “may I not send, I do not wish to send.” In the Neo-Assyrian dialect the distinction between the prohibitive and the vetitive is blurred, and often the two are used interchangeably.

#### 4.2.2 N stem

The N stem is characterized by a prefixed *n* before the root; in forms in which *n* stands directly before a consonant, it assimilates completely to that consonant (see §3.8.1.4, 12). The N stem, being based on the G stem, shares the same set of personal affixes, and, as in the G, the middle radical is doubled in the N durative and the theme vowel of the N perfect is that of the durative. However, the N differs from the G in its organization of vowel classes: verbs of the *a-u* and *a* classes in the G are collapsed into an *a-i* class in the N; G *i* class verbs are unchanged in the N. G *u* class verbs, which are rare in the N, sometimes remain *u*-class and sometimes join the dominant *a-i* pattern. When vocalic suffixes are added to the preterite, the vowel between C<sub>2</sub> and C<sub>3</sub> is lost because of syncope (see §3.8.2.2). Note the following singular and plural forms for the verb *parāsum* “to cut”:

## (19) Third person and imperative forms of the N stem

<i>Durative</i> (3rd com. sg./3rd fem. pl.)	ipparras/ipparrasā
<i>Perfect</i> (3rd com. sg./3rd fem. pl.)	ittapras/ittaprasā
<i>Preterite</i> (3rd com. sg./3rd fem. pl.)	ipparis/ipparsā
<i>Imperative</i> (fem. sg./com. pl.)	naprisī/naprisā

The N stem serves as the passive of active-transitive G verbs: for example, *ipparis* “it was separated”; *ṭuppum iššebir* “the tablet was broken.” Stative verbs, although rarely attested in the N, assume an ingressive nuance: for example, *šumšu immassik* “his name will become bad” (cf. *maskum* “bad” [G verbal adj.]). Several verbs have a reflexive meaning in the N: for example, *ittalbaš* “he has clothed himself” (< *labāšum* “to put on clothing”). A few verbs have N rather than G as their basic form (i.e., lexical N verbs): for example, *ippalis* “he looked” (< *palāsum* “to see,” rare in G).

## 4.2.3 D stem

The D stem (*Doppelungsstamm*) is distinguished by a double middle radical in all finite and nonfinite forms. The personal prefixes of the D (and Š, see below) all have *u* where the G and N stems have *a* or *i*; thus, the first-person and third-person singular are formally identical: for example, *udammīq* “I/she/he made good.” The distribution of U-signs suggests that in Old Akkadian there was a difference in the pronunciations of the first- and third-person prefixes, i.e., the U-sign is used fairly consistently for the third-person prefix (probably representing *yu-*), whereas the first person is usually written with the Ú- and Û-signs (probably for *’u-*). All D verbs belong to the *a-i* class (as do Š verbs; see below), regardless of their vowel classes in the G or N; significantly, the theme-vowel of the perfect follows the preterite and not the durative, as in the G and N. Hence, the theme-vowel of all D duratives is *a*; the theme-vowel of all preterites, perfects, and imperatives is *i*. The Assyrian D and Š imperatives, as well as the other prefixless forms, i.e., the verbal adjective and infinitive, differ from their Babylonian counterparts in exhibiting *a* between C<sub>1</sub> and C<sub>2</sub> rather than *u*. Below is a sample of singular and plural forms in the D stem (Assyrian forms are in parentheses):

## (20) Second person and imperative forms of the D stem

<i>Durative</i> (2nd masc. sg./2nd com. pl.)	tuparras/tuparrasā
<i>Perfect</i> (2nd masc. sg./2nd com. pl.)	tuptarris/tuparrisā
<i>Preterite</i> (2nd masc. sg./2nd com. pl.)	tuparris/tuparrisā
<i>Imperative</i> (masc. sg./com. pl.)	purris/purrisā (parris/parrisā)

A recent study concludes that “the basic function of the D stem is that of underlining an increase in transitivity vis-à-vis the corresponding G stem” (Kouwenberg 1997:445). For intransitive G verbs the D normally has a factitive function, as in *mātam urappiš* “he widened the land” (cf. G *mātum irpiš* “the land became wide”), whereas for transitive G verbs the D connotes “plurality and salience, mostly plurality of the direct object and the action itself” (Kouwenberg 1997:445), as in *nārātim upetti* “I opened canals” (cf. G *nāram epte* “I opened a canal”). The D may denote an activity performed on plural objects: for example, *ušebber ṭuppātim* “he broke many tablets” (cf. G *išber ṭuppam* “he broke the tablet”). The D also serves to form denominative verbs: for example, *ruḡgubum* “to roof something” (< *ruḡbum* “roof”). Additionally, there are some lexical verbs for which the D is the basic stem: for example, *kullumum* “to show.”

#### 4.2.4 Š stem

The Š stem is characterized by a prefixed *š(a)-* before the verbal root. The vowel class and the personal prefixes correspond precisely to those of the D; in other words, all verbs belong to the *a-i* class and *u* replaces the *a* and *i* vowels of the G (and N) prefixes. As noted under the D stem, prefixless forms of the Š stem (i.e., the imperative, verbal adjective, and infinitive) have *a* in the first syllable in Assyrian and *u* in Babylonian. Note the following forms (Assyrian forms are in parentheses):

##### (21) First person and imperative forms of the Š stem

<i>Durative (1st com. sg./1st com. pl.)</i>	<i>ušapras/nušapras</i>
<i>Perfect (1st com. sg./1st com. pl.)</i>	<i>uštapis/nuštapis</i>
<i>Preterite (1st com. sg./1st com. pl.)</i>	<i>ušapris/nušapris</i>
<i>Imperative (fem. sg./com. pl.)</i>	<i>šuprisī/šuprisā</i> (šaprisī/šaprisā)

The main function of the Š stem is to form the causative of G verbs, particularly active-intransitive verbs: for example, *ušamqit* “I/she/he caused (someone, something) to fall” (cf. G *imqut* “she/he fell”). For some adjectival verbs, the Š rather than the D serves as the factitive stem: for example, *tušamrišā* “you (pl.) made sick, caused trouble” (cf. *nimraš* “we became sick”). As with the other derived stems, some verbs occur only in the Š and have no G counterpart: for example, *šuklulum* “to complete.”

#### 4.2.5 Infixed -ta- stems

For each of the stems presented above there is an infixed -ta- sub-stem: Gt, (Nt), Dt, Št<sub>1</sub>, and Št<sub>2</sub>. The theme-vowels of the -ta- stems are those of the corresponding basic stems (however, G *a-u* verbs are *a* class verbs in the Gt). The infixed -ta- is inserted after the first radical in the G and D, or after the characteristic preformative of Š (and N) stems. The preterites of the -ta- infix stems are formally indistinguishable from the G, D, N, and Š perfects, so that the two can be distinguished only by context; the perfect of the -ta- stem infixes -tat(a)-.

The Gt stem denotes (i) a reciprocal meaning (e.g., *mithušum* “to strike one another,” *qitribum* “to draw close to one another”); (ii) a reflexive nuance (e.g., *piššušum* (< \**pitšušum*) “to anoint oneself”); or (iii) a separative sense with verbs of motion (e.g., *atlukum* “to go away”). While the first two functions of the Gt are known from elsewhere in Semitic, the last-named, the separative sense, represents an innovation within Akkadian.

The Dt and Št<sub>1</sub> stems serve principally as the passive of their respective stems: for example, *uštālpit* “he was destroyed” (cf. *ušālpit* “he destroyed”). The Št<sub>2</sub> stem, which is distinguished from the Št<sub>1</sub> stem only in the durative (by a characteristic doubled middle radical, e.g., Št<sub>1</sub> *uštāpras*, Št<sub>2</sub> *uštāparras*), has a variety of functions, including reflexive and passive of the Š, causative of the Gt and N, and denominative (Streck 1994).

The Nt stem is exceptionally rare and its existence may very well be questioned, especially given that attested forms are identical to the Ntn (see §4.2.6). Where it does seem to appear, it has a reciprocal or separative nuance.

#### 4.2.6 Infixed -tan- stems

Similarly, each of the main stems has an infixed -tan- stem associated with it – Gtn, Ntn, Dtn, and Štn; the theme-vowel is that of the corresponding -ta- stem. As with the

-*ta*- stems, the -*tan*- morpheme is inserted immediately after the first consonant in the G and D, or, in the N and Š, after the characteristic preformative. Only in the durative is the -*tan*-morpheme completely preserved: for example, Gtn *iptanarras*, Ntn *ittanapras*, Dtn *uptanarras*, Štn *uštanapras*. In all other forms, the *n* either assimilates to the following consonant (in the Gtn), or is dropped (in the Ntn, Dtn, and Štn), as in the following preterite forms: Gtn *iptarras*, Ntn *ittapras*, Dtn *uptarris*, Štn *uštapis*. Therefore, with the exception of the durative, the Ntn, Dtn, and Štn forms are identical to their -*ta*- counterparts. For all stems, the -*tan*- infix serves as an iterative to the meaning of the corresponding main stem: for example, *aštanappar* “I am continually sending word” (cf. *ašappar* “I am sending word”); *ištatakkān* “he has placed repeatedly.”

#### 4.2.7 Rare stems

In addition to the stems described above, several other stems of rare or restricted occurrence are also attested. The ŠD stem – which combines the double middle radical of the D stem and the š- preformative of the Š stem (e.g., *ušpazzer* “he saved”) – is used for poetic effect and is limited to Standard Babylonian and the Old Babylonian literary dialect where it may replace either the D or the Š stem. The very rare R and Rt stems, which are characterized by reduplication of the third radical (e.g., *iprassas* 3rd com. sg. durative), denote intensification of the verbal root and are, therefore, similar to the D. The Dtt stem (e.g., *uptatarras* 3rd com. sg. durative) is attested only in Neo-Assyrian where it acts as the passive of the D stem (the expected Dt stem probably does not occur in Neo-Assyrian). Additional stems have been suggested, but such forms are so sparsely attested that they may very well represent scribal errors.

#### 4.2.8 Quadriradical verbs

As noted above, a few roots have four radicals; for some the fourth radical is weak. The vast majority of these verbs are not attested in the G stem, but instead have the N as their basic stem. Causatives are formed with the Š, and iterative Ntn and Štn are attested as well: for example, *nabalkutum* N (*a*) “to jump, rebel”; *šubalkutum* Š causative. Most quadriradical verbs have *l* or *r* as the second radical.

#### 4.2.9 The ventive morpheme

The ventive morpheme is closely related to the first common singular dative suffix; in fact, the two are identical in terms of morphological shape, -*am*, -*m*, and -*nim*, and distribution (see §4.1.6.2). The ventive is a directional element that denotes motion or activity in the direction of the speaker or writer; it is most frequently found with verbs of movement and of sending. The element may be suffixed to any finite verbal form including the imperative: for example, *nillikam* “we came here” (< *alākum* “to go”); *šūbīlam* “send (it) here!” (< Š *wabālum* “to carry”). As with the separative sense of the Gt stem, the ventive seems to represent an innovation within Akkadian; note the opposition created by the two with the verb *alākum* “to go”: *atlak* “go away!” (Gt) against *alkam* “come here!” (G + ventive).

#### 4.2.10 Subordination markers

The morpheme -*u* is suffixed to all verbal forms that occur in subordinate clauses, provided that the verb does not have an ending that is part of the subject marker (2nd fem. sg. -*ī*,

pl. *-ū* and *-ā*) or the ventive morpheme. In Babylonian, verbs that have one of these endings are unmarked in subordinate clauses. In Old Assyrian, *-ni* is attached to forms that cannot take *-u*. In Middle Assyrian and Neo-Assyrian *-ni* is attached even to those forms which are already marked with *-u*; in these dialects a pronominal suffix may intervene between the *-u* and the *-ni* subordination markers. The variations in the form of the subordination marker may be summarized as follows (after Huehnergard 1997:284; *iprus* = 3rd masc. sg. preterite of *parāsum* “to cut”):

## (22) Forms of the subordination markers

	MAIN CLAUSE	SUBORDINATE CLAUSE		
	<i>Bab./Old Ass.</i>	<i>Old Bab.</i>	<i>Old Ass.</i>	<i>Mid/Neo-Ass.</i>
<i>Pret. 3rd masc. sg.</i>	<i>iprus</i>	ša <i>iprusu</i>	ša <i>iprusu</i>	ša <i>iprusū-ni</i>
+ <i>Ventive</i>	<i>iprusam</i>	ša <i>iprusam</i>	ša <i>iprusan-ni</i>	ša <i>iprusan-ni</i>
+ <i>3rd masc. sg. suff.</i>	<i>iprusšu</i>	ša <i>iprusūšu</i>	ša <i>iprusūšu</i>	ša <i>iprusūšū-ni</i>
+ <i>Vent.+suff.</i>	<i>iprusaššu</i>	ša <i>iprusaššu</i>	ša <i>iprusaššū-ni</i>	ša <i>iprusaššū-ni</i>
<i>Vbl. Adj.+3rd fem. sg.</i>	<i>parsat</i>	ša <i>parsat</i>	ša <i>parsat-ni</i>	ša <i>parsatū-ni</i>

In Old Akkadian the normal subordination marker is *-u*; however, in addition to *-u*, sometimes the suffix *-ni* (or *-na*) is used as in Assyrian. In Old Akkadian texts from the Diyala region a unique subordination suffix *-a* is attested. For the syntax of subordinate clauses see §5.12.

## 4.2.11 Nonfinite verbal forms

Akkadian exhibits three nonfinite forms: infinitive, participle, and verbal adjective, all of which are attested for the G, N, D, Š and their respective *-ta-* and *-tan-* stems.

### 4.2.11.1 Verbal adjective

The functions of the verbal adjective and its G form have already been discussed above (see §4.1.5.1). The verbal adjective for the derived stems can similarly be used attributively or predicatively. The morphological shape of the verbal adjective is formally identical to the infinitive for all derived stems: Gt *pitrus-* (Ass. *pitars-*); Gtn *pitarrus-*; N *naprus-*; Ntn *itaprus-* (with loss of initial *n*); D *purru-* (Ass. *parrus-*); Dt, Dtn *putarrus-*; Š *šuprus-* (Ass. *šaprus-*); Št<sub>1-2</sub>, Štn *šutaprus-*.

### 4.2.11.2 Participle

The participle is a declined adjective that is very often substantivized. The shape of the participle in the G is *pāris-*; for the other derived stems the participle is characterized by the prefix *mu-*: Gt *muptaris-*; Gtn *muptarris-*; N *mupparis-*; Ntn *muttapris-*; D *muparris-*; Dt, Dtn *muptarris-*; Š *mušapris-*; Št<sub>1-2</sub>, Štn *muštapis-*. The participle is active in voice for transitive verbs while the verbal adjective is passive: for example, *šābitum* “captor” versus *šabtum* “captive.” For active intransitive verbs the participle is characterized by an imperfective aspect, while the verbal adjective imparts a perfective nuance: for example, *wāšibum* “sitting down” versus *wašbum* “having sat down, seated.” Participles of stative verbs do not occur. As a substantive, the participle is most often found in the bound form with a dependent genitive: for example, *pāris purussē* “the one who decides decisions”; *wāšib ālim* “the one who dwells in the city, city-dweller.”

### 4.2.11.3 Infinitive

The Akkadian infinitive is a verbal noun and is always masculine and singular. The shape of the infinitive in the G stem is *parās-*. For the derived stems the morphological shape of the infinitive is identical to that of the verbal adjective (see §4.2.11.1).

In Eblaite, the D and Š stem infinitives have the same form as in Assyrian Akkadian, namely, D *parrus* and Š *šaprus*. (The forms *purruš* and *šurruš* are also attested in copies of one recension of a large Sumerian–Eblaite vocabulary text; see Conti 1996. It is likely that this recension reflects the influence of a contemporary dialect of Old Akkadian.) In addition to these forms, Eblaite attests infinitives of *t*-stems that, in addition to an infixed *-t-*, also exhibit a prefixed *t-*: G iterative /tartappidum/ (cf. G /rapādum/ “to run”); D iterative(?) /tuḏtaqqinun/ (cf. G /ḏaqqānum/ “to be bearded,” D /ḏaqqunum/); Š iterative /tuštaʔkilum/ (cf. Akkadian G *akālum* “to eat,” Š *šūkulum* “to feed”). Such forms are otherwise unknown in Semitic.

For the syntax of the infinitive see §5.5.

## 4.3 Particles

Prepositions, adverbs, and the particle *lū* are treated below; for the conjunctions and enclitic particles see Syntax.

### 4.3.1 Prepositions

Common prepositions in Akkadian are the following: *ina* “in”; *ana* “to, toward”; *ištu*, Middle Babylonian *ultu*, Neo-Assyrian *issu* “from, out of,” temporal “since, after” (from Proto-Semitic \**wiṣtu(m)* “in(side)”; *adi* “up to”; *kī*, *kīma* “like, as”; *lāma* “before” (temporal) (from *lā* “not” + enclitic *-ma* [see §5.7]); *aššu(m)*, Old Assyrian *aššumi* “on account of, for the sake of” (from *ana šum(i)* “for the name (= sake) of”); *ašar* “where” (originally the bound form of a noun *ašrum* “place,” i.e., “(in the) place of”); *mala* “as many as, as much as” (originally the bound form of the infinitive *malûm* < \**malā’um* “to become full,” i.e., “fullness of”). The prepositions *itti* “with” and *eli* “upon, over, against” regularly take possessive suffixes: for example, *elīya* “upon me”; *ittišu* “with him.” Nouns governed by prepositions are always in the genitive case.

Of the above-mentioned prepositions, Eblaite shares the prepositions *in* “in” (Akk. *ina*), *ʾa<sub>5</sub>(NI)-na* “to” (Akk. *ana*), which are not found elsewhere in Semitic outside of these two languages. Eblaite also attests a preposition *ʾaštā* or *ʾaštī* “with, from” (corresponding to the Old Akkadian preposition *ištē/ištī*), *ʾaštu(m)* (Oakk. *ištum*) “from, after,” and *bali* “without, without the knowledge/consent of” (Akk. *balu(m)*). Additionally, Eblaite contains the prepositions *si-in* “to, for the sake of” (not attested in Akkadian, but found in South Arabian *s<sub>1</sub>n*); *min* “in, at” (not attested in Akkadian, but cf. Hebrew *min-* “from”); *mi-nu* “from, to”; *si-gi* “with”; GABA “before” (probably for *maḥaras* in Akk.; construct of *maḥrum* “front (part)”; *iš-ki* “in favor of” (which has been compared with Ethiopic *ʾaska* “until”); and *al* “on” found elsewhere in Semitic (Akk. *eli*). Significantly, Eblaite, like Akkadian, does not display the common West Semitic proclitic prepositions \**bi-* “in” and \**la-* “to.”

### 4.3.2 Adverbs

It was noted above that the locative ending *-um* and terminative ending *-iṣ* remained productive after the Old Akkadian period in the formation of adverbs (see §4.1.1.3): for example,

*elēnum* “above, in addition”; *mādiš* “much, greatly.” When followed by the morpheme *-am*, possibly the accusative case, the terminative often assumes a distributive force: for example, *ūmišam* “daily”; *warḥišam* “monthly.” Additional adverbial endings include *-i* (e.g., *ali* “where”) and *-Ø* (i.e., the absolute form [see §4.1.4]; e.g., *zamar* “quickly, suddenly”).

The accusative case may be used in a wide range of adverbial functions: (i) accusative of place (e.g., *šumēlam* “on the left”); (ii) accusative of time (e.g., *urram* “tomorrow”); (iii) ablative accusative (e.g., *niḷqēšunūti* “we took from them”); (iv) accusative of respect, manner, or means (e.g., *ḥamuttam alkam* “Come quickly!”).

An adverbial use of the accusative is attested for Eblaite as well: for example, (i) accusative of place (e.g., *zi-il NE-na-áš<sup>ki</sup> mu-DU É ma-tim* “at the junction of N., we will enter into the mausoleum”; cf. var.: *zi-il NE-na-áš<sup>ki</sup> mu-DU si-in É ma-tim*); (ii) accusative of time (e.g., 5 UD GIBIL... [TUŠ] “on the fifth day... they sat”); (iii) accusative of respect (e.g., *wa-a PAD ma-lik-tum ba-na-sa* “she veiled the face of the queen”; lit.: “she veiled the queen with respect to her face”; *ma-lik-tum* represents a fixed logographic writing for /māliktam/ (see §2.6 and §4.1.2)).

There are two negative adverbs in Akkadian, *ul* and *lā*. The former also has the form *ula* in early Old Babylonian and Old Assyrian. The particle *ul* is used to negate independent declarative sentences and clauses, both verbal and verbless. It is also used to negate interrogative sentences in which no interrogative pronoun or adverb occurs. Elsewhere *lā* is used to negate: all subordinate clauses; injunctions, both verbal and verbless (see §§4.2.1.5, 4.3.3); interrogative nouns and pronouns; and individual words, including infinitives and adjectives: for example, *ṭemum lā damqum* “an unfavorable report”; *dabāb lā kittim* “untrue speech.”

Although some adverbs may stand at the beginning of their sentence, their regular position is directly before the verb (in verbless clauses, the negative adverb *ul* stands likewise before the predicate).

### 4.3.3 The particle *lū*

This particle has three functions in Akkadian: (i) to express alternatives (e.g., *abum lū aḥum lū aššatum* “father, brother, or wife”); (ii) to denote injunctions in verbless clauses (e.g., *lū awilāta* “be (masc. sg.) a man”; negative injunctions are formed with *lā*: e.g., *lā ina ekallim šina* “they (fem.) must not be in the palace!”); and (iii) as an asseverative particle (e.g., *lū ēpuš* “I verily built”).

## 4.4 Compounding

As with the other Semitic languages, Akkadian is characterized by a poverty of real word compounds. However, certain compound noun phrases are expressed by a bound form governing a genitive: for example, *mār(i) šiprim* “messenger” (lit. “son of the message”); *bēl ḥubullim* “creditor” (lit. “lord of the debt”). Only occasionally are these expressions treated as a morphological unit; in other words, their plurals are most often formed by pluralization of the governing noun (e.g., *mārū šiprim* (nom.) “messengers”), rather than by marking the end of the phrase (e.g., *mār šiprī*). Rarely, however, a bound form with accompanying genitive evolved into a type of word compound, especially when the final consonant of the bound form was an *n* and was therefore susceptible to assimilation to the first consonant of the following genitive: for example, *būn pānī* (lit. “features of the face”) > *buppānu* “face”; *šaman šammim* (lit. “oil of the plant”) > *šamaššammum* “sesame seed oil”; *mūr nisqim* (lit. “foal of choice quality”) > *murnisqum* “(select) horse or donkey.”



## 4.5 Numerals

Cardinal and ordinal numbers are usually expressed logographically in Akkadian. Thus, the pronunciation and construction of many numbers are unknown.

### 4.5.1 Cardinal numbers

The cardinal numbers from 1 to 10, except for 2, occur in the absolute and, less often, free forms. There was no “zero” in the Akkadian numerical system. The numbers from 11 to 19 occur only in the absolute state. Below are the Old Babylonian forms for numbers 1 through 10:

#### (23) The old Babylonian cardinal numbers (1–10)

	DECLINED STATE		ABSOLUTE STATE	
	<i>Masculine</i>	<i>Feminine</i>	<i>Masculine</i>	<i>Feminine</i>
1	(ištēnum)	(ištētum)	ištēn	išteat, ištēt
2	šinā	šittā	—	—
3	šalāšum	šalāštum	šalāš	šalāšat
4	erbūm	erbettum	erbe/erba	erbet(ti)
5	ḥamšum	ḥamištum	ḥamiš	ḥamšat
6	šeššum	šedištum	šediš?	šeššet
7	sebūm	sebettum	sebe	sebet(ti)
8	samānūm	samāntum	samāne	samānat
9	tišūm	tišitum	tiše	tišīt
10	eš(e)rum	ešertum	ešer	eš(e)ret

The cardinals 20 through 50 are expressed in the feminine plural of the absolute form; they may modify masculine or feminine nouns: *ešrā* “20”; *šalāšā* “30”; *erbeā/erbā* “40”; *ḥamšā* “50”. In compound numbers, higher-order components precede lower ones.

The numerical system, as inherited from Sumerian, was based on both the sexagesimal and the decimal systems. Higher numbers are expressed in both systems; the following forms modify both masculine and feminine nouns:

#### (24) Higher-ordered Old Babylonian cardinal numbers: sexagesimal and decimal systems

SEXAGESIMALS			DECIMALS		
	<i>Absolute</i>	<i>Free</i>		<i>Absolute</i>	<i>Free</i>
60	šūš(i)	šūšum	100	meat	(meatum)
600	nēr	nērum	1,000	līm(i)	limum
3,600	šār	šārum			

##### 4.5.1.1 Agreement with cardinal numbers

The numbers 1 and 2 agree in gender with the item counted; however, the numbers 3 through 19 are subject to the phenomenon of chiasmic concord encountered elsewhere in Semitic, whereby the gender of these numbers is the opposite of that of the item counted: for example, *ištēn wardum* “one male slave,” *ištēt amtum* “one female slave,” but *šalāšat wardū* “three male slaves”; *šalāš amātum* “three female slaves.” While numbers are usually expressed in the absolute case as noted above, the item counted or measured is usually in the free form, its case determined by the context.



### 4.5.2 Ordinal numbers

The ordinal numbers are adjectives which always agree with the gender of the modified noun (i.e., chiasitic concord is not observed; see §4.5.1.1). The base of the cardinal numbers in Babylonian is *parus-*, in Assyrian *paris-*. There are several terms for “first”; additionally, “first” is the only ordinal regularly to precede its noun. Note the following Old Babylonian forms:

#### (25) The old Babylonian ordinal numbers (1–10)

	<u>Masculine</u>	<u>Feminine</u>
<i>First</i>	pānûm maḥrûm (išti <sup>3</sup> um) ištēn	pānītum maḥrītum (ištītum) išteat
<i>Second</i>	šanûm	šanītum
<i>Third</i>	šalšum	šaluštum
<i>Fourth</i>	rebûm	rebūtum
<i>Fifth</i>	ḥamšum	ḥamuštum
<i>Sixth</i>	šeššum	šeduštum
<i>Seventh</i>	sebûm	sebūtum
<i>Eighth</i>	samnum	samuntum
<i>Ninth</i>	tešûm	tešūtum
<i>Tenth</i>	ešrum	ešurtum

## 5. SYNTAX

To date there have been no comprehensive studies concerning the historical development of Akkadian syntax; thus, the description given below is based largely on Old Babylonian, the most thoroughly studied dialect. However, even a fairly cursory review of the other dialects reveals that Akkadian is remarkably conservative not only in its morphology, but also in its syntax, despite the fact that the written language in all probability lagged behind the spoken language.

Beyond the brief comments made below regarding word order and use of the conjunction *wa*, almost nothing can be stated with certainty regarding the syntax of Eblaite. Note: In the word-for-word renderings of Eblaite, the verb forms are given simply as bare lexical forms, such as “go,” “dwell,” since the logograms used to write them do not specify person or tense.

### 5.1 Word order

One of the most significant innovations within Akkadian is the adoption of an SOV (Subject–Object–Verb) word order, under Sumerian influence, for both main and subordinate clauses. This development is in sharp contrast to the VSO order of most other Semitic languages, with the notable exception of modern Ethiopic, which similarly adopted an SOV order under Cushitic influence. In literary genres, however, Akkadian word order was much less restricted, the verb often preceding the object, or even the subject, for poetic effect.

Word order in Eblaite appears to be much freer than in Akkadian. In addition to the SOV order encountered in Akkadian and Sumerian, SVO order is frequently attested as well, as in other Semitic languages.

## 5.2 Agreement

As noted above, the Akkadian and Eblaite verbs agree with their subject in gender and number; attributive adjectives follow their head nouns and must agree with them in gender, number, and case. The predicative construction is not subject to the same rules of agreement as the attributive form; the base is indeclinable and not marked for gender, number, or case, while the enclitic subject pronoun has a clearly defined gender (except for the first person), number, and case, which is invariably nominative. For example, *šarr-āku* means both “I am king” and “I am queen” (i.e., a feminine form *\*\*šarrat-āku* does not exist); similarly, *šarr-ānu* literally means both “we are king” and “we are kings.”

## 5.3 Apposition

Nouns and phrases are very frequently found in apposition, particularly with titles and epithets. Nouns in apposition agree in number and in case with their antecedent: for example, *ana Marduk, bēlim rabīm, rubēm maḥrīm* “for Marduk, the great lord, the foremost prince,” where all substantives and modifying adjectives are governed by the preposition *ana*, and are thus in the genitive case.

## 5.4 Genitive constructions

In Akkadian the genitive may be expressed either with or without the determinative pronoun, with no ascertainable difference in meaning. The determinative pronoun *ša* (see §4.1.6.4) is placed in apposition to the preceding governing noun; the governed noun appears after *ša* and is in the genitive case (26A). Occasionally *ša* appears before its antecedent, in which case the governed noun is re-expressed with a resumptive possessive pronoun (26B–C):

- (26) A. *šarrum ša mātim*  
king, the one of land-GEN.  
“The king of the land”
- B. *ša bēlim kussīšu*  
the one of lord-GEN. his throne  
“The throne of the lord”
- C. *ša PN aštakan dabdašu*  
the one of PN-GEN. I brought about his defeat  
“I brought about the defeat of PN”

The latter construction, often referred to as the *anticipatory genitive*, is used for poetic effect and is most often found in literary texts.

However, the use of the bound form (see §4.1.2) directly before a dependent noun is the more common construction for expressing the genitive. The determinative pronoun is deleted and the governing noun appears in the bound form, juxtaposed to the governed noun in the genitive: for example, *qaqqad awilim* “the head of the man.” The resulting construction, the so-called *genitive chain*, represents an inseparable unit. Adjectives modifying the governing noun occur after the chain in the appropriate case (27A–B). Only the negative adverb *lā* may intervene between the bound form and its genitive (27C–D):

- (27) A. mār awīlim šeḥrum  
 son-of man.GEN. young  
 “The young son of a man”  
 B. mār bīte rabû  
 son-of house.GEN. great  
 “The oldest son of the house” (Middle Assyrian)  
 C. eršet lā târi  
 land-of not return.GEN.  
 “Land of no return”  
 D. bēl lā ilim.  
 owner-of not god.GEN.  
 “Owner of no god” (i.e., “irreligious person”)

Normally, only one genitive noun can be dependent on a governing noun; an exception occurs when two genitives form a logical unit: for example, *bēl šamê u eršetim* “lord of heaven and earth.” It is impossible for more than one bound form to govern a single genitive; in instances of multiple governing nouns *ša* must be used:

- (28) ina eqlim kirīm ū bītim ša ilkīšu  
 in field.GEN. orchard.GEN. or house.GEN the one of his service obligation.GEN.  
 “Among a field, orchard, or house belonging to his service obligation”

When chains of more than two elements occur, all but the last element appear in the bound form:

- (29) A. qurun šalmāt ummānātīšu  
 pile-of corpses-of his troops.GEN.  
 “The pile of corpses of his troops”  
 B. ina qāt mār awīlim  
 from hand-of son-of man.GEN.  
 “From the hand of the son of a man”

## 5.5 Syntax of the infinitive

The syntax of the Akkadian infinitive is very complex, and only the basic aspects can be mentioned here; for details see Aro 1961. The infinitive, as a verbal noun, may occur in any case. As the subject or direct object of its clause, it appears in the nominative and accusative cases, respectively: for example, *erēšum qerub* “planting is near”; *erēbam ul iddišim* “he did not allow her to enter” (lit. “entering he did not give to her”). When governed by a preposition or bound noun, the infinitive appears in the genitive: for example, *ašar lā amārim* “a place that cannot be found” (lit. “a place of not finding”). When governed by certain prepositions, particularly *ina* “in,” the resulting prepositional phrase is often equivalent to a temporal clause: for example, *ina kašādīm* “when arriving, upon arrival”; *itti zikarim šanīm ina utūlim lā ittašbat* “(if) she has not been caught while lying with another man.” With other prepositions, particularly *ana* “to,” the construction may express purpose: for example, *ana lā enê* “so that it cannot be changed” (Neo-Babylonian/Late Babylonian); *ana amār bēliya šarik* “it was donated so that my lord sees it” (lit. “for the seeing of my lord”; Middle Babylonian). Many other prepositions may be construed with the infinitive, resulting in a wide range of meanings and nuances: for example, *aššum elēm aštaprakkum* “I have written to you (masc. sg.) about coming up”; *kīma lā ragāmim epuš* “act (masc. sg.) so that there be no legal contest” (lit. “act according to not contesting”). Additionally, as a

noun, the infinitive may take possessive suffixes: for example, *ṭēmka ina šemêya* “upon my hearing your report”; *adi târīšu šibā* “stay (pl.) until his return.”

The infinitive may also be construed as a verb, taking a subject or object: for example, with a pronominal suffix serving as the logical subject:

- (30) *tuppam ina šemêka*  
 tablet in hear.INF.=your  
 “When you hear the tablet”

With the infinitive taking both a subject and an object (rare), one finds, for example:

- (31) *dannum enšam ana lā ḥabālim*  
 strong weak to NEG. oppress.INF.  
 “So that the strong do not oppress the weak”

The subject or object of an infinitive may be expressed in the genitive if it follows a preposition:

- (32) A. *ana šemê bēliya ašpuram*  
 to hear.INF.-of lord=my I=wrote=VENT.  
 “I have written so that my lord hears it”  
 B. *ana awātīšu kašādīm ēgurakka*  
 to words=his achieve.INF. he=hired=VENT.=you  
 “Has he hired you in order to achieve his goals?”

As verbs, infinitives may also govern prepositional phrases and adverbial complements:

- (33) A. *ina Kaniš ina erābīšu*  
 in Kaniš in enter.INF.=his  
 “When he entered into Kaniš” (Old Assyrian)  
 B. *lā alāka iqbīšu*  
 NEG. go.INF. he=spoke=him  
 “He ordered him not to go” (Neo-Assyrian)

Infinitives may also enter into paranomastic constructions which serve to intensify the verbal form; in such constructions, the infinitive takes the locative-adverbial *-um* (§4.1.1.3) and often *-ma* (§§5.7, 5.9): for example, *šapārum-ma ašpur* “I have certainly sent.”

In a construction unique to Eblaite, infinitives placed in apposition to their objects may accept dative pronominal suffixes, which in Akkadian are restricted to finite verbal forms:

- (34) A. 1 DUG Ì <sup>giš</sup>GAB.LIŠ.ME na-ba-ba-šum (for /napāp-šum/)  
 1 vessel oil G... sprinkle.INF.=to-him  
 “A vessel of G.-oil, to be sprinkled on him”  
 B. ḥa-sa-nu BAR<sub>6</sub>.KÜG sa-ḥa-da-šum (for /šaḥād-šum/)  
 H. silver give.INF.=to-him  
 “An H. of silver, to be given to him”

## 5.6 Verbless clauses

There is no verb “to be” in Akkadian; instead, equational and existential clauses are verbless, expressed by juxtaposing the subject and the predicate. The tense of verbless clauses can

only be determined from context. In Old Babylonian, when the subject of a verbless clause is a noun, it precedes the predicate:

- (35) Ḫammurapi šarrum ša Bābilim  
 Hammurapi king the one of Babylon  
 “Hammurapi [is/was] king of Babylon”

If the subject is a pronoun, it follows the predicate:

- (36) šarrum ša Bābilim šū  
 king the one of Babylon he  
 “He [is/was] king of Babylon”

In other dialects, such as Old Assyrian (37A–B), and in Eblaite (37C) these rules of word order are not in force:

- (37) A. nēnu lā awīl gimillim  
 we NEG. man-of compliance  
 “We [are] not compliant people”  
 B. gāmer awātim nēnu  
 concluder-of matters we  
 “We are the concluder(s) of the legal matters”  
 C. an-da ŠEŠ ù an-na ŠEŠ  
 you brother and I brother  
 “You [are my] brother, I [am your] brother”(?)

In Akkadian, clauses of the type *Adverb (Phrase)–Noun (Phrase)* occur occasionally. Most often this construction is used for existential clauses; for example:

- (38) ina imitti ḫašim šēpum ina šumēl ḫašim piṭrum  
 in right-of lung foot in left-of lung fissure  
 “on the right of the lung [there] was a ‘foot,’ on the left [there] was a fissure”

Verbless clauses may also express simple possession when a noun phrase governed by *ša* (see §5.4) constitutes the predicate, as in:

- (39) kaspum u ḫurāšum ša ālim  
 silver and gold that of town  
 “The silver and gold belong to the town”

## 5.7 Topicalization

In the writing of Akkadian there are two methods of emphasizing a nonpredicate constituent: (i) left dislocation, and (ii) the addition of the emphatic particle *-ma*. With the former, the dislocated noun or noun phrase is placed at the beginning of the clause in the nominative case (sometimes referred to as the *nominative absolute* or *casus pendens*). Such clauses appear to have two subjects; however, the dislocated element is not part of the clause grammar. The noun or noun phrase that is dislocated is replaced in the clause by a corresponding pronominal suffix; for example:

- (40) A. *šumma awīlum ḥubullum elišu ibašši-ma*  
 if man.NOM. debt.NOM. against=him it=is-present=CONJ.  
 “If a man – a debt is lodged against him . . .”

- B. *šumma awīlum šārassu . . . šalmat*  
 if man.NOM. hair-NOM.=his . . . is black  
 “If a man – his hair is black”

With B compare:

- C. *šumma šarat awīlim . . . šalmat*  
 if hair-of man.GEN. . . is black  
 “If the hair of a man is black”

The enclitic particle *-ma* serves to mark the logical predicate of a clause (Rainey 1976); translation into English is usually facilitated by a cleft sentence, as in:

- (41) A. *aššum mārī Yā'ilānim ša maḥrika, tuša warkānum*  
 concerning sons-of Y. REL. before=you perhaps later  
*salimum ibbašši-ma ina qātim kullašunu aqbi.*  
 peace it=becomes-present=CONJ. in hand hold.INF.=them I=said  
*inanna mimma salimum itti Yā'ilānim ul ibašši, ša*  
 now anything peace with Y. NEG. it=is-present REL.  
*šabātīšū-ma adabbub*  
 seize.INF.=him=TOP. I=speak  
 “As for the Yā'ilānum tribesmen who are with you, I had said to hold them just in case peace should be established later. Now, there is no peace with Yā'ilānum; it is to seize them that I am planning”
- B. *Gimillum šū, dūršu nuḥatimmum; watriššu ana rēdim*  
 G. that status=his cook superfluously to r.-soldier  
*iššaṭer. inanna Gimillum šū, ina nuḥatimmī-ma illak!*  
 he=was-written now G. that in cooks=TOP. he=goes  
*pūḥšu, šani'am-ma ana rēdi mulli*  
 replacement=his other=TOP. to r.-soldiers assign.IMPV.  
 “As for this Gimillum, his permanent status is that of cook; he was registered as a rēdû-soldier superfluously. Now, as for this Gimillum, it is with the cooks that he will serve! In place of him, it is someone else that you must assign”

## 5.8 Cliticism

In addition to the emphatic particle *-ma* and the subordinate marker *-ni* in Assyrian (and *-na* in Old Akkadian), Akkadian possesses two other enclitics: *-mi* and *-man* (Old Assyrian *-min*). The particle *-mi* is used to indicate direct speech; it may be attached to one or more words within the speech:

- (42) *šāpirī iqbiām kīda šunū-mi lībalū-mi*  
 overseer=my he=said=VENT. outside they=QUOT. may=they=dry=QUOT.  
 “My overseer said to me, ‘Those should be dried outside’”

The relatively infrequent irrealis particle, *-man/-min*, is appended to *šumma* “if” to form unreal clauses: for example, *šumma-min mētāku* “if I had died.”

Akkadian does not possess any true proclitics; however, in some dialects, the prepositions *ina* “in,” *ana* “to,” and *eli* “upon” may lose their final vowels and become proclitic, with

assimilation of the consonant to the first consonant of the following word. Often the resultant consonantal doubling is not indicated in the script: for example, *a-pa-ni-ia* for *ap-pānīya* “towards me” (< *ana pānīya*); *i-li-bi-ša* for *il-libbiša* “within it (fem.)” (< *ina libbiša*); *e-ni-ši-i* for *en-niši* “above the people” (< *eli niši*).

## 5.9 Coordination

There are two independent coordinators in Akkadian, *u* and *ū* (*lū*), in addition to the enclitic *-ma*, which, besides its function as an emphasizing particle, is frequently used as a coordinator (Patterson 1970; Kraus 1987). The conjunction *u* (< *\*wa*) is used to connect both noun phrases and sentences. Clauses that are connected with *u* are of equal semantic stress and are reversible (i.e., a change in the order of the clauses does not affect their relationship or meaning): for example, ... *bītam inaššarū u šeḫrūtīm urabbū* “... they may keep the house and raise the children.”

The coordinating conjunction *-ma* is suffixed to verbs and is used only to connect clauses. Unlike *u*, clauses connected by *-ma* are logically or temporally related to one another and are therefore irreversible. Usually the first clause provides the conditions for the action expressed in the second clause; thus, *-ma* serves syntactically or logically to subordinate the first clause to the second. The following example demonstrates the various interpretations of clauses connected with *-ma*:

- (43) *ina nār GN mû maṭū-ma eḡel biltīni*  
 in river-of GN water.PL. diminished-are=CONJ. field-of tax=our  
*ul ikaššadū*  
 NEG. they=reach  
 “In the canal of GN the waters are (too) low and so they do not reach our taxable field”  
 “Because/ When/ If the waters are (too) low they will not reach our taxable field”

Akkadian did not possess a separate word for “but”; rather, both *u* and *-ma* can be used in this sense, especially when one of the clauses contains a negative, as, for example, in the following:

- (44) *BÜR.30.IKU šītāt eqlim šuāti bēlni ana ŠU.ḪA.MEŠ*  
 BUR=30(-DETV.) remainder-of field that lord=our to fishermen  
*UD.DA nadānam-ma ipiršunu lā šūšām iqbi*  
 collective give.INF.ACC.=and ration=their NEG. release.INF.ACC. he=said  
 “Our lord said to give the remaining 30 BUR of that field to the fisherman’s collective but not to release their rations”

The conjunction “or” is expressed by *ū* (< *\*ʾaw*) or *ū lū*. The conjunctions *u* and *ū* are identical in the writing system and, in the absence of *lū*, the two can only be distinguished by context.

Clauses may also be joined asyndetically, with the deletion of any of the conjunctions described above:

- (45) *Purattu ... miqtīša usuḫ ḡamiša šutbi*  
 Euphrates ... silt.PL.=its remove.IMPV. litter.PL.=its cause-to-arise.IMPV.  
*šutēšerši*  
 cause-to-be-in-order.IMPV.=it  
 “... as for the Euphrates, ... dredge its silt, remove its litter (and) set it in order!”

The conjunctions *wa* (Akk. *u*) and *-ma* are attested for Eblaite as well, although little is known of their syntactic ranges. In at least some instances *-ma* behaves as it does in Akkadian, connecting logically related clauses, for example:

- (46) <sup>d</sup>EN.KI... LUGAL iš<sub>11</sub>-gur-ma MAḤ(?) il-tum <sup>d</sup>EN.LÍL  
 Ea... king he=summoned=CONJ. exalted(?) gods Enlil  
 ʾa<sub>5</sub>(NI)-na <sup>d</sup>EN.KI INIM.DI  
 to Ea spoke  
 “he summoned Ea, ..., the king, and then the exalted one(?) of the gods, Enlil, spoke to Ea”

The conjunction *wa* is much more frequent in Eblaite than in Akkadian; and, as in Akkadian, it is often used to connect clauses and noun pairs. As in West Semitic, *wa* is often used to introduce clauses and sentences (and is presumably to be left untranslated):

- (47) A. wa Ì.NA.SUM-kum É in ba-da-a ša 2 li-im  
 CONJ. give=to-you house in B. REL. 2 thousand  
 “I am (herewith) giving you property in Baytān, (populated) by 2,000 (people)”  
 B. wa iš<sub>11</sub>-da-ga-sù 1 SUD MAŠKIM.E.GI<sub>4</sub>-ma si-in  
 CONJ. he=established(?)=him 1 star representative=TOP. toward  
 I-li-lu A.MU DINGIR.DINGIR.DINGIR  
 Enlil father gods  
 “The star established(?) him as representative to Illil, the father of the gods”  
 C. wa ÍL IGI.IGI EN wa NAM.KU<sub>5</sub>  
 CONJ. lift eyes lord CONJ. swear  
 “The lord lifted (his) eyes and swore”

Such a use of *u* in Akkadian is very rare. Eblaite also attests the conjunctions *šumma* (see §5.11); *ʾap* “and then, but then,” which occurs in Ugaritic and Hebrew, but not in Akkadian; and *ū-ma* (a compound of *u* [*< wa*] and *-ma*) “and then, and also, even.”

## 5.10 Sequence of tenses (*consecutio temporum*)

For Old Babylonian and Old Assyrian (i.e., those periods in which the perfect is used with a focusing nuance; see §4.2.1.2), past actions performed in sequence are often expressed by one or more preterites followed by a final perfect. This sequence of tenses is used to emphasize the final clause, the crucial event upon which the action in the subsequent clauses is based. Often the coordinator *-ma* connects the clause(s) containing preterite(s) to the following perfect clause; for example:

- (48) A. kaspam aknukam-ma uštābilakkum  
 silver I=sealed=VENT.=CONJ. I=have caused-to-carry=VENT.=to-you  
 “I sealed the silver and have sent it to you”  
 B. inanna milum illikam-ma nār Irnina ana dūr  
 now flood it=went=VENT.=CONJ. river-of I. to wall-of  
 kārim issaniq (< \*istaniq; see §3.8.1.4, 14)  
 quay it=has-reached  
 “Now, the flood has come, and the Irnina Canal has reached up to the wall of the quay”



C. ana Ili-inguranni tuppam ušābil-ma meher tuppi  
 to I. tablet I=caused-to-carry=CONJ. copy-of tablet  
 ušābilam-ma uštābilakkum  
 he=caused-to-carry=VENT.=CONJ. I=have-caused-to-carry=VENT.=to-you  
 “I sent a letter to Ili-inguranni, and he sent a response to me, and I have sent  
 (it) to you”

### 5.11 Conditional sentences

Akkadian exhibits both marked and unmarked conditional sentences. Unmarked clauses (with no word for “if”) are characterized by the conjunctive *-ma*, which serves to connect the protasis and the apodosis. The verbs of both clauses are usually in the durative: for example, *taša”al-ma iqabbākku* (Middle Babylonian) “If you ask, he will tell you.” Sometimes the protasis has instead the preterite or precativ; thus

- (49) mārī šanūtīm liršū-ma PN aḥūšunu rabūm  
 sons other may=they=acquire.PREC.=CONJ. PN brother=their big  
 “Even if they acquire (adopt) other children, PN will be their older brother”  
 lit. “Let them acquire other children . . .”

Marked conditional sentences, those introduced with *šumma* “if,” are more frequently attested. With such sentences, there is no conjunction between the protasis and apodosis; the two clauses are simply juxtaposed (with no intervening word for “then”). In *šumma* clauses, the negative adverb in the protasis is *lā*, whereas in the apodosis it is *ul*. Mesopotamian omens and laws are invariably expressed with marked conditional clauses, as in the following:

- (50) A. šumma ina birīt martim šilum šakin – šarram ina  
 if in midst-of gall-bladder depression situated-is king in  
 pānī pilšim idukkūšu  
 front-of breach they=kill=him  
 “If a depression is situated in the middle of the gall-bladder – they will kill the  
 king in front of a breach”  
 B. šumma mārūm abāšu imtaḥaṣ – rittašu inakkisū  
 if son father=his he=has-struck hand=his they=cut-off  
 “If a son strikes his father – they will cut off his hand”

Possibly under Sumerian influence, the subordination marker is not used with marked or unmarked conditional sentences.

Eblaite also exhibits both marked and unmarked conditional clauses. Marked clauses are likewise introduced by *šumma*; however, *wa* often serves to connect the protasis and apodosis (whereas in Akkadian they are normally joined asyndetically):

- (51) su-ma INIM ḪUL al PN PN<sub>2</sub> DUG<sub>4</sub> wa NAM.KU<sub>5</sub>  
 if word evil against PN PN<sub>2</sub> speak CONJ. swear  
 “If PN<sub>2</sub> utters an evil word against PN, then he will swear . . .”

With an unmarked clause, *-ma* may or may not be used to connect the protasis and apodosis; for example:

- (52) si-a-ma      MÍ.DUG<sub>4</sub>.GA      áš-da      DUMU.NITA      DUMU.NITA      AL.TUŠ      ap  
          she=TOP.    want                              with      son                              son                              dwell      also  
                       NU.MÍ.DUG<sub>4</sub>.GA      É              EN      AL.TUŠ  
                       NEG.=want                              house    lord    dwell  
          “If she wants, she will live with her two sons; but if she does not want to,  
          she will live in the house of the lord”

## 5.12 Subordinate clauses

Akkadian subordinate clauses are traditionally grouped into three categories: (i) relative clauses; (ii) temporal clauses; and (iii) other types of clauses, including local, causal, and object clauses. In all cases, where permissible, the verb is marked with the subordination marker (see §4.2.10). The negative adverb for all subordinate clauses is *lā*.

### 5.12.1 Relative clauses

For all Assyrian and Babylonian dialects the indeclinable determinative-relative pronoun (rel.), *ša*, is used to introduce relative clauses. Only for Old Akkadian and Eblaite is the pronoun fully declinable (see §4.1.6.4); for later periods the original accusative masculine nominative form (i.e., *ša*) is used, regardless of the environment.

The relative pronoun *ša* may occur without an antecedent, in which case the clause beginning with *ša* is syntactically equivalent to a noun and thus may serve as the subject (52A), direct object (52B), or indirect object of a main clause verb:

- (53) A. *ša*      iṣṣabtū-ma                              ilikšu              ittalku  
          REL.    he=has-taken=SUBORD.=CONJ.    service=his    he=has-gone=SUBORD.  
               *šu*-ma      illak  
               he=TOP.    he=goes  
          “the one who has taken possession and performed his service obligation shall  
          be the one to continue to perform the obligation”  
       B. *ša*      īn-ka              maḥru              ana PN idin  
          REL.    eye=your    receives=SUBORD.    to      PN    give.IMPV.  
          “that which seems just to you, give to PN”  
          lit. “That which your eye receives, ...”

Most often, however, *ša* is preceded by an antecedent, as in the following:

- (54) A. *kaspum*    *ša*      PN    ilqe’ūni  
          silver      REL.    PN    he=took.SUBORD.  
          “The silver that PN took” (Old Assyrian)  
       B. *ana*      māriša      *ša*      irammu  
          for    son=her    REL.    she=loves.SUBORD.  
          “For her son whom she loves”

Since the relative pronoun must follow its antecedent noun directly, it may not be preceded by a preposition. When *ša* expresses the genitive (55A) or dative (55B), it must be resumed by the appropriate pronominal suffix:

- (55) A. *šarrūtum*    *ša*      iṣdā-ša              šuršudā  
          kingship      REL.    foundations=its    firm-are  
          “A kingship whose foundations are firm”

- B. mannum awīlum ša ṭuppaka ana maḥrišu tašpuru  
 who man REL. tablet=your toward=him you=sent=SUBORD.  
 “Who is the man to whom you sent your tablet?”

When *ša* represents the direct object of a verb, the pronominal suffix is optional; for example, compare (56A) and (56B):

- (56) A. šēnū ša šarrum iddinu  
 flocks REL. king he=gave=SUBORD.  
 “The flocks (sheep and goats) that the king gave”  
 B. ana <sup>kur</sup>Uišidiš ša Ursa ēkimu-š(u) aqṭirib  
 to U. REL. U. he=conquered=SUBORD.=it I=have-approached  
 “I approached the land of Uišidiš, which Ursa had conquered” (Standard Babylonian)

As with the genitive construction, *ša* may be deleted, the antecedent noun then appearing in the bound form (§4.1.2): for example, *awāt niqabbū ul išemme* “he does not listen to the words we say” (= *awātīm ša niqabbū ul išemme*). Such constructions, inherited from Common Semitic, are already comparatively rare in Old Akkadian and Old Babylonian, and are virtually unknown in Old Assyrian and in the late Assyrian and Babylonian dialects, with the exception of certain literary texts in Babylonian.

Relative clauses may also be verbless; in Assyrian texts the particle *-ni* is suffixed to the predicate, as in the following Old Assyrian example:

- (57) ṭuppam ša kaspum kaspī-ni ukāl  
 tablet REL. silver silver=my=SUBORD. I=have  
 “I have a tablet (that proves) that the silver [is] my silver”

The indefinite pronouns *mamman* “whoever,” *mimma* “whatever,” as well as *mala*, Ass. (*am*)*mar* (bound form of *malūm* < \**malā’um* “to become full”) “as much/many as” may all be used as relative pronouns with the omission of *ša*: for example, *ām mala ērišūki idnīm* “Give (fem. sg.) me as much grain as I requested of you.”

In Eblaite, relative clauses may be either introduced by the declineable pronoun *θa* or simply juxtaposed to their main clauses; in the latter case the morphological shape of the preceding noun (i.e., whether or not it assumes the bound form as in Akkadian) cannot be determined because of the ambiguity of the writing system:

- (58) <sup>giš</sup>GIGIR-sum ša-ti U<sub>5</sub> <sup>d</sup>Ku-ra wa <sup>d</sup>Ba-ra-ma  
 chariot=his REL. ride K. CONJ. B.  
 “(Concerning) the chariot upon which K. and B. rode”

(cf. variant <sup>giš</sup>GIGIR-sum U<sub>5</sub> <sup>d</sup>Ku-ra ù <sup>d</sup>Ba-ra-ma).

### 5.12.2 Temporal clauses

Akkadian temporal clauses are introduced by a number of subordinating conjunctions. Frequently encountered conjunctions include: *inūma* (*inu*) “when”; *ūm* “on the day that, when”; *ištu* (Standard Babylonian *ultu*) “after, as soon as, since”; *kīma* (Middle Babylonian *kī*) “as soon as, when”; *warka/i* “after”; *adi* “until, as long as, while”; *lāma* “before”; and *adi* (...) *lā* “before.” Temporal clauses usually precede their main clause; the tenses of both the main and the temporal clause verbs are determined by fairly predictable patterns. For

example, with many of the conjunctions (specifically, *inūma*, *ūm*, *ištu*, *kīma*, *warka/i*, and *adi*), when the main clause action is completed in the past (i.e., with a verb in the preterite or perfect), the temporal clause, if verbal, stands in the preterite:

- (59) mārum šū warki abūšu imūtu irgum  
 son that after father=his he=died.PRET.=SUBORD. he=brought-suit.PRET.  
 “That son brought suit after his father died”

If the main clause action is completed in the present or future (i.e., with a verb in the durative, imperative, precative, prohibitive, or verbless clause), the temporal clause, if verbal, stands in the perfect or durative. The perfect is used to mark the anteriority of the action of the temporal clause when compared to that of the main clause; for example:

- (60) inūma âm taštāmu alkam  
 when grain you=have-bought.PERF.=SUBORD. go.IMPV.=VENT.  
 “When you have bought the grain, come here”

The durative is used either when the anteriority of the temporal clause is unmarked or to express the coincidence of the two actions; thus:

- (61) inūma âm tašammu alkam  
 when grain you=buy.DUR.=SUBORD. go.IMPV.=VENT.  
 “When you buy the grain, come here”

In Middle Babylonian the conjunction *kī* (< *kīma*) “as soon as, when, after” is particularly common. When the verb of the temporal clause is in the preterite and the main clause verb is in the perfect (or preterite with negative clauses, see 4.2.1.2 and 4.2.1.3), *kī* is positioned immediately before the verb rather than at the beginning of the clause:

- (62) šēpīšu kī unakkisu itūšu  
 feet=his when he=cut-off.PRET.=SUBORD. beside=him  
 iktalāšu  
 he=has-detained.PERF.=him  
 “After he cut (the enemy’s) feet off, he kept him prisoner”

Several *kī* clauses may stand next to one another asyndetically, as in the following:

- (63) ṭēm muršīša kī iš’alūši riksa  
 report-of illness=her when he=asked.PRET.=SUBORD.=her bandaging  
 kī ēsiḥu urakkasūši  
 when he=obtained.PRET.=SUBORD. they=bind.DUR.=her  
 “After he has inquired about the report of her illness (and) obtained material for bandages, they bind her”

When the temporal clause verb is in the perfect and the main clause verb is in the present, imperative, or precative, *kī* does not necessarily stand immediately before the verb. In such clauses, the perfect is used to mark anteriority or temporal clause action in the future; often *kī* is translated as “as soon as”:

- (64) A. kī PN DUMU šipriya iktaldakku  
 when PN son-of message=my he=has-reached.PERF.=VENT.=to-you  
<sup>giš</sup>GIGIR.MEŠ liššām-ma  
 wagons may=he=carry.PREC.=VENT.=CONJ.  
 “as soon as PN, my messenger, has come to you, let him deliver the wagons”

- B. PN ana pānika kī altaprakku  
 PN to front-of=you when I=have-sent.PERF.=VENT.=to-you  
 šitālšū-ma liqbâkku  
 ask.IMPV.=him=CONJ. may=he=say.PREC.=VENT.=to-you  
 “As soon as I have sent PN to you, ask him so that he may tell you”
- C. ana mē mūti kī taktalda  
 to water.PL-of death when you=have-reached.PERF.=VENT.  
 teppuš minu  
 you=do.DUR. what?  
 “When you have reached the waters of death, what will you do then?”

When no such marking is intended, the temporal clause verb is in the present; similarly, *kī* may stand either before the verb or at the beginning of the temporal clause; thus:

- (65) kī DUMU šipriya u DUMU šiprika  
 when son-of message=my CONJ. son-of message=your  
 illaka itti aḥāmiš lilqūni  
 he=goes.DUR.=VENT. with each-other may=they=take.PREC.=VENT.  
 “If my messenger and your messenger are coming, may they deliver (it) together”

In Eblaite, temporal and main clauses may be juxtaposed without a conjunction, as in the following:

- (66) A. <sup>d</sup>A-NI-ru<sub>12</sub> U<sub>5</sub> GABA <sup>d</sup>Ku-ra DU si-in NE-na-áš<sup>ki</sup>  
 A. ride front-of K. go toward N.  
 “The god A. rides before K. (when) they go to N.”
- B. BA<sub>4</sub>.TI <sup>d</sup>Ku-ra ù <sup>d</sup>Ba-ra-ma si-in É ma-dím MU.DU  
 arrive K. CONJ. B. toward house-of death go  
<sup>d</sup>Ku-ra ù <sup>d</sup>Ba-ra-ma si-in DURU<sub>5</sub>:É<sup>ki</sup>  
 K. CONJ. B. toward chamber  
 “(When) K. and B. arrive at the mausoleum, K. and B. enter into [their] chamber”

Alternatively, the conjunction *wa* may connect a temporal clause to a main clause, as in

- (67) ZÀ.ME UD sa-ba-da-su-ma wa PAD.TÚG ba-na-a <sup>d</sup>Ku-ra wa <sup>d</sup>Ba-ra-ma  
 rite(s) day seven=his=TOP. CONJ. veil(verb) faces-of K. CONJ. B.  
 “(When) their rites which last seven days are carried out, the faces of K. and B. are veiled”

Additionally, Eblaite attests several temporal conjunctions: *a-ti* “until” (cf. Akk. *adi*); *in* UD “when” (cf. Akk. *in(a) ūm*); *ù-lu(-um)* “after” (cf. the Akk. demonstrative *ullûm* “that, distant” < \**’ullay-*): for example, NU TÚG-ZI:ZI *a-ti-ma* MU.DU É <sup>d</sup>Ku-ra “She did not put on (the ceremonial garments) when she entered the temple of K.”; 4 NITA:UDU SIKIL.SIKIL... *ma-lik-tum* NÍDBA *in* UD BA<sub>4</sub>.TI É <sup>d</sup>Ku-ra “Four pure male sheep... the queen offered in sacrifice, when she arrived at the temple of K.”; *ù-lu* BA<sub>4</sub>.TI EN ù *ma-lik-tum*... *A-ma-za-ù* NÍDBA “After the king and queen arrived... A. offered a sacrifice.”

### 5.12.3 Other subordinate clauses

The remaining subordinate clause types likewise precede the main clause; however, unlike the instance of the temporal clauses, there are no predictable patterns for the use

of tenses. The perfect, however, generally does not occur. Local clauses are usually expressed with the conjunction *ašar* (bound form of *ašrum* “place”) or *ēm(a)* “where(ever)”: for example, *imtaši ašar iwwaldu* “He forgot where he was born” (Standard Babylonian); *ēm tammarūšu šabassu* “Wherever you see him, seize him!” Causal clauses often use the conjunctions *aššum* “because,” *ana ša* “because (of the fact that),” or *ištu* “because, since”:

- (68) A. *ana ša* *lā* *ḥabbulākūšunnī-ma* *kaspam*  
 because (to+REL.) NEG. indebted=I=to-him=SUBORD.=CONJ. silver  
*ilqe'u* *šabtāšu*  
 he=took=SUBORD. seize.IMPV.=him  
 “Because I am not indebted to him and he took the silver, seize him!”  
 (Old Assyrian)
- B. *ištū-ma* *dīnam* *ušāḥizūkā-ma...* *lā*  
 because judgment I=caused-to-take=SUBORD.=you=CONJ. ... NEG.  
*tešmû* *ul* *wašrāta*  
 you=heard=SUBORD. NEG. obedient=you  
 “Because I have uttered the judgment and ... you have not listened, you are disobedient”

The conjunction *kīma*, in addition to its temporal use, often introduces object clauses (i.e., “that, the fact that” clauses); for example:

- (69) *šāpirum* *kīma* *immerī* *nēmettaka* *ana* *ekallim* *lā*  
 overseer that sheep tax=your to palace NEG.  
*tublam* *ulammidanni*  
 you=brought=VENT. he=informed=VENT.=me  
 “The overseer informed me that you had not brought the sheep, your tax, to the palace”

In Middle Babylonian the conjunction *kī* serves a variety of functions. Its position in its clause is semantically significant: in causal or object clauses it stands first, like other conjunctions, as in *kī annīta amāta iqbūni* “because they told me this matter.” Temporal clauses, however, are marked by the placement of *kī* immediately before the verb, as in *ana muššurīni kī illika* “when PN came to release us.”

## 6. LEXICON

The majority of the Akkadian lexicon is inherited from Proto-Semitic; however, there is also a great deal of Sumerian lexical interference in Akkadian as well. Although there are no comprehensive treatments of the lexical integration of Sumerian loanwords into Akkadian, recent statistical analyses suggest that one-tenth of the lexicon is borrowed from Sumerian; most of the words are nouns covering a wide semantic range. It must be noted, however, that this figure reflects the total percentage of lexical entries for the nearly three millennia of attested Akkadian and not the frequency of use or the degree of lexical integration for particular dialects. For Old Akkadian the number of Sumerian loanwords in Akkadian is quite small. A study of Sumerian loanwords in Old Babylonian Akkadian has found 529 loanwords, of which, however, 102 are found only in lexical lists (i.e., learned texts containing Akkadian words and their Sumerian counterparts) and were therefore probably not part of the spoken language. Only four of the Old Babylonian Sumerian loanwords are also attested for Old Akkadian (Edzard 1970:157, n. 2; Lieberman 1977:7).

Early Sumerian loanwords are characterized by voiceless consonants where the original Sumerian word has what, in later Sumerian at least, seems to have been a voiced consonant: for example, *ikkarum* “farmer” (< ENGAR); *laputtum* “captain” (< NU.BANDA); *parakkum* “dais” (< BARAG); *asum* “physician” (< A.ZU). This shift is due no doubt to the differences in the Sumerian and Akkadian phonological systems and, consequently, the way in which Akkadian speakers heard Sumerian phonemes. Late Sumerian loanwords, borrowed most likely after the death of Sumerian as a spoken language in the Old Babylonian period, display a one-to-one correspondence between Sumerian and Akkadian voiced and voiceless consonants: for example, *guzalû* “throne bearer” (< GU.ZA.LÁ); *agubbû* “basin for holy water” (< A.GÚB.BA); *bandudû* “bucket” (< BA.AN.DU<sub>8</sub>.DU<sub>8</sub>).

During the long period in which Akkadian is attested, Mesopotamians came into contact with many peoples, through either the filtration of various semi-nomadic groups into the region, trade and diplomacy with neighboring areas, or the conquests of the imperial periods. Thus, within the lexicon of Akkadian, loanwords and foreign phrases are attested in various periods and dialects from such diverse languages as Amorite, Egyptian, Elamite, Greek, Hittite, Hurrian, Kassite, Persian, Subarian, Urartian, as well as various West Semitic languages. With the rise of Aramaic during the first millennium and the beginning of its eventual eclipse of Akkadian as the spoken language of Mesopotamia, many Aramaic loanwords filtered into Akkadian as well.

Eblaite, similarly, derives a great deal of its lexicon from Proto-Semitic, such as the following words, here grouped semantically:

1. *Body parts*: for example, /ʔammatum/ “forearm,” /ʕaθ'mum/ “bone”
2. *Kinship terms*: for example, /ʔumm-um/ “mother,” /kallatum/ “wife”
3. *Clothing*: for example, /kusitum/ “garment,” /šaʔnā(n)/ “pair of sandals”
4. *Building*: for example, /baytum/ “house,” /libittum/ “brick”
5. *Tools and techniques*: for example, /magazzu(m)/ “blade for shearing,” /t'ahānum/ “to grind”
6. *Social organization*: for example, /maliktum/ “queen,” /mayšarum/ “justice”

However, Eblaite also shares many lexical isoglosses with Akkadian which find no parallel in the other Semitic languages: for example, /zaʔ<sub>x</sub>ārum/ “to hate” (Akk. *zêrum*), /ramānum/ “self” (Akk. *ramānum*). Eblaite also attests many isoglosses with West Semitic that are not encountered in Akkadian: for example, /bak'aru(m)/, Hebrew and Arabic *baqar* “cow”; /mabt'ah-/ , Hebrew *mibtāḥ* “confidence”; /ʔarzatum/, Hebrew, Aramaic, and Arabic *'arz* “cedar.”

Eblaite also contains a number of Sumerian loanwords, which may have made their way into Eblaite by way of Akkadian: for example, /malāxum/ (Sum. MĀ.LAḪ<sub>4</sub>, Akk. *malāḫu*) “sailor”; /melammu/ (Sum. ME.LĀM, Akk. *melammu*) “divine radiance,” as well as numerous words and names from unknown language(s).

## 7. READING LIST

The standard reference grammar of Akkadian is von Soden 1995 (3rd edition, GAG). A more concise, but nonetheless excellent, overview of Akkadian is Ungnad 1992 (English translation). Dialect-specific treatments include Aro 1955 (Middle Babylonian), Deller 1959 (Neo-Assyrian), Gelb 1961 (Old Akkadian), Hecker 1968 (Old Assyrian), Hueter 1996 (Late Babylonian), Mayer 1971 (Middle Assyrian), von Soden 1931, 1933 (Standard



Babylonian), Streck 1995 (Late Babylonian), de Vaan 1995 (Neo-Babylonian), Woodington 1982 (Neo-Babylonian).

A recent teaching grammar of (Old Babylonian) Akkadian is Huehnergard 1997. Riemschneider's grammar of Akkadian (1974, English translation) includes useful chapters outlining the distinguishing characteristics of the various dialects. Other introductory texts covering the essential grammar include Caplice 1988 and Marcus 1978; Miller and Shipp 1996 includes a useful sign list and glossary, as well as paradigms for introductory study.

Major studies of the various peripheral dialects include Adler 1976 (Mitanni), Huehnergard 1989 (Ugarit), Izre'el 1991 (Amurru), Labat 1932 (Boghaz-köi), and Wilhelm 1970 (Nuzi). Linguistically oriented treatments of Akkadian include Buccellati 1996, Gelb 1969, Groneberg 1987 (Standard Babylonian), and Reiner 1966.

There are two extensive dictionaries for Akkadian, von Soden's three-volume *Akkadisches Handwörterbuch* (1965–1981; *AHw*), which includes many attestations for each entry, but without extensive citation or translation, and *The Assyrian Dictionary of the University of Chicago* (also referred to as the *Chicago Assyrian Dictionary*, or *CAD*), an encyclopedic reference work nearing completion (to date, seventeen of the twenty-one volumes have been published).

There are three standard sign lists in common use. Labat's *Manuel d'épigraphie akkadienne* (1988, *MEA*) illustrates the diachronic development of the individual sign forms and provides their phonetic and logographic values. Borger's *Assyrisch-babylonische Zeichenliste* (1988, *ABZ*) presents the same information, although there is greater concentration on sign values and less on the evolution of the individual sign forms. Von Soden and Röllig's *Das akkadische Syllabar* (1991, *AS*) is the authoritative reference for phonetic sign values; however, logographic values and the sign forms for the individual dialects are not given.

Editions of Akkadian (and Sumerian) texts published through 1973, with cross-references to subsequent commentaries, can be found in Borger's three-volume *Handbuch der Keilschriftliteratur* (1967–1975, *HKL*). Texts published after this date can be located in the annual "Register Assyriologie" of the periodical *Archiv für Orientforschung* and in the annual "Keilschriftbibliographie" of the journal *Orientalia*.

There are no comprehensive treatments of Eblaite. Studies of the major grammatical features and classification of Eblaite include Cagni 1981, 1984, 1987; Diakonoff 1990; Edzard 1984; Fronzaroli 1984, 1990, 1992, 1996; Gelb 1981; Krebernik 1992, 1996; and Lambert 1992.

## Bibliography

- Adler, H.-P. 1976. *Das Akkadische des Königs Tušratta von Mitanni*. Kevelaer: Butzon and Bercker.
- Aro, J. 1955. *Studien zur mittelbabylonischen Grammatik*. Studia Orientalia 20. Helsinki: Societas Orientalis Fennica.
- . 1961. *Die akkadischen Infinitivkonstruktionen*. Studia Orientalia 26. Helsinki: Societas Orientalis Fennica.
- Borger, R. 1988. *Assyrisch-babylonische Zeichenliste* (AOAT 33/33A, 4th edition; *ABZ*). Neukirchen: Butzon and Bercker.
- Buccellati, G. 1996. *A Structural Grammar of Babylonian*. Wiesbaden: Otto Harrassowitz.
- Cagni, L. (ed.). 1981. *La lingua di Ebla: atti del convegno internazionale (Napoli, 21–23 aprile 1980)*. Istituto Universitario Orientale, Dipartimento di Studi Asiatici, Series Minor 14. Naples.
- . (ed.). 1984. *Il bilinguismo a Ebla: Atti del convegno internazionale (Napoli, 19–22 aprile 1982)*. Istituto Universitario Orientale, Dipartimento di Studi Asiatici, Series Minor 22. Naples.
- . (ed.). 1987. *Ebla 1975–1985: Dieci anni di studi linguistici e filologici. Atti del convegno internazionale (Napoli, 9–11 ottobre 1985)*. Istituto Universitario Orientale, Dipartimento di Studi Asiatici, Series Minor 27. Naples.



- Caplice, R. 1988. *Introduction to Akkadian* (3rd edition). Studia Pohl, Series Maior 9. Rome: Pontifical Biblical Institute.
- Conti, G. 1996. "Thèmes 'assyriens' et thèmes 'babyloniens' Ebla." In P. Zemánek (ed.), *Studies in Near Eastern Languages and Literatures. Memorial Volume of Karel Petráček*, pp. 193–202. Prague: Oriental Institute, Academy of Sciences of the Czech Republic.
- Deller, K. 1959. "Lautlehre des Assyrischen." Dissertation, Universität Wien.
- Diakonoff, I. 1990. "The importance of Ebla for the history of linguistics." In C. Gordon (ed.), *Eblaïtica: Essays on the Ebla Archives and Eblaïte Language*, vol. 2, pp. 3–31. Winona Lake, IN: Eisenbrauns.
- Edzard, D. 1970. Review of *The Chicago Assyrian Dictionary*. *Zeitschrift für Assyriologie* 60: 157–160.
- . 1984. "Zur Syntax der Ebla-Texte." *Quaderni di Semitistica* 13:101–116.
- Fronzaroli, P. (ed.). 1984. *Studies on the Language of Ebla*. Quaderni di Semitistica 13. Florence: Università di Firenze, Istituto di Linguistica e di Lingue Orientali.
- . 1990. "La langue d'Ebla." *Les Annales Archéologiques Arabes Syriennes* 40:56–63.
- . (ed.). 1992. *Literature and Literary Language at Ebla*. Quaderni di Semitistica 18. Florence: Università di Firenze, Dipartimento di Linguistica.
- . 1996. "Notes sur la syntaxe éblaïte." In J.-M. Durand (ed.), *Amurru 1: Mari, Ébla et les hourrites – dix ans de travaux*, pp. 125–34. Paris: Editions Recherche sur les Civilisations.
- Gelb, I. 1955. "Notes on von Soden's grammar of Akkadian." *Bibliotheca Orientalis* 12:93–111.
- . 1957. *Glossary of Old Akkadian*. Materials for the Assyrian Dictionary 3. Chicago: University of Chicago.
- . 1961. *Old Akkadian Writing and Grammar of Akkadian* (2nd edition). Materials for the Assyrian Dictionary 2. Chicago: University of Chicago.
- . 1969. *Sequential Reconstruction of Proto-Akkadian*. Assyriological Studies 18. Chicago: University of Chicago.
- . 1981. "Ebla and the Kish civilization." In Cagni 1981, pp. 9–73.
- Gelb, I. et al. 1956–. *The Assyrian Dictionary of the Oriental Institute of the University of Chicago*. Chicago: University of Chicago.
- Groneberg, B. 1987. *Syntax, Morphologie und Stil der jungbabylonischen "hymnischen" Literatur* (2 vols.). Freiburger Altorientalische Studien 14. Stuttgart: Steiner.
- Haayer, G. 1986. "Languages in contact: the case of Akkadian and Sumerian." In H. Vanstiphout, K. Jongeling, F. Leemhuis, et al. (eds.), *Scripta signa vocis: Studies about Scripts, Scriptures, Scribes and Languages in the Near East, presented to J. H. Hespers*, pp. 71–84. Groningen: Forsten.
- Harper, R. 1904. *The Code of Hammurabi King of Babylon about 2250 BC*. Chicago: University of Chicago.
- Hecker, K. 1968. *Grammatik der Kültepe-Texte*. Analecta Orientalia 42. Rome: Pontifical Biblical Institute.
- Hetzron, R. 1968. "Main verb markers in Northern Gura." *Africa* 38:156–172.
- Huehnergard, J. 1986. "On verbless clauses in Akkadian." *Zeitschrift für Assyriologie* 76:218–249.
- . 1987. "'Stative,' predicative, pseudo-verb." *Journal of Near Eastern Studies* 46:215–232.
- . 1989. *The Akkadian of Ugarit*. Harvard Semitic Studies 34. Atlanta: Scholars Press.
- . 1997a. *A Grammar of Akkadian*. Harvard Semitic Studies 45. Atlanta: Scholars Press.
- . 1997b. Review of *Grundriß der akkadischen Grammatik* (3rd edition), by W. von Soden and W. Mayer. *Orientalia* NS 66:434–444.
- Hueter, G. 1996. "Grammatical studies in the Akkadian dialects of Babylon and Uruk 556–500 BC." Ph.D. dissertation, University of Oxford.
- Izre'el, S. 1991. *The Akkadian Dialect of the Scribes of Amurru in the 14th–13th Centuries BC* (2 vols.). Harvard Semitic Studies 41. Atlanta: Scholars Press.
- Knudsen, E. 1980. "Stress in Akkadian." *Journal of Cuneiform Studies* 32:3–16.
- Kouwenberg, N. 1997. *Gemination in the Akkadian Verb*. Studies Semitica Neerlandica. Aasen: van Gorcum.
- Kraus, F. 1987. *Sonderformen akkadischer Parataxe: Die Koppelungen*. Mededelingen der Koninklijke Nederlandse Akademie van Wetenschappen, Afd. Letterkunde, Nieuwe Reeks, 50/1. Amsterdam: North-Holland Publishing.
- Krebernik, M. 1982–1983. "Zu Syllabar und Orthographie der lexikalischen Texte aus Ebla." *Zeitschrift für Assyriologie* 72:178–236; 73:1–47.

- . 1992. "Mesopotamian myths at Ebla: ARET 5, 6 and ARET 5, 7." In Fronzaroli 1992, pp. 41–62.
- . 1996. "The linguistic classification of Eblaite: methods, problems, and results." In J. Cooper and G. Schwartz (eds.), *The Study of the Ancient Near East in the Twenty-First Century*, pp. 233–251. Winona Lake, IN: Eisenbrauns.
- Labat, R. 1932. *L'Akkadian de Boghaz-köi. Étude sur langue des lettres, traités et vocabulaire trouvés à Boghaz-Köi*. Bordeaux: Librairie Delmas.
- . 1988. *Manuel d'épigraphie akkadienne* (6th edition by F. Malbran-Labat; MEA). Paris: Paul Geuthner.
- Lambert, W. 1992. "The language of ARET V, 6 and 7." In Fronzaroli 1992, pp. 41–62.
- Lieberman, S. 1977. *The Sumerian Loanwords in Old-Babylonian Akkadian. Volume 1: Prolegomena and Evidence*. Missoula, MT: Scholars Press.
- . 1979. "The phoneme /o/ in Sumerian." In M. Powell, Jr. and R. Sack (eds.), *Studies in Honor of Tom B. Jones*, pp. 21–28. Neukirchen-Vluyn: Butzon and Bercker.
- Marcus, D. 1978. *A Manual of Akkadian*. Washington, DC: University Press of America.
- Mayer, W. 1971. *Untersuchungen zur Grammatik des Mittelassyrischen*. AOATS 2. Rome: Pontifical Biblical Institute.
- Miller, D. and R. Shipp. 1996. *An Akkadian Handbook*. Winona Lake, IN: Eisenbrauns.
- Patterson, R. 1970. "Old Babylonian parataxis as exhibited in the royal letters of the Middle Old Babylonian period and in the Code of Hammurapi." Ph.D. dissertation, University of California, Los Angeles.
- Rainey, A. 1976. "Enclitic *-ma* and the logical predicate in Old Babylonian." *IOS* 6:51–58.
- Reiner, E. 1966. *A Linguistic Analysis of Akkadian*. The Hague: Mouton.
- Riemschneider, K. 1974. *An Akkadian Grammar*. Translated by T. Caldwell, J. N. Oswalt, J. F. Sheehan, et al. Milwaukee: Marquette University Press.
- Soden, W. von. 1931, 1933. "Der hymnisch-epische Dialekt des Akkadischen." *Zeitschrift für Assyriologie* 40:163–227; 41:90–183.
- . 1965–81. *Akkadisches Handwörterbuch* (3 vols.). Wiesbaden: Otto Harrassowitz.
- . 1995. *Grundriss der akkadischen Grammatik* (3rd edition). *Analecta Orientalia* 33. Rome: Pontifical Biblical Institute.
- Soden, W. von and W. Röllig. 1991. *Das akkadische Syllabar* (AnOr 42, 4th edition; AS). Rome: Pontifical Biblical Institute.
- Streck, M. 1994. "Funktionsanalyse des akkadischen Št<sub>2</sub>-Stamms." *Zeitschrift für Assyriologie* 84: 161–197.
- . 1995. *Zahl und Zeit: Grammatik der Numeralia und des Verbal Systems im Spätbabylonischen*. Cuneiform Monographs 5. Groningen: Styx.
- Ungnad, A. 1992. *Akkadian Grammar* (5th edition by L. Matouš). Translated by H. Hoffner, Jr. SBL Resources for Biblical Study 30. Atlanta: Scholars Press.
- Vaan, J. de. 1995. "*Ich bin eine Schwertklinge des Königs*": *Die Sprache des Bēl-ibni*. AOAT 242. Neukirchen-Vluyn: Butzon and Bercker.
- Westenholz, A. 1991. "The phoneme /o/ in Akkadian." *Zeitschrift für Assyriologie* 81:10–19.
- Wilhelm, G. 1970. *Untersuchungen zum Hurro-Akkadischen von Nuzi*. Kevelaer: Butzon and Bercker.
- Woodington, N. 1982. "A Grammar of the Neo-Babylonian Letters of the Kuyunjik Collection." Ph.D. dissertation, Yale. Ann Arbor, MI: University Microfilms.

# The cuneiform script

𒀭	aš, dil, rum, rù	𒀭	la
𒀬	ḫal	𒀭	bin, pin; <sup>(giš)</sup> APIN = <i>epennu</i> “plow”
𒀭	mug/k/q, b/puk	𒀭	maḫ
𒀭	ba, pá	𒀭	tu, tú
𒀭	zu, šú	𒀭	li, le
𒀭	su, kuš; SU = <i>zumru</i> “body”; KUŠ = <i>mašku</i> “skin, hide”	𒀭	kúr, bab/p, pap; KÚR = <i>aḫû</i> “strange, foreign, hostile,” <i>nakāru</i> “to be hostile”
𒀭	rug/k/q, šin, šun	𒀭	mu; MU = <i>nīšu</i> “life,” <i>šattu</i> “year,” <i>šumu</i> “name”
𒀭	bal, pal; BALA = <i>palû</i> “reign”	𒀭	qa; SILA <sub>3</sub> = <i>qû</i> (unit of capacity)
𒀭	ád/t/t, gir	𒀭	kád/t, šíd
𒀭	búl, púl	𒀭	kád/t
𒀭	tar, ṭar, tír, ṭír, kud/t, qud/t, ḫaz/s/š, ḫaš, sil, šil; SILA = <i>sūqu</i> “street”	𒀭	gil, kíl, qíl
𒀭	an, il; AN = <i>šamû</i> “sky, heaven”; DINGIR = <i>ilu</i> “god” (also determinative before deities)	𒀭	ru, šub/p
𒀭	ka, qà; KA = <i>pû</i> “mouth”	𒀭	be, pè, bad/t/t, til, mid/t/t, ziz/s; BE = <i>šumma</i> “if”
𒀭	nag/k/q; NAG = <i>šatû</i> “to drink”	𒀭	na
𒀭	KÚ = <i>akālu</i> “to eat”	𒀭	šir
𒀭	rí, ré, iri <sub>4</sub> ; ere <sub>4</sub> ; URU = <i>ālu</i> “town, city”	𒀭	k/qul; NUMUN = <i>zēru</i> “seed,” “progeny”
𒀭	ÌR = <i>ardu</i> “slave”	𒀭	ti, ṭi
𒀭	ITI = <i>arḫu</i> “mouth” (also determinative before names of months)	𒀭	bar, pár, maš; MAŠ = <i>mišlum</i> “half, middle”; <i>šumma</i> “if”; MAŠ.GAG.EN or MAŠ.EN.GAG = <i>muškēnu</i> “dependent, commoner”
𒀭	šaḫ, šiḫ; ŠAH = <i>šaḫû</i> “pig”	𒀭	nu

Tables © Ecological Linguistics 2002. All rights reserved. Used by permission.

𒄠	MÁŠ = <i>šibtu</i> “interest,” <i>puḫādu</i> “lamb”; <sup>(ú)</sup> MÁŠ.ŠU.GÍD = <i>bārû</i> “diviner, haruspex”	𒂗	en; EN = <i>bēlu</i> “lord”
𒄠𒄡𒄢	kun	𒄠𒄡	ṭar
𒄠	ḫu, pag/k/q; MUŠEN = <i>iššūru</i> “bird” (also determinative following names of birds)	𒄡	šur
𒄠𒄡𒄢	nam	𒄡𒄢	suḫ
𒄠𒄡	ig/k/q, eg/k/q	𒄡𒄠	mùš; <sup>d</sup> INANA (deity, Sum. Inana, Akk. Ištar)
𒄠𒄡𒄢	mud/t/ṭ	𒄡𒄢	sa
𒄠𒄢	rad/t/ṭ	𒄡𒄢	gán, kán, gà; GANA <sub>2</sub> = <i>eqlu</i> “field”; IKU = <i>ikû</i> (surface measure)
𒄠𒄢𒄢	zi, ze, sí, šé, šé	𒄡𒄢	kár
𒄠𒄢𒄢	gi, ge	𒄡𒄢	gú, tik/q; GÚ = <i>kišādu</i> “neck”; GÚ.UN (or GUN) = <i>biltu</i> “weight, tribute, load”
𒄠𒄢𒄢	ri, re, dal, tal ṭal	𒄡𒄢𒄢	dur, ṭur
𒄠𒄢𒄢	nun, zil, šil; NUN = <i>rubû</i> “prince”	𒄡𒄢𒄢𒄢	GUN (or GÚ.UN) = <i>biltu</i> “weight, tribute, load”
𒄠𒄢𒄢	gáb/p, kab/p, qáb/p ḫúb/p	𒄡𒄢𒄢	lâl; LÂL = <i>dišpu</i> “honey”
𒄠𒄢𒄢𒄢	ḫub/p	𒄡𒄢	gur, qur; GUR = <i>kurru</i> (measure of capacity)
𒄠𒄢	kad/t/ṭ, qād/t; GADA = <i>kitû</i> “linen”; NA.GADA “shepherd”	𒄡𒄢	si, se
𒄠𒄢𒄢	dim, tim	𒄡𒄢𒄢	dar, tár, ṭar
𒄠𒄢𒄢	mun	𒄡𒄢𒄢𒄢	sag/k/q, šag/k/q, riš, ris, res; SAG = <i>rēšu</i> , “head”; SAG.İR = <i>ardu</i> “male slave”; SAG.GEME <sub>2</sub> = <i>amtu</i> “female slave”
𒄠𒄢𒄢	ag/k/q	𒄡𒄢𒄢	má; <sup>(giš)</sup> MÁ = <i>eleppu</i> “boat”
𒄠𒄢𒄢	MÈ = <i>tāḫāzu</i> “battle”	𒄡𒄢𒄢𒄢	dir, ṭir
𒄠𒄢	tab/p, ṭab/p, dáb/p; TAB.BA = <i>tappû</i> “business associate, partner”	𒄡𒄢𒄢	in
𒄠𒄢	šum, tag/k/q	𒄡𒄢𒄢	rab/p
𒄠𒄢	ab/p	𒄡𒄢𒄢𒄢	šar; LUGAL = <i>šarru</i> “king”
𒄠𒄢	nab/p	𒄡𒄢𒄢	šir, ḫir, sar, šar; KIRI <sub>6</sub> = <i>kirû</i> “garden, orchard”
𒄠𒄢𒄢	mul	𒄡𒄢𒄢	bât; BÂD = <i>dûru</i> “wall”
𒄠𒄢𒄢	ug/k/q	𒄡𒄢	si, sè
𒄠𒄢𒄢	az/s/š		











	URUDU = <i>erû</i> “copper, bronze”		kas, raš/s; KASKAL = <i>ḥarrānu</i> “road, path, journey”
	ká; KÁ = <i>bābu</i> “gate, opening”		
	um		gab/p, qab/p; GABA = <i>irtu</i> “chest, breast”
	dub/p, tub/p, ṭup; DUB = <i>ṭuppu</i> “tablet”; DUB.SAR = <i>ṭupšarru</i> “scribe”		duḥ, taḥ, ṭuḥ
	dá, ta ṭá		ru <sub>6</sub> ; EDIN = <i>ṣēru</i> “plain, steppe”
	i		daḥ, taḥ, ṭaḥ
	ia		
	gan, kan, kám (also determinative following numbers)		am; AM = <i>rīmu</i> “wild bull”
	tur, ṭur; DUMU = <i>māru</i> “son”; DUMU.MUNUŠ = <i>mārtu</i> “daughter”; TUR = <i>ṣeḥēru</i> “to be small, young”		šir <sub>4</sub> ; UZU = <i>šīru</i> “meat, flesh” (also determinative before body parts)
	ad/t/ṭ; AD = <i>abu</i> “father”		ne, ṭe, bil, pil, kúm, bí; IZI = <i>išātu</i> “fire”
	ši, še, zí, zé		bíl, píl
	šàm (variant of šám)		NA <sub>4</sub> = <i>abnu</i> “stone” (also determinative before stone objects)
	ram		kak, qaq
	šám; SA <sub>10</sub> = <i>šāmu</i> “to buy, purchase”		ni, né, zal, šal, lí, lé, ì; Ì (or Ì.GIŠ) = <i>šamnu</i> “oil fat”
	zik/q		ir er
	gum, kum, qum, qu		mal, gá, mà
	gaz, gaš		DAGAL = <i>rapāšu</i> “to be wide, large”; AMA = <i>ummu</i> “mother”
	SUḤUŠ = <i>išdu</i> “base, foundation”		SILA <sub>4</sub> = <i>puḥādu</i> “lamb”
	kas <sub>4</sub> ; <sup>lú</sup> KAS <sub>4</sub> = <i>lasīmu</i> “courier”		ùr
	úr; ÚR = <i>sūnu</i> “lap,” <i>pēmu</i> “thigh”		dag/k/q, tág/k/q, ṭak
	il, él		
	du, ṭu, gub/p, kub/p, qub/p; DU = <i>alāku</i> “to go”; GUB = <i>uzuzzu</i> “to stand”		
	dum, tum, ṭum, tu <sub>4</sub>		pa, ḥad/t/ṭ; UGULA = <i>aklu</i> “overseer, inspector”
	ANŠE = <i>imēru</i> “donkey”		šab/p, sab/p
	EGIR = <i>arki</i> “behind, in back of, after”		síp; <sup>(lú)</sup> SIPA = <i>rē'û</i> “shepherd”
	GEŠTIN = <i>karānu</i> “wine”		
	uš, nid/t/ṭ		

𒀭	iš, íz/s/š, mil; SAĦAR = <i>eperu</i> “earth, dust, soil”	𒄩	giš, iz/s/š, ez/s/š; GIŠ = <i>išu</i> “wood” (also determinative before wooden objects)
𒀬	bi, bé, pí, kaš; KAŠ = <i>šikaru</i> “beer”	𒄩𒀭	GUD = <i>alpu</i> “ox, bull”
𒀭𒀬𒀭	šim, rig/k/q	𒄩𒀭	al
𒀭𒀬	kib/p, qib/p	𒄩𒀭	ub/p, ár
𒀭𒀬	mar	𒄩𒀭	gàr, qar
𒀭𒀬	e	𒄩𒀭𒀬	id/t/š; ed/t/š; Á = <i>idu</i> “side, arm, strength”
𒀭𒀬𒀭	dug/k/q, lud/t/š; DUG = <i>karpātu</i> “pot, container” (also determinative before vessels)	𒄩𒀭𒀬	lil
𒀭𒀬	un; UN = <i>nišū</i> “people”; KALAM = <i>mātu</i> “land, country”	𒄩𒀭𒀬	MURUB <sub>4</sub> = <i>qablu</i> “hip, waist, middle”
𒀭𒀬	gid/t/š, kid/t/š, qid/t, saḥ, líl	𒄩𒀭𒀬	ṭes; <sup>(lú)</sup> SIMUG = <i>nappāḥu</i> “smith, metal worker”
𒀭𒀬	šid/t/š, lag/k/q	𒄩𒀭	áš
𒀭𒀬	rid/t/š, mis; KIŠIB = <i>kunukku</i> “cylinder seal”	𒄩𒀭	ma
𒀭𒀬	ú, šam; Ú = <i>šammu</i> “grass, herb, plant” (also determinative before plants)	𒄩𒀭	gal, qal; GAL = <i>rabû</i> “great”
𒀭𒀬𒀭	ga, kà, qá	𒄩𒀭	BARAG = <i>parakku</i> “cult, dais, sanctuary”
𒀭𒀬	luḥ, làḥ, liḥ, raḥ, riḥ	𒄩𒀭	gir, kir, qir, biš, piš
𒀭𒀬	kal, dan, tan, rib/p, lab/p	𒄩𒀭𒀬	mir; AGA = <i>agû</i> “crown”; NIMGIR = <i>nāgīru</i> “herald”
𒀭𒀬𒀬	bid/t/š, pid/t; É = <i>bītu</i> “house, temple”	𒄩𒀭	bur, pur
𒀭𒀬𒀬	nir	𒄩𒀭𒀬	BALAG = <i>balaggu</i> “a musical instrument (drum)”
		𒄩𒀭	ša
𒀭𒀬𒀬	gi <sub>4</sub> , ge <sub>4</sub>	𒄩𒀭	šu, qad/t; ŠU = <i>qātu</i> “hand”
𒀭𒀬	ra	𒄩𒀭	lul, lib/p, lup, nar
𒀭𒀬	LÚ = <i>awīlu</i> “man” (also determinative before male professions)	𒄩𒀭𒀬	sa <sub>6</sub> ; GIŠIMMAR = <i>gišimmaru</i> “date-palm”
𒀭𒀬	šiš, sis, siš; ŠEŠ = <i>aḥu</i> “brother”	𒄩𒀭𒀬	ALAN = <i>šalmu</i> “statue”
𒀭𒀬	zag/k/q; ZAG = <i>idu</i> “side, border”	𒄩𒀭𒀬	URI = <i>Akkadû</i> “Akkadian”
𒀭𒀬	gam	𒄩𒀭	zib/p, šib/p, šip

𒌦	kur, mad/t/ṭ, nad/t, lad/t/ṭ, šad/t/ṭ, sad/t/ṭ; KUR = <i>šadû</i> “mountain,” <i>mātu</i> “country, land”		
𒌷	še; ŠE = <i>ûm</i> “barley, grain” (also determinative before grains)	𒌷	ḫi, ḫe
𒌵	bu, pu, sîr, šîr, gíd/t/ṭ, qíd/t, šúd	𒌵	a', i', e', u', 'a, 'i, 'e, 'u
𒌶	uz/s/š	𒌶	aḫ, iḫ, eḫ, uḫ
𒌷𒌶	šud/t/ṭ, sir, sù	𒌷𒌶	kam (also determinative following numbers)
𒌷𒌶𒌶	muš, šir	𒌷𒌶𒌶	im, em
𒌷𒌶𒌶𒌶	tir	𒌷𒌶𒌶𒌶	bir, pîr
𒌷𒌶	te, ṭe <sub>4</sub> , de <sub>4</sub> ; TE = <i>ṭeḫû</i> “to approach”	𒌷𒌶	ḫur, ḫar, mur; gišHUR = <i>uṣurtu</i> “design, plan”
𒌷𒌶𒌶	kar	𒌷𒌶𒌶	ḫuš
𒌷	liš, lis		
𒌷	u <sub>4</sub> , ud/t/ṭ, tam, tú, par, pir, liḫ, ḫiṣ; UD/U <sub>4</sub> = <i>ûmu</i> “day”; <sup>d</sup> UTU = Sum. Utu, Akk. Šamaš (deity)		
𒌷𒌶	pi, pe, tál; GEŠTU = <i>uznu</i> “ear, wisdom, understanding”		
𒌷𒌶𒌶	lib, lip; ŠAG <sub>4</sub> = <i>libbu</i> “heart, mind, thought, inside”		
𒌷𒌶	šab/p, zab/p; ERIN <sub>2</sub> = <i>šābu</i> “gang, army, troops”		
𒌷	u	𒌷𒌶	ši, lim; IGI = <i>īnu</i> “eye”
𒌷𒌶𒌶	muḫ; UGU = <i>muḫḫu</i> “skull, top”; <i>eli</i> “on, upon, over, above”	𒌷𒌶𒌶	ar
𒌷𒌶	lid/t/ṭ; ÁB = <i>arḫu</i> “cow”	𒌷𒌶𒌶	SIG <sub>5</sub> = <i>damāqu</i> “to be good, favorable”
𒌷𒌶	kiš, kis, qiš, qis	𒌷𒌶𒌶	ù
𒌷𒌶	mi, mé, šil, gi <sub>6</sub> ; GI <sub>6</sub> = <i>mušītu</i> “night, nighttime”	𒌷𒌶𒌶	ḫul
𒌷𒌶	gul, qúl, sún	𒌷𒌶𒌶	di, de, ṭi, ṭe, sá; DI = <i>dīnu</i> “decision, judgment”; DI.KUD = <i>dānu</i> “to judge”
𒌷𒌶	nim, num, nù, tum <sub>4</sub>	𒌷𒌶𒌶	dul, tul
𒌷𒌶	lam	𒌷𒌶𒌶	ki, ke, qí, qé; KI = <i>eršetu</i> “earth, land, district” (also determinative following names of countries)
𒌷𒌶	zur, šur	𒌷𒌶𒌶	din, tin

	pan, ban		dun, šul
	gim, kim, qim, tím		KUG = <i>ellu</i> “pure”; KUG.SIG <sub>17</sub> = <i>hurāšu</i> “gold”; KUG.BABBAR = <i>kaspu</i> “silver”
	ul		pad/t/ṭ, šug/k/q
	GÌR = <i>šēpu</i> “foot”	«	man, mìn, niš
		««	eš, sin
		┘	diš, tiš, ṭiš, tiz (also determinative before male proper names)
		┘┘	lal, lá
	kil, qil, rim, ḥab/p		šal, sal, rag/k/q, mán, mín; MUNUS = <i>sinništu</i> “woman” (also determinative before female proper names and occupations)
	ENGUR = <i>apsû</i> “abyss, subterranean ocean”		zum, súp, šum, šu, ríḡ/k/q
	(ḡiš)GIGIR = <i>narkabtu</i> “chariot”		nin; NIN = <i>aḥātu</i> “sister,” <i>bēltu</i> “lady, mistress”
	zar, šar		dam, ṭam; DAM = <i>mutu</i> “husband,” <i>aššatum</i> “wife”; DAM.GÀR = <i>tamkāru</i> “merchant”
	ùʾ		GEME <sub>2</sub> = <i>amtu</i> “female slave”
	bul, pul		gu, qù
	sug/k/q		NAGAR = <i>nagāru</i> “carpenter”
	NENNI = <i>annanna</i> “so-and-so, such-and-such”		nig/k/q
┘	me, mì, šib/p, sib/p		el, il <sub>5</sub>
┘┘	meš (also a marker of plurality following logograms)		lum, ḥum
	ib/p, eb/p		SIG <sub>4</sub> = <i>libittu</i> “(mud) brick”
	ku, qú, dúr, tuš; TÚG = <i>šubātu</i> “garment” (also determinative before garments)		dúk, tug/k/q
	lu; UDU = <i>imмерu</i> “sheep”		ur, lig/k/q, daš, das, taš, tas, tíz, tís, tíš
	dib/p, tib/p, ṭib/p, dab/p		a; A = <i>mú</i> “water”



	kin, qin, qi, qe; KIN = <i>šipru</i> “message, work, labor”		za, sà, ʒa
	šík, šíq; SÍG = <i>šīpātu</i> “wool” (also determinative before objects made of wool or types of wool)		ḫa, ku <sub>6</sub> ; KU <sub>6</sub> = <i>nūnu</i> “fish” (also determinative following names of fish)
	ERIN = <i>erēnu</i> “cedar”		sig/k/q, šik/q
	šú		tu
	ÉN = <i>šiptu</i> “incantation”		šá, níg/k/q, gar; NINDA = <i>akalu</i> “bread, food”

# Ugaritic

DENNIS PARDEE

## 1. HISTORICAL AND CULTURAL CONTEXTS

Ugaritic is the only well-attested example known today of the West Semitic languages spoken in the Levantine area in the second millennium BC. The position of Ugaritic among the Semitic languages has been a matter of dispute, in part because of a confusion of categories, namely between literary and linguistic criteria. Literarily, the poetic texts show strong formal (poetic parallelism), lexical, and thematic affinities with Biblical Hebrew poetry. Linguistically, however, Ugaritic is considerably more archaic than any of the well-attested Northwest Semitic languages, and probably descends directly from a Levantine “Amorite” dialect. All indications are that it is not more directly related to East Semitic (Akkadian) than to West Semitic. Within the latter branch, it shares certain important isoglosses with Northwest Semitic as opposed to Arabic (e.g., roots  $Iw \rightarrow Iy$ ) and with Canaanite as opposed to Aramaic (e.g.,  $/d/ \rightarrow /s/$ ). The isoglosses shared with Arabic (e.g., consonantal inventory) represent for the most part features commonly inherited from Proto-Semitic.

Ugaritic is a one-period language, attested only for the last part of the Late Bronze Age, approximately 1300–1190 BC. This is because the writing system in which known Ugaritic texts are inscribed was not invented (at least according to present data) until the early thirteenth century, whereas the city of Ugarit – virtually the only site where Ugaritic texts have been discovered – was destroyed early in the twelfth century. In recent years it has become clearer that the greatest number of texts date from the last few decades of the site and there is, therefore, no basis on which to define a “late” Ugaritic over against the main body of texts (contra Tropper 1993b), for the main body of texts *is* late Ugaritic. The only clear strata of the language are the poetic dialect in which most mythological texts are written and the prose dialect used for everyday communication and administration.

Virtually all Ugaritic texts have been discovered at the site of the ancient city of Ugarit, modern Ras Shamra, excavated by the French more or less continuously since 1929 (Yon 1997). The site had been inhabited since the Neolithic period (Contenson 1992), but texts are presently attested only for the Late Bronze Age; the Middle Bronze levels, where finds of Akkadian texts are to be expected, have hardly been penetrated. In recent years, Ugaritic texts have been discovered at neighboring Ras Ibn Hani, a suburb of Ugarit (Bordreuil *et al.* 1987). From rare mentions of Ugarit in texts from other sites (Mari, el-Amarna), it is clear that the inhabitants of the city were of so-called Amorite stock, for they bear Amorite names and maintained cultural relations with the other Amorite kingdoms of the eighteenth century BC.

The area under the control of Ugarit was limited on the north and east by important natural boundaries (the Jebel al-Aqra<sup>c</sup> on the north and the Jebel Ansariyeh on the east), with occasional control of areas bordering these boundaries (e.g., southern portions of the

state of Mukish to the north). The southern boundary was at the southern extremity of the Gebleh Plain, and also varied (e.g., at times including the kingdom of Siyannu). The average territory may have been approximately 2,000 sq. km. (Saadé 1979:33).

There are approximately 50 mythological texts in poetry and some 1,500 texts in prose (including decipherable fragments). The primary types of prose texts are (i) religious (ritual, pantheon, votive); (ii) ominological (astral, malformed births, extispicy); (iii) medical (hippiatric); (iv) epistolary; (v) administrative (contracts, lists of many sorts); and (vi) didactic (abecedaries, exercises).

The prose texts originated largely from the palace administration of the city of Ugarit. The administration was headed by a king, often in vassal position to a king of a larger political entity, particularly the Hittite king in the period documented. Many of the letters emanate directly from the royal family; many of the ritual texts specifically mention the king; most of the administrative texts deal with one aspect or another of royal control of the resources of the kingdom (real estate, taxes, management of royal goods, working of royal raw materials, etc.). The hundred-plus epistolary documents, in particular, reveal the Ugaritic that was in everyday use in the city.

Because they provide a mythical and literary background to the Hebrew Bible, the poetic texts have made Ugarit famous. They are, however, comparatively few in number and the poetic dialect presents many difficulties of interpretation. Several of the tablets bearing the major mythological texts are signed by a scribe named Ilimilku who some now suspect may have lived near the end of the kingdom of Ugarit, rather than nearly a century earlier, the generally accepted view (Pardee 1997:241 note 3). The poems that he and other scribes wrote down had in all likelihood been passed down by oral tradition for centuries.

The nature of the corpus and of the writing system places limits on our ability to describe the language. The number of texts is relatively small and virtually all are damaged to some degree, leaving few long stretches of text for analysis. This is especially true of the prose texts, which were usually written on tablets smaller than those bearing the major mythological texts. There are no prose narrative texts as yet from which to derive a narrative prose syntax. The poetic texts are largely narrative rather than lyrical, but are of little use, because of their archaic form, for projecting a prose syntax. The upshot is that phonology is described largely in terms of graphemes; morphology is to a significant degree reconstructed; reasonably comprehensive descriptions of morphosyntax and of poetic syntax are possible; the prose discourse syntax particular to letters is reasonably well known, while narrative prose syntax is known primarily from narrative sections of letters.

The Ugaritic language was only one of at least eight languages (and/or writing systems) in use at Ugarit. The one other Semitic language attested is Akkadian, the international lingua franca of the time, in which approximately 2,000 texts are written in syllabic cuneiform, primarily epistolary, legal, and administrative. A number of texts have also been found in Sumerian, Hittite (written in standard syllabic cuneiform and in hieroglyphic), Egyptian, Hurrian (written in Ugaritic consonantal cuneiform and in standard syllabic cuneiform), and Cypro-Minoan (not fully deciphered).

## 2. WRITING SYSTEMS

### 2.1 The consonant alphabets

The Ugaritic writing system is unique in that it adapts the cuneiform principle (wedges inscribed in clay) to represent graphemes of a consonantal type for the purpose of writing a

West Semitic language. The Semitic consonantal writing system had been devised some two to four centuries before the earliest attested Ugaritic texts, and there is no particular reason to believe that it was not in use at Ugarit before the invention of the Ugaritic cuneiform characters. Indeed, it is not unlikely that the cuneiform system is a representation in clay of a linear alphabet (i.e., one written with ink), though presently available data do not allow a precise description of the origin of the cuneiform alphabet of Ugarit.

At present, three consonantal systems are attested at Ugarit: (i) the *long alphabet*, well attested by abecedaries; (ii) the *short alphabet*, very rarely attested and of uncertain composition (no abecedy has yet been discovered representing this script); (iii) a South Semitic type alphabet, presently attested at Ugarit by a single abecedy (RS 88.2215), showing South Arabian character order (i.e., *h, l, h, m...*), very similar to an abecedy discovered in 1933 at Beth-Shemesh in Palestine but only recently deciphered (bibliography in Bordreuil-Pardee 1995b; 2001, text 32).

The long alphabet was clearly intended for writing Ugaritic, for virtually all texts, whether in prose, in poetry, or of a didactic nature, are written in it. The short alphabet shows merging of phonemes (and thus graphemes) on the Phoenician model (e.g., /š/ and /t/ written *t*), and the few texts in consonantal cuneiform discovered beyond the borders of Ugarit appear to be written in variants of the alphabet script (Dietrich and Loretz 1988; cf. Bordreuil 1981). It seems, therefore, to be an adaptation of the long alphabet to a Phoenician-type consonantal repertory. The language of at least one text written in this system has been identified as Phoenician (Greenstein 1976; Bordreuil 1979). Though the abecedy in South Arabian order consists of the same number of signs as the basic consonantal repertory of the long alphabet, it shows several variant sign forms and was not, therefore, a simple reorganization of the Ugaritic script along South Arabian lines. Because only abecedaries are attested in this version of the script, one can only speculate as to the language that it was used to convey.

Several examples of the (long) consonantal alphabet written out partially or in full (i.e., abecedaries) provide our oldest witnesses to the concept of a repertory of consonants existing in a fixed order. The Ugaritic abecedy consists of twenty-seven symbols denoting the consonants of the language, plus an additional three characters appended to the end. The Ugaritic symbols follow the order customary for the later Northwest Semitic alphabets, which, however, contain only twenty-two signs:

### Semitic abecedaries

Northwest Semitic																										
ʾ	b	g		d	h	w	z	ḥ	ṭ	y	k		l	m	n		s	ʿ	p	ṣ	q	r	š		t	
Ugaritic																										
ʾ	b	g	ḥ	d	h	w	z	ḥ	ṭ	y	k	š	l	m	<u>d</u>	n	ṣ	ʿ	p	ṣ	q	r	<u>t</u>	ḡ	t	
ī	ū	š																								

The five extra signs of Ugaritic (*ḥ, š, d, ṣ, ḡ*) are dispersed at apparent random within the order, seemingly suggesting the invention of the Northwest Semitic alphabet for a language, such as Ugaritic, which had a larger consonantal inventory than those of the well-known first-millennium languages.

The origin of the three additional signs (*ī, ū, š*) appended to the end of the abecedy is in dispute. The patent similarity of *form* between the Ugaritic symbol transliterated *š*, and the *s*-character of the later Northwest Semitic script makes a common origin likely, but the reason for the addition of this sign to the Ugaritic alphabet is unclear (compare Segert

**Table 9.1 The Ugaritic cuneiform consonantal script**

Character	Transcription	Character	Transcription
	<i>a</i>		<i>d</i>
	<i>b</i>		<i>n</i>
	<i>g</i>		<i>z</i>
	<i>h</i>		<i>s</i>
	<i>d</i>		<i>c</i>
	<i>h</i>		<i>p</i>
	<i>w</i>		<i>s</i>
	<i>z</i>		<i>q</i>
	<i>h</i>		<i>r</i>
	<i>t</i>		<i>t</i>
	<i>y</i>		<i>g</i>
	<i>k</i>		<i>t</i>
	<i>s</i>		<i>i</i>
	<i>l</i>		<i>u</i>
	<i>m</i>		<i>s</i>

1983:201–218; Dietrich and Loretz 1988). In *function*, *š* is like Ugaritic *s*, but only in certain words – other *s*-words are never written with *š*.

## 2.2 The syllabic characters

The typification of the Ugaritic script as “consonantal” requires some qualification. The initial character *a* and the two “supplemental” characters *i* and *u* function as *syllabic* symbols, having the CV value of glottal stop plus the vowel *a*, *i*, or *u*. The reason for the presence of these syllabic *alif* (the name of the Northwest Semitic character for the glottal stop) signs is uncertain (perhaps they were added for the purpose of writing a language such as Akkadian, which permits syllables to begin with vowels; Akkadian texts written with the Ugaritic script have been found, but they are rare). To represent a syllable-final glottal stop, *i* is used. The situation presents difficulties, however, for a syllable-final glottal stop seems sometimes to quiesce, sometimes to be followed by a very brief vowel (compare “secondary opening” in Biblical Hebrew). See Verreert 1983:223–258; another hypothesis is proposed by Tropper 1990b.

## 3. PHONOLOGY

### 3.1 Consonants

The Ugaritic consonantal system is typically described in terms of *graphemes* rather than in phonetic terms. By comparison with the later West Semitic languages, and in comparison with other contemporary languages (Akkadian, Egyptian, Hurrian), however, the phonetic system can be approximated (see Tropper 1994a; Gordon 1997):

## Ugaritic obstruents

	<i>Bilabial</i>	<i>Inter-dental</i>	<i>Dental</i>	<i>Palato-alveolar</i>	<i>Velar</i>	<i>Pharyngeal</i>	<i>Glottal</i>
<i>Stops</i>							
<i>Voiceless</i>	p		t		k		ʔ (/ʔ/)
<i>Voiced</i>	b		d		g		
<i>Emphatic</i>			ṭ (/t'/)		q (/k'/)		
<i>Fricatives</i>							
<i>Voiceless</i>		ṯ (/θ/)	s	š	ḫ (/x/)	ḥ (/ħ/)	h
<i>Voiced</i>		ḏ (/ð/)	z	ḡ (/ɣ/)		ʕ (/ʕ/)	
<i>Emphatic</i>		ẓ (/ð'/)	ṣ (/s'/)				

The fricative transcribed š may be lateral fricative /ɬ/ instead.

In addition, the following sonorants occur:

	<i>Bilabial</i>	<i>Dental</i>	<i>Palatal</i>
<i>Nasals</i>	m	n	
<i>Liquids</i>		r, l	
<i>Glides</i>	w		y

In comparison with Arabic, Ugaritic had one fewer consonantal phoneme, there being no sign for \*q, which had shifted to š. The Ugaritic writing system made no distinction between š and ś. Indeed, there being no evidence from graphic confusions within Ugaritic for the survival of \*ś (unlike Hebrew), it appears likely that it had merged with /š/ (Blau 1977:106; Tropper 1994a:29–30).

The graphic system does not correspond precisely to the phonetic one. The symbol ẓ is used for etymological ṯ (/θ'/), but certain words containing etymological ṯ are regularly written with symbol ḡ (e.g., *nḡr* “guard” from the root *NZR*), probably expressing a phonetic shift, itself reflective of a dual articulation of ẓ (dental and laryngeal; cf. Aramaic /ð'/ ≈ < q > → /ʕ/; Segert 1988). The use of the symbol ẓ for /t'/ is not nearly as widespread as has been claimed (see Freilich and Pardee 1984), appearing only in *CTA* 24 and probably in *RIH* 78/14 (Bordreuil and Caquot 1980:352–353; Tropper 1994b; Pardee 2000:859–71).

Etymological /ð/ poses particular problems: it is sometimes written with the character ḏ, but usually with *d*. Apparent confusion of /ð/ and /z/ characterizes certain roots: for example, *nḏr/nzr* “vow” (both in Ugaritic); *ḏmr/zmr* “sing”; *ḏr/ʕr* “seed/arm.” Though there is, therefore, certainly evidence for disparities between the graphic and phonetic systems, the situation was probably not as confused as some have thought. Examination of the confusions claimed by Tropper 1994a reveals that the interpretations of the texts, and hence of the phoneto-semantic identifications, are sometimes either dubious or faulty: for example, *šr* and *ṯr* are not the same word (Tropper 1994a:38) – the first is “flesh, meat,” while the second denotes a kinship status; the two terms only become homophonous in Hebrew with the coalescence of /š/ and /θ/.

### 3.2 Vowels

Because the Ugaritic writing system does not include vowel characters, Ugaritic vocalic phonology represents an uneasy truce between description and reconstruction. It has this feature in common with all of the pre-Christian era Northwest Semitic languages; however, those attested in the first millennium BC either make use of *matres lectionis* (“mothers of reading,” consonant characters used to signal the presence of a vowel) and have later

vocalization systems on the basis of which some retrojection can be done (Aramaic, Hebrew), or else have later congeners in which *matres lectionis* are used (late Phoenician, Punic, Neo-Punic). The reconstruction of the Ugaritic vocalic system must rely, therefore, on two types of internal sources: (i) the “extra” *alif* signs in the Ugaritic script (see §2.2); and (ii) Ugaritic words in syllabically written texts. The latter appear in three distinct forms: (i) the so-called polyglot vocabularies (Ugaritic words written in ancient “dictionary” entries); (ii) Ugaritic words in Akkadian texts; and (iii) proper names. For the first two types, see Nougayrol 1968: texts 130–142 and indices pp. 351–352, and Huehnergard 1987; the third type is more difficult to use for reliable results because of the presence of archaic elements in Ugaritic names and of non-Ugaritic names. If one wishes to reconstruct a form or a word where these internal sources are silent, one must rely on comparative Semitic considerations.

The Ugaritic vocalic system is assumed to have consisted of the same six phonemes reconstructed for Proto-Semitic, /a/, /i/, /u/, /ā/, /ī/, /ū/, to which two secondary long vowels were added by monophthongization, /ê/ < \*/ay/ and /ô/ < \*/aw/. There is no evidence for secondary lengthening of the short vowels (e.g., /a/ → *qameṣ* in Biblical Hebrew), nor for any shifts of the long vowels (e.g., the “Canaanite shift” /ā/ → /ō/). Apparent anomalous uses of the *alif* signs may indicate the presence of glide vowels following certain of the laryngeal and pharyngeal consonants (Verreert 1983), though these data are susceptible to other interpretations (Troppner 1990b).

## 4. MORPHOLOGY

### 4.1 Word formation and word classes

Like the other Semitic languages, Ugaritic morphology is of the inflecting (or fusional) type. The traditional view according to which a Semitic word consists of a consonantal root + internal vowel(s) + additional morphemes still has merit today. Though there are clearly nominal roots, which include a vocalic element (e.g., *kalb*- “dog”), and verbal roots in which vocalic variation is the rule and which serve as the basis for nominal derivation (see below), both types of roots generate derivatives. Morphology thus consists of an abstract entity known as a root, which exists in concrete form as a set of consonants, usually two or three, which in a nominal root may include a vowel, and which is modified by internal vowel change (*ablaut*), by suffixation, and/or by prefixation. Thus, a Ugaritic dictionary, organized by root (according to the tradition of Semitic-language dictionaries), will begin with the simplest form attested, either a verb or a noun, and will proceed from this simple form through the attested verbal forms (if any such exist), then through entries characterized by suffixation, then through those characterized by prefixation and/or by further suffixation: for example, *MLK* “to rule,” *mlk* “king,” *mlkt* “queen,” *\*mmlkt* “kingdom.”

Though it is not a useless thing to analyze an old West Semitic text according to the grammatical categories commonly used for the modern languages of scholarship, a descriptive analysis of these languages gives three primary categories of words: *nouns* (see §4.2), *verbs* (see §4.4), and *particles* (see §4.6). There is, nonetheless, a significant degree of overlap within these categories (e.g., verbal nouns and particles derived from nouns) and there are clearly definable subcategories (e.g., adjectives and adverbs). The three-division description is nevertheless important, for the elements belonging to overlapping categories and to subcategories are clearly definable according to one or other of the primary categories (e.g., verbal nouns will have nominal morphology along with certain syntactic and lexical

features of verbs; adjectives will have nominal morphology not verbal morphology; verbal adjectives will have nominal morphology along with certain syntactic and lexical features of verbs, etc.).

Nouns and adjectives are marked for gender, number, and case, but not for definiteness and only partially for state. These grammatical categories are expressed by affixation. Internal vowel variation and prefixation function primarily in nouns to mark lexical categories rather than grammatical ones.

Verbs are marked for aspect/tense, for person, for voice, and for mood. There are (i) two aspects – perfective and imperfective, the first marked only by suffixation, the second by prefixation and suffixation; (ii) three voices – active, middle, and passive, marked by internal vowel change and by prefixed consonantal morphemes; and (iii) five moods, all marked by suffixation to the imperfective verb. The position of the person markers indicates aspect/tense; in other words, person is expressed by suffixation in the perfective, by prefixation in the imperfective.

Particles are characterized by the absence of the morphological markers of nouns and verbs. This is completely true, however, of only the most basic particles, for many are secondarily derived from nouns or pronouns and may thus include markers characteristic of the nominal system.

The following presentation of the morphological categories will follow this three-way division, with an attempt to delineate clearly the overlapping categories and the subcategories. In the ensuing discussions and tables Ø is used to indicate forms that are expected to exist but that are not attested in the texts presently extant, while -ø is used for forms without a consonantal indicator of a morpheme otherwise indicated consonantally in the paradigm or for a form ending with hypothetical *zero* vowel.

## 4.2 Nominal morphology

### 4.2.1 Nominal formation

Nominal forms may consist of the following:

1. *ROOT + internal vowel(s)*: for example, *MaLK*- “king”; *DaKaR*- “male.”
2. *Nominal prefix + ROOT + internal vowel(s)*: for example, *maL* *ʾaK*- “messenger.”
3. *ROOT + internal vowel(s) + nominal suffix*: for example, *RaʾaB ān*- “famine.”
4. *Combinations of 2 and 3*: for example, *ʾaLʾiYān*- “mighty.”

There are also a certain number of reduplicated (e.g., *qdqd* “top of head,” *ysmsm* “beauteous”) and quadriconsonantal (e.g., *ʾrgz* “walnut”?) nominal forms.

The most common nominal prefixes are *m*- (concrete entities), *t*- (abstract entities); much rarer are *ʿ*- and *y*- (both for concrete entities).

The most common nominal suffixes are *-n* (*-ān*-, more rarely *-an*-) and *-t* (perhaps, as in the later Northwest Semitic languages, *-īt*- and *-ūt*- for abstracts).

The data are inconsistent on the matter of whether nouns of the *qatl/qitl/qutl* types had monosyllabic or bisyllabic stems in the plural (as in Hebrew: *melek* < *malk*, *mālākīm* < *malak*-). Either the bisyllabic plural base was in the process of development from an originally monosyllabic one (Sivan 1992), or else the plural stem was already bisyllabic in Proto-Ugaritic and the second vowel was inconsistently elided in Ugaritic (Huehnergard 1987:304–307).



### 4.2.2 Case

Case-markers are suffixed and consist of a combination of vocalic and consonantal elements. A *triptotic* case system – nominative, genitive, accusative – is used in the singular, a *diptotic* one – nominative, oblique – in the dual and plural. This system is consistent with case systems known from fully vocalized languages and is demonstrated internally by the reasonably consistent use of the appropriate *alif* sign (see §2.2) in writing nouns of which ʾ ([ʔ]) is the final consonant: for example, sg.masc.nom. *ksū*=[kussaʔu]; sg.masc.gen. *ksī*=[kussaʔi]; sg.masc.acc. *ksā*=[kussaʔa]; pl.masc.nom. *rpūm*=[rapaʔūma]; pl.masc.obl. *rpīm*=[rapaʔīma].

There is no separate case for the expression of the vocative. There are two lexical vocative markers, *l* and *γ* (cf. Arabic *ya*), but a noun may be vocative without the use of a lexical marker. There is some evidence that the oblique case was used in the plural (Singer 1948) and one datum (*ksī* “O throne”) for the genitive in the singular, perhaps by analogy with the case that normally follows the preposition *l* (Bordreuil and Pardee 1991:158).

The accusative case is used both for the object(s) of transitive verbs and for various adverbial notions.

There are some nouns, particularly those bearing a nominal suffix containing a long vowel (e.g., *-ān*, *-īt*) that have a diptotic singular system: *-u* nominative, *-a* oblique (Liverani 1963; Huehnergard 1987:299.)

### 4.2.3 Gender

Gender is marked by suffixed morphemes: the singular masculine by *-ø*; singular feminine by *-t* (= [-(a)t-]); plural masculine by lengthening of case-vowel (lengthened genitive singular = plural oblique); plural feminine by *-t* (= [-āt-]). The dual morpheme was probably attached to the singular stem, masculine or feminine.

Several nouns that take feminine agreement do not bear the *-t-* morpheme (e.g., *ūm* “mother”); while the plural morphemes do not correspond in every case to the sex/gender of the entity devoted (e.g., *grnt*, pl. of *grn* “threshing-floor,” probably masculine as in Hebrew).

### 4.2.4 Number

Singular, dual, and plural are productive number categories, marked by variations in the case-vowel, with affixation of *-m* to the dual and plural (for the problem of the quality of the vowel after this *-m* on the dual, see Huehnergard 1987:298, who posits that it was originally *i* on the dual, *a* on the plural).

### 4.2.5 Definiteness

There is no quasi-lexical marker of definiteness in Ugaritic (cf. *h-* in Hebrew), though the unusually frequent use of *hn* in one text may be a precursor of such a development (Liverani 1964:181–182; Pardee, 1984a:218, n. 23).

### 4.2.6 State

A fifth grammatical category, morphosyntactic in nature, is useful in describing the ancient Semitic languages; this is the category of state. There are two primary states, absolute and construct; a third, the pronominal state, is useful in describing some of the later Northwest Semitic languages where vowel reduction is prevalent, and will be referred to briefly here.

**Table 9.2 The Ugaritic noun: absolute state**

	Singular		Dual	Plural
<i>Masculine</i>				
<i>Nominative</i>	malku	<i>Nominative</i>	malkāmi or malkāma	malakūma or malkūma
<i>Genitive</i>	malki	<i>Oblique</i>	malkêmi or malkêma	malakīma or malkīma
<i>Accusative</i>	malka			
<i>Feminine</i>				
<i>Nominative</i>	malkatu	<i>Nominative</i>	malkatāmi or malkatāma	malakātu or malkātu
<i>Genitive</i>	malkati	<i>Oblique</i>	malkatêmi or malkatêma	malakāti or malkāti
<i>Accusative</i>	malkata			

*Absolute* describes a noun in unbound form, *construct* a noun bound to a following one in the genitive relationship, and *pronominal* a noun bound to a following pronoun also in the genitive relationship.

An example of typical masculine and feminine nouns in the absolute state, indicating the markers of case, gender, and number, is presented in Table 9.1. Note that in the dual and plural numbers, variant forms occur. The vowel /ê/ is from earlier \*/ay/ (see §3.2).

#### 4.2.6.1 Construct state

In Ugaritic, the case-vowel is preserved in the first word(s) of genitive phrases (in traditional grammar the head noun is called the *nomen regens*, the second noun the *nomen rectum*). Thus, in the singular, the genitive relationship is marked only by the genitive case-vowel on the second element of the phrase. This feature is shared with, for example, classical Arabic, whereas in other Semitic languages the first word also shows some form of modification (e.g., Akkadian *šarru* becomes *šar* in construct, Hebrew *dābār* becomes *dabar*; for another view of the Ugaritic data, see Zevit 1983; refutation by Huehnergard 1987:300–301). In the dual and the plural the *-m* of the *nomen regens* is usually dropped in construct.

<i>Singular</i>	malku qarīti	“The/A king (NOM.) of the/a city”
<i>Dual</i>	malkā qarīti	“[The] two kings (NOM.) of the/a city”
<i>Plural</i>	mal(a)kū qarīti	“[The] kings (NOM.) of the/a city”

#### 4.2.6.2 Pronominal state

The case-vowel is also preserved in the pronominal state, again in contrast with Akkadian where the case-vowel drops; here Hebrew shows remnants of a system similar to the Ugaritic one (*dābār* + *kā* for *dabar* + *V* + *ka*).

<i>Singular</i>	malkuhu	“his king” (NOM.)
<i>Dual</i>	malkāhu	“his two kings” (NOM.)
<i>Plural</i>	mal(a)kūhu	“his kings” (NOM.)

### 4.2.7 Adjectives

Adjectival morphology is identical to that of nouns. An adjective used independently (“substantivally,” according to the traditional grammatical term), not as a modifier of a noun,

functions itself as a noun. When an adjective modifies a noun, it agrees in gender, number, and case with the noun. It is by this morphosyntactic feature that adjectives are most clearly differentiated from nouns, for a noun used to modify another noun does not vary in gender (e.g., the phrase “the woman is a man” in Ugaritic would be *ḏtt mt hy* (lit. “[The] woman, a man [is] she”), where *ḏtt* retains its feminine marker and *mt* its masculine marker). Attributive adjectives normally follow the noun they modify; predicate adjectives either precede or follow the noun.

The primary adjectival suffix is the so-called gentilic or *nisbe* ending consisting of *vowel* + *-y* (= [-yy-]) + *case-vowel*. The quality of the first vowel is uncertain. The only apparently explicit indication shows [u], *qnūlym* “people who work with royal purple dye or with lapis lazuli” (CAT 2.73:17 [line 39 in Pardee 1983–1984]).

Comparative and superlative adjectival markers do not exist and such notions must thus be expressed lexically (e.g., by forms of the root *M'D* “much”) or syntactically (e.g., *n'mt šnt il*, “the best years of El” [CAT 1.108:27], a substantified adjective in construct with a noun, lit. “the good ones of the years of El”).

A nominal genitive formation is often used in place of an adjectival one, e.g., *ḏtt šdqh* (= [ʔaθθatu šidqihu]) “the wife of his legitimacy” = “his legitimate wife” (CTA 14:12 [Gordon 1965:113, §13.22]).

#### 4.2.8 Numerals

In Ugaritic, numerals belong to nominal categories: cardinal numbers are nouns, ordinals adjectives. Numbers in texts may be either fully written out or expressed by number signs, using the same system as is used in Akkadian texts (a single vertical wedge = “1,” a single oblique wedge = “10,” etc.).

The Ugaritic repertory of numerals is largely similar to the standard West Semitic inventory:

	<i>Cardinals</i>	<i>Ordinals</i> (where different)
1	ḥd/ḥt and šty	?
2	tn/tt	
3	tl̄t/tl̄tt	
4	arb'/arb't	rb'
5	hmš'/hmšt	
6	tt/ttt	td̄t
7	šb'/šb't	
8	tmn(y)/tmnt	
9	tš'/tš't	
10	šr'/šrt	
11	šty šr'/šrh	
12	tn šr'/šrh etc.	
20	šrm etc.	
100	mīt (sing.)/maīt (pl.)	
1,000	ālp	
10,000	rbt	

With the exception of words containing an *alif* sign, the vocalism of numerals can be reconstructed only from comparative data.

The primary distinctive feature of the Ugaritic numerals is in their morphosyntax: as opposed to the other ancient Semitic languages, where the numerals 3 through 10 observe

*chiastic concord* (i.e., incongruent gender agreement, feminine-looking numbers with masculine nouns and vice versa), the distribution of numbers marked with  $-\phi$  versus  $-(a)t$  shows less regularity.

Other features of numerals deserving special comment:

1. The formant  $\text{šty}$  is used for the number “one,” as in Akkadian, not just in the number “eleven” as in Hebrew.
2. The only attested forms of the absolute case of the number 2 are  $\text{tn}$  and  $\text{tt}$  ( $\text{tnm}$  is adverbial, “twice,” in CTA 18 IV 22, 33; 19 II 78; CAT 1.104:18, 20). This form constitutes an isogloss with Akkadian ( $\text{šine}$ ) against the other West Semitic languages (e.g., Hebrew  $\text{šnayim}$ ). See Pardee 2000:195.
3. The alternate form with  $-h$  of the 10-word found in the cardinals of the teens is not used only to modify feminine nouns as in Hebrew. Moreover, the presence of  $h$  in the Ugaritic spelling shows that the origin of the element was consonantal, though its form (i.e., the vowel[s] with which the consonant is associated) and its function are uncertain.
4. The ordinals may have a long vowel between the second and third radicals, though its quality is unknown; hence the difference between 6 and 6th: respectively  $\text{tittu}$  ( $< *tidtu$ ) versus  $\text{taditu}$ , or the like. The ordinals are certainly not formed with the *nisbe* suffix (as in Hebrew), for that morpheme appears in Ugaritic as  $-\gamma$  (see §4.2.7).

Fractions are very poorly known:  $\text{hšt}$  appears in prose in the meaning “half” of a given quantity (CTA 34:10) while  $\text{nšp}$  apparently means “half” of a (sheqel-)weight in administrative texts.

In a mythological text (CTA 14 I 16–20) one finds a series of D-stem passive feminine participles of denominative verbs formed from numbers, designating a series of women:  $\text{mtltt}$ ,  $\text{m'rbt}$ ,  $\text{mhmšt}$ ,  $\text{mtdtt}$ ,  $\text{mšb't}$  “the third one . . . the seventh one.” From context these forms refer back to  $\text{mtrht}$  (line 13) “the married one,” namely “the third woman (taken in marriage),” “the fourth . . .” etc. These words are thus neither fractions nor multiplicatives, as has often been claimed.

In the number phrase, the noun denoting the counted entity may be either in the same case as the number (i.e., the numeral and the noun are in apposition) or in the genitive case (Blau 1972:78–79).

In poetry, several cases are found of the ordinal number preceding the noun it modifies, in apparent contradiction to the rule that attributive adjectives follow the noun they modify (Gordon 1965:48–49, §7.44; Blau 1972:79). It is likely that such constructions were genitival; in other words, the adjective was in construct with the noun, rather than appositional, as is the case when the attributive adjective follows the noun it modifies, though the semantic nuance of the genitival construction is unknown. One encounters, for example,  $\text{b šb' ymm}$  (CTA 17 I 16), probably [bi šabiʿi yamīma] “on the seventh of days.” Rarer is a prepositional formulation:  $\text{hn šb[ʿ] b ymm}$  (CTA 17 V 3–4), probably [hanna šabiʿa bi yamīma], literally “Behold on the seventh among days.”

The preposition  $l$  is often used to join the unit to the ten in compound numbers, as in  $\text{tn l'šrm}$  “twenty-two” (Pardee 1976:302).

### 4.3 Pronouns

In their function as replacing nouns, pronouns share features with nouns, though they are not as consistently marked for case, gender, number, and state as are nouns and adjectives.

### 4.3.1 Personal pronouns

Ugaritic possesses both independent and clitic personal pronouns.

#### 4.3.1.1 Independent personal pronouns

The primary function of independent personal pronouns is to express the grammatical concept of person on the noun side of the grammar (person is expressed grammatically in verbs, but not in nouns); this function entails the marking for gender. Case is also marked, apparently diptotically, though the oblique forms are rarely attested. Nominative case forms are as follows:

	<i>Singular</i>	<i>Dual</i>	<i>Plural</i>
<i>1st com.</i>	ʾnk/ʾn	Ø	Ø
<i>2nd masc.</i>	ʾt	ʾtm	ʾtm
<i>2nd fem.</i>	ʾt	Ø	Ø
<i>3rd masc.</i>	hw	hm	hm
<i>3rd fem.</i>	hy	Ø	Ø

In the oblique case, separate forms are attested for only the following:

	<i>Singular</i>	<i>Dual</i>	<i>Plural</i>
<i>3rd masc.</i>	hwt	hmt	hmt
<i>3rd fem.</i>	hyt	Ø	Ø

These forms function both as accusatives (i.e., direct object of a transitive verb: *kbd hyt* “honor her”; *kbd hwt* “honor him” (CTA 3 III 7, VI 20)) and as genitives (*tbr dīy hwt* “he broke the pinions of him”; *tbr dīy hyt* “he broke the pinions of her” (CTA 19 III 122, 144)).

The first- and the second-person forms consist, as in most of the Semitic languages, of a deictic element *ʾn* followed by the pronominal element proper. The vocalization of these forms can then be approximated as follows:

	<i>Singular</i>	<i>Dual</i>	<i>Plural</i>
<i>1st com.</i>	[ʾanāku] (<[ʾan + āku])		
<i>2nd masc.</i>	[ʾatta] (<[ʾan + ta])	[ʾattumā] (<[ʾan + tumā])	[ʾattumu] (<[ʾan + tumu])
<i>2nd fem.</i>	[ʾatti] (<[ʾan + ti])		

The optional first-person singular form *ʾn* already shows the dropping of the consonantal element *-k-*, though its vocalization is unknown ([ʾanā], as in Aramaic, or [ʾanī], by analogy with other first-person pronominal forms, as in Hebrew?).

The third-person singular forms consist of an augmented form of the primitive pronoun: [hu] > [huwa], [hi] > [hiya].

#### 4.3.1.2 Clitic personal pronouns

Proclitic and enclitic pronouns, clearly related historically to the independent forms just cited, are also attested. Historically speaking, finite verbal forms (see §4.4.2) are made up of a pronominal element providing the notion of person, plus the verbal element. These pronominal elements were *suffixed* in the *perfective*, essentially *prefixed* in the *imperfective*:

	PERFECTIVE			IMPERFECTIVE		
	Singular	Dual	Plural	Singular	Dual	Plural
<i>1st com.</i>	-t	-ny	-n	ʾ	n-	n-
<i>2nd masc.</i>	-t	-tm	-tm	t-	t-	t-
<i>2nd fem.</i>	-t	Ø	-tn	t-	t-	t-
<i>3rd masc.</i>	-φ ([ -a ])	-φ ([ -ā ])	-φ ([ -ū ])	y-	y-/t-	y-/t-
<i>3rd fem.</i>	-t	-t	-φ ([ -ā ])	y-	t-	t-

As it is absent in the other Semitic languages while being attested in Egyptian, the first common dual *-ny* (also attested as a genitive enclitic) appears to be an archaic retention in Ugaritic. Other dual forms indicated were apparently differentiated from identically written plural forms (or singular in the case of the 3rd fem. perf.) by vocalic pattern.

Enclitic personal pronouns are also attached to nouns, with a genitive function, and to verbs, with a primarily accusative function (occasionally dative). Here the second person is marked by *-k-* rather than by *-t-*:

	<i>Singular</i>	<i>Dual</i>	<i>Plural</i>
<i>1st com.</i>	-y/-φ/-n	-ny	-n
<i>2nd masc.</i>	-k	-km	-km
<i>2nd fem.</i>	-k	Ø	-kn
<i>3rd masc.</i>	-h	-hm	-hm
<i>3rd fem.</i>	-h	-hm	-hn

The forms indicated for the first person are distributed according to function: *-y/-φ* is genitive (i.e., attached to nouns); *-n* accusative (i.e., attached to transitive verbs). The former set is distributed according to the case of the singular noun to which the genitive suffix is attached (nom. = *-φ*; gen./acc. = *-y*); the *-φ* form is assumed to have arisen through syncope (\*[ -uya ] → long vowel, usually identified as [ -i ]). This distribution differs from early Phoenician, where the suffix on nominative/accusative nouns is identical (i.e., orthographic *-φ*), *-y* only appearing in the genitive. As with the independent and prefixed pronominal elements, most of the dual forms were apparently differentiated from identically written plural forms by vocalic pattern.

Accusative enclitic pronouns on imperfect verbs show a great deal of variation because of assimilation to *-n* verbal forms and apparent reanalysis. For example, singular third masculine can appear as *-h* (= [ -hu ]); as *-n* (= [ -annu ] < [ -an ] + [ hu ]); as *-nh* (= [ -annahu ] < [ anna ] + [ hu ]); as *-nn* (= [ -annannu ]; apparently from [ -anna ] + [ nnu ], through reanalysis of [ nnu ] as a pronominal suffix); and finally even *-nnn* (apparently = [ -annannannu ], through double reanalysis). See Pardee 1984b:244–245, n. 14.

#### 4.3.2 Relative pronouns

The relative pronoun is composed of \**d̄* ([ ð ]) + vowel, nearly always written with *d*, marked for gender and number, though the forms are not used consistently. This particle is directly related to the *dū dā dī* series in Arabic and to the *ze<sup>h</sup>/zōt* series in Hebrew (used sporadically as a relative pronoun there), and its basic function is therefore deictic, as is shown in Ugaritic by the enclitic use of *-d* in demonstrative pronouns and adjectives (see §4.3.3) and in adverbials. The gender and number categories indicated here represent agreement between the relative pronoun and its antecedent:

*d* (sg. masc.)    *dt* (pl. masc. and fem.; not used consistently,  
interchangeable with *dt*)  
*dt* (sg. fem., also interchangeable with *dt*)

### 4.3.3 Demonstrative pronouns

The primary demonstrative pronouns and adjectives are compounds consisting of the deictic particle *hn* (probably essentially the same particle as the Hebrew definite article and as the deictic particle *hēn/hinnē<sup>h</sup>* in that language), to which explicating elements are joined: either the relative pronoun *d* (cf. Arabic *illādī*) in the case of the proximal demonstrative; or *k*, of uncertain origin, in the distal. The forms are identical to those of the demonstrative adjectives, and the two categories are defined, therefore, by their syntactic characteristics:

*Proximal*    *hnd* ~ *hndt*  
*Distal*        *hnk* ~ *hnkt*

The forms with and without *-t* are not distributed consistently according to gender, and the *-t* may thus be the enclitic particle and not the feminine morpheme *-t*.

Though the usage is rare and to date attested only in the oblique case, the third-person independent pronouns could also be used as demonstrative adjectives, apparently, as in Hebrew, with a distal connotation: for example, *mlk hwt* “that king” (CAT 1.103:43); *hwt hyt* “that land” (CAT 45', 55', 56'; for the reading of line 45', see Pardee 1986:119, 124).

### 4.3.4 Other pronouns

The other pronominal elements do not show the primary morphological characteristics of nouns and thus overlap with the category of particles. They are included here in order to provide a complete picture of pronouns.

#### 4.3.4.1 Interrogative pronouns

The attested interrogative pronouns are *my* “who?” and *mh* “what?” Comparing *mh*, of which the *-h* is consonantal, with Biblical Hebrew *mah* leads to the conclusions that (i) the gemination following the Hebrew pronoun represents assimilation of the *-h*; and (ii) the presence of the <h> in the orthography is therefore historical writing (this solution appears more likely than positing a Proto-Hebrew form *man* and identifying the orthographic <h> as a secondary *mater lectionis*).

#### 4.3.4.2 Indefinite pronouns

The indefinite pronouns and adjectives are *mn/mnk* and *mnm*. As presently attested, *mn* and *mnk* denote human entities (“whoever”), *mnm* inanimate ones (“whatever”). The basic particle was plausibly [mVn] with the distinction between human and nonhuman referents expressed by ablaut (e.g., [min-] for humans, [man-] for nonhumans); *-k* and *-m* are expanding elements of uncertain semantic content. Because “enclitic” *-m* may be attached to any part of speech, it would not be surprising to encounter the form *mnm* applied to humans; it would have been distinguished from the nonhuman reference by its characteristic vowel.

## 4.4 Verbal morphology

The verbal system represents an archaic form of West Semitic, one with an N-stem; a D-stem (characterized by the doubling of the middle radical); a causative stem in Š; *t*-stems built off the G-, D-, and Š-stems; as well as some less well-attested stems. For discussion of the conventional classification of Semitic verb-stems, see Chapter 6, §3.3.5.2.

As in the other Semitic languages, the basic verbal form can itself express various sorts of action. The primary opposition is *transitive* versus *intransitive*. Of the latter sort, there are two primary types: verbs of *motion* and *stative* verbs. Verbs of motion are themselves of two primary types: verbs that express only motion and those that express either the motion or the state achieved (e.g., *qm* “arise” or “be standing”). Stative verbs also can denote either the state itself or the attainment thereof (e.g., *qrb* “be near” or “become near,” i.e., “approach”). These distinctions are reflected in the verbal system: only transitive verbs can be passivized and tend to take double accusatives in the causative, single accusatives in the D-stem. Stative verbs are factitivized in the D-stem, cannot be passivized in the G-stem, and have a stative participial form rather than the active one. Verbs of motion cannot be passivized in the G-stem, appear rarely in the D-stem, and are transitivized in the Š-stem, where they take the single accusative construction. There are of course, a certain number of verbs that either cross-categorize or defy classification.

### 4.4.1 Verb-stems

The attested verbal stems are as follows:

1. *G-stem*: base stem, or simple stem; active and passive voices.
2. *Gt-stem*: *-t-* infixed after first radical of G-stem; middle/reflexive in function.
3. *D-stem*: doubled middle radical; factitive in function; active and passive voices.
4. *tD-stem*: *t-* prefixed to D-stem (see Huehnergard 1986); middle/reflexive in function.
5. *N-stem*: preformative *n-*; middle/passive in function.
6. *Š-stem*: preformative *š-*; causative in function; active and passive voices.
7. *Št-stem*: *-t-* infixed after *š-* of causative stem; middle/reflexive in function; the few forms attested indicate that the form may no longer have been productive.
8. *L-stem*: lengthened vowel after first radical and reduplicated second or third radical; intensive or factitive in function.
9. *R-stem*: reduplication of both radicals of biconsonantal root, of second and third radicals of triconsonantal root; factitive in function.
10. *tR-* or *Rt-stem*: *t* prefixed to first root consonant or infixed after first root consonant of R-stem; factitive-reflexive in function.

The following examples are given with vocalization in order to illustrate the phonetic distinctions between the forms (see below). Many details of the vocalizations are, however, still uncertain. Here, an asterisk before a G-stem form indicates that the verb is only attested in Ugaritic in the following derived stem (and does not indicate that the form is reconstructed).

1. *LḤM* “to eat (something)” (G-stem transitive, *laḥama*); *LḤM* “to provide (someone) with food” (D-stem, *liḥḥama*); *ŠLḤM* “to cause (someone) to eat (something)” (Š-stem, *šalḥima*).
2. *RḤṢ* “to wash” (G-stem transitive, *raḥaṣa*); *(ʾ)RTH Ṣ* “to wash oneself” (Gt-stem, *ʾirtaḥaṣa*).



3. NTK “to pour out” (G-stem transitive, *nataka*); NTK “to pour forth” (N-stem, *nattaka* < *nantaka*).
4. \*BKR “to be the first born” (G-stem stative, *bakura*); BKR “to promote (someone) to the status of first born” (D-stem, *bikkara*).
5. \*KMS “to squat” (G-stem intransitive, *kamasa*); TKMS “to collapse” (tD-stem, *takam-masa*).
6. ‘RB “to enter” (G-stem verb of movement, *‘araba*); ŠRB “to cause (someone) to enter” (Š-stem, *ša‘riba*).
7. RĤQ “to be far off or to move far off” (G-stem stative, *raḥuqa*); ŠRĤQ “to cause to be far off” (Š-stem, *šaḥiqa*).
8. QL “to fall” (G-stem intransitive, *qāla?*); ŠQL “to cause (something) to fall” (Š-stem, *šaqla*); (ʾ)ŠTQL “to cause oneself to fall” → “to arrive” (Št-stem, *ʾištaqla*).
9. RM “to be or become high” (G-stem stative, *rāma*); RMM “to raise” (L-stem, *rāmama*).
10. \*KR(R) “to turn” (G-stem verb of movement, *karra*); KRKR “to turn, twist, snap” (said of what one does with the fingers) (R-stem *karkara*); cf. the adjectival form YSM SM “beautiful” < YSM (G-stem stative *yasuma* “to be beautiful”).
11. \*YPY “to be beautiful” (G-stem stative, *yapiya*); TTPP “she makes herself beautiful” (only form attested of Rt- or tR-stem, *tītēpēpī* < \**tītaypaypiyu* or *tītāpēpī* < \**tiytapaypiyu*).

#### 4.4.2 Verb conjugations (aspect/tense)

There are two verbal conjugations marked for person, gender, and number: one is characterized by STEM + PRONOMINAL ELEMENT and expresses acts viewed as complete (*perfective*, often called the “perfect” though the term is technically incorrect), the other is characterized by PRONOMINAL ELEMENT + STEM (+ AFFIX) and expresses acts not viewed as complete (*imperfective*, often called the “imperfect”). The pronominal elements (see §4.3.1.2) were joined to the verbal elements in an archaic stage of the language. This description of the form and function of the two verbal conjunctions is accurate for the prose texts. In poetry the distribution of the two forms just described has thus far defied complete description. Usage seems to reflect an older stage of the language, when the zero-ending imperfect form (see §4.4.6, 4) functioned as a preterite, like Akkadian *iprus*. In the West Semitic verbal system, the permansive came to function as perfective, the “subjunctive” (*iprusu*) as an imperfective, and the preterite as a jussive (and, particularly in Biblical Hebrew, as a frozen preterite after *wa-*).

In spite of the problems of description and categorization of the verbal system in the poetic texts, many scholars, e.g., Tropper 1995, have preferred to classify the Ugaritic verbal system on the basis of poetic usage, rather than on that of the prose texts (similar attempts, of course, have been made in the classification of Biblical Hebrew). It appears legitimate to see in the poetic texts remnants of a previous stage of the language (plausibly closer to East Semitic), remnants that seem not to be used consistently because they are no longer representative of the spoken language, while the prose texts reflect spoken Ugaritic in the thirteenth–twelfth centuries BC. Only in these prose texts is a reasonably consistent system visible (cf. Mallon 1982).

The Ugaritic verbal system is here classified as *aspectual*, rather than *tensed*, primarily because of its similarity to the prose system of Biblical Hebrew (Pardee 1993a, 1993b, 1995). While tense is a real-world phenomenon (past–present–future), aspectual systems include a greater degree of subjectivity; in other words, the speaker may express a situation as complete or incomplete according to several criteria. Because of the nature of tense, aspectual systems cannot ignore temporal considerations; accordingly, a language may not be identified as a

tensed language simply because it reflects real-world temporal considerations. On the other hand, a language may be classed as aspectual if it ignores real-world temporality, as in the use of the imperfect in Biblical Hebrew prose to express past-tense iteratives (e.g., *yišma* ‘he used to hear’).

The perfective may have been characterized by internal ablaut for active (*katab-*) versus stative (*katib-*, *katub-*), but all internal evidence is for the *katib-* type (writings of the middle radical with : *lik*=[laʔika] ‘he sent’; *šil*=[šaʔila] ‘he asked’). Syllabic writings attest some *katab-* forms (Huehnergard 1987:319–320).

The imperfective was characterized by internal ablaut, perhaps for active (*yaktub-*) versus stative (*yiktab-*, *yaktib-*). There are few data for these differentiations, but those that do exist tend to agree with the data from the later West Semitic languages, making the conclusions plausible. In addition, the imperfective is also marked, by affixation to the stem, for mood (see §4.4.4). The Barth–Ginsberg Law of *a*-dissimilation (*yaktab* → *yiktab*) was operative in Ugaritic.

No certain evidence exists for a present-future form corresponding to Akkadian *iparras* (Fenton 1970).

### 4.4.3 Voice

Active verbs are of two primary types, transitive and intransitive (e.g., *halaka* ‘he went’; *maḥaša ʿêba* ‘he smote the enemy’). The concept of transitivity is not a useless one in Semitics, for not only do certain verbs take complements that correspond to what in other languages would be direct objects, but distinctively marked passive forms, used almost exclusively for verbs that in other languages would be qualified as transitive, are common. Though lack of vocalization in Ugaritic makes identification difficult, it is likely that all transitive forms (i.e., G-stem transitive verbs, D-stem, and Š-stem) had passive forms that were differentiated from the active by ablaut (for a contrary view on the G-passive finite forms, see Verreet 1986:324–330; brief refutation in Tropper 1993a:478–479). In addition, the N-stem, basically an intransitivizing and deagentifying stem, can be used as a passive (such a usage of the *t*-stems, which became common in Hebrew, is not clear in Ugaritic.) Passives are attested for finite forms (e.g., *tšt išt* ‘fire is placed’ [CTA 4 VI 22]) as well as for participles. There is as yet no evidence for ablaut-passive imperatives, though there was almost certainly an N-stem imperative (*RS 34.126:13 ibky* and *ibid.* 18 *išḥn*, the first of which appears to function as a passive ‘be bewept’ [Bordreuil and Pardee 1991:157–159]). On the basis of comparative data, one would not expect a passive infinitive necessarily to have existed.

Between the two extremes marked by the clearly transitive and passive forms, there is a whole middle range of forms denoting reflexivity, reciprocity, advantage or disadvantage to actor and so forth. These notions are clearest in the *t*-stems (Gt, tD, and Št). The primary function of the N-stem in Ugaritic, as in several of the Semitic languages, was for the expression of patient-oriented acts and it is thus used for both the passive and the middle.

### 4.4.4 Mood

Verbal mood was in Ugaritic, as in the other West Semitic languages, marked by variations to the imperfective stem.

#### 4.4.4.1 Imperative

The imperative in Ugaritic does not have the preformative element characteristic of the imperfective, but the fact that its stem-vowel is identical to that of the imperfective leaves

no doubt as to the historical linkage of the imperative to the imperfective. Its form is thus ROOT + stem-vowel (+ additional PRONOMINAL ELEMENT). The question of an additional vowel between the first two radicals is unresolved: imperfective *yaktub-*; imperative *kVtub-* or *ktub-*. In the case of *kVtub-*, the quality of the first vowel is unknown. Is it always identical to the stem-vowel or sometimes different? To the basic imperative element may be added the *-a(n)(na)* elements; see §4.4.4.2. The imperative existed only in the second person and only for positive commands (negative commands are expressed by *dl* + jussive).

#### 4.4.4.2 Other moods

All other moods are marked by affixation to the full imperfective stem (the stem YKTB/*yaktub-* will be used below for STEM):

#### Nonimperative moods of Ugaritic

<i>Jussive</i>	YKTB + $\phi$	<i>yaktub</i>
<i>Indicative</i>	YKTB + u	<i>yaktubu</i>
<i>Volitive</i>	YKTB + a	<i>yaktuba</i>
<i>Energic 1</i>	YKTB + (a)n	<i>yaktubVn</i>
<i>Energic 2</i>	YKTB + anna	<i>yaktubanna</i>

The morphosemantic values of these moods are largely derived from comparison with other Semitic languages, for the forms are not used consistently in the poetic texts and the prose texts have not yet furnished sufficient material to establish usage with certainty. Because of the absence of vowel indicators, the usage of one mood or another can be determined only when the root ends in *alif* or *yod*: the form of *alif* will indicate the quality of the following vowel, while the presence or absence of *yod* may indicate the presence or absence of a following vowel (*yabniyu* = <ybny>; \**yabniy* → *yabni* = <ybn>). These III-weak roots (see §4.4.6) thus provide us with the primary internal data on the aspectual and modal systems in Ugaritic, but inconsistency of usage, particularly in the case of III-*y* roots, also creates a significant degree of uncertainty.

The *-a* form does not function primarily as a marker of syntactic dependency (Verreert 1988), but as a *volitive* (Tropper 1991; 1993a:473–474; Pardee 1993b), and its traditional classification – namely “subjunctive” – borrowed from Arabic, is thus not appropriate (this is to be understood not as a claim that the *-a* form cannot appear in subordinate clauses, but as a denial that such is its principal function).

The two *energic* forms are only distinguishable when followed by a suffix (see §4.3.1.2) and their semantic import is uncertain. The distribution of these suffixed forms clearly indicates the existence of two *energic* forms, *-an* and *-anna* (as in Arabic). Whether there also existed a similar form built on the “indicative” (*-u+n(a)*), as apparently in old Canaanite (Rainey 1996, II:234–244), has not been determined.

Mood distinction in forms containing a suffixed pronominal subject element (e.g., pl. 3rd masc. *yaktub* + *ū*) is variable in the later languages and impossible to determine in Ugaritic (except where the distinction was marked by consonantal *-n* – there the problem is the precise function of the *-n*).

#### 4.4.5 Strong verb paradigm

The G-stem of the Ugaritic strong verb is illustrated in Table 9.3 (particularly doubtful reconstructions are indicated with one or more question marks); KTB is the root meaning to “to write.” More extensive paradigms, with proposed vocalizations, can be found in Segert 1984.

**Table 9.3 The Ugaritic verb: G-stem**

	Perfective	Imperfective	Jussive	Imperative
Singular				
1st com.	katabtu	ʾaktubu	ʾaktub	
2nd masc.	katabta	taktubu	taktub	kutub(a)
2nd fem.	katabti	taktubīna	taktubī	kutubī
3rd masc.	kataba	yaktubu	yaktub	
3rd fem.	katabat	taktubu	taktub	
Dual				
1st com.	katabnayā (?)	naktubā (?)	naktubā?	
2nd masc.	katabtumā	taktubā(ni)	taktubā	kutubā
2nd fem.	Ø	Ø	Ø	Ø
3rd masc.	katabā (?)	yaktubā(ni)	yaktubā	
		taktubā(ni)	taktubā	
3rd fem.	katabtā (?)	taktubā(ni)	taktubā	
Plural				
1st com.	katabnū	naktubu	naktub	
2nd masc.	katabtum(u)	taktubū(na)	taktubū	kutubū
2nd fem.	kababt(n)a	taktubna (?)	taktubna (??)	kutubā (?)
3rd masc.	kababū	yaktubūna	yaktubū	
		taktubūna	taktubū	
3rd fem.	kababā	taktub(ā)na (?)	taktubā (??)	

The third-person dual and plural imperfectives often have preformative *t*-, rather than *y*- (Verreet 1988). The presence of different forms in similar texts appears to show that *t*-preformative cannot in and of itself mark a distinction either of gender (masc. vs. fem.) or of number (dual vs. pl.): for example, *t ʿrbn gtrm* “the *gtrm* will enter” (CTA 33:9); *yrdn gtrm* “the *gtrm* will descend” (CAT 1.112:18); cf. *tʿln ilm* “the gods will ascend” (CAT 1.112:8).

Second-person feminine dual forms are not attested, but the graphic identity of third-person masculine and feminine pronominal forms indicates that a distinction would, in any case, have been vocalic and thus indeterminable in the consonantal orthography.

The N-stem imperative had *i* in the preformative syllable: for example, *išḥn* ([ʔiṣṣaxin-] < \* [ʔiṣṣaxin-]) “be hot!” (RS 34.126:18, cf. *ibky* “be bewept!” in line 13; Bordreuil and Pardee 1991:157–158). The same holds for the Gt perfective: thus, *itdb*, generally taken as a scribal error for *itbd* ([ʔiṭabada] < \* [ʔiṭabada]) “it has perished” (CTA 14 I 8).

The Gt and tD were apparently characterized by different stem-vowels in the imperfective, *i* versus *a*: *yštīl* (Gt) versus *yštāl* (tD) “ask, importune” (Huehnergard 1986).

It is highly unlikely that there existed an H-causative (Hebrew *Hiphil*) or a ʾ-causative (*Aphel*) alongside the Š-causative (Merrill 1974; Tropper 1990a).

#### 4.4.6 Weak verbs

In Ugaritic, a *weak verb* is in essence one that contains an *alif*, one that at a proto-stage contained \**y* or \**w* in any of the root positions; or one which contains a geminate (i.e., C<sub>1</sub>C<sub>2</sub>C<sub>2</sub>). Some peculiarities of the weak verb roots of Ugaritic are outlined below. Roman numerals are used to designate the position of the weak consonant in the root (1st, 2nd, 3rd).

1. Some I-*alif* roots show vagaries in orthography that indicate some form of mutation of the *alif* (quiescence, “secondary opening”?): for example, *yīḥd* versus *yūḥd*, both meaning “he seizes” (see Verreet 1983; Tropper 1990b).
2. I-*y/w* roots have all (with very rare exceptions) become I-*y* in the perfective. Most imperfectives show a bisyllabic stem, with *a* in the prefix syllable: thus, *ārd* ([ʔarid-]) “I descend”; *YD* “to know” has *i* in the prefix syllable, *idʿ* ([ʔidaʕ-]) “I know,” probably reflecting an *a* stem-vowel, because of the final guttural, and the Barth–Ginsberg Law (see §4.4.2).
3. *Hollow roots* have no consonantal element in the slot occupied by consonant II in triconsonantal roots. Most attested imperfectives have preformative vowel *a*: *ābn* ([ʔabīn-]) “I understand.” The root *Bʿ* “to enter” is written with *ū*, apparently representing [u]: *ūbū* ([ʔubūʔu]) “I enter” (indicative), *ūbā* ([ʔubūʔa]) “that I might enter” (*-a* volitive). See Pardee 1988:221.
4. III-*y/w* roots have shifted almost entirely to III-*y*; exceptions are attested for *āšlw* “I relax” (CTA 14 III 149) and *ātwt* “you have come” (CTA 4 IV 32). The zero-ending imperfective (jussive, historical “preterite”) has apparently monophthongized (*\*yabniy* → *yabni*) but, as noted above, usage is not consistent in the poetic texts, and use of historical writing (i.e., [yabni] = <ybn/ybny>) may be at the origin of some forms. See Verreet 1988 (and Sivan 1982 for III-weak nominal forms).

#### 4.4.7 Nonfinite verbals

There are two productive forms, the infinitive and the participle, which are associated with the verb but not marked for aspect or person. These forms belong by their morphology to the noun side of the grammar, by their syntax to both the noun and the verb (i.e., complementation can be either accusative or genitive).

##### 4.4.7.1 Infinitives

There was one abstract verbal noun (infinitive). The pattern in the G-stem does not seem to have been fixed (Huehnergard 1987:320), though it is likely that *katāb-* was the most common for strong roots; compare *bšāl* ([bi šaʔāli], the preposition *b* + infinitive). The infinitive in the derived stems was formed by ablaut: no *m*-preformative infinitives are attested. The nominal character of the infinitive will, of course, have appeared also in its case morphology and morphosyntax.

Though there is a syntactic usage corresponding to the so-called *infinitive absolute* construction, there does not seem to have been in Ugaritic a productive separate form so used in contradistinction to the verbal noun. One will note that it is the *katāb-* form that became the infinitive absolute in Biblical Hebrew, whereas this form functions frequently as a verbal noun in Ugaritic. Where discernible (i.e., in III-ʾ roots), the infinitive in “absolute” usage ends in *u*, homophonous with the nominative, though its origin may be different: *hm ḡmū ḡmūt* ([himma ḡamāʔu ḡamiʔti]) “If you are indeed thirsty” (CTA 4 IV 34 [Gordon 1965:79, 121, §§9.27; 13.57]).

##### 4.4.7.2 Participles

Each verbal stem has at least one corresponding verbal adjective (participle). If the stem is transitive, there will be a participle for each voice, the active and the passive. In addition, it is likely that the G-stem had two *stative verbal adjectives*, for a total of four: thus, active *kātib-*; stative *katib-* and *katub-*; passive *katūb-*.

All the derived stems except the N-stem form the participle with a prefixed *m*-. The D-stem had *u* in the preformative of the participle (cf. *mu-na-aḥ-lī-mu*, the syllabic writing of the personal name *mnḥm* “the one who brings comfort”).

The morphology of the verbal adjectives, is like that of the other adjectives and the nominal case system could in most instances specify a participle where ambiguity was potential. For example, *raḥuqu*, with final *-u*, could only be a stative participle, while *raḥuqa* could be either verbal or adjectival, but only the latter if the word could be construed as in the accusative case.

Several nouns, nonparticipial in form, are built from the Š-stem: for example, *šʔqt* “she who causes to pass on”; *šmrr* “that which causes bitterness” (i.e., “venom”).

## 4.5 Adverbs

Adverbials may be expressed by adverbial lexemes or by adverbialization of a noun, that is, by prefixing a preposition, by suffixation of an adverbial morpheme (see §4.6.2), or by using a particular form of the noun. Adverbial lexemes are either etymological nouns of which the derivation is clear (e.g., *ʔ* “now,” *ʔn* (*ʔ*+*n*) “above”) or particles (e.g., *tm*, “there”). The accusative case was the primary case used for adverbialization of nouns: for example, *qdqd* “on the head,” *ym* “for a day,” *šmm* “to the heavens.”

## 4.6 Particles

### 4.6.1 Deictic particles

The standard presentative particle is *hn* (conventional translation “behold”). The basic element is *h*-, for alongside *hn* one finds *hl*, *hln*, *hlny* (on expanding particles see §4.6.5). It is likely that this particle *hn* is at the origin of the Phoenician/Hebrew definite article (*ha* + gemination), while variant forms thereof appear in other West Semitic languages (e.g., Arabic *ʔil*- and the Aramaic postpositive article, if from *h*’ or the like).

In epistolary usage, the functions of *hn*- and *hl*- are distinct in that only the latter is used in a clearly local sense of “here” (RS 15.174:7; RS 29.093:11) whereas both function deictically, “behold.” This analysis of previously known texts is reinforced by the following recent examples in which *hl*- appears immediately before *hn*:- RS 92.2005:9 *hln hn ʔmn* (“Here, behold with me”); RS 94.2497:5 *hlny hnn b bt mlk* (“Here, behold in the house of the king”).

Rhetorical “now” is expressed by a form of this deictic particle with affixed *-t* (see §4.6.2).

The deictic element *-d*- (< *l*-*ḏ*/) was quite productive, functioning independently as a relative pronoun (see §4.3.2) and enclitically as part of the demonstrative pronoun and adjective (see §4.3.3), and as an adverbial (see §4.6.2).

### 4.6.2 Adverbial particles

As noted in §4.5, adverbials may be expressed by adverbial lexemes or by adverbialization of a noun (i.e., by prefixing a preposition, by use of the accusative case, or by suffixation of an adverbial morpheme).

The following are examples of adverbial particles: *hn*, *hnn*, *hnnny* “here”; *hl*, *hlh*, *hlny* “here,” *tm*, *tmn*, *tmny* “there”; *ht* rhetorical “now” (probably *hn* + *-t*), and *dp* “also.”

Interrogative adverbs are *iy* and *an* “where?”; *ik(y)* “how?”; *lm* (probably *l* “to/for” + *m* “what?”) “why?” The particle *ik* is often used as a rough equivalent of *lm*: for example, *ik mgy gpn wigr* “how is it that *gpn-w-igr* have come?” (not: “how have *gpn-w-igr* come?”) (CTA 3 III 33). The interrogative particles normally come at the head of the sentence. Judging from

passages difficult to understand if taken as declarative, it is likely that interrogation could also be indicated by voice inflection. There is no interrogative particle in Ugaritic such as Hebrew *hă-* which marks a following phrase as a question.

Negative adverbs are *l* (primarily indicative) and *āl* (primarily volitive). The particle *in* is used, as in Hebrew, primarily to negate nominal phrases; *bl* is rare, attested primarily in poetry and with nouns.

The primary asseveratives and negatives were written the same but probably had different vocalizations: *l* = [lā] “not” and [la] “indeed” (Huehnergard 1983:583–584); *āl* = [ʔal] “must not” and [ʔallu] (?) “must.”

Prepositional adverbialization is extremely common: for example, *l* (preposition) + *lm* (noun) = “for a long time.”

The two most common adverbial suffixes attached to nouns are *-m* and *-h*. The first cannot be defined precisely, for it appears on virtually all parts of speech. One common occurrence is on adverbial nouns, perhaps only augmenting the adverbial accusative. The second corresponds to the locative/directive *he* in Biblical Hebrew and is used both locally and temporally: thus, *šmmh* “to the heavens,” *lmh* “for a long time.” Note that, in contrast to Hebrew where the *hê* is written without *mappîq*, the Ugaritic *-h* is consonantal.

### 4.6.3 Conjunctions

The most common coordinating conjunction is *w-*, capable of linking phrases at all levels (word, clause, sentence, paragraph). The conjunction *p* (cf. Arabic *fa*) occurs more rarely, usually with a notion of cause-and-effect linkage; *dp* “also” (and expanded forms) functions most commonly at the paragraph level, and is in all probability a form of *p* produced by prefixing [ʔ]. The conjunction *u* functions both independently and correlatively (*u...u* “either...or”) and probably covers two lexemes: (i) [ʔū] “and”; (ii) [ʔô] (<\*[au]) “either/or.”

The most common subordinating conjunction is *k* “because, when, if” (comparable to Hebrew *kî*), expanded with *-y* and with *-m* (same meaning), and rarely with *-d* (the same particle as the relative pronoun), with no appreciable change of meaning. Both *im* and *hm* are attested as conditional conjunctions meaning “if.”

### 4.6.4 Prepositions

Ugaritic overlaps significantly with the other West Semitic languages in its prepositional system. Some of these are primitive particles (e.g., *b* “in,” *k* “like,” *l* “at”); others are derived from clearly identifiable verbal or nominal roots (e.g., *ʔ* “upon,” *tḥt* “under,” *dlḥr* “after”); still others are combinations of the two processes (e.g., *l* + *pn* “in front of,” *b* + *yḏ*, “in the hand/control of,” *b* + *tk* “in the midst of”). One also finds similarities in nuances and translation values (e.g., *b* = “in, within, through, by the intermediary of, by the price of,” etc.). The status of compound prepositions (i.e., those formed of two primary prepositions) is as yet uncertain: the only example attested to date is *l* + *b*, apparently meaning something like “within,” though the identity of the first element is uncertain (Rainey 1973:56; Freilich 1986:119–130).

The primary peculiarity of Ugaritic is the absence of a prepositional lexeme expressing the ablative notion “from, away from.” This absence is compensated by a complex system of verb + preposition combinations, where the translation value of the preposition can be determined only by usage and by context (Pardee 1975, 1976, with a discussion of prepositional semantic ambiguity). The prepositional system as a whole appears to function primarily to denote position rather than direction, a stative notion rather than a motional one. Directionality and motion were supplied primarily by the verb. What this means in



practice is that virtually any preposition may appear in expressions of the ablative and the modern reader must depend on elements other than the preposition itself to reach a proper interpretation of a passage. The following passage is instructive, for it includes a preposition with “opposite meanings” in the expression of a “from . . . to” situation, but along standard Ugaritic lines – that is, by means of different verb + preposition combinations (*yrd l* “descend from,” *yṯb l* “sit upon”): *yrd l ks ṯyṯb l hdm w l hdm yṯb l drṣ* “he descends from the throne, he sits upon the footstool, and [he descends] from the footstool, he sits upon the earth” (CTA 5 VI 12–14).

There are also certain functional differences between Ugaritic and the other Semitic languages: for example, the increased use of *ʾm* “with” to denote the end-point of a trajectory; *l* used to form compound numbers. Moreover, different lexemes occur: for example, *zr* “back,” yielding “on top of”.

The substantive following a preposition is, as nearly as can be determined, always in the genitive case (as in Akkadian, Arabic, etc.). This is shown for Ugaritic by nominal phrases spelled with a final *alif* character: for example, *l ksī* [lê kussaʾi] “to the chair/throne”; *b nšī* [bi našāʾi] “in his lifting, when he lifts.”

The case system still being in force, no prepositional particle has developed in Ugaritic to mark the direct object of a transitive verb, such as, for example, Phoenician *ʾyt*, Hebrew *ʾōt*- and *ʾet/ʾēt*).

#### 4.6.5 Enclitic particles

Ugaritic makes use of a baroque array of enclitic particles (Aartun 1974, 1978), the disentanglement of which is made all the more difficult by the absence of vocalized texts. These particles can be joined to all parts of speech and are capable of accretion one to another (e.g., *h+n+n+y*). Particles that apparently have little more than an “emphatic” function may develop a paradigmatic function alongside particles of more precisely definable origin: for example, *hnd* “this” = *h* (deictic particle) + *n* (particle) + *d* (relative pronoun), alongside *hnk* “that” = *h* (deictic particle) + *n* (particle) + *k* (particle). The principal enclitic particles are these:

1. *-d* = relative pronoun that can function as a compounding element with nouns (e.g., *šbʿd*, “sevenfold”) and with other particles (e.g., *hnd* “this”), and is expandable (e.g., *šbʿīd*, also “sevenfold”).
2. *-h* = adverbial (see §4.6.2).
3. *-y* = enclitic particle, particularly as expander to another particle (e.g., *hn+n+y*).
4. *-k* = enclitic particle, particularly in *hnk* “that.”
5. *-m* = enclitic particle used on all parts of speech (see §4.6.2 for use with adverbials).
6. *-n* = enclitic particle used on all parts of speech. One particularly striking usage is the “*n* of apodosis” (Hoftijzer 1982); in certain omen texts characterized by a repetitive protasis–apodosis structure, the first word in the apodosis, if a singular noun in the absolute state, has enclitic *-n* (Pardee 1986:126, 129; Tropper 1994b:466–469).
7. *-t* = enclitic particle, particularly as expander to another particle (e.g., *ht < hn+t* with assimilation; *hn+d+t*).

#### 4.7 Compounds

Compound verbs are virtually unknown in old West Semitic, and compound nouns are rare (the primary case cited for Ugaritic is *bl mt* “not death” used in parallel with *ḥym* “life” in



CTA 17 VI 27). Complex prepositional phrases, made up of a preposition and a common noun, are certainly well attested (see §4.6.4, and the list and discussion in Pardee 1976:306–310), but it is in most cases dubious that the complex phrase had evolved as a lexical entity of which the compositional elements were no longer perceived.

## 4.8 Derivational processes

Because Ugaritic is a poorly attested one-period language, it is hardly possible to describe synchronic derivational processes. Viewing the language comparatively, however, it appears clear that the known state of the language reflects a number of such processes, for one can spot certain morphemes of which the function is best described as derivational.

Within categories, the generating of new particles by particle accretion is perhaps the clearest derivational process (better so termed perhaps than as compounding), though the semantics of the process are unclear in most cases.

Across categories, the nominal system, particularly the *m*- and *t*-prefixes and the *-n* suffix noted in §4.2.1, and certain ablaut forms (e.g., *qattāl* to express a *nomen professionalis*), usually reflect a deverbal notion rather than an inner-nominal process. The suffixing of particles to nominal elements, to the extent that these particles were not perceived by native speakers as lexical items, also represents a form of derivation.

Across subcategories, the case of the *nisbe* ending, by which nouns are transformed into adjectives (see §4.2.7), is the clearest case of a derivational morpheme.

## 5. SYNTAX

The relative dearth of prose texts mentioned in the introduction makes it difficult to ascertain a normative prose syntax, while the lack of vocalized texts makes some aspects of morphosyntax difficult to ascertain precisely.

### 5.1 Word order

On the phrase level, there are two primary nominal phrases: the genitival and the adjectival.

The *genitival phrase* is the common Semitic construct state: X of Y (see §4.2.6.1). The first element is in the case required by context, the second in the genitive. It can denote the various relationships well known elsewhere (subjective genitive, objective genitive, genitive of identification, genitive of material, etc.). No lexical or pronominal element may intervene between the members of a construct chain, only enclitic particles.

The *adjectival phrase* is of two types, (i) the phrase-level or attributive, in which the adjective follows the noun and agrees in gender, number, and case; and (ii) the sentence-level or predicative, in which the adjective may either precede or follow the noun and agrees in gender, number, and case. An attributive adjective modifying any member of a construct chain must come at the end of the chain (e.g., *ḥbr ktr ṭbm* “the companions of Kothar, the good ones” [CAT 1.108:5]). Apparent attributive adjectives preceding the noun they modify are most frequently substantives in construct with the noun (*nʾmt šnt il* “the excellent ones of the years of El” = “the most excellent years of El” (CAT 1.108:27)).

In nominal sentences, word order is essentially free with fronting used for topicalization. Thus *hw mlk* will denote “he, not someone else, is king” (an “identifying” sentence), *mlk hw* “he is king, he is not something else” (a “classifying” sentence).

In the simplest verb phrase, consisting of verb + pronoun, the subject pronoun is part of the verbal form itself, suffixed in the perfective, prefixed in the imperfective. The primary variation occurs through addition of an independent pronoun for “emphasis,” creating a formal *casus pendens* (e.g., *ātm bštm w ān šnt* “as for you, you may tarry but as for me, I’m off” (CTA 3 IV 77)). The independent pronoun may precede or follow the verbal unit. The simple verb phrase is by definition a sentence: SUBJECT + PREDICATE (imperfective) or PREDICATE + SUBJECT (perfective).

In verbal sentences one finds fronting for topicalization as in, for example, the following (RS 34.124:25–28 [Bordreuil and Pardee 1991:148]):

ybnm hlk ‘m mlk āmr wybl hw mīt hrš  
 Yabninu went to king of Amurru and he took he one hundred of gold  
 SUBJECT : VERB :: VERB : SUBJECT  
 “Yabninu (not someone else) went to the king of Amurru, and he took, did he,  
 one hundred [pieces of] gold”

According to one study, there is a strong tendency in poetry to place the object phrase close to the verb, either before it or after it (Wilson 1982:26).

The verb is usually fronted in subordinate clauses where the subject is known (CAT 2.16: 6–8):

ūmy td‘ ky ‘rbt l pn špš  
 My mother must know that I entered to face of Sun  
 “May my mother know that I have entered before the ‘Sun’”

The *subject–verb(–object–modifier)* order is regular in the first clause of apodoses in texts of the omen and hippiatric genres (the basis structure of sentences in both genres is protasis–apodosis). This order cannot be demonstrated to be the result of influence from another language (Pardee 1986:128–129), and probably reflects, therefore, systematized topicalization (Troppe 1994b:469–471), though the general absence of *w*- of apodosis (see §5.3.2) and the presence of *-n* of apodosis (see §4.6.5) in these texts must be included in an explanation of the phenomenon.

On the basis of present evidence, therefore, it is impossible to say that Ugaritic is a primarily VSO language, though, as in Biblical Hebrew, this is certainly the case in subordinate clauses.

## 5.2 Coordinate clauses

Coordination is indicated most commonly by *w*-, by *p*- when effect is denoted (for coordinating conjunctions see §4.6.3). Asyndesis is fairly frequent at the sentence (and paragraph) level, common at the phrase level.

## 5.3 Subordinate clauses

The principal types of subordinate clauses are (i) relative, (ii) conditional, and (iii) a variety of temporal/circumstantial, causal, resultative, and completative (object) clauses most commonly introduced by *k* ([kī]) when lexically marked (the conjunction is written both *k* and *ky* [Pardee 1984a:214–215]). The whole concept of “subordinate” clause is rendered murky by the frequent use of the so-called *w* (or more rarely *p*) of apodosis (see §5.3.2) – that is, heading the main clause with *w* or *p* when it follows the “subordinate” clause. The details have not been worked out for Ugaritic, and the state of the corpus renders a comprehensive

view difficult; points of similarity with Biblical Hebrew indicate that the overall situation in Ugaritic may not have been dissimilar (cf. Gross 1987).

### 5.3.1 Relative clauses

Explicit relative clauses are marked by a preceding *d/dt*. Relative adverbials are usually marked: for example, *ādr̥m d b gr̥n* “the leaders who are at the threshing floor” (CTA 17 V 7). Unmarked relative verbal clauses are difficult to spot because the notion of person is marked in the verb, and SUBJECT is by definition included in both verbs. An example upon which there is general agreement is *yd mḥṣt āqht ḡzr tmḥṣ ālpm ib* “The hand [that] struck Hero Aqhat will strike the enemy by thousands” (CTA 19 IV 220–221).

The relative pronoun functions at both the phrase level – *il d pīd* “god of mercy” (CTA 4 II 10 and frequent) – and at the sentence level – subject (A), object (B), adverbial (C) below:

- A. *il ... d yšr*  
god- ... who he sings  
“the god ... who sings” (CAT 1.108:2–3);
- B. *skn d šlyt tryl*  
sacred stone that she caused to ascend T̥arriyelli  
“Sacred stone which T̥arriyelli offered” (CAT 6.13:1–2)
- C. *mt hr̥nmy d in bn lh*  
Man Harnamite who there is not son to him  
“the Harnamite man to whom there is no son” = “who has no son” (CTA 17 I 19)

Note the relative genitive construction.

- ḥry ... d k n'm 'nt n'mh*  
Ḥurraya ... who like beauty of Anat her beauty  
“whose beauty is like Anat’s” (CTA 14 VI 289–292)

The relative pronoun may either have an explicit antecedent, as in the examples just cited, or be used “absolutely”: for example, *p d in b bty ttn* “For what is not in my house shall you give” (CTA 14 III 142).

The conjunction *k(y)* does not function as a relative particle (see §5.3.3).

### 5.3.2 Conditional clauses

Conditions may be marked by *hm* or (less frequently) *im* and tend to precede the main clause. Conditional clauses may be unmarked. A lexical distinction between real and unreal conditions is as yet unknown. The main clause following the conditional clause may or may not be preceded by the so-called *w* or *p* of apodosis (for [A] below see Bordreuil and Caquot 1980:359–360; Pardee 1984a:222; and for [B], see Bordreuil and Pardee 1991:147):

- A. *hm ymt w ilḥmn ank*  
If he dies and I indeed fight I  
“If he should die, I will go on fighting on my own” (RIH 78/12:19–22)
- B. *im ht l b mṣqt yṭbt qrt p mn likt*  
If behold to in distress she/it is sitting city and what I sent

ʾn̄k lḥt bt mlk ʾmr  
 I tablet of daughter of king of Amurru  
 “So if the city is remaining undecided, then for what reason did I send a letter  
 regarding the daughter of the king of Amurru?” (RS 34.124:20–24).

### 5.3.3 Other subordinate clauses

Temporal/circumstantial phrases may be expressed as a true clause, that is, as conjunction + finite verb (*k tḏbr* “when you speak”), or as a prepositional phrase consisting of preposition + infinitive (*b šāl* “in [his] asking” = “when he asks”).

Causal and resultative clauses are not nearly so frequent as in Biblical Hebrew. Causal clauses, particularly, are often difficult to distinguish from temporal/circumstantial clauses. A reasonably clear example of each, respectively, follows:

- A. tšmḥ... ʾṯrt ...k mt ʾliyn bʿl  
 She rejoices... Athirat ...that is dead/has died most mighty Baal  
 “May Aṯirat rejoice because Mighty Baal is dead” (CTA 6 I 39–42)  
 B. mn (!) krt k ybky  
 What Kirta that he weeps  
 “Who/what is Kirta that he should weep?” (CTA 14 I 38–39)

The principal marker of completative (object) clauses is *k(y)*:

ʾmy tḏ ky ʿrbt l pn špš  
 My mother must know that I entered to face of Sun  
 “May my mother know that I have entered before the ‘Sun’ ”

A particularly common word order in letters is a construction in which a *casus pendens* is followed by a subordinate clause marked by *k(y)*, with the main clause coming only after these two clauses

lḥt šlm k likt ʾmy ht ʿmny  
 Tablet of well being that CONJ. she sent my mother behold with me  
 kll šlm  
 everything is well  
 “As for the letter of greeting, as for the fact that my mother sent [it], behold with  
 me everything is fine” (CAT 2.34:5–7)

For this interpretation of the structure, see Pardee 1977:7–8.

## 5.4 Agreement

Personal pronouns agree in (i) person, gender, and number with an appositional verbal form (*ʾn̄k ḏḥwy* “I give life” [CTA 17 VI 32]); and (ii) gender, number, and case with an appositional or predicate noun (*ʾt ʾmy* “you, my mother” [CAT 2.30:20–21]; *ʾt ḏḥ* “you are my brother” [CTA 18 I 24]) and with predicate adjectives (*ḏbḥn ndbḥ hw* “the sacrifice [-*n* of apodosis], sacrificed is it”) (where *ndbḥ* is an N-stem participle; CTA 40:9 and parallels). The relative pronoun agrees in gender and number with its antecedent; whether the case of the relative pronoun itself is decided by the case of the antecedent or by the function of the relative pronoun in its clause cannot be determined (cf. Arabic, where case agreement is decided in the relative clause).

Adjectives agree in gender, number, and case with the modified noun. Demonstrative pronouns agree in gender and number with the antecedent (case unknown), while demonstrative adjectives agree in gender, number, and case with the modified noun.

Interrogatives and indefinite pronouns do not show agreement.

## 6. LEXICON

Ugaritic fits the common Semitic and common West Semitic pattern in kinship terms (*ḏb* “father,” *ūm* “mother,” etc.), tree names (*ḏrz* “cedar,” etc.), geographical terms (*nhr* “river,” etc.), with some notable peculiarities: for example, *ḥwt* ([*huwwat*-]) “land (geographical–political entity),” alongside *ḏrṣ* “earth, ground” and *bld* “homeland.”

When deciphering a Ugaritic text, one finds points of lexical contact with all of the Semitic languages. Because of the small number of texts, the image of the Ugaritic scholar deciphering a text on the basis of various Semitic dictionaries is not totally false, though with the increase in number of reasonably well-understood texts, inner-Ugaritic lexicography is becoming more practicable. The apparent heterogeneity of the Ugaritic lexicon may be explained in two ways: (i) the archaic nature of the language (cognates with other Semitic languages will thus be largely with retentions in those languages); (ii) the relatively poor corpus of texts in the languages with which Ugaritic appears most closely related linguistically (thus, if Hebrew and Phoenician were attested more extensively, there would be fewer exclusive isoglosses between Arabic and Ugaritic).

The principal motion verbs are useful language/dialect isoglosses (e.g., for all the similarities between Hebrew and Aramaic, the systems of motion verbs are quite different in the two languages). Here Ugaritic falls directly in the Hebrew/Phoenician group: *hlk* “go,” *yrd* “descend,” *ly* “ascend,” *b* “enter” (alongside *rb*), *yṣ* “exit,” *tb* “return.” Some verbs of movement that can also denote the state attained are: *qm* “arise,” *škb* “lie down,” *md* “stand,” *rkb* “mount.” Primary motion verbs peculiar to Ugaritic are *tb* “go away,” *mgy* “go to, arrive at” (apparently < *MZY*), and *yšql* “he arrives,” used only in poetry and in the imperfective.

Expressions of existence resemble most closely the later Northwest Semitic pattern: there are positive and negative quasi-verbs, *lt* and *in*, respectively, corresponding, for example, to Hebrew *yēš* and *ʾayin/ ʾēʾn*, as well as the verb *kn* (*nʾmn ykn* “there will be prosperity” [*RIH* 78/14:3, Bordreuil and Caquot 1980:352–353]), which corresponds to the regular verb “to be” in Phoenician (and Arabic) and to the more strongly marked verb “to be stable” in Hebrew.

In spite of the cosmopolitan nature of the city of Ugarit, there are relatively few readily identifiable loanwords: *ḥtt* “silver” is an apparent example from Hittite, *khṭ* “chair, throne” an example from Hurrian.

## References and abbreviations

- Aartun, K. 1974. *Die Partikeln des Ugaritischen*, part 1. *Alter Orient und Altes Testament* 21/1. Kevelaer: Butzon and Bercker; Neukirchen-Vluyn: Neukirchener Verlag.
- . 1978. *Die Partikeln des Ugaritischen*, part 2. *Alter Orient und Altes Testament* 21/2. Kevelaer: Butzon and Bercker; Neukirchen-Vluyn: Neukirchener Verlag.
- Blau, J. 1972. “Marginalia Semitica.” *Israel Oriental Studies* 2:57–82.
- . 1977. “‘Weak’ phonetic change and the Hebrew *šin*.” *Hebrew Annual Review* 1:67–119.
- Bordreuil, P. 1979. “L’inscription phénicienne de Sarafand en cunéiformes alphabétiques.” *Ugarit-Forschungen* 11:63–68.

- Bordreuil, P., A. Bounni, N. Saliby, E. Lagarce, and J. Lagarge. 1987. "Les fouilles de Ras Ibn Hani (Syrie) en 1984 et 1986." *Académie des Inscriptions et Belles-Lettres, Comptes Rendus*, pp. 274–301.
- Bordreuil, P., and A. Caquot. 1980. "Les textes en cunéiformes alphabétiques découverts en 1978 à Ibn Hani." *Syria* 57:343–373.
- Bordreuil, P., and D. Pardee. 1991. *Une bibliothèque au sud de la ville (Quartier "Sud-Centre"). Les textes de la 34<sup>e</sup> campagne (1973), deuxième partie: les textes en cunéiformes alphabétiques*. Ras Shamra-Ougarit VII. Paris: Editions Recherche sur les Civilisations.
- . 1995a. "L'épigraphie ougaritique: 1973–1993." In M. Yon, M. Sznycer, and P. Bordreuil (eds.), *Le pays d'Ougarit autour de 1200 av. J-C*, pp. 27–32. Ras Shamra-Ougarit XI. Paris: Editions Recherche sur les Civilisations.
- . 1995b. "Un abécédaire du type sud-sémitique découvert en 1988 dans les fouilles archéologiques françaises de Ras Shamra-Ougarit." *Académie des Inscriptions et Belles-Lettres, Comptes Rendus*, pp. 855–860.
- . 2001. "Textes alphabétiques enougaritique." In M. Yon, D. Arnaud (ed.), *Études Ougaritiques. I. Travaux 1985–1995*. Ras Shamra-Ougarit XIV. Paris: Editions Recherche sur les Civilisations.
- CAT 1995 = Dietrich, M., O. Loretz, and J. Sanmartin. *The Cuneiform Alphabetic Texts from Ugarit, Ras Ibn Hani and Other Places* (KTU: 2nd enlarged edition). Abhandlungen zur Literatur Alt-Syrien-Palästinas und Mesopotamiens 8. Münster: UGARIT-Verlag.
- Contenson, H. de. 1992. *Préhistoire de Ras Shamra: Les sondages stratigraphiques de 1955 à 1976*. Ras Shamra-Ougarit VIII/1–2. Paris: Editions Recherche sur les Civilisations.
- CTA 1963 = Herdner, A. *Corpus des tablettes en cunéiformes alphabétiques découvertes à Ras Shamra-Ougarit de 1929 à 1939*. Mission de Ras Shamra 10. Bibliothèque Archéologique et Historique 79. Paris: Imprimerie Nationale; Paul Geuthner.
- Cunchillos, J.-L. 1993. *Banco de Datos Filológicos Semíticos Noroccidentales. Primera Parte: Datos Ugaríticos. I. Textos Ugaríticos*. Madrid: Instituto de Filología.
- . 1995. *Banco de Datos Filológicos Semíticos Noroccidentales. Primera Parte: Datos Ugaríticos. II/1–3. Concordancia de Palabras Ugaríticos en Morfología Desplegada*. Madrid: Instituto de Filología.
- Dietrich, M., and O. Loretz. 1988. *Die Keilalphabet. Die phönizisch-Kanaanäische und altarabischen Alphabete in Ugarit*. Abhandlungen zur Literatur Alt-Syrien-Palästinas 1. Münster: UGARIT-Verlag.
- Fenton, T. 1970. "The absence of a verbal formation \**yaqattal* from Ugaritic and Northwest Semitic." *Journal of Semitic Studies* 15:31–41.
- Freilich, D. 1986. "Is there an Ugaritic deity *Bbt*?" *Journal of Semitic Studies* 31:119–130.
- Freilich, D., and D. Pardee. 1984. "{z} and {t} in Ugaritic: a re-examination of the sign-forms." *Syria* 61:25–36.
- Gordon, C. 1965. *Ugaritic Textbook: Grammar, Texts in Transliteration, Cuneiform Selections, Glossary, Indices*. Analecta Orientalia 38. Rome: Pontifical Biblical Institute.
- . 1997. "Ugaritic phonology." In A. S. Kaye (ed.), *Phonologies of Asia and Africa*, pp. 49–54. Winona Lake, IN: Eisenbrauns.
- Greenstein, E. 1976. "A Phoenician inscription in Ugaritic script?" *Journal of the Ancient Near Eastern Society of Columbia University* 8:49–57.
- Gross, W. 1987. *Die Pendenskonstruktion im Biblischen Hebräisch. Arbeiten zu Text und Sprache im Alten Testament* 22. St. Ottilien: EOS Verlag.
- Hoftijzer, J. 1982. "Quodlibet Ugariticum." In G. van Driel, Th. J. H. Krispijn, M. Stohl, et al. (eds.), *Zikir Šumim. Assyriological Studies Presented to F. R. Kraus on the Occasion of his Seventieth Birthday*, pp. 121–127. Nederlands Instituut voor het Nabije Oosten Studia Francisci Scholten Memoriae Dicata 5. Leiden: Brill.
- Huehnergard, J. 1983. "Asseverative \**la* and hypothetical \**lu/law* in Semitic." *Journal of the American Oriental Society* 103:569–593.
- . 1986. "A Dt stem in Ugaritic?" *Ugarit-Forschungen* 17:402.
- . 1987. *Ugaritic Vocabulary in Syllabic Transcription*. Harvard Semitic Studies, 32. Atlanta: Scholars Press.
- Liverani, M. 1963. "Antecedenti del diptotismo arabo nei testi accadici di Ugarit." *Rivista degli Studi Orientali* 38:131–160.



- . 1964. "Elementi innovativi nell'ugaritico non letterario." *Atti della Accademia Nazionale dei Lincei, Rendiconti della Classe di Scienze morali, storiche e filologiche*, vol. VIII, fascicle 19, pp. 173–191.
- Mallon, E. 1982. *The Ugaritic Verb in the Letters and Administrative Documents*. Ph.D. dissertation, The Catholic University of America. Ann Arbor, MI: University Microfilms.
- Merrill, E. 1974. "The Aphel causative: does it exist in Ugaritic?" *Journal of Northwest Semitic Languages* 3:40–49.
- Nougayrol, J. 1968. "Textes suméro-accadiens des archives et bibliothèques privées d'Ugarit." In J.-C. Courtois (ed.), *Ugaritica V. Nouveaux textes accadiens, hourrites et ugaritiques des archives et bibliothèques privées d'Ugarit, commentaires des textes historiques (première partie)*, pp. 1–446. Bibliothèque Archéologique et Historique 80. Mission de Ras Shamra 16. Paris: Imprimerie Nationale.
- Pardee, D. 1975. "The preposition in Ugaritic." *Ugarit-Forschungen* 7:329–378.
- . 1976. "The preposition in Ugaritic." *Ugarit-Forschungen* 8:215–322, 483–493.
- . 1977. "A new Ugaritic letter." *Bibliotheca Orientalis* 34:3–20.
- . 1983–1984. "The Letter of Puduhepa: the text." *Archiv für Orientforschung* 29–30:321–329.
- . 1984a. "Further studies in Ugaritic epistolography." *Archiv für Orientforschung* 31:213–230.
- . 1984b. "Three Ugaritic tablet joins." *Journal of Near Eastern Studies* 43:239–245.
- . 1986. "The Ugaritic *šumma izbu* text." *Archiv für Orientforschung* 33:117–147.
- . 1988. *Les textes para-mythologiques de la 24<sup>e</sup> campagne (1961)*. Ras Shamra-Ougarit IV. Paris: Editions Recherche sur les Civilisations.
- . 1993a. Review of *The Syntax of the Verb in Classical Hebrew Prose*, by Alviero Niccacci, 1990. *Journal of Near Eastern Studies* 52:313–314.
- . 1993b. Review of *Modi ugaritici: Eine morpho-syntaktische Abhandlung über das Modalsystem in Ugaritische*, by E. Verreet, 1988. *Journal of Near Eastern Studies* 52:314–317.
- . 1995. Review of *Studies in Verbal Aspect and Narrative Technique in Biblical Hebrew Prose*, by Mats Eskhult, 1990. *Journal of Near Eastern Studies* 54:64–66.
- . 1997. *The Ba'lu Myth: The Context of Scripture*. In W. Hallo and K. Younger (eds.), *The Context of Scripture. Vol. I: Canonical Compositions from the Biblical World*, pp. 241–274. Leiden: Brill.
- . 2000. *Les textes rituels*. Ras Shamra Ougarit. Paris: Editions Recherche sur les Civilisations.
- Rainey, A. 1973. "Gleanings from Ugarit." *Israel Oriental Studies* 3:34–62.
- . 1996. *Canaanite in the Amarna Tablets. A Linguistic Analysis of the Mixed Dialect Used by Scribes from Canaan* (4 vols.). Handbuch der Orientalistik 25. Leiden: Brill.
- RIH = text discovered at Ras Ibn Hani.
- RS = text discovered at Ras Shamra.
- Saadé, G. 1979. *Ougarit: Métropole cananéenne*. Beirut: Imprimerie Catholique.
- Segert, S. 1983. "The last sign of the Ugaritic alphabet." *Ugarit-Forschungen* 15:201–218.
- . 1984. *A Basic Grammar of the Ugaritic Language*. Berkeley: University of California Press.
- . 1988. "The Ugaritic voiced postvelar in correspondence to the emphatic interdental." *Ugarit-Forschungen* 20:287–300.
- Singer, A. 1948. "The vocative in Ugaritic." *Journal of Cuneiform Studies* 2:1–10.
- Sivan, D. 1982. "Final triphthongs and final *yul/a/i* – *wul/a/i* diphthongs in Ugaritic nominal forms." *Ugarit-Forschungen* 14:209–218.
- . 1992. "Notes on the use of the form *qatal* as the plural base for the form *qatl* in Ugaritic." *Israel Oriental Studies* 12:235–238.
- Tropper, J. 1990a. *Der ugaritische Kausativstamm und die Kausativbildungen des semitischen. Eine morphologisch-semantische Untersuchung zum Š-Stamm und zu den umstrittenen nichtsibilantischen Kausativstämmen des Ugaritischen*. Abhandlungen zur Literatur Alt-Syrien-Palästinas 2. Münster: UGARIT-Verlag.
- . 1990b. "Silbenschließendes Aleph im Ugaritischen – Ein neuer Versuch." *Ugarit-Forschungen* 22:359–369.
- . 1991. "Finale Sätze und *yqtl*-Modus im Ugaritischen." *Ugarit-Forschungen* 23:341–352.
- . 1993a. "Auf dem Weg zu einer ugaritischen Grammatik." In M. Dietrich and O. Loretz (eds.), *Mesopotamica-Ugaritica-Biblica. Festschrift für Kurt Bergerhof zur Vollendung seines 70.*

- Lebensjahres am 7. Mai 1992*, pp. 471–480. Alter Orient und Altes Testament 232. Kevelaer: Butzon and Bercker; Neukirchen-Vluyn: Neukirchener Verlag.
- . 1993b. “Morphologische Besonderheiten des Spätugaritischen.” *Ugarit-Forschungen* 25:389–394.
- . 1994a. “Das ugaritische Konsonanteninventar.” *Journal of Northwest Semitic Languages* 20(2):17–59.
- . 1994b. “Zur Grammatik der ugaritischen Omina.” *Ugarit-Forschungen* 24:457–472.
- . 1995. “Das altkanaanäische und ugaritische Verbalsystem.” In M. Dietrich, and O. Loretz (eds.), *Ugarit. Ein ostmediterranes Kulturzentrum im Alten Orient. Ergebnisse und Perspektiven der Forschung. Band I: Ugarit und seine altorientalische Umwelt*, pp. 159–170. Abhandlungen zur Literatur Alt-Syrien-Palästinas, 7. Münster: UGARIT-Verlag.
- Verreet, E. 1983. “Das silbenschiessende Aleph im Ugaritischen.” *Ugarit-Forschungen* 15:223–258.
- . 1986. “Beobachtungen zum ugaritischen Verbalsystem II,” *I17*:319–344.
- . 1988. *Modi Ugaritici. Eine Morpho-syntaktische Abhandlung über das Modalsystem im Ugaritischen*. Orientalia Lovaniensia Analect, 27. Leiden: Peeters.
- Wilson, G. 1982. “Ugaritic word order and sentence structure in KRT.” *Journal of Semitic Studies* 27:17–32.
- Yon, Marguerite. 1997. *La cité d'Ougarit sur le tell de Ras Shamra. Guides Archéologiques de l'Institut Français d'Archéologie du Proche-Orient, No. 2*. Paris: Editions Recherche sur les Civilisations.
- Zevit, Z. 1983. “The question of case endings on Ugaritic nouns in *status constructus*.” *Journal of Semitic Studies* 28:225–232.

*Nota bene:* A monumental grammar of Ugaritic appeared too late to be used in preparing this overview. It must be considered essential for the future study of Ugaritic.

Tropper, J. 2000. *Ugaritische Grammatik*. Alter Orient und Altes Testament 273. Münster: Ugarit-Verlag.



# Hebrew

P. KYLE McCARTER, JR.

## 1. HISTORICAL AND CULTURAL CONTEXTS

### 1.1 The position of Hebrew within the Semitic languages

Hebrew, the language of ancient Israel and Judah and their descendant Jewish communities, is a Northwest Semitic language. Northwest Semitic and Arabic constitute Central Semitic, which is a subgroup of West Semitic, one of the two primary divisions of the Semitic branch of the larger Afro-Asiatic family (Ch. 6, §§1–2). Within Northwest Semitic, Hebrew is classified as Canaanite as distinct from Aramaic. Other members of the Canaanite subgroup include the dialect of the city-state of Ugarit (cf. Ch. 9, §1) in the Late Bronze Age (c. 1550–1200 BC), and the languages of Israel's immediate neighbors in the Iron Age (c. 1200–586 BC), namely, Phoenician (Ch. 11) and the Transjordanian languages of Ammonite, Moabite, and Edomite (Ch. 12).

### 1.2 Stages in the development of Ancient Hebrew

Although linguistic features found in the limited surviving evidence for the Canaanite dialects of the Late Bronze Age anticipate some of the distinctive characteristics of Iron Age Hebrew, it is unlikely that Hebrew emerged as a discrete language before the end of the Late Bronze Age and the beginning of the Iron Age. Prosodic and linguistic studies suggest that the earliest poetry preserved in the Hebrew Bible may have been composed before the end of the second millennium BC, and this poetry represents the first identifiable phase of the language, which is called *Archaic* or *Archaic Biblical Hebrew* (before c. 1000 BC).

No extant inscription that can be identified specifically as Hebrew antedates the tenth century BC, and Hebrew inscriptions in significant numbers do not begin to appear before the early eighth century BC. Nevertheless, the Hebrew of the Iron Age inscriptions that do survive, especially those from Judah, is essentially the same as the Hebrew found in the biblical Primary History (Genesis–2 Kings) and the original portions of the books of the pre-exilic prophets. This form of Hebrew constitutes the classical phase of the language, which is known as *Classical* or *Biblical Hebrew* (BH) and corresponds to the speech of the kingdom of Judah from its formation to the Babylonian Exile (c. tenth–sixth centuries BC). The Hebrew of post-exilic Judah, which is represented by inscriptions of the Persian and Hellenistic periods and especially by the later biblical literature (c. sixth–second centuries BC), is called *Late Classical* or *Late Biblical Hebrew* (LBH). The Samaritan Pentateuch, which seems to have been independent of Jewish tradition by the late second century BC, is also an important witness to the Hebrew of this period.

The Hebrew of the early post-biblical period is represented by the Hebrew of the Dead Sea Scrolls and especially that of the Mishnah and other rabbinical literature. As noted below (§1.3), the literary documents from Qumran exhibit substantial continuity with Late Biblical Hebrew, while the few nonliterary documents stand much closer linguistically to Rabbinic Hebrew. From the viewpoint of the development of the language, there is a distinction between the Hebrew of the early rabbinical works – the Mishnah, the Tosefta and certain other, primarily halakhic compositions (c. first–third centuries AD) – and that of the later rabbinical works – the Jerusalem and Babylonian Talmuds and certain other, primarily haggadic compositions (fourth century AD and later). Viewed as a whole, this phase in the development of the language is called *Middle or Rabbinic Hebrew* (RH). Another important witness to Hebrew in late antiquity is the Hexapla, the six-column critical edition of the Old Testament compiled by the Church father Origen of Caesarea; in his second column (*Secunda*), Origen produced a Greek transliteration of the Hebrew text that reflects the pronunciation of the first half of the third century AD.

In this chapter, primary attention is given to the classical phase of Hebrew (BH), but important divergent or innovative features of the other ancient phases of Hebrew (LBH and RH) are noted. The subsequent phases of the language – *Medieval Hebrew* and *Modern or Israeli Hebrew* – fall outside the scope of the discussion.

### 1.3 The speech communities of Ancient Hebrew

In a general sense, the emergence of Hebrew as a discrete language corresponded to the emergence of Israel as a discrete polity in the central hill country of Palestine in the last centuries of the second millennium BC. By the tenth century BC, two Hebrew-speaking states had been established, Israel to the north in the Samarian hills and portions of central Transjordan and Galilee, and Judah to the south in the Judean hills with its capital at Jerusalem. The modest corpus of surviving inscriptions from the northern kingdom is sufficient to show that its dialect displayed features that were significantly different from that of Judah, as it is known from a more generous inscriptional corpus and, indeed, from the Hebrew Bible itself.

The two Iron Age states survived until 722 BC in the case of Israel, when its capital, Samaria, fell to the Assyrians (precipitating the extinction of the northern dialect), and until 586 BC in the case of Judah, when Jerusalem was destroyed by the Babylonians. Despite these catastrophes, Hebrew endured as a spoken and literary language in Palestine throughout the second half of the first millennium BC. During this period the use of Aramaic increased steadily in the larger region, becoming the regnant language of both Samaria and Galilee, and, beginning in the third century BC, Greek was introduced to many of the major cities of Palestine. Nevertheless, Hebrew persisted, alongside Aramaic, as a spoken language in Judah (or Judaea) proper into the rabbinic period.

Although Biblical Hebrew, Late Biblical Hebrew and the Hebrew of the literary manuscripts from Qumran constitute a unilinearly evolving dialect, descended from the language of pre-exilic Judah, Rabbinic Hebrew exhibits features that set it apart from this development. Since most of the literature of Rabbinic Hebrew is highly technical in character, it was once supposed that it was a language spoken only by scholars or even an artificially confected language that was never spoken at all. But the discovery and linguistic analysis of the nonliterary or quasi-literary documents from Qumran – especially the Copper Scroll and the Halakhic Letter (MMT) – and of the Bar Kochba correspondence from the Wadi Murabbaʿat and the Nahal Hever show that Rabbinic Hebrew was a popularly spoken language in the early centuries of the Common Era. Although many of the features of Rabbinic

Hebrew that diverge from Biblical Hebrew can be traced to contemporary influences, such as the prevalence of Aramaic and Greek, many others seem to be dialectal survivals from a much earlier period, when an ancestral form of Rabbinic Hebrew existed alongside Biblical Hebrew. The beginning of the demise of Rabbinic Hebrew as a spoken language is probably to be traced to the Roman suppression of the Second Jewish Revolt in AD 135 and the accompanying depredations, including the deportation of many Jews and the flight of others into the Aramaic-speaking Galilee. Even under these conditions Hebrew continued to be heard in some circles, but the primary language of Jews in the Roman diaspora was Greek just as the primary language of the long-established Babylonian diaspora was Aramaic. In Palestine, too, Rabbinic Hebrew was eventually replaced by Aramaic as a spoken language and survived only as the scholarly language of the Galilean exile community.

## 2. WRITING SYSTEMS

### 2.1 The Hebrew consonantal script

The earliest inscriptions unambiguously identifiable as Hebrew are written in a distinctive form of the consonantal writing system that served as the national script of both Israel and Judah in the Iron Age. This Hebrew script arose as a branch of the Phoenician, through which it was descended from the archaic consonantal script of the second millennium BC. The intermediary role of Phoenician is shown by the fact that the two scripts share a sign inventory that is fully representative of the consonantal phonology of Phoenician but insufficient to represent all the consonantal phonemes of Hebrew. In particular, only one sign corresponds to the Proto-Semitic phonemes /š/ and /ś/, a situation that is adequate for Phoenician, where the two consonants have merged (see Ch. 11, §3.1), but not for Hebrew, where they remain distinct (see §3.1 below).

After the Babylonian destruction of Jerusalem in 586 BC, the Hebrew script fell into disuse. Hebrew came to be written primarily in the Aramaic script, which, like the Aramaic language, was widely used in both the Neo-Babylonian and Persian Empires. Like the Hebrew writing system, the Aramaic had arisen as an early branch of Phoenician, so that it provided the same consonantal inventory as the old Hebrew script, and its adoption for writing Hebrew was straightforward. It was out of the Aramaic script tradition that the standard biblical book hand, known as the “square script” or simply the Jewish script, eventually developed. This writing system is shown in Table 10.1.

### 2.2 Vowel representation

Whereas Phoenician orthography was purely consonantal, the earliest Hebrew inscriptions exhibit a rudimentary form of vowel representation, with certain letter signs (*wāw*, *yôd* and *hē*) being assigned a secondary use as vowel markers. At first this use of *matres lectionis* (“mothers of reading”) was confined to final long vowels, with *wāw* representing final *ū*, *yôd* representing final *ī*, and *hē* representing final *ā*, *ē* or *ō*. Eventually, internal vowel letters began to be indicated on a sporadic basis, with *wāw* representing internal *ō* (contracted from *\*aw*) or *ū*, and *yôd* representing internal *ē* (contracted from *\*ay*) or *ī*. During the second half of the first millennium BC, *wāw* gradually replaced *hē* as the marker of final *ō*.

By the last century before the Common Era, the tendency to represent vowels *plene* (i.e., “fully” or with *matres*) reached its most elaborate development. Nevertheless, this development, though observable in the Samaritan Pentateuch and numerous biblical manuscripts

**Table 10.1 The Hebrew alphabet**

	Letter name	Transcription	Phonetic value
א	<i>ʾālep</i>	ʾ	[ʔ]
ב	<i>bêt</i>	<i>b</i>	[b], [v]
ג	<i>gîmel</i>	<i>g</i>	[g], [ɣ] or [ʁ]
ד	<i>dālet</i>	<i>d</i>	[d], [ð]
ה	<i>hēʾ</i>	<i>h</i>	[h]
ו	<i>wāw</i>	<i>w</i>	[w]
ז	<i>zāyin</i>	<i>z</i>	[z]
ח	<i>hêt</i>	<i>ḥ</i>	*[ħ], [H]
ט	<i>ṭêt</i>	<i>ṭ</i>	*[tʰ], [t]
י	<i>yôd</i>	<i>y</i>	[y]
כ	<i>kap</i>	<i>k</i>	[k], [x] or [χ]
ל	<i>lāmed</i>	<i>l</i>	[l]
מ	<i>mēm</i>	<i>m</i>	[m]
נ	<i>nûn</i>	<i>n</i>	[n]
ס	<i>sāmek</i>	<i>s</i>	[s]
ע	<i>ʾayin</i>	ʿ	[ʕ]
פ	<i>pēh</i>	<i>p</i>	[p], [f]
צ	<i>ṣādēh</i>	<i>ṣ</i>	*[sʰ], [tʃ] or [tʃʰ]
ק	<i>qôp</i>	<i>q</i>	*[kʰ], [k]
ר	<i>rēš</i>	<i>r</i>	[r]
ש	<i>śîn</i>	<i>ś</i>	*[ʃ], [s]
ת	<i>śîn</i>	<i>š</i>	[ʃ]
ת	<i>tāw</i>	<i>t</i>	[t], [θ]

from Qumran, is not reflected in the Hebrew Bible as transmitted in rabbinic tradition. In their efforts to standardize the sacred text, the rabbis elected a conservative tradition, giving authority to older manuscripts with “defective” spelling, so that the biblical books were preserved in an archaic orthography. In this way, rabbinic authority gave rise to the manuscript tradition that, in essential form, has survived into modern times. Although this tradition can safely be regarded as a faithful representation of the Hebrew language of the first millennium BC, the linguistic information it provides is accurate and complete only within the limits of the orthography of the Hebrew-Aramaic consonantal script.

## 2.3 Systems of biblical vowel notation

Because of its many ambiguities with regard to pronunciation, the biblical manuscript tradition was reinforced from an early date by an oral tradition that provided a guide to vocalization for use in liturgy and study. As Hebrew continued to develop regionally, the pronunciation traditions in the eastern (Babylonian) and western (Palestinian) Jewish communities began to diverge. By the second half of the first millennium AD these oral traditions had given rise to distinctive systems of “pointing” (*nîqûd*), graphic conventions for representing pronunciation fully by placing diacriticals above or below the text. The Babylonian tradition was fixed by a superlinear system developed in the sixth century AD and refined in the eighth–ninth centuries. The original Palestinian system, which was developed in the

**Table 10.2 The Tiberian representation of the principal Hebrew vowels**

Masoretic diacritical	Probable phonetic realization	Tiberian representation without <i>mater</i>	Tiberian representation with <i>mater</i>	Tiberian representation with final <i>mater</i>
<i>hîreq</i>	[i]	ה, <i>bi</i> or <i>bî</i>	הֿ, <i>bî</i> <sup>̇</sup>	
<i>šêrê</i>	[e]	שׁ, <i>bē</i>	שֿׁ, <i>bê</i>	הּ, <i>bēh</i>
<i>səgōl</i>	[ɛ]	ס, <i>be</i>	סֿ, <i>bê</i>	הּ, <i>beh</i>
<i>pataḥ</i>	[a]	פ, <i>ba</i>		
<i>qāmeš</i>	[ɔ]	ק, <i>bā</i> or <i>bo</i>		הּ, <i>bā</i>
<i>ḥōlem</i>	[o]	ח, <i>bō</i>	חֿ, <i>bô</i>	הּ, <i>bōh</i>
<i>qibbûš</i>	[u]	ק, <i>bu</i> or <i>bū</i>		
<i>šûreq</i>			שֿׁ, <i>bû</i>	

sixth–eighth centuries, was also superlinear. The extant documents using both of these systems provide important information about the development of Hebrew in late antiquity, although only a few manuscripts with Palestinian vocalization have survived. The older Palestinian system was superseded by a primarily infralinear and especially rigorous system developed in Tiberias, which enjoyed its most creative period between the late eighth and early tenth centuries AD. The Tiberian system of vowel notation is the only one that survives in active use, and it is regarded as authoritative in Jewish tradition, though a superlinear system developed for the Samaritan Pentateuch has a similar role in the Samaritan community. The Tiberian pointing is reinforced in its mission of safeguarding the integrity of the text by the Masora, a body of detailed annotations produced by scholars known as Masoretes (*ba’ālê hammāsōret*, literally, “masters of the tradition”); the text of the Hebrew Bible, when equipped with this apparatus, is called the Masoretic Text.

## 2.4 Tiberian vowel signs and modern transliteration

The representation in the Masoretic Text of the vowels and their morphophonemic varieties (see §3.2.1) was accomplished by the introduction of the Tiberian diacriticals into a text that, as explained in §2.2 above, already contained a minimal indication of vowels in the form of the *matres lectionis*. The present system of vowel representation is thus composite, and it is necessary in transliteration to indicate, as far as possible, both the *matres* and the diacritical marks of the Masoretes. It is also desirable to indicate vowel quantity because of the important light it sheds on the character of the ancient language and its historical, pre-Tiberian development. Information about vowel quantity cannot be deduced on the sole basis of the Tiberian vowel signs, however, since their purpose was to indicate quality rather than quantity. Nor are the *matres* a fully reliable guide. There was, to be sure, a tendency in the text to mark the ancient long vowels with *matres*, but in the conservative orthography of the Bible this was not carried through consistently or systematically. When vowels are marked for length in transliteration, therefore, they represent an interpretation made on the basis of an analysis of word structure and stress in light of modern research into the pre-Tiberian history of the language.

Table 10.2 lists the Tiberian spellings of the principal varieties of the seven vowels identified below in §3.2.1 together with their corresponding transliterations (for purposes of illustration the vowels are attached to the consonant *b*).

When using this type of transliteration it is important to keep its limitations and shortcomings in mind. Though it has the merit of highlighting information about the length of vowels, it can be misleading in this regard, since it gives the impression, for example, that *šērê*, transliterated <ē>, is the lengthened form of *səgōl*, <e>, when in fact *šērê* is an altogether different, higher vowel than *səgōl* ([e] ~ [ɛ]). The chief purpose of the transliteration system is to permit the reader to reconstruct the Tiberian spelling, but here, too, there are a few imperfections and unavoidable ambiguities. For example, both *šērê-yôd* (רֶ.) and *səgōl-yôd* (רֶ.) are transliterated <ê> (in some systems the latter is rendered <e(y)> or <ê> to avoid the ambiguity), and final *šērê-hē'* (רֶ.) is transliterated <ēh> to distinguish it from *šērê-yôd* (רֶ.) even though the *hē'* is a *mater* (see §2.2), that is, non-consonantal (in some systems *šērê-hē'* is rendered <ê> like *šērê-yôd* and *səgōl-yôd*, eliminating the misrepresentation but compounding the ambiguity).

### 3. PHONOLOGY

#### 3.1 Consonants

Table 10.3 illustrates the consonantal phonemes of Hebrew. As shown, the consonantal system consists of seventeen obstruents, including nine oral stops and eight fricatives; and six sonorants, including four approximants (glides and liquids) and two nasals.

##### 3.1.1 Obstruents

The set of stops comprises, in addition to the glottal stop /ʔ/, a symmetrical group of six consonants produced in two manners of phonation (voiced and voiceless), at three points of articulation (bilabial, alveolar and velar). This set is supplemented by two (dental and velar) ejective stops, the so-called “emphatics.” In Tiberian Hebrew the six non-emphatic stop phonemes, /b/, /p/, /d/, /t/, /g/ and /k/, possess a complete set of conditioned spirantized allophones, [v], [f], [ð], [θ], [ɣ] or [ʁ], and [x] or [χ], conventionally transliterated as *b̄*, *p̄*, *d̄*, *ḡ* and *k̄*, the development of which is discussed below (see §3.3).

The fricative group includes three voiceless, nonemphatic sibilants, /s/, /š/, and the sound conventionally transcribed as *ś*. Though the three were originally distinct, they were later reduced to two when *ś* lost its primitive character as a lateral (i.e., /l/), and merged with the other voiceless alveolar sibilant, /s/ (confusion of /s/ and *ś* is already present in Late Biblical Hebrew and becomes increasingly common at Qumran and in Rabbinic Hebrew). The sibilant inventory is completed by two other fricatives, voiced /z/ and emphatic /s'/ (conventionally written *ṣ*). All of these are alveolars except the post- or palato-alveolar /š/.

Biblical Hebrew has lost all three Proto-Semitic interdentals, \*ð, \*θ and \*θ̥ as well as the emphatic lateral \*ś or \*ð̥ and the velar or uvular fricatives \*ḡ and \*ḥ (see §3.6.1), though the interdentals \*ð and \*θ ([ð] and [θ]) and the velars \*ḡ and \*ḥ ([ɣ] and [x]) have been “revived” in the form of the spirantized allophones of /d/, /t/, /g/ and /k/, as noted above.

The original pronunciation of the three Hebrew ejectives or emphatics, *t̥*, *s̥* and *q̥*, is unknown. Although the nature of the emphatics in Ethiopic and Arabic is itself debated, it is usually argued on the basis of these cognate languages that the Hebrew emphatics were originally glottalic, as in Ethiopic and (probably) Old South Arabic – thus [tʔ], [sʔ] and [kʔ], the presumed Proto-Semitic situation – but later became pharyngealized ([tʕ], [sʕ] and [kʕ]) among Jews living in Arabic-speaking communities, and simplified to [t], [tʃ] or [tʰ]

**Table 10.3 The consonantal phonemes of Hebrew**

Manner of articulation	Place of articulation						
	Bilabial	Dental/Alveolar	Palato-alveolar	Palatal	Velar	Pharyngeal	Glottal
Stop							
<i>Voiceless</i>	p (פ)	t (ת)			k (כ)		ʔ (/ʔ/, א)
<i>Voiced</i>	b (ב)	d (ד)			g (ג)		
<i>Emphatic</i>		ṭ (/tʰ/, ט)			q (/kʰ/, ק)		
Fricative							
<i>Voiceless</i>		s (ס)	š (ש)			ħ (/ħ/, ח)	h (ה)
<i>Voiced</i>		z (ז)				ʕ (/ʕ/, ע)	
<i>Emphatic</i>		ṣ (/sʰ/, צ)					
<i>Lateral</i>		ṣ (/ʃ/, ש)					
Approximant							
<i>Glide</i>	w (ו)			y (י)			
<i>Rolled</i>		r (ר)					
<i>Lateral</i>		l (ל)					
Nasal	m (מ)	n (נ)					

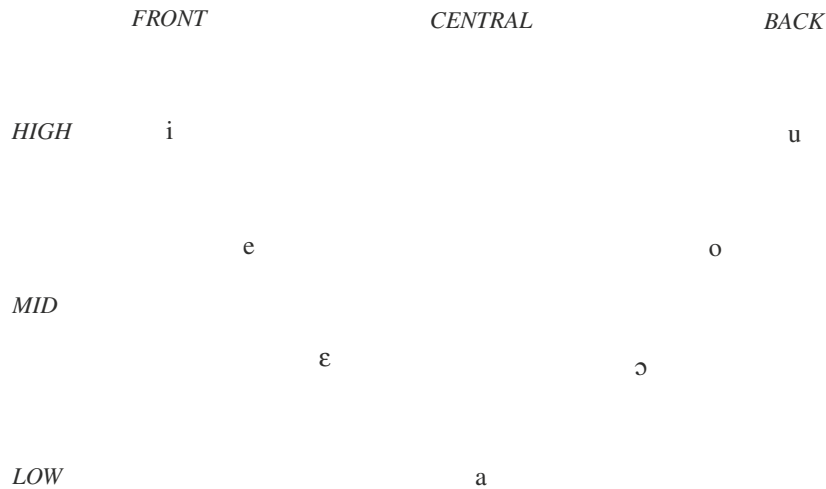
and [k] among European Jews. As shown by Tiberian tradition and confirmed by earlier Greek transcriptions, the emphatic stops, *ṭ* and *q*, did not share the secondary spirantized realization acquired by the six nonemphatic stops noted above.

Hebrew distinguishes four “guttural” consonants: two pharyngeals, one voiced /ʕ/ (conventionally transcribed as *ʕ*) and one voiceless /ħ/ (*ħ*), both of which are composite in origin (see §3.6.1), and two voiceless glottals, one stop /ʔ/ (*ʔ*) and one fricative /h/. As the language evolved, there was a tendency for these consonants to weaken and/or coalesce, a development with important secondary phonological consequences (see §3.3). While the glottals participated in this general pattern of weakening, they underwent, in addition, important changes of their own. In particular, the glottal stop, /ʔ/, was lost in syllable-final positions, a phenomenon that began very early and seems to have proceeded in stages (see §3.6.1) and in which the other glottal, /h/, may have participated in part.

### 3.1.2 Sonorants

Hebrew has two nasals, bilabial /m/ and alveolar /n/, both voiced. The tendency in Rabbinic Hebrew for these two consonants to alternate when final (especially *\*-m > -n*) is already in evidence in Septuagint transliterations and Qumran manuscripts but lacking in Biblical Hebrew itself, unless *\*šāllûm* is intended by the name *šāllûn* in Nehemiah 3:15 (for the related question of the replacement of the plural ending *-îm* with *-în*, see §4.2.2). When immediately followed by a non-guttural consonant, /n/ undergoes regressive assimilation (*\*nC > CC*), unless it follows the preposition *lā-* or is the third consonant in the stem: for example, *zākāntā*, “you have grown old” (1 Samuel 8:5).

Hebrew has four approximants, all voiced. Two of these, the bilabial and palatal semivowels /w/ and /y/, are glides. The other two are liquids; they include /r/, a rolled consonant, probably realized as either an alveolar [r] or uvular [ʀ] trill, and /l/, a lateral alveolar liquid.



**Figure 10.1** The seven full vowels of Tiberian Hebrew

## 3.2 Vowels

### 3.2.1 The quality of the Tiberian vowels

As explained below (see §3.2.2), ancient Hebrew in its early development probably preserved the basic triad of Proto-Semitic vowels, *\*i*, *\*a* and *\*u*, each of which could be long or short, and two “diphthongs” or vowel-glide sequences, *\*ay* and *\*aw*. The Tiberian system by which Biblical Hebrew is represented is much more complex, however, reflecting the medieval pronunciation that had evolved over the centuries from numerous phonological changes. There are Masoretic diacriticals for seven full vowels (*hîreq* [i], *šērê* [e], *səgōl* [ɛ], *pataḥ* [a], *qāmeṣ* [ɔ], *ḥōlem* [o] and *qibbûṣ/šûreq* [u]), and when vocal *šəwā* [ə] and the three other ultrashort or reduced vowels (the *ḥāṭēp* vowels) are added, the number of vowels rises to eleven. The approximate phonetic realization of the seven full vowels is illustrated in Figure 10.1, which presents Tiberian Hebrew as possessing a complete inventory of primary vowels.

### 3.2.2 The origin of the Tiberian vowels

As noted above (§3.2.1), Hebrew, in the early stages of its development, probably preserved the Proto-Semitic system of three vocalic phonemes, high front *\*i* and back *\*u* and low central *\*a*, which could occur either long or short, and two “diphthongs” or vowel-glide sequences, *\*ay* and *\*aw* (see Ch. 6, §§3.2.2 and 3.2.3). Though the phonological changes by which these sounds gave rise to the Tiberian system described above are numerous and often complex, constrained by the rules of syllabification and stress (see §§3.4 and 3.5 below), it is possible to describe the Masoretic vowels and diphthongs in relation to their ancient antecedents by taking historical and structural considerations into account.

#### 3.2.2.1 The development of the originally long vowels

The Proto-Semitic long vowels, *\*ī*, *\*ū*, and *\*ā*, undergo no special development in Hebrew. Proto-Semitic *\*ā* is realized as [o], but this is not an inner-Hebrew development but the result of a sound change (*\*ā* → *ō*) that Hebrew inherited from Proto-Canaanite (see §3.6.2). Proto-Semitic *\*ī* and *\*ū* remain unchanged, and they are most often represented



orthographically in the the Masoretic Text with *plene* spellings,  $\hat{i}$  (יִ) and  $\hat{u}$  (וִ), though this is by no means consistent (see §§2.2 and 2.4). In terms of their phonological behavior, the Hebrew vowels derived from the Proto-Semitic long vowels may be described as unchangeably long to distinguish them from reducible long vowels derived from originally short vowels (§3.2.2.2); they are not subject to reduction to  $\text{šewā}$  (ə), regardless of position.

### 3.2.2.2 The development of the originally short vowels

The development of the Hebrew short vowels is much more complex. Because of changes that occurred during the evolution of the language, an originally short vowel may be realized as long, short (not necessarily the same short vowel as the original) or reduced ( $\text{šewā}$  or one of the  $\text{ḥāṭēp}$  vowels). The possible morphophonemic variants of each of the short vowels are shown in (1):

(1) *Original*

<i>short vowel</i>	<i>Lengthened</i>	<i>Short</i>	<i>Reduced</i>
* $\hat{i}$	$\bar{e}$	i, a, e	ə, ä, ě
* $\hat{u}$	$\bar{o}$	u, o	ə, ō
* $\hat{a}$	$\bar{a}$	a, i, e	ə, ä, ě

The potential for an originally short vowel to lengthen or reduce is constrained by the type and position of the syllable in which it appears. To lengthen, it must be in an open syllable (CV) or an accented closed syllable (CV'C). To reduce, it must be in an unaccented open syllable (CV), since a closed syllable (CVC), like an open syllable containing an originally long vowel (CV:), is irreducible (for syllabification, see §3.4). In general, therefore, an originally short vowel tends to lengthen in a tonic syllable or in an open pretonic syllable, it tends to remain short in a closed unaccented syllable (though its quality may change), and it tends to reduce in an open propretonic syllable. In practice, however, the operation of these very general rules differs for nouns (including adjectives and verbal nouns) and finite verbs with pronominal suffixes, on the one hand, and finite verbs without pronominal suffixes, on the other. The rule of thumb for nouns and finite verbs with pronominal suffixes is that an originally short vowel reduces in a propretonic syllable if possible – that is, if a propretonic syllable is present and its vowel is reducible – while it lengthens in a pretonic syllable. The rule of thumb for finite verbs without pronominal suffixes is that an originally short vowel reduces in a pretonic syllable if possible, while it lengthens in a propretonic syllable. These rules, too, are generalizations, however, and a clearer picture emerges when the situation is reviewed for vowels in each of the three common syllabic stress positions: tonic, pretonic and propretonic.

Originally short vowels in *tonic* syllables are, in most circumstances, lengthened in both nouns and verbs. That is, the high vowels \* $\hat{i}$  and \* $\hat{u}$  are lowered to  $\bar{e}$  ([e]) and  $\bar{o}$  ([o]), and the low vowel \* $\hat{a}$  is backed to  $\bar{a}$  ([ɔ]). With certain exceptions, this pattern holds for tonic syllables of all kinds in nouns and finite verbs with pronominal suffixes when the short vowel in question is \* $\hat{i}$  or \* $\hat{u}$ . When the vowel is \* $\hat{a}$ , the pattern holds for open and singly closed (word-final) syllables but not for originally doubly closed syllables (-C<sub>1</sub>C<sub>1</sub># or -C<sub>1</sub>C<sub>2</sub>#). Since lengthening took place prior to the simplification of final doubled consonants, the vowel \* $\hat{a}$  before a final, originally doubled consonant (-CC#) remains: thus, \* $\text{amm}$  → 'am "people" (note, however, that \* $\hat{i}$  and \* $\hat{u}$  both lengthen before -CC#: \* $\text{libb}$  →  $\text{lēb}$  "heart"; \* $\text{uzz}$  → 'ōz "strength"). Also, in an originally word-final doubly closed syllable (see §3.4), when the tone vowel has become penultimate because of the insertion of an anaptyctic vowel to resolve the consonant cluster (-C<sub>1</sub>C<sub>2</sub># → -C<sub>1</sub>VC<sub>2</sub>#), an accented

short *\*a* is not lengthened (except in pause; see §3.5), though it retains its stress and is raised to *e* ([a] → [ɛ]). This pertains especially to nouns of the type *\*CaCC* – thus, *\*málk* → *mélek*, “king” (pausal *málek*). Note that with the high vowels there is no exception here (i.e., they usually lengthen in this situation), but sometimes, not consistently, before a word-final consonant cluster *\*i* ([i] → *e* ([ɛ]) instead of *ē* ([e]), especially in some nouns of the type *\*qitl*: for example, *\*śídq* → *śédeq*, “righteousness,” in contrast to *\*śípr* → *śéper*, “book.”

Similarly, the lengthening of *\*a* does not take place in the tonic syllable as a result of the triphthongization of some diphthongs, as in *\*báyṭ* → *báyit* (contrast *\*máwt* → *máwet*), or the formation of the dual ending *\*-áym* → *-áyim*. One other important exception where stressed *\*a* is not lengthened is the verbal suffix of the first-person singular: *-ání* “me” (but, again, pausal *-ání*).

The pattern of lengthening of originally short vowels in tonic syllables also holds true for finite verbs without pronominal suffixes, but only for *\*i* and *\*u* – thus, *\*yittín* → *yittén* “he gives”; *\*tiktúb* → *tiktób* “she writes.” Originally short *\*a* remains short in these circumstances – *yísmá* “he hears.” Again, however, the situation is different when an originally word-final doubly closed syllable is involved. In these cases, the original short vowel is retained without lengthening in the tonic syllable after anaptyxis (*\*yírb* → *yíreb* “may he become numerous”), though *\*a* ([a]) is raised to *e* ([ɛ]) (*\*yárb* → *yéreb* “may he make numerous”).

Finally, mention should be made here of the vowel shift described by F. W. M. Philippi, according to which *\*i* becomes *a* in originally closed accented syllables (*\*iCC#* → *áCC#*) – in short, “Philíppi → Philáppi.” Though this “law” seems to explain many Hebrew forms – such as (*\*bint* →) *\*bitt* → *\*batt* (→ *bat*) “daughter”; (*\*āmidt* →) *\*ōmidt* → *\*ōmádt* (→ *\*ōmédet*) “standing” (fem. sg. active participle); *\*zāqintī* → *zāqántī* “I am old” – its application admits of a very large number of exceptions, and it is inoperative in some witnesses (e.g., the Hexaplaric) to the developing Hebrew tradition.

Originally short vowels in open *pretonic* syllables are, in general, lengthened in nouns and reduced in unsuffixed verbs. More specifically, in nouns and finite verbs with pronominal suffixes, *\*i* and *\*u* are lengthened pretonically if there is a reducible propretonic (*\*šakinīm* → *šākēnīm* “neighbors”). If the propretonic is lacking or irreducible, however, the behavior of pretonic *\*i* and *\*u* depends on the quality of the tonic vowel. If the tonic vowel is also high, pretonic *\*i* and *\*u* reduce to *šəwā*: for example, *\*gibūl* → *gəbūl* “boundary”; *\*šōmirīm* → *šəmərim* “guards”; *\*yīšmurīhū* → *yīšməreḥū* “he guards him.” If the tonic vowel is not high, pretonic *\*i* and *\*u* lengthen (*\*i* → *ē*, *\*u* → *ō*): thus, *\*libáb* → *lēbáb* “heart”; *\*maššibā* → *maššēbā* “pillar.” Pretonic *\*a* always lengthens (*\*a* → *ā*) in nouns and suffixed verbs, whether the propretonic is reducible (*\*dabarīm* → *dəbārīm* “words”) or not (*\*kawkabīm* → *kōkabīm* → *kōkābīm* “stars”).

In contrast to the situation with nouns and suffixed verbs, the originally short vowels are usually reduced pretonically in finite verbs without pronominal suffixes – thus, for example, *\*yignubū* → *yignəbū* “they steal”; *\*yittinū* → *yittənū* “they give”; *\*yikbadū* → *yikbədū* “they are heavy.” An important exception is when the pretonic is the first syllable in a word; in such a case the vowel is lengthened: thus, *\*himītū* → *hēmīṭū* “they killed.”

Originally short vowels in *propretonic* syllables are, when possible, reduced in nouns and lengthened in unsuffixed verbs. The specific rule for nouns and finite verbs with pronominal suffixes is that an originally short vowel reduces propretonically if it is reducible, that is, if it appears in an originally open syllable. If the propretonic is irreducible, however, the pretonic reduces according to the rules (and exceptions) given above. In finite verbs without pronominal suffixes, an originally short vowel reduces when

possible in a pretonic syllable, as also explained above, and if this happens, *\*i*, *\*u*, or *\*a* in the prepretonic syllable lengthens: for example, *\*napalā* → *nāpələ* “she fell.” If, however, the pretonic is not reducible (that is, if it is closed or contains an originally long vowel), the prepretonic vowel reduces: *\*yudabbir* → *yədabbēr* “he speaks.”

To this point the discussion of the originally short vowels has been concerned primarily with their behavior in open syllables or closed accented syllables, both of which permit the lengthening or reduction of the vowel. In closed unaccented syllables, however, *\*i*, *\*u*, and *\*a* remained short despite occasional changes of vowel quality. This is true whether they appear in originally closed pretonic or prepretonic syllables, and it applies to both nouns and verbs. Examples of the former (pretonic) include the nouns *\*šibtuh* → *šibtō* “his tribe”; *\*kulluh* → *kullō* “all of it” (cf. *\*ḥudšah* → *ḥodšāh* “her new moon”); and *\*gapnī* → *gapnī* “my vine”; and the verbs *\*yimša* → *yimšā* “he finds”; *\*yuggad* → *yuggad* “it is reported”; and *\*yašbīt* → *yašbīt* “he causes to cease.” Examples of the latter (prepretonic) include the noun *\*milḥamāt* → *milḥāmōt* “wars” and the verb *\*yišmurū* → *yišmərū* “they watch.”

While the *quantity* of an originally short vowel remains the same in a closed unaccented syllable, however, its *quality* may be altered. Although a number of situations in which this occurs could be listed, the attenuation of *\*a* to *i* in the sequence *\*CaC<sub>1</sub>C<sub>2</sub>āC* → *CiC<sub>1</sub>C<sub>2</sub>āC* (where *C<sub>1</sub>* is not a guttural) is especially noteworthy. This phenomenon, commonly known as “*qatqat* → *qitqat* dissimilation,” operates in *m*-prefix nouns, such as *\*madbār* → *midbār* “wilderness” and *\*malḥāmā* → *milḥāmā* “battle” (see §4.2.5.4), and especially (with short *a* in the second syllable) in construct forms, such as *\*šadqat* → *šidqat* “righteousness (of)” and *\*mazbah* → *mizbah* “altar (of).” The historical distribution of *m*-prefix nouns with the form *miqṭal* suggests that *qatqat* → *qitqat* dissimilation took place at a relatively late date, since forms like *midbār* are found only in Tiberian Hebrew, in contrast to Hexaplaric and Babylonian *madbār*. On the other hand, verbal forms like *yiqṭal* (<*\*yaqṭal*) – for example, *\*yalmad* → *yilmad* “he learns” – and *niqṭal* (<*\*naqṭal*), the *Nip’al* perfect, developed much earlier, as shown not only by their attestation in all traditions of Hebrew vocalization but also by the presence of *\*yiqṭal* in cognate languages like Aramaic and Ugaritic. This suggests that the various forms that are often explained by appeal to *qatqat* → *qitqat* dissimilation are not in fact the result of a single phenomenon (for *\*yiqṭal* and the so-called Barth–Ginsberg Law, see §3.6.2).

### 3.2.2.3 The development of “diphthongs”

As noted above (see §3.2.2), it is customary to state that Proto-Semitic possessed two diphthongs, *\*aw* and *\*ay*, both of which were preserved, with modifications, in Hebrew. But since Proto-Semitic did not permit sequences of two (or more) vowels within a syllable (see Ch. 6, §3.2.3), the glides or semivowels, *\*w* and *\*y*, must be interpreted as consonants, and the two sequences (both [a + glide]) cannot be classified as true diphthongs. This sheds light on their realization in Tiberian Hebrew. When either of the “diphthongs” occurs in an accented syllable, *CáwC* or *CáyC*, it is “triphthongized,” or disyllabically resolved, before a final consonant by the insertion of an anaptyctic vowel, *e* in the case of *áw* (*CáwC* → *CáweC*) and *i* in the case of *áy* (*CáyC* → *CáyiC*) – thus, *\*máwt* → *máwet* “death,” and *\*báyṭ* → *báyit* “house.” In other words, the syllable containing the diphthong behaves like other syllables with final consonant clusters (see §3.4). Note, however, that when stressed *\*áy* occurs immediately before a syllable with the form *Cā*, it dissimilates to [ɛ], spelled *səgōl-yōd* (ֶ.) in the Masoretic Text – thus *\*áyCā* → *-éCā*, as in *\*ḥuqqáykā* → *\*ḥuqqé’kā* “your statutes.” In an unstressed syllable either diphthong is “monophthongized” or contracted: *\*aw* → *ō* or *\*ay* → *ê* – thus, *\*mawtó* → *môtó* “his death,” and *\*baytó* → *bêtó* “his house.” The vowels thus contracted merged phonetically with other long *ō*- and *ê*-vowels, regardless of their

historical origin, including  $\bar{o} < *ā$  (see §3.6.2) and  $\bar{o} < *u$  and  $\bar{e} < *i$  (see §3.2.2.2 and [1]; for the behavior of diphthongs in the dialects of Iron Age Hebrew, see §3.6.2).

### 3.3 Allophonic and morphophonemic variants

#### 3.3.1 Fricative allophones

At some point in the development of Tiberian Hebrew the six nonemphatic stops, /b/, /p/, /d/, /t/, /g/ and /k/, acquired a second, continuant realization, giving rise to six fricative allophones, [v], [f], [ð], [θ], [ɣ] or [ʁ], and [x] or [χ], conventionally transliterated as *b*, *p*, *d*, *t*, *g* and *k*. These forms arose as subphonemic or phonetic variants, originally restricted to nongeminated consonants in postvocalic positions. This development, which was shared by and probably influenced by Aramaic, is widely assumed to have taken place in the second half of the first century BC, but its precise chronology is unknown. The fricative allophones are fully represented in the Tiberian Masora, and there is evidence for their presence in the time of Rabbinic Hebrew, but their existence before the Common Era is not unambiguously documented.

#### 3.3.2 Gutturals

The so-called gutturals (pharyngeals and glottals or laryngeals) underwent a pattern of progressive but dialectically heterogeneous weakening that resulted in a special set of rules in Tiberian grammar governing these consonants, /ʕ/ (ʿ), /ħ/ (ḥ), /ʔ/ (ʾ), /h/, and the vowels in their environment. Though these rules are extensive and complex, three basic stipulations may be mentioned here. First, a guttural cannot be doubled (a rule that also applies to the liquid /r/), so that a doubled guttural was simplified ( $*GG \rightarrow G$ ), either with lengthening of the vowel in the newly opened preceding syllable (compensatory lengthening) – as in  $*yi$  “akil” →  $yēʾākēl$  “it is eaten”;  $*barrik$  →  $bārēk$  “to bless” – or without this lengthening (so-called virtual doubling) –  $*biʿir$  →  $biʿēr$  “he burned”;  $*yuraḥḥim$  →  $yəraḥēm$  “he has compassion.”

Second, a guttural cannot be followed by a simple  $šəwā$  ([ə]), requiring instead a “compound  $šəwā$ ,” a reduced or ultrashort variant of one of the short vowels (the *ḥātēp* vowels,  $\bar{e}$ ,  $\bar{o}$ , and  $\bar{a}$ ), as an auxiliary – thus,  $ʾēlōhīm$  “god,”  $ʾohōlī$  “my tent” and  $ḥālōm$  “dream.”

Third, when final, a guttural, other than /ʔ/ (ʾ), requires anaptyxis of *a* (“furtive *pataḥ*”) following a vowel other than *a* or  $\bar{a}$ : for example,  $*rūḥ$  →  $rūaḥ$  “wind”;  $*hišmīʿ$  →  $hišmīaʿ$  “he caused to hear.”

While it is difficult to date this pattern of weakening, and its progress is unlikely to have been uniform, it seems to have been well advanced by the time of the Samaritan Pentateuch and the Qumran literature, since occasional confusion of gutturals is found in both, and Qumran orthography exhibits conspicuous irregularities when the gutturals are involved, especially in nonformal manuscripts (i.e., those in which the scribes were not careful to reproduce the spelling practices of the biblical literature). On the other hand, it is clear that this development was primarily a matter of the weakening and coalescence of the gutturals rather than their disappearance, as shown by the mixed evidence of the Hexaplaric transcriptions. That the gutturals, in some configuration, were still a feature of Jewish speech c. AD 400 is shown by Jerome’s remark that the Jews ridiculed the Christians for their inability to pronounce them. It seems clear, then, that the gutturals were preserved in some communities and lost in others, most probably where Greek influence was strongest. Thus the Talmud (*Megillah* 24b) refers to a lack of distinction (coalescence) among the gutturals in the speech of certain Galilean villages, but not others (on the quiescence of /ʔ/, which,

though it played a part in the general phenomenon of guttural weakening, was of much earlier origin, see §3.6.1).

### 3.3.3 Vowel variation

For the development of vowel morphophonemic variation in Tiberian Hebrew, see §§3.2.2.1–3.2.2.2.

## 3.4 Syllable structure and phonotactic constraints

A Hebrew syllable must begin with a consonant. There is a single but important exception to this rule in Tiberian grammar, according to which the conjunction *wə-* “and” becomes *û-* before a syllable beginning with a consonant (not *y-*) plus *šəwā* – as in *ûd(ə)bārîm* “and words” – or a syllable beginning with a labial – such as *ûmélek* “and a king” (the Babylonian vocalization tradition also reflects the former situation, but not the latter, preserving the equivalent of *wə-* before a labial followed by a full vowel).

A syllable may contain only one vowel sound. The Hebrew diphthongs do not constitute an exception to this rule, since, as noted above (§3.2.2.3), they are not true diphthongs but vowel–glide combinations, and since, in any case, they are always either monophthongized to single vowel sounds – as in *\*baytuh* → *bêtô* “his house” – or triphthongized to vowel–glide–vowel combinations, thus forming parts of two distinct syllables – *\*bayt* → *báyit* “house.”

A syllable may be open or closed. A syllable ending with a vowel (long, short, or reduced) is described as open, while a syllable ending with a consonant is described as closed. Occasionally a syllable ends in two consonants, and in this case it is called doubly closed: for example, *kātābt* “you (fem. sg.) wrote.” Doubly closed syllables occur only at the ends of words, having arisen when a final vowel was lost (*kātābt* < *\*katabtī*). Such consonantal clusters were not permitted by the phonotactic rules of Proto-Semitic (see Ch. 6, §3.2.3), and Hebrew grammar exhibits a tendency to avoid them. When they do occur, the preceding vowel may be short (*wayyîšb* “and he captured”; *wayyašq* “and he watered”) or, with [i] lowered to [e] under the stress, long (*wayyēbk* “and he wept”; *wayyēšt* “and he drank”); but the medieval grammarians disagreed whether the final *šəwā* in such words was silent or vocal, and the Masoretes most often eliminated the problem by inserting an anaptyctic vowel, usually *səgōl* (*\*wayyipn* → *wayyîpen* “and he turned”; *\*yibn* → *yîben* “let him build”), but *pataḥ* before or after gutturals (*\*wayyîhr* → *wayyîḥar* “and he was angry”) and *ḥîreq* after *y* (*\*ayn* → *‘ayin* “eye”). “Segholation,” as this phenomenon is sometimes called, is most characteristic of nouns of the common type *\*CVCC* (“segholates”; see §4.2.5.2) – *\*arš* → *‘ereš* “earth”; *\*izr* → *‘ezer* “help”; *\*buqr* → *bôqer* “morning”; and with gutturals, *\*naḥl* → *nāḥal* “wadi” and so forth. Though anaptyxis in segholates is reflected in both the Babylonian and Tiberian traditions, its absence in the Hexaplaric materials suggests that it was a late phenomenon.

A syllable may be accented or unaccented (see §3.5). An accented syllable may be open or closed and contain a long or short vowel (CV(:), CV(:)C), though an accented syllable may not contain a reduced vowel. With rare exceptions, an unaccented syllable containing a long vowel will be open, while an unaccented syllable containing a short vowel is always closed (for the specific distribution of vowels in various types of syllables, which depends on rules of syllable formation deriving from the historical development of the language, see §3.2.2.2). In the Masoretic Text, when a closed unaccented syllable occurs in the middle of a word, the end of the syllable is indicated by the *šəwā* sign (·). The Masoretic diacritical for this syllable-dividing silent *šəwā* (*šəwā quiescens*) is the same as for the vocal *šəwā* (*šəwā*

*mobile*). In most cases, this will cause no difficulty for the reader since a consonant following an unaccented short vowel must be syllable final, so that it must close the syllable and the *šəwā* standing under it must be silent. Ambiguity arises only when the diacritical is *qāmeš*, which can indicate either long *ā* or short *o*. To resolve the ambiguity the Masoretes usually inserted the accent called *meteg*, a small perpendicular line (|), to the left of the *qāmeš* in an accented syllable, indicating that the *qāmeš* should be read as *ā* and thus that the following *šəwā* was vocal. In the absence of the *meteg*, the syllable should be read as unaccented and closed. Contrast *ʾakəlā* (אָכַלָּא) “she ate” and *ʾoklā* (אָכַלָּא) “food.”

According to the phonotactic rules of Tiberian grammar, only a consonant or a *full* vowel could constitute the coda of a syllable. In Masoretic sources, therefore, a consonant followed by a reduced vowel (simple or compound *šəwā*) was not regarded as an independent syllable. Thus, contrary to the guidelines given above, a word like *məlakim* “kings” would be analyzed as containing two syllables – *məlā-kim* – rather than three – *mə-lā-kim*. This rule explains, among other things, why the Masoretes chose the same sign (.) to represent both vocal and silent *šəwā*. Since most reduced vowels developed from vowels that were previously full, however, the medieval rule has the disadvantage of obscuring the historical development of the language, and it is not followed as a convention of syllabification by most modern grammarians.

Tiberian Hebrew does not tolerate two successive open syllables with the vowel /ə/. When such a sequence is produced in inflection or from a combination of morphemes, such as the prefixation of a preposition or suffixation of a pronoun to a noun, the phenomenon commonly called “the rule of *šəwā*” occurs. The sequence is simplified to a single closed syllable containing the vowel /i/ (\*CəCə → CiC) – thus, \**dəbərēʾ* → *dibrēʾ* “words (of),” \**bədəbārīm* → *bidbārīm* “with words,” and \**dəbərēhem* → *dibrēhem* “their words.”

### 3.5 Stress

In Hebrew the principal tone is usually, but not always, on the ultima – thus, *dābār* “word,” *dəbārīm* “words.” This situation is the result of a shift of stress to the ultima that took place in two phases early in the history (or prehistory) of the language. The original position of the stress in Proto-Hebrew is disputed. It seems clear, though, that it shifted to the ultima in two stages. The first shift affected all words except finite verbs without pronominal suffixes, and the second shift occurred in these verbs. This two-stage development gave rise to several distinctive features of Hebrew grammar, including some of the phonological features already noted, such as the tendency of vowels in open pretonic syllables to lengthen in nouns but reduce in unsuffixed verbs (see §3.2.2.2), as well as important morphological features to be noted, such as stem allomorphy for many noun-types (§4.2.6). Both of these shifts are reflected in the Hexaplaric, Babylonian, and Tiberian traditions of vocalization, and, in fact, they are likely to have been very early. In all likelihood the first, major shift closely followed the loss of final short vowels, which was shared by most of the Northwest Semitic languages, so that it was probably pre-Hebrew. Note in this regard that ultimate stress is also characteristic of Aramaic, as indicated, for example, by the Masoretic accentuation of Biblical Aramaic, and Phoenician, as can be inferred from vowel changes in the ultima that are likely to have been caused by stress.

Despite the preference for stress on the ultima, the penult receives the tone in a number of situations. For example, segholate nouns, as already noted, are characteristically paroxytonic – as in *ʿəmeq* “valley” – and the ultimate stress of the imperfects of certain types of verbs retreats in the production of jussives and the so-called “converted” or *wāw*-consecutive imperfects – *yiglēh* “he will uncover” ~ *yigel* “let him uncover” ~ *wayyigel*



“and he uncovered.” Also, a number of word-ending elements are for historical or structural reasons toneless. These include several verbal sufformatives (e.g., *ḥāšābtā* “you thought”), several nominal and verbal pronominal suffixes (*šāmərēnī* “protect me”), and the so-called locative *-h* (*báyṭāh* “to the house, homeward”; *’arṣāh* “to the earth”).

Both stress and consequent vowel quantity can be significantly affected by the so-called pause, a term for the increased stress placed on the tonic syllable of a word in the Hebrew Bible marked by one of the major accent diacritics, usually at the end of a verse or half-verse. In cases of the type just described, for example, where the stress of some imperfect verbs retreats from the ultima to the penult in the formation of converted imperfections, the tone returns under the pause to the ultima, which is lengthened accordingly – thus, *tāmū’t* “she will die” ~ *wattāmōt* “and she died” ~ *wattāmōt* “and she died.” Similarly, the tone may be restored under the pause to a vowel that lost its stress and was reduced to *šəwā* in the process of syllable formation according to one of the rules described above in §3.2.2.2, with the result that the original quality of the vowel returns and, if short, lengthens under the tone: thus, *\*yittinū* → *yittēnū* → *yittēnū* “they give.” In general, short vowels tend to lengthen under the pause (*\*qātāltā* → *qātāltā*), and often their newly lengthened status gives a clue to the pre-Masoretic quality of the underlying vowel, as in the case just cited of *yittēnū* (*ē* < *\*i*) and especially of segholate nouns, where, for example, an original *\*a/* realized as [ε] may be restored and lengthened under the pause (*\*gabr* → *géber* → *gáber* “man”).

Numerous minimal pairs can be cited to show that stress is phonemic in Hebrew: for example, *bā’ā* [bōʔa] “she came” ~ *bā’ā’* [bōʔa] “coming” (feminine singular active participle); *bānū* “they built” ~ *bānū* “in us.”

### 3.6 Diachronic phonological developments in relation to Proto-Northwest Semitic and Proto-Semitic

#### 3.6.1 Consonants

Of the twenty-three consonantal phonemes represented in Table 10.3, eighteen preserve Proto-Semitic consonants unaltered, and five – all fricatives – are the result of unconditioned mergers of two or three Proto-Semitic phonemes. These five include:

1. *z* (/z/) which arose from the merging of the voiced dental *\*z* (/z/) and the voiced interdental *\*ḏ* (or *\*ḏ*) (/ḏ/) – compare *zā’aq* (< *\*za’aq* < *\*za’aqā*) “he cried” to *zāhāb* (< *\*zahab* < *\*ḏahab*) “gold.”
2. *ḥ* (/ħ/) from the voiceless pharyngeal *\*ḥ* (/ħ/) and the voiceless velar *\*ḫ* (/x/) – compare *ḥēn* (< *\*ḥinn* < *\*ḥinn*) “favor” to *ḥārēš* (< *\*ḥariš* < *\*ḥariš*) “he is silent.”
3. *ʿ* (/ʕ/) from the voiced pharyngeal *\*ʕ* (/ʕ/) and the voiced velar *\*g* (/g/) – compare *ʿayin* (< *\*ʿayn* < *\*ʿayn*) “eye” to *ʿalmā* (< *\*ʿalmā* < *\*galmat*) “young woman.”
4. *š* (/ʃ/) from the emphatic dental *\*š* (/ʃ/) and the emphatic interdental *\*θ* (/θ/) and the emphatic lateral *\*ṣ* (or *\*ṣ*) (/ʃ/) – compare *šēdeq* (< *\*šīdq* < *\*šīdq*) “righteousness,” to *šēl* (< *\*šill* < *\*θill*) “shadow,” and *šēmer* (< *\*šāmr* < *\*ḏāmr*) “wool.”
5. *ś* (/ʃ/) from the voiceless palatal *\*ś* (/ʃ/) and the voiceless interdental *\*θ* (or *\*t*) (/θ/) – compare *śēm* (< *\*śim* < *\*śim*) “name” and *šōpēt* (< *\*šōpēt* < *\*θāpēt*) “judge.”

Proto-Semitic possessed a triad of dental/alveolar affricates: voiced *\*dʒ*, voiceless *\*tʃ* and ejective *\*tʃʼ*; see Chapter 6, §3.2.1.1. At an early date, these were deaffricated and merged with phonemes ancestral to the dental fricative triad in Hebrew – *\*dʒ* with *\*z*, *\*tʃ* with *\*s*, and *\*tʃʼ* with *\*ṣ* – so that it is not necessary to take them into account in a description of the Hebrew phonological system.

As noted in §3.3.2, \*ʾ (/ʔ/) participated in the general pattern of weakening that affected the other gutturals. In addition, however, it exhibits certain special characteristics suggesting that it lost consonantal force in certain conditions at a very early date. Though stable in initial positions, \*ʾ is lost frequently in syllable-closing positions, and always at the end of words. Quiescence of \*ʾ is attested for nouns of the type \*CaʿC- in Canaanite dialects as early as the fourteenth-century BC Amarna documents, as shown by the cuneiform spellings *ru-šu-nu* = \**rōšunu* (< \**rāšunu* < \**raʾšunu*) “our head” (EA 264:18), and *šu-ú-nu* = \**šōnu* (< \**šānu* < \**šaʾnu*) “flock” (EA 263:12). These glosses show that, at least in some Canaanite dialects, syllable-final \*ʾ was lost prior to the Canaanite Shift (\**ā* → *ō*; see §3.6.2), and the participation of Hebrew in this development is demonstrated by the noun forms *rō(ʾ)š* “head,” and *šō(ʾ)n* “flock,” in which the long vowels show that the /ʔ/, though preserved orthographically, has quiesced. When \*ʾ is lost in the related Hebrew sequences \*CiʿC- and \*CuʿC-, \**i* and \**u* are lengthened (lowered) to *ē* and *ō*, as in (*lā*)*šē(ʾ)t* (< \**šīʾt*) “to carry” and *bō(ʾ)r* (< \**buʾr*) “pit.” When syllable-final \*ʾ is lost in Hebrew verbs in which the third root consonant is ʾ (III-ʾ), a preceding *a* is lengthened to *ā*, but it does not shift to *ō*, showing that in this environment \*ʾ quiesced after the Canaanite Shift was completed: thus, *māšā(ʾ)* (< \**māšaʾ*) “he found”; *nāšā(ʾ)tā* (< \**nāšaʾtā*) “you carried.” In the same situation, a preceding *i* is, again, lengthened to *ē* – as in *yārē(ʾ)tī* (< \**yāriʾtī*) “I was afraid.” Though, in most cases, quiescent \*ʾ is preserved orthographically in Tiberian Hebrew, it is sometimes omitted altogether, as in *māšā(ʾ)tī* for \**māšā(ʾ)tī* “I found” in Numbers 11:11. In other cases, the consonantal force of \*ʾ has been restored by Masoretic hypercorrection, leading to grammatically artificial vocalizations, such as *zəʾēb* for \**zē(ʾ)b* (< \**ziʾb*) “wolf.”

### 3.6.2 Vowels

As noted in §3.2.2.1, Proto-Semitic \**ā* is realized in Hebrew as /o:/ as the result of an unconditioned sound change (\**ā* → *ō*) shared by the Canaanite languages. The *Canaanite Shift*, as it is called, is attested in Amarna glosses, such as those cited in §3.6.1 as well as *sú-ki-ni* for *sōkīni* (cf. Hebrew *sōkēn* “steward”), glossing Akkadian *rābiši* “inspector” (genitive), in EA 256:9 (cf. EA 362:69).

As noted in §3.2.2.3, the Hebrew diphthongs, \**aw* and \**ay*, are preserved and triphthongized under the tone but contracted in unaccented positions – thus, *yáyin* (< \**yayn*) “wine,” but *tēmān* (< *taymān*) “Teman, Southland.” Epigraphic evidence, however, shows that the diphthongs behaved differently in the northern and southern dialects of Hebrew. In Israelite or Northern Hebrew, \**aw* and \**ay* contracted in all positions (i.e., stressed or unstressed) – thus, *yn* (\**yēn* ~ Biblical Hebrew *yáyin*) “wine”; *tmn* (\**tēmān* ~ Biblical Hebrew *tēmān*) “Teman, Southland” – while in Judahite or Southern Hebrew, \**aw* and \**ay* were preserved in all positions – *yyn* (\**yayn* ~ Biblical Hebrew *yáyin*) “wine”; *tymn* (\**taymān* ~ Biblical Hebrew *tēmān*) “Teman, Southland.” It is clear that, as expected, the biblical pattern developed from that of the southern dialect of Jerusalem, in which diphthongs began to contract in unstressed positions during the last half of the first millennium BC.

As pointed out in §3.2.2.2, in the discussion of the phenomenon known as *qatqat* → *qītqat* dissimilation, which was generalized relatively late in the development of Hebrew, a change with this pattern (change of \**a* to \**i* in a closed unaccented syllable) occurred in prefixed verbal forms at an early date (\**yaqtal* → *yīqtal*). When final short vowels were lost in Proto-Hebrew, and the stress shifted to the ultima, the prefix vowels of singular and first-person plural verbs were most often left in closed, unaccented syllables – that is, \**yáqtulu* → \**yáqtul* → \**yáqtūl*. Whereas in Proto-Semitic the (indicative) verbal prefixes contained an *a*-vowel regardless of which of the three theme-vowels (*a*, *i*, *u*) the verb had – thus,



\**yaqtal*-, \**yaqtil*-, \**yaqtul*- – in Proto-Hebrew, and Northwest Semitic generally, the prefix-vowel of the *yaqtal*-type changed from *a* to *i*. This phenomenon was first described by Jacob Barth, and it was confirmed by H. L. Ginsberg, who showed that it was “fully operational” for Ugaritic. Thus according to the *Barth–Ginsberg Law*, as it is now commonly called, the prefix of *yqtl* in the simple active conjugation is vocalized with *i* when the thematic vowel is *a*; otherwise it is vocalized with *a* – thus *yiqtal*, but *yaqtul* and *yaqtil*. This is illustrated by Hebrew forms like \**yakbad* → *yikbad* “he is heavy,” \**yašlah* → *yīšlah* “he sends” and so forth. In Hebrew, however, the \**yi*- prefix is not limited to verbs with *a* as theme vowel, as shown by forms like *yīšpōṭ* (< \**yīšpuṭ* < \**yašpuṭ*) “he judges.” In contrast to \**yaqtal* → *yiqtal*, this change (\**yaqtul* → \**yiqtul* (→ *yiqṭōl*)) was not inherited from Proto-Northwest Semitic, as shown by syllabically written Ugaritic forms like *ia-aš-pu-tū-* for \**yašpuṭu*, corresponding to consonantal *yṭpṭ* (\**yaṭpuṭu*) “he judges.” In Hebrew, then, the form should probably be explained by simple pattern-leveling. That is, at an early stage the prefix vowel was *i* only in verbs with the stem-vowel *a*, as in Ugaritic. Subsequently, however, the *yi*-prefix was leveled through for other Hebrew verbs, namely, those with the stem-vowels *i* and *u*.

## 4. MORPHOLOGY

### 4.1 Morphological-type and word structure

Hebrew, like the other members of the Semitic family, is a fusional language. The meaning of a word is derived by inflection of a simple stem, commonly called the *root* on the basis of medieval usage – *šōreš* “root,” rendered into Latin as *radix*, hence *litterae radicales* “root letters” or “radicals,” as the individual consonants of the root (√) are still commonly called. As a rule, Hebrew words, whether verbs or nouns, are based on roots consisting of (usually) three radicals with a fixed sequence, which are inflected by affixes and/or some variation of additional morphological features, such as gemination and especially vowel patterning (*vowel gradation* or *ablaut*). The most important of these inflectional patterns are described below in subsequent sections.

Hebrew roots are predominantly triradical. Some evidence of originally biradical forms seems to survive, as in the case of certain verbs with *y* as first radical (I-*y*), which were originally \*I-*w*, a group having root allomorphs √wCC and √CC in Proto-Semitic and Afro-Asiatic (see Ch. 6, §3.3.1). As explained below in §4.5.4.2, this accounts for Hebrew forms like *šēb* (< \**šīb*) “sit!” the masculine singular imperative of √*yšb* (< \*√*yšb* < \*√[w]θb). Other Hebrew stem-types are sometimes interpreted as artificially triradical, altered from original biradicals, such as the so-called geminate roots (i.e., those with identical second and third radicals). At the same time, roots containing a glide as one of the stem consonants are often regarded as essentially biradical; these include not only the \*I-*w* and (less often) \*I-*y* roots, but also the so-called hollow or middle-weak roots (II-*w* and II-*y*) and the final-weak roots (III-*w* and III-*y*). Nevertheless, these “weak” types can also be explained as originally triradical, having developed from the partial or complete loss of one of the stem consonants by some process such as the elision of a glide in an intervocalic position. In short, the degree of biradicalism that is operative in Hebrew remains a debated point. What can be stated confidently is that, whatever the degree of biradicalism in its antecedent stages, Hebrew has been strongly conformed to a predominant triradical pattern.

Most of the small number of ostensible quadriradicals in Hebrew can be explained as products of augmentation or reduplication – for example, *garzen* “ax” (from √*grz* “cut”);

*galgal* “wheel” (from  $\sqrt{gl}$  “roll”) – and the same is true of the even rarer quinqueralicals – *səḥarḥar* “it palpitates” (from  $\sqrt{shr}$  “move around”) – when they are not in fact loanwords.

## 4.2 Noun morphology

Hebrew nouns have two genders, masculine and feminine; three numbers, singular, dual, and plural; and two states, free or absolute and bound or construct. Hebrew nouns are not marked inflectionally for case (see §4.2.3). In general, Hebrew adjectives (including verbal adjectives) are inflected like nouns.

The basic nominal paradigm is given in (2), using the nouns *yôm* “day” and *šānā* “year” as examples. Note that the plural *yāmîm* “days” is formed from a different root from that of the singular and dual, and that *šānôt* “years” has a more common alternative form – *šānîm*; these peculiarities do not obscure the inflection.

(2)		<i>Masculine</i>	<i>Feminine</i>
<i>Singular</i>	<i>Absolute</i>	<i>yôm</i>	<i>šānā</i>
	<i>Construct</i>	<i>yôm'</i>	<i>šənat'</i>
<i>Dual</i>	<i>Absolute</i>	<i>yômáyim</i>	<i>šənatáyim</i>
	<i>Construct</i>	<i>yômê'</i>	<i>šənatê'</i>
<i>Plural</i>	<i>Absolute</i>	<i>yāmîm</i>	<i>šānôt</i>
	<i>Construct</i>	<i>yāmê'</i>	<i>šənôt'</i>

### 4.2.1 Gender

As a rule, if the referent of a noun is naturally masculine, the noun will be masculine (*par* “bull”) and if the referent is naturally feminine, the noun will be feminine (*pārā* “heifer”). Nouns designating things without natural gender, such as inanimate objects or abstract ideas, may be either masculine or feminine – thus, *gēšem* “rain” (masculine), and *gib'ā* “hill” (feminine).

Though there are numerous exceptions, masculine nouns are, as a general rule, unmarked, while feminine nouns are marked. The feminine is marked by one of two endings, *-ā* (bound form *-at*) and *-t*. Although these two endings seem to have existed from an early stage in the language as unconditioned morphemic alternants, there are certain environments in which one or the other is preferred. Thus, feminine noun stems ending in a consonant cluster or a consonant preceded by a long vowel (-CC- or -V:C-) are marked by *-ā* – as in *'iššā* “woman” and *'ēšā* “counsel” – while *-t* follows forms ending with a vowel – *mišrît* “Egyptian” – and, very characteristically, is preferred on active participles, often leading to a “segholated” (cf. §3.4) ending – thus *\*yāšibt* → *yōšebet* “sitting” (the *-t* ending is used much more widely in Rabbinic Hebrew than in Biblical Hebrew). There are also many unmarked feminine nouns, including some with naturally feminine meaning, such as *'ēm* “mother,” and others designating inanimate objects, such as *'ēben* “stone” and *'ir* “city.”

### 4.2.2 Number

Plural nouns and adjectives in the unbound state are most often marked by the endings *-îm* and *-ôt* (for nouns in the bound state, see §4.2.4). The great majority of the former are masculine and the latter feminine, as suggested by (2). There are, however, numerous masculine nouns with the *-ôt* plural ending – thus, *'āb* “father,” *'ābôt* “fathers,” and *māqôm* “place,” *māqômôt* “places” – and a few that have both *-îm* and *-ôt* – for example, *nāhār* “river,”

*nəḥārîm* and (more often) *nəḥārôt* “rivers.” Similarly, several feminine nouns, whether or not they are marked as feminine in the singular (see §4.2.1) and whether or not they have natural feminine referents, take the *-îm* plural ending. Examples of marked feminine singular nouns with *-îm* plural endings include the natural feminine *’iššâ* “woman,” *nāšîm* “women,” but also *ḥiṭṭâ*, *ḥiṭṭîm* “wheat” (see also [2] above for *šānâ*, which usually forms its plural as *šānîm* but frequently as *šānôt*). Examples of unmarked feminine singular nouns with *-îm* plural endings include the natural feminine *rāḥēl* “ewe,” *rāḥēlim* “ewes,” but also *’îr* “city,” *’ārîm* “cities.” Examples of unmarked feminine singular nouns with *-ôt* plural endings include the natural feminine *’ēm* “mother,” *’immôt* “mothers,” but also *’éreṣ* “land,” *’ārāšôt* “lands.” Certain unmarked nouns that are construed sometimes as masculine and sometimes as feminine may have both plural endings – thus *’āb* “cloud” (usually masculine, but feminine in 1 Kings 18:44), *’ābîm* and *’ābôt* “clouds.”

In Late Biblical Hebrew the plural ending *-în* alternates with *-îm* (cf. *yāmîn* “days,” in Daniel 12:13), and in Rabbinic Hebrew *-în* is increasingly preferred. Though this development may have been influenced by Aramaic, it probably had its origin in dialect variation within Hebrew, since its distribution in the Bible is not exclusively confined to the latest literature and, in fact, occurs once in the most archaic poetry (*middîn* “carpets,” in Judges 5:10). Its ultimate explanation is the existence in Proto-Northwest Semitic of *\*-m* and *\*-n* allomorphs of the Proto-Semitic plural/dual boundness marker *\*-n* (see §4.2.4).

Although the dual is used in some Semitic languages, such as Ugaritic and Arabic, to refer to two of anything, its use in Biblical Hebrew is largely confined to natural pairs, such as *’oznáyim* “ears,” or *na’āláyim* “sandals,” or to numerals (*šəṇáyim* “two”) and double units of measurements of time or quantity: for example, *šəbū’áyim* “two (successive) weeks, a fortnight”; *’ammātáyim* “two cubits.” Probably as the result of a dialectal survival, the original broader use of the dual returns in Rabbinic Hebrew, where it can denote a pair of anything.

With unmarked nouns, the unbound dual ending, *-áyim*, is added directly to the base of the singular – thus, *ragláyim* ((*régel* < ) *\*ragl-* + *-áyim*) “feet” (masculine); and *yādáyim* ((*yād* < ) *\*yad* + *-áyim*) “hands” (feminine). With nouns marked as feminine, the ending is added to the singular base following one of the two types of marker (see §4.2.1), as follows. Nouns ending in *-â* (bound form *-at*) follow the pattern of *šāpâ* (bound form *šəpat*) “lip,” *šəpātáyim* “lips.” Nouns ending in *-t* follow the pattern of *nəḥōšet* (bound form *nəḥōšet* < *\*nuḥušt*) “bronze,” *nəḥuštáyim* “bronze fetters,” unless assimilated to the preceding pattern, as evidently in the case of *délet* (< *\*dalt*, but bound form *dəlat*) “door,” *dəlātáyim* “(double) doors.”

Adjectives follow more restricted rules with regard to number. The kind of variety displayed by nouns in forming *-îm* and *-ôt* plural, as described above, is lacking in adjectives (including participles), the masculine plurals of which are consistently marked by *-îm* and feminine plurals by *-ôt*. Also, dual endings do not occur with adjectives.

### 4.2.3 Case

In Proto-Northwest Semitic, the three short vowels were used to indicate case in singular nouns – *\*-u* for nominative, *\*-i* for genitive, and *\*-a* for accusative – and, following *\*-āt-*, in feminine plural nouns – *\*-ātu* for nominative and *\*-āti* for oblique. The loss of final short vowels and the leveling of the *\*-îm* ending on masculine plurals (see §4.2.2) left Hebrew nouns with no inflectional indication of case, except perhaps the bound–unbound opposition in genitive constructions. As a result, the case of nouns may be identified only from syntactical criteria.

#### 4.2.4 State

In Biblical Hebrew, as in other Semitic languages that have lost the Proto-Semitic system of case endings, the chief way to express a genitive relationship is the so-called construct chain (on the role of the construct chain in the determination of substantives, see §5.4). A construct chain consists of the juxtaposition of two or (rarely) more nouns in a sequence such as *dābar hammēlek* “the word of the king.” In this example, *hammēlek* “the king” is free in form like other nouns not forming parts of construct chains. It derives a genitive force, however, from its relationship to the preceding bound form *dābar*′. In traditional terminology, *dābar*′, the *nomen regens*, is said to “govern” *hammēlek*, the *nomen rectum*.

The two parts of a construct chain are closely associated accentually, with the principal stress moving ahead to the *nomen rectum*, which therefore remains morphologically unchanged and in what is called the *absolute state*. The *nomen regens*, however, becomes proclitic and often undergoes changes (especially including vowel shortening or reduction) in consequence of the loss of stress, so that it is said to be in the *construct state* – compare *dābār* “word” (absolute state) to *dābar*′ (construct state). The changes that affect singular nouns in the construct state include vowel reduction in newly unstressed syllables (*ā* → *ə* and *ē* → *ə*) and the shift of *ā* to *a* in final closed syllables (both illustrated, again, by the contrast *dābār* ~ *dābar*′). Nouns ending in a final stressed *səgôl* (–*eh* = 𐤌) become final *šērē* (–*eh* = 𐤌) in construct; for example, *maḥāneh* “camp” (absolute) ~ *maḥānēh*′ (construct).

As noted in §4.2.2, plural nouns in the absolute state normally end in –*ôt* (usually feminine) or –*īm* or, in Rabbinic Hebrew, –*in* (usually masculine). The –*īm* and –*in* endings are survivals of a Proto-Semitic boundness marker for plural and dual nouns, *\*-n(a)*. That is, free or unbound Proto-Semitic nouns ended in *\*-m* following short vowels and *\*-n(a)* following long vowels and diphthongs, so that nouns lacking these endings were “marked” as bound or construct (see Ch. 6, §3.3.2.1). In the evolution of the descendant languages the two endings were leveled and otherwise simplified. In the Northwest Semitic group the short-vowel ending, *\*-m*, disappeared, so that the bound–unbound contrast was lost in singular and feminine plural nouns until the later sound changes already described developed as the result of the proclisis of bound forms. On the other hand, the long-vowel ending, *\*-n(a)*, survived as a marker of the absolute plural and dual. Original *\*-n(a)* was realized, however, as –*n* in some Northwest Semitic languages (Aramaic, Moabite, the Deir ‘Alla dialect, and Rabbinic Hebrew) and as –*m* in others (Ugaritic, Phoenician, Ammonite, and Biblical Hebrew).

The bound or construct endings of plural nouns are –*ê* (..) corresponding to –*īm* in the absolute state, and –*ôt* corresponding to –*ôt* in the absolute. When pronominal suffixes are added to plurals ending in –*ôt*, the plural bound-form ending –*ê* (< *\*-ay*–) is interposed – thus, *mišwôtêkâ* “your commandments.” Not all of these forms can be readily explained in relation to the antecedent forms reconstructed for Proto-Northwest Semitic.

The Proto-Northwest Semitic forms of the unmarked, usually masculine, unbound plural were *\*-ūn* in the nominative and *\*-īn* in the oblique, corresponding to *\*-u* nominative, *\*-a* accusative, and *\*-i* genitive in the singular (see §4.2.3). When the loss of final short vowels caused the case system to collapse in the singular, the endings were leveled in the plural as well, and the oblique form, *\*-īn*, was generalized (as –*īn* or –*īm*, as explained above). At this point, the corresponding bound form in the plural must have been *\*-ī*, but for unknown reasons this form was abandoned in favor of the corresponding dual form, *\*-ay* (→ –*ê*; see below).

The Proto-Northwest Semitic forms of the marked, usually feminine, unbound plural were *\*-ātu* in the nominative and *\*-āti* in the oblique. With the loss of final short vowels these fell together as *\*-āt*, the expected antecedent form of –*ôt*. It is unknown, however, why

the newly formed masculine plural bound form *-ê-* (<\*-ay-) came to be inserted before suffixes added to these nouns.

For dual nouns the construct ending is *-ê* (ʿ..), originally *\*-ay*, corresponding to *-áyim* in the absolute state. In unmarked nouns *-ê* is added directly to the end of the base – thus *raglê'* “feet” (masculine), and *yəḏē'* “hands” (feminine). With nouns marked as feminine, the endings are added following the marker, as explained in connection with the dual absolute endings in §4.2.2 above – as in *šiptē'* “lips.”

In Rabbinic Hebrew, though the construct chain is still used frequently to express the genitive, it is increasingly replaced by a construction in which nouns are joined by the genitive particle *šel*, which arose from a combination of the relative particle *še-* (see §4.3.3) and the proclitic preposition *lə-* “belonging to, of” – thus, *haddābār šellamélek* or, more commonly, *haddābār šel hammélek* “the word of the king.” The *nomen regens* in such a construction is not in the construct state, and it may have an anticipatory pronominal suffix – thus already in Late Biblical Hebrew, *hinnēh miṭṭātô šellislōmōh* “There is the couch of Solomon” (Song of Songs 3:7).

## 4.2.5 Noun formation

The various Hebrew noun- and adjective-types are derived from the application of several kinds of operations to verbal roots, including vowel patterning, root consonant gemination and affixation. Though several noun-types have general or specific semantic associations, there are many others for which such associations cannot be identified. The following tabulation provides a selection of some of the most important noun-types. In arrangement it proceeds from the simpler to the more complex forms, and the paradigm root used is  $\sqrt{qtl}$  ( $\sqrt{ql}$  for biradical types). Except where indicated, the examples come from Biblical Hebrew.

### 4.2.5.1 Biradical types

The pattern CV:C (< \*CVC) includes a number of common nouns of the types *qāl* (< \**qal*) – thus, *dām* “blood”; *dāg* “fish.” The associated feminine forms are *qālā* (< \**qalat*; e.g., *bāmā* “high place,” *šānā* “year”) and *qélet* (< \**qalt*; e.g., *qéšet* “bow”; cf. Northern Hebrew *št* = \**šatt* (< \**šant*) “year”). The active participle of roots II-*w/y* is formed from this pattern – thus, *bā'* and (feminine) *bā'ā* “coming”; *šām* and (feminine) *šāmā* “placing.” Two members of this group, *'āb* “father” and *'āh* “brother” (plural *'aḥim* < \**'aḥīm*), have their construct form in *-ī* (*'ābī*), suggesting that these words had (anomalously) long singular case vowels in Proto-Semitic, the vowel of the genitive (\*-ī) having been leveled through the paradigm after the collapse of the case system. The CV:C pattern also includes nouns of the type *qēl* (< \**qil*): thus, *'ēl* “god,” *'ēš* “tree.” Again there are two associated feminine forms, namely, *qēlā* (< \**qilat*; e.g., *bēšā* “egg,” *mē'ā* “hundred”) and *qélet* (< \**qilt*), which forms the infinitive construct of roots *\*I-w* and some roots *I-n*: for example, *šebet* ( $\sqrt{yšb}$  <  $\sqrt{*wθb}$ ) “to sit”; *rēšet* (<  $\sqrt{yrs}$  <  $\sqrt{*wrθ}$ ) “to take possession of”; *géšet* ( $\sqrt{ngš}$ ) “to approach.” Though *qēl* is the absolute, presuffixal, and construct form for most members of this group, the common nouns *šēm* “name” and *bēn* “son” have the presuffixal forms *šəm-* and *bən-* and (sometimes) the construct forms *ben-* and *šem-* (the latter is rare). Another common noun, *bat* “daughter,” belongs to this pattern (*\*qilt*): *\*bint* → *\*bitt* → *\*batt* (by Philippi’s Law, see §3.2.2.2) → *bat*.

Nouns of similar form but deriving from a biradical type containing an originally long vowel, CV:C (< \*CV:C), include the patterns *qōl* (< \**qōl* < \**qāl*; e.g., *qōl* “voice,” *hōl* “sand”); *qīl* (< \**qīl*; e.g., *šīr* “song,” *qīr* “wall”); and *qūl* (< \**qūl*; e.g., *sūs* “horse,” *rūaḥ* “wind”). From *qīlā*, the feminine corresponding to *qīl*, come the nouns *bīnā* “understanding” and *qīnā* “dirge.” The infinitive construct of roots II-*w* is formed from the *qūl* pattern – thus, *qūm*

“to arise” – and that of roots II-*y* is formed from the *qîl* pattern – thus, *dîn* “to judge,” *rib* “to contend.”

#### 4.2.5.2 Triradical types without a doubled radical

The pattern CVCeC (<\*CVCC) constitutes an important group of nouns (the “segholates,” see §§3.4–3.5), which, when derived from sound roots, take the forms *qétel* (<\**qatl* or \**qitl*), *qétel* (<\**qitl*) and *qótel* (<\**qutl*). A distinctive feature of the segholates, including their feminine forms (\*CVCCat), is the formation of the plural from the base \*CVCaC (feminine \*CVCaCat); i.e., with *-a-* interposed between the second and third radicals. A large number of common nouns belong to the group *qétel* (<\**qatl*): *mélek* “king,” *késep* “silver,” *’éres* “earth,” *gépen* “vine,” *kéleb* “dog,” *’ébed* “slave,” and so forth. Almost as large is the group \**qitl*, including *qétel* (<\**qitl*; e.g., *séper* “book,” *šébeṭ* “rod”) – and *qétel* (<\**qitl*; e.g., *šédeq* “righteousness,” *qéreb* “midst”). The corresponding feminine is *qitlâ* (<\**qitlat*): for example, *šiphâ* “maidservant,” *gib’â* “hill,” *yir’â* “fear,” but also *herpâ* “reproach,” *’erwâ* “nakedness.” When based on an active verbal root, \**qitl(at)* nouns frequently have a passive sense – thus, *šéma* “report” (something heard) from √*šm* “hear”; *zéker* “memory” (something remembered); *zéba* “sacrifice” (something sacrificed); compare *’émeq* “valley” (something deep), from the stative verb √*mq* “be deep.” (Note: the presence of two types, *qétel* and *qétel*, from \**qitl*, and the convergence of *qétel* <\**qitl* with *qétel* <\**qatl* present problems in interpreting the Tiberian tradition, and when the evidence of the Babylonian [e.g., *málak* ~ Tiberian *mélek* and *qárab* ~ Tiberian *qéreb*] and Hexaplaric traditions is added, a number of ambiguities involving nouns of the type *qatl* and *qitl* emerge.) The third group of segholates, *qótel* (<\**qutl*), also includes some common nouns: for example, *bóqer* “morning,” *hódeš* “month,” *šóreš* “root”; *’óraḥ* “path.” Nouns of this group are frequently abstract (e.g., *qódeš* “holiness”), especially when derived from stative roots – thus, *róḥab* “width,” *góbah* “height,” *hóšek* “darkness.”

Another large, important group is represented by the pattern CV:CV:C (<\*CVCVC). This pattern is especially characteristic of adjectives, but it produces many common nouns as well. The group *qātāl* (<\**qatal*) includes a number of primary nouns having the form *qātāl* – such as *zāhāb* “gold,” *nāhār* “river” – but some of the nouns in this group are clearly collectives, such as *qāhāl* “assembly” and *bāqār* “cattle,” and it is possible to interpret many of the others in this way, including *’āpār* “dust,” *’ānān* “cloud,” *māṭār* “rain,” and possibly *’ādām* “man, person, humanity”; it has been suggested that some of these derive from a Proto-Semitic \**qatal* plural morpheme. The same type (*qātāl*) is especially productive of abstract nouns derived from verbs, which may be active (e.g., *hāmās* “distortion,” *nāqām* “vengeance”) or stative (e.g., *’āšām* “guilt,” *šābā* “satiety,” *rā’āb* “hunger,” *šāmā* “thirst”). The corresponding feminine form is *qetālâ* (<\**qatalat*) – for example, *’ādāmâ* “soil” – which, like *qātāl*, is characteristic of abstract nouns, such as *šedāqâ* “righteousness” and *bērākâ* “blessing.” Finally, and most typically, the group *qātāl* (<\**qatal*) contains numerous adjectives from stative roots, such as *hādāš* “new,” *rāšā* “evil,” *hāzāq* “strong,” *lābān* “white,” *šāpāl* “low,” and so forth. This is also true of the groups *qātēl* (<\**qatil*) – such as *zāqēn* “old,” *šāmēah* “joyous,” *ṭāmē* “unclean” – and *qātōl* (<\**qatul*): thus, *gādōl* “big,” *’āmōq* “deep,” *mātōq* “sweet,” *ṭāhōr* “clean,” *qārōb* “near,” *rāḥōq* “distant.”

The pattern CV:CV:C (<\*CVCV:C) is especially productive of adjectives, many of which are substantivized as nouns. The type *qātōl* (<\**qatāl*), however, is primarily nominal. Though it includes a few primary nouns – such as *šālōš* “three,” *’ātōn* “jenny” – it specializes as the form of the infinitive absolute of the simple verbal stem (Qal) – thus, *kātōb* “to write.” Other well-known nouns with this form, such as *šālōm* “peace” and *kābōd* “glory,” are like the infinitive in expressing the abstract idea of the verb. The type *qātīl* (<\**qatil*), though



it includes few primary substantives, frequently forms adjectives from verbs, whether from stative roots (*ḥāsîd* “pious,” *šā’îr* “little”) or active roots. Adjectives formed in this way from active roots tend to be passive in meaning and may be substantivized, such as *’āsîr* “bound,” substantivized as “prisoner,” and *šākîr* “hired,” substantivized as “hireling.” Many of these adjectives, when substantivized as passive, function as nouns of office – thus, *nāgîd* “prince” (i.e., “designee”); *māšîaḥ* “messiah” (“anointed one”); *nāšî* “chief” (“one who is lifted up”); *pāqîd* “commissioner” (“one who is appointed”). The type *qātûl* (< \**qatûl*), though again it includes few primary nouns, is a common adjectival pattern from stative roots – thus, *’arûm* “clever,” *’ašûm* “strong.” Most importantly, *qātûl* is generalized as the passive participle for active roots of the simple verbal stem (Qal) – thus, *kātûb* “written”; *pātûaḥ* “open.”

The particular importance of the pattern CōCVC (< \*CāCVC) is the role of the type *qôtêl* (< \**qātîl*), feminine *qôtêlâ* (< \**qātîlat*) and *qôtélet* (< \**qātîlt*), in forming the active participle of the simple verbal stem (Qal): for example, *yôrêd*, *yôrêdâ*, *yôrêdet* “going down.” These are often substantivized – thus, *kōhên* “priest,” *sōpêr* “scribe,” *yô’êš* “counsellor,” *gō’êl* “kinsman,” *ḥôtên* “father-in-law,” *yôlēdâ* “woman in labor” (with retention of *ê* in the substantive).

#### 4.2.5.3 Triradical types with doubling of the second radical

The pattern CVCCV:C (< \*CVCCVC) includes mostly adjectives, many of which may be substantivized. The type *qattâl* (< \**qattal*) is an adjectival pattern that usually denotes habitual action – thus, *qannâ* “jealous,” *ḥaṭṭâ* “sinful,” *naggâḥ* “accustomed to gore” (of the ox in Exodus 21:29 and 36), and *’awwâl* “unjust,” substantivized as “unjust person.” When substantivized, this form is especially characteristic of nouns of occupation – thus *dayyân* “judge,” *ṭabbāḥ* “cook,” *gannâb* “thief,” *ḥārâš* (< \**ḥarrâš*) “craftsman” (Rabbinic Hebrew adds to this category a number of examples not found in Biblical Hebrew: e.g., *baqqār* “cattle rancher,” *hārâg* (< \**harrâg*) “murderer,” *gammâl* “camel driver”). The type *qittêl* (< \**qattîl* by a pre-Hebrew sound change) belongs to a number of adjectives denoting physical conditions: thus, *’iwwêr* “blind,” *ḥērêš* (< \**ḥirrêš*) “deaf,” *gibbêaḥ* “bald,” *’ittêr* “disabled” (of the right hand → “left-handed” in Rabbinic Hebrew).

#### 4.2.5.4 Types with derivational affixes

Nouns with preformative mV- constitute a large group with a wide variety of meaning. Two of the most important types, \**maqtal* and \**miqtal*, have fallen together by *qatqat* → *qitqat* dissimilation (see §3.2.2.2) as *miqtâl*, with its feminine forms *miqtâlâ* and *miqtélet*. Examples include *midbâr* “pasture land,” *mišpâṭ* “judgment,” *mišpāhâ* “clan,” and *milḥāmâ* “battle.” In phonological situations involving a guttural, liquid, or nasal as the first root consonant, however, initial *ma-* may occur in nouns of either original type (\**maqtal* or \**miqtal*) – thus, *ma’ākāl* “food,” *ma’ārāb* “west,” *mal’āk* “messenger,” *mamlākâ* “kingdom,” *mattân* (< \**mantan*) “gift,” *maššâ* (< \**manšâ*) “burden, oracle.”

Among sufformatives may be mentioned (i) *-ôn* (< \**-ān*), which forms a number of substantives, especially from roots III-w/y – for example, *ḥāzôn* “vision,” *gā’ôn* “pride,” *hāmôn* “sound” – as well as adjectivals, such as *’aḥārôn* “behind, latter,” and *ḥiṣôn* “outer”; (ii) *-ût* (< \**-ūt*), which forms abstracts from concrete nouns – *malkût* “kingdom” (from \**malk* “king”), *’almānût* “widowhood” (cf. \**alman(at)* “widow”), *yaldût* “youth” (from \**yald* “child”), and (iii) *-î* (< \**-îy*), a common affix for forming adjectives from nouns, which is used especially to generate ordinals – such as *šālîšî* “third” – and gentilics, which may be substantivized – thus, *raglî* “on foot,” substantivized as “footman, foot-soldier” (from \**ragl* “foot”), *yəḥûdî* “of Judah, Jewish,” substantivized as “Judahite, Jew.”

### 4.2.6 Stem allomorphy

The early shift of stress to the final syllable (see §3.5) and the subsequent vowel changes that resulted in the course of inflection and suffixation (see §3.2.2.2) led to a wide variety in stem-form in many Hebrew nouns and adjectives. This stem allomorphy is among the most distinctive characteristics of the language in its later development. Note, for example, the variety of nominal stems found in the inflection of *dābār* “word”: unbound singular stem *dābār*; bound singular stem (with forward shift of stress) *dəbar*’; presuffixal singular stem *dəbār*- (light suffixes; see §4.3.1) or *dəbar*- (heavy suffixes); unbound plural stem (with forward shift of stress) *dəbār*-; bound plural stem *dibrē*’; presuffixal plural stem *dəbār*- or *dibrē*- (see §4.3.1).

## 4.3 Pronominal morphology

Hebrew has personal, demonstrative, relative, interrogative, and indefinite pronouns. There is no separate reflexive or resumptive pronoun, though the oblique cases of the pronominal suffixes may be used reflexively or resumptively (retrospectively) – the latter very commonly in relative clauses.

### 4.3.1 Personal pronouns

Hebrew personal pronouns occur in two forms, independent and enclitic (the pronominal suffixes). Both types are inflected for number, person, and gender. There are complete paradigms of singular and plural forms, but the Proto-Semitic dual forms, which may be reconstructed for the oblique cases at least (see Ch. 6, §3.3.3), have been generally lost (but see below). First-person personal pronouns have common gender, while second- and third-person personal pronouns have distinct masculine and feminine forms.

The standard forms of the independent personal pronouns, which serve as the nominative case (i.e., as subject or predicate nominative), are as follows.

(3)	Person	Gender	Number	
			Singular	Plural
	First	Common	’ānōkī, ’ānī	’ānāḥnū
	Second	Masculine	’attā	’attem
		Feminine	’att	’atten, ’attēnnā
	Third	Masculine	hū’	hēm, hēmmā
		Feminine	hī’	hēnnā

Although ’ānōkī and ’ānī are both widely used in Biblical Hebrew, the former is more common in earlier biblical literature, while the latter is predominant in the later literature, especially Late Biblical Hebrew, and survives alone in Rabbinic Hebrew. In Biblical Hebrew ’ānāḥnū has a rare variant, *nāḥnū*; in Rabbinic Hebrew (and already in Jeremiah 42:6 and at Qumran) both are replaced by ’ānū. The second-person singular forms exhibit some variety. Thus ’attā (masculine) is sometimes spelled ’t in Late Biblical Hebrew (vocalized as ’att or ’attā) and Qumran, while in Rabbinic Hebrew and the Hexapla the two forms alternate; ’att (feminine) is spelled ’ty occasionally in Biblical Hebrew (always vocalized as \*att) and regularly in the Samaritan Pentateuch. Both ’t (masculine) and ’ty (feminine) are likely to have arisen under Aramaic influence, though dialectal intrusion cannot be ruled out for the earlier examples, especially in the case of ’ty, which indicates the typologically earlier pronunciation \*attī. As with certain verb forms (see §4.5.4.1), the masculine and feminine forms of the personal pronouns show a tendency to merge in Rabbinic Hebrew, so that



'*attem* and '*atten*, on the one hand, and *hēm* and *hēn* (which has replaced Biblical Hebrew *hēnnā*), on the other, alternate in both the masculine and feminine.

The pronominal suffixes of the noun serve as the genitive of the personal pronoun when attached to substantives or prepositions (the latter corresponding most often to the dative or ablative in Indo-European and other languages), and the accusative when attached to verbs and certain particles:

#### (4) The pronominal suffixes on singular nouns

Person	Gender	Number	
		Singular	Plural
First	Common	-î	-ēnû
Second	Masculine	-ēkā	-kem
	Feminine	-ēk, ēkî	-ken
Third	Masculine	-ô, -ēhû	-ām
	Feminine	-āh, -'ēhā (תָּהָא)	-ān

As noted, these suffixes are genitive. They are inflected for singular and plural number. In Biblical Hebrew, however, there seem to be isolated survivals of the Proto-Semitic dual pronouns, as preserved, for example in Ugaritic (Ch. 9, §4.3.1.2) and Arabic. These occur in passages where apparently masculine plural pronominal suffixes of the second or third person have feminine pairs as antecedents, such as 2 Samuel 6:7, 10, and 12, where *-hem* and other ostensibly masculine suffixes occur in place of *-hen*, and so forth, referring to the feminine antecedent *pārôt* “(a yoked pair of) cows”; to *-hem* compare the corresponding dual pronouns in Ugaritic, *-hm*, and Arabic, *-humā*.

The (genitive) pronominal suffixes for dual and plural nouns are presented in (5):

#### (5) The pronominal suffixes on plural nouns

Person	Gender	Number	
		Singular	Plural
First	Common	-ay	-ēnû
Second	Masculine	-'ēkā (תָּכָא)	-ēkem
	Feminine	-āyik	-ēken
Third	Masculine	-āyw	-ēhem
	Feminine	-'ēhā (תָּהָא)	-ēhen

These suffixes are added to the noun stem, followed by the plural construct ending *-ê* (< \**-ay*), originally the dual stem (see §4.2.4). This applies both to masculine (*dabārāyik* “your (fem. sg.) words”) and feminine (*hômôtāyik* “your (fem. sg.) walls”) nouns.

In archaic and poetic contexts, the third-person masculine plural suffix has the variant *-āmô* on singular nouns and *-ēmô* on plural nouns. There is also evidence of variant traditions in the pronunciation of the second-person masculine singular pronominal suffix. Although this suffix is consistently vocalized *-ēkā* on both singular and plural nouns in Tiberian Hebrew, it is usually spelled with final *-k* (i.e., תָּ not תֵּ), and the Hexaplaric form is consistently *-akh* (-αχ); taken together, these things point to a non-Masoretic pronunciation *-āk*, which corresponds to the Rabbinic Hebrew form. On the other hand, the antiquity of the Tiberian vocalization is confirmed by the heavy predominance of the spelling תֵּ at Qumran.

When one of the genitive suffixes is added to a noun, the stress in the resulting word usually shifts to the suffix, causing an alteration in the form of the noun stem as the result of vowel

reduction in accordance with the rules summarized in §3.2.2.2. It follows that the form of the noun stem before suffixes is often similar or identical to the form of the noun stem in the construct state, which is typically altered by the same kind of shift of stress and consequent vowel reduction (see §4.2.4). Thus for the noun *dābār* “word,” the corresponding forms are construct singular *dābarʾ* “word (of)”; suffixed singular *dābarkem* “your (masc. pl.) word”; construct plural *dibrêʾ* “words (of)”; and suffixed plural *dibrêkem* “your (masc. pl.) words.”

In the suffixal forms of singular noun stems, variation may occur before the so-called *heavy* and *light* suffixes. The heavy suffixes are those beginning with a consonant, namely, *-kem* and *-ken*. In the case of the light suffixes, the noun ends with an open syllable, causing the stem-vowel to lengthen (cf. §3.4) – thus, *dābarkem* “your (masc. pl.) word,” but *dābārēkā* “your (masc. sg.) word”; *ḥômatkem* “your (masc. pl.) wall,” but *ḥômātēkā* “your (masc. sg.) wall.”

In the suffixal forms of plural noun stems, the double reduction leading to *dibrê-*, the form required by the “rule of šəwā” (see §3.4), occurs only with the second- and third-personal plural suffixes (i.e., those which are bisyllabic and accented on the final syllable).

Although the suffixal forms of most noun stems are produced by these rules, there are numerous other variations, many predictable on historical grounds – such as, *ʾōz* (< \**uzz-*) “strength,” suffixed form *ʾuzzakem* “your strength” – others simply irregular – for example, *yād* (< \**yad-*), heavy suffixed form *yedkem* “your hand.” A few noun stems are unchanged by suffixation – thus, *sūs* “horse,” suffixed form *sūsām*; *sūsīm* “horses,” suffixed form *sūsēhem*.

The attested forms of the pronominal suffixes when attached to the perfect verb are presented in (6):

(6) The pronominal suffixes on perfect verbs

Person	Gender	Number	
		Singular	Plural
First	Common	-ánî	-ánû
Second	Masculine	-əkā, -ėkā	-kem
	Feminine	-ėk	—
Third	Masculine	-ô, -ăḥû	-ām
	Feminine	-ăḥ	-ān

As noted, these are object suffixes. The forms shown are those used when the suffix is stressed and follows a verbal stem ending in a consonant, such as *šəlāḥáni* “he sent me.” The forms are slightly different when the suffix is unstressed and/or when following a stem ending in a vowel – thus, *šəlāḥátnî* “she sent me,” *šəlaḥtínî* “you (fem. sg.) sent me.”

The attested forms of the (accusative) pronominal suffixes when attached to the imperfect verb are presented below.

(7) The pronominal suffixes on imperfect verbs

Person	Gender	Number	
		Singular	Plural
First	Common	-énî, énnî	-énû, -énnû
Second	Masculine	-əkā, -ėkkā	-kem
	Feminine	-ėk	—
Third	Masculine	-ėḥû, énnû	-ēm
	Feminine	-ėḥā, -énnā	-ēn

In the case of the imperfect, the object pronouns follow *-ê-* or *-én-*, which is suffixed to the verbal stem. The forms with *-nn-* suggest a derivation from the Proto-Northwest Semitic energetic (see §4.5.2).

### 4.3.2 Demonstrative pronouns

In Hebrew the demonstrative pronouns are inflected for gender and number. The common forms of the near (“this, these”) and far (“that, those”) demonstratives are listed in (8):

(8)			<i>Singular</i>	<i>Plural</i>
<i>Near demonstrative</i>	<i>Masculine</i>		<i>zeh</i>	<i>’élleh</i>
	<i>Feminine</i>		<i>zō(’)t</i>	<i>’élleh</i>
<i>Far demonstrative</i>	<i>Masculine</i>		<i>hû’</i>	<i>hēm</i>
	<i>Feminine</i>		<i>hî’</i>	<i>hěnnâ</i>

Note that the far demonstratives are identical to the independent personal pronouns of the third person. The masculine and feminine singular far demonstratives showed an early tendency to merge, so that the feminine form is spelled *hw’* throughout the Pentateuch, though it is consistently vocalized *hî’* by the Masoretes. The forms *zōh* and *zô* appear in Biblical Hebrew as rare variants of *zō(’)t*, and *zô* became the regnant form in Rabbinic Hebrew. The longer forms *hallaz* (“this,” common), *hallázeh* (“this,” masculine), and *hallēzû* (“this,” feminine), which occur in Biblical Hebrew as rare synonyms of *zeh* and *zō(’)t*, constitute in Rabbinic Hebrew a full alternate paradigm of the near demonstrative, to which *hallálû* (“these,” common) provides the plural.

The demonstratives are used as both pronouns and adjectives, and, as adjectives, they are subject to the same rules of gender agreement and definiteness as other adjectives – compare *zeh hā’îš* “this is the man,” to *hā’îš hazzeh* “this man” (on the article see §4.4).

### 4.3.3 Relative pronouns

The common relative pronoun in Biblical Hebrew is *’ăšer*, which is indeclinable. Less often, in Archaic Hebrew and especially in Late Biblical Hebrew, the proclitic form *še-* (with gemination of the following consonant if possible) is found instead. In Rabbinic Hebrew this form replaces *’ăšer* almost entirely. Occasionally, and almost exclusively in poetry, *zeh* and *zû* are used as relatives (Psalm 74:2; Isaiah 42:24), recalling their derivation from the old relative-determinative pronoun *\*ð-* (see Ch. 6, §3.3.4).

These forms are of disparate origin. Voiceless and voiced relative particles, *\*θ-* and *\*ð-*, must be posited for Proto-Northwest Semitic. The former (*\*θ-*) is the base of the Hebrew relative *še-*, as well as Standard Phoenician and Ammonite *’š-* and Phoenician-Punic *š-* (see Ch. 11, §4.3.5). The latter (*\*ð-*), as noted, underlies the relative use of Hebrew *z-*. Hebrew *’ăšer* and Moabite *’šr* are thought to have arisen from a form of the substantive *\*’aθr-* “place.”

It is probable that the variation in Hebrew between *’ăšer* and *še-* was originally dialectal, the former, shared by Moabite, having been the southern (Judahite or at least Jerusalemite) form, and the latter, which has cognates in Phoenician and Ammonite, having been the northern (Israelite) form.

### 4.3.4 Interrogative and indefinite pronouns

The interrogative pronouns are *mî* “who?” and *mah* “what?” Neither is inflected for gender or number. In comparison to Common Semitic *\*man* “who,” Hebrew *mî* is an innovation

(\**mi:y-*) shared with Ugaritic (*my*), Old Canaanite (cf. *mi-ya* in EA 85:63; 94:12 and 116:67), Phoenician (*my*), and probably Ammonite (*m-*). The first consonant of the word following *mah* is doubled when possible (otherwise the vocalization of *mah* may be affected). This suggests that although the *-h* in the Tiberian spelling of *mah* (מַח) is a *mater* (see §2.2), the primitive form may have been \**mah* (with consonantal *-h*), especially in light of Ugaritic *mh* “what?” (see Ch. 9, §4.3.4.1).

Both *mî* and *mah* are used as indefinite pronouns in the sense of “whoever” and “whatever”: for example, *mî yārē’ wāḥārēd yāšōb* “whoever is fearful and trembling, let him turn back” (Judges 7:3). When *mî* and *mah* are used as indefinites in Rabbinic Hebrew they are usually augmented by the relative *še-* (see §4.3.3) and preceded by the proclitic substantive *kol’* – thus, *kol-mî še-* “whoever,” and *kol-mah še-* “whatever.”

The Proto-Semitic interrogative \**ʾayy-* (see Ch. 6, §3.3.4), from which a group of Hebrew interrogative adverbs is derived (*ʾayy-* + pronominal suffix “where?”; *ʾayyēh* “where?”; *ʾēk* “how?” etc.), was combined with the near demonstratives in Rabbinic Hebrew to produce another series of interrogative pronouns/adjectives: *ʾēzeh*, *ʾēzehû* “who? which?” (masculine singular); *ʾēzô*, *ʾēzôhî* “who? which?” (fem. sg.); *ʾēlû* “who? which? (common pl.) – compare *ʾē-zeh* “which?” already in Ecclesiastes 2:3 and 11:6.

#### 4.4 The article

The Hebrew definite article is prefixed directly to the noun it determines (on determination of substantives, see §5.4). The usual form of the article is *ha-* with gemination of the following consonant: for example, *hammélek* “the king.” When gemination is not possible, as in the case of nouns with initial guttural consonants or *r* (see §3.3.2), and in certain other circumstances, there is alternation of the length or quality of the vowel of the article itself. Like other Semitic languages, Hebrew lacks an indefinite article.

#### 4.5 Verbal morphology

Finite Hebrew verbs have two indicative forms, which contrast aspectually as perfective and imperfective (for the Proto-Northwest Semitic origins of the Hebrew indicatives, see §4.5.1). Both forms have three persons, two genders and two numbers (singular and plural). The *perfect* is inflected by the modification of a verbal stem through the addition of suffixes indicating person, gender, and number – thus, *stem* + *suffix*. The *imperfect* is inflected by modification of a related verbal stem through the addition of (i) prefixes indicating person and sometimes gender and (ii) suffixes indicating number and sometimes gender – thus, *prefix* + *stem* + *suffix*. The perfect stem for transitive-active verbs of the simple conjugation (Qal) is \**qātal*, while the imperfect stem is \**qtōl*; both of these change slightly when inflected (for the inflections, see §§4.5.4.1–2).

Like other Semitic languages, Hebrew verbs have a number of different stem patterns with a diversity of contrasting forms that signify semantic variations in relation to the basic meaning of the verbal root. These patterns (see §4.5.5) are conventionally called *conjugations*, and, more specifically, *derived* conjugations, since they are produced by the application of certain morphological and phonological changes to the simple stem, traditionally known as Qal (*qal* “light, easy, simple”) in Hebrew. Note that the term “conjugations” is retained here because of its conventional use in modern grammars, despite the lack of correspondence of the Hebrew verbal stems to the conjugations of the languages – principally Latin – from which the term derives; the term *binyānîm* “structures,” used by the medieval grammarians is more descriptive.

In addition to the indicatives, Hebrew has certain modal verb forms, including a command imperative as well as a cohortative and a jussive, both of which are primarily volitional in force (see §4.5.2). There are also a number of nonfinite verbal forms (see §4.5.3).

### 4.5.1 The aspects of the indicative verb

The perfect verb is *punctual* in aspect, while the imperfect is *durative*. In most cases, the perfect expresses a completed action, so that it may be translated with a verb in the simple past tense – thus, *kātābtī* “I wrote.” With verbs denoting dispositions or perceptions acquired in the past but still held or felt, a present-tense translation may be required – thus, *yādā’tī* “I know” (i.e. “I have come to know”); *bātāhtī* “I trust” (“I have come to trust”). With stative verbs, the best translation may employ a predicate adjective – thus, *zāqāntī* “I am old” (“I have grown old, aged”). The so-called *performative* perfect, employed in indirect speech and especially when the speaker is someone with authority, is used to indicate that the action expressed in the verb is accomplished by the very fact of its utterance – thus, *’āmārtī* “I say” (“proclaim, declare”). By contrast, the imperfect expresses an action that is incomplete and ongoing or still to be accomplished in the future, so that it may be translated with a verb in the present or future tense – thus, *’ektōb* might be rendered “I write,” “I will write,” or “I keep writing” (habitually or repeatedly). In Rabbinic Hebrew the aspectual character of the verbal system has weakened substantially, moving in the direction of a true tense system, with the perfect becoming predominantly a past-tense form and the imperfect taking on a modal character, while the principal burden of expressing the present and future tenses is assumed by the participle.

A verbal feature that is especially distinctive of Biblical Hebrew (though attested in early inscriptions in other Northwest Semitic languages) is the existence of the *converted* imperfect and perfect, which form the basic fabric of the narrative sequences in Biblical Hebrew (see §5.2.1). In these sequences converted imperfects, which are marked by a distinctive form of the conjunction (*wa-* + junctural doubling), have the punctual translational value of the perfect: thus, *watt’ō’mer sārāy ’el-’ābrām . . . wayyišma’ ’ābrām ləqōl sārāy wattiqah sārāy . . .* “and Sarai said to Abram . . . and Abram listened to the voice of Sarai, and Sarai took . . .” (Genesis 16:1–3). Converted perfects, which are also joined to the conjunction (in this case with its ordinary forms), have the durative translation value of the imperfect: for example, *w’ālā hā’iš* “and the man used to go up” (1 Samuel 1:3).

The converted imperfect exhibits a tendency, shared by the jussive (see §§3.5 and 4.5.2), to retract the tone from the final syllable of the verb (except in first-person forms), resulting in a shortening or collapse of the end of the word in certain forms found among the weak verbs (see §4.5.4.2) and the derived conjugations (see §4.5.5) – thus, indicative *yāqūm* “he arises”; jussive *yāqōm* “let him arise”; converted imperfect *wayyāqom* “and he arose.” There is a tendency in the converted perfect, operative in first- and second-person singular forms, to shift the tone forward to the ultima (without a corresponding change in vocalization) – thus, perfect *kātābtā* “you wrote,” converted perfect *wākātabtā* “and you will write.”

The origin of the converted verb forms can be explained with reference to distinctive developments that took place in early Hebrew in relationship to its antecedents. The indicative verbal system of Proto-Northwest Semitic had three forms: (i) *\*qatala*, a perfective, which expressed completed actions, usually in the past, but which (like its descendant, Hebrew *qātal*) also had a number of present-future uses; (ii) *\*yaqtulu*, an imperfective, which was used for habitual or durative actions but also served to express the present and future “tenses”; and (iii) *\*yaqtul*, a perfective, which functioned both as a jussive and as a

preterite, in other words, to express simple past actions (a past “tense”). With the loss of final short vowels, *\*yaqtulu* and the two types of *\*yaqtul* fell together as *\*yaqtul*, the antecedent of Hebrew *\*yiqṭōl* (see §3.6.2). This form became the ordinary Hebrew imperfect, retaining the present-future force of *\*yaqtulu*, but the jussive force of *\*yaqtul* was also preserved in *yiqṭōl*. The preterite force was lost, however, except in certain restricted environments, most characteristically the converted imperfect *wayyiqṭōl*. In most other situations the preterite role of *\*yaqtul* was appropriated by the perfect, *qātal* (< *\*qatal* < *\*qatala*). The converted perfect may have arisen by analogy with the converted imperfect, but it is unlikely that this would have happened were it not for the other present-future uses that *qātal* inherited from *\*qatala*.

Among the most important differences between Biblical Hebrew and Rabbinic Hebrew is the loss of the system of converted imperfects and perfects, completing a trend already observable in Late Biblical Hebrew.

#### 4.5.2 Command forms (the imperative and cohortative/jussive system)

In addition to the two indicatives, Hebrew has three principal modal forms, which are based on the imperfect and, when taken together, constitute a loose system expressing command and volition. The Hebrew imperative, which exists in the second person only, is formed by distinctive suffixes indicating gender and number attached to the imperfect stem without its prefixes. The imperative expresses direct command.

Both the cohortative and jussive express volition and resolve, though the jussive can also be described as an indirect command form, and, in combination with the adverbial particle *’al* (*’al* + *jussive*), it serves as the negative imperative. The cohortative (first person) and jussive (second and especially third person) are formed from the imperfect stem by the addition of distinctive prefixes expressing person and sometimes gender, and suffixes expressing number and sometimes gender. As noted above (§§3.5 and §4.5.1) there is a tendency in the jussive, observable in certain forms found among the weak verbs (see §4.5.4.2) and the derived conjugations (see §4.5.5), to retract the tone from the final syllable of the verb, resulting in a shortening or collapse of the end of the word in comparison to the indicative. The jussive-imperative-cohortative system for the simple stem (Qal) of the strong verbs is shown in (9) (the second-person jussive is not included).

(9)	Form	Gender	Number	
			Singular	Plural
Jussive		Masculine	yiktōb “let him write”	yiktābū “let them write”
		Feminine	tiktōb “let her write”	tiktōbnā “let them write”
Imperative		Masculine	kātōb “write”	kitbū “write”
		Feminine	kitbī “write”	kātōbnā “write”
Cohortative		Common	’ektābā “let me write”	niktābā “let us write”

In terms of their historical origin, the jussive and imperative are descended directly from the jussive and imperative of Proto-Northwest Semitic – thus, jussive *yiqṭōl* < *\*yaqtul* and imperative *qātōl* < *\*qutul* (the development of the former is described in §4.5.1). The cohortative is a partial survival of a volitional subjunctive: *’eqṭālā* < *\*aqtula*. Proto-Northwest Semitic also had an energetic with the form *\*yaqtulanna*, similar in force to the subjunctive and thus to the Hebrew jussive and cohortative. Relics of this form may survive in (i) the so-called *nūn energeticum*, a tone-bearing syllable with the form *-én-* (raised under stress

from \*-án-, which is sometimes preserved before the first-person singular suffix) that may be inserted before the pronominal suffixes of the imperfect (e.g., *wəʿšmərénna* “and I will keep it,” Psalm 119:34) and (ii) the -nā’ particle often used to strengthen cohortatives, jussives with optative force, and imperatives, especially in the rhetoric of courteous speech (e.g., *təḏabbēr-nā’ šiphātākā ’el-’ădōnî hammélek* “Let your maidservant speak to my lord the king,” 2 Samuel 14:12).

In Rabbinic Hebrew the special lengthened cohortative forms and shortened jussive forms disappear almost entirely (expanding a tendency already observable in the Samaritan Pentateuch), and the feminine plural imperative *kətóbnā* is lost, leaving *kitbû* as the common form. In general, the use of the imperative is much more restricted than in Biblical Hebrew.

### 4.5.3 Verbal nouns

Hebrew has two participles, active and passive. As noted in §4.2.5.2, the active participle of the simple verbal stem (Qal) has the form *qōtēl* (< \**qātil*), feminine *qōtālā* (< \**qātilat*) and *qōtélet* (< \**qātilt*) – thus, *kōtēb* (etc.) “writing.” The Qal passive participle is formed from active verbal roots using the form *qātûl* (< \**qatûl*) – thus, *kātûb* “written.”

As in certain other Semitic languages, such as Akkadian and Ugaritic, Hebrew forms an infinitive of the simple stem – the G-stem (*Grundstamm*) or Hebrew Qal – from the nominal pattern \**qatāl*. By normal phonological developments this infinitive, which is known as the *infinitive absolute*, has the form *qātōl* in Hebrew. In contrast to the situation in Akkadian (though in common with Ugaritic) the Hebrew reflex of this infinitive is not inflected, and it surrenders the ordinary infinitive functions to a second infinitive, known as the *infinitive construct*, which has the form *qəṭōl* (though the Qal infinitive construct has the form of the construct state of the Qal infinitive absolute (*qəṭōl* ~ *qātōl*), it does not function as its construct, and the terminology should not lead to confusion with the construct and absolute states of ordinary nouns). Thus, the infinitive construct is the true Hebrew infinitive, while the infinitive absolute is primarily adverbial in function, serving most characteristically to emphasize the verbal idea of the finite verb that it immediately precedes or follows: for example, *dārōš dāraš mōšeh* “Moses sought diligently” (Leviticus 10:16). Otherwise, the infinitive absolute is used to suggest the verbal idea in a general way, even occasionally serving as an uninflected substitute for a finite verb, in which case it derives its “inflection” from that of preceding verbs in a sequence: thus, *ūmāšā’tā ’et-ləbābō ne’ēmān ləpānēkā wəḱārōt ’immō habbārīt* “and you found his heart faithful before you and cut a covenant with him” (Nehemiah 9:8).

In Rabbinic Hebrew, the infinitive absolute is lost entirely, and the infinitive construct occurs almost exclusively with prefixed *lə-*.

### 4.5.4 Verb inflection

#### 4.5.4.1 The sound verb

The perfect and imperfect verbs of the simple stem (Qal) formed from sound roots are conjugated as shown in (10) and (11). Variations in these paradigms occur when one of the root consonants is a guttural, in accordance with the special phonological rules that obtain in the environment of gutturals (see §3.3.2):



## (10) The Qal perfect verb

Person	Gender	Number	
		Singular	Plural
Third	Masculine	kātab “he wrote”	kātəbû “they wrote”
	Feminine	kātəbā “she wrote”	kātəbû “they wrote”
Second	Masculine	kātābtā “you wrote”	kətabtem “you wrote”
	Feminine	kātabt “you wrote”	kətabten “you wrote”
First	Common	kātābtî “I wrote”	kātəbnû “we wrote”

## (11) The Qal imperfect verb

Person	Gender	Number	
		Singular	Plural
Third	Masculine	yiktōb “he writes”	yiktəbû “they write”
	Feminine	tiktōb “she writes”	tiktəbnā “they write”
Second	Masculine	tiktōb “you write”	tiktəbû “you write”
	Feminine	tiktəbî “you write”	tiktəbnā “you write”
First	Common	’ektōb “I write”	niktōb “we write”

Though it is always vocalized in the Masoretic Text as shown above, the ending of the second-person masculine singular perfect is most often written without a final *hē*’ in the Masoretic Text and Rabbinic Hebrew – thus, *ktbt* rather than *ktbth* (the usual Qumran form) – indicating a pronunciation *\*katabt*, which is also the more common form in the Hexapla (cf. the situation with the corresponding personal pronoun, §4.3.1). The second-person feminine singular perfect, though always vocalized as shown, is sometimes spelled with final *yōd*, indicating a pronunciation *\*katābtî* (cf., again, the corresponding personal pronoun, §4.3.1). In Rabbinic Hebrew, as part of the general tendency for final *-n* to replace final *-m* (see §3.1.2 and §4.2.2), the gender distinction in the second-person plural perfect is obscured, with *kətabten* becoming the common form. In Late Biblical Hebrew and Rabbinic Hebrew, the third- and second-person feminine imperfect forms coalesce with the corresponding masculine forms, *yiqṭəlû* and *tiktəbû*, and the older form, *tiqṭólnā*, is lost.

The paradigm verb used here (*kātab* “write”) belongs to the *a ~ u* vowel class, meaning that in its antecedent form the theme-vowel for the perfect was *\*a* (*\*kataba* → *kātab*) and the theme-vowel for the imperfect was *\*u* (*\*yaktub-* → *yiktōb*). As in other Semitic languages, however, Hebrew verbs are distributed among several vowel classes, which correspond generally to their semantic character. The principal theme-vowel patterns in Hebrew are listed in Table 10.4 (the paradigm verbs used are *kātab* “write,” *nātan* “give,” *šakab* “lie down,” *qārab* “draw near,” *zāqēn* “grow old,” and *qāṭōn* “be small”).

## 4.5.4.2 The weak verbs

The inflection of the Hebrew verb is modified under certain conditions: (i) when the second and third root consonants are identical (“geminate” verbs); (ii) when the initial root consonant is *n-* (I-*n*); (iii) when one of the root consonants is a guttural (I-, II-, or III-G); or (iv) when one of the original root consonants was a glide, *\*w* or *\*y* (*\*I-*, *\*II-* or *\*III-w*; I-, *\*II-* or *\*III-y*). The following synopsis enumerates the most important changes that occur during the inflection of these *weak verbs*, as they are customarily called.

The distinctive feature of the perfect of geminate verbs is the interposition of *-ō-* before verbal suffixes beginning with a consonant – thus, *sābəbā* “she went around,” but *sabbōtā* “you went around.” This feature is Proto-Semitic in origin. Though the imperfect displays



**Table 10.4 The vowel classes of Hebrew verbs**

Theme vowels	Aspect	Hebrew form	Antecedent form	Description
<i>a, u</i>	PERFECT	<i>kātab</i>	* <i>kataba</i>	A large class of primarily active-transitive verbs. III-guttural tended to become ( <i>a, a</i> <sup>1</sup> ).
	IMPERFECT	<i>yiktōb</i>	* <i>yaktub-</i>	
<i>a, i</i>	PERFECT	<i>nātan</i>	* <i>natana</i>	A large class with no semantic restrictions, but lost in Hebrew except for a few verbs.
	IMPERFECT	<i>yittēn</i>	* <i>yantin-</i>	
<i>a, a</i> <sup>1</sup>	PERFECT	<i>šākab</i>	* <i>šakaba</i>	An active-intransitive class, which falls together formally with the stative ( <i>a, a</i> ) class.
	IMPERFECT	<i>yīškab</i>	* <i>yīškab-</i>	
<i>a, a</i> <sup>2</sup>	PERFECT	<i>qārab</i>	* <i>qaraba</i>	A small stative class, enlarged by original ( <i>i, a</i> ) and ( <i>u, a</i> ) verbs with guttural roots.
	IMPERFECT	<i>yīqrab</i>	* <i>yīqrab-</i>	
<i>i, a</i>	PERFECT	<i>zāqēn</i>	* <i>zaqina</i>	A large, primarily stative-intransitive class. Many II- and III-gutturals became ( <i>a, a</i> ).
	IMPERFECT	<i>yizqan</i>	* <i>yizqan-</i>	
<i>u, a</i>	PERFECT	<i>qāṭōn</i>	* <i>qaṭuna</i>	A small stative class, originally *( <i>u, u</i> ), but transformed by resistance to stative * <i>yaqtul</i> .
	IMPERFECT	<i>yīqṭan</i>	* <i>yīqṭan-</i>	

wide variation, the basic forms are predictable from normal phonological changes – thus, *yāsōb* (< \**yasubbu*) “he goes around.”

I-*n* verbs are inflected normally in the perfect and in the imperfect indicative, except that in the latter case the expected assimilation of *n-* to the second radical occurs – thus, *yiddōr* (< \**yaddur* < \**yandur-*) “he vows” (*a, u*); *yittēn* (< \**yittin* < \**yantin-*) “he gives” (*a, i*); and *yiggaš* (< \**yiggaš* < \**yangaš-*) “he draws near” (*a, a*). In the (*a, a*) type, the imperative is usually shortened (*gaš*), and the “normal” form of the infinitive construct alternates with a short form with -*t* (*géšet*; see §4.2.5.1). The imperative and infinitive construct corresponding to *yittēn* are *tēn* and *tēt*. Perhaps because it is the antonym of *nātan* “give,” the common verb *lāqah* “take” has come to be inflected as if it were I-*n* in its imperfect and related forms – thus, *yīqqaḥ* (imperfect), *qaḥ* (imperative) and *qāḥat* (infinitive construct).

The perfect of I-G verbs presents no special problems, with the *ḥāṭēp*-vowel *ā* replacing simple *šəwā* (ə) as necessary (see §3.3.2) – thus, *‘amadtem* “you stood.” The imperfect appears in two forms according to the vowel classes of the verbal stems – thus, *ya‘āmōd* “he stands” (*a, u*) and *yehēzaq* “he is strong” (*a, a*). As noted in §3.6.2, the change of the imperfect prefix \**ya-* → *yī-* took place first in verbs with *a* as the imperfect theme-vowel (\**yaqtal* → *yīqtal*) and was subsequently extended to the other verbs. These two I-G forms reflect the intermediate stage – thus, *ya‘āmōd* < \**ya‘mud*, but *yehēzaq* < \**yihzaq* < \**yahzaq*. Many I-’ verbs generally follow the pattern of other I-G verbs, but with *ē* in imperfect prefixes for (*a, u*) as well as (*a, a*) stems – as in *ye‘ēsōp* “he gathers.” In some I-’ verbs, however, the /ʔ/ quiesced at an early date in postvocalic positions, leading to the lengthening of the prefix-vowel and the development of forms like *yō(‘)mar* “he says.”

Despite a few peculiarities, verbs II-G and III-G present no major divergences from the strong verb paradigm. In III-’ verbs the quiescence of word- or syllable-final /ʔ/ has led to the lengthening of the preceding *a* to *ā* (but not *ō*, as explained in §3.6.1) in perfect forms like *bārā(‘)* “he created” and *bārā(‘)tī* “I created.” Similarly in III-’ imperfects, the stem vowel, which is *a* as usual in gutturals, is lengthened after the loss of /ʔ/ – thus, *yibrā(‘)* “he creates.”

Most verbs I-*y* were originally \*I-*w*. As noted in §4.1, some of these, such as *yšb* “sit” (*a, i*), have very ancient root allomorphs, with and without *w-* – thus, *√\*wθb* and *√\*θb*, leading

to a mixture of forms like perfect *yāšab* (< \**waθaba*) “he sat,” imperfect *yēšēb* (< \**yiθib-*) “he sits,” and imperative *šēb* (< \**θib*) “sit.” In general, however, I-*y* verbs are regular in their inflection. As in I-*n* verbs of the (*a*, *a*) type, there is usually shortening of the imperative – thus, *rēd* “go down”; *šē* “go forth”; *dā* “know” – and the infinitive construct, which is augmented with *-t* – thus, *rédet* “to go down”; *šē(ʾ)t* “to go forth”; *dáʾat* “to know” (see §4.2.5.1).

When inflected, verbs II-*w/y*, the so-called hollow roots, behave as if biradical. In the perfect the inflectional endings are added to a biconsonantal stem – thus, *qām* “he arose,” *qāmā* “she arose,” *qāmtā* “you arose,” and so forth. In the imperfect, the distinction between verbs II-*w* and II-*y* becomes evident – thus, *yāqūm* “he arises,” but *yāšīm* “he places.” The jussive forms of these verbs are distinctive – *yāqōm* and *yāšēm* – and the converted imperfect employs the same forms, with retraction of the stress – *wayyāqom* and *wayyāšem*. The imperfect–jussive contrast is probably a survival of the Proto-Northwest Semitic situation (see §4.3.1 and §4.3.2), later vowel length being determined by whether the syllable was open or closed – thus, imperfect \**yaqūmu* → *yāqūm*, but jussive \**yaqum* → *yāqōm*. In later periods the hollow verbs tend to assimilate to triradical patterning, giving rise to forms like Late Biblical Hebrew Piʿel *qiyyam* “it established” (Esther 9:32) and Rabbinic Hebrew Piʿel *qiyyēm*.

III-*w/y* verbs are inflected according to a single paradigm regardless of the original final consonant (\**w* or \**y*) or vowel class. Thus, for example, the III-*y* (*a*, *i*) verb *bānā* “build” – *bānā* (< \**banaya*) “he built,” and *yibneh* (< \**yabniyu*) “he builds” – has the same Hebrew forms as the III-*w* (*i*, *a*) verb *ḥāyā* “live” – *ḥāyā* (< \**ḥayiwa*) “he lived,” and *yihyeh* (< \**yihyawu*) “he lives.” The jussive (and converted imperfect) form is apocopated with retracted stress and (variable) anaptyxis – thus, *yiben* (< \**yibn* < \**yabni* < \**yabniy*) “let him build”; and *tīreš* (< \**tirš* < \**tirša* < \**tarḏaw*) “let her be pleased”; but *yēbk* (< \**yibk* < \**yabki* < \**yabkiy*) “let him weep”; and *yēšt* (< \**yišt* < \**yišta* < \**yištay*) “let him drink.”

#### 4.5.5 The derived conjugations

As noted in §4.5, there are several stem patterns, known as “derived conjugations” or *binyānim*, by which semantic variety is derived from verbal roots. The most common *binyānim*, which are traditionally named for the corresponding third-person masculine singular perfect form of the verbal root √*pʿl*, are called Nipʿal, Piʿel, Puʿal, Hipʿil, Hopʿal and Hitpaʿel. Few, if any, Hebrew verbs are attested in all of these forms. In addition to these six, there is a special set used for II-*w/y* verbs, and a small additional group that occur relatively seldom. A synopsis of the forms of the derived conjugations in relation to the Qal verb is given in (12):

##### (12) Synopsis of the basic conjugations

	<i>Perfect</i>	<i>Imperfect</i>	<i>Imperative</i>	<i>Infinitive absolute</i>	<i>Infinitive construct</i>	<i>Participle</i>
<i>Qal</i>	<i>qātal</i>	<i>yīqtōl</i>	<i>qəṭōl</i>	<i>qātōl</i>	<i>qəṭōl</i>	<i>qōṭēl</i>
<i>Nipʿal</i>	<i>niqtal</i>	<i>yīqqātēl</i>	<i>hiqqātēl</i>	<i>niqtōl</i>	<i>hiqqātēl</i>	<i>niqtāl</i>
<i>Piʿel</i>	<i>qittēl</i>	<i>yəqattēl</i>	<i>qattēl</i>	<i>qattōl</i>	<i>qattēl</i>	<i>məqattēl</i>
<i>Puʿal</i>	<i>quttal</i>	<i>yəquttal</i>	—	<i>quttōl</i>	—	<i>məquttāl</i>
<i>Hipʿil</i>	<i>hiqtīl</i>	<i>yaqtīl</i>	<i>haqtēl</i>	<i>haqtēl</i>	<i>haqtīl</i>	<i>maqtīl</i>
<i>Hopʿal</i>	<i>hoqtal</i>	<i>yoqtal</i>	—	<i>hoqtēl</i>	—	<i>moqtāl</i>
<i>Hitpaʿel</i>	<i>hitqattēl</i>	<i>yitqattēl</i>	<i>hitqattēl</i>	<i>hitqattēl</i>	<i>hitqattēl</i>	<i>mitqattēl</i>

#### 4.5.5.1 *Nip'al*

The *Nip'al* is formed by the prefixation of *n-* to the verbal stem – thus perfect *niqtal* (< \**naqtala*) and imperfect *yiqqātēl* (< \**yanqatil-*). In addition to *niqtōl* (see [12]), which in light of the comparative Semitic evidence is probably the original form of the infinitive absolute, two other forms occur, *hiqqātōl*, the *ō* of which may have arisen by analogy with Pi'el *qattōl*, and *hiqqātēl*, which is identical to the form of the infinitive construct.

The meaning of the *Nip'al* is mediopassive. In origin it may have served to give intransitive-stative force to transitive-active verbs in Qal, to which it remains close inflectionally, and this early meaning is preserved in its frequently fientic character – as in Qal *rā'ā* “he saw” ~ *Nip'al nir'ā* “he appeared” (i.e., “he became visible”). With the loss or obscuration of the Qal passive, however, the *Nip'al* absorbed the role of the primary passive correspondent of Qal – thus, Qal *'āsar* “he bound, imprisoned” ~ *Nip'al (imperfect) yē'āsēr* “he will be bound, imprisoned” (Genesis 42:19). With other transitive Qal verbs, the voice of the corresponding *Nip'al* may be middle rather than passive: for example, Qal *'āsap* “he gathered” (transitive) ~ *Nip'al (plural) ne'ēspū* “they gathered”; compare *ūpəlišīm ne'ēspū ləhillāhēm 'im-yisrā'el* “And the Philistines gathered to fight with Israel” (1 Samuel 13:5). Finally, the *Nip'al* sometimes has reflexive force – thus, Qal *mākar* “he sold” ~ *Nip'al nimkar* “he sold himself.”

#### 4.5.5.2 *Pi'el*

The Pi'el is formed by doubling of the second radical – *qittēl* (< \**qattila* or \**qattala*), *yəqattēl* (< \**yuqattil-*). Predictable phonological changes occur when the second radical cannot be doubled because it is a guttural (see §3.3.2), and there is a special conjugational system for verbs II-*w/y* (see §4.5.5.7).

The basic and original meaning of the Pi'el is factitive (transitivizing), as applied to verbs that are intransitive or stative in the Qal – thus, Qal *hāzaq* “be strong” ~ Pi'el *hizzaq* “strengthen, fortify.” With active-transitive verbs, the Pi'el may pluralize the Qal meaning, so that the effect is intensive or iterative – thus, Qal *nātaq* “tear away, pull off” ~ Pi'el *nittēq* “tear apart, rip out”; Qal *šābar* “break” ~ Pi'el *šibbar* “shatter.” For many verbs that occur in both Qal and Pi'el, however, the difference in meaning is subtle or unclear, though the lexicons tend to try to specify an intensifying nuance for the Pi'el. With certain active-transitive verbs, the Pi'el seems to be the causative of the Qal: for example, Qal *lāmad* “learn” ~ Pi'el *limmad* “cause to learn, teach.” This is the role of the Hip'il with active-transitive verbs, however, and most such Pi'els may in fact be denominative. In any case, the Pi'el is especially productive of denominatives: thus, *qinnē* “be jealous” (from *qin'ā* “jealousy”); *'ippēr* “cast dust on” (from *'āpār* “dust”).

#### 4.5.5.3 *Pu'al*

The Pu'al, like the Pi'el, is formed by doubling of the second radical, but it is distinguished from the Pi'el by its *u-a* vowel patterning, which persists throughout the paradigm – thus, *quttal* (< \**quttala*), *yəquttal* (< \**yuquttal-*), and so forth. When the second radical cannot be doubled, the changes that occur are the same as those for the Pi'el (see §4.5.5.2).

The Pu'al functions as the passive of the Pi'el. It is used relatively infrequently except in its participial form, which serves as the passive participle of the Pi'el: for example, Pi'el infinitive construct *qaddēš* “to consecrate” ~ Pu'al participle *məquddāš* “consecrated.” In Rabbinic Hebrew the Pu'al survives only as a participle.

#### 4.5.5.4 *Hitpa'el*

The *Hitpa'el* is distinguished by prefixed *t-* and, like the *Pi'el* and *Pu'al*, the doubling of the second radical. In contrast to the other conjugations, however, the *Hitpa'el* seems to follow the pattern of the imperfect in the inflection of the perfect, imperative, and infinitives. In these same forms the preformative takes the shape *hit-*, the *hi-* possibly having arisen under the influence of the *Hip'il*. When the first root consonant is one of the dental stops (see §3.1.1), the prefixed *t-* is assimilated – as in *yittāmē* (< \**yittāmē*) “he defiles himself.” When the first root consonant is a sibilant (see §3.1.1), the *t-* metathesizes with it for the sake of euphony – *yistakkəḥū* (< \**yitšakkəḥū*) “they were forgotten” (Ecclesiastes 8:10). When the second radical cannot be doubled, the changes that occur are the same as those for the *Pi'el* (see §4.5.5.2).

The *Hitpa'el* is intransitive in meaning. Most characteristically, it gives reflexive (or reciprocal) force to an active form of same verb – thus, *Pi'el qiddēš* “consecrate” ~ *Hitpa'el hitqaddēš* “sanctify oneself.” In addition, it is often iterative – as in *Qal hālak* “walk” ~ *Hitpa'el hithallēk* “walk back and forth” – and sometimes denominative – *hitnabbē* “prophesy” (from *nābī* “prophet”).

Though the *Hitpa'el* is morphologically related to the *Pi'el* and *Pu'al* by the common feature of the doubled second radical, the active verbs to which it corresponds are not always *Pi'el* but may be *Qal* or *Hip'il* as well. This points to the likely historical background of the *Hitpa'el* as a composite conjugation produced by the merger of the prefixed *t-* forms of verbal roots of the simple, factitive, and causative stems. Remnants of an original *t-* form of the simple stem are recognizable in a few *Hitpa'els* that lack doubling of the second radical: for example, *hitpāqədū* “they mustered” (Judges 20:17)

In Rabbinic Hebrew the *Hitpa'el* was largely replaced, at least in the perfect, by the *Nitpa'el* (properly *Nitpa'al*), a new conjugation created by fusion of the *Hitpa'el* with the *Nip'al*, which could also have reflexive meaning (see §4.5.5.1).

#### 4.5.5.5 *Hip'il*

The formal marker of the *Hip'il*, found on the perfect, imperative, and infinitives, is prefixed *h-*. The long stem-vowel *-î-* is characteristic of both the perfect and imperfect, but the jussive has the expected *-ē-* (*yaqtēl* < \**yaqtīl* < \**yuhaqtīl*), and the *-î-* must have arisen by analogy with the *Hip'il* of verbs II-*w/y* – compare jussive *yāsēr* (< \**yasir* < \**yuhasir*) “let him remove,” to imperfect *yāsīr* (< \**yasir* < \**yuhasir-*) “he removes.”

In general the *Hip'il* serves as the causative of the *Qal*. With intransitive or stative verbs it is singly causative: for example, *Qal lābēš* “be dressed” ~ *Hip'il hilbiš* “cause to be dressed, clothe.” This is especially characteristic of verbs of motion – *Qal hālak* “go, walk” ~ *Hip'il hōlik* “bring, lead.” When the *Qal* is transitive, the *Hip'il* may be doubly causative: for example, *Qal yāda* “know” ~ *Hip'il hōdīa* “cause (someone) to know (something)” (cf. *wənōdī'ā 'etkem dābār* “and we will apprise you of something” [1 Samuel 14:12]); *Qal rā'ā* “see” ~ *Hip'il her'ā* “cause (someone) to see (something)” (cf. *wayyar'em 'et-məbō' hā'ir* “and they showed them the entrance to the city” [Judges 1:25]). Sometimes, especially when the *Qal* is stative, the *Hip'il* may be fientic or otherwise intransitive, even in verbs that also have causative *Hip'ils*: *Qal 'ārēk* “be long” ~ *Hip'il he'ērik* “become long,” but also “make long, prolong.” Many of these *Hip'ils* are inchoative or inceptive – such as *Qal bā'aš* “stink”; *Hip'il hib'īš* “begin to stink, become stinking,” but also “cause to stink.” Like the *Pi'el*, though less characteristically so, the *Hip'il* may form denominatives: for example, *he'ēzīn* “listen” (from *'ōzen* “ear”).

#### 4.5.5.6 *Hop'al*

Like the Hip'il, its active counterpart, the Hop'al is characterized by *h-* prefixed to the perfect. In contrast to the *hi-* preformative of the Hip'il, however, the Hop'al has the variants *ho-* and *hu-*; in Rabbinic Hebrew the option has been resolved in favor of the latter (often written plene, i.e., הוּ), probably by analogy with the Pu'al (Pi'el : Pu'al :: Hip'il : Hup'al).

Semantically, the Hop'al is the passive of the Hip'il – thus, Hip'il *hišlik* “he threw” ~ Hop'al *hošlak* “he was thrown.”

#### 4.5.5.7 *Polel, Polal, and Hitpolel*

Because most “hollow” verbs (II-*w/y*) are inflected as if they were biconsonantal (see §4.5.4.2 and note exceptions in Rabbinic Hebrew), they do not accept doubling of the second radical, the chief marker of the factitive conjugation group, Pi'el, Pu'al, and Hitpa'el. In hollow verbs like √*qwm* “rise up,” therefore, the functions of these conjugations are taken over by a group consisting of the Polel (active), Polal (passive), and Hitpolel (reflexive) conjugations. These are characterized formally by reduplication of the final stem consonant and *ō* in the first stem syllable – thus, Polel perfect *qômēm* “he raised up,” and imperfect *yəqômēm* “he raises up”; Polal *qômam* “he was raised up,” and *yəqômam* “he is raised up”; and Hitpolel *hitqômēm* “he raised himself up,” and *yitqômēm* “he raises himself up.” Geminate (§4.5.4.2) employ these forms on occasion, too, even in verbs for which the Pi'el group is also attested: for example, Pi'el imperfect *yəḥannēn qôlô* “he speaks favorably” (i.e., “makes his voice favorable”; Proverbs 26:25) ~ Polel imperfect *yəḥōnēnû* “they will treat favorably” (Psalm 105:15).

#### 4.5.5.8 *Other conjugations*

There are several other *binyānīm*, some very sparsely attested. Some of the more important and better understood are listed here.

The series Po'el (active), Po'al (passive), and Hitpo'el (reflexive) is similar to the Polel group (see §4.5.5.7), except that it forms verbs from sound roots – thus, Hitpo'el *yitgō'āšû māyim* “the waters surge” (√*gš* “shake”) in Jeremiah 46:8, a duplicate of the preceding line with Hitpa'el *yitgā'āšû* “[its waters] surge”; also, *šōrēš* “he took root” (Isaiah 40:24), a Po'el denominative from *šōreš* “root” (contrast the meaning of the Pi'el denominative *wəšēreškā* “and he will uproot you,” Psalm 52:7).

The series Pilpel (active), Polpal (passive), and Hitpalpel (reflexive) is characterized by reduplication of the two strong consonants of geminate and “hollow” verbs (II-*w/y*). Like the Polel and Po'el groups, they correspond in meaning to the factitive (Pi'el) group – thus, *gilgēl* “roll” (√*gll* “roll”); *kilkēl* “maintain” (√*kwl* “hold”).

The Pa'lal (active) and Pu'lal (passive) are quadriliterals formed by the reduplication of the third radical. Their meaning in either voice is stative – thus, *ša'ānan* “he has been at ease” (Jeremiah 48:11); *'umlal* “it is withered” (Joel 1:10).

### 4.6 Numerals

The Hebrew cardinals 1–10 are listed in (13).

(13)	Modifying a masculine noun		Modifying a feminine noun	
	Absolute	Construct	Absolute	Construct
1	'eḥad	'aḥad	'aḥat	'aḥat
2	šənáyim	šənē	štáyim	štē
3	šəlōšā	šəlōšet	šālōš	šəlōš
4	'arbā'ā	'arbā'at	'arba'	'arba'
5	ḥāmiššā	ḥāmēšet	ḥāmēš	ḥāmēš
6	šiššā	šēšet	šēš	šēš
7	šib'ā	šib'at	šéba'	šəba'
8	šəmənā	šəmənat	šəməneh	šəməneh
9	tiš'ā	tiš'at	tēša'	təša'
10	'āsārā	'áséret	'éser	'éser

The cardinals may be associated with the nouns they modify in one of two ways: (i) appositionally, using the absolute form; or (ii) genitively, using the construct form. The first two cardinals agree with the modified noun (the counted item) in gender. In the case of the cardinals 3–10, however, the form that is usually feminine elsewhere – that is, the form marked with *-ā* (bound form *-at*) or *-t* (see §4.2.1) – modifies masculine nouns, while the unmarked form modifies feminine nouns, a peculiarity shared with most other Semitic languages (cf. Ch. 6, §3.3.7). The 'teens are formed by placing the unit, which follows the gender rules stated above, before the word for ten (with special forms): for example, *šəlōšā 'āsār pārim* "thirteen bulls" (Numbers 29:14).

The cardinal 20 is expressed by the plural of 10 (*'érim*), and the other tens by the corresponding plurals of the units – thus, *šəlōšim* "30," *'arbā'im* "40," *ḥāmiššim* "50," *šiššim* "60," *šib'im* "70," *šəmənim* "80," and *tiš'im* "90." Note that the tens are not inflected for gender and occur only in the absolute state. The numbers 21 to 99 are formed by placing the unit, which follows the gender rules stated above, before or after the ten – thus, *šəlōšā wə'ešrim* 'iš or *'érim ūš(ə)lōšā* 'iš "23 men." The higher numbers include the following substantives: *mē'ā* (bound form *mē'at*) "(one) hundred"; *mā(')táyim* "200"; *šəlōš mē'ot* "300"; *'elep* "(one) thousand"; *rəbābā* "10,000."

The ordinal "first" is expressed by the adjective *rī(')šōn* (fem. *rī(')šōnā*). The ordinals from "second" to "tenth" are formed by adding the sufformatives *-ī* (masc.) and *-it* (fem.) to the cardinal (cf. §4.2.5.4), following the general pattern *\*qətilī* – thus, *šēnī* "second," *šəliši* "third," *rəbi'i* "fourth" (without the prothetic *'a-* of *'arba'*, "4"), *ḥāmiši* "fifth," *šišši* "sixth," *šəbi'i* "seventh," *šəmini* "eighth," *təši'i* "ninth" and *'āširi* "tenth."

## 5. SYNTAX

### 5.1 Word order

The usual word order in the Hebrew verbal clause is Verb–Subject–Object (VSO) followed by prepositional phrases or other adverbial elements – thus:

- (14) wayiṭṭa'                      yhwh        'ēlōhīm    gan-bə'ēden  
          and-he planted    Yahweh    God        garden-in-Eden  
          "And Yahweh-God planted a garden in Eden" (Genesis 2:8)

Although this generalization applies to subordinate as well as independent verbal clauses, exceptions are quite common, especially when some kind of emphasis is placed on the subject (→ SVO), for example,

- (15) hannāḥš      hiššī'ānī  
the-serpent    deceived-me  
"The serpent deceived me" (Genesis 3:13)

or on the object (→ OVS or VOS), as in:

- (16) 'et-qōləkā              šāmá'tī    baggān,  
DIR. OBJ.-voice-your    I heard    in-the-garden  
"I heard your voice in the garden" (Genesis 3:10)

As the preceding example shows, a pronominal subject, since it is inherent in the verb, is not usually expressed, except, again, for emphasis:

- (17) hī'      nātənâ-lī      min-hā'ēš  
she    she gave-me    from-the-tree  
"She gave me [fruit] from the tree" (Genesis 3:12)

As a rule, finite, indicative verbs are negated by *lō'*, while modal (cohortative or jussive) verbs are negated by *'al*. Regularly in prose and sometimes in poetry, the direct object is marked by the accusative particle *'et* (most often proclitic *'et-*), which precedes the accusative word or pronominal suffix (with the form *'ôti*, etc., but *'etkem* and *'ethen*). An indirect object, marked by the preposition *lə-*, normally follows the direct object, though this order is usually reversed when the indirect object is a pronoun and the object a noun.

In verbless clauses, in which the subject is nominal (a noun or pronoun) and the predicate is nominal, adjectival, or adverbial, the order, as a general rule, is subject–predicate in clauses identifying the subject (18A) but predicate–subject in clauses classifying the subject (18B):

- (18) A. šēm-              hannāhār    haššēnī    gîhôn  
the name of    the river    second    Gihon  
"The name of the second river was Gihon" (Genesis 2:13)  
B. 'ārūrâ    hā'ādāmâ    ba'ābûrēkā  
cursed    the soil    because of you  
"The soil is cursed because of you" (Genesis 3:17)

These rules operate fairly consistently in independent verbless clauses, whether they are declarative or interrogative, but less predictably if the clause is volitional; the word order of subordinate verbless clauses is not as consistent. The far demonstrative or third-person personal pronouns (see §4.3.2) are often used pleonastically to coordinate the two parts of a verbless clause – thus:

- (19) hannāhār    hārēbī'î    hū'      pōrāt  
the river    fourth    COPULA    Euphrates  
"The [name of the] fourth river was Euphrates" (Genesis 2:14)

## 5.2 Coordination and subordination

Like other Semitic languages, Hebrew exhibits a strong preference for paratactic constructions (coordination) over hypotactic constructions (subordination). Thus, in Hebrew prose narrative the great majority of clauses are joined with the conjunction *wə-*. This is true of coordinate clauses whether the relationship between the clauses being coordinated is one of conjunction or disjunction. Though subordinating conjunctions do exist, *wə-* is most often used even in the case of subordinate clauses, with subordination being signaled by word order and clause formation.



### 5.2.1 Conjunctive clauses

Conjunctive clauses describing sequential events most often employ the distinctive Hebrew narrative sequences, which are made up of clauses containing the so-called converted imperfect and perfect verbal forms (see §4.5.1). The converted imperfect, which is used for past narration, occurs in a sequence that typically begins with a clause containing a perfect verb followed by from one to several clauses introduced by converted imperfects, each of which requires a perfective (usually punctual) translation, as in:

- (20) *wəhannāḥāš      hāyâ    'ārûm      mikkol      ḥayyat      ḥāššādeh . . .*  
 and the serpent    was    shrewd    more than any    living thing of    the field  
*wayyô'mer    el-    hā'issâ . . .    wattô'mer    hā'issâ    el-    hannāḥāš . . .*  
 and it said    to    the woman    and said    the woman    to    the serpent  
 "Now the serpent was shrewder than any of the other wild animals . . . and it said  
 to the woman . . . and the woman said to the serpent . . ." (Genesis 3:1–2)

The converted perfect, which is used for present-future narration, operates in a reciprocal manner. It occurs in a sequence typically beginning with a clause containing an imperfect verb followed by from one to several clauses introduced by converted perfects, each of which requires an imperfective (present, future, or habitual-iterative) translation, for example:

- (21) *'al-kēn    ya'azob-    'iš    'et-'ābiw    wə'tet-'immô    wədābaq    bə'ištô*  
 therefore abandons a man his father and his mother and unites with his wife  
*wəḥayû      ləbāšār    'eḥād*  
 and they become    flesh    one  
 "Therefore a man abandons his father and mother and unites with his wife, and  
 they become one flesh" (Genesis 2:24)

In sequences belonging to either of these categories, the introductory verbal clause may be replaced by any of a variety of other clause types or, owing to the ubiquity of such sequences, it may be omitted altogether.

### 5.2.2 Disjunctive clauses

Disjunctive clauses are also coordinated most often with *wə-*, but they differ from conjunctive clauses in that they begin with a nonverbal element. These include (i) simple negative clauses, which typically begin with *lō'*,

- (22) *wayyihyû šənêhem 'ārûmîm . . . wəlō' yitbōšāšû*  
 "And the two of them were naked . . . but they were not ashamed" (Genesis 2:25)

(ii) contrastive clauses,

- (23) *mikkol 'eš-    haggān    'ākōl tō'kēl    ûmē'eš    haddā'at*  
 from any tree of the garden you may eat but from the tree of the knowledge of  
*tōb    wārā'    lō' tō'kal    mimmēnnû*  
 good and evil you may not eat from it  
 "From any of the trees of the garden you may eat, but from the tree of the  
 knowledge of good and evil you may not eat!" (Genesis 2:16–17)

as well as various kinds of (iii) explanatory and circumstantial clauses, which may be nominal or verbal. Note, for example, the three circumstantial clauses embedded in the following narrative sequence:

- (24) bəyôm ‘āsôt yhwh ’ēlōhîm” ’ereṣ wəšāmāyim wəkōl śiaḥ  
on the day of making Yahweh-God’s earth and sky and any shrub of  
haśśādeḥ ṭerem yiḥyeh bā’areṣ wəkōl- ‘ēseb  
the field not yet was on the earth and any herb of  
haśśādeḥ ṭerem yišmāḥ... wə’ādām ’āyin la’ābōd  
the field not yet had sprouted and a man there was not to till  
’et-hā’ādāmā... wayyīṣer yhwh ’ēlōhîm ’et-hā’ādām  
the soil and formed Yahweh-God man  
“When Yahweh-God made the earth and the sky, no wild shrub was yet on the  
earth, and no wild herb had yet sprouted... and there was no man to till the  
soil... and Yahweh-God formed man” (Genesis 2:4–7)

### 5.2.3 Subordinate clauses

Although clause subordination may also be expressed by word order and clause formation in clauses joined with *wə-*, there are, as noted, special subordinating conjunctions as well as a number of special constructions indicating subordination. Three of the most important types of subordinate clauses are discussed below.

#### 5.2.3.1 Conditional clauses

Conditional clauses may begin with the conjunction *’im*, *hēn* or *kî*:

- (25) kî ta’ābōd ’et-hā’ādāmā lō’-tōsēp tēt- kōḥāḥ lāk  
though you till the soil it will not again yield its strength to you  
“Though you till the soil, it will not yield its strength to you again” (Genesis 4:12)

When conditional clauses lack one of the subordinating conjunctions and are joined to the preceding clauses by *wə-*, they are often susceptible to either conditional or nonconditional translation, as in the following:

- (26) wēhāyā kol-mōṣē’î yahargēnî  
“If anyone finds me, he will kill me”  
or  
“And whoever finds me will kill me” (Genesis 4:14)

#### 5.2.3.2 Temporal clauses

Though temporal clauses often stand in simple coordination after the clause they modify –

- (27) wayyīqeṣ nōaḥ miyyênô wayyēda’ ’ēt ’āšer- ‘āsā-lô  
and awoke Noah from his wine and he realized that which had done to him  
bənô haqqāṭān  
his son young  
“When Noah awoke from his wine, he realized what his youngest son had done  
to him” (Genesis 9:24)

they are very frequently placed before the modified clause and introduced by a converted form of the verb “to be”:

- (28) wəhāyâ kî- yir'û 'ōtāk hammiṣrīm wə'amərû 'ištô  
and it will be that will see you the Egyptians and they will say his wife  
zō't  
this

“When the Egyptians see you, they will say, this is his wife” (Genesis 12:12)

This construction is also used routinely for temporal phrases, such as the following:

- (29) wayhî miqqēṣ 'arbā'im yôm wayyiptaḥ nōaḥ 'et-ḥallôn  
and it was at the end of forty days and opened Noah the window of  
hattēbāh  
the ark

“At the end of forty days Noah opened the window of the ark” (Genesis 8:6)

### 5.2.3.3 Relative clauses

Relative clauses, which are usually introduced by 'āšer (see §4.3.3), follow and further define nouns or their equivalent:

- (30) hā'āreṣ 'āšer 'ar'ekā  
the-land which I will show-you  
“The land that I will show you” (Genesis 12:1)

They may contain resumptive (retrospective) pronominal or adverbial elements. Although 'āšer itself is indeclinable, the resumptive pronouns in a relative clause are declined in agreement with the noun modified by the clause:

- (31) A. wə'ēṣ 'ōseh pərî 'āšer zar'ô- bô  
and trees making fruit which their seed in it  
“And trees making fruit in which is their seed” (Genesis 1:13)  
B. ūmin-habbəhēmā 'āšer lō' ṭəhōrā hî  
“And from the animal which is not pure” (Genesis 7:2)

Resumptive adverbials include especially šām “there,” and related forms:

- (32) hā'ādāmā 'āšer luqqaḥ miššām  
the soil which he was taken from there  
“The soil from which he was taken” (Genesis 3:23)

The so-called *independent* relative clauses are not true relatives. Rather than further define a governing substantive, they serve as one of the elements in a larger clause, as in the following.

- (33) wayiššā'er 'ak- nōaḥ wa'āšer 'ittô battēbā  
and was left only Noah and those who with him on the ark  
“Only Noah and those that were with him in the ark were left” (Genesis 7:23)

## 5.3 Agreement

In general, a predicate agrees with its subject in gender and number, and if the predicate is a verb, it agrees with its subject in gender, number, and person. There are, however, numerous exceptions to this general pattern. A collective subject, for example, is often construed with a plural verb. When the subject is a construct chain (see §4.2.4), the predicate may agree in number and gender with the *nomen rectum* rather than the *nomen regens*, which is properly the subject.

A verb preceding a compound subject, though often plural, may be singular, agreeing with the first member in the series:

- (34) wayyithabbē'            hā'ādām wə'istō  
and he hid himself the man and his wife  
“And the man and his wife hid themselves” (Genesis 3:8)

Perhaps in extension of the last category, a verb in the initial position is sometimes masculine singular regardless of the gender and number of the subject, so that the masculine singular performs, in effect, as an uninflected verbal form, as in the following:

- (35) yəhī                            mə'ōrōt  
let there be-MASC. SG. luminaries-FEM. PL.  
“Let there be luminaries” (Genesis 1:14)

Since there are no dual forms of verbs (see §4.5), adjectives (see §4.2.2), and pronouns (at least in the active language, see §4.3.1), dual subjects are construed with plural predicates.

## 5.4 Determination

Hebrew substantives are either definite or indefinite. Certain substantives, including proper nouns and most pronouns, are intrinsically definite. Common nouns are determined (become definite) when prefixed by the definite article (see §4.4) or when followed by a pronominal suffix or another definite noun in a genitive construction (i.e., when in construct state before another definite noun; see §4.2.4). According to the grammatical rules of Biblical Hebrew, a noun can be determined in only one of these ways, so that a proper noun cannot stand as the *nomen regens* in a construct chain, and neither a proper noun nor a noun in the construct state can have an article or a pronominal suffix. Although these rules apply generally to Northwest Semitic as a whole, they are by no means universal – the restrictions are much less severe in Ugaritic, for example. Iron Age inscriptional Hebrew provides clear exceptions, such as *yhwš šmrn* “Yahweh of Samaria,” at Kuntillet 'Ajrud, and several possible or certain exceptions are found in Biblical Hebrew itself: for example, *maḥšāš 'ōz* “my refuge of strength” (Psalm 71:7).

## 6. LEXICON

The core vocabulary of ancient Hebrew is an inventory of words shared with other Iron Age Canaanite languages – Phoenician, Ammonite, Moabite, and Edomite. Many are common Semitic, and most are common Northwest Semitic, though several characteristic entries in the lexicon represent preferences in Hebrew that were distinct from their Aramaic equivalents. Verbal examples include Hebrew √'ly versus Aramaic √slq “ascend”; Hebrew √yš' versus Aramaic √npq “go out”; Hebrew √bw' versus Aramaic √'ll “enter”; Hebrew √'zb versus Aramaic √šbq “leave”; and Hebrew √dbr versus Aramaic √mll “speak”; among many others. In most of these cases, the Hebrew preference seems to have been shared by the other members of the Canaanite family, though the evidence for the lexicons of these languages, especially those spoken in Transjordan, is scant. Within the Canaanite group itself, there are also examples of lexical specialization, which, taken together, suggest an isogloss between North and South Canaanite – thus Hebrew √hyy versus Ugaritic-Phoenician √kwn “to be” (narrowed to “be firm” in Hebrew); Hebrew *zāhāb* versus Ugaritic-Phoenician *ḥrṣ* “gold” (rare in Hebrew); Hebrew √šy versus Phoenician √p'l “do, make” (relatively rare and chiefly

poetic in Biblical Hebrew; Moabite also prefers √‘šy, though Ammonite seems allied with Aramaic √‘bd). Note also the retention in South Canaanite (Hebrew, including the Northern or Israelite dialect, and Ammonite) of √ntn “give” (cf. Amorite \*ntn and Akkadian *nadānu*) versus the North Canaanite (Ugaritic and Phoenician) innovation √ytn.

Throughout the history of ancient Hebrew there was a profound penetration of Aramaic vocabulary into the lexicon, a phenomenon that began to gain momentum in the period of Late Biblical Hebrew and steadily increased as Hebrew continued to be studied and spoken while Aramaic became the language of everyday discourse. The result is that, from an early date, there is a substantial Aramaic component to the Hebrew lexicon.

Less far-reaching but still significant is the number of loanwords that entered Hebrew from the speech of the peoples who dominated or controlled Judah (or Judaea) in antiquity. Biblical Hebrew contains a number of words derived from the languages of the major international powers of the Iron Age. There is a scattering of Egyptian words, such as *šēš* “linen” (Egyptian *šš* < \*ššr “linen”) and *ṭabbá‘at* “sealing ring” (Egyptian *ḏb‘wt* “signet, seal”). A number of words reflect Judah’s experience as a tributary of the Assyrian Empire. These include not only names of imperial institutions and officials, as found in the list in 2 Kings 18:17 – *tartān* (Neo-Assyrian *turtānu* “viceroy”), *rab-sārīs* (Neo-Assyrian *rab ša rēši* “chief eunuch”) and *rab-šaḡēh* (Neo-Assyrian *rab šaḡē* “chief butler”), but also words that became part of the general Hebrew vocabulary, such as *šōṭēr* “official, magistrate” (originally “scribe, registrar”?) from the Akkadian verb *šaṭāru* “write.”

In Late Biblical Hebrew many more Akkadian words entered the Hebrew lexicon from the Neo-Babylonian administration: for example, *‘iggéret* “letter” (Neo-Babylonian *egirtu*), *mékes* “tax” (Neo-Babylonian *miksu*), *middā* “tribute” (Neo-Babylonian *mandattu*), and \**ségen* “prefect” (Neo-Babylonian *šaknu* “provincial governor”). Other words were introduced from the bureaucracy of the Persian Empire: for example, *‘āḥašdarpān* “satrap” (Old Persian *hšaçaḡpāvan*; cf. Neo-Babylonian *aḥšadrapannu*), *dāt* “edict, law” (Old Persian *dāta*), and *pardēs* “park” (Old Persian; cf. Avestan *pairi-daēza* “enclosure”).

With the spread of Hellenization after Alexander’s conquest in the fourth century BC, Greek words began to appear in the Hebrew lexicon. Though at first the impact of Greek was felt primarily in the technical terminology of government, law, and commerce – *hipparkəyā* “provincial government” (ἐπαρχία), *būlē* “(city) council” (βουλή “council, senate”), *sanhedrīn* “Sanhedrin” (συνέδριον “council, congress”) – it expanded into the general Hebrew vernacular as Rabbinic Hebrew evolved – thus, *qāmīn* “furnace” (κάμινος), *pīlōn* “gateway” (πυλῶν), *zūg* “pair” (ζεύγος “yoke, pair”; cf. the denominative verb *ziwwēg*, “join”), and so forth. Under Roman administration, Hebrew-speaking Jews also adopted numerous Latin words, including especially, but not exclusively, military terms: for example, *qastrā* “camp” (*castra*), *ligyōn* “legion” (*legiō*), *mônīṭā* “coinage” (*monēta*), and so forth.

## Bibliography

- Anderson, F. 1970. *The Hebrew Verbless Clause in the Pentateuch*. Journal of Biblical Literature Monograph Series 14. Nashville/New York: Abingdon.
- . 1974. *The Sentence in Biblical Hebrew*. Janua Linguarum Series Practica 231. The Hague/Paris: Mouton.
- Bauer, H. and P. Leander. 1918–1922. *Historische Grammatik der hebräischen Sprache des alten Testamentes* (2 vols.). Halle: Niemeyer. Reprint 1962, Hildesheim: Olms.
- Bergsträsser, G. 1918–1929. *Hebräische Grammatik*. Wilhelm Gesenius’ *Hebräische Grammatik* (29th edition, 2 vols.). Leipzig: Vogel. Reprint 1986, Hildesheim: Olms.

- . 1928. *Einführung in die semitischen Sprachen: Sprachproben und grammatische Skizzen*. Munich: Max Hueber. Translated by P. Daniels as *Introduction to the Semitic Languages: Text Specimens and Grammatical Sketches*. 1983. Winona Lake, IN: Eisenbrauns.
- Blau, J. 1993 *A Grammar of Biblical Hebrew* (2nd edition). Porta Linguarum Orientalium 12. Wiesbaden: Otto Harrassowitz.
- . 1998. *Topics in Hebrew and Semitic Linguistics*. Jerusalem: Magnes Press.
- Brockelmann, C. 1956. *Hebräische Syntax*. Neukirchen: Buchhandlung des Erziehungsvereins.
- Garr, W. 1985. *Dialect Geography of Syria-Palestine, 1000–586 BC*. Philadelphia: University of Pennsylvania.
- Harris, Z. 1939. *Development of the Canaanite Dialects*. American Oriental Society 16. New Haven: American Oriental Society.
- . 1941. “Linguistic structure of Hebrew.” *Journal of the American Oriental Society* 61: 143–167.
- Huehnergard, J. 1992. “Languages (Introductory).” In D. N. Freedman (ed.), *Anchor Bible Dictionary*, vol. IV, pp. 155–170. New York: Doubleday.
- Joüon, P. 1923. *Grammaire de l’hébreu biblique*. Rome: Pontifical Biblical Institute. Translated and revised by T. Muraoka as *A Grammar of Biblical Hebrew* (2 vols.). Rome: Pontifical Biblical Institute, 1991.
- Kautsch, E. (ed.). 1909. *Wilhelm Gesenius’ Hebräische Grammatik* (28th edition). Leipzig: Vogel. Translated by A. Cowley as *Gesenius’ Hebrew Grammar* (2nd edition). Oxford: Clarendon Press, 1910.
- Khan, G. 1997. “Tiberian Hebrew phonology.” In A. Kaye (ed.), *Phonologies of Asia and Africa*, vol. I, pp. 85–102. Winona Lake, IN: Eisenbrauns.
- König, F. 1881–1897. *Historisch-kritisches Lehrgebäude der hebräischen Sprache* (3 vols.). Leipzig: Hinrichs.
- Kutscher, E. 1982. *A History of the Hebrew Language*. Jerusalem: Magnes Press; Leiden: Brill.
- Lambdin, T. 1971. *Introduction to Biblical Hebrew*. New York: Scribners.
- Lambdin, T. and J. Huehnergard. 1998. *The Historical Grammar of Classical Hebrew: An Outline*. Cambridge, MA: Harvard University.
- Lambert, M. 1946. *Traité de grammaire hébraïque* (2nd edition). Paris: Presses Universitaires. Reprint 1972, Hildesheim: Gerstenberg.
- Meyer, R. 1966–1972. *Hebräische Grammatik* (3rd edition, 4 vols.). Berlin: de Gruyter.
- Moran, W. 1961. “The Hebrew language in its Northwest Semitic background.” In G. E. Wright (ed.), *The Bible and the Ancient Near East: Essays in Honor of William Foxwell Albright*, pp. 54–72. New York: Doubleday.
- Murtonen, A. 1986–1990. *Hebrew in its West Semitic Setting: A Comparative Survey of Non-Masoretic Hebrew Dialects and Traditions* (4 vols.). Leiden: Brill.
- Pérez-Fernández, M. 1997. *An Introductory Grammar of Rabbinic Hebrew*. Translated by J. Elwolde from Spanish. Leiden: Brill, 1992.
- Polzin, R. *Late Biblical Hebrew: Toward an Historical Typology of Biblical Hebrew Prose*. Harvard Semitic Monographs 12. Missoula, MT: Scholars Press.
- Qimron, E. 1986. *The Hebrew of the Dead Sea Scrolls*. Harvard Semitic Studies. Atlanta: Scholars Press.
- Rabin, C. 1970. “Hebrew.” In T. A. Sebeok (ed.), *Current Trends in Linguistics*. Vol. VI, *Linguistics in South West Asia and North Africa*, pp. 304–346. The Hague: Mouton.
- . 1973. *A Short History of the Hebrew Language*. Jerusalem: Orot.
- . 2000. *The Development of the Syntax of Post-Biblical Hebrew*. Studies in Semitic Languages and Linguistics 29. Leiden/Boston: Brill.
- Rendsburg, G. 1997. “Ancient Hebrew phonology.” In A. Kaye (ed.), *Phonologies of Asia and Africa*, vol. I, pp. 65–83. Winona Lake, IN: Eisenbrauns.
- Richter, W. 1978–1980. *Grundlagen einer althebräischen Grammatik* (3 vols.). St. Ottilien: EOS.
- Sáenz-Badillos, A. 1993. *A History of the Hebrew Language*. Translated by J. Elwolde from Spanish. Cambridge/New York: Cambridge University Press, 1998.
- Schramm, G. and P. Schmitz. “Languages (Hebrew).” In D. N. Freedman (ed.), *Anchor Bible Dictionary*, vol. IV, pp. 203–214. New York: Doubleday.

- Segal, M. 1927. *A Grammar of Mishnaic Hebrew*. Oxford: Clarendon Press.
- Speiser, E. 1925–1934. “The pronunciation of Hebrew according to (based chiefly on) the transliterations in the Hexapla.” *Jewish Quarterly Review* 16:343–382; 23:233–265; 24:9–46.
- Sperber, A. 1937–1938. *Hebrew Based on Greek and Latin Transliterations*. Hebrew Union College Annual 12–13: pp. 103–274.
- . 1966. *A Historical Grammar of Biblical Hebrew*. Leiden: Brill.
- Steiner, R. 1992. “Hebrew.” In *International Encyclopedia of Linguistics*, vol. II, pp. 110–118.
- . 1998. “Ancient Hebrew.” In R. Hetzron (ed.), *The Semitic Languages*, pp. 145–173. London: Routledge.
- Waldman, N. 1989. *The Recent Study of Hebrew*. Bibliographica Judaica 10. Cincinnati: Hebrew Union College.
- Waltke, B. and M. O’Connor. 1990. *An Introduction to Biblical Hebrew Syntax*. Winona Lake, IN: Eisenbrauns.
- Yeivin, Y. 1985. משוורת הלשון הצברית המשתקפת בניקוד הבבלי. Text and Studies 12. Jerusalem: Academy of the Hebrew Language.



# Phoenician and Punic

JO ANN HACKETT

## 1. HISTORICAL AND CULTURAL CONTEXTS

Phoenician is a member of the Semitic language family, specifically the Northwest Semitic branch of Central Semitic. Within Northwest Semitic it is a Canaanite language, the closest relatives of which are Hebrew, Moabite, Ammonite, and Edomite.

### 1.1 Phoenicia

A description of the sources for the Phoenician language depends to a certain extent on what “Phoenician” is held to mean. The term “Phoenicia” is generally reserved for the strip of land sixty miles long (from Acco in the south to Tell Sukas in the north) and at most thirty miles wide, on the northern coast of the Levant, bounded on the west by the Mediterranean and on the east by the Lebanon Mountains – that is, the modern coast of Lebanon and part of the northern coast of modern Israel. As a scholarly convention, this area is referred to as *Phoenicia* after 1200 BC, the beginning of the Iron Age. In the early Iron Age, the ravages of the so-called Sea Peoples along the coast of ancient Canaan and into Egypt forced the withdrawal of Egyptian control over Canaan. This withdrawal allowed the Philistines and other Sea Peoples to gain control over the southern coastal plain, and even to expand eastward, where they met a westward-expanding Israel. The northern coastal plain, however, does not seem to have been invaded from the outside, nor do any disenfranchised or other “settling” peoples seem to have taken over, so that once Egyptian control was gone, the cities in this last remaining part of what had earlier been called Canaan flourished. It is this loose assembly of coastal cities that was called Phoenicia by the Greeks and by modern scholars. The cities were never united into a political entity, although in various periods one or another city was ascendant over the others; the people of Phoenicia continued to think of themselves as Canaanites, or to identify themselves according to their native city.

### 1.2 Textual evidence

Phoenician inscriptions have been found in and around the ancient Phoenician cities, but also throughout the Mediterranean world. The first inscriptions of any length are a series of royal inscriptions from tenth-century BC Byblos, but beginning in the ninth and lasting until the first century AD, there are inscriptions from Asia Minor, Cyprus, Sicily, Sardinia, Malta, Rhodes, Egypt, Greece, the Balearic Islands, and Spain.

A few texts dating even earlier than 1000 BC might be called Phoenician. Several dozen inscribed arrowheads come apparently from the Beqaa, the valley between the Lebanon and Antilebanon Mountains, and from farther south in Palestine (twelfth–eleventh centuries BC); and inscribed clay cones from Byblos date to the middle of the eleventh century. In both cases, the texts are almost entirely personal names and patronymics, so linguistic classification is difficult. The inscription on an eleventh-century fragmentary stela from Nora on Sardinia is most reasonably, given script and provenance, identified as Phoenician. Although the extant inscription contains parts of only four words, the stance of the letters indicates boustrophedon writing. The archaic Nora inscription is an artifact important for tracing the history of Phoenician expansion into the Mediterranean, but it is unfortunately not useful in a survey of the language.

The dialect of the Phoenician colony at Carthage and of inscriptions found throughout the Carthaginian empire is referred to as *Punic*, for which we have evidence beginning in the sixth century BC. Inscriptions dating after the fall of Carthage in 146 BC are said to be written in *Late Punic* or *Neopunic*, although the distinction is more one of script than of dialect. Neopunic inscriptions will be treated in this chapter as simply a late form of Punic that shows the drift that occurred, especially in phonology, after the stabilizing effect that Carthage's hegemony had had on the language was removed. Punic inscriptions date as late as the second century AD, and there are even later *Latino-Punic* inscriptions, Punic written in Latin script, that date to the fourth–fifth centuries AD. Punic inscriptions have been found in North African sites in modern Tunisia, Algeria, and Libya, as well as in Malta, Sardinia, Sicily, France, Spain, and the Balearics.

Besides Phoenician and Punic inscriptions proper, we have names transliterated into Hebrew, Akkadian, Greek, and Latin, plus a few transliterated lexemes found in Greek and Latin inscriptions and in classical sources, notably in Augustine (fourth century AD). Objects inscribed with personal names include seals, bowls, and ostraca. The *Poenulus* of Plautus includes some passages in Punic, but the process of transmission has garbled these passages badly enough that they must be used with circumspection.

Most of the Phoenician and Punic inscriptions can be described as royal inscriptions, tomb inscriptions (both royal and nonroyal), and cultic inscriptions (dedications of buildings or paraphernalia, votive inscriptions). The largest corpus consists of the hundreds of Punic child sacrifice (votive) inscriptions from North Africa, stelae which report that a *mulk*-sacrifice is presented to the god or gods who answered the prayer of the supplicant. Most of the stelae are no longer *in situ*, but beneath some of the stelae the burned remains of children, usually newborns, are found, and sometimes the remains of a substitute lamb or other animal.

### 1.3 Dialectal variation within Phoenician

In the linguistic discussion which follows this section, the focus will be on *Standard Phoenician*, with dialectal variants noted. In addition, a brief overview of dialectal differences occurring within Phoenician is presented here.

Even the earliest Phoenician inscriptions of the tenth–ninth centuries BC show evidence of dialectal differences. The dialect of Byblos is especially distinct from the other early inscriptions (said to be written in Standard Phoenician) and is treated separately in the grammars. Common Phoenician, then, must antedate the first millennium BC.

Old Byblian inscriptions from the tenth–ninth centuries retain the *-y* of III-*y* verbs (a type of “weak verb” in which the third consonant of the root was \**y*; see §4.1); use *z*

(proclitic on the verb) as the relative pronoun (see §4.3.5); and show *zn* as the masculine singular demonstrative (see §4.3.2). In the earliest of these inscriptions, the sarcophagus inscription of Ahiram, the third masculine singular possessive suffix on a genitive noun is *-h*, presumably */-ihu(:)/* (see §4.3.1.2).

After the Old Byblian period, our evidence for Byblian is lacking until the fifth century. At this point, however, Byblian looks more like Standard Phoenician, with relative *ʾš* and demonstrative *z*. The III-weak verbs have lost the third root consonant altogether.

Standard Phoenician inscriptions from the ninth century forward are reasonably homogeneous, with some local variants, especially in inscriptions from Cyprus, where consonant mergers seem to have taken place and the use of “prothetic *ʾ*” is more pronounced than elsewhere.

Punic, the dialect of the western colonies, is extant from the sixth century onwards, but only begins to diverge from Standard Phoenician in late texts, especially after the fall of Carthage in 146 BC. These divergences are largely phonological: modification and loss of the four pharyngeal and glottal obstruents */ħ/(<ḥ>)*, */ʕ/(<ʕ>)*, */h/*, and */ʔ/(<ʔ>)*; and confusion of sibilants (see §3.1). The Punic lexicon is also affected by the number of loanwords and foreign names that make their way into the inscriptions from Greek, Latin, and Numidian. The third masculine possessive suffix on nouns that end in a vowel is *-y* in Standard Phoenician, but *-m* in late Punic texts (see §4.3.1.2).

## 2. WRITING SYSTEM

Phoenician inscriptions are written in a consonantal alphabet, the form of which indicates that it actually developed in Phoenicia, whence it was borrowed by the Hebrews and Arameans, and eventually the Greeks. The Phoenician stage of the script is part of a long history of alphabetic development that can be traced in inscriptions from earlier Canaanite-speaking peoples.

The earliest known inscriptions using this alphabet are two graffiti recently found near Luxor that date from c. 1800 BC. That we have material from Serabit al-Khadem in the Sinai peninsula that is perhaps only slightly later, and other exemplars of fairly high date from Palestine (seventeenth–fifteenth centuries), suggests a date for the invention of the alphabet as far back as 2000 BC.

This writing system was entirely consonantal in origin and operated according to the acrophonic principle: drawing a picture, or pictogram, to represent the first consonant of the word which the picture depicts (such as drawing a bee to represent [b], and so on). In this early form of the alphabet the original *b* depicts a house, as the Canaanite word for house, *\*bayt*, begins with [b]; “(palm of) hand” is *\*kapp*, a word that begins with [k], and so the *k* symbol is a pictogram depicting a hand. The Canaanite-speaking people who invented this writing system would have been familiar with Egyptian writing (see Ch. 7, §2.1), but they simplified the process dramatically so that each of the original symbols corresponded to only one distinct consonantal phoneme.

Throughout the second millennium the consonantal script continued to develop. Whereas the earliest inscriptions were written both vertically and horizontally, horizontal came to predominate. A given early inscription could be written dextrograde, sinistrograde, or boustrophedon, but by 1000 BC, the direction of Phoenician writing had stabilized as sinistrograde. Since several of the pictograms changed stance according to the direction of the line of writing, when the direction stabilized, so did the stance of the characters.

**Table 11.1 The Phoenician consonantal script**

Character	Transcription	Character	Transcription
𐤀	ʾ	𐤁	l
𐤂	b	𐤃	m
𐤄	g	𐤅	n
𐤆	d	𐤇	s
𐤈	h	𐤉	c
𐤊	w	𐤋	p
𐤌	z	𐤍	ṣ
𐤎	ḥ	𐤏	q
𐤐	ṭ	𐤑	r
𐤒	y	𐤓	š
𐤔	k	𐤕	t

By the eleventh century BC, virtually all of the pictographic forms had developed into stylized “linear” descendants. This linear script is used through the first millennium to write Phoenician and Punic, while the Hebrew and Aramaic scripts had begun to follow separate paths by the tenth century. We know that both Hebrew and Aramaic borrowed their writing systems from elsewhere because the scripts they use do not provide an exact match for the consonant repertoire of either Hebrew or early Aramaic. Moreover, the letter names that we know from Hebrew and Aramaic actually correspond to the pronunciation of those words in Phoenician, another clue that the source script was Phoenician.

Though the linear Phoenician script was purely consonantal, a means was eventually developed, as in other consonantal Semitic scripts, to signal the presence of certain vowels consonantly. Consonants so used are conventionally termed the *matres lectionis* (“mothers of reading”). Thus, in late Punic inscriptions we see an inconsistent “vowel notation”; in fact, two systems of *matres lectionis* had merged by this time. The earlier system of Punic *matres lectionis*, named the “Domestic Orthography” by Menken (1981), was used for Semitic words in Punic inscriptions (sporadically from the third century BC): the character ʾ on the end of a word indicated that the word ended in some vowel; occasionally y was used explicitly for final /-ī/. In Phoenician, a final vowel usually marks a morphologically significant addition to a simpler form of the word in question – a pronominal suffix on a noun or verb, for example – with the result that this ʾ often served as a morpheme marker as well. A second system of *matres lectionis*, Menken’s “Foreign Orthography,” came into use slightly later than the Domestic Orthography (i.e., late second century BC). This system was used in Punic for spelling foreign names and words, then consistently in later inscriptions for many words, both foreign and Punic: ʾ for *o*-vowels and *e*-vowels; ʿ for *a*-vowels; w for *u*-vowels; y for *i*-vowels.

Because of their limited land resources, the people of the coastal cities who would eventually be called the Phoenicians early on turned to the sea and to mercantile activities, and it was such maritime occupation that brought the Phoenician people and script into contact with the Greek world. The Greek adaptation of the Phoenician writing system is generally dated at c. 800 BC, on the basis of the variety of scripts already evident in the earliest Greek inscriptions of the late eighth century, indicating both a common origin and some period of development to account for differences. But the antiquity of some Greek letter-forms and the amount of development beyond Phoenician forms suggest a long period of contact

between Phoenicians and the West before the final form of the Greek alphabet emerged. Like the early Canaanite inscriptions, the direction of writing of early Greek inscriptions can be dextrograde, sinistrograde, or boustrophedon; Greek eventually settled on dextrograde, in contrast to Phoenician from 1000 BC onward. All of these features argue for a complicated and extended process of the Greeks' acquiring their alphabet from the Phoenicians, rather than one date that can be proposed as the moment of transmission.

### 3. PHONOLOGY

Since Phoenician is no longer spoken, its phonology must be reconstructed on the basis of (i) transcriptions found in Hebrew, Assyrian, Greek, and Latin writings; and of (ii) comparative phonology of the Semitic languages.

#### 3.1 Consonants

In this chapter, the transliteration scheme commonly utilized in the philological study of Phoenician will be followed, for both consonants and vowels. In Table 11.2 these conventional symbols are used, but are followed by a phonetic transcription within parentheses, where such transcription differs from the conventional representation.

Twenty-nine consonants are reconstructed for Proto-Semitic (see Ch. 6, §3.2.1). Proto-Central Semitic retains all of them, as does Proto-Northwest Semitic. The following consonant mergers occur between Proto-Northwest Semitic and Canaanite (conventional transcription is given in parentheses):

**Table 11.2 The consonantal phonemes of Standard Phoenician**

Manner of articulation	Place of articulation					
	Bilabial	Dental/ Alveolar	Palatal	Velar	Pharyngeal	Glottal
<i>Stop</i>						
Voiceless	p	t		k		ʔ (/ʔ/)
Voiced	b	d		g		
Emphatic		ṭ (/t'/)		q (/k'/)		
<i>Affricate</i>						
Voiceless		s (/ʰs/)				
Voiced		z (/ʰdz/)				
Emphatic		ṣ				
<i>Fricative</i>						
Voiceless		š (/s/)			ħ (/ħ/)	h
Voiced					ʕ (/ʕ/)	
<i>Approximant</i>						
Voiced	w	r ?	y			
<i>Lateral approximant</i>						
Voiced		l				
<i>Nasal</i>						
Voiced	m	n				

(1)	<i>Proto-Northwest Semitic</i>		<i>Canaanite</i>
	*θ and *s (š)	→	/s/ (š)
	*ð and *d <sup>z</sup> (z)	→	/d <sup>z</sup> / (z)
	*θ' (θ) and *t's' (š) and *t' (š)	→	/t's'/ (š)

The following mergers then occur between Canaanite and Phoenician:

(2)	<i>Canaanite</i>		<i>Phoenician</i>
	/ʕ/ (ʕ) and /ɣ/	→	/ʕ/ (ʕ)
	/ħ/ (ħ) and /x/ (ħ)	→	/ħ/ (ħ)
	/tʃ/ (š) and /s/ (š)	→	/s/ (š)

Throughout Northwest Semitic, *n* assimilates to a following consonant, producing a geminate cluster. Geminate consonants are not indicated in the Phoenician script, however, and must be reconstructed, as with other features of the language, on the basis of Phoenician transcriptions into languages with scripts which do indicate gemination and by comparison with other Semitic languages.

There is no evidence for the spirantization of voiced and voiceless stops that is evident in Aramaic and Hebrew from the middle of the first millennium BC onward.

In Phoenician and Punic /ʔ/ (<'>) is often elided. In Punic, /h/ is modified (e.g., the definite article is sometimes written <'> rather than <h>) or omitted altogether. Pharyngeals and glottals are generally modified and eventually confused or lost.

### 3.2 Vowels

The vowels of Phoenician are less well understood than the consonants, since Phoenician inscriptions do not include any vowel notation until very late. Judging from related languages and from transcriptions into other scripts, the vowel phonemes of Figure 11.1 are identified for Standard Phoenician:

	<i>FRONT</i>	<i>CENTRAL</i>	<i>BACK</i>
<i>HIGH</i>	i:		u:
	i		u
<i>MID</i>	e:		o:
<i>LOW</i>		a	

**Figure 11.1** Vowel phonemes of Standard Phoenician

The vowels reconstructed for Proto-Semitic are \**a*, \**i*, \**u*, \**ā*, \**ī*, \**ū* (see Ch. 6, §3.2.2), as well as the diphthongs \**ay* and \**aw* (see Ch. 6, §3.2.3). In the development of Phoenician, however, the Proto-Semitic diphthongs became long mid vowels: \**ay* > /e:/ and \**aw* > /o:/.

### 3.2.1 High vowels

High vowels undergo several changes within the history of Phoenician. The short high-front /i/ (from PS \*i) shows three developments:

1. In syllables which had been originally open (see §3.2.5), accented /i/ > [ē]. Note the name βασιλληχ = *Balsilech* (CIL VIII 16) for /baʕl-sillík/ “Baal has sent,” among other evidence, all of which is late (Hellenistic or beyond).
2. In syllables which had been originally open (see §3.2.5), unaccented /i/ > [ε]. Consider the name Γεραστρατος (Josephus, *C. Ap.* 1, 157) for /gir-ʕastart/ “one bound to Astarte,” among other evidence, all late.
3. Elsewhere \*i is preserved (but see §3.2.3). Thus, Assyrian *ú-ru-mil-ki* (Senn. OI Prism, col. II, line 53, 8th century BC) for Phoenician /ʔo:r-milk/ “the [divine] king is light.”

The long vowel \*ī remains stable; we assume the length in names such as αβιβαλος (Josephus, *Ant.* 8, 5) for /ʔabī-baʕl/ “Baal is my [divine] father.”

Both the short and long high-back vowels, \*u and \*ū, were preserved, though appear to have been eventually fronted, and perhaps unrounded, in certain environments. The evidence for the shift is, however, meager, late, and rather unreliable (*Poenulus*), but it forms one end of a proposed chain (Fox 1996) that is otherwise well grounded. Thus, in *Poenulus*, we see evidence of /u/ > [ü] in the Latin transcription *chyl*, representing /kull/ “all,” and even perhaps of [ü] > [i] in *chil*, a transcription of the same word. For the fronting of /u:/ to [ü:]/[i:], *Poenulus* provides the (perhaps equally unreliable) evidence of the spellings *li* for /lu:/ “O that . . . !”; *hy* for /huʔ/ “he.”

### 3.2.2 Low-central vowels

The observed Phoenician development of Proto-Semitic \*ā > /o:/ (possibly with intermediate stage of \*/ɔ/) is actually a broader phenomenon known as the *Canaanite Shift*. This process occurs early in Canaanite, as is evidenced by the fourteenth-century Canaanite glosses in the Akkadian texts found at el-Amarna in Egypt (see Ch. 8, §1.1). The resulting /o:/ merges with the /o:/ reflex of Proto-Semitic \*aw, and both were eventually raised to /u:/ – note the Punic divine name in Greek transcription, χουσωρ for earlier /ko:sar/ (< \*kawθar; Eusebius *PE* 1.10.11), and feminine plural SANUTH for /sano:t/ “years” (KAI 180 c, e).

In syllables which had been originally open (see §3.2.5), Phoenician accented short /a/ (from PS \*a) > [o]. There is evidence that this change, known as the *Phoenician Shift*, had occurred by at least the eighth century BC. Note the eighth-century Assyrian transcription of the name *ḥi-ru-um-mu* for [ḥi:rom] < /ʔaḥi:rám/ “My [divine] kinsman is exalted” (T-P Annal 27, line 2; T-P Summary Inscription 9, reverse, line 5, has a variant difficult to assess: [ḥi-r]i-mu) and seventh-century *ba-ʕal-ma-lu-ku* for [baʕl-malok] < /baʕl-malák/ “Baal has ruled” (Assurb. Rassam, col. II, line 84). The [o] that was the result of the Phoenician Shift did not merge with /o:/ < \*ā and < \*aw and therefore was not raised to /u:/; recall the above χουσωρ < \*kōšár < \*kawθar, Eusebius, *PE* 1.10.11. The feminine of this same word, χουσαρτις for /ko:sart/ (Eusebius *PE* 1.10.43), provides evidence that the /á/ > [o] shift did not take place in originally closed syllables.

Elsewhere, Proto-Semitic short \*a is preserved in Phoenician (but see §3.2.3).



### 3.2.3 Vowel reduction

There is some evidence (again, *Poenulus*) that short vowels in open syllables are reduced to schwa pretonically in verbs and propretonically in nouns and adjectives, as in Biblical Hebrew.

### 3.2.4 Syllable structure

Syllables in Phoenician (again to the extent that such information can be reconstructed) appear to have the standard Semitic syllable shape: CV or CVC.

### 3.2.5 Accent

Accent also must be reconstructed, but there are clues. Earlier Northwest Semitic had short final case-vowels: \*-*u* for nominative, \*-*i* for genitive, and \*-*a* for accusative. At some point, short final vowels were lost in the Canaanite languages, although there is evidence (see §4.3.1.2) that the genitive case ending remained in Phoenician. As we saw in §§3.2.1–2, lengthening or raising of certain vowels occurred in the (newly) final syllable, as long as the syllable had been originally open. This situation suggests that the accent in Phoenician, as in Hebrew, was on the syllable preceding the case-vowel; then with loss of the case-vowel, it fell on the new final syllable of the word.

## 4. MORPHOLOGY

### 4.1 Word structure

Most Phoenician words, like those in all Semitic languages, are built around a triconsonantal root, which denotes a semantic field. The words themselves are discontinuous morphemes composed of a sequence of three consonants (the *root*) and the vowels and affixes that are morphologically significant. For instance, if the root *k-t-b* means “to write,” the Proto-Central Semitic (and Arabic) *\*katabat* would mean “she wrote,” *\*yaktubu* would mean “he will write,” *\*kātibūna* would mean “those who write,” and so forth. There is evidence for biconsonantal roots in Afro-Asiatic, the family of which Semitic is one branch; there are, furthermore, “weak” verbal roots, roots with first, second, or third root consonants which were originally *w* or *y*, and which had dropped out of the root (usually elided intervocally) in many of the languages, including Phoenician. But for most words, the triconsonantal root is still recoverable.

### 4.2 Nominal morphology

Many nouns are derived from verbal bases, such as participles, infinitives, agent nouns, nouns of place, time, instrument, *inter alia*. Such nouns are often formed with affixes and vocalic patterns that carry specific meanings.

#### 4.2.1 Case, gender, and number

Nominals in Phoenician are marked for gender and number: masculine singular (masc. sg.), masculine plural (masc. pl.), feminine singular (fem. sg.), feminine plural (fem. pl.). There is

some slight evidence of the retention of the Semitic dual. Proto-Northwest Semitic retained the three cases of Proto-Semitic (nominative, genitive, accusative), and there is evidence of at least the genitive in Phoenician (see §4.3.1.2), and possibly the accusative (see §4.4).

#### 4.2.2 State

Nouns occur in two states, the *absolute state* and the *construct state*. A noun in the construct state (called *nomen regens*) is “in construct” with (governs) a following noun in the genitive case (the *nomen rectum*). Together they make up a *construct chain*. If the *nomen rectum* is definite – that is, it includes the definite article; is written with a possessive pronominal suffix; or is a proper noun – the entire chain is definite. If *hbrk bʿl* in KAI 26 A I 1 means “the one blessed by Baal,” then we have an example of a construct chain modified in its entirety by one definite article written on the *nomen regens* (see Lambdin 1971).

#### 4.2.3 Noun endings

Masculine singular nouns have -Ø ending, in both the absolute and construct states.

Feminine singular nouns end in -t, in both absolute and construct states. This ending represents either -t or -ot (< /-at/): both occur in Semitic, and the unvocalized inscriptions do not allow us to make a distinction, except in rare cases such as *št* (/satt-/ < \*/sant-/; /sanat-/ would be written *šnt*). Note the original \*-at ending on the personal name *ab-di-mi-il-ku-ut-ti* (Esarh., p. 48, line 65) for /ʕabd-milkot/ “servant of the [divine] queen.” In late Punic, the final -t is apparently lost; witness the Latin transcription *Himilco* (CIS I 149; CIL VIII 10525) for /(?a)ħi:-milkot/ “brother of the [divine] queen,” and Punic *ħšdyqʿ* (KAI 154, 3) “the righteous one,” a feminine noun, and so vocalized [ʰsʰaddi:kʰo] < [ʰsʰaddi:kʰot] < /ʰsʰaddi:kʰat/.

Masculine plural nouns end in -m in the absolute state: -īm, as in *gubulim* “boundaries” and *alonim* “gods” in *Poenulus*; note also a rare late Punic *mater lectionis* in the ending, -ym of *khnym*, KAI 161, 6, meaning uncertain. Dual nouns apparently end in -ēm as in *iadem* “hands,” KAI 178, 1. Masculine dual and plural nouns in construct end in -ē, as in the goddess Tanit’s epithet *ῥανηβαλος* for /pane:-baʕl/ “face of Baal.”

Feminine plural nouns end in -t, in both absolute and construct states. This ending represents -ūt < /-ōt/ < \*-āt, as in *alonuth* “goddesses” (in *Poenulus*). In the late Latino-Punic inscriptions, the -t is sometimes missing. KAI 180 a and d have *sanu*, while c and e in the same inscription have *sanuth*, all meaning “years.” The feminine dual absolute *mʿtm* for the numeral 200 is probably *miʿatēm*, with ending -tēm (cf. the masculine dual absolute; colloquial modern Arabic -tēn).

#### 4.2.4 Adjectives

Adjectives in Semitic have the same external morphology as nouns. In Phoenician, then: masculine singular -Ø, feminine singular -t, masculine plural -m, feminine plural -t.

### 4.3 Pronouns

Phoenician attests personal pronouns, as well as demonstratives, interrogative pronouns, and relative pronouns.

### 4.3.1 Personal pronouns

Personal pronouns in Phoenician are of two kinds: independent and suffixed. Both sets occur in singular and plural forms, and both lack a gender distinction in the first person (but not in the second and third). There are also sometimes case distinctions, as we will see.

#### 4.3.1.1 Independent personal pronouns

Because Phoenician verbs are conjugated for person, number, and gender, a pronominal subject in a verbal clause is usually not expressed outside the verb itself; that is, an independent pronoun is not necessary, and when used is meant to emphasize the function of the pronominal subject. Independent pronouns can, in fact, be used to emphasize any nominal form in a sentence, such as the direct object of a verb, a pronominal suffix on a noun, or the object of a preposition. The standard forms of the independent personal pronouns and their reconstructed pronunciations are given in (3):

#### (3) Singular

1st com.	ʾnk	/ʾano:ki:/ (occasionally in Punic ʾnky, with -y for /-i:/)
2nd masc.	ʾt	/ʾatta(:)/
2nd fem.	ʾt	/ʾatti(:)/
3rd masc.	hʾ	/huʔ/
	hʾt	/huʔat/
3rd fem.	hʾ	/hiʔ/

#### Plural

1st com.	ʾnhn	/ʾanaħn(V)/
2nd masc.	not attested	
2nd fem.	not attested	
3rd masc.	hmt	/hummat/
3rd fem.	hmt	/himmat/

#### 4.3.1.2 Enclitic personal pronouns

The standard forms of the personal pronouns suffixed to nouns (as possessive) and to prepositions are presented in Table 11.3. The form of the enclitic pronouns attached to nouns shows some variation according to their morphophonetic context, those contexts being: (i) a nominative/accusative singular noun, or a feminine plural noun (i.e., occurring after a consonant); (ii) a genitive singular noun (i.e., occurring after *-i-*); and (iii) a masculine plural noun (i.e., occurring after some other vowel). Recall that the genitive singular ending was retained when other case endings were lost, so that the nominative/accusative enclitics are in effect forms following a consonant, while enclitics attached to genitives are forms occurring after a vowel.

Byblian third-person pronouns are different from the Standard Phoenician forms of Table 11.3. The attested Byblian forms are given in (4):

(4)	Enclitics on singular nouns and prepositions		Enclitics on plural nouns
3rd masc. sg.	h		w /-e:u/
	w	/-o:/	
3rd fem. sg.	h	/-aha(:)/	
3rd masc. pl.	hm	/-hum(ma)/	

**Table 11.3 The enclitic personal pronouns of Standard Phoenician**

	C _ _		-i- _ _		-V _ _	
Singular						
1st com.	Ø	/-i:/	y	/-iya(:)/	y	/-ayy/
	y	/-i:/ <sup>1</sup>				
2nd masc.	k	probably /-ka(:)/				
2nd fem.	k	/-ki(:)/				
3rd masc.	Ø	/-o:/ <sup>2</sup>	y	/-iyu(:)/ <sup>3</sup>	y	/-e:yu(:)/ <sup>3</sup>
3rd fem.	Ø	probably /-a:/ <sup>4</sup>	y	/-iya(:)/ <sup>3</sup>	y	/-e:ya(:)/ <sup>3</sup>
Plural						
1st com.	n	[-o(:)n] <sup>5</sup>				
2nd masc.		not attested				
2nd fem.		not attested				
3rd masc.	m	/-o:m/ <sup>6</sup>	nm	/-no:m/	nm	/-no:m/ <sup>7</sup>
3rd fem.	m	/-e:m/ <sup>8</sup>	nm	/-ne:m/ <sup>9</sup>		

Notes to Table 11.3

<sup>1</sup>The variant *-y* may be a *mater lectionis* (see §2) or by analogy with the genitive singular *-y*.

<sup>2</sup>We assume the nominative/accusative form is patterned on the accusative: \*/-ahu(:)/ > \*/-au(:)/ > /-o/.

<sup>3</sup>In these cases, /y/ arises as a palatal off-glide following a front vowel. The genitive ending on singular nouns is /-i/ and on plural nouns is /-e:/ thus, \*/-ihu(:)/ > /-iyu(:)/; \*/-ayhu(:)/ > \*/-e:hu(:)/ > /-e:yu(:)/; and so forth.

<sup>4</sup>Again, the nominative/accusative form is patterned on the accusative: \*/-aha(:)/ > /-a:/.

<sup>5</sup>See PΥBAΘΩN “our lady” KAI 175, 2.

<sup>6</sup>Again, assuming the accusative form has been taken over by the nominative: \*/-ahum/ > \*/-aum/ > /-o:m/.

<sup>7</sup>From an old plural verbal ending *\*-ūna-*, *\*yaqtulūnahum*. After loss of intervocalic /h/, *\*yaqtulūnaum* gives *\*/yak'tulu: no:m/*. Then /-no:m/ is extended to use on nouns as well; see Huehnergard 1991:190–194; Harris 1936:49–50.

<sup>8</sup>Amadasi Guzzo 1999 notes Krahmalkov's cautious approach (1993; either the 3rd masc. pl. *-m* was leveled through, or the 3rd fem. pl. comes from *\*/-ahim/* > *\*/-aim/* > */-e:m/*), but argues that the former is less likely than the latter.

<sup>9</sup>Guzzo argues that /e:m/ and /ne:m/ are to be differentiated from masculine plural /o:m/ and /no:m/; cf. n. 8.

The third masculine singular *h* is the earliest form and is only attested in the genitive */-ihu(:)/*. The interpretation of the third masculine singular form occurring on plural nouns, *w*, assumes a dual oblique ending before suffixes, as in Biblical Hebrew: thus, *\*/-ayhu(:)/* > *\*/-e:hu(:)/* > *\*/-e:u(:)/*, spelled <-w>.

Late Punic third-person forms are different in part. After a consonant, Punic shows the same enclitic forms as Phoenician proper (in third singular forms, the character ʾ functions as a *mater lectionis*; see §2):

- (5) 3rd masc. sg. ʾ /-o:/  
 3rd fem. sg. ʾ /-a:/  
 3rd masc. pl. m /-o:m/

After a vowel, early Punic texts show the same pronouns as Phoenician:

- (6) 3rd masc. sg. y /-iyu(:)/  
 3rd fem. sg. y /-iya(:)/  
 3rd masc. pl. nm /-no:m/

In later Punic texts, however, the third masculine singular usually appears as *-m* (*/-im/*). Huehnergard argues that */-iyu(:)/* would have been pronounced the same as */-iw/*, and that the *-m* suffix simply demonstrates a nasalization of the word-final */-w/* (for details, see Huehnergard 1991).

Phoenician and Punic enclitic pronouns suffixed to verbs are like those attached to nouns and prepositions with a few exceptions:

(7) *Singular*

<i>1st com.</i>	n	/-ni:/	
<i>2nd masc.</i>	k	/-ka(:)/	
<i>2nd fem.</i>	k	/-ki(:)/	
<i>3rd masc.</i>	h	/-hu(:)/	Old Byblian
	w		later Byblian
	ø		Standard Phoenician, after a consonant
	y		Standard Phoenician, after a vowel
	ʾ		Punic <i>mater lectionis</i>
	m	/-im/	Late Punic
<i>3rd fem.</i>	y		Standard Phoenician
	ʾ		Punic <i>mater lectionis</i>

*Plural*

<i>1st com.</i>	n	/-nu(:)/?	
<i>2nd masc.</i>	not attested		
<i>2nd fem.</i>	not attested		
<i>3rd masc.</i>	m		after a consonant
	nm		after a vowel

### 4.3.2 Demonstrative pronouns

The demonstratives in Phoenician are declined for person and number. They are used in conjunction with the definite article (see §4.4) only sporadically, even when modifying a definite noun; in other words, “this house” would be *hbt z* (“the house this”) or *hbt hz* (“the house the this”). Occasionally, even combinations like *bt z* (“house this”) are found when the phrase must be definite.

The various forms of the near demonstrative (“this, these”) are presented in (8):

(8)	<i>Phoenician</i>	<i>Byblian</i>	<i>Cypriot</i>	<i>Punic variants</i>
<i>Masc. sg.</i>	z	zn, z	ʾz	s, ʾz, hʾz, st, zt, <i>inter alia</i>
<i>Fem. sg.</i>	z	zʾt, zʾ	ʾz	st, zt
<i>Pl.</i>	ʾl	ʾl	ʾl	ʾl

Standard Phoenician *z* is from Proto-Semitic \**ḏ* and is also seen in other Semitic languages as the base for the near demonstratives. Prothetic ʾ is common in Cyprus before word-initial biconsonantal clusters (note that the use of prothetic ʾ suggests that Cypriot Phoenician *z* was pronounced as a double consonant sound, like Greek *zeta*; see Harris 1936:23–24; Woodard 1997:172); late forms with *s* indicate a confusion of sibilants. Vocalizations are unknown. The form extended with *-n* is known so far only at Byblos and on an inscribed ivory box found in Ur, origin unknown, *KAI* 29. Extension with *-n* is common on prepositions, however.

The far demonstrative (“that, those”) is identical to the independent third-person pronouns (see §4.3.1.1), as in Biblical Hebrew.

### 4.3.3 Interrogative pronouns

The interrogative pronouns in their use at the beginning of questions are known in Phoenician only from *Poenulus*. In Phoenician proper, *my* (probably /miya/) “who?” and

*m* (probably /mū/ < \*mō < \*mā) “what?” serve as indefinite relative pronouns as well: “whoever” effaces this inscription (KAI 24, 14); “whatever” (*m’š*) I did (KAI 24, 4). Note the occurrence of *ymu* in a Roman-era Punic inscription, IRT 873, 2, written in Latin characters, with a prothetic vowel.

#### 4.3.4 Indefinite pronoun

Phoenician attests the indefinite pronoun *mm*. Compare Peripheral Akkadian *mīnummē*.

#### 4.3.5 Determinative-relative pronouns

The pronouns *š*, *š*, along with late variations are probably equivalent to the Biblical Hebrew construction of *šē* + *gemination*, which replaces the more usual *’āšer* in very early and relatively late Biblical texts (perhaps denoting a dialectal difference rather than a chronological one).

The Semitic source of this relative pronoun (and its Biblical Hebrew cognate) is obscure. It might be the reflex of \*θ-, as known from Old Akkadian *θū*- and *θūt*, and from standard Akkadian *ša* (reflex of Old Akkadian accusative masculine *θā*). Phoenician and Hebrew *š*-, however, are the only West Semitic forms that can be so explained, all other West Semitic relative pronouns being derived from the voiced counterpart \*ð. An alternate interpretation is one which posits earlier Canaanite \*’*ašar* or the like, which was clipped to *’aš* or even *š*- in Phoenician and some Hebrew dialects (northern?), but developed into *’āšer* in the dialect of Hebrew most represented in the Bible (Judahite) (see Huehnergard forthcoming).

The Old Byblian relative pronoun *z* is, as in most other West Semitic languages, from \*ð (see §4.3.2).

### 4.4 Definite article

The Phoenician definite article, when written, appears as a prefixed *h*- accompanied by gemination of the ensuing consonant, as in Biblical Hebrew (in later texts, the glottal consonant sometimes appears as *’*, or is lost altogether). Though consonant gemination is not regularly indicated in Phoenician orthography, we know the following consonant was doubled because of the unusual spelling *’mmqm* for [ammak’u:m], earlier /hammak’o:m/ “the place” (KAI 173, 5). The origin of the definite article in West Semitic is, however, controversial, and the explanation for the Phoenician definite article is bound up with various theories. Of these, two theories predominate. The most common sees the West Semitic definite article as originating in a deictic particle, as in Indo-European. The second, championed by Lambdin 1971, identifies the origin of the West Semitic definite article in junctural doubling between a noun and a demonstrative, or between a noun and a relative, with the accusative ending of the noun (-*a*) leveled after final short vowels had been lost and the quality of the vowel between noun and demonstrative or relative no longer had meaning. According to Lambdin, in Arabic, Biblical Hebrew, and, we assume, Phoenician, Moabite, Ammonite, and Edomite (which we know only in consonantal texts), the chain [noun + /-a/ + doubling] is reanalyzed as [noun#] + [/a/ + doubling]. Since words in West Semitic ordinarily do not begin with a vowel, /h-/ or /ʔ-/ was added before /a/. Aramaic has a slightly different development, but one that gives Lambdin’s theory its explanatory force: in Aramaic, [noun + /a/ + doubling] became [noun + /a/ + /ʔ/#], where the glottal stop simply provides a boundary between the

short vowel and the next word, as can happen elsewhere in Semitic (this Aramaic sequence subsequently becomes [noun + /a:/]).

The definite article in Phoenician was lost after the inseparable prepositions *b-*, *l-*, *k-* (as in Biblical Hebrew), and after some free-standing prepositions, depending on dialect and chronology. Consider, for example, the Yehawmilk inscription from fifth-century Byblos (KAI 10), in which the definite article disappears after all prepositions; the Eshmunazor inscription from fifth-century Sidon (KAI 14), in which it is lost after all prepositions and after the direct object marker *ʔyt*; and the Karatepe inscription, from late eighth-century Asia Minor (KAI 26), for loss even after the *w-* “and” conjunction.

## 4.5 Verbal morphology

Phoenician verbs are inflected for person, gender, and number through the use of affixes and vowel patterns which are added to the (usually) triconsonantal root.

### 4.5.1 Verb-stems

All the Semitic languages have a verbal system that includes a basic stem (called the G-stem, from German *Grundstamm*), and several derived stems: passive, causative, reflexive, and so on. A general description follows, although the stems have individual histories in each of the Semitic languages (see also Ch. 6, §3.3.5.2):

1. *N-stem*: formed with a prefix *n-*, functioning as the passive of the G-stem, or as a reflexive.
2. *D-stem*: characterized by doubling of the middle root consonant; pluralizing or transitivizing (or raises the transitivity valence), or simply lexical.
3. *C-stem*: formed with a prefix *s-* (originally) or *h-* or *ʔ-*, functioning as a causative.
4. *t-stems* (Gt, tG, Dt, tD, and so on): built by either prefixing or infixing of a *t*; usually reflexive/reciprocal, and sometimes passive.

In addition, G, D, and C also have internal passives, in other words, related passive stems that are constructed by changes in the vowel pattern of the active stem. These are identified by the sigla G-, D-, C-.

The verbal morphology of Phoenician is fairly simple. The stems of which we have evidence are G, N, D, C, tG, Dt, and possibly some internal passives.

### 4.5.2 The Northwest Semitic system

The Northwest Semitic verbal system is characterized by the following constructions:

1. *A perfective*: the “Suffix-Conjugation”.
2. *A preterite/jussive*: the “Prefix-Conjugation” A.
3. *An imperfective*: the “Prefix-Conjugation” B (the only prefix-conjugation attested in Phoenician).
4. *Active and passive participles*: verbal adjectives indicating essential features or ongoing activity.
5. *An infinitive “construct”*: a verbal noun that serves as both infinitive and gerund.
6. *An infinitive “absolute”*: actually an adverb, which stands with a finite verb to emphasize the verb, or stands alone and can be interpreted as any verb form required.
7. *An imperative*.



There is no evidence of the preterite use of Prefix-Conjugation A in Phoenician; in its jussive use, it is indistinguishable in attested forms from Prefix-Conjugation B. In Proto-Canaanite, Conjugation A was *\*yaqtul* and Conjugation B was *\*yaqtulu*; when short final vowels were lost, the morphological distinction between A and B consequently disappeared for most verbs.

### 4.5.3 The Phoenician system

The Northwest Semitic verbal system with its Phoenician reflex, as far as the latter is known, is set out below. The root *q-t-l* is used; vocalization is given when it is secure, even if known solely by reconstruction.

#### 4.5.3.1 G-stem

The Suffix-Conjugation of the Phoenician G-stem is as follows:

(9)	<i>Singular</i>	<i>Plural</i>
1st com.	*qaltũ > qaltĩ	*qaltũ > qtl̄n
2nd masc.	*qaltā > qtl̄t	*qaltum(ũ) > not attested
2nd fem.	*qaltĩ > qtl̄t	*qaltin(ā,na) > not attested
3rd masc.	*qatala > qatal ([qatol])	*qalũ > qatalũ
3rd fem.	*qalat > qatala	*qalā > not attested

See Krahmalkov 1979 for the third feminine singular *qatala*, rather than expected *qatalo*; note that this *-a* is not from an originally open syllable.

The Prefix-Conjugation of the Phoenician G-stem is given in (10):

(10)	<i>Singular</i>	<i>Plural</i>
1st com.	*ʾaqtulu > ʾiqtul	*naqtulu > not attested
2nd masc.	*taqtulu > tiqtul	*taqtulūna > tiqtulū
2nd fem.	*taqtulīna > tqtl (tiqtulī?)	*taqtulnā > tqtl̄n (tiqtulna?)
3rd masc.	*yaqtulu > yiqtul	*yaqtulūna > yiqtulū
3rd fem.	*taqtulu > tiqtul	*yaqtulnā > not attested

The imperative (second person) is as follows:

(11)	<i>Singular</i>	<i>Plural</i>
Masc.	*qutul > qtl	*qutulū > not attested
Fem.	*qutulī > qtl	*qutulnā > not attested

The Northwest Semitic infinitive construct *\*qutul* gives Phoenician *qtl*, and the infinitive absolute *\*qatāl* becomes Phoenician *qatōl*. There is evidence that the infinitive construct of some weak verbs ends in “feminine” *-t*, as in Biblical Hebrew: thus, *l-qḥt* (preposition *l-* marking a purpose clause, and infinitive *qḥt*), from a root *l-q-ḥ* (which, although a strong verb in the perfect, behaves like a *I-n* verb in the imperfect, imperative, and infinitive construct); *l-dʿt*, from a root *y-d-ʿ*; *l-tt*, from a root *y-t-n*; *šbt*, from a root *y-š-b*.

Active and passive G-stem participles are presented in (12):

(12)	<i>Singular</i>	<i>Plural</i>
<i>Active participle</i>		
<i>Masc.</i>	*qātil- (+ case ending) > qōtil	*qātilūma/*qātilīma > qōtilīm
<i>Fem.</i>	*qātilat-/*qātīlt- > qtl̄t	*qātilāt- > not attested
<i>Passive participle</i>		
<i>Masc.</i>	*qatūl- > qatūl	*qatūlīm > qatūlīm
<i>Fem.</i>	*qatult-/*qatūlat- > qtl̄t	*qatūlāt > qatūlōt

No finite G-stem forms are attested in Phoenician.

#### 4.5.3.2 Derived stems

In the construction of the derived stems, the prefixes and affixes used are the same as those of the G-stem. The following are the most basic forms, (third) masculine singular, when appropriate, from Northwest Semitic to Phoenician, as far as can be determined:

The *N-stem* functions as a passive in Phoenician:

1. *Suffix-Conjugation*: \*naqtala > nqtl
2. *Prefix-Conjugation*: \*yiqqatilu > yqtl; note that the \*n- affix assimilates and doubles the first root consonant.
3. *Participle*: \*naqtal- > \*nqtl (only attested as fem. sg. nqtl̄t and masc. pl. nqtl̄m).

The *D-stem* is generally not distinguishable from the G by morphology alone.

4. *Suffix-Conjugation*: \*qattila > qittil
5. *Prefix-Conjugation*: \*yaqattilu > yaqattil
6. *Imperative*: \*qattil > qattil
7. *Infinitive construct*: \*qattil > qattil
8. *Infinitive absolute*: \*qattāl > qattōl
9. *Participle*: \*maqattil- > maqattil

One or two D passive (D-) Suffix-Conjugation forms are perhaps attested, recognized as such by context and by comparison with usage in related languages. There is some evidence for the special form used for the D of roots that are middle weak, that is, missing the middle consonant and therefore having nothing to double in this conjugation: thus, \*qālil, yaqallil from a root *q-w/y-l*; mtp̄p, “drummer,” participle from a root *t-w-p*.

Various Phoenician *C-stem* forms are attested:

10. *Suffix-Conjugation*: \*haqtila > yiqtil; ʾiqtil in late Punic. It is assumed that the *h*-prefix was lost by palatalization, which would have taken place in a high-vowel environment. One suggestion assumes *ha-* > *hi-* (as in Biblical Hebrew) with the addition of the negative *ī*. \*ī hiqtil > ī yiqtil and probably > yiqtēl.
11. *Prefix-Conjugation*: \*yahaqtilu > yqtl. Note the loss of intervocalic *h*; the Phoenician form is perhaps yaqtil.
12. *Infinitive*: \*yaqtīl > /yaqtīl/, as in Karatepe’s *yṯnʾ*, probably /yaṯniʾ/. There is some late evidence of a construct form with *l-*, but without *h-* or *y-* prefix, perhaps /laqtīl/.
13. *Participle*: \*mahaqtīl- > mqtīl. Note again the loss of intervocalic *h*. Phoenician perhaps has maqtīl, although late Punic texts have a *-y-* between the *m-* prefix and the root, representing either a high vowel (*miqtīl*), or the reanalysis to *mVyaqtīl*, by analogy with the Prefix-Conjugation (*yqtl*).

Again, one or two C passive (C-) stem-forms are perhaps attested.

Regarding the *t-stems*, two passive tG forms are attested at Byblos (*yitqatil?*), and two reflexive Dt forms elsewhere (*yiqattil?*).

## 4.6 Prepositions and particles

Phoenician, like many of the Semitic languages, has both free-standing and inseparable (proclitic) prepositions. Inseparable prepositions are *b-* “in,” “consisting of”; *l-* “to/for,” “at”; and *k-* “like/as.” The definite article is lost after these three inseparable prepositions. The preposition \**min-* “from” usually occurs as inseparable *m-*, with the *n-* assimilated to, and presumably doubling, the following consonant.

Many prepositions in Phoenician are extended, either by “prothetic ’”, as in *’b* for *b-*, or by the addition of *-n* or *-t* at the end, as in *bn* for *b-*, *ln* for *l-*, *’t* for *’l* “(up)on, over,” and *pnt* “before.” Prepositions are often combined with nouns to make new prepositions, such as *lpn* “in front of” from *l-* “at” and *pn* “face of”; and they are also combined with each other, even the proclitic prepositions – *lm* “from” from *l-* and *m-* < \**min*; even *lmb* “in,” “from,” “on account of” from *l-*, *m-*, and *b-*.

In Phoenician, the marker of a definite direct object is *’yt* from \**’iyyāt* > *’iyyōt*/([*’iyyūt*]), and is clearly distinct from the preposition *’t* “with” (*’itt*/). In Punic, the direct object marker is written *’t* or even *t*, indicating loss of the consonantal *y* and eventual elision of the *’*, as well. In *Poenulus*, the Latin transcription *yth* indicates that the vowel has become rounded.

The most interesting adverbs in Phoenician are the several negative adverbs, usually modifying verbs. The most common is *bl*, presumably /*bal*/ as in Biblical Hebrew, usually negating a verb but also used with nouns. There is also a negative *’y*, presumably /*’i*/ as in Biblical Hebrew, with *y* as a *mater lectionis* (see §2), which is used as both a particle of nonexistence and a verbal negative. The two can be combined, *’bl* or *’ybl*. For negative commands and prohibitions, *’l*/’*al*/ is used. For a negative purpose clause, *lm* “so that not, lest” is used, a combination of preposition *l-* and negative *m-*. There is no evidence for the negative *l’* so common in Hebrew, Arabic, and Aramaic.

There is evidence for the use of a locative /-a/ ending (originally \**-ah* with consonantal *h*), in some Punic forms with *’ mater lectionis* at the end of the word: *m’l’* and *mt’* [sic] for “above” and “below,” KAI 145, 14.

## 4.7 Conjunctions

The most common conjunctions in Phoenician are *w-* (/wa-/?; later /u/) “and”; *’m* (/’im/?) “if/when”; *k* (/ki/?) “that; because; when”; and *’p* /’ap/?) “moreover.” Prepositions can be used as conjunctions when paired with the relative *’š* (see §4.3.5).

# 5. SYNTAX

The survey of our sources for Phoenician (see §1) makes clear that very little of what we have in Phoenician provides evidence for the syntax of the language. Our longest inscription, from Karatepe in Asia Minor, is a translation of a Luwian inscription, and so must be used with caution as evidence for Phoenician syntax. A large percentage of our inscriptions are formulaic and simply identify the object on which they are written: “that which PN vowed to DN.” There are some clear features of syntax, however, that can be dealt with here.

## 5.1 Word order

Phoenician, like other Semitic languages, makes frequent use of verbless or “nominal” clauses. There is no verb “to be” in the present tense in Phoenician, so equational clauses/sentences are often written as subject + adverb or predicate adjective, and occasionally subject + predicate nominative. Verbal clauses – clauses that contain a conjugated verb – in Phoenician, as elsewhere in Semitic, are usually V–S–O. A switch in word order so that the subject precedes the verb is often a marker of emphasis on the subject.

## 5.2 Hendiadys

Verbal hendiadys is known in Phoenician, as in Biblical Hebrew. This conjoined construction takes one of two forms: (i) [finite verb A + *w*- “and” + finite verb B]; or (ii) [finite verb A + preposition *l*- + infinitive construct of verb B]. Such structural combinations, of course, need not be examples of hendiadys, but when they are, verb B is the main verb of the clause, and verb A is to be translated adverbially, as in:

- (13) *w*-kl    ʾdm    ʾš    ysp                    l-pʿl   mlʾkt   ʿlt   mzbh   zn...  
 and-any person who would increase to do work on altar this...  
 “And anyone who would do work again on this altar...” (KAI 10, 11–12).

The causative *ysp* is being used to denote repeated or continuous action, and is not interpreted literally.

## 5.3 Infinitive absolute

The infinitive absolute in Phoenician can be used to represent any verbal form if the context has made clear which form is expected (i.e., it functions as an unmarked verb form). This use of the infinitive absolute is especially pronounced in the Karatepe inscription, where infinitives absolute even take pronominal objective suffixes.

## 5.4 The vocative

Vocative *l*- is known in Phoenician, as in Ugaritic and Arabic, but is rare. To express a wish, Phoenician can use the particle *l*- /lu:/ “O that...!” proclitic on a verb, but that too is rare. Ordinarily a wish is conveyed by the volitive forms of the verb: (i) the first-person *cohortative*, which is indistinguishable from the imperfect (but note ʾ*pqn* in KAI 50, 3, where the *-n* seems to be a volitive particle, like Biblical Hebrew *nāʾ*); (ii) the imperative, in most cases indistinguishable from the perfect; and (iii) the third-person *jussive*, ordinarily indistinguishable from the imperfect.

## 5.5 Relative clauses

Relative clauses in Phoenician are generally introduced by the relative pronoun ʾš (*z* in Old Byblian), and occasionally by the “interrogative” pronouns (see §4.3.3). There are rare occurrences also of a resumptive pronoun after ʾš:

- (14) ʾnk   yḥawmlk...   ʾš   pʿlt                    hrbt   bʿlt                    gbl   mmlkt   ʿl   gbl  
 I   Yehawmlk... who she made me the lady the Lady of Byblos sovereign over Byblos  
 “I am Yehawmlk whom the lady, Lady of Byblos, made sovereign over Byblos”  
 (KAI 10, 1–2).

## 6. LEXICON

The Phoenician lexicon is, for the most part, typically Semitic, but the Phoenicians spread throughout the Mediterranean as merchants and eventually colonists. Those Phoenicians would, of course, have had exposure to other languages and would have adopted words and names from other cultures. These loanwords come from a number of other languages and language families. The Kilamuwa inscription from Anatolia, where an Aramaic dialect is the local language, describes Kilamuwa as *br* “son of” Ḥayya’, using Aramaic *br* rather than the Canaanite *bn* that is usual in Phoenician. There are also Luwian personal and place names in Phoenician inscriptions from Anatolia, such as the name Kilamuwa itself, and several in the Azatiwada inscription from Karatepe. We also see Egyptian personal and place names in Phoenician inscriptions found in Egypt.

Greek and Latin names and their (usually nominative) case endings are fairly common in later inscriptions, plus a few words like *drachma*, *imperator*, *senator*, and *podium*. Numidian words and personal and place names are known from the North African inscriptions: *mynd* “ruler,” from Numidian *mnkd* “head, chief”; personal names Massinissa (*msnsn*) and Micipsa (*mkwsn*); and the place name Thugga (*tbgg*).

## 7. READING LIST

Ward 1997 is a good, standard overview of Phoenician history and culture. Markoe 2000 also provides an overview, but stresses material culture. Moscati 1968 and Harden 1962 are classic book-length descriptions.

McCarter 1975 traces the development of the Canaanite/Phoenician alphabet, as does Naveh 1982, more generally. Woodard 1997 is an excellent source for early Greek alphabets and their relationship to the Phoenician and Phoenicians.

Amadasi Guzzo 1997 is a nice summary of the Phoenician language. Huehnergard 1992 and 1995 place Phoenician within the Semitic languages. Harris 1936 is still a useful structuralist introduction, although dating of inscriptions is especially out of date, and recent finds are, of course, not included. Segert 1976 is more up to date and includes more about the use of classical and other sources for our knowledge of Phoenician, something between an introductory and reference grammar. Friedrich, Röllig, and Amadasi Guzzo 1999 is a sound reference grammar, with abundant citations to evidence for Phoenician language and grammar outside the Phoenician corpus itself.

### Abbreviations

Assurb. Rassam	Rassam Cylinder of Assurbanipal. Streck, M. <i>Assurbanipal und die letzten assyrischen Könige bis zum Untergang Nineveh's</i> . Leipzig: Hinrichs, 1916.
CIL	<i>Corpus Inscriptionum Latinarum</i> . 1862–. Berlin: Reimer.
CIS	<i>Corpus Inscriptionum Semiticarum</i> . Pars Prima, Inscriptiones Phoeniciae. 1881–. Paris: Klincksieck.
Esarh.	Borger, R. <i>Die Inschriften Asarhaddons Königs von Assyrien</i> . Archiv für Orientforschung Beiheft 9; Graz, 1956.
Eusebius PE	Eusebius, <i>Præparatio evangelica</i>
IRT	<i>The Inscriptions of Roman Tripolitania</i> . Edited by J. Reynold and J. Ward. Rome: British School at Rome, 1952.

Josephus <i>Ant.</i>	Josephus, <i>Jewish Antiquities</i>
Josephus <i>C. Ap.</i>	Josephus, <i>Contra Apion</i>
KAI	<i>Kanaanäische und aramäische Inschriften</i> (3rd edition, 3 vols.). Edited by H. Donner and W. Röllig. Wiesbaden: Otto Harrassowitz, 1971–1976.
Senn. OI Prism	Oriental Institute Prism of Sennacherib's Annals. Luckenbill, D. <i>The Annals of Sennacherib</i> . Chicago: University of Chicago Press, 1924.
T-P	Tadmor, H. <i>The Inscriptions of Tiglath-Pileser III, King of Assyria</i> . Jerusalem: Israel Academy of Sciences and Humanities, 1994.
DN	divine name
PN	personal name

## Bibliography

- Amadasi Guzzo, M. 1996. "L'accompli à la 3e personne du féminin singulier et le pronom suffixe à l'accusatif de la 3e personne du singulier: note de grammaire phénicienne." In B. Pongratz-Leisten, H. Kühne, and P. Xella (eds.), *Ana šadī Labnāni lū allik: Festschrift für Wolfgang Röllig*, pp. 1–9. Neukirchen-Vluyn: Neukirchener Verlag.
- . 1997. "Phoenician-Punic." In Meyers 1997, vol. 4, pp. 317–324.
- . 1999. "Plural feminine personal suffix pronouns in Phoenician." *Eretz Israel* 26 (Cross volume): 46\*–51\*.
- Cross, F., and D. Freedman. 1951. "The pronominal suffixes of the third person singular in Phoenician." *Journal of Near Eastern Studies* 10:228–230.
- . 1952. *Early Hebrew Orthography: A Study of the Epigraphic Evidence*. American Oriental Series 36. New Haven: American Oriental Society.
- Fox, J. 1996. "A sequence of vowel shifts in Phoenician and other languages." *Journal of Near Eastern Studies* 55:37–47.
- Friedrich, J., and W. Röllig. 1999. *Phönizisch-punische Grammatik* (3rd edition, revised by M. Amadasi Guzzo). *Analecta Orientalia* 55. Rome: Pontifical Biblical Institute. [Original publication: Johannes Friedrich. *Phönizisch-punische Grammatik*. 1951.]
- Garr, R. 1985. *Dialect Geography of Syria-Palestine, 1000–586 BCE*. Philadelphia: University of Pennsylvania.
- Hackett, J. 2002. "The study of partially documented languages." In S. Izre'el (ed.), *Semitic Linguistics: The State of the Art at the Turn of the Twenty-first Century*, pp. 57–75. *Israel Oriental Studies* 20. Winona Lake, IN: Eisenbrauns.
- Harden, D. 1962. *The Phoenicians*. London: Thames and Hudson.
- Harris, Z. 1936. *A Grammar of the Phoenician Language*. American Oriental Series 8. New Haven: American Oriental Society.
- Hetzron, R. 1974. "La division des langues sémitiques." In A. Caquot and D. Cohen (eds.), *Actes du premier Congrès international de linguistique sémitique et chamito-sémitique, Paris 16–19 juillet 1969*, pp. 181–194. The Hague/Paris: Mouton.
- Huehnergard, J. 1991. "The development of the third person suffixes in Phoenician." *Maarav* 7:183–194.
- . 1992. "Languages: introductory survey." In D. N. Freedman (ed.), *The Anchor Bible Dictionary*, vol. IV, pp. 155–170. New York: Doubleday.
- . 1995. "Semitic languages." J. Sasson (ed.), *Civilizations of the Ancient Near East*, vol. 4, pp. 2117–2134. New York: Scribners.
- . forthcoming. "On the etymology of the Hebrew relative *šē-*." In Avi Hurvitz and Steven E. Fassberg (eds.), *In Biblical Hebrew in Its Northwest Semitic Setting: Typological and Historical Perspectives*. Jerusalem: Magnes.
- Kouwenberg, N. 1997. *Gemination in the Akkadian Verb*. *Studia Semitica Neerlandica*. Assen: Van Gorcum.

- Krahmalkov, C. 1970. "Studies in Phoenician and Punic grammar." *Journal of Semitic Studies* 15:181–188.
- . 1972. "Comments on the vocalization of the suffix pronoun of the third feminine singular in Phoenician and Punic." *Journal of Semitic Studies* 17:68–75.
- . 1974. "The object pronouns of the third person of Phoenician and Punic." *Rivista di studi fenici* 2:39–43.
- . 1993. "The third feminine plural possessive pronoun in Phoenician-Punic." *Journal of Near Eastern Studies* 52:37–41.
- . 2000. *Phoenician-Punic Dictionary*. Leuven: Peeters.
- . 2001. *A Phoenician Punic Grammar*. Leiden, Boston: Brill.
- Lambdin, T. 1971. "The junctural origin of the West Semitic definite article." In H. Goedicke (ed.), *Near Eastern Studies in Honor of William Foxwell Albright*, pp. 315–333. Baltimore: Johns Hopkins University Press.
- Markoe, G. 1997. "Phoenicians." In Meyers 1997, vol. 4, pp. 324–331.
- . 2000. *Phoenicians*. Berkeley, Los Angeles: University of California Press.
- McCarter, P. 1975. *The Antiquity of the Greek Alphabet and the Early Phoenician Scripts*. Missoula, MT: Scholars Press.
- Menken, D. 1981. "Neopunic Orthography." Dissertation, Harvard University.
- Meyers, E (ed.). 1997. *Oxford Encyclopedia of the Archaeology of the Near East* (5 vols.). New York: Oxford University Press.
- Moscatti, S. 1968. *The World of the Phoenicians*. Translated by A. Hamilton. London: Weidenfeld and Nicolson.
- Naveh, J. 1987. *Early History of the Alphabet* (2nd edition). Jerusalem: Magnes Press.
- Segert, S. 1976. *A Grammar of Phoenician and Punic*. Munich: Beck.
- Voigt, R. 1998. "Der Artikel im Semitischen." *Journal of Semitic Studies* 43:221–258.
- Ward, W. 1997. "Phoenicia." In Meyers 1997, vol. 4, pp. 313–317.
- Woodard, R. 1997. *Greek Writing from Knossos to Homer*. New York: Oxford University Press.



# Canaanite dialects

DENNIS PARDEE

## 1. HISTORICAL AND CULTURAL CONTEXTS

The term *Canaanite* has two primary usages: (i) to designate the dialects of Northwest Semitic spoken in the region called *Canaan* in the second half of the second millennium BC; and (ii) to differentiate the “Canaanite” dialects of the first millennium, primarily Phoenician and Hebrew, from other Northwest Semitic languages spoken in Canaan after c. 1000 BC, primarily the Aramaic dialects. The principal feature defining Canaanite is the so-called *Canaanite shift*, that is, Proto-Semitic \**ā* realized as *ō* (e.g., Hebrew *ṭōb* “good” corresponds to Aramaic *ṭāb*).

For the Canaanite of the second millennium BC, there are two primary sources: (i) the texts written in Akkadian, the lingua franca of the time, by Canaanite scribes and which contain both Canaanitisms and explicit glosses, i.e., words written in cuneiform script as a gloss in the local language on a preceding Akkadian word; (ii) the *Proto-Canaanite* inscriptions, that is, inscriptions written in archaic linear script and apparently recording the local language.

Some controversy surrounds what “Canaan” meant, both politically and geographically, in the second millennium BC (Na’aman 1994). In the second half of the millennium, the term was used to designate the area of Asia under Egyptian control, including a number of city-states. It comprised an area stretching roughly from what is today northern Lebanon to the border of Egypt, perhaps including some of the arable lands of Transjordan. The term is already attested in the first half of the millennium (eighteenth-century BC texts from Mari) in regard to cities located in the same general area, and there is no reason to doubt that the geographical extent of Canaan was already similar to that known several centuries later. The origin of the term is, however, still unclear.

On the possibility of dividing Canaanite into *North Canaanite* and *South Canaanite*, with the former comprised by Ugaritic, see Tropper 1994 and Pardee 1997c.

For the Canaanite of the first millennium BC and later, there are nearly continuous bodies of inscriptions beginning shortly before 1000 BC. In the case of Phoenician, these inscriptions are found from Anatolia to Egypt to Mesopotamia during roughly the first half of the millennium, then throughout the western Mediterranean as late Phoenician and Punic until the latter dies out well into the Christian era. In the case of Hebrew, a long series of dialects is attested from the tenth century BC down to the present. Canaanite languages distinct from Hebrew and Phoenician were also spoken in Transjordan during the first millennium BC, i.e., Moabite, Ammonite, and Edomite. The sources for these languages are very sparse and they cease in the Persian period, replaced by Aramaic; there are thus few data by which to determine how long they survived as living languages.

Because separate articles are devoted to detailed presentations of Hebrew and Phoenician, this article will deal with the earlier manifestations of Canaanite.

## 2. WRITING SYSTEMS

The two principal bodies of evidence for Canaanite in the second millennium BC correspond to two writing systems.

### 2.1 Cuneiform

The greater number of data come from Canaanite features in Akkadian documents that date for the most part from the early fourteenth century. For the description of Akkadian cuneiform as a writing system, see Chapter 8, §2.

The vast majority of these documents, which total nearly four hundred, were discovered at Tell el-Amarna in Egypt (see Ch. 8, §1.1). They represent the international correspondence directed to Egyptian pharaohs of the early fourteenth century, from as far away as Hattusas, the capital of the Hittite empire (north-central Anatolia), and Babylon (southern Mesopotamia). The Akkadian of these letters varies according to the local scribal schools; that used by the scribes of the various cities of Canaan is so marked by local features that it has been described as a scribal “code,” a hybrid language that, though basically Akkadian and thus incomprehensible to speakers of the local language, would have been understood only by Akkadian speakers trained in its use (Moran 1987:27; 1992:xxi–xxii; Rainey 1996, II:1–16, 31–32).

The Canaanite substratum may be derived by triangulation between the written forms, normative Akkadian of the period, and later Canaanite. The primary difficulties with this derivation are two: (i) problems stemming from the writing system itself, which permits multiple values for a given sign; and (ii) the very process of describing an unknown language by assumed parallels from other languages that are only attested half a millennium and more later. These difficulties are palliated in part by the presence of explicit glosses: an Akkadian word or a Sumerian logogram of known meaning may be followed by one or two oblique wedges (German *Glossenkeil* is the technical term for such a wedge) and then by a Canaanite word. The most famous of these is perhaps *ŠU* : *zu-ru-uḫ* in EA 287:27, where *ŠU* is the logogram for “hand/arm” and *zu-ru-uḫ* is the Canaanite gloss, corresponding to Hebrew *z-rōʿ*<sup>c</sup>, Aramaic *d-rāʿ*<sup>c</sup>, and Arabic *ḍirāʿ*<sup>c</sup> and illustrating the shift of \**ā* to *ō* (Sivan 1984: 29), and perhaps of \**d* to *z* (Sivan 1984: 41).

As fraught with difficulties as the above described derivational process is, we know much more about Canaanite from these Akkadian texts than we do from the so-called Proto-Canaanite inscriptions. That is because the latter are far fewer in number and poorly preserved.

### 2.2 Proto-Canaanite

The problem of the Proto-Canaanite inscriptions is directly linked with that of the *Proto-Sinaitic* inscriptions. The latter are a group of inscriptions, numbering about thirty, discovered near Egyptian turquoise mines in the Sinai, dated variously to the eighteenth or fifteenth centuries BC, which have been only partially deciphered but which seem to represent a form of early West Semitic (for a recent overview with bibliography, see

Pardee 1997b). Corresponding to these texts are a group of about twenty texts discovered in southern Canaan and spread over about five centuries, from the seventeenth century BC to the twelfth (Sass 1988, 1991).

The state of preservation of these latter, Proto-Canaanite, inscriptions is even poorer than is that of the Proto-Sinaitic inscriptions. The identification of Proto-Canaanite as a West Semitic script rests on (i) formal similarity with the earlier Proto-Sinaitic script; (ii) the decipherment of a minority of these texts; and (iii) the formal evolution towards Phoenician script. Because of these difficulties, the state of decipherment of these inscriptions is even less advanced than in the case of the Proto-Sinaitic inscriptions. The principal text of one of the best-preserved of Proto-Canaanite inscriptions, that from 'Izbet Šarṭah, seems not to be Semitic in spite of the fact that it contains a Hebrew/Phoenician-type abecedary. On the other hand, the well-known Lachish Ewer inscription has been very plausibly deciphered as West Semitic (for an overview, with bibliography, see Pardee 1997a). Unfortunately, the state of preservation of most of the other inscriptions and their broad geographical and temporal spread make reliable decipherment in most cases impossible. These inscriptions, to the extent that they are Semitic, are written in a purely consonantal script, with no use of *matres lectionis*; and this feature coupled with the problems posed by the paucity and state of the texts make it difficult to define the language represented. The presence of a Hebrew/Phoenician-type abecedary dating to c. 1200 BC in the 'Izbet Šarṭah inscription may be seen as indicating, even if the actual text accompanying the abecedary is in another language, that the script was used in other cases to write texts in a language of the Canaanite type. This conclusion is borne out by the Lachish Ewer inscription.

In addition to these texts from southern Canaan, there are a group of arrowhead inscriptions discovered in southern Canaan and Phoenicia and a very limited number of archaic Phoenician inscriptions from Byblos that seem to provide a bridge between Proto-Canaanite and Phoenician. Unfortunately, the small number of texts and the states of preservation again interfere in determining origins and filiations of the scripts as well as of the languages represented.

Finally, there is at least one inscription in the Ugaritic cuneiform script that has been identified as Phoenician in nature (see Ch. 11, §2.1).

### 3. GRAMMAR

From inscriptions that predate the abecedaries of the 'Izbet Šarṭah ostrakon (twelfth century BC), some fifteen Proto-Canaanite signs representing consonantal phonemes are identifiable with some degree of certainty. As these match the Proto-Sinaitic data, as well as the data from the later West-Semitic languages, it may be assumed that the original Proto-Canaanite consonantal inventory was similar to, if not identical with, the Proto-Sinaitic inventory and that the two groups of texts represent the same language, or two or more languages/dialects descended from a common ancestor.

Virtually all other aspects of the linguistic description of Canaanite dialects are derived from the texts written in Akkadian cuneiform. After a century of research, comprehensive studies of these data have been produced by Sivan 1984 (phonology, morphology, and lexicon of the Northwest Semitic words in western Akkadian texts of the fifteenth–thirteenth centuries); Rainey 1996 (a study of the Akkadian of the Amarna texts, with special emphasis on Canaanite features, particularly verbal morphosyntax); and Moran 1987 and 1992 (comprehensive translations of the Amarna texts into French and English). Sivan spread his

net a bit wider than he might have done (see Huehnergard 1987); his work is thus useful as a collection of all data furnished by texts written in Akkadian on the various Northwest Semitic languages from Antioch to the border of Egypt in the period covered, but it is more difficult to use as a source for defining Canaanite. Rainey 1996, on the other hand, is specifically a study of the Akkadian texts written by Canaanite scribes; its goal, however, was not to present exclusively the extracted Canaanite data as a grammar of Canaanite, but to present the larger picture, of which the Canaanite part is sometimes quite small. All the relevant data are, however, gathered in these two works, accompanied by expert analyses and extensive bibliographical information (including proper credit to earlier scholarship, particularly Moran's basic studies).

The following are some of the primary characteristics of Canaanite of c. 1400 BC:

1. The Canaanite shift of \**ā* to *ō*.
2. A consonantal inventory that is smaller than that of Ugaritic and different from that of Aramaic (e.g., \**d* → *ṣ*).
3. A case system marked primarily by suffixed vowels, like that of Ugaritic (see Ch. 9, §4.2.2). Case-vowels have generally disappeared or acquired other functions in the first-millennium Northwest Semitic languages.
4. A verbal system of which the morphology and morphosyntax are very similar to those of Ugaritic (see Ch. 9, §4.4). The first-millennium languages have evolved beyond this stage, often retaining only remnants of the earlier systems.
5. The probable absence of a *š*-causative stem (like Phoenician and Hebrew).
6. Dissimilation of the vowel *a* in *YaQTaL*- verbal forms, giving *YiQTaL*, the so-called *Barth–Ginsberg Law*.
7. Many details of the lexical inventory are known (Sivan 1984), but pieces of systems – for example, primary verbs of movement – are missing, making comparisons with later systems difficult.

One may speak of these features as defining Canaanite; it is likely, however, that constellations of less important features characterized a number of local Canaanite dialects.

## References and Abbreviations

- EA = text citation according to Knudtzon, J. 1915. *Die el-Amarna Tafeln mit Einleitung und Erläuterungen*. Leipzig: Hinrichs.
- Huehnergard, J. 1987. Review of Sivan 1984. *Journal of the American Oriental Society* 107:713–725.
- Moran, W. 1987. *Les lettres d'el Amarna. Correspondance diplomatique du pharaon*. Littératures Anciennes du Proche-Orient 13. Paris: Les Editions du Cerf.
- . 1992. *The Amarna Letters*. Baltimore: Johns Hopkins University Press.
- Na'aman, N. 1994. "The Canaanites and their land – a rejoinder." *Ugarit-Forschungen* 24:397–418.
- Pardee, D. 1997a. "Proto-Canaanite." In E. Meyers (ed.), *The Oxford Encyclopedia of Archaeology in the Near East*, vol. 4, pp. 354–355. Oxford: Oxford University Press.
- . 1997b. "Proto-Sinaitic." In E. Meyers (ed.), *The Oxford Encyclopedia of Archaeology in the Near East*, vol. 4, pp. 352–354. Oxford: Oxford University Press.
- . 1997c. Review of Tropper 1994. *Journal of the American Oriental Society* 117:375–378.
- Rainey, A. 1996. *Canaanite in the Amarna Tablets. A Linguistic Analysis of the Mixed Dialect Used by Scribes from Canaan* (4 vols.). Handbuch der Orientalistik, 25. Leiden: Brill.
- Sass, B. 1988. *The Genesis of the Alphabet and Its Development in the Second Millennium BC*. Ägypten und Altes Testament 13. Wiesbaden: Otto Harrassowitz.
- . 1991. *Studia Alphabetica. On the Origin and Early History of the Northwest Semitic, South Semitic and Greek Alphabets*. Orbis Biblicus et Orientalis 102. Freiburg: Universitätsverlag; Göttingen: Vandenhoeck and Ruprecht.

- Sivan, D. 1984. *Grammatical Analysis and Glossary of the Northwest Semitic Vocables in Akkadian Texts of the 15th–13th C. BC from Canaan and Syria*. *Alter Orient und Altes Testament*, 214. Kevelaer: Butzon and Bercker; Neukirchen-Vluyn: Neukirchener Verlag.
- Tropper, J. 1994. "Is Ugaritic a Canaanite language?" In G. Brooke, A. Curtis, and J. Healey (eds.), *Ugarit and the Bible*. Proceedings of the International Symposium on Ugarit and the Bible, Manchester, September 1992, pp. 343–353. *Ugaritisch-Biblische Literatur* 11. Münster: UGARIT-Verlag.

# Aramaic

STUART CREASON

## 1. HISTORICAL AND CULTURAL CONTEXTS

### 1.1 Overview

Aramaic is a member of the Semitic language family and forms one of the two main branches of the Northwest Semitic group within that family, the other being Canaanite (comprising Hebrew, Phoenician, Moabite, etc.). The language most closely related to Aramaic is Hebrew. More distantly related languages include Akkadian and Arabic. Of all the Semitic languages, Aramaic is one of the most extensively attested, in both geographic and temporal terms. Aramaic has been continuously spoken for approximately 3,500 years (c. 1500 BC to the present) and is attested throughout the Near East and the Mediterranean world.

Aramaic was originally spoken by Aramean tribes who settled in portions of what is now Syria, Lebanon, Jordan, Turkey, and Iraq, a region bounded roughly by Damascus and its environs on the south, Mt. Amanus on the northwest and the region between the Balikh and the Khabur rivers on the northeast. The Arameans were a Semitic people, like their neighbors the Hebrews, the Phoenicians, and the Assyrians; and unlike the Hittites, Hurrians, and Urartians. Their economy was largely agricultural and pastoral, though villages and towns as well as larger urban centers, such as Aleppo and Damascus, also existed. These urban centers were usually independent political units, ruled by a king (Aramaic *mlk*), which exerted power over the surrounding agricultural and grazing regions and the nearby towns and villages. In later times, the language itself was spoken and used as a lingua franca throughout the Near East by both Arameans and non-Arameans until it was eclipsed by Arabic beginning in the seventh century AD. Aramaic is still spoken today in communities of eastern Syria, northern Iraq, and southeastern Turkey, though these dialects have been heavily influenced by Arabic and/or Kurdish. These communities became increasingly smaller during the twentieth century and may cease to exist within the next few generations.

### 1.2 Historical stages and dialects of Aramaic

The division of the extant materials into distinct Aramaic dialects is problematic due in part to the nature of the writing system (see §2) and in part to the number, the kinds, and the geographic extent of the extant materials. Possible dialectal differences cannot always be detected in the extant texts, and, when differences can be detected, it is not always clear whether the differences reflect synchronic or diachronic distinctions. With these caveats in mind, the extant Aramaic texts can be divided into five historical stages to which a sixth

stage may be added: *Proto-Aramaic*, a reconstructed stage of the language prior to any extant texts.

### 1.2.1 Old Aramaic (950–600 BC)

Though Aramaic was spoken during the second millennium BC, the first extant texts appear at the beginning of the first millennium. These texts are nearly all inscriptions on stone, usually royal inscriptions connected with various Aramean city-states. The corpus of texts is quite small, but minor dialect differences can be detected, corresponding roughly to geographic regions. So, one dialect is attested in the core Aramean territory of Aleppo and Damascus, another in the northwestern border region around the Aramean city-state of Samʿal and a third in the northeastern region around Tel Fekheriye. There are a few other Aramaic texts, found outside these regions, most of which attest Aramaic dialects mixed with features from other Semitic languages, for example, the texts found at Deir ʿAlla.

### 1.2.2 Imperial or Official Aramaic (600–200 BC)

This period begins with the adoption of Aramaic as a *lingua franca* by the Babylonian Empire. However, few texts are attested until c. 500 BC when the Persians established their empire in the Near East. The texts from this period show a fairly uniform dialect which is similar to the “Aleppo–Damascus” dialect of Old Aramaic. However, this uniformity is due largely to the nature of the extant texts. Nearly all of the texts are official documents of the Persian Empire or its subject kingdoms, and nearly all of the texts are from Egypt. It is likely that numerous local dialects of Aramaic existed, but rarely are these dialects reflected in the texts, one possible exception being the Hermopolis papyri (see Kutscher 1971).

### 1.2.3 Middle Aramaic (200 BC–AD 200)

This period is marked by the emergence of local Aramaic dialects within the textual record, most notably Palmyrene, Hatran, Nabatean, and the dialect of the Aramaic texts found in the caves near Qumran (the Dead Sea Scrolls). However, many texts still attest a dialect very similar to Imperial Aramaic, but with some notable differences (sometimes called *Standard Literary Aramaic*; see Greenfield 1978).

### 1.2.4 Late Aramaic (AD 200–700)

It is from this period that the overwhelming majority of Aramaic texts are attested, and, because of the abundance of texts, clear and distinct dialects can be isolated. These dialects can be divided into a western group and an eastern group. Major dialects in the west include Samaritan Aramaic, Jewish Palestinian Aramaic (also called Galilean Aramaic) and Christian Palestinian Aramaic. Major dialects in the east include Syriac, Jewish Babylonian Aramaic, and Mandaic. This period ends shortly after the Arab conquest, but literary activity in some of these dialects continues until the thirteenth century AD.

### 1.2.5 Modern Aramaic (AD 700 to the present)

This period is characterized by the gradual decline of Aramaic due to the increased use of Arabic in the Near East. Numerous local dialects, such as Ṭuroyo in southeastern Turkey and



Maṣlulan in Syria, were attested in the nineteenth century, but by the end of the twentieth century many of these dialects had ceased to exist.

## 2. WRITING SYSTEM

### 2.1 The alphabet

Aramaic is written in an alphabet which was originally borrowed from the Phoenicians (c. 1100 BC). This alphabet represents consonantal phonemes only, though four of the letters were also sometimes used to represent certain vowel phonemes (see §2.2.1). Also, because the Aramaic inventory of consonantal phonemes did not exactly match the Phoenician inventory, some of the letters originally represented two (or more) phonemes (see §3.2). During the long history of Aramaic, these letters underwent various changes in form including the development of alternate medial and final forms of some letters (see Naveh 1982). By the Late Aramaic period, a number of distinct, though related, scripts are attested. Below are represented two of the most common scripts from this period, the Aramaic square script (which was also used to write Hebrew) and the Syriac Estrangelo script, along with the standard transliteration of each letter. Final forms are listed to the right of medial forms. In Christian Palestinian Aramaic an additional letter was developed to represent the Greek

**Table 13.1 Aramaic consonantal scripts**

Square script	Estrangelo	Transliteration
א	ܐ	ʾ
ב	ܒ	b
ג	ܓ	g
ד	ܕ	d
ה	ܗ	h
ו	ܘ	w
ז	ܙ	z
ח	ܚ	ḥ
ט	ܛ	ṭ
י	ܝ	y
כ	ܟ	k
ל	ܠ	l
מ	ܡ	m
נ	ܢ	n
ס	ܣ	s
ע	ܥ	ʿ
פ	ܦ	p
צ	ܥ	ṣ
ק	ܩ	q (or k)
ר	ܪ	r
ש	ܫ	š
ת	ܬ	t

**Table 13.2 Aramaic vowel diacritics**

Tiberian	Transliteration	Jacobite	Transliteration
בֿ or בֿי	bī or bî	ܒ̈ or ܒ̈י	bī or bî
בֿ or בֿי	bē or bê	ܒ̈	be
בֿ	be		
בֿ	ba	ܒ̈	ba
בֿ	bā or bo	ܒ̈	bā
בֿ or בֿי	bō or bô		
בֿ or בֿי	bu or bū	ܒ̈ or ܒ̈י	bū or bû

letter π in Greek loanwords. It had the same form as the letter *p* of the Estrangelo script, but was written backwards.

## 2.2 Vowel representation

### 2.2.1 Matres lectionis

Prior to the seventh or eighth century AD, vowels were not fully represented in the writing of Aramaic. Instead, some vowels were represented more or less systematically by the four letters *ʾ*, *h*, *w*, and *y*, the *matres lectionis* (“mothers of reading”). The first two, *ʾ* and *h*, were only used to represent word-final vowels. The last two, *w* and *y*, were used to represent both medial and final vowels. The letter *w* was used to represent /u:/ and /o:/. The letter *y* was used to represent /e:/ and /i:/. The letter *ʾ* was used to represent /a:/ and /e:/, although its use for /a:/ was initially restricted to certain morphemes and its use for /e:/ did not develop until the Middle or Late Aramaic period. The letter *h* was also used to represent /a:/ and /e:/. The use of *h* to represent /e:/ was restricted to certain morphemes and eventually *h* was almost completely superseded by *y* in the texts of some dialects or by *ʾ* in others. The use of *h* to represent /a:/ was retained throughout all periods, but was gradually decreased, and eliminated entirely in the texts of some dialects, by the increased use of *ʾ* to represent /a:/. Originally, *matres lectionis* were used to represent long vowels only. In the Middle Aramaic period, *matres lectionis* began to be used to represent short vowels and this use increased during the Late Aramaic period, suggesting that vowel quantity was no longer phonemic (see §3.3.2 and §3.3.3).

### 2.2.2 Systems of diacritics

During the seventh to ninth centuries AD, at least four distinct systems of diacritics were developed to represent vowels. These four systems were developed independently of one another and differ with respect to the number of diacritics used, the form of the diacritics, and the placement of the diacritics relative to the consonant. Two systems were developed by Syriac Christians: the Nestorian in the east and the Jacobite in the west. Two systems were developed by Jewish communities: the Tiberian in the west and the Babylonian in the east. The symbols from two of these systems, as they would appear with the letter *b*, are represented in Table 13.2 along with their standard transliteration.

The Tiberian system also contains four additional symbols for vowels, all of which represent “half-vowels.” The phonemic status of these vowels is uncertain (see §3.3.3.1) and one of the symbols can also be used to indicate the absence of a vowel:

(1)	Symbol	Transliteration
	Ⲛ	ə or no vowel
	Ⲛ̇	ë
	Ⲛ̈	ä
	Ⲛ̉	ö

### 2.3 Other diacritics

The Tiberian system and the two Syriac systems contain a variety of other diacritics in addition to those used to indicate vowels. The Tiberian system marks two distinct pronunciations of the letter *š* by a dot either to the upper left or to the upper right of the letter, and it indicates that a final *h* is not a *mater lectionis* by a dot (*mappiq*) in the center of the letter. The Syriac systems indicate that a letter is not to be pronounced by a line (*linea occultans*) above that letter. Both the Tiberian and the Syriac systems also contain diacritics that indicate the alternate pronunciations of the letters *b*, *g*, *d*, *k*, *p*, and *t* (see §3.2.3). The pronunciation of these letters as stops is indicated in the Tiberian system by a dot (*daghesh*) in the center of the letter, and in the Syriac system by a dot (*quššāyā*) above the letter. The pronunciation of these letters as fricatives is indicated in the Tiberian system either by a line (*raphe*) above the letter or by the absence of any diacritic, and in the Syriac system by a dot (*rukkākā*) below the letter (see also Morag 1962 and Segal 1953).

## 3. PHONOLOGY

### 3.1 Overview

The reconstruction of the phonology of Aramaic at its various stages is complicated by the paucity of direct evidence for the phonological system and by the ambiguous nature of the evidence that does exist. The writing system itself provides little information about the vowels, and its representation of some of the consonantal phonemes is ambiguous. Transcriptions of Aramaic words in other writing systems (such as Akkadian, Greek, or Demotic) exist, but this evidence is relatively fragmentary and difficult to interpret. The phonology of the language of the transcriptions is not always fully understood and so the effect of the transcriber's phonological system on the transcription cannot be accurately determined. Furthermore, no systematic grammatical description of Aramaic exists prior to the beginning of the Modern Aramaic period. So, the presentation in this section is based upon (i) changes in the spelling of Aramaic words over the course of time; (ii) the information provided by the grammatical writings and the vocalized texts from the seventh to ninth century AD; (iii) the standard reconstruction of the phonology of Proto-Aramaic; and (iv) the generally accepted reconstruction of the changes that took place between Proto-Aramaic and the Late Aramaic dialects.

### 3.2 Consonants

The relationship of Aramaic consonantal phonemes to Aramaic letters is a complex one since the phonemic inventory underwent a number of changes in the history of Aramaic. Some of these changes took place after the adoption of the alphabet by the Arameans and produced systematic changes in the spelling of certain Aramaic words.

**Table 13.3 Old Aramaic consonantal phonemes**

Manner of articulation	Place of articulation							
	Bilabial	Inter-dental	Dental/ Alveolar	Palato- alveolar	Palatal	Velar	Uvular	Pharyngeal
<i>Stop</i>								
Voiceless	p		t			k		ʔ ( ʔ )
Voiced	b		d			g		
Emphatic			t' (t)			k' (q)		
<i>Fricative</i>								
Voiceless		θ (š)	s	š				ħ (h)
Voiced		ð (z)	z					ʕ ( ʕ )
Emphatic		θ' (š)	s' (š)					
<i>Trill</i>							r (r)	
<i>Lateral cont.</i>								
Voiceless			ɬ (š)					
Voiced			l					
Emphatic			ɬ' (q)					
<i>Nasal</i>	m		n					
<i>Glide</i>	w				y			

### 3.2.1 Old Aramaic consonantal phonemes

Table 13.3 presents the consonantal phonemes of Old Aramaic with the transliteration of their corresponding symbols in the writing system (see Table 13.1). Only one symbol is listed in those cases in which the transliteration of the written symbol is identical to the symbol used to represent the phoneme. In all other cases, the transliteration of the written symbol is placed in parentheses. Phonemes listed as “Emphatic” are generally considered to be pharyngealized. Note that three letters (*z*, *š* and *q*) each represented two phonemes and that one letter (*š*) represented three phonemes, although in one Old Aramaic text (Tel Fekheriye) the /θ/ phoneme was represented by *s* rather than *š* each of which, therefore, represented two phonemes. That the letter *š* has /ɬ/ as one of its values and *q* has /ɬ'/ as one of its values is likely (see Steiner 1977), but not certain. An alternative for *q* is /ð'/. No satisfactory alternative has been proposed for *š*.

In texts of the Samʿal dialect of Old Aramaic and in the Sefire texts found near Aleppo, the word *npš* is also spelled *nbš*. The occasional spelling of words with *b* rather than *p* also occurs in Canaanite dialects and Ugaritic and suggests that voicing may not have distinguished labial stops in some of the dialects of Northwest Semitic.

### 3.2.2 Imperial Aramaic consonantal phonemes

By the Imperial Aramaic period, three changes had taken place among the dental consonants: (i) /ɬ/ had become /s/; (ii) /ɬ'/ had become /ʕ/; and (iii) /ð/, /θ/, and /θ'/ had become /d/, /t/, and /t'/, respectively. These changes reduced the phonemic inventory of dentals to the following:

(2)	Stop	Fricative	Lateral continuant	Nasal
Voiceless	t	s		
Voiced	d	z	l	n
Emphatic	t' (ṭ)	s' (ṣ)		

These changes in the phonemic inventory produced changes in the spelling of Aramaic words. For example, words containing the phoneme /ð/ and spelled with the letter *z* became spelled with the letter *d* because the phoneme /ð/ had become /d/. Similar spelling changes took place in words spelled with the letters *š*, *ṣ* and *q*. For some time, both spellings are attested in Aramaic texts, but the change is complete by the Late Aramaic period, except in Jewish Aramaic dialects in which the letter *š* is retained for the phoneme /s/ in a few words, perhaps under the influence of Hebrew which underwent the same sound change but which consistently retained the older spelling.

### 3.2.3 Stop allophony

At some time prior to the loss of short vowels (see §3.3.2), the six letters *b*, *g*, *d*, *k*, *p*, and *t* each came to represent a pair of sounds, one a stop, the other a fricative. For example, *b* represented [b] and [v] (or, possibly, /β/); *p* represented [p] and [f] (or, possibly, /φ/); and so forth. At this stage, the alternation between the stop and fricative articulations was entirely predictable from the phonetic environment. The stop articulation occurred when the consonant was geminated (lengthened) or was preceded by another consonant. The fricative articulation occurred when the consonant was not geminated and was also preceded by a vowel. This alternation was purely phonetic in the case of the four pairs of sounds represented by *b*, *p*, *g*, and *k*. In the case of the two pairs of sounds represented by *d* and *t* the alternation was either phonetic or morphophonemic. If the development of this alternation occurred prior to the shift of /ð/ to /d/ and /θ/ to /t/ (see §3.2.2), then the presence of these two phonemes would have made the alternation morphophonemic. If it occurred after this shift, then the alternation was phonetic. At a later stage of Aramaic, short vowels were lost in certain environments and, as a result, the environment which conditioned the alternation was eliminated in some words. The fricative articulation, however, was not eliminated and so the alternation between the two articulations became phonemic in all six cases.

## 3.3 Vowels

The inventory of Aramaic vowel phonemes is more difficult to specify than that of consonantal phonemes, since vowels are not fully represented in the writing system until the beginning of the Modern Aramaic period. Prior to that time, the *matres lectionis* (see §2.2.1) were the only means by which vowels were represented. In the Old and Imperial Aramaic periods, the *matres lectionis* were only used to indicate long vowels. During the Middle Aramaic period they began to be used to indicate short vowels as well, and this expansion of their use continued into the Late Aramaic period. This change in the use of the *matres lectionis* suggests that vowel quantity was not phonemic by the Middle Aramaic period and that vowel quality was the only relevant factor in their use. Given this evidence and the data provided by the four systems of vowel diacritics that were developed at the beginning of the Modern Aramaic period, three distinct stages of the phonology of Aramaic vowels can be distinguished: Proto-Aramaic, Middle Aramaic, and Late Aramaic.

### 3.3.1 Proto-Aramaic

The reconstructed Proto-Aramaic inventory of vowel phonemes is equivalent to the reconstructed Proto-Semitic inventory of vowel phonemes:

(3)	<i>Front</i>	<i>Central</i>	<i>Back</i>
<i>High</i>	/i/ and /i:/		/u/ and /u:/
<i>Low</i>		/a/ and /a:/	

In addition, when /a/ was followed by /w/ or /y/, the diphthongs /au/ and /ai/ were formed.

### 3.3.2 Middle Aramaic

A number of vowel changes took place between the Proto-Aramaic and the Middle Aramaic periods; providing a relative chronology, much less an absolute chronology, of these changes is problematic. Questions of chronology aside, these changes can be divided into three groups:

1. Changes which did not affect the system of vowel phonemes, such as the shift of /a/ to /i/ (“attenuation”) in some closed syllables.
2. Changes which occurred in every dialect of Aramaic:
  - (i) Stressed /i/ and /u/ were lowered, and perhaps lengthened, to /e/ or /e:/ and /o/ or /o:/.
  - (ii) In all dialects, but differing from dialect to dialect as to the number and the specification of environments, /ai/ became /e:/ (or possibly /ei/) and /au/ became /o:/ (or possibly /ou/).
  - (iii) In the first open syllable prior to the stressed syllable and in alternating syllables prior to that, short vowels were lost. In positions where the complete loss of the vowel would have produced an unacceptable consonant cluster, the vowel reduced to the neutral mid-vowel [ə]. Because the presence of this vowel is entirely predictable from syllable structure, it is not analyzed as phonemic.
  - (iv) Quantity ceased to be phonemic.
3. Changes which apparently occurred in some dialects, but not others:
  - (i) The low vowel /a:/ was rounded and raised to /ɔ/.
  - (ii) Unstressed /u/ was lowered to /ɔ/ in some environments.
  - (iii) Unstressed /i/ was lowered to /ɛ/ in some environments.
  - (iv) Unstressed /a/ was raised to /ɛ/ in some environments.

A dialect in which all of these changes occurred would have the vowel system of (4), along with the diphthongs /ai/ (or /ei/) and /au/ (or /ou/), if they had been retained in any environments:

(4)	<i>Front</i>	<i>Central</i>	<i>Back</i>
<i>High</i>	/i/		/u/
<i>Mid</i>	/e/		/o/
	/ɛ/		/ɔ/
<i>Low</i>		/a/	

A dialect in which only the first two sets of changes occurred would have the same system but without the vowels /ɛ/ and /ɔ/.

### 3.3.3 Late Aramaic

At the beginning of the Modern Aramaic period, four sets of diacritics were independently developed to represent Aramaic vowels fully. These sets of diacritics represent the phonemic distinctions relevant to four dialects of Late Aramaic. The distinctions indicated by these systems are qualitative, not quantitative, indicating that vowel quantity was not phonemic by this time. In all of these systems, the pronunciation of the low vowel(s) is/are uncertain and so two options are usually given. Also indicated in (5)–(8) are the standard transliteration equivalents in the writing system.

#### 3.3.3.1 The Tiberian system

(5)	Front	Central	Back
High	/i/ = <i> and <î>		/u/ = <u> and <û>
Mid	/e/ = <ē> and <ê> /ɛ/ = <e>		/o/ = <ō> and <ô> /ɔ/ = <o> and <ā>
Low		/æ/ or /a/ = <a>	

The phonemic status of the /ɛ/ vowel is uncertain, because its alternation with other vowels in the system is nearly always predictable. If /ɛ/ is not a phoneme, then this system would be equivalent to the Babylonian system (see §3.3.3.2).

The Tiberian system also contains four additional symbols for vowels (see §2.2.2), all of which represent vowels of very brief duration: the neutral mid vowel /ə/, and very brief pronunciations of /ɛ/, /ɔ/, and /a/. Diachronically, these vowels are the remnants of short vowels which were reduced in certain syllables (see §3.3.2). They are only retained in positions where the complete loss of the vowel would produce an unacceptable consonant cluster and so they represent a context-dependent phonetic (rather than a phonemic) phenomenon.

#### 3.3.3.2 The Babylonian system

(6)	Front	Central	Back
High	/i/ = <i> and <î>		/u/ = <u> and <û>
Mid	/e/ = <ē> and <ê>		/o/ = <ō> and <ô>
Low	/æ/ (or /a/) = <a>		/a/ (or /ɔ/) = <ā>

This system is essentially equivalent to the Tiberian system, but without /ɛ/. It is probable that /ɛ/ is absent in this dialect because it never developed from /i/ and /a/, rather than because it first developed and then was subsequently lost. This system also contains a symbol for the neutral mid vowel /ə/ but, unlike the Tiberian system, the diacritic is not ambiguous (i.e., it does not also represent the absence of a vowel; see §2.2.2).

#### 3.3.3.3 The Nestorian system

(7)	Front	Central	Back
High	/i/ = <i> and <î>		/u/ = <u> and <û>
Mid	/e/ = <ē> and <ê> /ɛ/ = <e>		/o/ = <ō> and <ô> /ɔ/ = <ā>
Low		/æ/ or /a/ = <a>	

This system is essentially the same as the Tiberian and the Middle Aramaic system, though the /ɛ/ vowel is much more common and is certainly a phoneme in this system.



### 3.3.3.4 The Jacobite system

(8)	Front	Central	Back
High	/i/ = <î> and <ê>		/u/ = <û> and <ô>
Mid	/e/ = <e>		/o/ = <â>
Low		/a/ = <a>	

This system has the smallest of all inventories and is a result of two changes from the Middle Aramaic (= Nestorian) system: (i) the raising of /e/ and /o/ to /i/ and /u/ respectively; and (ii) the raising of /ɛ/ and /ɔ/ to /e/ and /o/ respectively.

## 3.4 Syllable structure

Aramaic has both closed (CVC) and open (CV) syllables. During the time that vowel quantity was phonemic in Aramaic, a closed syllable could not contain a long vowel, whereas an open syllable could contain either a long or a short vowel. After vowel quantity was no longer phonemic, such restrictions were no longer relevant to the phonemic system, although vowels in closed and open syllables very likely differed phonetically in quantity.

The only apparent restriction on vowel quality in Aramaic syllables occurs in connection with the consonants /ʔ/, /ʕ/, /ħ/, /h/, and /r/. At an early stage in Aramaic, a short high vowel preceding these consonants became /a/. A preceding long high vowel retained its quality, but, in some dialects, /a/ was inserted between the high vowel and the consonant.

## 3.5 Stress

There is one primary stressed syllable in each Aramaic word (with the exception of some particles; see §§4.6, 4.7.4, and 4.8.1). In Proto-Aramaic, words having a final closed syllable were stressed on that syllable; and words having a final open syllable were stressed on the penultimate syllable, regardless of the length of the word-final vowel. At a very early stage, word-final short vowels were either lost or lengthened and so the stressed, open penultimate syllable of words with a final short vowel became the final stressed, closed syllable. Stress remained on this syllable and the rules regarding stress were not altered. These rules remain unaltered throughout most of the history of Aramaic, though in some dialects of Late Aramaic, stress shifted from the final syllable to the penultimate syllable in some or all words which had a closed final syllable.

## 3.6 Phonological processes

### 3.6.1 Sibilant metathesis

In verb forms in which a /t/ is prefixed (see §4.4.1) to a root which begins with a sibilant, the sibilant and the /t/ undergo metathesis: for example, /ts/ → /st/ and /tš/ → /št/. If the sibilant is voiced /z/ or pharyngealized /sʕ/, /t/ also undergoes partial assimilation: /tz/ → /zd/ and /tsʕ/ → /sʕtʕ/.

### 3.6.2 Assimilation of /t/

In verb forms in which a /t/ is prefixed (see §4.4.1) to a root which begins with /d/ or /tʕ/, the /t/ completely assimilates to this consonant. This assimilation also takes place in a few roots whose first consonant is a labial – /b/, /p/, and /m/ – or the dental/alveolar /n/.

### 3.6.3 Assimilation and dissimilation of /n/

Historically, the phoneme /n/ completely assimilates to a following consonant when no vowel intervenes between the two: \*nC → CC. During and after the Imperial Aramaic period, some geminated (lengthened) consonants dissimilate to /n/ plus consonant, CC → nC, even in cases in which no /n/ was present historically. This dissimilation is the result of Akkadian influence and appears more commonly in the eastern dialects.

### 3.6.4 Dissimilation of pharyngealized consonants

In some Aramaic texts, words which have roots that historically contain two pharyngealized consonants show dissimilation of one of the consonants to its nonpharyngealized counterpart. In a few Old Aramaic texts, progressive dissimilation is shown: for example, *qtl* (i.e., /kʔtʔl/) → *qtl*. In some Imperial Aramaic texts the dissimilation is regressive: for example, *qtl* → *ktl* and *qṣʔ* (i.e., /kʔsʔ/) → *kṣʔ*. These dissimilations may have been the result of Akkadian influence, which attests similar dissimilations.

### 3.6.5 Elimination of consonant clusters

At various stages of Aramaic, phonotactically impermissible consonant clusters were eliminated in various ways.

#### 3.6.5.1 Anaptyxis

In Proto-Aramaic, all singular nouns ended in a short vowel, marking case (see §4.2.2). When this final short vowel was lost, some nouns then ended in a cluster of two consonants: as in \*/málku/ → /málk/. In order to eliminate this cluster, a short anaptyctic vowel (usually /i/, sometimes /a/) was inserted between the two consonants: /málk/ → /málik/. Stress then shifted to this vowel from the preceding vowel: /málik/ → /malík/. At a later stage, the vowel of the initial syllable was lost and the anaptyctic vowel was lowered (see §3.3.2): /malík/ → /mlík/ → /mlék/.

#### 3.6.5.2 Schwa

The loss of short vowels in some open syllables (see §3.3.2) created the possibility of consonant clusters at the beginning and in the middle of words. In positions where the complete loss of the vowel would have produced an unacceptable consonant cluster, the cluster was avoided by reducing the short vowel to the neutral mid-vowel /ə/.

#### 3.6.5.3 Prothetic aleph

When a word begins with a cluster of two consonants, sometimes the syllable /ʔa/ or /ʔε/ is prefixed to it: for example, the word /dmɔ/ is sometimes pronounced /ʔadmɔ/.

## 4. MORPHOLOGY

### 4.1 Morphological type

Aramaic is a language of the fusional type in which morphemes are unsegmentable units which represent multiple kinds of semantic information (e.g., gender and number). On the basis of morphological criteria alone, Aramaic words can be divided into three categories: (i) nouns, (ii) verbs, and (iii) uninflected words. The final category includes a variety of

words such as adverbs (see §4.5), prepositions (see §4.6), particles (see §4.7), conjunctions (see §4.8), and interjections (see §4.9). As the name suggests, words in this category are distinguished from words in the first two categories by the absence of inflection. Words in the first two categories can be distinguished from each other by differences in the categories for which they are inflected and by the inflectional material itself.

## 4.2 Nominal morphology

Under this heading are included not only nouns and adjectives, but participles as well.

### 4.2.1 Word formation

Excluding inflectional material, all native Aramaic nouns, adjectives, and participles (as well as verbs; see §4.4.1) consist of (i) a two-, three-, or four-consonant root; (ii) a vowel pattern or ablaut; and, optionally, (iii) one or more prefixed, suffixed, or infixes consonant(s). Multiple combinations of these elements exist in the lexicon of native Aramaic words, and earlier and later patterns can be identified within the lexicon.

In Old and Imperial Aramaic, the patterns found are ones that are common to the other Semitic languages. Many patterns are characterized by differences in ablaut only: for example, *qal*, *qāl*, *qil*, *qall*, *qitl*, *qutl*, *qatal*, *qatāl*, *qatīl*, and *qātīl*. Additional patterns are characterized by the gemination (lengthening) of the second root consonant: for example, *qattal*, *qittal*, *qattīl*, and *qattāl*. Still others display prefixation – for example, *maqtal*, *maqtīl*, *maqtāl*, *taqtīl*, and *taqtūl*; or suffixation – for example, *qatlūt*, *qutlīt*, and *qitlāy*; or reduplication – for example, *qatlal* and *qataltāl*. The semantics of some of these patterns or of individual suffixes is clear and distinct: for example, the pattern *qattāl* indicates a profession (*nomen professionalis*), the pattern *qatīl* is that of the passive participle of the Pəʿal stem; and the suffix *-āy* (the *nisbe* suffix) indicates the name of an ethnic group.

In Late Aramaic, the use of suffixes increased, apparently as a result of two historical factors. First, the loss of short vowels in open syllables prior to the stressed syllable often eliminated the single vowel which distinguished one vowel pattern from another. So, the use of suffixes may have been increased to compensate for the loss of distinct vowel patterns. Second, the contact of Aramaic with Indo-European languages, especially Greek, may have increased the use of suffixes since the morphology of those languages largely involves suffixation rather than differences in vowel patterns.

One notable nonsuffixing pattern that developed in the Middle or Late Aramaic period is the *qātōl* pattern which indicates an agent noun (*nomen agentis*). The older agent noun pattern, *qātēl* (< *qātīl*), is also the pattern of the active participle of the Pəʿal stem, and by the Middle Aramaic period the participle came to be used almost exclusively as a verbal form, and so a new, purely nominal, agent noun form was developed.

### 4.2.2 Inflectional categories

Nouns, adjectives, and participles are inflected for gender, number, and state. There are no case distinctions in any extant dialect of Aramaic, though such distinctions did exist in Proto-Aramaic. There are also no comparative or superlative forms of adjectives at any stage of the language. There are two genders, masculine and feminine, and nouns can be distinguished from adjectives and participles in that nouns have inherent gender whereas

adjectives and participles do not. There are two numbers, singular and plural, and although a few words retain an ancient dual form, there is no productive dual in Aramaic. There are three states: absolute, construct, and emphatic. The absolute and the emphatic states of a noun are free forms and the construct state is a bound form. In earlier stages of Aramaic, the absolute state represented an indefinite noun, the emphatic state represented a definite noun and the construct state represented a noun the definiteness of which was determined by the noun to which it was bound. In Late Aramaic, the absolute state was almost entirely lost and the emphatic state became used for both definite and indefinite nouns. Definiteness was then determined contextually or by the use of the numeral “one” as a kind of indefinite article. At this stage, the construct state was retained only in frozen forms and was not productive, with the exception of a few words such as *br* “son-of.” However, adjectives and participles retained the absolute state throughout all periods because of their use as predicates to form clauses (see §5.2.1).

The transliterations of the written forms of the inflectional suffixes for nouns, adjectives, and participles are presented in (9). The forms of each suffix are represented both with and without vowel diacritics (see §§2.1, 2.2.2). The symbol  $\emptyset$  represents the absence of an inflectional suffix. The letters  $\text{'}^{\text{h}}$  and  $\text{'}^{\text{h}}$  are *matres lectionis* (see § 2.2.1). On the phonemic values of the transliteration of vowel diacritics see §3.3.3:

	<i>Masculine</i>		<i>Feminine</i>	
	<i>Singular</i>	<i>Plural</i>	<i>Singular</i>	<i>Plural</i>
<i>Absolute</i>	$\emptyset$	-yn (= -în)	-h (= -ā)	-n (= -ān)
<i>Construct</i>	$\emptyset$	-y (= -ay or -ê)	-t (= -at)	-t (= -āt)
<i>Emphatic</i>	- $\text{'}^{\text{h}}$ (= -ā)	-y $\text{'}$ (= -ayyā)	-t $\text{'}$ (= -tā)	-t $\text{'}$ (= -ātā)

Several points should be noted regarding these inflectional suffixes:

1. The masculine singular emphatic is also sometimes attested as *-h*.
2. The feminine singular absolute, in some dialects, is also rarely attested as *- $\text{'}^{\text{h}}$* . In Syriac, it is consistently attested as *- $\text{'}^{\text{h}}$* .
3. The *y* of the masculine plural absolute is a *mater lectionis* and so is sometimes omitted in writing, especially in early texts.
4. The *y* of the masculine plural construct is either a consonant, representing the diphthong /ai/, or a *mater lectionis* representing /e:/ which had developed from /ai/ in some dialects.
5. The Samʿal dialect of Old Aramaic attests *-t* (= -āt) as the feminine plural absolute form, the usual form in Canaanite dialects.
6. In eastern dialects of Middle and Late Aramaic, the masculine plural emphatic appears as *- $\text{'}^{\text{h}}$*  or *-y $\text{'}$*  (= -ê), perhaps under Akkadian influence.

Many Aramaic nouns, adjectives, and participles show two (or more) vowel patterns which alternate depending on the phonological form of the inflectional material. These multiple patterns are the result of the phonological changes that took place during the history of Aramaic. However, explaining these alternating patterns synchronically requires a rather complex set of rules and will not be attempted here. In two groups of nouns, adjectives, and participles (those with a final consonant which was historically /w/ or /y/), these phonological changes also produced changes in the forms of some of the inflectional suffixes. Nouns, adjectives, and participles with a final consonant /w/ developed the vowel /u/ or /o/ in both the masculine singular absolute and construct as well as in the three

feminine singular forms (the /w/ remained a consonant in the other seven forms). In the feminine singular absolute and construct forms, this vowel replaced the vowel of the inflectional suffix.

Nouns, adjectives, and participles with a final consonant /y/ show even more changes. The inflectional suffixes for these words are given in (10):

(10)	<i>Masculine</i>		<i>Feminine</i>	
	<i>Singular</i>	<i>Plural</i>	<i>Singular</i>	<i>Plural</i>
<i>Absolute</i>	-ʾ or -y (= -ê)	-yn (= -ayin or -ên) or -n (= -an)	-y (= -î) or -yʾ (= -yā)	-yn (= -yān)
<i>Construct</i>	-ʾ or -y (= -ê)	-y (= -ay or -ê)	-yt (= -ît or -yat)	-yt (= -yāt)
<i>Emphatic</i>	-yʾ (= -yā)	-yʾ (= -ayyā or -yê)	-ytʾ (= -îtā)	-ytʾ (= -yātā)

In the masculine singular emphatic and the feminine plural forms, /y/ remains a consonant and the inflectional suffix is standard. In the other forms, /y/ generally becomes a vowel, sometimes fusing with the inflectional ending, although in some nouns it remains a consonant and the suffix is standard.

### 4.3 Pronouns

#### 4.3.1 Personal pronouns

Personal pronouns occur in both independent and bound (i.e., enclitic) forms.

##### 4.3.1.1 Independent personal pronouns

Independent forms of the personal pronouns vary slightly from dialect to dialect and from period to period. All but the rarest of forms are listed in (11):

(11)	<i>Singular</i>	<i>Plural</i>
<i>1st common</i>	ʾnh, ʾnʾ	ʾnhn, ʾnhnn, ʾnhnʾ, ʾnhnh, nhnʾ, hnn, ʾnn
<i>2nd masculine</i>	ʾnt, ʾt, ʾnth, ʾth	ʾntm, ʾntwn, ʾtwn
<i>2nd feminine</i>	ʾnty, ʾnt, ʾt, ʾty	ʾntn, ʾntyn, ʾtyn
<i>3rd masculine</i>	hʾ, hwʾ, hw	hm, hwm, hmw, hmwn, ʾwn, hnwn, ʾynwn, hynwn
<i>3rd feminine</i>	hʾ, hyʾ, hy	ʾnyn, hnyn, ʾynyn, hynyn

The first- and second-person pronouns all have an initial ʾn, and the remainder of each form generally resembles the inflectional suffix of the perfect verb (§4.4.2.1). Forms written without *n* are those which have undergone assimilation of /n/ to /t/ (see §3.6.3). The third-person singular forms have an initial *h*, and the plural forms have an initial *h* or ʾ. The masculine has a back vowel /o/ or /u/, and the feminine has a front vowel /i/ or /e/. Most of the spelling differences reflect the presence or absence of *matres lectionis*, though some reflect historical developments. Of particular note is the replacement of the earlier final /m/ of the second and third masculine plural forms with the later /n/ under the influence of the feminine forms.

In the Samʿal dialect of Old Aramaic the first common singular is the Canaanite ʾnk(y).

##### 4.3.1.2 Bound personal pronouns

These forms are used for the possessor of a noun, the object of a preposition, the subject or object of an infinitive, or the object of a verb and they vary depending on the type of word to which they are suffixed.

The bound forms that are suffixed to nouns, prepositions, particles, and infinitives can be divided into two sets: Set I is used with masculine singular nouns, all feminine nouns, infinitives, and some prepositions; Set II is used with masculine plural nouns, the other prepositions, and the existential particles:

(12) Bound pronouns suffixed to nouns, prepositions, particles, and infinitives

	Set I		Set II	
	Singular	Plural	Singular	Plural
1st common	-y	-n <sup>3</sup> , -n	-y	-yn, -yn <sup>3</sup>
2nd masculine	-k	-km, -kwn	-yk	-ykm, ykwn
2nd feminine	-ky, -yk	-kn, -kyn	-yky	-ykn, -ykyn
3rd masculine	-h, -yh	-hm, -hwm, -hwn	-wh, -why, -wy	-yhm, -yhwm, -yhwn
3rd feminine	-h	-hyn	-yh	-yhn, -yhyn

Note the following:

1. The first common singular suffix occurring on the infinitive is more commonly *-ny* than *-y*. In Syriac, the infinitive also occurs with alternate forms of the third masculine singular (*-ywhy*) and third feminine singular (*-yh*).
2. In Set I, the third masculine singular *-yh* and the second feminine singular *-yk* reflect the presence of an internal *mater lectionis* in Late Aramaic texts.
3. The differences in the second- and third-person plural forms of both sets are a result of the presence or absence of *matres lectionis* and the shift of final /m/ to /n/ in the masculine forms. In Samaritan Aramaic, the third plural forms of both sets are also attested without *-h*. In Jewish Babylonian Aramaic, the second- and third-person plural forms of both sets are also attested without the final *-n*.
4. In Sets I and II, the first common plural form without <sup>3</sup> reflects the absence of a *mater lectionis* in earlier texts and the absence of a final vowel in later texts.
5. In Jewish Palestinian Aramaic, the third feminine singular, second masculine singular, second feminine singular, and the first common plural forms in Set II are also attested without the initial *y*, suggesting a shift of /ai/ to /a/. The first common singular, first common plural, and third feminine singular forms in Set II are also attested as *-ʿy*, *-ynn*, and *-yh<sup>3</sup>* respectively, in this dialect.
6. The second feminine singular form of Sets I and II is also written without the final *y* in Jewish Palestinian Aramaic, suggesting the loss of the final vowel, and in Syriac the *y* is written but not pronounced.
7. The third masculine singular form *-wh* of Set II probably reflects the absence of a *mater lectionis* in earlier texts. The *-wy* form reflects the loss of the intervocalic /h/ in later texts. The Sam<sup>2</sup>al dialect of Old Aramaic attests *-yh*, suggesting the diphthong /ai/ rather than /au/. This diphthong is the historically earlier vowel which became /au/ in all other dialects.

The bound forms of the pronouns that are attached to verbs will vary depending on three factors: (i) the tense of the verb; (ii) the phonological form of the verb; and (iii) the dialect. Most variation is a result of the phonological form of the verb rather than verb tense, although the forms used with the imperfect frequently show an additional *-n-* (= /inn/). In some dialects of Late Aramaic, this additional *-n-* is also found in forms that are used with the perfect. Other differences in bound pronouns across dialects tend to reflect broader phonological changes in the language, such as the loss of word-final vowels or consonants.

Bound forms of the third-person plural pronouns are generally not suffixed to verbs, although there are attested forms in Old Aramaic, particularly in the Samaritan dialect, and in Jewish Babylonian Aramaic and Jewish Palestinian Aramaic. More commonly, an independent form of the pronoun is used instead. However, in some dialects, these forms are not stressed and so they are phonologically enclitic to the preceding verb form, even though they are written as separate words.

In (13)–(15), *y*, *w*, and *'* are all *matres lectionis*, but *h* represents a true consonant:

**(13) Bound pronouns suffixed to verbs: perfect tense**

	<i>Singular</i>	<i>Plural</i>
<i>1st common</i>	-ny, -y, -n	-n, -nn, -n'
<i>2nd masculine</i>	-k	-kn, -kwn
<i>2nd feminine</i>	-ky	-kyn
<i>3rd masculine</i>	-h, -yh, -hy, -yhy	
<i>3rd feminine</i>	-h, -h'	

Note the following:

1. The first common singular form *-y* is attested in Jewish Palestinian Aramaic and Samaritan Aramaic. In Jewish Babylonian Aramaic, the form *-n* is attested and it represents the loss of the final vowel. The final vowel is also lost in Syriac, but the form is still written *-ny*.
2. Syriac also attests the third masculine singular forms *-why* and *-ywhy*.
3. The first common plural form *-n* represents the loss of the final vowel, and the form *-nn* represents the additional *-n*. Both forms are only attested in Late Aramaic dialects.
4. Jewish Babylonian Aramaic also attests a second masculine plural form *-kw*, as well as second masculine singular (*-nk*), second masculine plural (*-nkw*), and third feminine singular (*-nh*) forms with the additional *-n*.
5. Old Aramaic attests the third masculine plural forms *-hm* and *-hmw*.
6. Jewish Babylonian Aramaic attests the third masculine plural forms *-ynwn*, *-ynhw* and the third feminine plural forms *-nhy* and *-ynhy*. Samaritan Aramaic attests the third masculine plural form *-wn* and third feminine plural form *-yn*.

**(14) Bound pronouns suffixed to verbs: imperfect tense**

	<i>Singular</i>	<i>Plural</i>
<i>1st common</i>	-n, -ny, -nny	-n, -nn
<i>2nd masculine</i>	-k, -nk, -ynk	-kwn, -nkwn
<i>2nd feminine</i>	-ky, -yk	-kyn, -nkyn
<i>3rd masculine</i>	-h, -hy, -nh, -nhy	
<i>3rd feminine</i>	-h, -nh	

Note the following:

1. In Old and Imperial Aramaic, forms with and without the additional *-n* are attested. In Jewish Palestinian Aramaic, Jewish Babylonian Aramaic, and Samaritan Aramaic, the forms with *-n* are much more commonly attested than the forms without *-n*. In Syriac, the forms with *-n* are not attested at all.

2. In Old Aramaic, the first common singular form *-n* is pronounced with a final vowel but is written without a *mater lectionis*. In Jewish Babylonian Aramaic, the form *-n* represents the loss of the final vowel. In Syriac, the final vowel is also lost, but the form is still written *-ny*.
3. No second feminine singular forms with additional *-n* happen to be attested in the extant texts. The form *-ky* is pronounced with a final vowel in Old and Imperial Aramaic, but in Jewish Babylonian Aramaic and Syriac the final vowel is lost, though in Syriac the form is still written *-ky*.
4. The third masculine singular forms *-hy* and *-nhy* are only found in Old and/or Imperial Aramaic.
5. Syriac also attests the third masculine singular forms *-yhy* and *-ywhy* and the third feminine singular form *-yh*.
6. Jewish Babylonian Aramaic attests the third masculine singular forms *-yh* and *-ynyh*, the third feminine singular form *-ynh*, and the second masculine plural form *-ynkw*.
7. In Jewish Palestinian Aramaic, the first common plural form *-nn*<sup>?</sup> is also attested.
8. Old Aramaic attests the third masculine plural forms *-hm* and *-hmnw*.
9. Jewish Babylonian Aramaic attests the third masculine plural forms *-ynwn*, *-ynhw* and the third feminine plural form *-ynhy*. Samaritan Aramaic and Jewish Palestinian Aramaic attest the third masculine plural form *-nwn*. Jewish Palestinian Aramaic also attests the third feminine plural form *-nyn*.

(15) Bound pronouns suffixed to verbs: imperative

	<i>Singular</i>	<i>Plural</i>
<i>1st common</i>	<i>-ny</i> , <i>-n</i> , <i>-yny</i> , <i>-yn</i> , <i>-y</i>	<i>-n</i> , <i>-yn</i> , <i>-n</i> <sup>?</sup> , <i>-yn</i> <sup>?</sup> , <i>-nn</i>
<i>3rd masculine</i>	<i>-h</i> , <i>-hy</i> , <i>-yh</i> , <i>-why</i> , <i>-yhy</i>	
<i>3rd feminine</i>	<i>-h</i> , <i>-yh</i> , <i>-h</i> <sup>?</sup>	

Note the following:

1. The first common singular form *-ny* is attested in all dialects. The first common singular form *-y* is only attested in Jewish Palestinian Aramaic and Samaritan Aramaic. In Jewish Babylonian Aramaic, the forms *-yn* and *-n* are attested in addition to *-ny* and represent the loss of the final vowel. In Syriac, the forms *-ny* and *-yny* are attested, but the *y* is not pronounced.
2. The third masculine singular form *-h* is attested in Old Aramaic, Imperial Aramaic, Jewish Babylonian Aramaic, and Samaritan Aramaic. This form is also written with a *mater lectionis* as *-yh* in Jewish Babylonian Aramaic and Jewish Palestinian Aramaic. The form *-hy* is attested in Old Aramaic, Imperial Aramaic, Jewish Palestinian Aramaic, and Syriac, although in Syriac the *h* is not pronounced. The forms *-why* and *-yhy* are only attested in Syriac and the *h* is not pronounced.
3. Only Syriac attests the third feminine singular form *-yh* and only Jewish Palestinian Aramaic attests the third feminine singular form *-h*<sup>?</sup>.
4. First common plural bound pronouns are only attested in Late Aramaic. Syriac attests *-n* and *-yn*. Jewish Palestinian Aramaic attests *-n* and *-n*<sup>?</sup>. Samaritan Aramaic attests *-n* and *-nn*. Jewish Babylonian Aramaic attests *-yn*<sup>?</sup>.
5. Jewish Babylonian Aramaic attests the third masculine plural form *-nhw* and the third feminine plural form *-nhy*. Jewish Palestinian Aramaic attests the third masculine plural form *-nwn* and the third feminine plural form *-nyn*.



In Late Aramaic, as a result of the use of the participle as a verb form, shortened forms of the first- and second-person independent pronouns became suffixed to the participle to indicate the subject. In Syriac and Jewish Babylonian Aramaic, third-person forms developed alongside the first- and second-person forms, and all of these forms are commonly used in a variety of nonverbal clauses, not just those with participles. In these uses, the pronouns are written as separate words, but are phonologically enclitic to the preceding word (see §5.2.1).

### 4.3.2 Demonstrative pronouns

#### 4.3.2.1 Near demonstratives

In Old, Imperial, and Middle Aramaic, the singular forms of the near demonstratives are characterized by an initial *z* or *d* (= historical /ð/; see §3.2.2) followed, in the masculine forms, by *n* and a final *mater lectionis* -*h* or -*ʾ*. The forms are as follows: masculine singular *znh*, *znʾ*, *dnh*, *dnʾ* and feminine singular *zʾ*, *zh*, *dʾ*. In Middle Aramaic, the masculine singular forms *dn* and *zn* are also attested, suggesting that the final vowel was being lost in this period. Gender is not distinguished in the plural forms of the near demonstrative. These forms are all characterized by an initial ʾ. They are ʾl, ʾh, ʾln.

In the Late Aramaic period, the near demonstratives are often attested with an initial *h*. This *h* generally replaces the initial *d* of the singular forms and the initial ʾ of the plural form. However, some singular forms in some dialects attest both the *h* and the *d*. For example, Syriac attests masculine singular *hn* and *hnʾ*, feminine singular *hdʾ*, and plural *hlyn*. Jewish Palestinian Aramaic attests masculine singular *dyn*, *dnʾ*, *hyn* and *hn*, feminine singular *dʾ*, and plural *hlyn* and *ʾlyn*. Jewish Babylonian Aramaic attests many forms including masculine singular *dyn* and *hdyn*, feminine singular *hdʾ* and *hʾ*, and plural *ʾlyn* and *hlyn*. Samaritan Aramaic attests masculine singular *dn*, feminine singular *dh*, and plural *hlyn* and *ʾlyn*.

#### 4.3.2.2 Far demonstratives

In Old, Imperial, and Middle Aramaic, the far demonstratives are like the near demonstratives in that the singular forms are characterized by an initial *z* or *d* and the plural forms by an initial ʾ, but, unlike the near demonstratives, this initial element is followed by *k*. The forms are as follows: masculine singular *znk*, *zk*, *dk*; feminine singular *zk*, *zkʾ*, *dk*, *zky*, *dky*; and plural *ʾlk*, *ʾlky*. In addition to these forms, there are sporadic attestations of the third-person independent personal pronouns being used as demonstratives. This usage is common in the Canaanite dialects, and these attestations are generally found in Aramaic dialects influenced by Canaanite such as the Samʾal dialect of Old Aramaic and some Middle Aramaic dialects influenced by Hebrew.

In the Late Aramaic period, the third-person independent personal pronouns become more commonly used as far demonstratives, although in most dialects they do not displace the earlier forms, but are simply attested alongside them. In Syriac, the earlier forms are lost entirely and the far demonstratives are distinguished from the personal pronouns by the vowel of the first syllable of the singular forms and by the presence of *h* rather than ʾ as the initial consonant of the plural forms.

### 4.3.3 Reflexive pronouns

The equivalent of a reflexive pronoun is expressed by suffixing a bound form of a personal pronoun to *npš* “life, soul” or *grm* “bone.”

#### 4.3.4 Possessive pronouns

Possessive pronouns are usually expressed by bound forms of the personal pronouns, but in Middle and Late Aramaic the particle *z/dyl* (= particle *z/dy* + preposition *l*) with a suffixed bound form became used as a possessive pronoun.

### 4.4 Verbal morphology

#### 4.4.1 Word formation

Excluding inflectional material, all native Aramaic verbs (as well as nouns; see §4.2.1) consist of (i) a two-, three-, or four-consonant root; (ii) a vowel pattern or ablaut; and, optionally, (iii) one or more prefixed or infixes. The root provides the primary semantic value of the verb form. The other two elements (ii and iii) provide semantic distinctions of voice, causation, and so forth; and variations in these two elements define a system of verbal stems or conjugations which are morpho-semantically related to each other. Of these two elements, the vowel pattern is less important than the additional consonant(s) since vowels frequently change from one inflected form to another. The distinctions between the stems are generally, but not always, maintained despite these vowel changes. Furthermore, some of these vowel patterns differ slightly from one dialect to another. For these reasons, the vowel patterns will not be treated in the following discussion.

##### 4.4.1.1 Major verb stems

Numerous verb stems exist in Aramaic, but there are only six primary stems. They can be defined morphologically as follows, assuming in each case a three-consonant root.

1. *Pə'al*: This stem is the most frequently attested of the six. It is also the simplest stem morphologically, characterized by the absence of any consonants other than the root consonants. For this reason, it is considered the basic stem. This stem attests multiple vowel patterns in both of the primary finite forms of the verb, and it is the only stem with multiple vowel patterns.
2. *ʾEthpə'al* or *ʾIthpə'al*: This stem is characterized by the presence of a prefixed *ʾt*-. Historically, this prefix is *ht*-, and forms with *ht*- are sporadically attested in all periods.
3. *Pa'al*: This stem is characterized by the gemination (lengthening) of the second root consonant.
4. *ʾEthpa'al* or *ʾIthpa'al*: This stem is characterized by the gemination (lengthening) of the second root consonant and by a prefixed *ʾt*-. Historically, this prefix is *ht*-, and forms with *ht*- are sporadically attested in all periods.
5. *Haph'al* or *ʾAph'al*: This stem is characterized by the prefixation of the consonant *h*- or the consonant *ʾ*-. The forms with *h*- are historically earlier than the forms with *ʾ*- and had almost entirely disappeared by the Middle Aramaic period, though a few forms with *h*- survive into the Late Aramaic period.
6. *ʾEttaph'al* or *ʾIttaph'al*: This stem is characterized by a prefixed *ʾtt*-. The second *t* is historically the *h*- or *ʾ*- of the *Haph'al*/*ʾAph'al* which has been assimilated to the preceding *t*.

Certain modifications of these stems occur when there are two or four root consonants rather than three. Verbs with four root consonants only have forms corresponding to the

*Paʿʿel* and the *ʾEthpaʿʿal/ʾIthpaʿʿal* stems, the two middle root consonants taking the place of the geminated (lengthened) second root consonant of a verb with three root consonants. Verbs with two root consonants develop a middle root consonant *-y-* in the *Paʿʿel* and the *ʾEthpaʿʿal/ʾIthpaʿʿal*, and the distinction between the *ʾEthpəʿel/ʾIthpəʿel* and the *ʾEttaphʿal/ʾIttaphʿal* forms is completely lost, with the retention of the latter forms only.

#### 4.4.1.2 Voice and other semantic distinctions

This system of stems expresses a variety of semantic distinctions, and a variety of relationships exist between the stems. One of the primary distinctions is that of voice. The *Pəʿal*, the *Paʿʿel*, and the *Haphʿel/ʾAphʿel* stems all express the active voice. The three stems with prefixed *ʾt-* all express the passive voice. Each of the passive stems is directly related only to its morphologically similar active stem, and the relationships of the passive stems to one another simply mirror the relationships of the active stems to one another. In Proto-Aramaic, it is likely that the stems with prefixed *ʾt-* were reflexive, but in the extant dialects of Aramaic, reflexive uses of these stems are only sporadically attested.

The relationships of the active stems to one another are more complex. The *Paʿʿel* and the *Haphʿel/ʾAphʿel* are directly related to the *Pəʿal*, but not to each other. The *Haphʿel/ʾAphʿel* expresses causation. A *Haphʿel/ʾAphʿel* verb of a particular root is usually the causative of the *Pəʿal* verb of that same root. For example, the *Haphʿel/ʾAphʿel* verb *hkšl/ʾkšl* “to trip someone up” is the causative of the *Pəʿal* verb *kšl* “to stumble.” There are, however, a number of *Haphʿel/ʾAphʿel* verbs, some of which are denominative, for which there is no corresponding *Pəʿal* verb or which do not express causation.

The relationship of the *Paʿʿel* stem to the *Pəʿal* stem varies depending on the semantic class into which the verb in the *Pəʿal* stem falls. The verbs in the *Pəʿal* stem exhibit a number of semantic distinctions, the two most important of which are (i) the distinction between stative verbs and active verbs, and (ii) the distinction between one-place predicates (usually syntactically intransitive) and two-place predicates (usually syntactically transitive). As a general rule, to which there are exceptions, if the *Pəʿal* verb is stative and/or a one-place predicate, the *Paʿʿel* verb of that same root is “factitive” (i.e., causative). If there is a *Haphʿel/ʾAphʿel* verb of that same root, it is roughly synonymous with the *Paʿʿel* verb or there is a lexically idiosyncratic difference in meaning; for example, *Pəʿal qrb* “to come near,” *Paʿʿel qrb* “to bring near, to offer up,” *Haphʿel/ʾAphʿel hqrb/ʾqrb* “to bring near,” or, in some dialects only, “to fight.” If the *Pəʿal* verb is a two-place predicate, the *Paʿʿel* verb of that same root will be “intensive,” though in some cases, the two verbs are synonymous or there is a lexically idiosyncratic difference in meaning; for example, *Pəʿal zmr* “to sing,” *Paʿʿel zmr* “to sing.” There are, furthermore, numerous *Paʿʿel* verbs, many of which have four root consonants and for which there is no corresponding *Pəʿal* verb.

By the Late Aramaic period, the relationships between the stems had broken down through the process of lexicalization. Although some of the relationships still held between individual verbs of the same root, in many cases they did not. This breakdown was aided by the similarity in meaning of some pairs of verbs and, in the case of the *ʾEthpəʿel/ʾIthpəʿel* and the *ʾEthpaʿʿal/ʾIthpaʿʿal*, by their increasing morphological similarity due to vowel changes in the language.

#### 4.4.1.3 Minor stems

In Old Aramaic, it is possible that a set of passive stems existed, corresponding to each of the three major active stems, and differing from them in vowel pattern only. Possible attestations of such stems are quite rare and many are disputed.

In all periods of Aramaic, and especially in Late Aramaic, a number of still additional stems are attested, but these are limited, occurring in no more than a few roots. One notable pair of stems is the *šaphʿel* and its passive, the *ʿEštaphʿal/ʿIštaphʿal*. These stems correspond in form and meaning to the *Haphʿel/Aphʿel* and the *ʿEttaphʿal/ʿIttaphʿal*, but with a prefixed *š-* rather than *h-* or *ʿ-*. In the *ʿEštaphʿal/ʿIštaphʿal*, metathesis of /š/ and /t/ has taken place (see §3.6.1). The forms of these stems that are attested in Aramaic are apparently loanwords from two possible sources: (i) Akkadian in the Imperial and Middle Aramaic periods, and (ii) (an) other Northwest Semitic language(s) in which the *šaphʿel* was the standard causative stem in the Old and/or Proto-Aramaic periods. Neither of these stems is productive in any extant Aramaic dialect.

#### 4.4.2 Inflectional categories

Verbs are inflected for three persons, two genders (not distinguished in the first person), two numbers, and two primary “tenses,” the perfect and the imperfect. There is also a set of second- and third-person jussive forms (attested in Old and Imperial Aramaic only), a set of second-person imperative forms, and an infinitival form, which is not inflected. In the active stems, there are two sets of participial forms, an active set and a passive set. In the passive stems, there is one set of (passive) participial forms. Participles are inflected like adjectives (see §4.2.2). The perfect and the imperative are characterized by inflectional suffixes, and the imperfect is characterized primarily by prefixes, though some forms have both prefixes and suffixes. The vowels that are associated with the root consonants of these forms will vary depending on the stem of the verb, the phonological form of the inflectional material, and the position of stress. As with nouns, variations in these vowels are the result of the phonological changes that took place during the history of Aramaic. However, explaining these alternating patterns synchronically requires a set of rather complex rules and will not be attempted here.

The exact semantic value of the two primary tenses is uncertain. It is likely that at the earliest stages of Aramaic, the perfect and the imperfect expressed distinctions of aspect and, secondarily, distinctions of tense and modality. The perfect was used to express perfective aspect, and tended to be used to express past tense and realis mode; whereas the imperfect was used to express imperfective aspect, and tended to be used to express non-past tense and irrealis mode. However, as early as the Imperial Aramaic period, tense began to be the primary distinction between the two forms and the participle began to be used more commonly as a verbal, rather than a nominal, form. By the Late Aramaic period, the perfect had become the past tense, the participle had become the non-past tense, and the imperfect was used to express contingency, purpose, or volition and occasionally to express future action. In conjunction with this shift, the system was augmented by “composite tenses” (see §4.4.2.6) that were used to express further distinctions of aspect and modality.

##### 4.4.2.1 Perfect tense

The perfect is characterized by inflectional suffixes. In (16), the written forms of these inflectional suffixes are represented in transliteration, both with and without vowel diacritics (see §§2.1, 2.2.2). Earlier or more broadly attested suffixes are listed above later or more narrowly attested suffixes. The symbol *ø* represents the absence of an inflectional suffix, either graphically and phonologically or only phonologically. In these forms, only *t* and *n* represent true consonants; all other letters are *matres lectionis* (see §2.2.1). On the phonemic values of the transliteration of the vowel diacritics, see §3.3.3. Verbs with a final root consonant that was historically /w/ or /y/ attest slightly altered forms of some of these suffixes.

(16)	Singular	Plural
3rd masculine	-ø	-w (= -û or -ø) -wn (= -ûn)
3rd feminine	-t (= -at)	- <sup>3</sup> or -h (= ā) -n (= -ān) or -yn (= -ên) or -y (= -ø or -î)
2nd masculine	-t <sup>3</sup> or -th or -t (= -tā) -t (= -t)	-tn or -twn (= -tôn or -tûn)
2nd feminine	-ty (= -tî) -t or -ty (= -t)	-tn or -tyn (= -tên or -tîn)
1st common	-t or -yt (= -et, -êt, or -ît)	-n <sup>3</sup> or -n (= -nā) -n (= -n) or -nn (= -nan)

Note the following:

1. The third feminine singular suffix is also sometimes attested as -<sup>3</sup> or -h (= -ā) in Jewish Babylonian Aramaic and in Samaritan Aramaic.
2. The second masculine singular suffix -t<sup>3</sup> or -th always represents -tā and is attested in all periods, although in Late Aramaic it is only attested in Jewish Palestinian Aramaic as a rare form. The spelling -t is also attested in all periods. In earlier periods, when *matres lectionis* were less frequently used, -t represents -tā written without a *mater lectionis*. In later periods, when *matres lectionis* were more frequently used, it represents -t.
3. The second feminine singular suffix -ty (= -tî) is an earlier form. In Late Aramaic, -ty is only found in Syriac and Samaritan Aramaic, where it represents -t.
4. The first common singular suffix is written with a *mater lectionis* only in some Late Aramaic texts. Its pronunciation varied from dialect to dialect and sometimes within individual dialects.
5. The third masculine plural suffix -w is attested in all periods and all dialects. It represents -û in all dialects except Syriac where its value is -ø. The suffix -wn (= -ûn) is a later alternate form found in Syriac and Jewish Palestinian Aramaic.
6. There are no distinct forms of the third feminine plural suffix attested in Old or Imperial Aramaic. In a few texts, third masculine plural forms are used with feminine plural subjects. The suffix -<sup>3</sup> or -h (= -ā) is attested in most dialects of Middle and Late Aramaic. The suffix -n (= -ān) is attested in Jewish Palestinian Aramaic and in Jewish Babylonian Aramaic. The suffix -y (= -î) is attested in Samaritan Aramaic, and the suffixes -yn (= -ên) and -y (= -ø) are attested in Syriac. These last two forms may have developed by analogy to the second feminine plural suffix.
7. The second masculine plural suffix is also attested as -tm (= -tûm or -tôm) in Old Aramaic. The suffixes -tn and/or -twn are attested in all periods.
8. No forms with a second feminine plural suffix are attested in Old Aramaic. The suffixes -tn and/or -tyn are attested in all other periods.
9. The first common plural suffix -n<sup>3</sup> always represents -nā and it is attested in all periods, but not in all dialects. The suffix -n is also attested in all periods. In earlier periods, it represents -nā written without a *mater lectionis*. In later periods, it represents -n. The form -nn (= -nan) is an alternate form only found in some dialects of Late Aramaic.

#### 4.4.2.2 Imperfect tense

The imperfect is characterized by inflectional prefixes, and, in some forms, suffixes as well. In the *ʾAphʿel* and the three stems with prefixed *ʾt-*, a prefixed consonant replaces the *ʾ* of the stem. In the earlier forms of these stems with prefixed *h-* or *ht-*, the *h-* remains and the consonant is prefixed to it. In (17), forms which are almost exclusively attested in eastern Late Aramaic are listed below forms which are attested in western Late Aramaic and all earlier dialects. All letters represent true consonants except *y* in the second feminine singular suffix, and *w* in the second and third masculine plural suffixes, which are *matres lectionis*. Verbs with a final root consonant that was historically /w/ or /y/ attest slightly altered forms of the suffixes.

(17)	<i>Singular</i>	<i>Plural</i>
<i>3rd masculine</i>	y- . . . -∅ n- . . . -∅ or l- . . . -∅	y- . . . -n or -wn (= -ûn) n- . . . -wn (= -ûn) or l- . . . -wn (= -ûn)
<i>3rd feminine</i>	t- . . . -∅	y- . . . -n (= -ân) n- . . . -n (= -ân) or l- . . . -n (= -ân)
<i>2nd masculine</i>	t- . . . -∅	t- . . . -n or -wn (= -ûn)
<i>2nd feminine</i>	t- . . . -n or -yn (= -în)	t- . . . -n (= -ân)
<i>1st common</i>	ʾ- . . . -∅	n- . . . -∅

Note the following:

1. The vowel following the prefix of each of these forms is determined by the stem and/or the initial root consonant of the particular verb.
2. In Syriac, the third masculine singular and plural, and the third feminine plural prefix is *n-* rather than *y-*.
3. In Jewish Babylonian Aramaic, the third masculine singular and plural, and the third feminine plural prefix is *l-* rather than *y-*. This prefix also occurs sporadically in other dialects.
4. In Syriac, there is an alternate third feminine singular form with the suffix *-y* (= -∅).
5. In Samaritan Aramaic, the second feminine singular suffix is *-y* (= -î), and in Jewish Babylonian Aramaic this suffix is attested as an alternate form.
6. In the Samʿal dialect of Old Aramaic, the third masculine plural suffix is attested as *-w* (= -û).
7. In Samaritan Aramaic and Jewish Babylonian Aramaic, the second and third masculine plural suffixes each have an alternate form *-w* (= -û).

#### 4.4.2.3 Jussive

In Old and Imperial Aramaic, quasi-imperative forms of the second and third persons, called “jussive forms,” are attested. These forms can be distinguished from the imperfect by the absence of the final *-n* in the plural forms as well as in the second feminine singular form. No distinction between the imperfect and the jussive is found in the other forms. By the Middle Aramaic period, no distinct jussive forms remained, although forms without the final *-n* were retained in some dialects either as the only imperfect form or as an alternate imperfect form (see §4.4.2.2).

#### 4.4.2.4 Imperative

The four imperative forms are closely related to the corresponding second-person imperfect forms. They differ from the imperfect forms in two ways: (i) they lack the prefix of the imperfect form (in the *ʾAphʿel* and the three stems with prefixed *ʾt-* the *ʾ* is present); and (ii) in most dialects, they lack the final *-n* of the imperfect forms, and what remains is a *mater lectionis* indicating the final vowel. Verbs with a final root consonant that was historically /w/ or /y/ attest slightly altered forms of these suffixes.

(18)	Singular	Plural
2nd masculine	-ø	-w (= -û)
2nd feminine	-y (= -î)	-h or -ʾ (= -ā)

Note the following:

1. In Jewish Palestinian Aramaic, the final *-n* is retained in the feminine singular and the two plural forms.
2. In Samaritan Aramaic, the final *-n* is optionally retained in the feminine plural.
3. In Syriac, the feminine singular suffix *-y* represents -ø, as does the masculine plural suffix *-w*. There is also an alternate form of the masculine plural suffix with final *-n* (*-wn* = -ûn). Finally, the standard feminine plural suffix is not attested in this dialect. Instead the feminine plural suffixes *-y* (= -ø) and *-yn* (= -ên) are attested.

#### 4.4.2.5 Infinitive

Each of the stems has a single infinitive form and this form is not inflected, although bound forms of the personal pronoun may be suffixed to it to indicate its subject or object (see §4.3.1.2). The infinitive is an action noun (*nomen actionis*) and, as such, it commonly occurs as the object of a preposition, especially the preposition *l* (see §5.3).

The *Pəʿal* infinitive has the historical form *\*maqtal* which becomes *miqtal* or *meqtal*, or remains *maqtal*, depending on the dialect and/or the first root consonant of the word. When a bound form of a personal pronoun is attached to one of these forms and the bound form begins with a vowel, the vowel preceding the final root consonant is reduced to /ə/ (e.g., *miqtəli*). Other, less common, forms of the *Pəʿal* infinitive are attested in a number of periods and dialects. For example, in Old Aramaic, a few infinitives without the prefixed *m-* are attested, and in Old and Imperial Aramaic a few infinitives with final *-at* or *-ût* or *-ā* (written with a *mater lectionis*) are attested. The form with final *-ā* resembles one of the common forms of the infinitives in the other stems and it is also attested in Jewish Palestinian Aramaic, Jewish Babylonian Aramaic, and Samaritan Aramaic. Also noteworthy is the form *miqtôl* attested in Jewish Palestinian Aramaic and Jewish Babylonian Aramaic.

The infinitives of the other stems are all formed in the same way. In every period and nearly every dialect, the infinitive has *ā* preceding and following the final root consonant (the second *ā* being written with a *mater lectionis*). In Syriac, the forms have final *-û* (written with a *mater lectionis*) rather than *-ā*. When a pronominal suffix is attached to any of these forms, *-ā* becomes *-at* or, more commonly, *-ût*, and *-û* becomes *-ût*. Sporadically throughout all periods of Aramaic, forms with final *-at* or *-ût* also occur without a suffix attached. In Old, Imperial, and Middle Aramaic, the infinitives of these stems do not have any kind of prefix, but in most dialects of Late Aramaic the prefixed *m-* of the *Pəʿal* stem is also found on the other stems (this prefix replaces the *ʾ* of the *ʾAphʿel* and the three stems with prefixed *ʾt-*). Jewish Babylonian Aramaic is one dialect that does not attest the prefix *m-* and,



furthermore, it attests an additional set of infinitive forms which are the common forms in this dialect. These forms have *ô* preceding the final root consonant and *ê* following the final root consonant (both vowels are written with a *mater lectionis*). These forms are also sporadically attested in Jewish Palestinian Aramaic.

#### 4.4.2.6 Composite tenses

As early as the Imperial Aramaic period, “compound” or “composite” tenses are attested which consist of an active participle combined with a finite form of the verb *hwʾh* “to be.” An active participle in combination with a perfect form of *hwʾh* is used to express past progressive or habitual action, and an active participle in combination with an imperfect form of *hwʾh* is used to express future progressive or habitual action. By the Late Aramaic period, these tenses had become much more commonly used, and additional tenses had developed in some of the dialects. For example, in Syriac, the perfect of *hwʾh* is used with the perfect of another verb either as a pluperfect or as a stylistic variant of the perfect verb.

### 4.5 Adverbs

In earlier dialects of Aramaic, there are relatively few adverbs and adverbial modification was frequently accomplished by the use of the absolute forms of nouns and adjectives: for example, *š/sgyʾ* “much, very.” In some cases, the noun or adjective may have retained an old accusative suffix */-a/*. One possible example is *klʾ* “completely” a form of the noun *kl* “all, every.” A few examples of adverbs which are not related to nouns are: *tnh*, *tnn* “here”; *tmh*, *tmn* “there”; *kn* “thus, so”; and *ʾdn*, *ʾdyn* “then.”

In Late Aramaic, these adverbs were retained and others were added to the lexicon through the increased use of adverbial suffixes such as *-ʾyt* in Syriac, which can be suffixed to any adjective to form an adverb.

### 4.6 Prepositions

All prepositions may have bound forms of the personal pronouns suffixed to them (see §4.3.1.2), and some prepositions are attested in combination with the particle *z/d(y)* (see §4.7.4), forming subordinating conjunctions (§4.8.2). Morphologically, prepositions can be divided into three categories:

1. *Inseparable prepositions*: Three prepositions, *b* “in,” *l* “to,” and *k* “like, as” (the last only attested in a few dialects) are phonologically and graphically proclitic to the following word. The preposition *mn* “from,” in some of its forms, also falls into this category.
2. *Independent unstressed prepositions*: These prepositions are written as separate words but receive no stress and so are phonologically proclitic to the following word. Some common prepositions are *ʾl* “over, to,” *ʾm* “with,” and *ʾd* “up to, until.” Also included in this group are the preposition *mn* “from,” in some of its forms, and the marker of the direct object, *ʾyt* in Old Aramaic, *yt* in Imperial Aramaic, Middle Aramaic, and Jewish dialects of Late Aramaic (see §5.2.2).
3. *Independent stressed prepositions*: These prepositions are written as separate words and are not phonologically proclitic to the following word. Some examples are: *ngd* “opposite,” *qdm* “before, in front of,” and *ʾhry* “behind, after.”



## 4.7 Particles

### 4.7.1 Existential particles

The particle *ʔt(y)* “there-is/are” expresses existence. The particle *lyt(y)* “there-is/are-not,” a contraction of the negative particle *lʔ* (see §4.7.2) and the existential particle *ʔt(y)*, expresses nonexistence. Both of these particles may have bound forms of the personal pronouns suffixed to them (see §4.3.1.2).

### 4.7.2 Negative particles

The particle *lʔ* “not” is used to negate verbs, clauses and, rarely, nouns. The particle *ʔl* “not” is used in prohibitions, which are expressed in Aramaic not by imperative verbs, but by jussive or imperfect verbs.

### 4.7.3 Interrogative particles

Numerous interrogative particles are attested in each of the Aramaic dialects, and the forms frequently vary from dialect to dialect. However, *mn*, *mʔn* “who,” and *mh*, *mʔ* “what” are constant throughout nearly all dialects. In texts influenced by Hebrew (the Biblical Aramaic texts and the Targums), a particle *h* is attested which may be prefixed to the first word of a clause to indicate that it is a question. In Jewish Babylonian Aramaic, the particles *my* and *ʔtw* have this function.

### 4.7.4 The particle *z/d(y)*

This particle is spelled *zy*, *z*, or *dy* in earlier texts and *d* or *dy* in later texts (see §3.2.2). In some dialects and periods, it is phonologically and graphically proclitic to the following word. It is an extremely important particle which indicates that the following noun or clause stands in some subordinate relationship to what precedes it. It has five primary uses: (i) to express a “genitive” relationship between two nouns; (ii) to introduce a relative clause modifying a preceding noun; (iii) to indicate the object clause of a verb; (iv) to introduce direct or indirect speech; (v) to express purpose or result. This particle is also used in combination with prepositions to form subordinating conjunctions (see §4.8.2).

## 4.8 Conjunctions

### 4.8.1 Coordinating conjunctions

A number of coordinating conjunctions are attested. Most notable is the ubiquitous *w* “and, but, or” which is always phonologically and graphically proclitic to the following word. Also attested are the less common *ʔw* “or,” (*ʔ*)*p* “also,” and *brm* “but,” which are neither phonologically nor graphically proclitic to the following word. In Syriac, the conjunction *dyn* “but, and then,” equivalent to Greek *δέ*, is quite common.

### 4.8.2 Subordinating conjunctions

A number of prepositions are used with the particle *z/d(y)* to form subordinating conjunctions: for example, *mn* “after,” *d* “until,” and *k* “when.” Other widely attested subordinating conjunctions are: *dln* “lest, perhaps”; *ʔl*, *ʔlw* “except that, however”;

*bdyl d* “so that, because”; *hn*, *ʾn* “if”; and *kl qbl* “because, on account of, inasmuch as.” In Syriac, the conjunction *gyr* “for, because,” equivalent to Greek γάρ, is quite common.

## 4.9 Interjections

Examples of the few attested interjections are: *ʾrw*, *hn*, *hʾ* “behold,” and *hy*, *ʾy*, *wy* “alas.”

## 4.10 Numerals

### 4.10.1 Cardinals

The cardinal numerals 1 through 10 are not inflected for number, only for gender and state, and they rarely occur in the construct and emphatic states. The numeral 2, in both the masculine, *tryn*, and the feminine, *trtyn*, forms, retains the Proto-Aramaic dual inflectional suffix *-yn*. In (19) the most common absolute forms of the numerals 1 through 10 are listed. The forms listed as “masculine” are those which modify masculine nouns, and those listed as “feminine” modify feminine nouns, despite the fact that the masculine forms of the numerals 3 through 10 are morphologically feminine, and the feminine forms are morphologically masculine (cf. §4.2.2 and §5.1).

(19)	<i>Masculine</i>	<i>Feminine</i>
1	ḥd	ḥdh, ḥdʾ
2	tryn	trtyn
3	tlth, tltʾ	tlt
4	ʾrbʿh, ʾrbʿʾ	ʾrbʿ
5	ḥmšh, ḥmšʾ	ḥmš, ḥmyš
6	šth, štʾ, ʾšth, ʾštʾ	št, šyt
7	šbʿh, šbʿʾ	šbʿ
8	tmnyh, tmnyʾ	tmnh, tmnʾ, tmny
9	tšʿh, tšʿʾ	tšʿ, tyšʿ
10	ʿšrh, ʿšrʾ, ʿsrh, ʿsrʾ	ʿšr, ʿsr

Note the following:

1. In these forms, final *-h* or *-ʾ* is a *mater lectionis*. Forms with *-h* occur in earlier dialects and forms with *-ʾ* occur in Late Aramaic, except Jewish Palestinian Aramaic and Samaritan Aramaic, which attest *-h*.
2. The final *-y* in the feminine form of 8 is a *mater lectionis* as is the medial *-y-* in the feminine forms of 5, 6, and 9, but not in the masculine form of 8. In that form, it is a consonant.
3. The medial *-y-* in both forms of the numeral 2 represents the Proto-Aramaic diphthong *\*/ai/*, which may have been retained in these forms as late as the Imperial Aramaic period. By the Middle or Late Aramaic period, this diphthong in this particular form had become */e:/* (see §3.3.2) in all dialects and so the *y* then functions only as a *mater lectionis*.
4. In some dialects, the masculine form of 6 is sometimes written with a prothetic *aleph* (see §3.6.5.3).
5. The numeral 10 is written with *š* in earlier dialects and with *s* in later ones (see §3.2.2).

The numerals 11 through 19 are inflected only for gender and consist of a combination of a form of the relevant digit (absolute, construct, or alternate) and an alternate form of

the numeral 10. The forms of these numerals vary across the Aramaic dialects, and in some dialects multiple forms of some of these numerals are attested.

The numerals 20, 30, 40, 50, 60, 70, 80, and 90 are not inflected. They each have a single form which is characterized by a suffixed *-în*. These forms are essentially equivalent to the masculine plural absolute form of the corresponding digit, except for the numeral 20 which is equivalent to the masculine plural absolute form of 10: for example, *tlāt* “3,” *tlātîn* “30”; and ‘*sar* “10,” ‘*asrîn* “20.”

The numeral 100 is a feminine noun and the numeral 1,000 is a masculine noun. They are fully inflected for number and state, their plural forms being used in combination with the digits 3 through 9 to form 300, 3,000, and so forth. The numerals 200 and 2,000 are formed using the dual inflectional suffix rather than the digit 2.

Bound forms of the personal pronouns can be suffixed to the numerals 2 through 10, though they are rarely attested.

#### 4.10.2 Ordinals

There are distinct ordinal forms of the numerals 1 through 10. These forms have the same root consonants as the corresponding cardinals, except for the numeral 1, and, except for the numerals 1 and 2, they are characterized by the vowel *î* preceding the final root consonant and the suffix *āy* following the final root consonant: for example, *tlāt* “3,” *tlitāy* “3rd.” In some dialects of Middle and Late Aramaic, the suffix is *āʾ*. These numerals are adjectives and can be fully inflected for gender, number, and state, although they are most commonly attested in the absolute state. For ordinal numerals higher than 10, the corresponding cardinal numeral is used.

In some dialects of Late Aramaic, cardinal numerals with the prefixed particle *z/d(y)* are also used as ordinals: for example, *dtryn* “who [is] 2” = “2nd.”

## 5. SYNTAX

### 5.1 Noun phrase structure

Any noun or adjective can constitute a noun phrase by itself. An adjective which stands alone is interpreted as a concrete noun meaning “one who has the quality designated by the adjective.”

Adjectives can be either attributive or predicative (see §5.2.1). An attributive adjective stands in an appositional relationship to a noun. The adjective nearly always follows the noun and agrees with it in gender, number and state:

- (20) A. ʾnš̄ ṭb  
           man good  
           “a good man”  
       B. ʾnš̄' ṭb'  
           the-man the-good  
           “the good man”

With the decreased use of the absolute state in Late Aramaic, the second example came to mean either “the good man” or “a good man” (see §4.2.2).

Demonstrative pronouns may be used either attributively or predicatively (see §5.2.1), but these uses cannot be distinguished by the form of the demonstrative itself, except in Jewish dialects of Late Aramaic in which an attributive demonstrative has a prefixed *h-* (this *h* is in addition to the *h* which is characteristic of some forms of the demonstrative pronouns in Late Aramaic; see §4.3.2.1). An attributive demonstrative may either precede or follow the noun it modifies, which must be in the emphatic state:

- (21) A. *byt' dnh*  
           the-house this  
           “this house”  
       B. *dnh byt'*  
           this the-house  
           “this house”

Though the position of the pronoun is not fixed, one position or the other tends to be preferred in each dialect and/or time period. With the increased use of the emphatic state, the demonstrative came to be used in some instances as little more than a definite article (see §4.2.2).

The modification of nouns by cardinal numerals shows a number of idiosyncrasies which differ from dialect to dialect. There are a few features that all cardinal numerals show in all dialects.

1. The numerals 1 to 19, which are the only numerals that distinguish gender, must agree in gender with the noun they modify. However, the numerals 3 to 10 show “chiastic concord” – the morphologically masculine form modifies feminine nouns and the morphologically feminine form modifies masculine nouns (see §4.10.1).
2. Numerals other than 1 may either precede or follow the noun, and the noun is plural.
3. The numeral 1 nearly always follows the noun and, of course, the noun is singular.
4. The numerals 2 to 10 can occur in either the absolute or the construct state with a following noun, but there is no difference in meaning: for example, (i) *tryn* (absolute) *'nšyn*; and (ii) *try* (construct) *'nšyn* – both meaning “two men.”
5. The numerals 100 and 1,000 are nouns which may be modified by other numerals.

The ordinal numerals are adjectives and have the syntax of adjectives (see [20] above).

Modification of a noun by a prepositional phrase, an adverb, or a clause is accomplished through the use of the particle *z/d(y)* “who, which”; for example:

- (22) *'nš' dy bbyt'*  
           the-man who in-the-house  
           “the man who [is] in the house”

The particle *z/d(y)* can be omitted in this construction, though this is extremely rare.

The relationships between two noun phrases that are expressed by the genitive case in some languages are expressed in Aramaic in two different ways.

On the one hand, genitive relationships can be expressed by a construct chain in which a noun in the absolute or emphatic state is preceded by one or more nouns in the construct state. The definiteness of all nouns in a chain is determined by the definiteness of the final noun:

- (23) A. *br* *mlk*  
           son-of.CONSTRUCT king.ABSOLUTE  
           “a king’s son”  
       B. *br* *mlk’*  
           son-of.CONSTRUCT the-king.EMPHATIC  
           “the king’s son”

Most construct chains consist of two nouns, though construct chains of three nouns are not uncommon and chains of four nouns are attested. The use of the construct chain decreased over time, and by the Late Aramaic period the construction is only attested in chains that had been reanalyzed as compound nouns or in chains formed with a few words such as *br* “son-of” and *byt* “house-of.”

On the other hand, genitive relationships can be expressed by a construction using the particle *z/d(y)* in which one noun is followed by the particle and a second noun. The second noun may be in either the absolute or emphatic state. The first noun may appear in one of three forms: (i) in the absolute state; (ii) in the emphatic state; or (iii) it may be suffixed with a bound form of the personal pronoun that agrees in gender and number with the second noun, although this form may only be used if the second noun is in the emphatic state:

- (24) A. *br’* *dy* *mlk’*  
           the-son.EMPHATIC of the-king.EMPHATIC  
           “the king’s son”  
       B. *brh* *dy* *mlk’*  
           son-his (= the king) of the-king.EMPHATIC  
           “the king’s son”

Constructions in which one or the other or both nouns are in the absolute state are rare and occur most commonly in constructions expressing the “genitive of material”:

- (25) *tr‘n* *zy* *’bn*  
       gates.ABSOLUTE of stone.ABSOLUTE  
       “stone gates”

## 5.2 Clause structure

### 5.2.1 Nonverbal clauses

Nonverbal clauses in Aramaic can be formed by the juxtaposition of a noun (phrase) or a pronoun used as a subject with an adjective, participle, prepositional phrase, adverb, or noun (phrase) used as a predicate. In such a clause, the predicate usually precedes the subject, except for the participle, which usually follows the subject. A predicative adjective or participle must agree with its subject in gender and number, and must also be in the absolute state, regardless of the state of its subject:

- (26) *ṭb* *khn’*  
       good.ABSOLUTE the-priest  
       “The priest is good.”

When a noun (phrase) is the predicate, an additional personal pronoun is often used, either preceding or following the subject:

- (27) A. ywḥnn hw' khn'  
           John     he     the-priest  
           "The priest is John."  
       B. ywḥnn khn'         hw'  
           John     the-priest   he  
           "John is the priest."

In Syriac and Jewish Babylonian Aramaic, the use of such pronouns was greatly expanded and they became used in all kinds of nonverbal clauses. In connection with this use, additional bound forms of the personal pronoun were developed (see §4.3.1.2).

### 5.2.2 Verbal clauses

In Aramaic, a finite verb form, by itself, can constitute a verbal clause. Since the verb is inflected for person as well as gender and number, no other element is necessary to constitute a clause.

A verbal clause may contain a subject noun (phrase), although the subject is commonly omitted in Aramaic if it is contextually identifiable. The verb agrees in gender and number with its subject. If a plural subject is of mixed gender, the verb is masculine. Not uncommonly, a singular verb will occur with a plural subject or a masculine verb will occur with a feminine subject. Such disagreements between subject and verb are much more commonly attested when the subject follows the verb; when the subject precedes the verb, the verb rarely disagrees with it.

An indefinite direct object of a verb is not specially marked in Aramaic. A definite direct object of a verb is sometimes marked in Old Aramaic by the particle *ʾyt* (see §4.6). A later form of this particle, *yt*, is sometimes used in Imperial Aramaic, Middle Aramaic, and Jewish dialects of Late Aramaic, often in imitation of the Hebrew particle *ʾt*. More commonly in these periods and dialects, and exclusively in all other dialects of Late Aramaic, the preposition *l* is used to mark the definite direct object of a verb. In Late Aramaic, a definite direct object often occurs both as a bound pronoun suffixed to the verb and as a noun (phrase) marked with the preposition *l*:

- (28) ktbh                 lktb'  
       he-wrote-it   the-book  
       "He wrote the book."

Finally, the direct object of a verb may be omitted from a clause if it is identifiable from the immediate context.

The indirect object of a verb is also marked by the preposition *l* "to" which often leads to ambiguity. The indirect object may also be omitted from a clause if it is identifiable from the immediate context.

Two kinds of verbal adjuncts are particularly noteworthy. First, the agent of a passive verb is rarely indicated in most Aramaic dialects; however, in Syriac, the agent is more commonly expressed and when it is, the preposition *mn* marks it. Second, the absolute form of a noun or adjective can be used within a clause as an adverb rather than as a verbal complement. This use of nouns and adjectives is more common in earlier dialects and it decreases in later dialects as the number of true adverbs increases (see §4.5).

### 5.2.3 Subordinate clauses

There is no difference in the structure of a subordinate clause as compared with a main clause, except, of course, for the presence of a subordinating conjunction. However, this particle is sometimes omitted and the subordinate nature of the clause must then be inferred.

At times, a subordinate relationship exists between two formally coordinate clauses. There are two notable examples of such a relationship. The first is the conditional clause. In general, the protasis of a conditional clause begins with a conditional particle and will either precede or follow the apodosis to which it is subordinate. However, sometimes the protasis and the apodosis will be joined by the coordinating conjunction *w* (the so-called *waw* of apodosis) in which case, the protasis will always precede the apodosis; for example:

- (29) *hn kn 'bdw . . . wšdqh yhwh lk*  
 if thus you-do . . . and-merit will-be to-you  
 "If you act in this way . . . (then) you will have merit."

The second is verbal hendiadys, a construction in which two verbs are conjoined and share all verbal complements, but the first verb expresses a modification of the second rather than an independent action, as in the following:

- (30) *'sgy wqr' lhwn*  
 he-increased and-he-called to-them  
 "He called to them often."

This construction tends to occur in dialects and texts which are influenced by Hebrew, where the construction is more common.

### 5.2.4 Word order

The word order of the elements in a clause is not grammatically fixed in Aramaic and varies in part by the place of any given clause within the larger discourse. However, there are certain orders which can be considered "standard" and appear to have no special discourse function. In most dialects of Aramaic, this standard order is VSO (verb, subject, object, indirect object), although a pronominal object or indirect object will frequently precede a nominal subject. In Imperial Aramaic, the verb is often the final element of the clause, a result of Akkadian influence. Verbal adjuncts usually follow verbal complements within a clause.

A subordinate clause usually follows, but sometimes precedes, all of the elements of the main clause to which it is subordinated, although there are occasional examples of a subordinate clause being followed by complements or adjuncts of the main clause. These examples are most common when the elements of the main clause are particularly long and/or the subordinate clause is particularly short. In general, though, each clause is a discrete unit.

Negative particles, interrogative particles, coordinating conjunctions, and subordinating conjunctions will nearly always occur as the first element of a clause. Two regular exceptions to this tendency are the Syriac particles *gyr* "for, because" and *dyn* "but, and then" which are postpositive, like their Greek counterparts *γάρ* and *δέ*.

## 5.3 Infinitival syntax

The infinitive has aspects of nominal syntax and aspects of verbal syntax. As a verb, the infinitive can occur with its own complements and adjuncts. As a noun, it and its associated elements can occur as a complement or an adjunct of a verb. As a complement, it most

commonly occurs as an object (usually marked with *l*), though its use as a subject, especially the subject of a nonverbal clause, is not uncommon. As an adjunct, it nearly always occurs as the object of the preposition *l*.

The functions of the infinitive as an adjunct are numerous and they parallel the functions of subordinate clauses. Frequently the same function can be expressed either by an infinitive or by a subordinate clause and there are even attestations of infinitives and subordinate clauses being conjoined with *w* “and.” Two of the more common functions of the infinitive, both with the preposition *l*, are purpose/result and “epexegetic” or explanatory. There are also a few isolated examples of the temporal use of the infinitive with prepositions such as *k* “as, when” and *b* “in, when.” This use of the infinitive was never common in Aramaic, and all of the examples of this use after the Old Aramaic period are in texts influenced by Hebrew, where the temporal use of the infinitive is quite common.

Because the infinitive most commonly occurs with the preposition *l* prefixed to it, this *l* became reanalyzed, apparently as early as the Imperial Aramaic period, as part of the infinitive form itself rather than as a preposition indicating the function of the infinitive within a clause. As a result, the word order of the complements of the infinitive became less rigid. In Old Aramaic, the infinitive precedes all of its complements, but in Imperial Aramaic and many dialects of Middle and Late Aramaic, the object of the infinitive commonly precedes it, even though the infinitive has *l* prefixed to it.

In dialects of Aramaic influenced by Hebrew and in the Old Aramaic Sefire texts, the infinitive is sometimes used in the same way as the Hebrew infinitive absolute, a use in which the infinitive occurs with a verb of the same root and stem to express the certainty of the action:

- (31) *mbnʾ      bnʾ*  
       to-build   he-builds  
       “He will certainly build.”

In this use, the infinitive never occurs with prefixed *l*.

## 5.4 Additional syntactic constructions

### 5.4.1 Possession

To express the notion of possession, the particle *ʾyt(y)* “there-is/are” or the verb *hw ʾh* “to be” is used in combination with the preposition *l* “to.” The thing possessed is the subject of the verb or the particle, and the possessor is the object of the preposition:

- (32) *ʾyt      lʾnšʾ      ksp*  
       there-is   to-the-man   silver  
       “The man has silver.”

### 5.4.2 Comparison

A comparative construction is formed by the use of a predicative adjective in combination with the preposition *mn* “from.” One of the compared objects is the subject of the clause, and the other is the object of the preposition:

- (33) *ṭb      ʾbdʾ      mn      mlkʾ*  
       good   the-servant   from   the-king  
       “The servant is better than the king.”



### 5.4.3 Impersonal constructions

Two impersonal constructions are commonly attested. In the first, a masculine plural (or, less commonly, singular) active verb is used without an explicit or contextually understood subject to express the equivalent of a passive verb:

- (34) lk                      trdyn                      mn                      'nš'  
       you.OBJ.    they-will-drive-out    from    humanity  
       “You will be driven out from human society.”

In the second, a passive participle is used in combination with the preposition *l* “to” to express the equivalent of an active finite verb:

- (35) ṭwr'                      bšyn                      lh  
       the-mountains    be-searched.PASS.PART.    to-him  
       “He searched the mountains.”

This construction can even be used with an intransitive verb which normally would not have a passive participle:

- (36) qym                      ly                      qdm                      šlytn'  
       be-stood.PASS.PART.    to-me    in-front-of    powerful-men  
       “I have stood in front of powerful men.”

This construction was borrowed from Persian where it is commonly attested.

## 6. LEXICON

Because of its use as a lingua franca and its contact with many other languages throughout its history, Aramaic contains numerous loanwords in addition to its core lexicon of native words. Nearly all of these loanwords are nouns. Aramaic borrowed very few verbs directly from other languages, although sometimes denominative verbs were created from loaned nouns. In the Imperial Aramaic period, Aramaic acquired words from Akkadian, Persian, and Egyptian. In the Middle Aramaic period, Greek words were added to the lexicon and these additions increased in the Late Aramaic period. Latin words were also added in the Late Aramaic period, as were a second group of Persian words in the eastern dialects. Finally, Hebrew was a constant source of loans in Jewish dialects of Aramaic.

### 6.1 Akkadian

Most Akkadian loanwords are administrative or architectural terms such as *sgn* (< *šaknu* “prefect”), *pḥt* (< *pīḥātu* “governor”), *grh* (< *egirtu* “letter”), and *trbš* (< *tarbišu* “court-yard”); though other terms such as *mlḥ* (< *malāḥu* “boatman”) and *šp* (< *āšipu* “enchanter”) are also attested (see Kaufman 1974). Another notable loanword is the *Šaph'el* verb *šyzb* (< *ušēzib* “to save”). Akkadian loanwords are completely assimilated to Aramaic, both phonologically and morphologically.

### 6.2 Persian

Like Akkadian, many Persian loanwords are administrative terms, reflecting the Persian rule of the Near East, and these words are all completely assimilated to the Aramaic inflectional system (despite the fact that Persian is an Indo-European language). Some examples are: *prtrk* (< *frataraka* “governor”), *hmrkry'* (< *hmārakara* “accountant”), and *'zdkr'* (< *azdākara* “messenger”). A number of Persian words for very common items or concepts

became the common Aramaic terms as well, for example: *ptgm* (< *patigāma* “word”), *rzʾ* (< *rāza* “secret”), and *znʾ* (< *zana* “kind”); see Muraoka and Porten 1998.

### 6.3 Egyptian

Egyptian loanwords are very rare in Aramaic and are restricted to Imperial Aramaic texts from Egypt. These words do not become part of the broader Aramaic lexicon. For whatever reason, a considerable number of these words relate to boats, though commodities and other terms are also attested. Some examples are: *tqm* (< *tgm* “castor oil”), *qntḥntr* (< *qnḥ-ntr* “divine shrine”), *ṭp* (< *ḏp* “part of a ship’s mast”), and *šntʾ* (< *šnt* “linen robe”); see Muraoka and Porten 1998.

### 6.4 Greek

Greek loanwords, which total over two thousand from various dialects, represent the largest group of non-native words in the Aramaic lexicon. They are not always completely assimilated to the Aramaic inflectional system. Many loanwords show multiple forms which reflect Greek rather than Aramaic inflectional suffixes. In some cases, forms with Aramaic inflectional suffixes coexist with forms that reflect Greek suffixes. Some examples are: *ʾrtyqyʾ*, *ʾrtyqws* (< αἰρετικός “heretic”), *ʾwsyʾ*, *ʾwsyʾs* (pl.) (< οὐσία “essence”), and *ṭksʾ*, *ṭksyn* (pl.) (< τάξις “order, row”); see Krauss 1898–1899.

### 6.5 Latin

Latin loanwords are relatively rare and are mostly restricted to dialects of western Late Aramaic. They are similar to Greek loanwords in that they are not always fully assimilated to the Aramaic inflectional system. Some examples of Latin loanwords are: *dwnʾtybʾ* (< *donativa* “imperial gift”), *ṭblh* (< *tabula* “board, tablet”), and *qlnds* (< *kalendas* [acc.] “first day of the month”); see Krauss 1898–1899.

### 6.6 Hebrew

Hebrew loanwords are only attested in Jewish dialects of Aramaic, and their status in those dialects is not always clear. This uncertainty is a result, in part, of the similarity of Hebrew and Aramaic. Frequently, words in the two languages only differed by a single vowel or by an inflectional suffix. Also, Hebrew and Aramaic coexisted for a very long time in Jewish communities, and literate members of those communities would have been well acquainted with both languages. So, when a Hebrew word appears in an Aramaic text, it may be a loanword, or it may simply be a Hebrew word which is being used because the writer of the text could assume that the readers of the text would be acquainted with it.

## Bibliography

The bibliography contains all the special studies cited in the chapter as well as other bibliographies, surveys of Aramaic dialects, and at least one grammar for each of the major dialects of Aramaic.

Brock, S. 1996. *Syriac Studies: A Classified Bibliography, 1960–1990*. Kaslik, Lebanon: Université Saint-Esprit de Kaslik.

Cantineau, J. 1930. *Le nabatéen I. Notions générales – écritures, grammaire*. Paris: Ernest Leroux.

———. 1935. *Grammaire du palmyrénien épigraphique*. Cairo: L’Institut Français d’Archéologie Orientale. Reprint 1987, Osnabrück: Otto Zellar Verlag.

Crichton, W. (trans.). 1904. *Compendious Syriac Grammar*. London: Williams and Norgate.

Dalman, G. 1905. *Grammatik des jüdisch-palästinischen Aramäisch nach der Idiomen des palästinischen Talmud, des Onkelostargum und Prophetentargum und der jerusalemischen*

- Targume* (2nd edition). Leipzig: Hinrichs. Reprint 1981, Darmstadt: Wissenschaftliche Buchgesellschaft.
- Degen, R. 1969. *Altaramäische Grammatik der Inschriften des 10-8 Jh. v. Chr.* Abhandlungen für die Kunde des Morgenlandes 38/3. Wiesbaden: Steiner.
- Drijvers, H. 1973. "XIII. Syriac and Aramaic." In J. Hespers (ed.), *A Basic Bibliography for the Study of the Semitic Languages*, vol. I, pp. 283–335. Leiden: Brill.
- Fitzmyer, J. 1966. *The Genesis Apocryphon of Qumran Cave I*. Biblica et Orientalia 18. Rome: Pontifical Biblical Institute.
- . 1979. "The phases of the Aramaic language." In *A Wandering Aramean. Collected Aramaic Essays*, pp. 57–84. SBL Monograph Series 25. Missoula, MT: Scholars Press.
- Fitzmyer, J. and S. Kaufman. 1992. *An Aramaic Bibliography, Part I: Old, Official, and Biblical Aramaic*. Baltimore and London: Johns Hopkins University Press.
- Folmer, M. 1995. *The Aramaic Language in the Achaemenid Period*. Orientalia Lovaniensia Analecta 68. Louvain-La-Neuve: Peeters.
- Greenfield, J. 1978. "Aramaic and its dialects." In H. Paper (ed.), *Jewish Languages. Theme and Variation*, pp. 29–43. Cambridge, MA: Association for Jewish Studies.
- Grossfeld, B. 1972. *A Bibliography of Targum Literature*, vol. I. Bibliographica Judaica 3. Cincinnati: Hebrew Union College Press; New York: KTAV Publishing House.
- . 1977. *A Bibliography of Targum Literature*, vol. II. Bibliographica Judaica 8. Cincinnati: Hebrew Union College Press; New York: KTAV Publishing House.
- . 1990. *A Bibliography of Targum Literature*, vol. III. New York: Sepher-Hermon Press.
- Hug, V. 1993. *Altaramäische Grammatik der Texte des 7. Und 6. Jhs. v. Chr.* Heidelberger Studien zum Alten Orient 4. Heidelberg: Heidelberger Orientverlag.
- Kaufman, S. 1974. *The Akkadian Influences on Aramaic*. Assyriological Studies 19. Chicago: University of Chicago Press.
- Krauss, S. 1898–1899. *Griechische und lateinische Lehnwörter in Talmud, Midrash und Targum* (2 vols.). Berlin: S. Calvary. Reprint 1987, Hildesheim: Olms.
- Kutscher, E. 1970. "Aramaic." In T. A. Sebeok (ed.), *Current Trends in Linguistics*, vol. VI: *Linguistics in South West Asia and North Africa*, pp. 347–412. The Hague: Mouton.
- . 1971. "The Hermopolis papyri." *Israel Oriental Studies* 1:103–119.
- Levias, C. 1900. *A Grammar of the Aramaic Idiom Contained in the Babylonian Talmud*. Cincinnati: Bloch Publishing and Printing.
- Macuch, R. 1965. *Handbook of Classical and Modern Mandaic*. Berlin: Walter de Gruyter.
- . 1982. *Grammatik des samaritanischen Aramäisch*. Berlin: Walter de Gruyter.
- Margolis, M. 1910. *A Manual of the Aramaic Language of the Babylonian Talmud*. Munich: Beck.
- Morag, S. 1962. *The Vocalization Systems of Arabic, Hebrew, and Aramaic; Their Phonetic and Phonemic Principles*. Janua Linguarum. Series Minor 13. The Hague: Mouton.
- Muraoka, T. and B. Porten. 1998. *A Grammar of Egyptian Aramaic*. Handbuch der Orientalistik 32. Leiden: Brill.
- Naveh, J. 1982. *Early History of the Alphabet*. Leiden: Brill.
- Nöldeke, T. 1875. *Mandäische Grammatik*. Halle: Buchhandlung des Waisenhauses. Reprint 1964, Darmstadt: Wissenschaftliche Buchgesellschaft.
- . 1898. *Kurzgefasste syrische Grammatik* (2nd edition). Leipzig: Weigel.
- Rosenthal, F. 1939. *Die aramaistische Forschung seit Th. Nöldeke's Veröffentlichungen*. Leiden: Brill. Reprint 1964.
- Rosenthal, F. (ed.). 1967. *An Aramaic Handbook* (4 vols.). Wiesbaden: Otto Harrassowitz.
- Schulthess, F. 1924. *Grammatik des christliche-palästinischen Aramäisch*. Tübingen: Mohr. Reprint 1965, Hildesheim: Olms.
- Segal, J. 1953. *The Diacritical Point and the Accents in Syriac*. Oxford: Clarendon Press.
- Segert, S. 1975. *Altaramäische Grammatik mit Bibliographie, Chrestomathie und Glossar*. Leipzig: VEB Verlag Enzyklopädie.
- Steiner, R. 1977. *The Case for Fricative-Laterals in Proto-Semitic*. New Haven, CT: American Oriental Society.
- Stevenson, W. 1962. *Grammar of Palestinian Jewish Aramaic* (2nd edition). Oxford: Clarendon Press.

# Ge'ez (Aksum)

GENE GRAGG

## 1. HISTORICAL AND CULTURAL CONTEXTS

There are four great kingdoms on earth: the first is the kingdom of Babylon and Persia; the second is the kingdom of Rome; the third is the kingdom of the Axumites; the fourth is the kingdom of the Chinese. Mani, *Kephalaia* LXXVII

### 1.1 Historical background

This third-century Manichaean text (cf. Kobishchanov 1979:59) shows the reputation enjoyed, at least sporadically, by the shadowy East African kingdom which was just over the horizon of the known classical universe, and which from time to time impinged on the consciousness of the world of late antiquity. Its capital city, Aksum, was located near the northern edge of the great Ethiopian plateau which rises abruptly behind the Red Sea coast, a mirror image to the coastal escarpment on the Yemeni side, and slopes down gently into East Africa, its steep southeast flank forming the northern wall of the Great Rift Valley in Africa.

The Kingdom of Aksum itself arose toward the end of the second century AD, and continued to play an important role in the region until the rise of Islam. But prior to that, the region of Aksum had been the site of a South Arabian colony, centered at Yeha, about 30 kilometers east of Aksum, paleographically dated to around 500 BC by monumental inscriptions of the classical Sabean type. Presumably, given the intensely commercial orientation of the South Arabian city-states and the geographical proximity (Arabia and the Horn of Africa are separated by only 40 kilometers in the Bab-el-Mandab straits at the southern end of the Red Sea), contacts based on trade relations go back even further. It is clear that many of the political and cultural traditions of the kingdom stemmed out of this colonization or its antecedents, and linguistically, as will be seen, Ethiopic Semitic is a close cousin of Old South Arabian. However, it is not possible to derive Ethiopic Semitic, or any of its constituent branches and languages, from any single attested form of Old South Arabian.

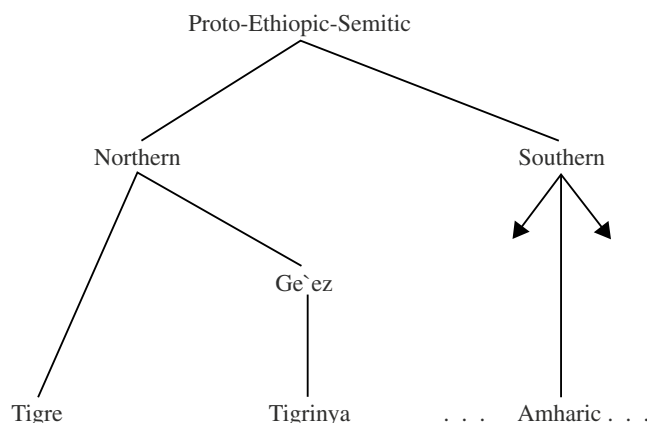
### 1.2 Linguistic history

Classical Ethiopic, the language of Aksum, whose self-designation is Ge'ez ([gəʕəz], etymology uncertain), is presumably derived from one or more forms of South Semitic brought from Yemen, probably in the first half of the first millennium BC. In all likelihood, Ethiopic Semitic evolved out of a South-Arabian-based trade lingua franca, perhaps passing through stages of piginization and creolization familiar from differentiation and development of language families elsewhere in the world (e.g., Romance). The substratum languages in this

development presumably belonged to the Cushitic language family, and a number of important early loanwords from Cushitic are evident in Ge'ez – but at present it is not possible to reconstruct the mechanisms of this development. Ge'ez disappeared as a spoken language probably some time before the tenth century AD, but continued as the liturgical language of the Ethiopian Orthodox Church, and as the only official written language of Ethiopia practically up to the end of the nineteenth century.

As has been implied, Ge'ez is one of a number of Semitic languages spoken in Ethiopia. Although Ge'ez is the earliest attested Ethio-Semitic language, it cannot be taken as identical with Proto-Ethio-Semitic. One simple example – the word for “not” had to be *\*ʔal* in Proto-Ethio-Semitic (e.g., Amharic, *al* “not”), and does indeed have the form *ʔl* in Old South Arabian. In Ge'ez, however, it has become everywhere *i* (< *\*ay* by palatalization of the /l/), except for a fossilized remnant preserved in the word *albo* (< *al-bä-hu* “not in it”) “not exist; not have.” The other Ethio-Semitic languages could not have inherited this lexical item from Ge'ez in its present form.

Within the Ethio-Semitic subfamily, Ge'ez, together with Tigre and Tigrinya, falls into the Northern Ethio-Semitic branch. Here it is closely related to modern Tigre (northern highlands and Red Sea coastal plain), and perhaps stands in a more or less proximate ancestral relationship to Tigrinya (Northern Ethiopia and Eritrea), which probably started to emerge as a distinct entity on the home territory of Ge'ez from around the tenth century on. The remaining Ethio-Semitic languages (a dozen or so, including Amharic) belong to a separate (Southern) group, which cannot be derived from any attested Northern Ethio-Semitic language. Northern and Southern Ethio-Semitic, however, do seem to constitute a distinct genetic node in the Semitic family tree, a node which has thus the following structure:



**Figure 14.1** The Ethio-Semitic subfamily

On the level of Semitic as a whole, Ge'ez is most closely related to a Southern group of Semitic languages that includes Epigraphic and Modern South Arabian. The exact historical relationships among Epigraphic South Arabian (see Ch. 15), Modern South Arabian, and Ethio-Semitic have been difficult to establish because of what was until quite recently insufficient data on the Modern South Arabian languages, and because of the phonological indeterminacy and morphological poverty of the textual evidence for Epigraphic South Arabian (an extensive corpus, but written in one of the more resolutely vowelless of the Semitic writing systems, and in a discourse format which, in spite of a respectable diversity of subject matter, managed to restrict itself almost entirely to third-person pronominal and verbal forms).

Even more difficult has been the relation of these three to the quite distinct Northern (Classical) Arabic, which we will henceforth refer to simply as “Arabic.” At one time it

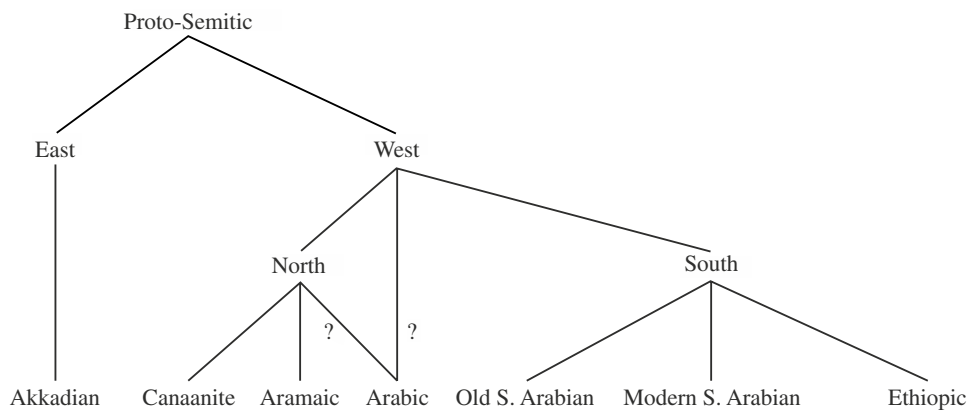
had been common to group South Arabian, Ethiopic, and Arabic into one *South Semitic* subfamily, and to appeal to a small inventory of shared features:

1. On the level of phonology – the presence of a voiceless labial continuant /f/, instead of the stop /p/ as is the case in Canaanite, Aramaic (frequently grouped together in a subfamily referred to as *Northwest Semitic*), and Akkadian (usually considered to represent by itself a separate *East Semitic* branch).
2. On the level of morphology – the presence of a highly developed system of *internal* plurals, formed by changes in stem syllabicity and vocalism (a feature occurring only restrictedly in Canaanite, and apparently absent in Aramaic and Akkadian; see §4.1.1.2 for the patterns attested in Ge'ez).

More recently, however, there has been a tendency to give more weight to an innovation which Arabic shares with Northwest Semitic. Through this innovation, an ancestral *present stem* characterized by a bi-syllabic CVCVC pattern (and perhaps also by gemination of the middle consonant), found in Akkadian (*i-parras* “he decides”) and Ethiopic (*yəqattəl* “he kills”), is dropped in favor of a CCVC (Arabic *ya-qtulu*, Hebrew *yi-qtol* “he kills”) pattern, which either coexisted with the bisyllabic pattern in Proto-Semitic, or was innovated on the basis of an inherited jussive stem pattern (Ethiopic *yəqtəl*, Arabic *ya-qtul*).

As opposed to this northern innovation, there is a genuinely southern feature, shared by Ethiopic and South Arabian, not present in Arabic. This has to do with the initial consonant of the first- and second-person subject suffixes of the past tense (the relevant forms for Arabic and Ge'ez, together with those of the Akkadian stative, are given in §4.3.1). In the (i) first-person singular and (ii) second-person singular and plural forms, Arabic, like Hebrew and Aramaic, has /t/, whereas Ge'ez, like Modern South Arabian, has /k/ (and we now know that Old South Arabian also had /k/, certainly in the second singular, and probably in the other forms also, see Ch. 15, §4.3.1). Since 1969 many Semitists have subscribed to a theory first enunciated by Robert Hetzron: namely, that the earliest form of this paradigm can be seen in the Akkadian stative paradigm (where the consonant is preceded by a long /ā/). Here, as in the independent pronoun, the first person is marked by a /k/ and the second person by a /t/. In Western Semitic, the northern languages would have generalized the non-third-person consonant to /t/, whereas the southern languages would have generalized the /k/.

Giving a family tree interpretation to this line of argument, Arabic would be taken out of South Semitic, and either joined to Hebrew and Aramaic as a third branch of Northwest Semitic, or put into a separate Central (Western) branch – resulting in a Semitic family tree with a shape like that of Figure 14.2:



**Figure 14.2** The Semitic language family

Using a more wave-like model of language differentiation, of course, one could simply observe the obvious correlation of geographical location and linguistic subgrouping and postulate that in the area (roughly noncoastal Syro-Palestinian) in which Semitic, having split from Afro-Asiatic, began to differentiate into various branches, different features were spreading from different centers of innovation. The central position of Arabic in this network of changes enabled it to share both in innovations originating from northern (reformation of the present tense, generalization of /t/ in the past) and in innovations from the south (\*p > f, elaboration of internal plurals).

### 1.3 Ge'ez literature

The corpus of written Ge'ez material can be conveniently divided into three groups of texts.

#### 1.3.1 Axumite and Pre-Axumite monumental inscriptions

The “prehistory” of this corpus is formed by about 160 Old South Arabian texts, many of them occurring in or near the core of the later Aksumite Ethiopian kingdom, and attested from around the sixth century BC. About 13 of these are “royal” in content. The Aksumite corpus itself is formed by about a dozen longish royal inscriptions in Ge'ez, the most important of which concern at least two kings called Ezana (perhaps mid-fourth century and late fifth century AD). Six of the Ge'ez inscriptions are written in the Old South Arabian script, 3 in nonvocalized Ethiopic, and 4 in the earliest attestation of vocalized Ethiopic script (see §2). The inscriptions of Ezana I are pagan, as are the first ones of Ezana II. The last ones of Ezana II attest to the introduction of monotheism (presumably Christian) to the court at Aksum, while those of his successors are explicitly Christian. The last few inscriptions may be as late as the ninth to eleventh century. There are also about 230 other short Aksumite inscriptions in vocalized and unvocalized Ethiopic – at least 9 of them from a period before Ezana I. Related to this corpus are 18 Greek inscriptions found on Aksumite territory, of which 6 are royal. At least 3 of the major royal inscriptions exist in three versions, Greek, Ge'ez written in Old South Arabian script, and Ge'ez written in Ethiopic script.

#### 1.3.2 Early Christian texts

Although there are few, if any, extant manuscripts earlier than the twelfth century, scholars have isolated a corpus of texts which represent the earliest layer of Ge'ez literature – the first texts drawn up to define and propagate Christianity in Ethiopia. The process of translation of these texts from Greek had begun by the fifth century, and later royal inscriptions contain explicit citations from the Book of Psalms. This body of texts includes the Ge'ez translation of the Bible and related apocrypha, liturgical texts, some lives of saints, some patristic fragments, and a version of the monastic Rules of Pachomius. Although the original translations date from a period when Ge'ez was still a spoken language, we know that many of the texts were revised in the light of standard Arabic redactions (particularly the Arabic “Vulgate” Bible). In view of this revision, especially given the lack of a long manuscript tradition, it is sometimes difficult to establish exact details of some aspects of this earliest manuscript corpus.

#### 1.3.3 Ge'ez: post-1000 AD

After a very obscure period of isolation starting with the collapse of Byzantium in the Near East, and continuing during the first centuries of Islam, the Ethiopian church reestablished around the year 1000 an official contact with Egypt that would last until the end of World War II. There was a new flourishing of ecclesiastical literature of all genres (much of it



translated from the Arabic, in turn translated from Greek, Coptic, Syriac, or other sources). In addition, an original secular or court literature arose in the form of royal chronicles, legal texts, even a sort of national epic (the *Kəbrä Nəgäst* "Glory of Kings," an elaboration of the legend of Solomon and Sheba). A more popular magic literature also took shape, centered around the production of amulets and "magic scrolls" – a productivity that continued into the present century.

## 2. WRITING SYSTEMS

The Ethiopian consonantal script is derived more or less proximately from some version of the South Semitic writing system which also appears in the Epigraphic South Arabian inscriptions (see Ch. 15, §2). The earliest Aksumite inscriptions are written without indication of vocalization. In a major innovation, introduced with apparent abruptness in the later inscriptions of Ezana II, an approach to vocalization appears which is unique among Semitic scripts. Alone among these scripts, Ethiopic represents vowels, not by a separate set of (superlinear or sublinear) vowel signs, but by means of a more or less uniform modification of the basic letter-shape. The base form of the consonant sign (the so-called *first-order* form) is taken to represent the consonant followed by the *unmarked* short vowel /ä/. Six other alterations of the basic shape are introduced to represent the consonant followed by the vowels /i, u, a, e, ə, o/ – the *second* through *seventh orders*. Note that the sixth order is used in addition to represent the consonant in isolation (with no following vowel). Each sign thus represents a CV sequence, and a 26-character consonantal script is thereby transformed into a 182-character syllabary. To this inventory are added five forms for each of the consonants *k, g, q,* and *h* to represent the labiovelars /k<sup>w</sup>, g<sup>w</sup>, q<sup>w</sup>, h<sup>w</sup>/ plus the vowels /ä, i, a, e, ə/, for a grand total of 202 signs. The resulting basic syllabary, in the traditional Ethiopic order (which is also the order followed by dictionaries of Ethiopic languages which give head words in Ethiopic script), can be seen in Tables 14.1 and 14.2.

In fact, an influence from the Indian subcontinent cannot be excluded a priori in the development of this vocalized system – some of the first mentions of the Aksumite state are in itineraries of voyagers between the Mediterranean, the Red Sea and Indian Ocean spice coasts, India, and Ceylon. On the other hand it is not far-fetched to see some of the character modifications (note for example the *omicron*-like shape of many of the "seventh-order" character modifications) as vocalic subscripts influenced by the shape of the corresponding Greek vowel, as in Syriac vocalization. There is also a mystery concerning the historical origin of the order of consonants in the syllabary, which is quite different from that of the Hebrew, Aramaic, and Arabic scripts. However, there is some evidence that at least the beginning of this order – *h, l, ḥ, m* . . . – was known and used in ancient Ugarit, and also in Southern Arabia.

As an addition to the syllabary as such, Table 14.3 gives the Ethiopic number notation system. Although the general ductus and appearance of the numerals is influenced by those of the syllabary, the system is clearly borrowed from the Greek α, β, γ, δ, . . . = 1, 2, 3, 4, . . . and so forth.

Although the Ethiopic writing system provided a generally adequate representation of Ge'ez words (and continues to provide the same service for Amharic and Tigrinya), there are three aspects of phonological shape which are not directly represented: (i) stress, frequently not noted in practical orthographies in any case; (ii) lack of a way of indicating the phonologically prominent and morphologically important feature of gemination (a two-dot diacritic, introduced to represent gemination of consonants in European grammars and dictionaries of Ethiopic since the seventeenth century, never became part of the manuscript or printed Ethiopic orthographic tradition); and (iii) lack of an unambiguous



**Table 14.1 The Ge'ez syllabary**

	ä	u	i	a	e	ə	o
<i>h</i>	ሀ	ሁ	ሂ	ሃ	ሄ	ህ	ሆ
<i>l</i>	ለ	ሉ	ሊ	ላ	ሌ	ል	ሎ
<i>ḥ</i>	ሐ	ሑ	ሒ	ሓ	ሔ	ሕ	ሖ
<i>m</i>	መ	ሙ	ሚ	ማ	ሜ	ም	ሞ
<i>ś</i>	ሠ	ሡ	ሢ	ሣ	ሤ	ሥ	ሦ
<i>r</i>	ረ	ሩ	ሪ	ራ	ራ	ረ	ሮ
<i>s</i>	ሰ	ሱ	ሲ	ሳ	ሴ	ስ	ሶ
<i>q</i>	ቀ	ቁ	ቂ	ቃ	ቄ	ቅ	ቆ
<i>b</i>	በ	ቡ	ቢ	ባ	ቤ	ብ	ቦ
<i>t</i>	ተ	ቲ	ቲ	ታ	ቲ	ት	ቶ
<i>ḥ</i>	ሳ	ሳ	ሳ	ኃ	ኄ	ኅ	ኆ
<i>n</i>	ነ	ኑ	ኒ	ና	ኔ	ን	ኖ
<i>ʔ</i>	ኦ	ኦ	ኦ	ኦ	ኦ	ኦ	ኦ
<i>k</i>	ከ	ከ	ከ	ካ	ኬ	ክ	ኮ
<i>w</i>	ወ	ወ	ወ	ወ	ወ	ወ	ወ
<i>ʕ</i>	ዐ	ዐ	ዐ	ዐ	ዐ	ዐ	ዐ
<i>z</i>	ዘ	ዘ	ዘ	ዘ	ዘ	ዘ	ዘ
<i>y</i>	የ	የ	የ	የ	የ	የ	የ
<i>d</i>	ደ	ደ	ደ	ደ	ደ	ደ	ደ
<i>g</i>	ገ	ገ	ገ	ገ	ገ	ገ	ገ
<i>ṭ</i>	ጠ	ጠ	ጠ	ጠ	ጠ	ጠ	ጠ
<i>p̣</i>	ጸ	ጸ	ጸ	ጸ	ጸ	ጸ	ጸ
<i>ṣ</i>	ሸ	ሸ	ሸ	ሸ	ሸ	ሸ	ሸ
<i>ḍ</i>	ፀ	ፀ	ፀ	ፀ	ፀ	ፀ	ፀ
<i>f</i>	ፈ	ፈ	ፈ	ፈ	ፈ	ፈ	ፈ

**Table 14.2 The labiovelar symbols**

	ä	i	a	e	ə
<i>q</i>	ቁ	ቁ	ቁ	ቁ	ቁ
<i>ḥ</i>	ሳ	ሳ	ሳ	ሳ	ሳ
<i>k</i>	ከ	ከ	ከ	ከ	ከ
<i>g</i>	ገ	ገ	ገ	ገ	ገ

way of representing a consonant not followed by a vowel, i.e., in word- or syllable-final position (note the parallel ambiguity inherent in the Hebrew schwa symbol). As a consequence, an orthographic representation  $k_1l_6b_6$  (where  $C_n$  is the  $n$ th-order shape of the consonant  $C$ ) might conceivably stand for any of the following values:

- (1)  $kälb$      $käləb$      $kälbə$      $källəb$   
 $käləbb$      $käləbbə$      $källəbb$      $källəbbə$

(excluding many values such as those including  $/\#kk\dots/$  or  $/\dots llbb\#/$  or  $/\dots ləbə\#/$  on general phonotactic grounds).

**Table 14.3 The Ethiopic numerals**

1	ᐃ	11	ᐃᐃ	20	ᐃᐃ
2	ᐃᐃ	12	ᐃᐃᐃ	30	ᐃᐃᐃ
3	ᐃᐃᐃ	13	ᐃᐃᐃᐃ	40	ᐃᐃᐃᐃ
4	ᐃᐃᐃᐃ	14	ᐃᐃᐃᐃᐃ	50	ᐃᐃᐃᐃᐃ
5	ᐃᐃᐃᐃᐃ	15	ᐃᐃᐃᐃᐃᐃ	60	ᐃᐃᐃᐃᐃᐃ
6	ᐃᐃᐃᐃᐃᐃ	16	ᐃᐃᐃᐃᐃᐃᐃ	70	ᐃᐃᐃᐃᐃᐃᐃ
7	ᐃᐃᐃᐃᐃᐃᐃ	17	ᐃᐃᐃᐃᐃᐃᐃᐃ	80	ᐃᐃᐃᐃᐃᐃᐃᐃ
8	ᐃᐃᐃᐃᐃᐃᐃᐃ	18	ᐃᐃᐃᐃᐃᐃᐃᐃᐃ	90	ᐃᐃᐃᐃᐃᐃᐃᐃᐃ
9	ᐃᐃᐃᐃᐃᐃᐃᐃᐃ	19	ᐃᐃᐃᐃᐃᐃᐃᐃᐃᐃ	100	ᐃᐃᐃᐃᐃᐃᐃᐃᐃᐃ
10	ᐃᐃᐃᐃᐃᐃᐃᐃᐃᐃ			200	ᐃᐃᐃᐃᐃᐃᐃᐃᐃᐃᐃ
				1,000	ᐃᐃᐃᐃᐃᐃᐃᐃᐃᐃᐃᐃ
				10,000	ᐃᐃᐃᐃᐃᐃᐃᐃᐃᐃᐃᐃᐃ
				100,000	ᐃᐃᐃᐃᐃᐃᐃᐃᐃᐃᐃᐃᐃᐃ

### 3. PHONOLOGY

Our principal source of explicit information for phonology is the living pronunciation tradition of ecclesiastical Ge'ez – a basically oral tradition which is only sporadically, and for the most part relatively recently, recorded in any written form. This tradition appears to be fairly uniform, and represents prestige pronunciation of Ge'ez in the central Ethiopian plateau (thus in a largely Amharic milieu) where royal residences and many centers of ecclesiastical influence have tended to be located since the decline of Aksum. This pronunciation tradition is thoroughly Amharicizing in its treatment of consonantal values, although more “Tigrinya-izing” pronunciation patterns seem to exist in the north. To the extent that the traditional pronunciation preserves stress, gemination, and syllable-structure patterns, which are at least in part distinct from those found either in Amharic or in Tigrinya, they may well reflect the state of affairs in an earlier stage of Ge'ez itself.

In any case, although the basic features of the pronunciation tradition are relatively clear, there is still much to be done by way of scholarly investigation and evaluation of this tradition. It must also be kept in mind that, apart from the relatively small corpus of Aksumite inscriptions, apparently from the early and formative period of classical Ge'ez, almost all Ge'ez texts were either produced in a period when Ge'ez was no longer a spoken language, or are preserved in a long, poorly studied manuscript tradition, with a gap of many centuries between the period of formation of the core classical corpus (Bible translation, key liturgical, hagiographic, and monastic texts – perhaps sixth century) and the oldest extant manuscripts (rarely older than the fourteenth century).

#### 3.1 Consonants

In evaluating phonological representations of Ge'ez, it is important to keep a number of things distinct:

1. The conventional scholarly transliteration of the Ge'ez writing system. This is largely governed by conventional Semitist, largely Arabicizing notation under the constraint of providing one transliteration symbol for each consonant in the Ge'ez writing

system. Western scholarly pronunciations of Ge'ez tend to be unduly influenced by this transliteration notation. All Ge'ez citations in this chapter are in this traditional notation.

2. A largely Amharicizing traditional pronunciation that simply ignores consonantal distinctions which do not exist in Amharic.
3. The most completely preserved Ethiopian Semitic inventory of the original consonantal distinctions (not including, of course, new consonantal distinctions introduced since the disappearance of Ge'ez – for example, many new palatalized spirants and affricates). Here Tigrinya will serve.
4. The pronunciation of the corresponding consonants as reconstructed for South Arabian.
5. The corresponding Arabic consonants.
6. For control, the system as it appears in Hebrew.

A number of Ge'ez consonants will have the same representation in all five systems: *t*, *d*, *k*, *g*, *f*, *ʔ*, *m*, *n*, *r*, *l*, *w*, *ɣ*. Another series of consonants, the labiovelars *kʷ*, *qʷ*, *gʷ*, *hʷ*, are unique as such in Semitic, but correspond in cognates with the nonlabialized *k*, *q*, *g*, *h*. For the others, the cognate sets yield the correspondence series of Table 14.4:

**Table 14.4 Semitic consonantal correspondence series**

South				Central	North
Ethio-Semitic					
Ge'ez (translit.)	Ge'ez (trad.)	Tigrinya	Old South Arabian	Arabic	Hebrew
P	P	P	—	—	—
Ṗ	P'	P'	—	—	—
f	f	f	f	f	p
ṭ	t'	t'	t' <ṭ>	ṭ	ṭ
q	k'	k'	k' <q>	q	q
s	s	s	s <s <sub>3 <td>s</td> <td>s</td> </sub>	s	s
s	s	s	š <s <sub>1 <td>s</td> <td>š</td> </sub>	s	š
s	s	s	θ	θ	š
ś	s	s	ṣ <s <sub>2 <td>š</td> <td>ś</td> </sub>	š	ś
ṣ	s'	s'	s' <ṣ>	ṣ	ṣ
ḍ	s'	s'	ṭ' <ḍ>	ḍ	ṣ
ḍ	s'	s'	θ' <ḍ>	ḍ	ṣ
z	z	z	z	z	z
z	z	z	ḏ	ḏ	z
ḥ	h	ḥ	x	x	ḥ
ḥ	h	ḥ	ḥ	ḥ	ḥ
ʕ	ʔ	ʕ	ʕ	ʕ	ʕ
ʕ	ʔ	ʕ	ɣ	ɣ	ʕ

Details of interpretation are given below, but it is important to note the following regarding Table 14.4: the traditional pronunciation (column 2) gives consonantal signs essentially their Amharic value – pharyngeals *ʕ* and *ḥ* (/ħ/) merge with *ʔ* and *h* respectively; <ṣ> and <ḍ> are both pronounced *ṣ*, and <s> and <ś> are both pronounced *s*. Tigrinya (as well as Tigre) preserves the distinction between *ʔ* and *ʕ* and between *ḥ* and *h*. But *no Ethiopic language*

and *no element of the pronunciation tradition* provides the least bit of information about the pronunciation of the Ge'ez graphemes <ś>, <ḏ>, and <ḥ>. Nevertheless, since the earliest pre-Aksumite writing system did adopt these consonant signs from the parent South Semitic alphabet, while excluding a large number of other signs representing consonants which had already merged in Ethiopic, they must have represented distinct consonants in early and classical Ge'ez. It is not clear when the mergers took place. The graphemes seem to be used consistently in the earlier monumental inscriptions, and in some strands of the manuscript tradition (recall, however, what was already said concerning the primitive state of the study of this tradition). However, variant writings of the same word with <ḥ> and <ḥ> begin to appear already in some late monumental inscriptions, and, in the low end of the manuscript tradition, <ḥ, ḥ> (and sometimes <h>), <ḏ, ṣ>, and <ś, s> are used as virtual allographs. In Table 14.5 the consonants corresponding to <ś>, <ḏ>, and <ḥ> are interpreted with the help of data from cognate languages (note especially the South Arabian correspondences in Table 14.4).

### 3.1.1 Voiceless labials (the graphemes < p, p̣, f >)

Ge'ez is unique among the Semitic languages in having not only a voiceless labial stop and continuant, but an *emphatic* labial stop as well (see §3.1.2). Akkadian, Aramaic, and Canaanite have only a /p/, while Arabic and other South Semitic languages have only /f/. Ge'ez with /f/ thus patterns historically, as expected, with South Semitic and Arabic. A large number of the occurrences of /p/ and /p̣/ occur in loanwords, mostly from or by way of Greek: for example, *pilās* “temple, gate” (from Greek πύλη (*pýlē*) “gate”); *pāpp* as “metropolitan, patriarch” (from πάππας (*pāppās*) “father, title of priests”). However, there are a large number of other occurrences, fully integrated into the native grammar and vocabulary of Ge'ez, where the origin of the stop is much less clear: for example, *heṗä* “strike, throw, shoot with an arrow”; *hāppälä* “wash clothes.”

### 3.1.2 Emphatic consonants (the graphemes < p̣, ṭ, q, q<sup>W</sup>, ṣ, ḏ >)

As can be seen from Table 14.5, a coarticulatory feature usually called “emphasis” in Semitic is realized as *glottalization* in modern Ethiopian Semitic. Since the Arabic realization of this feature, *velarization* or *pharyngealization*, was once automatically imputed to early Semitic, glottalization in Ethio-Semitic was ascribed to the influence of language contact with earlier Cushitic languages in Ethiopia. However, the relatively recent discovery that emphatic consonants are also glottalized in the Modern South Arabian languages makes it possible that this might be a common South Semitic feature, and perhaps even, as some have argued, common Semitic. Note that /ṣ/ (ṣ), the glottalized version of /s/, in modern Ethio-Semitic languages, as in many languages worldwide, tends to be realized phonetically as an affricate [t̪ʃ]. The phonetic value of Ge'ez ḏ is uncertain; however, since it merged with the fricative ṣ (see §3.1.3), it had also become a continuant by the time of merger.

### 3.1.3 Sibilants (the graphemes < s, ś, ṣ, ḏ, z >)

The consonants represented by the graphemes <ś> and <ḏ> have merged respectively with *s* and *ṣ* in the phonological system represented by the traditional pronunciation – and, indeed, in all modern Ethiopian Semitic. These two consonants are reflexes of a lateralized series (voiceless and glottalized) in Proto-Semitic, also attested in South Arabian. There is, however, no evidence either in the tradition or in Ethiopian Semitic as to what value these consonants may have had in Ge'ez. For <ḏ> the transcription value *ḏ* comes from the conventional representation of the etymologically corresponding segment in Arabic and Old

South Arabian; while *ṣ* is an older conventional representation of the Proto-Semitic voiceless lateral, and also of the grapheme which represents its Hebrew reflex. In some grammars and dictionaries, Ge'ez <ṣ> is transcribed as *š*, since it corresponds etymologically to Arabic /š/. There are, however, some major problems with this practice. In the first place, it is not certain whether, or at what periods, Ge'ez <ṣ> might have been pronounced as [š]. More seriously, this transcription could lead to confusion, since a genuine /š/ did in time develop in Ethiopic Semitic (mostly from palatalization of /s/), and a new grapheme for it, properly transcribed as *š*, was created by adding a diacritic to the grapheme <s>. Moreover, this <š> grapheme *can* occur in late Ge'ez texts, usually in modern personal or place names.

### 3.1.4 Laryngeals (the graphemes <ʔ, ʕ, ḥ, ḥ, ḥ>)

As can be seen from Table 14.4, in the pronunciation tradition the only accepted values for these consonants are [ʔ] for the first two, and [ḥ] for the remainder. Moreover, as already noted, no Ethio-Semitic language has kept *ḥ* distinct from *ḥ*. Finally, in Ge'ez, as in many varieties of Semitic, there is no phonological distinction in word-initial position between simple vocalic (#[V-]) and glottal (#[ʔV-]) onset, even though the writing system has to use a glottal-stop symbol (<ʔ>) to “carry” the vowel.

### 3.1.5 Labiovelars (the graphemes <k<sup>w</sup>, q<sup>w</sup>, g<sup>w</sup>, ḥ<sup>w</sup>>)

All of the velars of Ge'ez developed a corresponding labiovelar phoneme. In some cases, there is an unambiguous conditioning environment with a (long) rounded vowel: thus, the denominal verb *tārg<sup>w</sup>āmā* “translate” comes ultimately from the Aramaic loanword *targūm* “translation”; *ḥ<sup>w</sup>* “brother” shows the influence of the Proto-Ethiopic, and Proto-Semitic, long stem-vowel in \**aḥū-*. In other cases, derivation from a form with such an environment must be assumed.

### 3.1.6 Summary of consonantal features

With the foregoing reservations, the articulatory features of Ge'ez consonants will be interpreted as in Table 14.5:

**Table 14.5 Consonants of Ge'ez**

Manner of articulation	Place of articulation						
	Labial	Dental/Alveolar	ʔ	Velar	Labiovelar	Pharyngeal	Glottal
<i>Stop</i>							
<i>Voiceless</i>	p	t		k	k <sup>w</sup>		ʔ
<i>Glottalized</i>	p̰	t̰		q	q <sup>w</sup>		
<i>Voiced</i>	b	d		g	g <sup>w</sup>		
<i>Fricative</i>							
<i>Voiceless</i>	f	s	ś	ḥ	ḥ <sup>w</sup>	ḥ	h
<i>Glottalized</i>		š		ḏ			
<i>Voiced</i>		z				ʕ	
<i>Sonorant</i>							
<i>Nasals</i>	m	n					
<i>Liquids</i>		r, l					
<i>Glides</i>		y			w		

## 3.2 Vowels

We can assume that Proto-Ethio-Semitic possessed the common Semitic vowel system, with three short vowels, three long, and two diphthongs:

(2)	Short	Long	Diphthong
	i	ī	
	u	ū	
	a	ā	ai au

The Ge'ez system then resulted from a series of changes:

- (3) A. \*i, \*u → ə  
 B. \*a → ä  
 C. \*ai → e, \*au → o

As a consequence of A and B, the quantity of the long vowels was made redundant, and they became simply the unmarked low, high-front, and high-back vowels of the system: in other words, \*ā > a, \*ī > i, \*ū > u. The monophthongizing of the diphthongs to /e, o/ rounded out the system, which is attested in the earliest vocalized texts, and remains remarkably stable even in many of the modern Ethiopian Semitic languages:

	FRONT	CENTER	BACK
HIGH	i	ə	u
MID	e		o
		ä	
LOW		a	

**Figure 14.3** Vowels of Ge'ez

Here /ä/ is a low-central vowel, higher and more forward than /a/, secondarily perhaps also shorter. Note that in historicizing transliterations, what are here noted as the vowels /ä/ and /a/ are written as /a/ and /ā/ respectively.

## 3.3 Phonotaxis and syllable structure

The preferred syllable-type is (C)V(C), thus with no initial or final clusters; maximally, clusters of two consonants are allowed intervocalically.

### 3.3.1 Word-initial clusters

Word-initial clusters are resolved by epenthetic ə. Thus the imperative, whose systematic form for the simple stem of the verbs *ngr* “speak” and *lbs* “wear” would be /ngər/, /lbäs/, becomes *nəgər*, *ləbäs*. There is no productive rule for breaking up clusters with an initial vowel, but there are some isolated lexical patterns such as *əgzīʔ* “lord” from the root *gzʔ* “rule”; note also the form of some common complementizers and conjunctions *əsmä* “because,” *ənzä* “while,” *əskä* “until” (of uncertain etymology, although the first may be connected to *səm* “name”).

### 3.3.2 Word-final clusters

Word-final clusters are less clear. Most studies of traditional pronunciation seem to agree that word-final clusters of two consonants are not invariably broken up by an epenthetic vowel. Thus, Makonnen (1984) transcribes: *baḥrə* “sea”; *ṭäbbäbtə* “wise men” (masculine plural of *ṭäbib*); *yəblə* “he says, he will say” (irregular imperfective of *bəhlä*). The most common Western scholarly practice would be to pronounce the first two with a final cluster ([*baḥr*], [*ṭäbbäbt*]), and the last with epenthetic or reduced [ə], on the strength of the presumed underlying pattern: [*yəbəl*] reduced from *\*yəbəhhəl*, from *\*yəbahhəl* (by laryngeal rule [4B] below).

### 3.3.3 Gemination

Gemination is a widely employed inflectional and derivational process in Ge’ez, as in the rest of Ethio-Semitic. In the traditional pronunciation, all consonants can geminate except the laryngeals. In general, gemination seems to be limited to vowel-pattern environments in which the introduction of gemination will not give rise to problematic consonant clusters.

### 3.3.4 Laryngeal effects

A series of rules affect the vowels /ä, ə/ in the vicinity of laryngeals (ʔ, ʕ, ḥ, ʕ, h/), which find their most productive application in the conjugation of verbs containing a laryngeal as a radical (see §4.3.3). The following is simply a descriptive statement of the phenomena, with some illustrative examples (L = laryngeal):

- (4) A. ə → ä / \_ L<sub>ä</sub>  
 Ex.: *yäḥärrəs* “he plows” (cf. *yənäggər* “he speaks”)  
 B. ä → ə / \_ L [V, + high] ([V, + high] = /i, ə, u/)  
 Ex.: *yəməḥər* “he is merciful” (cf. *yənäggər*)  
 C. ä → ə / \_ L<sub>final</sub> V  
 Ex.: *asməʕ-ä* “he caused to hear” (cf. *albäsä* “he clothed”)  
 D. ä → a / \_ L {C, #}  
 Ex.: *sämaʕ-ku* “I heard” (cf. *läbäsku* “I wore”)  
 E. ä → a / L \_  
 Ex.: *ḥaqäfä* “he embraced” (cf. *nägärä* “he spoke”)

### 3.3.5 Glide effects

There are a number of interactions of vowels and glides. As can be seen from the chart of labiovelar graphemes (Table 14.2), a labiovelar cannot be followed by a rounded vowel. On the other hand, there is a neutralization of the distinction between a labiovelar followed by /ə, ä/ and a velar followed by /u, o/: for example, *qwal* ~ *qul* “bunch of grapes”; *qwäṭiṭ* ~ *qoṭiṭ* “slender.” Most of the vowel-glide configurations which arise in the morphology can be handled by the following rules:

- (5) A. äw → o (*fätäwkä* ~ *fätokä* “you loved”)  
 B. äy → e (*sätäykä* [sätekä rare] “you drank”)  
 C. \*əw → u (*\*yəfättəw* → *yəfättu* “he loves”)  
 D. \*əy → i (*\*yəsättəy* → *yəsätti* “he drinks”)

Of these rules, the first is optional, the second is rare, while the third and fourth are obligatory.

### 3.3.6 Consonant assimilation

This involves only the two consonants /t/ and /k/, widely utilized in affixation in Ge'ez. The velar /-k/, which occurs in suffixes marking first- and second-person subject in the verb, assimilates to a preceding (i.e., stem-final) velar stop: for example, *ädäg+kä* > *ädäggä* “you left”; *säräq+ku* > *säräqqu* “I stole.” The dental /t-/ occurs as a prefix marking passive-reflexive derived stems in the imperfective tense (the shape of the prefix in the perfective is *tä-*). It assimilates fully to a following dental stop or sibilant: thus, *yässämmä-* “he is heard”; cf. *yəṭqättäl* “he is killed.” As a suffix /t-/ marks feminine (and also plural) forms of nouns and participles. Here it assimilates to a preceding dental stop: for example, *kəbud+t* > *kəbədd* “heavy” (fem.). Note, however, idiosyncratically “one” (fem.) *ahatti* < *ahad+ti*, “daughter” +t. *wälätt* < *wäläd*.

## 3.4 Stress

Our knowledge of stress depends completely on the still inadequately studied traditional pronunciation. At present some general patterns seem to hold:

1. Verbs are stressed on the penult except in the second-person plural feminine: *yəngər* “may he speak,” *yənäggər* “he speaks,” *nägärä* “he spoke,” *yənäggəru* “they speak”; but *nägärkən* “you (2nd pl. fem.) spoke.”
2. Nouns and pronouns have stem-final stress (i.e., not on the suffix vowel of the accusative; see §4.1.1.3): *nəgús* (nom.), *nəgüsä* (acc.) “king.”
3. Personal pronouns as well as verbs and nouns with pronominal suffixes follow special patterns, giving rise to minimal pairs like *yənäggərə* (< *yənäggər+ha*) “he speaks to her” versus *yənäggəra* “they (fem.) speak.”

## 4. MORPHOLOGY

### 4.1 Nominal morphology

Nominal morphology in Ge'ez expresses three morphosyntactic categories: gender (masculine, feminine), number (singular, plural), and case (absolute, accusative-construct). Under this heading will be treated nouns as well as adjectives and participles.

#### 4.1.1 Nouns

##### 4.1.1.1 Gender

In comparison with other classical Semitic languages, gender is less systematically marked in the morphology of the noun (see §4.1.2.1 for adjectives). There is a suffix *-t* that occasionally marks feminine nouns which are paired with a masculine noun using the same stem: for example, *bəʔsi* “man,” *bəʔsit* “woman”; *əgziʔ* “lord,” *əgziʔt* “lady”; *əḥiʔ* “brother,” *əḥət* “sister.”



4.1.1.2 *Number*

Plural can be marked by suffixation or by internal vowel change. The common pluralizing process is by suffixation of *-at* (from common Semitic feminine plural, used in Ge'ez for both genders): thus, *may*, *mayat* (plural) “water”; *ṣəge*, *ṣəgeyat* (plural) “flower”; *ʕaša*, *ʕašat* (plural) “fish.” Nouns with feminine formative *-t* may or may not drop this before the suffix: *ṣərəṣt*, *ṣərəṣt-at* (plural) “law,” but *ʕäzäq-t*, *ʕäzäq-at* (plural) “well.”

A great many Ge'ez nouns, however, form their plural according to one of the internal (so-called broken) plural patterns, if necessary using “underlying” glides or supplemental consonants to make up the canonical consonants of the pattern. The vast majority of triconsonantal internal plurals follow one of six patterns:

1. *CVCäC*: while probably the most archaic, this pattern is not the most productive, but does include some of the most basic lexical items. The pattern is probably to be connected with the one internal plural pattern that can be assigned to Proto-Semitic, exemplified in Hebrew by the so-called *shegolate plurals* (*melek* ~ *mlakim* < \**malk-* ~ *malak-*), and indeed to Afro-Asiatic: for example, *əzn* ~ *əzän* “ear”; *əgr* ~ *əgär* “foot”; *əd* ~ *ədäw* “hand,” *ab* ~ *abäw* “father,” *əḥ* ~ *aḥäw* “brother.”
2. *aCCaC*: this is the most productive pattern for triconsonantal nominal stems. Many biconsonantal stems that can be analyzed as *CwC* or *CyC* take this pattern, as well as a few *CC* stems that become *CCt* in the plural: for example, *ləbs* ~ *albas* “garment”; *färäs* ~ *afras* “horse”; *bet* ~ *abyat* “house”; *ṣom* ~ *aṣwam* “fast”; *səm* ~ *asmat* “name.”
3. *aCCuC*: this pattern, a special Ethiopian development, seems to occur most frequently with initial laryngeal stems: for example, *adg* ~ *aʔ dug* “ass”; *hägär* ~ *ahgur* “city.”
4. *aCCəCt*: this pattern is a variant of the above, with the addition of a final *-t* and a reduction of *u* to *ə*, which originally had to be a shortening in closed syllable (i.e., \**ū* → \**u* / \_\_\_\_ CC: this is one of the few synchronic reflexes of the original length distinction between *u* < \**ū* and *ə* < {\**i*, \**u*}): for example, *rəʔs* ~ *arʔəst* “head”; *gäbr* ~ *agbərt* “slave.”
5. *CäCaCəC(t)*: this very productive (and common Semitic) quadriliteral pattern occurs on almost all noun stems with four consonants, as well as with a number of nouns having three consonants and at least one so-called “long” stem-vowel – *i*, *e*, *o*, or *u* (see §3.2). Consider the following examples: *dəngəl* ~ *dänagəl* “virgin”; *mäsfən* ~ *mäsafənt* “prince”; *kokäb* ~ *käwakəbt* “star”; *mäskot* ~ *mäsakut* (< *mäsakəwt*) “window”; *dorho* ~ *därawəh* “chicken”; *lelit* ~ *läyaləy* “night”; *bəḥer* ~ *bähawərt* “earth”; *wəḥiz* ~ *wähayəzt* “river”; *qäsis* ~ *qäsawəs* “priest.”
6. *aCaCəC(t)*: this is another way of extending the quadriliteral pattern just discussed to triconsonantal stems: thus, *bägʔ* ~ *abagəʔ* “sheep”; *ganen* ~ *aganənt* “devil.”

Note that the glide inserted to fill out the *triliteral* (CCC) or *quadriliteral* (CCCC) pattern is not generally predictable from the nature of the vowel. Although there are numerous exceptions, there is a tendency toward a polarity pattern in forms with optional *-t*: *-t* is added in the plural if it is absent in the singular (unless the noun is feminine), and absent in the plural if it is present in the singular.

Finally, an additional morphological plural marking occurs with all plural forms (suffix or internal) followed by possessive suffixes: an *-i-* is inserted between the noun and the suffix. Thus corresponding to *ṣəgeyat* “flowers,” *abyat* “houses,” we have *ṣəgeyatina* “our flowers,” *abyatina* “our houses.”

#### 4.1.1.3 Case

Common Semitic had a three-case system: a nominative in \*-u, a genitive in \*-i, and an accusative in \*-a (cf. Akkadian *kalb-u-m*, *kalb-i-m*, *kalb-a-m*; Arabic *kalb-u-n*, *kalb-i-n*, *kalb-a-n*). In Ethiopic the merger of short, high vowels ( $\{*i, *u\} \rightarrow \text{ə}$ ), plus the eventual disappearance of final ə ( $\text{ə} \rightarrow \phi / \_\_\_ \#$ ), automatically neutralized the distinction between nominative and genitive, and reduced the nominal case system to a single morphologically marked case form in Ge'ez, the *accusative-construct*, henceforth simply *accusative*. It is opposed to an unmarked *absolute* form, identical with the stem, which we will refer to as *nominative*. The accusative is formed by suffixation of -ä to the unmarked form of the noun: thus, nominative *bet* “house,” accusative *betä*. This form continues the inherited function of the direct object of a verb, as in: *särhā nəgus betä* “The/a king built the/a house” (lit. “built king-NOM. house-ACC.”). It is also used for the head (first) noun in the so-called *construct* configuration, as in: *betä nəgus* “the/a house of the/a king” (lit. “house-ACC. king-NOM.”).

In both object and possessive constructions, morphological indication of case can be replaced by syntactic paraphrase. In the case of the direct object, the construction *Verb Noun-ACC.* can be replaced by *Verb+object suffix lä Noun-NOM.*, where *lä* is the preposition “to.” Thus, instead of *särhā betä* “he made the/a house,” one can have *särho* (< *särhā+hu*) *lä bet* “he made the house” (note that the prepositional paraphrase tends to be preferred for definite direct objects). In lieu of the construct *Noun<sub>1</sub>-ACC. Noun<sub>2</sub>-NOM.*, there are two possibilities: either *Noun<sub>1</sub> zä Noun<sub>2</sub>*, where *zä* is the relative pronoun; or *Noun<sub>1</sub>+possessive suffix lä Noun<sub>2</sub>*. Thus, instead of *betä nəgus* one can have either *bet zä nəgus* or *betu* (< *bet+hu*) *lä nəgus* (where the latter variant may be preferred for a definite head noun).

#### 4.1.2 Adjectives and participles

The gender- and number-marking systems of this morphological class have undergone less simplification than was the case with the noun. The class includes the following:

1. General adjectives of many canonical shapes, of which we will use *sānay* “beautiful” as typical.
2. A special class of quality adjectives of the well-known Semitic form masculine singular *CāC(C)iC*: *hāddis* “new,” *ṣābiyy* “big.”
3. The present or active participle of the verb, having the form *CāCaCi* (*qātali* “killer”). This replaces an older *\*CāCiC* common Semitic pattern with a compound reflex of *\*CaC(C)āC+ī* (a habitual agent nominalization, plus a relational – so-called *nisbe* – denominal adjective formation).
4. The passive or intransitive participle of the verb, of the form *CəCuC* (*qətul* “killed,” Proto-Semitic *\*qutūl*).

##### 4.1.2.1 Gender

Generally gender is marked in conjunction with the number category (see below for paradigms); most adjectives and participles form the feminine singular by suffixing *-t*, and keep the *-at* suffix for the feminine plural (as opposed to the noun, which generalizes the suffix to masculines). The masculine plural is marked by suffixing *-an*. Some masculine (and feminine) plurals are formed by internal vowel change, plus suffix *-t*; and adjectives of the *CāCCiC* form take a special *CāCCaC* shape.

4.1.2.2 *Number*

The paradigms for the classes of §4.1.2 are as follows:

(6)		<i>Singular</i>	<i>Plural</i>
	<i>Masc.</i>	śānay	śānayan
	<i>Fem.</i>	śānayt	śānayat

The active participle has a special masculine plural form:

(7)		<i>Singular</i>	<i>Plural</i>
	<i>Masc.</i>	nāgari	nāgärt
	<i>Fem.</i>	nāgarit	nāgariyat

The passive participle and the *CäCCiC* adjectives have a special feminine singular form:

(8)		<i>Singular</i>	<i>Plural</i>
	<i>Masc.</i>	nəgur	nəguran
	<i>Fem.</i>	nəgərt	nəgurat
	<i>Masc.</i>	ḥāddis	ḥāddisan
	<i>Fem.</i>	ḥāddas	ḥāddisat

Some *CäCCiC* forms have a common plural like the active participle:

(9)		<i>Singular</i>	<i>Plural</i>
	<i>Masc.</i>	ʕābiyy	ʕābbäyt
	<i>Fem.</i>	ʕābbay	ʕābbäyt

4.1.2.3 *Case*

Case is marked as on the noun.

4.2 *Pronouns*

Within Semitic, the personal pronouns of Ge'ez offer interesting pattern similarities and contrasts:

(10) *Independent pronouns*

	<i>Akkadian</i>		<i>Arabic</i>		<i>Ge'ez</i>	
	<i>Sg.</i>	<i>Pl.</i>	<i>Sg.</i>	<i>Pl.</i>	<i>Sg.</i>	<i>Pl.</i>
<i>1st</i>	anāku	nīnu	ʔana	naḥnu	anä	nəḥnä
<i>2nd masc.</i>	atta	attunu	ʔanta	ʔantum	antä	antəmu
<i>2nd fem.</i>	atti	attina	ʔanti	ʔantunna	anti	antən
<i>3rd masc.</i>	šū	šunu	huwa	hum	wəʔətu	wəʔətomu/əmntu
<i>3rd fem.</i>	šī	šina	hiya	hunna	yəʔəti	wəʔəton/əmantu

As can be readily seen, the first- and second-person independent pronouns of Ge'ez are fairly straightforward representatives of Common Semitic, whereas a certain amount of idiosyncratic innovation has taken place in the third-person independent pronouns. The suffix pronouns, on the other hand, object and possessive, show predictable Semitic forms:

## (11) Suffix pronouns

	Akkadian		Arabic		Ge'ez	
	Sg.	Pl.	Sg.	Pl.	Sg.	Pl.
1st	-ī, -nī	-ni	-ī, -nī	-nā	-Vyä/-Vni	-Vnä
2nd masc.	-ka	-kunu	-ka	-kum	-Vkä	-Vkəmu
2nd fem.	-ki	-kina	-ki	-kunna	-Vki	-Vkən
3rd masc.	-šu	-šunu	-hū	-hum	-hu~-u~-o	-homu~-omu
3rd fem.	-ši	-šina	-hā	-hunna	-ha~-a	-hon~-on

Note that the first singular suffix is -Vyä with nouns and -Vni with verbs. The stressed or unstressed vowel (V) with first- and second-person forms is, for nouns, the stem-final vowel, or vocalic suffix, if there is one, or ə if the noun form ends in a consonant. For verbs it is ä (except for second-person object suffixes with jussive verb forms, compare *yənäggərākkä* “he speaks to you” and *yəngörkä* “may he speak to you”). In the third person, in nouns and verbs ä+hú/há/hómu/hón gives ó/á/ómu/ón (as in *nägärä+hú* > *nägäro* “he spoke to him”; *betä+hú* > *betó* “his house” [acc.]) In nouns C+hV gives CV (as in *bet+hú* > *betú* “his house” [nom.]).

The deictic-relative paradigms are built on the stem-series *z-* (singular, mostly masculine), *ənt-* (most feminine singulars), *əll-* (plural), corresponding to \**ḏ-* (masculine), \**t-* (feminine), and \**l-* (plural) in Common Semitic. Far deixis adds the element -*ku*. Both near and far have a “long” form with suffix -*tu*~*ti*~*tä*. The paradigm of the whole deictic–relative–interrogative series is as follows:

(12)		Singular		Plural	
		Nom.	Acc.	Nom.	Acc.
“this”	Masc.	zə-	zä	əllu	
	Fem.	za	za	əlla	
“this” (long)	Masc.	zəntu	zäntä	əllontu	əllontä
	Fem.	zati	zätä	əllantu	əllantä
“that”	Masc.	zəku	zək <sup>w</sup> ä, zəku	əlləku	
	Fem.	əntəku	əntək <sup>w</sup> ä əntäku	əlləku	
“that” (long)	Masc.	zəktu	zəktä	əlləktu	əlləktä
	Fem.	əntakti	əntaktä	əllaktu	əllaktä
Relative	Masc.	zä-	əllä		
	Fem.	əntä		əllä	

The members of the interrogative series have the \**mVn-* and \**ʔay-* shapes also known from Common Semitic:

(13)		Singular		Plural	
		Nom.	Acc.	Nom.	Acc.
“who?”		männu	männä		
“what?”		mənt	məntä		
“which?”		ay		ayat	ayatä

## 4.3 Verbs

Although each Semitic language has its own elaboration and adaptation of the system of verbal inflection, there is a common core of categories and formal processes which is visible

in the different systems, and which permits a fairly accurate approximation of the essential features of the proto-system. In all Semitic languages, finite verb paradigms can be analyzed into two “subparadigms”: one marks *person* by the use of subject affixes; the other accounts for the *stem-form* – marks the combination of root, derived stem, and tense-mode.

#### 4.3.1 Person morphology: subject affixes

As can be seen from (14), the prefixing subparadigms differ very little from one another within the major branches of Semitic, and presumably all continue fairly directly an ancestral Proto-Semitic system:

##### (14) Prefixing/Suffixing

	Akkadian		Arabic		Ge'ez	
	Sg.	Pl.	Sg.	Pl.	Sg.	Pl.
1st	a-	nī-	a-	na-	ʔə-	nə-
2nd masc.	ta-	ta-...-ā	ta-	ta-...-ū	tə-	tə-...-u
2nd fem.	ta-...-ī	ta-...-ā	ta-...-ī	ta-...-na	tə-...-i	tə-...-ā
3rd masc.	i-	i-...-ū	ya-	ta-...-ū	yə-	yə-...-u
3rd fem.	ta-	i-...-ā	ta-	ta-...-na	tə-	yə-...-ā

The suffixing subparadigms, however, seem to have been subject to a certain amount of transformation (see §1.2):

##### (15) Suffixing

	Akkadian		Arabic		Ge'ez	
	Sg.	Pl.	Sg.	Pl.	Sg.	Pl.
1st	-āku	-ānu	-tu	-nā	-ku	-na
2nd masc.	-āta	-ātunu	-ta	-tum	-ka	-kəmmu
2nd fem.	-āti	-ātina	-ti	-tunna	-ki	-kən
3rd masc.	φ	-ū	-a	-ū	-a	-u
3rd fem.	-at	-ā	-at	-na	-at	-ā

On the one hand, the first- and second-person suffix forms, especially in the Akkadian, look very much like reduced enclitic forms of the independent pronouns given in (10) above, with the *-ā-* perhaps as some linking element. The third-person forms, on the other hand, look like elements from the nominal inflection – as a matter of fact, the Akkadian stative can be formed on the basis of any noun or adjective (e.g., *damq-at* “she is good,” *šarr-āku* “I am king”). The major difference in the Arabic-like languages (Canaanite, Aramaic) is that the *-k- ~ -t-* alternations in the paradigm have been leveled out in favor of *-t-*, whereas in Ge'ez and Modern South Arabian they have been leveled out in favor of *-k-*. As was pointed out by Hetzron 1972, this is important evidence in favor of the temporal priority of an Akkadian-like system over the other two. It is now usually supposed, therefore, that the Proto-Semitic system resembled the Akkadian one, and that the Western Semitic past tense evolved out of something like a verbal adjective with enclitic reduced pronouns.

#### 4.3.2 Stem morphology

The second series of subparadigms governs the stem-form to which the subject markers are affixed. There are three basic morphological categories involved, which we will refer to as (i) *tense*, (ii) *derivational class*, and (iii) *lexical class*.

#### 4.3.2.1 *Tense*

The primary “tense” categories are (i) past, (ii) non-past, (iii) jussive, and (iv) imperative. The imperative stem is identical with the jussive stem in Ge'ez, and hence will not be noted in the paradigms below. The infinitive is not a tense in any usual morphosyntactic sense of the term, but it will be useful to display infinitives in the paradigm as a fourth “tense” form.

#### 4.3.2.2 *Derivational class*

The derivational classes are (i) *base* (zero-affix); (ii) *causative* (prefix *ä-*); (iii) *passive-reflexive* (prefix *tä-*, if not preceded by a subject prefix, otherwise *t-*); and (iv) *causative-passive* (prefix *ästä-*). The prefixes of (ii)–(iv) continue the common Semitic derivational stems: thus, (ii) causative \**š/h* giving *ʔ* in Arabic and Ethiopic, with all three possibilities existing in Old and Modern South Arabian; (iii) prefix and infix *t(a)* in Akkadian, Cananite, Aramaic, Arabic, and South Arabian; (iv) the combination \**š+t(a)* is attested as such in Arabic, South Arabian, and Akkadian, and as \**h+t(a)* in Northern Semitic. As in all Semitic languages, these derivational classes are formed, with more or less idiosyncratic semantics, with all verbs in the lexicon, although not all verbs occur in all derivational classes.

#### 4.3.2.3 *Lexical class*

Historically related to the various Semitic verbal derivation systems, but almost completely lexicalized in Ge'ez are the categories of lexical class, conventionally designated in Ge'ez with the letters A, B, and C. A is the unmarked class. In the base past and jussive there are two subclasses: A1 has stem-vowel *ä* in the past and stem-vowel *ə* in the jussive; whereas A2 has stem-vowel *ə* in the past and *ä* in the jussive. Recalling that Ge'ez *ə* represents Proto-Semitic \**i* and \**u*, this clearly corresponds to the common stem-vowel alternations (e.g., in Arabic): past *CaCaC* ~ present-jussive *CCuC*, past *CaCiC* ~ present-jussive *CCaC*. Note that some verbs can be A1 in the past and A2 in the jussive, and vice versa. B is the class of verbs with geminating middle radical (Piel in Hebrew, D-stem in Akkadian, Second Form in Arabic); C is the class of verbs with stem-vowel /*ā*/ after the first radical consonant (Third Form in Arabic). Unlike other Semitic languages, these do not occur in Ge'ez as derived forms of the unmarked base, but as a lexically determined class. A verbal root must be marked in the lexicon as class A, class B, or class C in Ge'ez, and if it occurs in one class, it will not occur in another (the few cases where this occurs are usually counted as homophonous). An exception to this general rule is the class of passive-reflexive C (*tānagärä*, cognate with the Sixth Form of Arabic) and causative-passive C (*ästānagärä*) which occur with many verbs, the former frequently as a reciprocal, the latter with widely varying semantics.

Given the existence of this set of lexically determined categories, it is convenient to enter quadriradical verbs (verbs with four root consonants) under this general heading as lexical class D. These verbs are especially frequent in Ethiopic Semitic. There are only a few cases where they can be etymologically linked to triradical verbs, but phonologically most of them are either of the form  $C_1RC_2C_3$  ( $R = /n, l, r/$ ) or of the form  $C_1C_2C_3C_3$ , where  $C_1, C_2, C_3$  otherwise follow the co-occurrence constraints of triconsonantal roots. As is often the case in Semitic, these D (quadriradical) verbs closely resemble B (middle-geminating) verbs in their morphological structure.

#### 4.3.2.4 *Sample paradigms*

Table 14.6 gives the stem-paradigms for the so-called “strong” verbs (verbs with roots that do not have a glide or vocalic radical; see §4.3.3.1). The lexical items used are *ngr* “speak”

**Table 14.6 Strong verb-stem paradigms**

		Base	Causative	Pass.-refl.	Caus.-pass.
<i>Past</i>	A1	nägär-	angär-	tänägr-	astängär-
	A2	läbs-	albäs-	täläbs-	astälbäs-
	B	fäššäm-	afäššäm-	täfäššäm-	astäfäššäm-
	C	masän-	amasän-	tämasän-	astämasän-
	D	dängäš-	adängäš-	tädängäš-	astädängäš-
<i>Pres.-fut.</i>	A	-näggər(-)	-anäggər(-)	-tnäggär(-)	-astänäggər(-)
	B	-feššəm(-)	-afeššəm(-)	-tfeššäm(-)	-astäfeššəm(-)
	C	-masən(-)	-amasən(-)	-tmasän(-)	-astämasən(-)
	D	-dänäggəš(-)	-adänäggəš(-)	-tdänäggəš(-)	-astädänäggəš(-)
<i>Jussive</i>	A1	-ngər(-)	-angər(-)	-tnägär(-)	-astängər(-)
	A2	-lbäs(-)	-albəs(-)	-tläbäs(-)	-astälbəs(-)
	B	-fäššəm(-)	-afäššəm(-)	-tfäššäm(-)	-astäfäššəm(-)
	C	-masən(-)	-amasən(-)	-tmasän(-)	-astämasən(-)
	D	-dängəš(-)	-adängəš(-)	-tdängəš(-)	-astädängəš(-)
<i>Infinitive</i>	A	nägir(ot)	angəro(t)	tänägro(t)	astänägro(t)
	B	fäššəmo(t)	afäššəmo(t)	täfäššəmo(t)	astäfäššəmo(t)
	C	masno(t)	amasno(t)	tämasno(t)	astämasno(t)
	D	dängəšo(t)	adängəšo(t)	tädängəšo(t)	astädängəšo(t)
<i>Gerund</i>	A	nägir-	angir-	tänägir-	astänägir-
	B	fäššim-	afäššim-	täfäššim-	astäfäššim-
	C	masin-	amasin-	tämasin-	astämasin-
	D	dängiṣ-	adängiṣ-	tädängiṣ-	astädängiṣ-

(Class A1); *lbs* “wear” (Class A2); *fšm* “finish” (Class B); *msn* “perish” (Class C); *dngš* “surprise” (Class D, quadriradical).

### 4.3.3 Strong and weak verbs

The distinction strong versus weak root is bound up with the fact that Semitic inflection and derivation typically involves a triconsonantal *strong* root, such as *lbs* “wear,” into which are inserted different vowel patterns (combined with different prefixes, suffixes, and infixes): for example, *läbsä* “he wore”; *ləbs* “clothes”; *ləbsät* “dressing”; *mälbäs* “clothing”; *albas* “articles of clothing”; *albäsä* “he clothed”; *täläbsä* “get dressed”; *tälabäsä* “disguise oneself”; *astälabäsä* “clothe several persons”; and so forth. A number of words, however, do not have three true consonants, and the place of the “missing” consonant is occupied by a glide *w* or *y* (sometimes also *n* and <sup>?</sup>), yielding *weak* roots of the form *wCC*, *yCC*, *CwC*, *CyC*, *CCw*, *CCy*, and so forth. Since some of the oldest and most widespread Semitic vocabulary (*wld* “bear child,” *qwm* “stand,” *bky* “cry”) is of this form, it has been suggested that the unique phonological organization involved in triconsonantal strong roots is a tendency (already incipient in Afro-Asiatic) that only gradually overtook large portions of the Semitic lexicon and morphology, and that the weak roots are an attempt to fit older, nontriconsonantal lexical items onto the innovative triconsonantal patterns. Semitic languages differ in the extent to which weak verbs are assimilated to the strong patterns – in general, Akkadian tends to be the most conservative (i.e., least assimilating), and Ge’ez the most assimilating.

Arabic is somewhere in the middle, and Hebrew is somewhat more conservative than Arabic. Reflexes of the archaic weak root \**mwt* “die” are illustrated in (16):

(16)	Akkadian		Arabic		Ge'ez	
	Present	Past	Present	Past	Present	Past
3rd sg. masc.	i -mūat	i -mūt	ya -mūt- u	māt- a	yə -mawwət	mot- a
2nd sg. masc.	ta -mūat	ta -mūt	ta -mūt- u	mut- ta	tə -mawwət	mot- ka
3rd pl. masc.	i -mutt- ū	i -mūt- ū	ya -mūt- ū- na	māt- ū	yə -mawwət- u	mot- u

The finite base A form (see §§4.3.2.2; 4.3.2.3) of triconsonantal glide roots is illustrated in (17) using the lexical items *wrd* “descend,” *wdq* “fall” (W1); *mwt* “die” (W2); *śym* “appoint” (Y2); *ftw* “love,” *bdw* “be desert” (W3); *bky* “cry,” and *sty* “drink” (Y3). The other lexical and derivational classes are straightforward extensions of the base A form. Note that for medial *w/y* verbs (W2/Y2) of the A class, there is no distinction between subclass A1 and A2, and that initial-*y* verbs (Y1) are very few in number, and have been largely regularized.

(17)			<i>Past</i>	<i>Present</i>	<i>Jussive</i>
	W1	A1	wäräd-	-wärrəd(-)	-räd(-)
		A2	wädq-	-wäddəq(-)	-däq(-)
	W2	A	mot-	-mäwwət(-)	-mut(-)
	Y2	A	śem-	-śäyyəm(-)	-śim(-)
	W3	A1	fätäw/fäto-	-fättu(-)	-ftu(-)
		A2	bädw-	-bäddu(-)	-bdäw(-)
	Y3	A1	bäkäy-	-bäkki(-)	-bki(-)
		A2	säty-	-sätti(-)	-stäy(-)

Verbs which have a laryngeal (*L*; i.e., *ʔ*, *ʕ*, *ħ*, *h*) as a radical largely follow the strong pattern, as modified by the special vowel–laryngeal sequence constraints noted above (see §3.3.4): thus for LCC, third masculine singular – past *ʕäqäbä* “he kept,” present *yäʕäqqəb* (< \**yäʕäqqəb*, by laryngeal vowel harmony). Many CLC verbs are also completely “regular,” as *sähäbä*, *sähäbkä* “he, you (sg. masc.) pulled”; however, an important subclass of these verbs displays a past stem-vowel pattern with *ə*: *səḥtä*, *səḥətkä* “he, you (masc. sg.) erred.” CCL verbs of the A class are idiosyncratic in that they are all of the A2 (*läbsä*) subclass: *wädʔä*, *wädakä* “he, you (masc. sg.) left.” In addition, CCL verbs of the B, C, and D class have the unique property in the Ge'ez conjugation system of also having a “*läbsä*-like” pattern in the past: *läqqəḥä*, *läqqəḥkä* “he, you (masc. sg.) lent”; *baləḥä*, *baləḥkä* “he, you (masc. sg.) rescued,” *zängəʕä*, *zängəʕkä* “he, you (masc. sg.) raved.”

Finally, there are a dozen or so verbs, most with glide or laryngeal radicals, which show one or more idiosyncratic irregularities in stem-paradigm. An important one from the historical point of view is the unique (and archaic) conjugation pattern of the verb *bhl* “say” in its base form (the derived class forms are conjugated regularly). Instead of an expected past tense \**bəhlä* (compare *kəhlä* from *khl* “be able”), this verb has a prefixing past tense, the only survival of this archaic form in Ethiopic Semitic, with stem *-be* in nonsuffixed forms, *-bel-* in suffixed: *yəbe*, *yəbelu* “he, they said.” The present stem of this verb is *-bəl(-)*, and its jussive is *-bal(-)*: *yəbəl* “he says,” *yəbal* “let him say” (compare *yəkəl* “can,” *yekal* “let him be able”).



## 4.4 Adverbs and prepositions

A number of adverbs are productively formed from accusatives (suffix *-ä*) of nouns and adjectives: for example, *lelitä* “by night”; *qādmä* “in front”; *rəḥuqä* “afar”; *mārirä* “bitterly.” Note that – although there is no direct etymological relation – an accusative-like form is also the norm for many conjunctions: thus, *əmä* “if”; *sobä* (earlier *sobe*) “when”; *ənbalä* “except”; *ənzä* “while”; *askä* “until.”

On the other hand, an adverbial (hence nominal) origin is clear for many prepositions: for example, *māngälä* “towards,” *maʾkälä* “between.” Most of the usual prepositions end in *-ä* before nouns, and *-e* before pronominal suffixes: thus, *ḥäbä* ~ *ḥäbe-* “to, towards”; *dibä* ~ *dibe-* “on”; *mäslä* ~ *mäsle-* “with” (e.g., *mäsläsäbʾ* “with (the) man,” *mäslehu* “with him”). Note the following special cases: *əmännä* (proclitic form *əm-*) ~ *əmännē-* “from”; *kämä* ~ *käma-* “like”; *wästä* ~ *wēstet-* “in”; *ḥäwdä* ~ *ḥäwdä-* “around”; *askä* (does not occur with pronominal suffixes) “until, up to.” The monosyllabic prepositions are proclitic: *bä-* “in” (*bä-* before first- and second-person pronouns, third-person singular *bo* ~ *bottu*, *ba* ~ *batti*); *lä-* “to” (except *litä* “to me,” *lottu* “to him,” *latti* “to her,” *lon* ~ *latton* “to them” [fem. pl.]).

## 4.5 Numerals

The Geʿez number system shows a number of archaisms. For one thing, even though gender marking is considerably reduced in the noun, the Geʿez cardinal numbers show a faithful continuation of the common Semitic gender polarity switch, with a *t*-suffixed form in the masculine, and an unmarked form in the feminine. In addition to these forms, Geʿez has a great variety of derived forms, the most important of which are the ordinals, chiefly of the pattern *CaCəC* (< \**CāCiC*). There are also day-of-week/month forms (*CāCuC* pattern), and adverbial forms (“once,” “twice,” etc.; of the *CəCC* pattern). For the numbers 1 to 10, these forms are as follows:

(3)	Cardinal		Ordinal	Day	Adverbial
	Masc.	Fem.			
1	aḥädu	aḥati	qādami	əḥud	məʾrā, aḥätä
2	kəʾe, kəʾetu	kəʾeti	dagəm, kaləʾ, kaʾəb	sānuṣ	kaʾbä, dagmä
3	šälästu	šälas	šaləs	šäluṣ	šəlsä
4	arbaʾtu	arbaʾ	rabəʾ	räbuʾ	rəbʾä
5	ḥäməstu	ḥäms	ḥaməs	ḥämuṣ	ḥəmsä
6	sədəstu	səssu	sadəs	säduṣ	sədsä
7	säbʾätu	säbʾu	sabəʾ	säbuʾ	səbʾä
8	sämäntu	sämani	samən	sämun	səmnä
9	təsʾätu, täsfätu	təsʾu, täsfu	tasəʾ	täsuʾ	təsfä
10	ʾäśərtu	ʾäśru	ʾäśər	ʾäśur	ʾäśrā

Except for the day nominalization, which uses the inherited Semitic root *sny* (< \**ṯny*), the inherited root of the numeral “two” has been replaced by \**kil* ʾ-, the Semitic word for “both.” Other nominalizations involving “two” call upon the lexical items *dgm* “repeat” and *kfb* “double.” The ordinal for “one” uses the lexical item *qdm* “precede.” The ordinals provide the only remnant in Geʿez morphology of the common Semitic active participle of the form \**CāCiC* > *CaCəC*, replaced by a new form in Geʿez (see §4.1.2). The masculine cardinals have an accusative form in *-tä*, and have *-ti-* before suffix pronouns: *šälästä* “three” (acc.), *šälästihomu* “the three of them.” Feminine cardinals are usually treated as invariants.



Nominal modifiers generally follow the head noun, as in the following illustrations of adjective, numeral, and relative constructions (and see *passim* for possessive):

- (21) A. noba qäyḥ (189, 37)  
Nubian red  
“red Nubians”  
B. mägäbtä kəlʔetä (189, 23)  
leader [ACC. PL.] two [ACC.]  
“two leaders”  
C. ḥäwarya-nä zä-fänäwku l-ottu (189, 11)  
messenger=our which=send [PAST, 1ST SG.] to=him  
“our messenger which I sent to him”

However the order modifier–noun is also possible, and for numerals seems to be even more frequent:

- (22) A. ärbäʔtu ʔängadä (188, 15)  
four [MASC.] tribe [ACC. PL.]  
“four tribes”  
B. ʕəsra wä-sälusä mäwaʕälä (189, 16)  
twenty and=three day [ACC. PL.]  
“twenty-three days”

The use of enclitics such as *-ssa* for fronting major sentence constituents, widely used in the standard translation literature to correspond to Greek constructions with μέν (*mén*) and δέ (*dé*), appears already in a sixth-century Geʾez text from Marib, where it, in fact, is a citation of Psalm 19, 8:

- (23) əmuntu-ssä bā-ʔäfras wä-bä-särägälat wä-nəḥnä-ssä nəʕabi  
they=indeed in=horses and=in=chariots and=we=indeed glory  
bä-səmə əgziʔä bəḥer (195, 27)  
in=name [ACC.] lord [ACC.] earth  
“They indeed in horses and chariots, but we glory in the name of God”

## 5.2 Subordination

### 5.2.1 Relative clauses

As exemplified in (21), relative clauses are introduced by the relative pronoun *zä-* (singular and proclitic), *ʔällä* (plural). Word order follows the general pattern of main clauses, although with more of a tendency for the verb to come at the end:

- (24) ʔəgziʔä sämay zä-wähäbä-ni ʔəgziʔä kwəlu  
lord [ACC.] heaven who=give [PAST, 3RD MASC. SG.]=me lord [ACC.] all  
zä-b-ottu ʔämänku  
who=in=him believe [PAST, 1ST SG.]  
“the Lord of Heaven who gave me (to be) lord of all, in whom I have believed”

### 5.2.2 Adverbial clauses

Adverbial clauses are introduced by the subordinating conjunctions mentioned in §4.4 and behave in general like relative clauses:

- (25) ḍäbäʔku                      noba      sobe      ʔäḍrärä                      ʔəḥzabä  
 go to war [PAST, 1ST SG.]    Nubia    when    revolt [PAST, 3RD MASC. SG.]    people  
 noba (189, 8)  
 Nubia  
 “I went out to battle against Nubia when the peoples of Nubia revolted”

Note the use of ʔanzä and the past tense use of the imperfective (not present-future):

- (26) baṣṣu                      dāwālā                      ʔägädä      ʔə<n>zä  
 come [PAST, 3RD MASC. PL.]    region [ACC.]    Agada    while  
 yəqattəlu                      wa-yəḍewwəwu  
 kill [IMPF., 3RD MASC. PL.]    and=take captive [IMPF., 3RD MASC. PL.]  
 wa-yəmähärrəku (187, 20)  
 and=plunder [IMPF., 3RD MASC. PL.]  
 “They came to the region of Agada killing and taking captives and plundering”

A subordinate clause expressing sequential action (*conjunctive*, i.e., “and then”) is formed by the infinitive + suffix pronoun – the so-called *gerund* construction – before the main verb, found frequently in later Ge'ez, and indeed in all modern Ethiopic Semitic:

- (27) wä-baṣiḥ-omu                      ʔängäbo      bä-həyā      rākäbä-nä  
 and=come[INF.]=them    Angabo    in-there    find[PAST, 3RD MASC. SG.]=us  
 ʔäbäʔälkəʔo (187, 5)  
 Abaʔälkəʔo  
 “And they having come to Angabo, Abaʔälkəʔo found us there”

In the Aksumite texts this construction also occurs after the main verb:

- (28) wä-däḥnä                      ʔätäwu                      ʔäfrih-omu                      ḍar-omu  
 and=safe [ACC.]    return [PAST, 3RD MASC. PL.]    fear[CAUS., INF.]=them    enemy=their  
 wä-mäwiʔ-omu                      bä-ḥaylä                      ʔəgziʔä      bəḥer (189, 33)  
 and=conquer[INF.]=their    in=power [ACC.]    lord[ACC.]    earth  
 “And they returned safely having terrorized their enemy and conquered by the power of the Lord of the Earth (i.e., ‘God’)”

Simultaneity and adverbial modification can be expressed also paratactically (note the identity of tense in the *main* and *modifying* verbs):

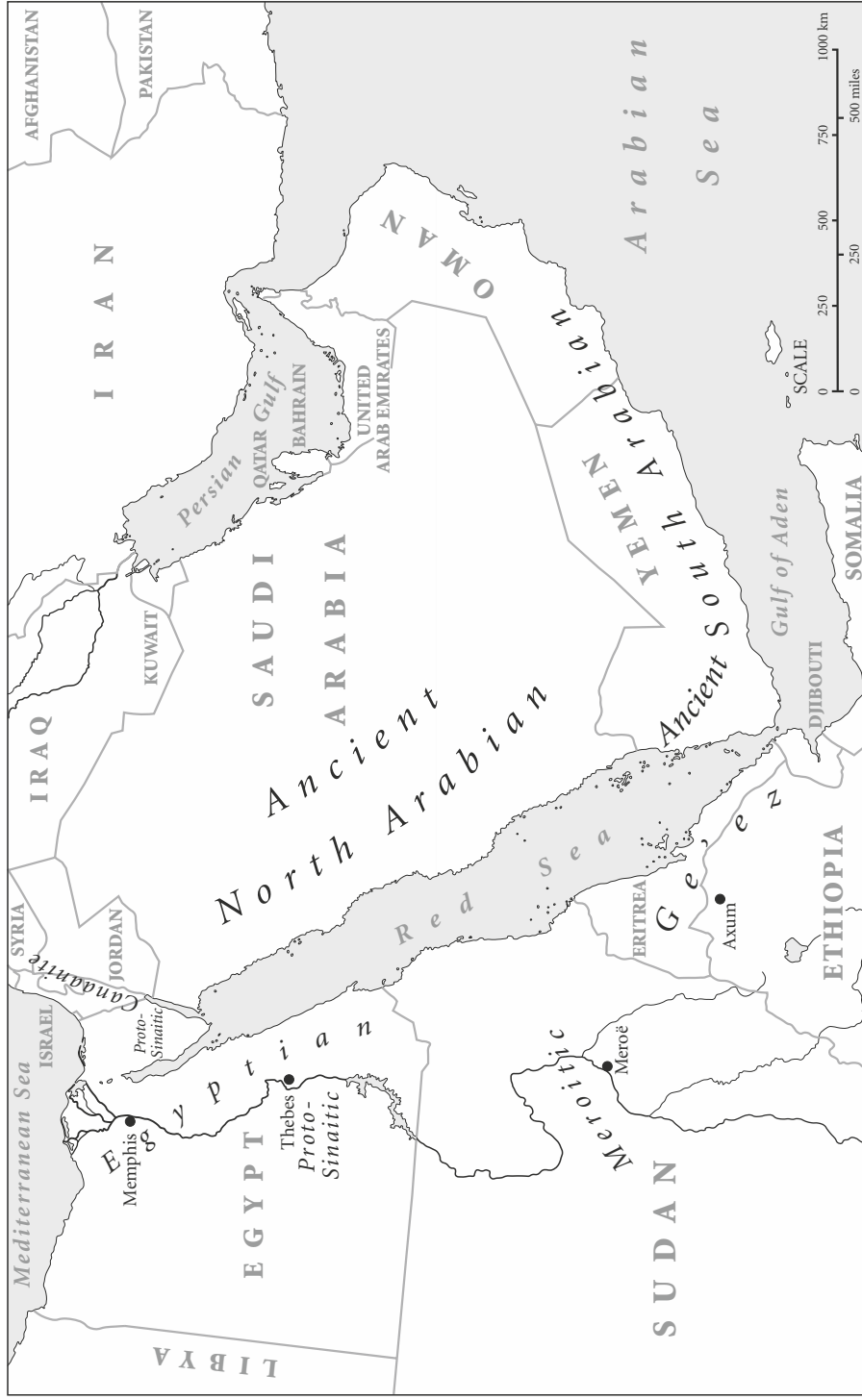
- (29) konäna-homu                      ... kämä yəḥoru  
 order[PAST, 3RD MASC. SG.]=them ... that    go [JUSS., 3RD MASC. PL.]  
 wä-yəʔälu                      wä-yəḥoru  
 and=spend day [JUSS., 3RD MASC. PL.]    and=go [JUSS., 3RD MASC. PL.]  
 wä-yəbitu (187, 13)  
 and=spend night [JUSS., 3RD MASC. PL.]  
 “He ordered them ... that they travel day and night”

In addition to its use in the so-called *gerund* construction – infinitive+pronominal suffix (see [27] and [28]) – the infinitive can also occur as a simple verbal complement:



- 
- Hetzron, R. 1972. *Ethiopian Semitic: Studies in Classification*. Journal of Semitic Studies Monograph 2. Manchester: Manchester University Press.
- . 1997. *The Semitic Languages*. London: Routledge.
- Kobishchanov, Y. 1979. *Axum*. Edited by Joseph Michels; translated by Lorraine Kapitanoff. University Park: Pennsylvania State University Press.
- Lambdin, T. 1978. *Introduction to Classical Ethiopic (Ge'ez)*. Missoula: Scholars Press.
- Leslau, W. 1987. *Comparative Dictionary of Ge'ez*. Wiesbaden: Otto Harrassowitz.
- Makonnen, A. 1984. *Matériaux pour l'étude de la prononciation traditionnelle du guèze*. Paris: Editions Recherche sur les Civilisations.
- Mittwoch, E. 1926. *Die Traditionelle Aussprache des Äthiopischen*. Abessinische Studien, 1. Berlin/Leipzig: de Gruyter.





**Map 1.** The Ancient languages of Northeastern Africa and Arabia



# Ancient South Arabian

NORBERT NEBES AND PETER STEIN

## 1. HISTORICAL AND CULTURAL CONTEXTS

Ancient (or Epigraphic) South Arabian (for terminology see Macdonald 2000:30), which is considered part of the southern branch of the Semitic language family, is divided into four main dialects, Sabaic, Minaic, Qatabanic, and Hadramitic, which are named after the most important peoples of southwest Arabia in the first millennium BC. These peoples founded their towns at the eastern edges of the central Yemeni highlands, in the wadi deltas that lead into Ramlat as-Sabʿatayn, the desert edge of the Rubʿ al-ḥālī, where favorable natural and geographical conditions prevail. Since Ramlat as-Sabʿatayn is also called Ṣayḥad by the medieval Yemenite geographers, the term Ṣayḥadic, coined by A. F. L. Beeston, has also been used in Sabaic scholarship recently as a generic term for the Ancient South Arabian dialects.

The dialect attested for the longest period and by the most inscriptions by far is Sabaic, the core area of which comprises the region of Mārib and Ṣirwāḥ, but which later also extended to a large part of the highland.

The first Sabaic inscriptions begin in the eighth century BC; the first Sabaic monuments of any length that can be dated reliably by a synchronism with Assyrian sources are to be placed in the early seventh century BC (Wissmann 1982:148). Sabaic is documented for a period of over 1,400 years and may be periodized into three main phases: (i) *Early Sabaic*, with mainly boustrophedon inscriptions dated from the eighth to the fourth century BC, and to which also the texts of the following two centuries from the area of Mārib and the highland are assigned; (ii) *Middle Sabaic*, from the first century BC until the end of the fourth century AD – most of the Sabaic documents, in which the dedicatory inscriptions from the Awām-temple in the oasis of Mārib comprise the largest self-contained text corpus, come from this period; (iii) *Late Sabaic*, of the monotheistic period, which ends in the sixth century AD. In the inscriptions from this period the traditional gods are no longer called upon, but only a single divinity (Raḥmānān). The last inscription dated according to the Himyarite calendar comes from the year AD 554/9.

Under Sabaic are also generally subsumed the inscriptions composed in the *Haramic* dialect, which exhibit linguistic influences from North Arabic. Another group of inscriptions in Sabaic were written by the Himyar, a people who first appeared in the southern highlands in the late second century BC; during the second and third centuries AD, they played an ever more important role in South Arabia, until from the fourth century they controlled large parts, and finally all, of Yemen from their capital Ṣafār.

The epigraphic documentation of Minaic, Qatabanic, and Hadramitic begins to increase from the middle of the first millennium BC, as the Sabaeans lose their dominance over South Arabia.

The first inscriptions written in the *Minaic* dialect appear at about the same time as the earliest Sabaic written evidence (eighth century BC) – though in smaller numbers – and come from the ancient cities along the large Wadi Maḏāb, which lies to the northwest of Mārib. Minaic trading colonies, and thus Minaic inscriptions, are also found outside South Arabia, as in the ancient oasis of Dedān (the present-day al-‘Ulā in Saudi Arabia), and even beyond the Arabian peninsula, as on the island of Delos and in Egypt, in testimony to the presence of Minaic merchants far to the north. Minaic disappears as a dialect around the end of the second century BC.

To the southeast of Mārib, in the wadis Ḥarīb and Bayḥān, lies the *Qatabanian* heartland and its main city Timnaʿ. The Qatabanic dialect area, following the area controlled by the Qatabanian kings, extends far to the southwest, to ʿGabal al-ʿAwd (not far from Zafār), and, according to Latin and Greek authors, to the Bāb al-Mandab on the Red Sea. Around the end of the second century AD, 150 years after the destruction of Timnaʿ, Qataban is finally crushed by Sabaʿ and Hadramawt, after which the epigraphic documentation of this dialect ceases.

*Hadramitic* inscriptions are concentrated in the ancient region along the large Wadi Hadramawt in the eastern part of southwest Arabia and in the royal city of Šabwa which, situated at the southwestern entrance to the wadi, plays a significant role in antiquity as the starting point of the incense route. Hadramitic inscriptions are also found a few hundred kilometers southeast of Šabwa in Samārum (modern Ḥawr Rūrī near Salāla in Oman on the coast of the Indian Ocean), which was founded by Hadramitic colonists toward the end of the first century BC. At the beginning of the fourth century AD the Himyar incorporate Hadramawt into their area of control, after which the epigraphic documentation of Hadramitic likewise ends and is replaced by Sabaic.

## 2. WRITING SYSTEM

The Ancient South Arabian writing system, which is also commonly called the *Ancient South Arabian monumental script* and which is used for all Ancient South Arabian dialects equally, is a segmental script of twenty-nine graphemes that primarily, but not exclusively, serve to represent consonants. A striking feature of the Ancient South Arabian script is the geometric form of the letters, which, in the early period, stand in a fixed relationship of height and width to one another and can reach a height of over 30 centimeters in monumental exemplars. In contradistinction to the later North Arabic script the individual letters are not joined to one another, each letter standing rather on its own. Words are separated from each other by a vertical dividing line. The Ancient South Arabian script has no punctuation marks. The direction of writing is horizontal, from right to left. A characteristic of the inscriptions of the Early Sabaic period from the Mārib area is boustrophedon writing, in which the direction of writing changes, and which is later given up in favor of the sinistroke style. A peculiarity specific to the Late Sabaic inscriptions is letters carved out of the stone in relief.

Inscriptions are written primarily on well-worked stone surfaces, stone blocks, or smoothed rock faces. Inscriptions can also, however, be cast in bronze or prepared on iconographic objects of bronze or on coins or amulets, and the like.

At the beginning of the 1970s, the first instances of writing on wooden sticks, in a hitherto unknown minuscule script, were discovered in Yemen. The understanding of these sticks, which come from the Yemenite Ġawf and of which several thousand have come to light in the

**Table 15.1 The Ancient South Arabian consonantal script**

Character	Transcription	Character	Transcription
𐩦	h	𐩨	s <sub>3</sub>
𐩧	l	𐩩	f
𐩪	ḥ	𐩫	ʾ
𐩬	m	𐩭	ʿ
𐩮	q	𐩯	ḏ
𐩰	w	𐩱	g
𐩲	s <sub>2</sub>	𐩳	d
𐩴	r	𐩵	ḡ
𐩷	b	𐩸	t
𐩹	t	𐩺	z
𐩻	s <sub>1</sub>	𐩼	ḏ
𐩽	k	𐩾	y
𐩿	n	𐩿	ṭ
𐱁	ḥ	𐱂	z
𐱃	š		

meantime, is made especially difficult because of the script and the unknown vocabulary. Concerning the contents of the roughly thirty examples published thus far, probably dating to the second/third centuries AD, it can be said at present that they are documents partly written in the form of letters that have to do with legal and economic matters (Ryckmans, Müller, and Abdallah 1994).

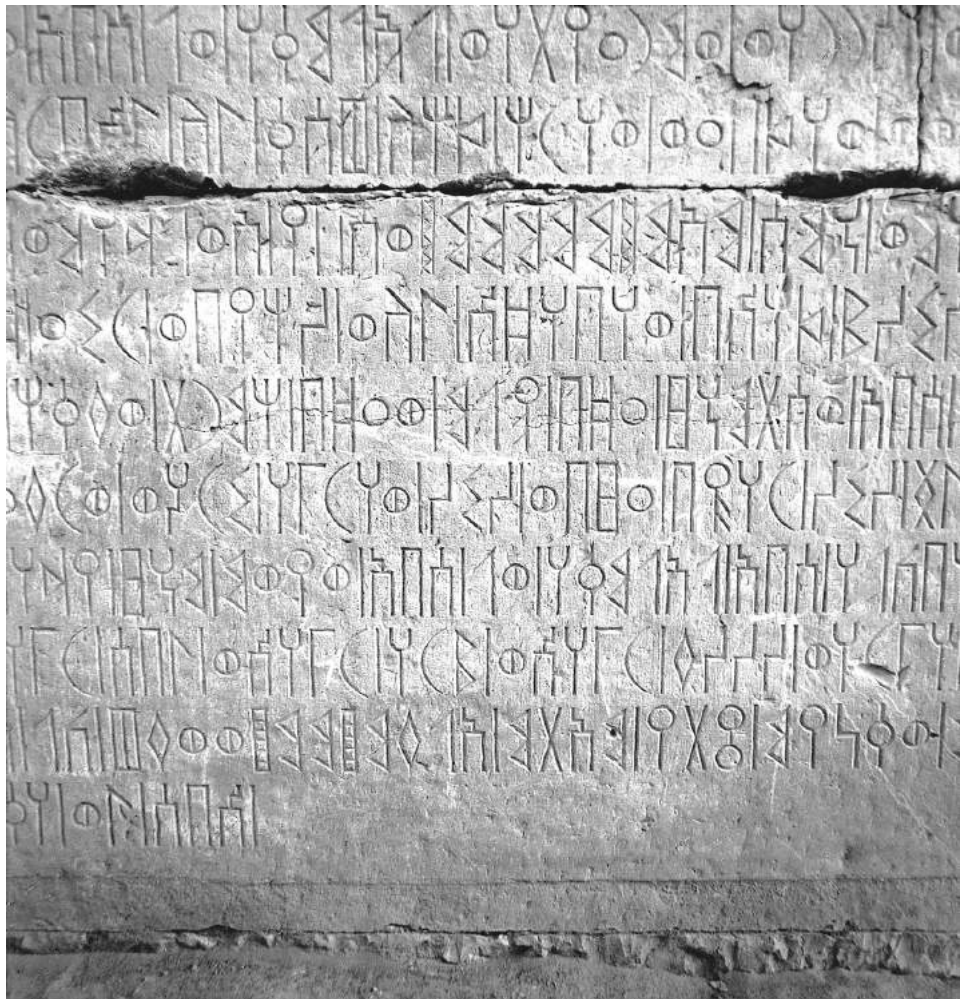
Apart from a large number of graffiti, mostly of personal names, the inscriptions written in the monumental script can be assigned to quite varied text genres. The most widely attested group in all Ancient South Arabian dialects is that of the dedicatory inscriptions, which sometimes contain reports of entire military campaigns. Besides these, building inscriptions, irrigation regulations, grave inscriptions, law texts, and other types of legal documents, as well as so-called penitential and expiatory inscriptions have been found (Müller 1994: 307–312).

In view of the record of documentation, it is principally Sabaic that will be treated in the following summary, with examples from the earlier or later periods noted. The abbreviations used to identify inscriptions are those of *Sab. Dict.* (pp. xx–xxv) and Stein 2003 (pp. 274–290).

### 3. PHONOLOGY

#### 3.1 Preliminary remarks

The Ancient South Arabian writing system is, like that of many other Semitic languages, primarily devised for the representation of consonants, and expresses vowels only in very restricted cases. In the absence of an oral tradition, the precise pronunciation of its graphemes is unknown, and a conventional reconstruction of the sound values is possible only by comparison with other Semitic languages.



**Figure 15.1** The great inscription of Karib'il Watar (c. 685 BC) R 3945: section of lines 11-20

## 3.2 Consonants

### 3.2.1 Consonantal inventory

The following overview gives a rough classification of the Ancient South Arabian consonants according to manner and place of articulation, modeled on the reconstructed Proto-Semitic phonological system.

### 3.2.2 Sibilants

The classification of the graphemes represented above by  $s_1$ ,  $s_2$ , and  $s_3$  was long debated. The usual transcription in the older literature –  $s$ ,  $\text{š}$ , and  $\text{ś}$  – is modeled on the form of the letters and on parallels in classical Arabic, and quickly leads to confusions in etymological comparisons with other Semitic languages which likewise exhibit three distinct “s”-sounds that are, however, transcribed differently. The following shows the correlation between the older and newer transcription systems (for comparison the corresponding Arabic and Hebrew sounds are also given):

(1)	Ancient South Arabian		Classical Arabic	Hebrew
	New	Old		
	s <sub>1</sub>	s	s	š
	s <sub>2</sub>	š	š	ś
	s <sub>3</sub>	ś	s	s

**Table 15.2 The consonantal phonemes of Ancient South Arabian**

Manner of articulation	Place of articulation								
	Labial	Inter-dental	Dental/alveolar	Palato-alveolar	Palatal	Velar/uvular	Pharyngeal	Glottal	
<i>Stop</i>									
<i>Voiceless</i>			t			k		ʔ (/ʔ/)	
<i>Voiced</i>	b		d			g			
<i>Emphatic</i>			ṭ (/tʰ/)			q			
<i>Fricative</i>									
<i>Voiceless</i>		f	ṯ (/θ/)	s <sub>3</sub>	s <sub>1</sub>	ḫ	ḥ (/ħ/)	h	
<i>Voiced</i>			ḏ (/ð/)	z		ġ	ʕ (/ʕ/)		
<i>Emphatic</i>			ẓ (/θʰ/)						
<i>Affricate</i>									
			š (/tʰsʰ/)						
<i>Nasal</i>									
<i>Voiced</i>	m		n						
<i>Lateral continuant</i>									
<i>Voiceless</i>				s <sub>2</sub>					
<i>Voiced</i>				l					
<i>Emphatic</i>				ḏ (/lʰ/)					
<i>Liquid</i>									
				r					
<i>Glide</i>									
<i>Voiced</i>	w				y				

### 3.2.3 Glides

The graphemes *w* and *y* represent primarily consonants – as in other Semitic languages – but also serve as so-called *matres lectionis* (see, *inter alia*, Ch. 16, §2) to indicate vowels (cf. Nebes 1997:114f.). Thus, parallel writings (sometimes in one and the same text, e.g., J 651/12–13, 20) such as *ywm* and *ym* “day” or *byt* and *bt* “house” allow conclusions to be drawn about the occurrence of the monophthongs /o:/ and /e:/ respectively.

## 3.3 Vowels

The few statements that can be made about the vocalization of Ancient South Arabian are based on the use of the glides *w* and *y*. The final writings of plural forms of verbs (Sab. *hqn̄yw* “they dedicated”), personal pronouns (*h̄mw*; *-hw*), and enclitic particles (*-mw*; *-my*) are in all likelihood to be read as vocalic (presumably as long /u:/ and /i:/). The same holds for imperfect forms of verbs II-*w/y*, such as, *ykw̄n* “he will be”, of which defective writings (*ykn̄*) are also attested (see also above, §3.2.3, on monophthongization). Apart from these few hints, practically nothing is known about the vocalization of the Ancient South Arabian

texts preserved for us. The question of a possible differentiation of syllables according to vowel quantity, therefore, likewise cannot be answered with certainty, nor are any broad conclusions possible concerning the accentual relationships of Ancient South Arabian.

### 3.4 Sabaic phonological variation and change

#### 3.4.1 Sound change

Orthographic evidence suggesting various sound changes occurs. The letters *ṣ* and *ẓ* commonly alternate (and also strongly resemble each other in the script), for example in the word for “statuette”, which in the Middle Sabaic dedicatory inscriptions from the Awām temple in Mārib appears as both *ṣlm* and (rarely) *ẓlm* (e.g., J 688/3). In Late Sabaic the sibilant *s*<sub>3</sub> is increasingly replaced by *s*<sub>1</sub>, for example in *ms<sub>1</sub>nd* instead of *ms<sub>3</sub>nd* “inscription”. In some dialects, *w* and *y* alternate in verbal and nominal forms of weak roots in comparison with the Sabaic “Standard”, for example in derivations of *rḏw* “(to have) pleasure”, so that the root is also listed in the dictionaries under *rḏy*, or in *qwl/qyl* “tribal leader”, the plural of which always appears as *ʿqwl* (cf. *ḏwln* beside *qyln* in ‘Abadān 1/40).

#### 3.4.2 Assimilation

As in Hebrew, *n* can assimilate to a following consonant (see Ch. 10, §3.1.2). No firm rules for this phenomenon in Sabaic have thus far been discerned, however, since unassimilated forms are attested just as often in apparently identical contexts (compare, e.g., *hkrrn* “to alter, damage” [infinitive] and *hnkrrn*; *ʿfs<sub>1</sub>* “souls” and *ʿnfs<sub>1</sub>*). In Middle Sabaic, assimilation of *n* seems to be the rule.

#### 3.4.3 Metathesis

In some texts, from the southern dialect area, metathesis is a common phenomenon, which nevertheless appears to be restricted to relatively few words, and particularly to the plurals *ʿywn* (instead of *ʿwyn*) from *wyn* “vineyard” and *ʿlwd* (instead of *ʿwld*) from *wld* “child”.

#### 3.4.4 Regional variation

The texts of certain regions exhibit certain peculiarities that indicate some dialectal coloring of Sabaic. Our grammatical “standard” is based on the texts from Mārib and the central Yemeni highlands.

### 3.5 Non-Sabaic phonological features

In Hadramitic the sounds *s*<sub>3</sub> and *ṭ* have fallen together to a large extent, a development that is expressed in the alternation of the corresponding graphemes. Thus the number “3” appears as *s<sub>2</sub>ls<sub>3</sub>* (e.g., R 2687/5; cf. Sab. *s<sub>2</sub>ṭt*), and the pronominal suffix of the third person feminine as *-s<sub>3</sub>* and *-ṭ*. Similarly, Minaic writes the phoneme /s/ in foreign proper names as *ṭ* (e.g., *dṭt* “Delos” in R 3570/3), but nevertheless keeps the phoneme distinct in the language proper.

Particularly distinctive of Minaic is the insertion of an etymologically unexplained *h* in certain nominal endings, pronouns, and particles (see the forms in the relevant section).



Here too, probably, also belong the Minaic plural forms *bhn* and *bhnt* of *bn* “son” and *bnt* “daughter”. The meaning of these spellings is uncertain. Perhaps they are plene-writings of a long vowel different from /u:/ and /i:/, as is suggested especially by comparison within Semitic. The same phenomenon can be observed in the hadramitic ending *-hn* marking the determinate state.

## 4. MORPHOLOGY

### 4.1 Word structure

Ancient South Arabian shares the fundamental common feature of the Semitic language family: the inflectional morphological system based on a (usually triconsonantal, or *triradical*) root. This means that from a basic scaffold of, as a rule, three consonants, a whole variety of verbal and nominal forms are built by means of the affixing and infixing of a few formative elements (as is common, the root *fʿl* “to do, make” will serve as the paradigmatic root in what follows); semantically, such forms can always be traced back in some way to the basic meaning of the root (e.g., verbal forms of various stems, such as *yfʿlmm*, *fʿl*, *sʿtfʿl*, or nouns such as *mfʿl*, *fʿlt*, etc.). Additional types of morphological differentiation (such as by the lengthening of vowels and consonants) are not visible in the consonantal script, but should be assumed.

As in other Semitic languages, there are also a number of biradical nouns (e.g., *sʿlm* “name”, *yd* “hand”) and a few quadriradical roots (e.g., *kwkb* “star” *sʿ₂ml* Robin-al-Lūmī 1/2 “left, northern”). Verbs that go back originally to biradical roots (so-called weak verbs; see §4.4.3) have largely been brought into line with the triradical system by means of the insertion of “weak” radicals.

Ancient South Arabian distinguishes three numbers (singular, dual, and plural) and two genders (masculine and feminine).

### 4.2 Nominal morphology

#### 4.2.1 Noun patterns

The system of noun patterns in Ancient South Arabian can be only incompletely reconstructed because of the inadequacy of the script; it ought for the most part, however, to have been similar to that of the other Semitic languages in its essential features.

The heavy use of *broken plurals* (see Ch. 6, §3.3.2.4) in Ancient South Arabian is noteworthy. Most of these have the pattern *ʿfʿl*: for example, *ʿbyt* from *byt* “house”; *ʿqwl* from *qwl/qyl* “tribal leader”. In addition to these there are many other forms of broken plurals, such as *mfʿlt* from singular *mfʿl* (e.g., *mḥfdt* from *mḥfd* “tower”; *mqymt* from *mqm* “might”) and the converse (e.g., *mṣnʿ* from *mṣnʿt* “fortress”), or *fʿl* from singular *fʿlt* (e.g., *ʿnt* from *ʿntt* “woman”), as well as the so-called *nisbe*-plural *ʿfʿl(n)* (e.g., *ḥbs₂n* from *ḥbs₂y* “Abessynian”). Further, several different plurals can be formed from one noun (cf. the numerous plural forms of *ḥrf* “year” in *Sab. Dict.*).

In contrast, the external “sound” plural is markedly rare and apparently restricted to a few words.

#### 4.2.2 Noun state

As in other Semitic languages, nouns in Ancient South Arabian exhibit, in addition to number and gender, three states (for the forms in Sabaic see [2] below); a fourth state, the so-called “absolute”, is limited to a few syntactic contexts, mainly numerals (see §4.6.1.1):

1. *Indeterminate state (status indeterminatus)*: marked in the singular as a rule by “mimation” (-m occurring in final position); denotes an indefinite noun: for example, *šlmm d-dhbm* “a(ny) statuette of bronze”; *kl qntnm* Gl 1142/9 “each [type of] flock”.
2. *Determinate state (status determinatus)*: corresponds to the form of a noun marked with a definite article in other languages, for example, *dn šlmm d-dhbn* “this statuette of bronze”; *šlmnhn* “the two statuettes”; *s<sub>2</sub>ltn šlmtn* “the three statuettes [of women]”; *hgrn s<sub>2</sub>bwt* “the city [of] Šabwa”; *kl qwln* “all of the tribal leaders.” Proper names are naturally definite (e.g., *s<sub>2</sub>bwt* in the example above).
3. *Construct state (status constructus)*: in possessive phrases, the form of the governing noun (*nomen regens*) joined to an immediately following genitive; the second (genitive) member of the construction may be a pronominal suffix, a governed noun (*nomen rectum*), or an asyndetic relative clause (see §5.4.2). The accompanying *nomen rectum* is usually definite, but may also be indefinite. As examples consider the following: *ʿwld-hw* “his children”; *ʿdy hlḥhgrnhn ns<sub>2</sub>qm w-ns<sub>2</sub>n* J 643/25 “in the vicinity of the two cities [of] Našqum and Naššān”; *mlky s<sub>1</sub>bʿ* “the two kings of Saba”; *nḏʿ w-s<sub>2</sub>šy s<sub>2</sub>n ʾm* C 407/33 “harm and malice of a[ny] enemy”; *kl s<sub>1</sub>bʿ t w-dbyʿ w-tqdmṯ s<sub>1</sub>bʿy w-dḅʿ w-tqdmn mrʾy-hmw* J 581/6–7 “all expeditions, battles, and attacks, which their two lords led” (the three verbs cannot be meaningfully rendered literally in the translation); compare also §5.4.2. As the last two examples show, several *nomina regentes* may appear in succession in the construct state.

In several instances a genitive relationship is expressed by means of a relative pronoun (see §4.3.3), only rarely, however, by means of apposition, as in *tlṯt šlmm d-dhbm* J 567/9 “three statuettes of bronze” (similarly in rare cases by a following asyndetic relative clause; see §5.4.2).

Because of the lack of vocalization, the Ancient South Arabian case system can only be reconstructed on the basis of the construct state of the external plural of *bn* “son”, which, especially in Early Sabaic texts, appears both as *bnw* (nominative syntactically) and as *bny* (oblique; see Stein 2002a).

The inflectional endings which mark each of the states are summarized below:

(2)	Masculine			Feminine		
	<i>Constr.</i>	<i>Indet.</i>	<i>Det.</i>	<i>Constr.</i>	<i>Indet.</i>	<i>Det.</i>
<i>Singular</i>	-ḥ	-m	-n	-t	-t-m	-t-n
<i>Dual</i>	-ḥ/-y	-n	-nhn	-t-y	-t-n	-t-nhn
<i>External plural</i>	-w/-y	-n	-nhn	-t	-t-m	-t-n

Note the following:

1. The feminine endings presented above are the “regular” forms. There is also a set of “natural” feminines which are formed like the masculine, that is, without the ending -t.
2. The masculine singular endings are likewise those of the broken plurals of masculine forms.
3. In Masculine dual constr. the first ending is attested in Early Sabaic, the second in Middle Sabaic and Late Sabaic.
4. In Masculine constr. plural the first ending is nominative, the second oblique.



## 4.3 Pronominal morphology

### 4.3.1 Personal pronouns

Forms of the first and second persons are only very sparsely attested thus far, the latter primarily in the minuscule inscriptions of the wooden sticks (see §2). Typical of a Semitic language, personal pronouns occur in both independent and suffixed (clitic) forms. A gender distinction is not attested in the first person and in the dual number (see Ch. 6, §3.3.3).

The independent personal pronouns serve as subjects of nominal clauses and, more often, of verbal clauses. In the latter case, the pronoun is placed at the beginning of the clause and the following verb is usually separated by *f*- (see §5.1). Consider the following examples: *mr' ʔ Ry* 508/11 (Late Sabaic) “you are lord”; *w-ʔ s<sub>3</sub>hln* A-40-4/3 (beside *w-ʔnt f-s<sub>3</sub>hln*; see §4.4.4); *w-hmw f-hmdw* C 2/7–8 “and they thanked”; *w-t ʔwlw bn-hw b-wfym hw' w-kl s<sub>2</sub>w'-hmw* J 631/13–14 “and they returned from there safely, [namely] he himself and all their retinue”. The personal pronouns of the third person are identical with the second group of demonstrative pronouns (see §4.3.2).

The clitic personal pronouns (pronominal suffixes) appear on both noun and verb forms: for example, *bn-y-hw* “his sons”; *hmr-hw* “he granted to him”; *l-kmw* YM 11729/3 “to you (pl.)”; *l-krbn-kmw* (< *l-ykrbn-kmw*) YM 11733/2 “may he bless you (pl.)”; and so forth.

The independent pronouns and and pronominal suffixes are summarized in (3):

(3)	Independent pronouns	Pronominal suffixes
<i>Singular</i>		
1st com.	ʔn	(-n)
2nd masc.	ʔnt, ʔt	-k
2nd fem.		-k
3rd masc.	hʔ, hwʔ	-hw
3rd fem.	hʔ, [hyʔ]	-h, -hw
<i>Dual</i>		
2nd com.	ʔtmy	-kmy
3rd com.	hmy	-hmy
<i>Plural</i>		
1st com.		(-n)
2nd masc.	ʔntmw	-kmw
2nd fem.		
3rd masc.	hmw	-hmw, (-hm)
3rd fem.	[hn]	-hn

Regarding the pronouns of (3), note the following:

1. The first common singular pronoun ʔn is attested in a few Late Sabaic texts: *ʔtrw ʔn ms<sub>3</sub>ndn ʔn ʔbrh* C 541/3–4 “I, Abraha, wrote this inscription” (the verb as *plural maiestatis*); see also §4.4.2 on VL 24/3 = J 2353/3.
2. First com. sing. -n (accusative) so far attested only in feminine personal names such as *s<sub>2</sub>f<sub>nn</sub>s<sub>1</sub>r* (= *s<sub>2</sub>f-n ns<sub>1</sub>r*) “[the god] Nasr protected me”.
3. The second feminine singular clitic -k occurs in the “Sun Hymn” (Robin 1991:122), and in Oost. Inst. 14/5f.
4. The third feminine clitic of the form -hw is attested only in Middle Sabaic.

### 4.3.2 Demonstrative pronouns

These may be divided into two groups according to their form and function: *Demonstrative 1* indicates the immediate situation of the speaker or reader of an inscription, whereas *Demonstrative 2* points back to something mentioned previously in the text. A demonstrative pronoun precedes the noun it modifies, which appears in the determinate state: for example, *ḏn ṣlmn ḏ-ḏhbn* J 578/4–5 “this (i.e., the present) statuette of bronze”; *h' fṛwtn* R 4815/5,7 “this (i.e., the aforementioned) canal”; *hmw ḥmrn* J 576/10,16 “these (i.e., the aforementioned) Himyarites”. The demonstratives of the second group distinguish special forms for the oblique case: for example, *b-hwt bytn* E 13 §10 “in that castle”.

(4)		<i>Demonstrative 1</i>	<i>Demonstrative 2</i>	
			<i>Nominative</i>	<i>Oblique</i>
	<i>Singular</i>			
	<i>Masculine</i>	<i>ḏn</i>	h', hw'	hwt
	<i>Feminine</i>	<i>ḏt, (ḏtn')</i>	h', hy'	hyt
	<i>Dual</i>			
	<i>Common</i>	<i>ḏyn</i>	hmy	hmt, (hmyt)
	<i>Plural</i>			
	<i>Masculine</i>	<i>ḥn</i>	hmw	hmt
	<i>Feminine</i>	<i>ḥt</i>	hn	hnt

### 4.3.3 Relative pronouns

The relative pronoun, either inflected or in the frozen form *ḏ-*, appears before independent or attributive relative clauses (see §5.4.1 and §5.4.2) or before nouns in a circumlocution for a construct chain, as in *ṣlmn ḏ-ḏhbn* “the statuette of bronze” (see also the example from J 657 in §4.6.1.1), in which the nouns often agree in definiteness (thus *ṣlmn ḏ-ḏhbn* vs. *ṣlmm ḏ-ḏhbm*). The standard forms of the relative pronoun are as follows:

(5)	<i>Masculine</i>	<i>Feminine</i>
<i>Singular</i>	<i>ḏ-</i>	<i>ḏt, t-</i>
<i>Dual</i>	<i>ḏy</i>	<i>ḏty</i>
<i>Plural</i>	<i>ḥ, ḥw, ḥy, ḥt</i>	<i>ḥt</i>

Note the following:

1. In feminine singular, the second form is Late Sabaic.
2. In masculine plural, the first form is Early Sabaic, the second and third ones are Middle Sabaic nominative and oblique respectively, the last one is Late Sabaic.

### 4.3.4 Indefinite pronouns

For the indefinite pronouns *mn* “someone” and *mhn* “something” see §5.4.1 and §5.3.6.

## 4.4 Verbal morphology

### 4.4.1 Verb-stems

The following verb-stems may be distinguished graphically (on the fundamental nature of Semitic verb-stems, see Ch. 6, §3.3.5.2):

(6)  $0_1(f^1l)$ ,  $0_2(f^1l)$ ,  $H(hf^1l)$ ,  $T_{in}(ft^1l)$ ,  $T_{Pr}(tf^1l)$ ,  $ST(s_1tf^1l)$ .

For the stem  $0_2$  see below, §4.4.5.

These are inflected regularly throughout, retaining their formative elements even with prefixes (e.g., *yhf<sup>1</sup>l<sup>nn</sup>*). Possible additional stems, marked by vocalic or consonantal lengthening, may also be assumed, but are not yet clearly ascertained on the basis of the script.

Likewise active and passive forms of the verb cannot be outwardly distinguished. However, a differentiation of voice can be demonstrated for many verbs on syntactic grounds (see the example from J 669 in §5.3.6), a distinction that must have been marked vocally.

#### 4.4.2 Verb inflection

As in other Semitic languages, two conjugation types exist in Ancient South Arabian: the suffix-conjugation, usually termed the *perfect*, and the prefix-conjugation, or *imperfect*. The latter is further divided into a simple, unaugmented “short form” and a form augmented by *-n* called the “long form” or *N-imperfect* (see Nebes 1994b). The base of the prefix-conjugation of the simple stem (for both short and long forms) has the shape  $f^1Vl$  as in Arabic (see Nebes 1994a).

The forms of the two conjugations are presented in (7). The final *-w* or *-y* of dual and plural forms usually disappears (in the orthography) before a following suffix, as in *ln hbrw b-ly ḥḍrn w-hs<sub>1</sub>ḥt-hmw* E 13 §9 “when they set out against the Hadramites and defeated them”. Likewise the prefix *y-* is sometimes not written after a preceding precativ particle *l-*, as in *l-hṣbḥnn* beside *l-yḥṣbḥnn* “may [ʿAṭtar and ʿAlmaqah] keep [you happy]” in the minuscule inscriptions YM 11729/2–3 and YM 11732/2 (cf. also Ghul Document A/1–2); the ending *-nn* clearly marks these forms as finite. Such defective imperfect forms are to be distinguished from the precativ infinitives discussed in §5.6.2.

#### (7) Summary of finite verb forms

		Prefix-conjugation	
		Short form	Long form
<i>Singular</i>			
3rd masc.	$f^1l$	$y-f^1l$	$y-f^1l-n$
3rd fem.	$f^1l-t$	$t-f^1l$	$t-f^1l-n$
2nd masc.	$f^1l-k$	$t-f^1l$	$t-f^1l-n$
2nd fem.	$f^1l-k$		
1st com.	$(f^1l-k)$		
<i>Dual</i>			
3rd masc.	$f^1l, f^1l-y$	$y-f^1l-y$	$y-f^1l-nn$
3rd fem.	$f^1l-ty, (f^1l-tw)$	$[t-f^1l-y]$	$t-f^1l-nn$
<i>Plural</i>			
3rd masc.	$f^1l-w$	$y-f^1l-w$	$y-f^1l-nn$
3rd fem.	$f^1l-y, f^1l-n (?)$	$t-f^1l-n (?)$	$t-f^1l-nn$
2nd masc.	$f^1l-kmw$		$t-f^1l-nn$

Regarding these finite verb forms, note the following:

1. The forms of the second-person masculine are thus far attested almost exclusively in the minuscule inscriptions on wooden sticks (cf. Ryckmans, Müller, and Abdallah 1994 with publication of several texts and further bibliography).
2. The second-person singular  $f^1l-k$  is attested as feminine with certainty only in the “Sun Hymn” (Robin 1991:122).

3. The first-person singular is not certainly attested in texts published to date. Perfect forms probably occur in VL 24 = J 2353: for example, line 3, *w-br 'k-h ḥ* “and I built them (the irrigation works)”.
4. Both of the third feminine plural forms, *f<sup>l</sup>-y* and *f<sup>l</sup>-n*, are attested only rarely; also perhaps *f<sup>l</sup>-tw* (cf. Nebes 1985:34–38).
5. In dual 3rd masc., the first form is Early Sabaic, the second one is later.

#### 4.4.2.1 The suffix-conjugation

The suffix-conjugation, or perfect, (*f<sup>l</sup>*) may appear in both main and subordinate clauses and is primarily used for the past and the pluperfect. It may occur in statements that, according to their meaning, denote *duration* as well as those that describe *punctual* actions. Compare, for example, the following: (i) *w-thrgw b- 'mhmw bn s<sub>2</sub>f s<sub>2</sub>rqm 'dy mqt<sup>t</sup> s<sub>2</sub>ms<sub>1</sub>n w-l<sup>y</sup>l lylm 'dy s<sub>2</sub>r<sup>q</sup> kwkbn d-sb<sup>h</sup>n J 649/32–34* “And they fought with them from daybreak until sunset and (from) the falling (?) of night until the rising of the morning-star”; (ii) *w-wdqy hmy bt<sup>n</sup>hn d-hmdn w-bt<sup>c</sup> bn hwt d<sup>n</sup>mn J 651/20–21* “And both these houses of Hamdān and of Bata<sup>c</sup> collapsed because of this rain”.

The basic time reference for the perfect is anteriority to a given “relative moment”. In the past, as in both of the preceding examples, the relative moment lies in the temporal sphere of the author; in the pluperfect, it lies in the syntactically superordinate clause, as in *w-hmdw mqm 'lmqh k-ḥlhmw bn qblm d-wdq 'l-hmw YM 440/6–8* “And they expressed their gratitude for the power of Almaqah, that he had saved them from a misfortune that had come upon them”.

In the protasis of conditional sentences and in relative clauses with conditional connotation, the suffix-conjugation has a present meaning. This is explained, as in Arabic, by the specific relationship of anteriority of the apodosis and protasis; see the examples in §5.3.6.

#### 4.4.2.2 The prefix-conjugation

Concerning the distribution of the short (*yf<sup>l</sup>*) and long (*yf<sup>l</sup>n*) forms of this imperfect, it should be noted that *yf<sup>l</sup>n* forms constitute over three-quarters of the attestations. A rigorously consistent differentiation of functions between the two types cannot be established. A historical consideration of the uses reveals, however, that in Early Sabaic *yf<sup>l</sup>* is attested considerably more often than *yf<sup>l</sup>n*. To be noted as well is that in the Middle Sabaic period *yf<sup>l</sup>* appears considerably more often than *yf<sup>l</sup>n* in narrative contexts, where both long and short types are used to describe a “progression of action” as, for example, in the following: *w-bnhw f-ygb<sup>w</sup> 'dy hgrn n' d w-bnhw f-yh<sup>s</sup>rn mlkn 'ls<sup>r</sup>h yḥd<sup>b</sup> w-d-bn ḥms<sup>1</sup>hw w-'f<sup>s</sup>hw 'dy 'r<sup>d</sup> mh<sup>n</sup>f<sup>m</sup> w-yqm<sup>w</sup> w-hb<sup>l</sup>n hgrnhn J 576/7–8* “From there they went to the city of Na<sup>ḏ</sup>. From there King 'Išaraḥ Yaḥdib, along with part of his main army and his riders, set out for the region of Muha<sup>n</sup>ifum. [Then] they destroyed and seized the both cities”.

As a common denominator to which the overwhelming majority of examples may be reduced, the terms that suggest themselves for the relative time reference of the imperfect are *simultaneity* and *posteriority*. The “relative moment” is either the present moment of the writer/speaker or to be found in the immediate syntactic context (e.g., a superordinate clause).

The long form of the imperfect (*yf<sup>l</sup>n*), seldom the short form (*yf<sup>l</sup>*), occurs in statements with the present and future reference: for example, *mḏb<sup>h</sup>t b-h ydb<sup>h</sup>n mlkn twrm b-ywm ts<sup>1</sup> 'm d-twr C 671/1–4 = R 3104/1–4* “altar on which the king on the 9th day [of the month of] dū Tawr offers a bull”.

In subordinate clauses introduced by conjunctions and in relative clauses which have a syntactically superordinate clause situated in the past, *yf<sup>l</sup>n* and, less often, *yf<sup>l</sup>* may have

modal nuances: for example, *bkn wqhww... l-s<sub>1</sub>b' w-qtdmn... w-l-s<sub>2</sub>ym l-hw mdrfn s<sub>1</sub>wn tmlhyn d-yhmynhw bn d'bn* J 651/28–33 “When [their lord] ... commanded him to carry out and direct [the work] ... [and] to erect the dam-works for it [i.e., the city of Mārib] further up the wadi, which would protect it from flooding” (potentiality or intentionality).

The prefix-conjugation also forms the basis for the production of other modal verb forms: (i) *l-yf'ln* denotes the *precative* (used to express wishes); (ii) *l-yf'l* serves as the *jussive* (expressing indirect commands; for the imperative, see §4.4.4); and (iii) *'l yf'l* functions as the *vetitive* (used to express negative wishes). These can be respectively illustrated by the following: (i) *w-l-yhmrnhw lmqhw hzy w-rdw mr'hmw* J 667/14–15 “And may Almaqahū grant him the goodwill and the pleasure of their lord”; (ii) *w-l ydbhw bn ms<sub>2</sub>mnhn 'itr w-s<sub>2</sub>ms<sub>1</sub>m w-dbh<sub>1</sub>m b-hrn* C 74/11–13 “And let them offer [an animal sacrifice] to 'Aṭtar and to Šamsūm and an animal sacrifice [to Almaqah] in Hirrān from [the yield of] both cultivated areas”; and (iii) *w-'l yhwfd b-h 'md w-'lbm* C 610/3 (Early Sabaic) “And neither vines nor 'ilb-trees may be planted there”.

#### 4.4.3 Weak verbs

Both tri- and biradical spellings of verbs II-*w/y* occur, the latter being the more common, as in *ykwnn* and *ykn* “he will be”; *hqwḥ* and *hqḥ* “he completed” (on triradical roots, see §4.1). Since no semantic distinction is generally discernible between the bi- and triradical forms of the verbs in question, it may be assumed in principle that these are purely graphic variants of one and the same verb form, and not forms of different verbal stems (such as *kwn* as a stem with doubling of the second radical beside *kn* as the simple stem). This does not preclude in any way the existence of derived verbal stems, but the identification of the latter can only be made on the basis of comparative contexts (many verbs are attested only in one or the other written form).

Verbs II-*geminatae* (i.e., with the second consonant doubled) are written either tri- or biradically, as in *hbrr* “to come forth”, versus *hg* “to make a pilgrimage”; alternative spellings of individual roots are only rarely attested thus far (see the entries under *ḍrr* and *kl* in *Sab. Dict.*), which suggests the existence of different verbal stems.

Verbs III-*w/y* exhibit sound forms for the most part, as in *hwfy-hw* “he granted to him”; *yhrḍwn* “he will satisfy”; shortened forms are rare: compare *w-hrḍ-hw* C 365/5 “and [because] he satisfied him”; *l-yṭ* (< *l-yṭy*) R 4176/10,11 (Early Sabaic) “may it come”. Note-worthy is the alternation of *w* and *y* in a few roots; with *yhrḍwn* above compare *yhrdyn* and, in general, §3.4.1.

Verbs I-*w* lose their first radical in the prefix-conjugation (thus *l-yz'n* “may he continue [... to do]”, etc.). The few instances of verbs I-*y* exhibit both spellings (*yyf'n* or *yf'n* “it will be proclaimed”). The *hf'l* and *s<sub>1</sub>tf'l* stems are as a rule regularly formed (e.g., *hwfy* “he fulfilled”, rarely *hfy*; *s<sub>1</sub>twfy* “he was protected”). In the *ft'l* stem the first radical is missing in the orthography, as in *tqhw* “they complied/completed” (cf. in contrast the sound form of the *tf'l* stem, as in *tws<sub>3</sub>'w* “they attacked”). Since verbs I-*n* exhibit a similar appearance (e.g., *tḏ'n* “to harm”, infinitive), an assimilation of the first radical to the infix, as in Arabic, suggests itself.

As a rule, the first consonant of verbs I-*n* is assimilated to a following consonant, as in *hkl* beside *hnkl* “he carried out”. (see also §3.4.2).

#### 4.4.4 Imperative

The imperative is attested in the minuscule script of the wooden sticks. It has the form *f'l(-n)*: for example, *w-ḥt f-s<sub>3</sub>hln* YM 11742/2 “and you (sg.), look after ...!”

#### 4.4.5 Infinitive and participle

From the Middle Sabaic period onwards, the infinitive is morphologically divided into two forms: a non-augmented form (*fʿl*) of the basic stem 0<sub>1</sub>, and a form augmented by *-n* for all derived stems (e.g., *hfʿln*, *tfʿln*, etc.). This rule, restricted to the region of Mārib and the central Yemeni highlands, also allows a morphological distinction between the basic stem 0<sub>1</sub> (infinitive *fʿl*) and a derived stem 0<sub>2</sub> (infinitive *fʿln*). See Stein 2002b. This *-n* is to be distinguished from that of the determinate state (see §4.2.2), which, like other formative elements (mimation, etc.), may not appear on the infinitive.

Characteristic of the usage of the individual Ancient South Arabian dialects, and in particular of Sabaic, are the various functions of the infinitive. Two basic uses underlie these: (i) the infinitive stands as the predicate, in which position it is interchangeable with a finite verb form; and (ii) the infinitive assumes the role of a part of the clause dependent on the predicate (in this position it is no longer interchangeable with a finite verb form; see Nebes 1988). On infinitival syntax, see §5.6.

The (active) participle of the basic stem has the form *fʿl*, as in *wḏm ʾw bhʾm* C 548/2 “going out or in”; *bn nkrm w-mhbʾm* C 29/5 “[they placed their house under the protection of ‘Attar], against anyone who would alter or harm it”. The inscription Šilwī-aš-Šuzayf 1, written in the Haramic dialect, exhibits a participle without mimation (lines 3–5): *b-hn gwz bʾtʾtn w-hwʾ ʾbr* “because he passed through [the region of?] Baḥatān, crossing [a border]”. A passive participle of the basic stem of the form *mfʿl* is difficult to confirm. The participles of the derived stems have a prefix *m-*, as in *mhbʾm* “who harms” (cf. the example from C 29 above); active and passive forms cannot be distinguished outwardly, and in general it is difficult to distinguish between participles and other nominal forms.

### 4.5 Particles

In addition to the conjunctions (see §§5.2–5.3) and object clause marker (see §5.3.1), the particles of Ancient South Arabian include prepositions, particles of negation, and enclitics.

#### 4.5.1 Prepositions

The most important prepositions, with their primary meanings, are as follows:

1. *b-* “in, at” (local); “in, on, during” (temporal); “with, by” (instrumental/sociative).
2. *l-* “to(ward)” (local and temporal); expression of the dative. Sometimes there is overlap with the semantic range of *b-*, as in *l-tʾtm ywmm* J 631/28 “on the third day”; versus *b-tʾtm ywmm* J 577/12.
3. *bn* “(away) from” (local and temporal); also partitive and explanatory (e.g., *kl s2ʾmt... bn ʾns1m w-bʾlm w-tʾwm w-bʾrm* R 3910/2–3 “every purchase of person, camel, bull, or [other] cattle”). In texts in the Haramic dialect *bn* is replaced by *mn*, otherwise unknown in Ancient South Arabian.
4. *ʾbr* “in the direction of”. This preposition has a wide range of meaning and often occurs with other prepositions (e.g., *b-ḏr hs2tʾw b-ʾbr mrʾ-hmw* E 13 §2 “In the war that [that people] had fomented against their lord”).
5. *ʾd(y)* “up to” (local); “until” (temporal). In addition to expressing direction, this preposition also expresses the goal of an action and sometimes also stands simply

for local “in” (e.g., *ʔmr w-ʔql ʂdqm ʿdy kl ʕ<sub>1</sub>rr-hmw* E 18/21 “[May the deity grant] proper crops in all their valleys”).

6. *ʔ(y)* “(up)on”, frequently combined with *b-*. This preposition often serves to express enmity (e.g., *s<sub>1</sub>bʔ s<sub>1</sub>bʔ w-dbʾ b-ʔy hbs<sub>2</sub>t* E 19/6–7 “the campaign that they undertook and carried out against Abessynia”).

Other common prepositions are *(b-)'m* “(together) with”; *bʿd(n)* “after” (local and temporal); *b(y)n* “between”; *hg(n)*, *b-hg* “like, corresponding to”; *(l-/b-)qbl* “before” (local and temporal).

#### 4.5.2 Negative particles

The negative in all applications is *ʔ*. In addition to verbal clauses it also appears in nominal clauses, particularly in the negation of existence, often with jussive force, as in *ʔ ʕ<sub>1</sub>s<sub>1</sub>ʔ* G/1379/3,7 “let there be no one who lays claim”. In Late Sabaic the negation is *dʔ*, in the Haramic dialect, however, it is *lm* (followed by the short form of the prefix-conjugation like in Arabic).

#### 4.5.3 Enclitics

Occasionally the particle *-m* or *-mw* (less often *-my*) is appended to an individual word, particularly on prepositions and on certain (for the most part lexically determined) conjunctions, yet hardly ever on verbs (see Nebes 1991). The function of the particle is probably emphasis; compare, for example, *w-b-mw hwt hrfn* E 69/20 “and in the very same year” and *b-hwt hrfn* J 751/8–9 “in that year”. A second enclitic, *-n*, is mainly attached to a preposition or conjunction and causes an inversion of the original meaning, e.g., *ln* “from” opposite *l-* “to(wards)”, *ʕmn* “from” opposite *(b)'m* “(together) with”.

### 4.6 Numerals

In the written record preserved for Ancient South Arabian, numbers are usually written out; only rarely are they expressed with special numeric characters.

#### 4.6.1 Cardinals

The numbers from 1 to 9 each have a masculine and a feminine form, the latter augmented with *-t*:

(8)	Masculine	Feminine
1	ʔd	ʔt
2	ʔny	ʔty
3	s <sub>2</sub> ʔt, ʔʔt	s <sub>2</sub> ʔtt, ʔʔtt
4	ʔrbʿ	ʔrbʿt
5	hms <sub>1</sub>	hms <sub>1</sub> t
6	s <sub>1</sub> dʔ, s <sub>1</sub> ʔ	s <sub>1</sub> dʔt, s <sub>1</sub> ʔt
7	s <sub>1</sub> bʿ	s <sub>1</sub> bʿt
8	ʔmny, ʔmn	ʔmnyt, ʔmnt
9	ts <sub>1</sub> ʿ	ts <sub>1</sub> ʿt
10	ʕs <sub>2</sub> r	ʕs <sub>2</sub> rt

Regarding the above cardinal numbers, note the following:



1. In addition to the common forms of 2 there also occurs *kl̥y* (Early Sabaic; Middle Sabaic *kly*), feminine *kl̥ty*, for “both”, which is always definite: for example, *kl̥y m̥hfdnhn yzl w-drʿ J 557* (Early Sabaic) “both of the towers Yaʿzil and Dara”.
2. The first cited form of the numbers 3, 6, and 8 is attested in Early Sabaic, the second in the Middle Sabaic and Late Sabaic periods.
3. A number *ʿhdy* “1” and *s<sub>1</sub>t* “6” are attested in some late inscriptions from south-eastern regions, e.g., in ‘Abadān 1/23 (cf. also *s<sub>1</sub>ty* “60” in R 5085/11 [Late Sabaic]).

The numbers from 11 to 19 are composed of the relevant unit (masculine or feminine) and *ʿs<sub>2</sub>r* (unchangeable). The numbers 11 and 12 are thus far only attested in their masculine forms (*ʿhd ʿs<sub>2</sub>r* and *ṭny ʿs<sub>2</sub>r*), the other numbers, conversely, almost exclusively in the feminine form (e.g., *ṭlt ʿs<sub>2</sub>r* “13”; *s<sub>1</sub>dṭt ʿs<sub>2</sub>r* [Early Sabaic] “16”).

The number 20 has the form *ʿs<sub>2</sub>ry*, while the other tens up to 90 have the form of the unit with a suffixed *-y*: for example, *s<sub>2</sub>lty* or *ṭlty* “30”.

The number 100 in the singular is *mṭ*, in the dual (“200”) *mṭn*, in the plural *mʿ* (Early Sabaic), *mṇ*, or *mṭ* (e.g., *s<sub>1</sub>bʿ mṭ* “700”). The word for 1,000, *ʿlf*, has a broken plural, *ʿlf*.

In compound numbers the elements (units, tens, etc.) go from smallest to largest, connected by *w-*.

#### 4.6.1.1 Construction of the cardinals

The gender polarity of the numbers 3 to 10 that is common throughout the Semitic languages is also found in Ancient South Arabian. That is, a counted masculine noun takes the feminine form of the relevant number, a feminine noun the masculine form of the number.

The numeral appears before the thing counted, and agrees with it in definiteness. With an indefinite noun the number appears in the absolute state (see §4.2.2; exceptions are *mṭ* and *ʿlf*, which regularly exhibit mimation); with definite nouns the number is likewise definite. The thing counted is usually in the dual with 2, in the plural with higher numbers. A few examples will illustrate the construction: *ḥms<sub>1</sub> w-ʿs<sub>2</sub>ry w-mṭ ʿfrs<sub>1</sub>m J 665/30–31* “125 riders”; *ḥms<sub>1</sub> mʿnm w-ʿlfm ʿs<sub>1</sub>dm J 576/15* “1,500 soldiers”; *ṭltt ʿs<sub>2</sub>r ywmtm E 13 §10* “for 13 days”; *ṭlttn ʿlmm ʿly dḥbn J 657/3* “the 3 statuettes of bronze” (definite).

#### 4.6.2 Ordinals

A special form is *qdm* “first”. The other ordinal numbers up to 10 differ outwardly from the cardinals only in that the masculine forms always have three consonants, thus *rbʿ* “fourth”; *s<sub>1</sub>dṭ* “sixth”. Feminine forms have the ending *-t*, as in *ṭnyt* (also *ṭnt*) “second”.

Attributive ordinals are placed after the thing counted and agree with it in state and gender, as in *drm ṭntm Ja 576/11* “a second time”; *ḥrf wddl... rbʿn J 618/9–10* “the fourth year of the [eponym] Wadad’il” (definite). Several temporal expressions are constructed differently, such as *b-ywm ts<sub>1</sub>m* “on the ninth day” (for reference see §4.4.2.2; construct state); *b-ṭltm ywmm J 577/12* “on the third day”.

#### 4.6.3 Other numerals

Fractions have the same outward form as the ordinals: thus *s<sub>2</sub>lt* (Early Sabaic) “a third”, *rbʿ* “a fourth”, and so on (e.g., *kl ṭmn qbrn ygr DAI FH Awām 1997-2/2* (Early Sabaic) “the entire eighth of the grave Yagur”). In compounds the fraction looks like the singular, as in *s<sub>2</sub>lt rbʿ kl qbrn ygr DAI FH Awām 1997-5/2* (Early Sabaic) “three-fourths of the entire



grave Yagur". For "half" there is a specific word, *fqlh*: for example, *w-kwn fqlhm l-šbhm w-ḥmym w-fqlhm l-b'ttr* DAI FH Awām 1997-5/4 (Early Sabaic) "and one half [of the grave] belongs to Šubḥum and Ḥamyum and one half to Bi'aṭtar". In addition, fractions can also be expressed periphrastically, as in *šb'm bn tmny šb' C 640/2* "a finger out of eight fingers" (= "one-eighth").

Multiplicatives are only rarely attested; they are formed by the addition of *ʾd* ("time(s)") to the numeral, as in *s<sub>2</sub>lttʾd C 366* (Early Sabaic) "three times" or "for the third time"; *s<sub>1</sub>dtʾd Schm/Mārib 19/A4* (Early Sabaic; fragmentary context).

## 4.7 Non-Sabaic morphological features

In contrast to Sabaic as an "H-language", all other Ancient South Arabian dialects are so-called "S-languages"; in other words, they form the causative stem and the pronouns with *s<sub>1</sub>*, thus *s<sub>1</sub>fʿl* (Sab. *hfʿl*), *-s<sub>1</sub>w* (Sab. *-hw*), and so forth.

### 4.7.1 Nominal morphology

Outside of Sabaic, external plurals are met with more commonly, especially in Minaic. In particular, an *h* often appears word-finally in construct state forms – in Minaic even in the singular (see §3.5).

An *h* can also be inserted in the external plural ending of the feminine, as in Minaic *ʾnthtn R 3306A/3* = as-Sawdā' 37/3 "the women"; *w-ʾdhty M 275/3* "and the lands of ...".

### (9) Summary of non-Sabaic nominal endings

		<i>Construct</i>	<i>Indeterminate</i>	<i>Determinate</i>
<i>Singular/Broken plural</i>	Minaic	-h, - <i>φ</i>	(-m)	-n
	Qatabanic	- <i>φ</i>	-m	-n
	Hadramitic	- <i>φ</i>	-m	-hn, -n
<i>Dual</i>	Minaic	-y, -hy	-ny	-nhn, -nyhn
	Qatabanic	-y, -w, -h(y)	-myw	-nyhn
	Hadramitic	-y, -hy	-nyw	-yhn, -yn
<i>External plural</i>	Minaic	-hw, -hy	-hn	
	Qatabanic	-w, -y, (-h)		
	Hadramitic	(-hy)		(-yhn)

Many of the forms given are attested only rarely on account of the limited extent of what has been preserved. Note that the interpretation of the endings *-hy* and *-yhn* as plural is not completely certain; for discussion of the attestations see Beeston 1984: §H 13:2, 3.

### 4.7.2 Pronominal morphology

#### 4.7.2.1 Personal pronouns

With a few exceptions in Qatabanic inscriptions (e.g., *bd-k J 367* "your (sg.) servant"), only third-person forms are attested. The distinctive long forms of the third-person masculine singular pronominal suffix in Qatabanic and Hadramitic (*-s<sub>1</sub>ww*) are attached to external plurals and duals of nouns, but not to verbal forms (so Beeston 1984: §Q 23:2, H 23:2). The forms of the suffixed personal pronouns are as follows:

(10)	<u>Minaic</u>	<u>Qatabanic</u>	<u>Hadramitic</u>
<i>Singular</i>			
3rd masc.	-s <sub>1</sub> , -s <sub>1</sub> w	-s <sub>1</sub> , -s <sub>1</sub> ww	-s <sub>1</sub> , -s <sub>1</sub> ww
3rd fem.	-s <sub>1</sub>	-s <sub>1</sub> , -s <sub>1</sub> yw	-ṭ, -tyw, -s <sub>3</sub> , -s <sub>3</sub> yw
<i>Dual</i>			
3rd com.	-s <sub>1</sub> mn	-s <sub>1</sub> my	-s <sub>1</sub> my
3rd masc.			-s <sub>1</sub> mn, -s <sub>1</sub> myn
<i>Plural</i>			
3rd masc.	-s <sub>1</sub> m	-s <sub>1</sub> m	-s <sub>1</sub> m
3rd fem.	-s <sub>1</sub> n	-s <sub>1</sub> n	

#### 4.7.2.2 Demonstrative pronouns

Whereas the forms in Qatabanic for the most part correspond to those of Sabaic (thus *dn*, *dt*, *s<sub>1</sub>mt*, *-s<sub>1</sub>myt*, etc.; exceptions are the masculine plural of the first group, *dt<sub>n</sub>*, and the nominative masculine of the second group, sg. *s<sub>1</sub>w*, pl. *s<sub>1</sub>m*), in Minaic the demonstrative pronouns of the second group are essentially not attested at all, and those of the first group only very rarely (one of the few plural forms is *hlt mhf<sub>d</sub>tn* R 3015/2 = M 239/2 “these towers”; cf. R 2965/2 = M 185/2).

#### 4.7.2.3 Relative pronouns

Qatabanic exhibits *d-* as a frozen relative particle as well as the form *dn*, as in *s<sub>2</sub>n’ dn qnyw w-bqnyn* Ry 367/9 = NAM 483/9 “[may *dū* Samāwī take revenge . . .] on every enemy of that which they have acquired and will acquire”. In the following summary, uncertain and markedly rare forms are not listed:

(11)	<u>Minaic</u>		<u>Qatabanic</u>		<u>Hadramitic</u>	
	<u>Masc.</u>	<u>Fem.</u>	<u>Masc.</u>	<u>Fem.</u>	<u>Masc.</u>	<u>Fem.</u>
<i>Singular</i>	<i>d-</i>	<i>dt</i>	<i>d-</i> , <i>dw</i>	<i>dt</i>	<i>d-</i>	<i>dt</i>
<i>Dual</i>	<i>dy</i>	<i>dtyn</i>	<i>dw</i>			
<i>Plural</i>	<i>hl</i> , <i>hl</i>		<i>dtw</i>			

#### 4.7.2.4 Indefinite pronouns

Qatabanic exhibits in addition to *mn* also *ʔy* “who(ever)”.

### 4.7.3 Verbal morphology

#### 4.7.3.1 Verb-stems

In Minaic a few verbs exhibit a spelling *f<sup>c</sup>l* (such as *lly* “raise,” e.g., M 203/2). Since consonantal length in Ancient South Arabian is not expressed in the script, such forms are probably to be understood as another verbal stem with reduplicated second radical, to be distinguished from a possible stem *f<sup>c</sup>l* with doubling.

#### 4.7.3.2 Suffix-conjugation

The dual and plural ending is not usually written in Minaic; the forms are thus identical in appearance with the singular (e.g., *s<sub>3</sub>l’* both “he dedicated” and “they dedicated”). The plural of the third-person feminine is attested in Qatabanic and Hadramitic as *f<sup>c</sup>ln* (see Robin 1983:181–184; Nebes 1985:34).

#### 4.7.3.3 Prefix-conjugation

The indicative forms in Qatabanic, in contrast to the other dialects, are formed with a prefix *b-*, as in *kl mngw byktrbwn* AM 757/11 “all things that they will request” (vs. jussive *w-l-yqny* R 3688/4 “and may he acquire”). Here too the prefix *y* can be lost in writing; see the example in §4.7.2.3. The form of the third-person masculine plural of the prefix-conjugation in Qatabanic is *y-fʿl-wn* (see the example above).

Qatabanic *b-yfʿl* for the most part corresponds to Sabaic *yfʿln*, in being used for indicative statements of the present and the future: *w-kl s<sub>1</sub>hmm w-qnym bykn w-yks<sub>3</sub>ʿ ws<sub>1</sub>t dtn bytn* Folkard 1/5–6 “And all of the servants and flocks that are present and live in these houses”; *w-l yfth dn fthn w-mhrt n b-ʿdm ʾw ʾbnm kn-m byhrg mlkn* R 3566/21 “And this decree and decision is to be published on wood or stone, as the king will command”. The use of the prefix-conjugation for the past in a narrative context is only very rarely attested, with a few certain examples thus far only in Minaic: *w-yfqr zydʿl b-wrhh hthr w-yfnnw kb bn kl bytth ʾlt mʿr* R 3427/2 = M 338/2 “And Zaydʿil died in the month of Hathor, and they sent linen from all the temples of the gods of Egypt”.

Occasionally, prefixed verbal forms augmented with *b-* are also found in Minaic. Imperfects of the form *yfʿln* (“long form”) are rarely attested in Minaic and are often of uncertain number; the other dialects exhibit no such forms at all.

#### 4.7.3.4 Infinitive

The infinitive is regularly formed without an *n*-augment; in Qatabanic, however, mimation may appear in certain cases (cf. Nebes 1988:70f., 73, and §5.8.3 below).

### 4.7.4 Non-Sabaic particles

The prepositions exhibit a number of distinctive features in comparison with those in Sabaic. Thus, for Sabaic *l-* Minaic usually has the preposition *k-*, Hadramitic *h-* (for further specifics on Hadramitic see Beeston 1984: §H 33:3). The forms that end in *-y* in Sabaic end in *-w* in Qatabanic, thus *ʿdw*, *ʿlw*, and so forth.

In Minaic the particle *k-* in its various functions has a preposed *s<sub>2</sub>*, as in *bn s<sub>2</sub>-kd* R 2980/13 = Shaqab 19/13 “from (the possibility) that”. In contrast to Sabaic the other dialects have a temporal conjunction *mty* (Hadramitic *mt*) “when”.

The negative in Minaic (only sparsely attested) is the particle *lhm*.

Enclitic *-m(w)* is common to all dialects; in addition, Minaic and Qatabanic also exhibit a particle *-y* (Minaic also *-m-y*), while Hadramitic has *-hy* (see Nebes 1991). In Hadramitic the particle *-m* also occurs on verb forms, as in *b-ʿbr dt ynsf-m* Rb I/84 no. 196/2–3 “because he will perform a ritual(?)”.

### 4.7.5 Numerals

Different forms for the number 1 are found in Qatabanic (*ʿtd*, fem. *ʿtt*; also *ʿs<sub>1</sub>tn*) and Minaic (*ʿs<sub>1</sub>t*). The words for 3 and 6 correspond in these dialects to the Early Sabaic forms (thus *s<sub>2</sub>lt(t)*, *s<sub>1</sub>dt(t)*). Hadramitic exhibits both *s<sub>2</sub>lt(t)* and the spelling *s<sub>2</sub>ls<sub>3</sub>(t)*.

The tens in Minaic and Hadramitic may exhibit an *h* in the ending: for example, *ʿrbʿhy* (Minaic, also *ʿrbʿy*) “40”; *ʿtmnhy* “80”.

Distributives are expressed in Qatabanic by repetition of the numeral: *b-ʿs<sub>2</sub>r ʿs<sub>2</sub>r hbs<sub>2</sub>tm mʿm l-ʿtt ʿtt ywmm* R 3854/6–7 “ten full Ḥabṣat-coins each for each day”.

## 5. SYNTAX

### 5.1 Word order

The first clause of an inscription begins with the subject, less commonly (though often in legal documents) with an adverb such as *kn*, *hgn*, among others, “thus”. In all other main clauses, which are usually introduced by the conjunction *w-* (see §5.2), as well as in subordinate clauses introduced by a conjunction, the verbal predicate normally precedes (VS).

In main clauses introduced by *w-*, the subject, object, or a locative or temporal prepositional phrase may appear at the beginning:

- (12) A. *w-ʿws<sub>1</sub>ʾl*      *f-ḥmd*      *mqm*      *ʾlmqh*  
 and==Awsʾil and==he-thanked power-of Almaqah  
 “And Awsʾil expressed his gratitude for the power of Almaqah” (J 644/7)
- B. *w-bythmw*      *nʾmn*      *f-ḏbw*  
 and==house=their Nuʾmān and==they-repaired  
 “And their house Nuʾmān they repaired” (C 648/4)
- C. *w-bn*      *hgrn*      *nʾḏ*      *f-ytʾwlw*  
 and==from city=DET. Naʾḏ and==they-returned  
 “And from the city of Naʾḏ they returned” (J 576/10)

The predicate, as the examples show, is introduced by *f-*, although there are also many cases without *f-* (e.g., *w-ʾlmqh l-yḥmrnhw* J 692/4–5 “and may Almaqah grant to him”). It is rare, when a nominal element is preposed, that the predicate is introduced by *w-* (e.g., *w-frs<sub>1</sub>hw ndf w-zḥn* J 649/20–21 “and his horse Nadif was wounded”; Nebes 1995:22–45; 218–219; 221–231).

The preposing of nominal elements is less common in verbal subordinate clauses, except for resumptive pronouns in relative clauses (see §5.4.2):

- (13) *k-hʾ*      *mtʾhw*  
 that==he he-saved=him  
 “That he [i.e., Almaqah] had saved him” (J 619/10–11)

Resumptive constructions, in which a preposed nominal or prepositional element is resumed by a pronoun elsewhere in the sentence, are uncommon:

- (14) *w-ḏmhw*      *frʾm...*      *w-ʿs<sub>1</sub>d*      *b-ʿmhw*  
 and==servants=his Fārīʾum... and==soldiers-of in==with=him  
*wkb*      *b-wfym*      *blthmw*  
 it-found in==success=INDET. mission=their  
 “And as for his servants, [namely] Fārīʾum... and the soldiers with him, their mission had a successful conclusion” (E 13 §11)

The predicate of a nominal clause may consist of a noun or a prepositional phrase; nominal clauses may be main or subordinate clauses. The subject normally stands first, as in (15A); when the predicate consists of a prepositional phrase, it often stands before an indefinite subject, as in (15B):

- (15) A. *w-ḏn-m*      *wtn*      *mṣdqm*  
 and==this==ENCL. document-of-transfer=DET. binding=INDET.  
 “And this document of transfer is binding” (Gl 1572/7)

- B. *w-ʾl*      *l-hmw*      *b-hw*      *kl*      *mwm*  
 and==not to==them in==it any water=INDET.  
 “while they had no water in it [i.e., the castle Šaqīr]” (E 13 §10)

## 5.2 Coordination

The coordinating conjunction is the particle *w-* “and”; in addition, there is a disjunctive particle (*f-*)*w* “or”. Main clauses and syntagms of equal syntactic status are connected by *w-*. The use of *f-* between clauses of equal rank is rare in Sabaic; it is found primarily in inscriptions in the Haramic dialect:

- (16) *f-ḥtʾt*      *w-tḥlʾn*      *f-ḥḏrʾt*      *w-ʿnw*  
 and==she expiated and==will-pay-fine and==she-submitted and==  
 be-humbled.INF.  
 “Then she expiated and will pay a fine. Then she submitted and humbled herself” (C 568/5–7)

## 5.3 Subordination

A subordinate clause introduced by a conjunction follows its main clause. Exceptions are conditional sentences and complex sentences with a conditional connotation. In the latter sentence types, as well as in other occasional instances of preposed hypotactic clauses introduced with a conjunction, the following main clause is often introduced by *f-*, though also with *w-* or *ϕ* (Nebes 1995:46–53; 219–221; 231–234).

### 5.3.1 Object clauses

Object clauses are introduced by the particle *k-*. Depending on the temporal relationship, they may contain the conjugational form *fʾl* (perfect tense) for anteriority and *yfʾln* (imperfect tense; see §4.4.2) for posteriority:

- (17) A. *w-ys<sub>1</sub>mʿw*      *k-nblw*      *hmw ʾgrn*      *b-ʿbr*      *ʾhzb*  
 and==they-heard that==sent those Nagranites in==direction-of bands-of  
 ḥbs<sub>2</sub>t Abyssynians  
 “And they heard that the aforesaid Nagranites had sent [a delegation] to the Abyssynian bands” (J 577/10)
- B. *w-tbs<sub>2</sub>rw*      *b-ʿm*      *ʾlmqh*  
 and==they-received-good-news in==with Almaqah  
*k-yḥmrnhmw*      *s<sub>1</sub>qym*      *mlym*  
 that==he-would-grant=them irrigation=INDET. winter(?)=INDET.  
 “And they received from Almaqah the good news that he would grant them irrigation in the winter(?)” (J 653/7–8)

### 5.3.2 Temporal clauses

For the temporal notion “when,” the conjunctions *ywm* (properly: “on the day when”; Early Sabaic/Middle Sabaic), *bkn* (Middle Sabaic), and *k-* (Late Sabaic) are used, followed invariably by *fʾl* (perfect) as predicate:

- (18) A. *yḏʾl*      *ḏrḥ*      *bn*      *s<sub>1</sub>mhʾly*      *mkrb*      *s<sub>1</sub>bʾ*      *gnʾ*      *ʾwm*  
 Yadaʿil Ḍarīḥ son-of Sumuhūʿalī mukarrib-of Sabaʾ he-walled Awām

byṭ      ʾlmqh      ywm      ḏbh      ʿttr  
 temple-of Almaqah day-of he-sacrificed ʿAttar  
 “Yadaʿil Dariḥ, son of Sumuhūʿalī, *mukarrib* of Sabaʿ, surrounded Awām,  
 the temple of Almaqah, with a wall [on the day] when he offered an animal  
 sacrifice to ʿAttar” (C 957; Early Sabaic)

B. b-ḏt      hws<sub>2</sub>ʿhmw      ʾlmqh      b-wḏʿ      s<sub>2</sub>ʿbn  
 in==REL. he-granted=them Almaqah in==subjugate.INF. tribe=DET.  
 ngrn      bkn      qs<sub>1</sub>dw      w-nzʿ      ydm  
 Nagrān when they-rose-up and==withdraw.INF. hand=INDET.  
 bn      ʾmrʿhmw      ʾmlk      s<sub>1</sub>bʿ      b-ʿbr      ʾḥbs<sub>2</sub>n  
 from lords=their kings-of Sabaʿ in==direction-of Abessynians=DET.

“Considering that Almaqah granted them [i.e., both kings of the Sabeans] the  
 subjugation of the tribe of Nagrān, when they [i.e., the Nagrānites] rose up  
 and withdrew from their lords, the kings of Sabaʿ, their support against the  
 Abessynians” (J 577/8)

C. w-ts<sub>1</sub>ṭrw      ḏn      ms<sub>1</sub>ndn      qyln      s<sub>2</sub>rḥʾl      yqbl ḏ-yzʿn  
 and==he(!)-put-up this inscription=DET. *qayl*=DET. Šarahʾil Yaqbal REL.==  
 Yazʿan  
 k-qrn      b-ʿly      ngrn  
 when==he-took-up-position in==against Nagrān  
 “The *qayl* Šarahʾil Yaqbal of the clan Yazʿan put up this inscription when he had  
 taken up a position against Nagrān” (J 1028/6; Late Sabaic)

Other temporal relationships are expressed by the conjunctions *bʿd(n)* ḏ- (and the like)  
 “after”; *ln*, *ln* ḏ- “from the time that, since”; *d(y)* ḏ-/ḏt, *tḥw* “until”; *brtn* “when”; and ḏ  
 “when” (Haramic only):

(19) A. f-yṣnʿw      b-hwt      bytn      s<sub>2</sub>qr      ḥms<sub>1</sub>t  
 and==they-took-up-a-defensive-position in==that castle=DET. Šaqīr five  
 ʿs<sub>2</sub>r ymtm... ʿdy ḏt nfṣ      mrʿhmw      s<sub>2</sub>ʿrm      ʾwtr  
 ten days=INDET. until REL. he-arrived lord=their Šāʿirum Awtar  
 w-mṣrhw      bʿd ḏt      s<sub>1</sub>bṭw      mṣr      ḥḏrmwt  
 and==troops=his after REL. they-defeated troops-of Hadramawt  
 b-ḥlf      ḏt      ḡylm  
 in==district-of ḏāt Ḡaylim

“They took up a defensive position in the aforementioned castle Šaqīr for 15  
 days... until finally their lord Šāʿirum Awtar and his troops arrived, after  
 they had defeated the troops of Hadramawt in the district of ḏāt Ḡaylim”  
 (E 13 §10)

B. w-l-h[ʿ]nnhw      bn      ḥlḏ      ḥlḏ      ln  
 and==for==save.INF.=him from illness-of he-suffered-illness since  
 ḏ-ʿtw      bn      mqmn      ḏ-lḡm  
 REL.==he-retained from observation-post=DET. REL.==Laḡum  
 “And so that he [i.e., Almaqah] would save him from the illness from which  
 he suffered since the time that he had returned from the observation post of  
 Laḡum” (J 633/4–6)

C. b-ḏt      ḥmrhw      ṣdqhw      b-ms<sub>1</sub>ʾlhw      brtn      blthw  
 in==REL. he-granted=him right=his in==oracle=his when he-sent=him

mr'hw s<sub>2</sub>mr yhr's<sub>2</sub>  
 lord=his Šammar Yuhar'iš

“Considering that he [i.e., Almaqah] granted him [i.e., the author] what was fitting, in his oracle, when his lord Šammar Yuhar'iš despatched him”

(BR M.Bayhān 5/3–4)

- D. ḡs<sub>2</sub>nm bn gnyt ḡlwnyn 'd bny  
 Ġašnum son-of Gāniyat Ġulwānite=DET. when he-built  
 w-qyḥ b'ry 'lhw d-s<sub>1</sub>m[wy d-]yḡ[rw]  
 and==he-plastered wells.DUAL-of god=his dū-Samāwī REL.=Yaḡruw  
 “Ġašnum, the son of Gāniyat, the Ġulwānite, [wrote this] when he built  
 and plastered the two wells of his god dū Samāwī of Yaḡruw” (Ko 4/1–6)

### 5.3.3 Circumstantial clauses

Circumstantial clauses expressing simultaneity with the verbal predicate, analogous to the Arabic type *wa-huwa yaf'alu*, cannot be identified in Old South Arabian with certainty. With a nominal predicate, however, such syntagms are attested in Middle and Late Sabaic and in the inscriptions in the Haramic dialect (Nebes 1990):

- (20) A. w-s<sub>3</sub>mkw bn ḥyrthmw mhs<sub>1</sub>knm  
 and==they-went-up from encampment=their Muhaskanum  
 w-' frs<sub>1</sub>hmw b-'nḥrm  
 and==riders=their on==fast-horses=INDET.  
 w-ṭrydm  
 and==well-conditioned-horses=INDET.  
 “And they went up from their encampment Muhaskanum, their riders on fast, well-conditioned horses” (J 576/15–16)
- B. bhn qrbh mr' ywm tlt ḥgtn  
 because he-approached=her man day-of third pilgrimage=DET.  
 w-h' ḥyḍ  
 and==she menstruating  
 “Because on the third day of the pilgrimage a man had approached her, when she was menstruating” (C 533/2–4; Haramic)

The nominal clause that is simultaneous with what precedes may also be introduced by the temporal conjunction *bkn* or *k-*

- (21) A. w-[b-d]t s<sub>2</sub>fthw rmn b-mqmtm bkn  
 and==in==REL. he-promised=him Rummān in==power=INDET. when  
 'bhw dn[m y]zfr ws<sub>1</sub>t ḍr ḥmyrm  
 father=his Danam Yaḡfur in war-of Himyar.  
 “And considering that Rummān promised him with power, when his father Danam Yaḡfur found himself at war with Himyar” (C 140/10–12)
- B. w-qds<sub>1</sub>w b't mrb k-b-hw  
 and==they-consecrated church-of Mārib while==in==it  
 qs<sub>1</sub>s<sub>1</sub>m  
 priest=INDET.  
 “And they consecrated the church of Mārib, while a priest was there”  
 (C 541/66–67; Late Sabaic)

### 5.3.4 Causal clauses

Causal relationships are formed with the conjunctions (*l-*)*qbl(y)* *d-/dt*; less often *ʾln*, *ʾln d-/dt*; in the inscriptions in the Haramic dialect with *bhm*:

- (22) A. *hqny ʾlmqh d-hrn dn ms<sub>3</sub>ndn l-qbl dt*  
 he-dedicated Almaqah REL.=Hirrān this tablet=DET. because  
*s<sub>1</sub>ʾlhw ʾlmqh b-ms<sub>1</sub>ʾlhw*  
 he-asked=him Almaqah in==oracle=his  
 “He dedicated this (bronze) tablet to Almaqah of Hirrān, because Almaqah had asked him in his oracle” (C 79/1–4)
- B. *w-ʾl hrb b-hwt wrhn ʾln d-ʾl*  
 and==not he-undertook in==this month=DET. on-account-of REL.=not  
*tqr<sup>c</sup> s<sub>1</sub>l<sup>tm</sup>*  
 he-drew lots=INDET.  
 “And in this month he did not undertake this procedure [to obtain an oracle in the temple], because he had not drawn [appropriate] lots” (NNAG 12/7–8)
- C. *tnhy w-tndrn l-d-s<sub>1</sub>mw*  
 he-publicly-confessed and==do-penance.INF. to==dū-Samāwī  
*bhn qrb mrʾtm*  
 because he-approached woman=INDET.  
 “He publicly confessed and did penance before dū Samāwī, because he had approached a woman” (C 523/1–3; Haramic)

### 5.3.5 Comparative clauses

Comparative clauses are introduced by *hgn*, *hngn*, *hg(n) dt*, or *hg(n) k-*:

- (23) *w-ḥmdw b-dt s<sub>1</sub>tkml ʾh[wnhm]w b-ʿm*  
 and==they-thanked in==REL. it-was-accomplished alliance=their in==with  
*mlk ḥbs<sub>2</sub>tn hgn s<sub>1</sub>tkml ʾhwnhmw b-ʿm*  
 king-of Abessynia=DET. just-as it-was-accomplished alliance=their in==with  
*yd<sup>c</sup>b ḡyl[n ml]k ḥḍrmwt b-qdmy dt hqnytn*  
 Yadaʿab Ḡaylān king-of Hadramawt in==before this dedication=DET.  
 “And they thanked [Almaqah] that their alliance with the king of Abessynia came into being, just as their alliance with Yadaʿab Ḡaylān, the king of Hadramawt, had come into being before this dedication” (C 308/14–16)

### 5.3.6 Conditional sentences

The conditional particles of the protasis are *hm* and *hmy*; the apodosis is introduced by *f-*, *w-*, or *φ*:

- (24) A. *w-hm ʾl tʾḥd f-ḥlt nfs<sub>1</sub>hw*  
 and==if not he-is-seized and==it-is-at-the-mercy-of life=his  
*l-d-yhrghnw*  
 to==REL.=he-kills-him  
 “And if he is not seized, then his life is at the mercy of him who kills him”  
 (R 4088/4–8)



B. w-hmy hfkn f-t'lmn b-hmy  
 and==if you-sent and==sign.IMPERATIVE in==them  
 “And if you send [the two copies of the contract], then sign them”  
 (YM 11749/2)

The temporal *bkn* has a conditional nuance when the predicate is *yfʿln* (imperfect; see Nebes 1994b: 49):

- (25) w-bkn ymtn bʿrm b-ʿm d-ys₂'mnhw  
 and==when it-dies head-of-cattle=INDET. in==with REL==he-buys=it  
 w-ygzn s₁bʿm ywmm f-brʿm  
 and==it-passes seven=INDET. day=INDET. and==free-of-responsibility=INDET.  
 mhs₂'mn bn mwthw w-btłthw  
 buyer=DET. from death=its and==loss=its  
 “And if a head of cattle dies on the one who buys it, and seven days have already passed, then the seller is not responsible for its death and loss” (R 3910/5–6)

In an inscription in the Haramic dialect, *hn* appears as a conditional particle:

- (26) hn l-yngs₁n s₁lḥhw... l-yzlʿn l-ʿlt  
 if it-defiles weapons=his JUSS.==he-pays to==these of  
 ʿttr w-ʿrs₂wwn ʿs₂r ḥyʿlym  
 ʿAttar and==priests=DET. ten Ḥayyʿil-coins=INDET.  
 “If his weapons are defiled... then he should pay ten Ḥayyʿil-coins to the congregation of ʿAttar and to the priests as penance” (C 548/2–5)

In addition, a conditional connotation is expressed by sentences introduced by *mʿn-mw* and *mhn-mw* when the predicate has *yfʿln* (imperfect):

- (27) A. ḥgn s₂ftthw ʿmthw mbs₂mt k-mʿn-mw  
 as she-promised=him maidservant=his Mubaššimat that==as-soon-as  
 yḥmrnhw ḥyw lhw wldm thqnynhw  
 he-will-grant=her live=INF. to==her child=INDET. she-will-dedicate=him  
 “As his maidservant Mubaššimat promised him [i.e., Almaqah] that, as soon as he would grant her that a child would survive for her, she would dedicate to him” (J 717/4–7)
- B. w-s₂ftw ʿlmqhw k-mhn-mw yldn l-hmw  
 and==they-promised Almaqahū that==as-soon-as it-is-born to==them  
 bnm w-yḥywn f-yhqnynn šlmm  
 son=INDET. and=he-survives and==they-will-dedicate statuette=INDET.  
 “And they promised Almaqahū that, as soon as a son were born to them and he survived, they would dedicate a statuette” (J 669/9–12)

Iterative expressions are introduced by *ʿhnn* (-mw), *(b-)'hn* (-mw), and *hn-mw*. The subordinate clause may precede the main clause, as in the “publication-clause” found in legal contexts:

- (28) ʿhnn ʿkr w-l-yyf ʿn bn  
 whenever it-is-contested and==JUSS.==it-will-be-made-known among  
 byt d-ḥbb w-ʿqyn šrwḥ  
 house-of REL==Ḥubāb and==administrators-of Širwāḥ



- (31) A. mn-mw                      d-ys<sub>2</sub>'mn                      'b<sub>dm</sub>                      f-<sup>3</sup>w  
 whoever==ENCL. REL.==he-buys male-servant=INDET. and==or  
 'mtm                      w-b'rm                      w-s<sub>2</sub>'mtm  
 female-servant=INDET. and==cattle=INDET. and==purchase=INDET.  
 f-l-yknn                      m'dhw                      'hd wrhm  
 and==JUSS.==it-will-be period=its one month=INDET.  
 "Whoever buys a male or female servant or cattle, or makes any purchase [at all],  
 its period of time [i.e., in which the purchase price must be paid and in which  
 complaints may be registered] is to be [at most] one month" (R 3910/3–4)
- B. w-d-yrh<sub>dn</sub>                      b-hw                      l-ys<sub>1</sub>b<sub>tn</sub>                      hms<sub>1</sub>y  
 and==REL.==he-washes in==it JUSS.==he-receives fifty  
 s<sub>1</sub>b<sub>tm</sub>                      b-mqmn  
 blows=INDET. in==place=DET.  
 "And whoever washes in it [i.e., in the cistern reserved for the goddess Nawšam]  
 is to receive fifty blows on the spot" (Rob Maš 1/11–12)

### 5.4.2 Attributive relative clauses

These may be syndetic or asyndetic. Regarding the former, the rule in Sabaic is that relative clauses must be introduced by the relative pronoun if the antecedent is marked by the definite (-n) or indefinite (-m) article:

- (32) A. hqny                      'lmqh... šlmn                      d-s<sub>2</sub>fthw  
 he-dedicated Almaqah statuette=DET. REL.==he-promised=him  
 "He dedicated to Almaqah... the statuette that he had promised him"  
 (C 409/2–4)
- B. w-hmrhw                      mr'hw                      'lmqh                      hyw                      l-hw                      ḡlmm  
 and==he-granted=him lord=his Almaqah live.INF. to==him boy=INDET.  
d-ys<sub>1</sub>tmyn                      mrs<sub>1</sub>'m  
 REL.==he-is-named Marsū'um  
 "And his lord Almaqah granted him that a son, who is named Marsū'um, survived  
 for him" (J 655/7–10)

In Sabaic, asyndetic relative clauses normally require the construct state of the antecedent:

- (33) w-htb                      'bd'                      whbhw                      mlk  
 and==he-gave-back districts-of.CONSTR. he-gave=him king-of.CONSTR.  
 s<sub>1</sub>b'                      l-'lmqh                      w-l-s<sub>1</sub>b'  
 Saba' to==Almaqah and==to==Saba'  
 "And he [i.e., Karib'il] gave back to Almaqah and Saba' the districts that the king  
 of Saba' had given to him [i.e., Sumuhūyafa']" (R 3945/14–15; Early Sabaic)

These constructions, which are very common in Sabaic, and also known from Akkadian, frequently occur in connection with paronomastic expressions:

- (34) A. s<sub>1</sub>b't                      s<sub>1</sub>b'  
 campaign-of.CONSTR. he-undertook  
 "The campaign that he undertook"
- B. mrḍ                      mrḍ  
 illness-of.CONSTR. he-became-ill  
 "The illness with which he became ill"

- C. 'ml' s<sup>1</sup>tml'  
 requests-of.CONSTR. he-requested-fulfillment  
 "The requests whose fulfillment he requested"

It rarely happens that the relative pronoun is missing with a definite or indefinite antecedent:

- (35) w-kwn h' mt'tn mt' bn hwt tyln  
 and==it-was this saving-event=DET. he-saved from this lava-flow  
 b-wrḥ...  
 in==month-of  
 "And this saving event by means of which he [i.e., Ta'lab] protected [them] from  
 this lava-flow, took place in the month..." (C 323/8–9)

It is also exceptional that the relative pronoun is used with a preceding antecedent in the construct state:

- (36) bkn mt'hmw bn 'ws<sub>1</sub> d-kwn b-'rdn  
 when he-saved==them from plague-of.CONSTR. REL==it-was in==land=DET.  
 "When he [i.e., Almaqah] saved them from the plague that raged in the land"  
 (C 81/3–4)

The resumptive personal pronoun, which indicates the syntactic integration of the antecedent into the relative clause, is obligatory in genitive constructions, and sometimes also appears in the case of adverbial constructions in which the collocation *preposition + pronoun* stands before the verb of the relative clause:

- (37) A. ṣlmn d-ṣrfn d-mdlthw 'rb' m'n<sup>m</sup>  
 statuette=DET. REL==silver=DET. REL==value=its four hundred=INDET.  
 w-'ḥd 'lfm rḍym  
 and==one thousand=INDET. coins-of-good-quality=INDET.  
 "The silver statuette, whose value corresponds to 1,400 coins of good quality"  
 (J 609/4–6)
- B. ṣlmn d-d[h]bn d-b-hw ḥmd ḥyl  
 statuette=DET. REL==bronze=DET. REL==in==it he-thanked power-of  
 w-mqm 'lmqh  
 and==might-of Almaqah  
 "The bronze statuette, with which he expressed his gratitude for the power  
 and might of Almaqah" (J 739/4–5)

## 5.5 Asyndetic constructions

It should be noted that apart from asyndetic relative clauses, verbal asyndeton is markedly rare in Sabaic, confined to a few uncertain cases:

- (38) bkn rkby bn s<sub>1</sub>rn bryn yrt'nn 'dy  
 when they.DUAL-were-ridden from wadi=DET. Bāriyān they-will-graze until  
 ḥb<sup>tn</sup>  
 Ḥabtān  
 "When they [the two horses] were ridden from Wadi Bāriyān to Ḥabtān, in order  
 (?) to graze there" (J 745/9–11)

## 5.6 Infinitival syntax

As noted above (see §4.4.5), the infinitive appears in two basic constructions.

### 5.6.1 Replaceable by a finite verb

If the infinitive can be replaced by a finite verb, it continues a preceding verb paratactically with *w-*. The statement denoted by the infinitive corresponds to the preceding verb in person, tense, and mode. As a rule, the infinitive follows the verb immediately, and several infinitives may join together in an “infinitive chain”:

- (39) *w-y'ttmw*                      *w-tqdmn*                      *w-rtḏḥn*                      *b-‘m*  
 and==they-regrouped and==confront.INF. and==join-battle.INF. in==with  
*hmt ḥbs<sub>2</sub>n*  
 those Habashites  
 “And they [i.e., the Sabaeans] regrouped, came to confrontation, and joined battle with those Abessynians” (J 575/5)

### 5.6.2 Not replaceable by a finite verb

In the positions in which the infinitive cannot be replaced by a finite verb, it occurs primarily as the object. In this function it is found especially after verbs with certain meanings: for example, after verbs of granting (e.g., *ḥmr*, *ḥwfy*, *ḥws<sub>2</sub>ʿ*); of promising (e.g., *s<sub>2</sub>ft*) and of commanding (e.g., *wqh*); of preventing and hindering (e.g., *mnʿ*). In these cases the infinitive may or may not be introduced by a preposition (*ḥmr φ-fʿl(n)*, *ḥwfy φ-fʿl(n)*, *ḥws<sub>2</sub>ʿ b-fʿl(n)*; *s<sub>2</sub>ft l-fʿl(n)*; *wqh l-fʿl(n)*; *mnʿ bn fʿl(n)*), according to what the individual verb governs:

- (40) A. *b-ḏt*                      *ḥws<sub>2</sub>ʿ*                      *ʾlmqh*                      *mrʾyhmw*                      *b-s<sub>2</sub>kr*  
 in==REL. he-granted Almaqah lords.DUAL.=their in==defeat.INF.  
*w-nqm*                      *w-qtł*                      *w-ḥtlʾn*  
 and==take revenge.INF. and==kill.INF. and==subjugate.INF.  
*w-ḥs<sub>1</sub>ḥtn*                      *ḏ-rydn*                      *w-mṣrhw*  
 and==rout.INF. ḏū-Raydān and==troops=its  
 “Considering that Almaqah granted to their two lords to defeat, take revenge on, kill, subjugate, and rout ḏū Raydān and his troops” (J 2107/8–9=NAM 429/8–9)
- B. *f-ʾl*                      *ymnʿw*                      *bny*                      *gdnm...*                      *bn*                      *hyʿ l-hmw*  
 and==not they-may-prevent Banū Gadanim from perform.INF to==them  
*[h]ʾ fnwtn*                      *ms<sub>1</sub>bʾ*                      *mwn*  
 this canal=DET. watercourse-of water=DET.  
 “They may not prevent the Banū Gadanim from having this canal serve them as a watercourse” (C 611/7–8)

Less often *fʿl(n)* functions as subject, as for example in conjunction with the legal formula *ʾl s<sub>3</sub>n*:

- (41) *f-ʾl*                      *s<sub>3</sub>n*                      *qs<sub>2</sub>bn*                      *mḥmyn*  
 and==not it-is-permitted reconstruct.INF. field-irrigated-by-a-dam-canal=DET.  
 “Therefore it is not permitted to reconstruct a field irrigated by a dam-canal”  
 (C 380/4)

Infinitives with *l-* are common for purpose and result:

- (42) bkn blthmw mr'hmw 'ls<sub>2</sub>rḥ yḥḍb... l-gzmn  
 when he-sent=them lord=their Ilšaraḥ Yaḥḍib to=extirpate.INF.  
 hmt ḥbs<sub>2</sub>n  
 those Habashites  
 “When their lord Ilšaraḥ Yaḥḍib... sent them to extirpate those Abessynians”  
 (J 575/2)

Likewise the request formulas of the form *w-l-f<sup>l</sup>l(n)* that appear in the closing clauses of votive inscriptions (as in *w-l-ḥmr*, *w-l-ḥwfy<sup>n</sup>*, *w-l-ḥws<sub>2</sub>'n* “and may [the deity] grant”; *w-l-mt'n* “and may [the deity] save”; *w-l-ḥ'nn* “and may [the deity] help”, etc.) must be considered as infinitives expressing purpose in relation to the introductory *ḥqny*, albeit, in many cases, the syntactic construction of the whole inscription can only be understood if these syntagms are taken as independent clauses.

The complements of a dependent infinitive are not construed “nominally”, in the form of a construct chain, but rather “verbally” – in other words, by the use of case endings, the logical subject or object of the infinitive would be put in the nominative or accusative (Nebes 1987). This is apparently so, for instance, in the cases in which the infinitive is followed by an independent personal pronoun that distinguishes between nominative and genitive/accusative forms:

- (43) b-ḏt ḥmrhmw t'wln ḥmw w-'frs<sub>1</sub>ḥmw  
 in=REL. he-granted=them return.INF. they and==cavalry=their  
 w-gys<sub>2</sub>ḥmw b-wfym  
 and==army-their in==safety=INDET.  
 “Considering that he [i.e., Almaqah] granted them that they, their cavalry, and their army returned safely” (J 616/28–29)

## 5.7 Agreement

As a rule, the predicate agrees with a preceding subject in gender and number:

- (44) ḥt'mhw w-s<sub>2</sub>fnrm... s<sub>2</sub>mt<sub>y</sub> wṭnn  
 Uḥt'ummuhū and==Šāfnīrām they-set-up boundary-stone=DET.  
 l-lmqh  
 for==Almaqah  
 “Uḥt'ummuhū and Šāfnīrām... set up the boundary stone for Almaqah”  
 (C 389/1–5)

In the Middle Sabaic period especially, the verb often appears in the plural for an expected dual:

- (45) 's<sub>1</sub>dm ṣḥḥ w-'ḥyhw rb'wm bnw ḏ-'s<sub>2</sub>rm  
 Asadum Aṣḥaḥ and==brother=his Rabb'awām sons-of REL.==‘Āšīrum  
 ḥqnyw  
 they-dedicated  
 “Asadum Aṣḥaḥ and his brother Rabb'awām, members of the clan ‘Āšīrum, dedicated” (NAM 2659/1–2)

When the verb *kwn* “to be” forms the predicate, the rules of agreement are frequently not adhered to:

- (46) ḥwm w-ʿws<sub>1</sub> w-mwt kwn b-ʾrḏn  
 epidemic-of and==plague-of and==death-of it-was in==land=DET.  
 “Epidemic, plague, and death, which prevailed in the land” (J 645/13–14)

## 5.8 Non-Sabaic syntactic features

While in the areas of phonology and morphology the other Ancient South Arabian dialects exhibit significant differences from Sabaic and can also be clearly distinguished from one another, specific observations in the area of syntax are possible only to a very limited extent. This is connected with the fact that, in comparison with Sabaic, the textual basis for the other dialects is extremely meager, and elaborate narrative contexts on the basis of which syntactic relationships could be described are lacking. Moreover, many longer Qatabanic and Minaic inscriptions, especially in the case of legal documents, offer serious difficulties of interpretation at present because of their vocabulary. Specific differences from Sabaic and from the other dialects can be noted primarily for Qatabanic.

### 5.8.1 Attributive relative clauses

Like Sabaic, Qatabanic distinguishes three constructions: syndetic relative clauses with *d-* when the antecedent is marked as definite, and asyndetic relative clauses when the antecedent is in the construct state (especially in paronomastic constructions). If, however – as the third possibility – the antecedent is indefinite, with mimation, then as in Arabic the relative pronoun is not used:

- (47) b-kl mngwm b-yktrbw[n] ʿmn thrgs<sub>1</sub>  
 in==all-of matters they-will-ask from authority=his  
 “In all matters which they will ask from his [i.e., Warafū’s] authority”  
 (AM 177+208/10–11)

### 5.8.2 Asyndetic constructions

Qatabanic exhibits asyndetic coordination to a larger extent than Sabaic, both in nominal phrases, as in the titulature of Qatabanic rulers, *qzr qyn ršw* “treasurer, administrator, and priest”, and with finite verbs, as in:

- (48) ʿs<sub>1</sub>yw zrbw bnyw qbrs<sub>1</sub>m nfs<sub>1</sub>m  
 they-bought they-acquired they-built tomb=their Nafisum  
 “They have bought, acquired, and built their tomb Nafisum” (J 343/2)

### 5.8.3 Infinitival constructions

Dependent infinitives may exhibit an *-m* in Qatabanic:

- (49) w-hmw ys<sub>1</sub>s<sub>1</sub>lb kbrn bn lšq  
 and==if he-neglects Kabīr=DET. from prosecute.INF.  
 w-qrw w-ʿthdm w-s<sub>1</sub>ʿdbm  
 and==accuse.INF. and==look-after.INF. and==punish.INF.  
 ḥg-ḏn ḏ-mḥrn  
 according-to==this REL==ordinance=DET.  
 “And if the Kabīr neglects to prosecute, to accuse, to look after, and to punish  
 according to this ordinance” (R 3854/8–9)

## 6. LEXICON

In addition to the normal common Semitic words such as kinship terms, parts of the body, numbers, and so forth, Ancient South Arabian possesses a very independent vocabulary, which seems to be relatively isolated within the Semitic lexicon. In many cases a semantic comparison with other Semitic languages, even when the root and the corresponding derivative are attested in them, is scarcely helpful, and rarely leads to a satisfactory solution in a specific epigraphic context. As an example may be mentioned the wooden sticks, the interpretation of which is made extremely difficult not only because of the minuscule script, but primarily because of the partly unknown vocabulary.

Nevertheless, because of their geographical and chronological proximity there exist a number of lexical connections not only with North Arabian, as shown by the inscriptions in the Haramic dialect, but also with classical Ethiopic (see Müller 1983). Yet Ancient South Arabian is clearly distinct from its neighboring sister languages in vocabulary as well as in grammar. It can practically be stated that an Ancient South Arabian inscription with the (extensive) lexicon of classical Arabic or Ethiopic cannot be translated and understood properly.

Nor does Ancient South Arabian have close lexical connections with the Modern South Arabian languages, a fact that confirms the discovery, already made on the basis of morphology and syntax, that the Modern South Arabian languages in no way represent the linguistic continuation of Ancient South Arabian.

Many words, especially terms from agriculture and irrigation technology, are found in the works of Yemenite writers of the Arabic Middle Ages, and continue in part to survive today in Yemenite Arabic dialects (see al-Selwi 1987).

In the monotheistic period, the vocabulary of the Sabaic inscriptions is augmented by some Greek and Jewish Aramaic expressions, especially in the religious sphere (see Beeston 1994).

## 7. READING LIST

An informative cultural and historical survey of the present state of research into Ancient South Arabia is presented in the catalog of the Vienna Yemen-Exhibition (Seipel 1998), in which additional literature is also cited. A tightly packed, informative summary of the individual dialects is given in Beeston 1984; the detailed review of Müller 1986 should be consulted for corrections. The grammars of Höfner 1943 and Bauer 1966 contain much useful information, particularly as far as the older material is concerned, but for recently published texts, the number of which has increased sharply in the last two decades, they are no longer up to date. Recently, a detailed analysis of Sabaic phonology and morphology based on the entire epigraphic material has been prepared by Stein 2003. The relevant dictionary is *Sab. Dict.*, in which the epigraphic material published up to 1981 is critically reviewed in very succinct form. The other dictionaries are helpful only for the advanced student. Still lacking are detailed monographic presentations of the phonology, morphology, or syntax, as well as a concordance that would systematically make the vocabulary of the Ancient South Arabian dialects accessible.

Since 1973, W.W. Müller has produced an annual annotated bibliography on Ancient South Arabia in the journal *Archiv für Orientforschung* (Vienna), now available as Müller 2001, and since 1985, in *Bibliographie linguistique*, a bibliography on the South Arabian languages, in which the Ancient South Arabian dialects are also covered. A comprehensive bibliography for the ancient source material has been published recently by Kitchen 2000.



### Acknowledgments

Sections 1, 2, 4.4.2.1, 4.4.2.2, 5, 6, and 7 were written by Norbert Nebes, sections 3 and 4 by Peter Stein. The authors thank Prof. Dr. Walter W. Müller of Marburg for checking the manuscript and for valuable suggestions and proposals.

### Selected Bibliography

#### *Grammar of Ancient South Arabian*

- Avanzini, A. 1992. "H-forms in Qatabanian inscriptions". In *Yemen. Studi archeologici, storici e filologici sull'Arabia meridionale*, vol. 1, pp. 13–17. Rome: Istituto Italiano per il Medio ed Estremo Oriente.
- Bauer, G. M. 1966. *Jazyk yužnoaravijskoj piš' mennosti*. Moscow: Nauka.
- Beeston, A. F. L. 1984. *Sabaic Grammar*. Journal of Semitic Studies Monograph 6. Manchester: University of Manchester.
- Gruntfest, Y. B. 1965. "Konsektivnye konstrukcii v južnoarabskom yazyke". In *Kratkie Soobščeniia Instituta Narodov Azii*, 86, *Istorija i Filologija Bližnego Vostoka, Semitologija*, pp. 129–147. Moscow: Institut narodov Azii.
- Höfner, M. 1943. *Altsüdarabische Grammatik*. Porta Linguarum Orientalium XXIV. Leipzig: Otto Harrassowitz. Reprint 1976, Osnabrück: Zeller.
- Kogan, L. E., and A. V. Korotayev. 1997. "Sayhadic (Epigraphic South Arabian)". In R. Hetzron (ed.), *The Semitic Languages*, pp. 220–241. London: Routledge.
- Müller, W. W. 1986. "Rez. von A. F. L. Beeston, Sabaic Grammar". *Journal of Semitic Studies* 31:270–275.
- Nebes, N. 1985. "Zwei Miszellen zur sabäischen Verbmorphologie". *Bibliotheca Orientalis* 42: 27–39.
- . 1987. "Zur Konstruktion von Subjekt und Objekt abhängiger Infinitive im Sabäischen". In C. Robin and M. Bâfaqih (eds.), *Şayhadica – Recherches sur les inscriptions de l'Arabie préislamique offertes par ses collègues au Professeur A.F.L. Beeston (L'Arabie préislamique 1)*, pp. 75–98. Paris: Paul Geuthner.
- . 1988. "The infinitive in Sabaean and Qatabanian inscriptions". *PSAS* 18:63–78.
- . 1990. "Gibt es im Sabäischen 'Zustandssätze' analog dem arabischen Schema *wa-huwa yaf'alu* und *wa-huwa fi l-bayti*?" In W. Diem and A. Falaturi (eds.), *XXIV. Deutscher Orientalistentag vom 26. bis 30. September 1988 in Köln*, pp. 61–69. ZDMG Supplement VIII. Stuttgart: Steiner.
- . 1991. "Die enklitischen Partikeln des Altsüdarabischen". In *Études sud-arabes. Recueil offert à Jacques Ryckmans*, pp. 133–151. Publications de l'Institut Orientaliste de Louvain T. 39. Louvain-la-Neuve.
- . 1994a. "Zur Form der Imperfektbasis des unvermehrten Grundstammes im Altsüdarabischen". In W. Heinrichs and G. Schoeler (eds.), *Festschrift Ewald Wagner zum 65. Geburtstag*, vol. I, pp. 59–81 (BTS 54). Beirut/Stuttgart: Steiner.
- . 1994b. "Verwendung und Funktion der Präfixkonjugation im Sabäischen". In N. Nebes (ed.), *Arabia Felix. Beiträge zur Sprache und Kultur des vorislamischen Arabien. Festschrift Walter W. Müller zum sechzigsten Geburtstag, unter Mitarbeit von R. Richter, I. Kottsieper und M. Maraqtan*, pp. 191–211. Wiesbaden: Otto Harrassowitz.
- . 1995. *Die Konstruktionen mit /FA-/ im Altsüdarabischen. Syntaktische und epigraphische Untersuchungen*. Veröffentlichungen der Orientalischen Kommission der Akademie der Wissenschaften und der Literatur. Mainz 40. Wiesbaden: Otto Harrassowitz.
- . 1997. "Stand und Aufgaben einer Grammatik des Altsüdarabischen". In R. Stiegner (ed.), *Aktualisierte Beiträge zum 1. Internationalen Symposium Südarabien, interdisziplinär an der Universität Graz, mit kurzen Einführungen zu Sprach- und Kulturgeschichte*, pp. 111–131. Graz: Leykam.
- Robin, C. 1983. "Compléments à la morphologie du verbe en sudarabique épigraphique". In *Matériaux arabes et sudarabiques*, pp. 163–185. Paris: Groupe d'Études de Linguistique et de Littératures Arabes et Sudarabiques.

- Sima, A. 2001. "Altsüdarabische Konditionalsätze". *Orientalia*, 70:283–312. Rome: Pontifical Biblical Institute.
- Stein, P. 2002a. "Gibt es Kasus im Sabäischen?" In N. Nebes (ed.), *Neue Beiträge zur Semitistik. Erstes Arbeitstreffen der Arbeitsgemeinschaft Semitistik in der Deutschen Morgenländischen Gesellschaft vom 11. bis 13. September 2000 an der Friedrich-Schiller-Universität Jena*, pp. 201–222. Wiesbaden: Otto Harrassowitz.
- . 2002b. "Zur Morphologie des sabäischen Infinitivs". *Orientalia*, 71:393–414. Rome: Pontifical Biblical Institute.
- . 2003. *Untersuchungen zur Phonologie und Morphologie des Sabäischen*, Rahden/Westf.: Marie Leidorf.

### Dictionaries

- Arbach, M. 1993. "Lexique madhābien, comparé aux lexiques sabéen, qatabanite et ḥaḍramawtique". Dissertation, Aix-en-Provence.
- Avanzini, A. 1980. *Glossaire des inscriptions de l'Arabie du Sud 1950–1973. II* (ʿh). Quaderni di Semitistica 3. Florence: Istituto di Linguistica e di Lingue Orientali, Università di Firenze.
- Biella, J. C. 1982. *Dictionary of Old South Arabic. Sabaean Dialect*. Harvard Semitic Studies No. 25. Chico, CA: Scholars Press.
- Ricks, S. D. 1989. *Lexicon of Inscriptional Qatabanian*. Studia Pohl 14. Rome: Pontifical Biblical Institute.
- Sab. Dict.*: A. F. L. Beeston, M. A. Ghul, W. W. Müller, J. Ryckmans. 1982. *Sabaic Dictionary*. Louvain-la-Neuve: Peeters/Beyrouth: Librairie du Liban.

### Other literature cited

- Al-Selwi, I. 1987. *Jemenitische Wörter in den Werken von al-Hamdānī und Naṣwān und ihre Parallelen in den semitischen Sprachen*. Berlin: Reimer.
- Beeston, A. F. L. 1994. "Foreign loanwords in Sabaic". In N. Nebes (ed.), *Arabia Felix. Beiträge zur Sprache und Kultur des vorislamischen Arabien. Festschrift Walter W. Müller zum 60. Geburtstag*, pp. 39–45. Wiesbaden: Otto Harrassowitz.
- Kitchen, K. A. 2000. *Documentation for Ancient Arabia. Part II. Bibliographical Catalogue of Texts*. Liverpool: University Press.
- Macdonald, M. C. A. 2000. "Reflections on the linguistic map of pre-Islamic Arabia". *Arabian archaeology and epigraphy* 11:28–79.
- Müller, W. W. 1983. "Äthiopische Marginalglossen zum sabäischen Wörterbuch". In S. Segert and A. Bodrogligeti (eds.), *Ethiopian Studies. Dedicated to Wolf Leslau on the Occasion of his Seventy-Fifth Birthday*, pp. 275–285. Wiesbaden: Otto Harrassowitz.
- . 1994. "Die altsüdarabische Schrift". In H. Günther and O. Ludwig (eds.), *Schrift und Schriftlichkeit. Ein interdisziplinäres Handbuch internationaler Forschung*, pp. 307–312. Berlin: de Gruyter.
- . 2001. *Südarabien im Altertum. Kommentierte Bibliographie der Jahre 1973 bis 1996 unter Mitarbeit von E.-M. Wagner hrsg. von N. Nebes*. Rahden/Westf.: Marie Leidorf.
- Robin, C. (ed.). 1991. *L'Arabie antique de Karib'il à Mahomet. Nouvelles données sur l'histoire des Arabes grâce aux inscriptions*. Revue du Monde Musulman et de la Méditerranée 61. Aix-en-Provence: Edisud.
- Ryckmans, J., W. W. Müller and Y. M. Abdallah. 1994. *Textes du Yémen antique inscrits sur bois*. Publications de l'Institut Orientaliste de Louvain, 43. Louvain-la-Neuve.
- Seipel, W. (ed.). 1998. *Jemen, Kunst und Archäologie im Land der Königin von Saba*. Vienna: Kunsthistorisches Museum.
- Wissmann, H. von. 1982. *Die Geschichte von Saba* II. *Das Großreich der Sabäer bis zu seinem Ende im frühen 4. Jh.v.Chr.* Hrsg. von W. W. Müller. Vienna: Österreichische Akademie der Wissenschaften.

# Ancient North Arabian

M. C. A. MACDONALD

## 1. HISTORICAL AND CULTURAL CONTEXTS

In the western two-thirds of the Arabian Peninsula, from southern Syria to Yemen, inscriptions testify to the use of a number of different ancient languages and scripts. In the southwest, these inscriptions may date from as early as the thirteenth century BC and continue up to the seventh century AD, while in central and north Arabia they seem to be concentrated in the period between the eighth century BC and the fourth century AD. Some languages, like Aramaic and, later, Greek, came to the region from outside, but the rest were indigenous tongues expressed in scripts developed locally.

Literacy seems to have been extraordinarily widespread, not only among the settled populations but also among the nomads. Indeed, the scores of thousands of graffiti on the rocks of the Syro-Arabian desert suggest that it must have been almost universal among the latter (see Macdonald 1993:382–388). By the Roman period, it is probable that a higher proportion of the population in this region was functionally literate than in any other area of the ancient world.

### 1.1 North Arabian

The ancient languages in the southwest of the Peninsula are known as Ancient (or Old) South Arabian (see Ch. 15), while those in central and northern Arabia and in the desert of southern Syria are classed as North Arabian. This latter category is divided into two subgroups. The first of these is *Arabic*, which is subdivided into (i) Old Arabic (that is Arabic attested in pre-Islamic texts which have survived independently of the early Arab grammarians, thus the Namārah inscription but not the “Pre-Islamic poetry,” see Macdonald, forthcoming); (ii) Classical and Middle Arabic; and (iii) the vernacular dialects. The second subgroup is called *Ancient North Arabian*. The most striking difference between the two subgroups lies in the definite article, which is *ʾal-* in Arabic, but is *h-* or zero in Ancient North Arabian (see §4.3.1). Until recently, this division was largely unrecognized by linguists working outside the field, and Ancient North Arabian (which was sometimes misleadingly called “Proto-Arabic”) was usually treated as a collection of early dialects of Arabic. However, it is now clear that Ancient North Arabian represents a linguistic strain which, while closely related to Arabic, was distinct from it (Macdonald 2000:29–30).

#### 1.1.1 Arabic

Arabic, and thus by implication the North Arabian group as a whole, has traditionally been classified, along with the Ancient South Arabian, Modern South Arabian and Ethiopic



**Figure 16.1** Pre-Islamic Arabia

languages, as *South West Semitic* (e.g., Brockelmann 1908–1913: i, 6). However, more recently, it has been grouped instead with Canaanite and Aramaic, under the rubric *Central Semitic* (e.g., Faber 1997; see Ch. 6, §2.3), and this classification is certainly more appropriate for Ancient North Arabian.

Old Arabic seems to have coexisted with Ancient North Arabian throughout north and central Arabia but, in contrast to Ancient North Arabian, it remained a purely spoken language. The earliest Old Arabic inscriptions in what we think of as the Arabic script (in fact the latest development of the Nabataean Aramaic alphabet) date from the early sixth century AD. Before that, Old Arabic was written only on very rare occasions and then, necessarily, in a “borrowed” script (Ancient South Arabian, Dadanitic, Nabataean, or Greek). At present, seven such documents in Old Arabic have been identified, and in a number of others, Old Arabic features occur in texts which are otherwise in Sabaic (an Ancient

South Arabian language), Dadanitic, Safaitic, Nabataean, and possibly East Arabian Aramaic (see Macdonald 2000:50–54 and forthcoming).

### 1.1.2 Ancient North Arabian

Ancient North Arabian is made up of a number of interrelated dialects, attested only in inscriptions. These are dated roughly between the eighth century BC and fourth century AD, after which the language disappears from the record. Well over forty thousand of these texts have been discovered so far and it is known that scores of thousands remain to be recorded. However, approximately 98 percent of these are graffiti, informal inscriptions the majority of which consist only of names. The amount of linguistic evidence they can provide is therefore relatively meager and our knowledge of the structure of these dialects is extremely fragmentary – a situation exacerbated by the nature of the writing systems used (see §2). Despite this, a surprising amount of information is to be found in these inscriptions, and more is being identified every year.

Ancient North Arabian was used by the settled peoples and nomads of central and north Arabia and by the nomads in what is now southern Syria and eastern and southern Jordan. It is attested in the following dialects (see Macdonald 2000:29–30, 32–36, 40–46): (i) Oasis North Arabian (ONA), consisting of Taymanitic, Dadanitic, Dumaitic, and Dispersed Oasis North Arabian; (ii) Safaitic; (iii) Hismaic; (iv) Thamudic B, C, D, and “Southern Thamudic”; and, possibly, (v) Hasaitic.

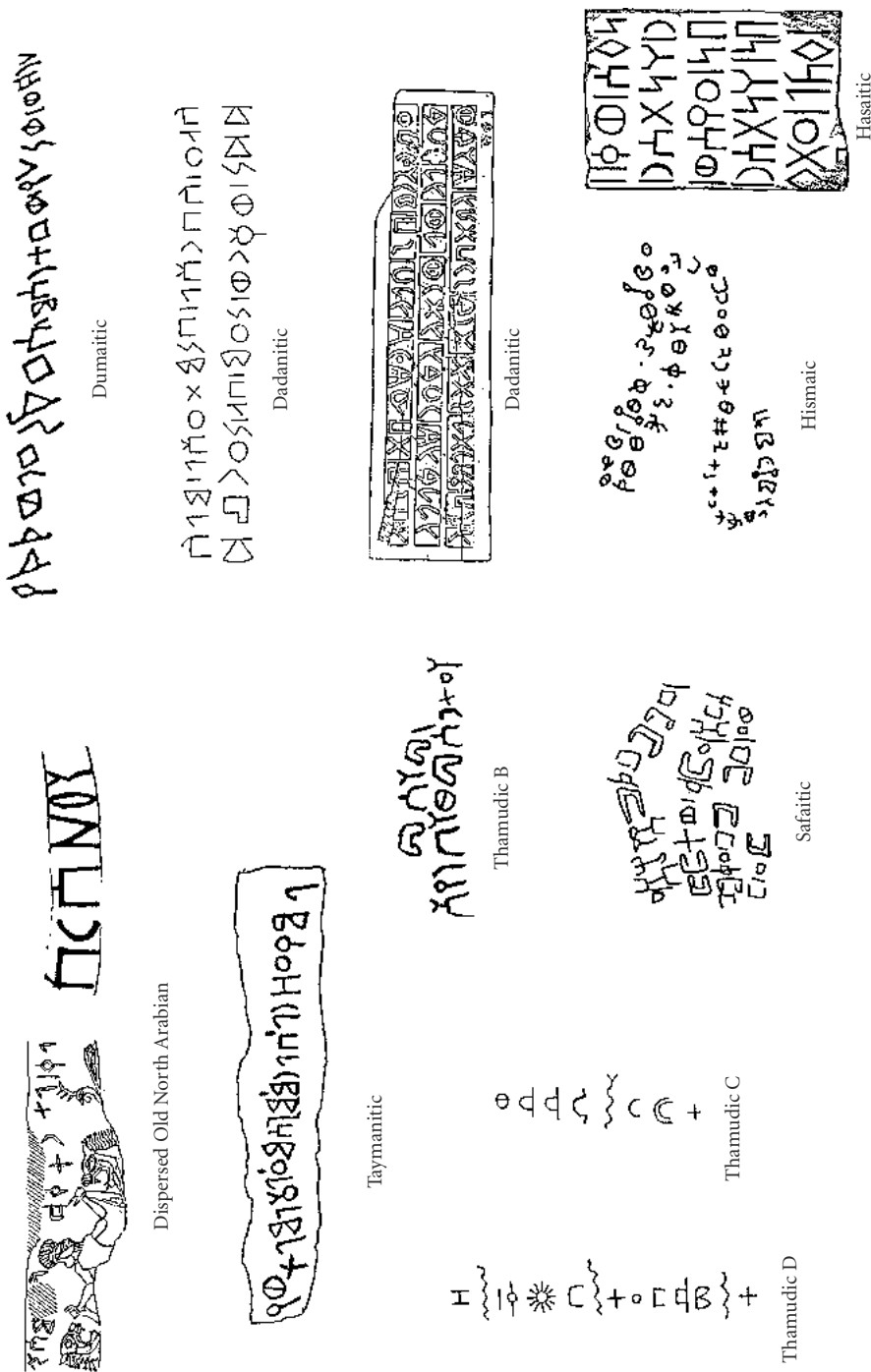
#### 1.1.2.1 Oasis North Arabian

Of these dialects, the earliest attested are those belonging to the group known as *Oasis North Arabian*. From at least the middle of the first millennium BC, local dialects of Ancient North Arabian were spoken in the major oases of northwest Arabia: Taymāʾ, Dadan (modern al-ʿUlā; for the spelling Dadan, see Sima 2000 and Macdonald 2000, n. 1) and probably Dūmā (modern al-Ġawf); see Figure 16.1. The populations of these settlements were heavily involved in the trade in frankincense and other aromatics which were brought from South Arabia to Egypt, the Mediterranean coast, Syria, and Mesopotamia where there seems already to have been a considerable Arab presence. It is therefore not surprising that brief texts in scripts similar to those used in these oases have been found outside Arabia, principally in Mesopotamia. In the past they have been known by such misnomers as “Chaldaean” and “Old Arabic,” but I have recently suggested that a better term would be *Dispersed Oasis North Arabian* (Macdonald 2000:33), a label which I hope emphasizes the fact that they are a heterogeneous collection of texts which have in common only the fact that they are written in varieties of the Oasis North Arabian alphabet and that they were found outside Arabia.

*Dumaitic* is so far represented by only three brief texts found near Sakākā in northern Saudi Arabia (Winnett and Reed 1970:73, 80–81 [WTI 21–23], 207, 216, where they are called “Jawfian”). They are in a distinctive variety of the Oasis North Arabian script (see Fig. 16.3) which differs in certain important respects from Taymanitic and Dadanitic. At present they are undatable, but they may be from the middle of the first millennium BC.

*Taymanitic* refers to the dialect and script used in the oasis of Taymāʾ and its surroundings, probably in the sixth and fifth centuries BC. It is represented by short inscriptions with very distinctive linguistic and orthographic features. The number of known Taymanitic texts has recently been doubled (from c. 200 to c. 400) by Kh. M. Eskoubi’s edition of new texts, including two which mention *nbnd mlk bbl* “Nabonidus king of Babylon,” who spent ten years of his reign 552–543 BC, in Taymāʾ (Eskoubi 1999: nos. 169 and 177; Müller and Said 2001).

*Dadanitic* is a new term which covers the inscriptions in the local language and script of the oasis of Dadan. These were formerly divided into “Dedanite” and “Lihyanite,” following



**Figure 16.2** Examples of the Ancient North Arabian scripts

the nomenclature of successive kingdoms in the oasis, but, needless to say, linguistic and paleographical developments did not necessarily parallel political changes, and this particular subdivision has proved misleading. Dadanitic is the only Ancient North Arabian dialect and script in which large numbers of monumental inscriptions were written. These are concentrated in and around the oasis, with only occasional examples found elsewhere. In addition, there are hundreds of Dadanitic graffiti in and around the settlement. There is no firm dating evidence for the inscriptions of Dadan, though dates ranging from the sixth century BC through the first century AD have been proposed. Dadan was also the site of a South Arabian (Minaean) trading station and there are numerous monumental inscriptions and graffiti in *Maḏhābic*, the South Arabian language used by the Minaeans (see Ch. 15). The prosperity of Dadan may have been eclipsed in the first century AD by the neighboring oasis of Ḥegrā (modern Madā'in Šālīḥ), some twenty kilometers to the north, which became an important city of the Nabataean kingdom.

#### 1.1.2.2 *Safaitic*

This is the language of most of the graffiti found in the deserts of black, broken-up lava in southern Syria, northeastern Jordan, and northern Saudi Arabia. The vast majority were written by the nomads who lived in this area between roughly the first century BC and the fourth century AD. So far, some twenty thousand Safaitic inscriptions have been recorded, and there are many times this number still awaiting study, as can be seen by any visitor to these desert areas.

#### 1.1.2.3 *Hismaic*

Hismaic was the language of the nomads of the Ḥismā sand-desert of southern Jordan and northwest Saudi Arabia, and some of the inhabitants of central and northern Jordan. They were contemporaries and close neighbors of the Nabataeans, whose capital, Petra, was not far away from the northern end of the Ḥismā in Wādī Ramm, southern Jordan. Thus, they probably date to the first centuries BC/AD and possibly a little later. In the past, Hismaic has been called "Thamudic E" (see below), and misleadingly "Tabuki Thamudic" and "South Safaitic." The last-mentioned is a complete misnomer since the dialect and script are quite distinct from those of Safaitic.

#### 1.1.2.4 *Thamudic*

Thamudic is not the name of a dialect or script but of a sort of "pending" category into which are placed all texts which appear to be Ancient North Arabian but which are not Oasis North Arabian, Safaitic, or Hismaic. Both Taymanitic (formerly "Thamudic A") and Hismaic (formerly "Thamudic E") were originally included in this category until the advent of properly recorded texts and intensive studies made it possible to define them as distinct dialects with their own scripts (see Macdonald and King 1999). The rubrics "B," "C," "D," and "Southern Thamudic" represent relatively crude subdivisions of those texts still in this "pending" category. There is no way of dating most of these inscriptions, though one Thamudic B inscription (Ph 279 aw) appears to mention a "king of Babylon" and so presumably dates to a time before the fall of the Babylonian Empire in 539 BC. By contrast, a Thamudic D inscription (JSTham 1) at Madā'in Šālīḥ (ancient Ḥegrā) gives a summary of an adjacent Nabataean tomb inscription which is dated to AD 267. The vast majority of the Southern Thamudic texts remains unpublished, but for an excellent summary presentation see Ryckmans 1956.

#### 1.1.2.5 *Hasaitic*

This term refers to the language of a number of inscriptions, almost all gravestones, most of which have been found in northeastern Arabia. They consist almost entirely of genealogies



and exhibit very few linguistic features. The language is regarded (provisionally) as Ancient North Arabian because of certain characteristic expressions such as  $\underline{d}^{\prime}l$  “of the lineage of” (see §3.1.1). They are written in the Sabaic (Ancient South Arabian) script, with certain minor adaptations.

## 1.2 Sources of Ancient North Arabian

A large number of the Safaitic, and the vast majority of the Thamudic, inscriptions published so far, were recorded in the nineteenth and early twentieth centuries and are known only from hand copies, often by copyists who could not read the script. Many of these copies are inaccurate, and, in the case of the texts classed as Thamudic, this has proved a major obstacle to their successful interpretation. It is only since large numbers of texts have been photographed that the study of Taymanitic, Safaitic, and Hismaic has been placed on a secure footing.

The dialects of Ancient North Arabian on which we have most information are Dadanitic and Safaitic. The discussion below will therefore concentrate mainly on these, with details from the others where they are available.

The principal resource in the interpretation of the Ancient North Arabian inscriptions has always been the grammar and vocabulary of Classical Arabic and this has been both a blessing and a curse. On the credit side, Classical Arabic has provided a model against which the linguistic phenomena attested in Ancient North Arabian can be evaluated, though there is always a temptation to interpret the, often enigmatic, data in such a way as to make them fit this model, thus obscuring real differences (as is the case in Caskel 1954). Moreover, it should never be forgotten that, unlike most languages, Classical Arabic represents a conscious choice and amalgam of dialects and, to a greater or lesser extent, a systematization of grammatical structures by Arab scholars of the eighth and ninth centuries AD.

Similarly, it should be remembered that the concept of a descriptive dictionary of a living language is no older than the nineteenth century. Prior to that, the purpose of a dictionary was prescriptive, fixing the language in what was considered to be its most “correct” form. Thus, even the immensely rich vocabulary of Classical Arabic represents a choice by the grammarians and lexicographers of what was available to them, and much that might have helped in the reconstruction of Ancient North Arabian was no doubt excluded. Arabic dictionaries can anyway be a trap to the unwary, since they contain meanings which have developed over a wide geographical area and many centuries of intense literary activity, but with little or no indication of when and where a particular sense is first attested. Moreover, as in all languages, words can have meanings which are restricted to certain contexts, and, unless these are quoted (as they are in the great Arabic-Arabic lexica, but not in shorter European compendia), a completely false interpretation can be given. The widespread misapprehension that Ancient North Arabian texts can be read simply by using an Arabic dictionary has led many astray and has resulted in a far greater degree of uncertainty in the interpretation of Ancient North Arabian than in most other ancient languages.

One further point should be noted. In the past, some discussions of Ancient North Arabian grammar have sought to identify linguistic features in the personal names found in Ancient North Arabian inscriptions and have then treated these as if they represented the language of the texts (e.g., Littmann 1943:xii–xxiv; Caskel 1954:68–71; and even sporadically in Müller 1982). Not surprisingly, this has led to confusion, with marked differences appearing between the apparent linguistic features of the names and those of the language used by their bearers. It is important to remember that a name does not “mean” anything except the person, group, place, and so forth to which it refers. It is usually only in exceptional



circumstances that parents invent one (e.g., the seventeenth-century English Puritan called "Praise-God Barebones"). Names often continue in use over a very long period and can travel extensively, so the vast majority of names available to parents in any particular society at any particular time have been inherited, often from a linguistic environment very different from their own. The etymology of a name, while interesting in itself, is therefore linguistically irrelevant to the text in which it appears.

In this chapter, the following conventions will be used: /d/ = the etymological phoneme; [d] = the sound; *d* = the letter in a particular script. Letters between { } are doubtful readings. Many Ancient North Arabian texts have been reread or reinterpreted since their original publication, so in some cases the readings and interpretations quoted here will differ from those in the original editions. All examples quoted have been checked on photographs whenever these are available.

## 2. WRITING SYSTEMS

It is generally held that the Semitic consonantal alphabet was invented in the first half of the second millennium BC (see Ch. 12, §2.2). Later in the same millennium, two separate traditions developed out of the proto-alphabet, each with its own letter-forms, letter-order and (possibly) letter-names. One was the Phoenico-Aramaic (or Northwest Semitic), from which are ultimately derived almost all traditional alphabetic scripts in use today. The other was the Arabian (or South Semitic) alphabetic tradition, which was used almost exclusively in Arabia in the pre-Islamic period and which was the basis of the Ethiopic syllabary (see Ch. 14, §2), the only form in which it survives today (Macdonald 2000:32).

The Arabian alphabetic tradition is subdivided into two families: (i) the Ancient South Arabian, of which Sabaic is the most famous and from which the Ethiopic syllabary was developed; and (ii) the Ancient North Arabian. While the Ancient North Arabian scripts are clearly related to each other and to the Ancient South Arabian, the exact relationship has not yet been established. One problem is the lack of securely dated texts from both North and South Arabia; a second has already been touched on – the fact that so many Ancient North Arabian inscriptions are known only from unreliable hand copies. However, the major obstacle to a paleographical analysis of the Ancient North Arabian inscriptions is the fact that the vast majority of them are informal texts written by innumerable individuals who learned to write, not in schools, but casually from a companion, and whose letter-forms were not therefore part of a slowly evolving tradition, but represent a multiplicity of individual choices (Macdonald 1993:382–388; 2004a).

An indication of this is provided by the four Safaitic abecedaries which have been discovered so far. Each is in a different letter-order and none of them bears any relation to the inherited orders of the Northwest and South Semitic alphabets. The letters have simply been arranged according to the writers' differing perceptions of similarity in their shapes (see Macdonald 1993:386 and Macdonald *et al.* 1996:439–443). By contrast, the only known Dadanitic abecedar is in the South Semitic letter-order, while the unique Hismaic example more or less follows the Northwest Semitic order, but with significant differences which suggest that it was unfamiliar to the writer (Macdonald 1986:105–112).

The alphabets of Dadanitic, Hismaic, and Safaitic are each made up of twenty-eight letters. This is probably also true of Thamudic B, C, and D and Hasaitic, though some signs have yet to be identified in these scripts. Taymanitic seems to have had a slightly different phonemic repertoire from the other Ancient North Arabian dialects (see §3.1.2), and only twenty-six or twenty-seven letters have been identified with certainty.

Figure 16.3 shows the most common letter-forms in the different Ancient North Arabian scripts. With the exception of the sign for *ḡ* and the leftmost sign for *f*, the forms in the Hasaitic row are those of the South Arabian alphabet. It will be noted that the forms of some letters are remarkably stable throughout all the scripts: for example, , ' *t*, *w*, and *y*. On the other hand, in some cases the same, or very similar, shapes are used in different alphabets to represent quite different phonemes. Thus, the sign used for *g* in Hismaic is identical to that for *t* in Thamudic B, Safaitic, and South Arabian/Hasaitic; while the sign for *d* in South Arabian (and Hasaitic) is used for *d* in Thamudic B, C, and D and in Safaitic, but for *t* in Hismaic. The reasons for this are not yet understood.

In the scripts used by the inhabitants of the great oases, namely, Dumaitic, Taymanitic, and Dadanitic, the direction of writing is almost always right-to-left. In Taymanitic, texts of more than one line were often, but by no means always, written boustrophedon (i.e., continuously, with the lines running in alternate directions). However, the practice of breaking at the end of the line and placing the beginning of the next line under that of the one before is also quite common in Taymanitic and is the norm in Dadanitic. Texts were written without spaces between the words, but word-dividers are the norm in Dadanitic monumental texts and are commonly, though not consistently, used in Dadanitic graffiti and in Taymanitic and Dumaitic. Hasaitic is written either in separate lines or boustrophedon and, since it uses the South Arabian script, employs word-dividers.

By contrast, the scripts used primarily by nomads (Thamudic B, Hismaic, and Safaitic) can be written in any direction (left to right, right to left, downwards, upwards, in a circle or coil, etc.). They meander across the uneven surfaces of the rocks on which they are carved, over the edge onto an adjacent face and occasionally onto an adjacent rock. They are written continuously without word-dividers (Macdonald 2004c). This absence of word-dividers applies equally to Thamudic C and D, which were probably also written by nomads, though these show a marked preference for writing in vertical columns.

In common with all Semitic alphabets, the letters of the North Arabian scripts represent consonants only. However, in contrast to most of the Northwest Semitic scripts, none of the South Semitic alphabets, with the exception of Dadanitic, developed *matres lectionis*, letters which, in addition to their consonantal values, can in certain contexts represent a long vowel. It has been suggested that in Safaitic the letters , ' *w*, and *y* were occasionally used to represent long vowels (Winnett and Harding 1978:12; Robin 2001:553), but this is incorrect and the handful of examples quoted can all be more convincingly explained in other ways.

However, in Dadanitic, final /a:/ was usually represented by *-h* (as in Hebrew) and final /u:/ by *-w*, though the evidence for other *matres lectionis* is less convincing (Drewes 1985). In contrast to the Northwest Semitic scripts, the letter ' *alif* does not seem to have been used to mark a vowel in Ancient North Arabian.

The diphthong /ai/ is represented in final position in Dadanitic (*pace* Drewes 1985:170–171), though the representation of final /au/ is much less certain. However, diphthongs (if they existed) are rarely if ever represented in the other Ancient North Arabian scripts. Thus, in Safaitic the word for “death” appears as *mt* (cf. Arabic *mawt*), that for “raiding party” as *gs*<sup>2</sup> (cf. Arabic *ḡayš*), and so forth. Littmann claimed that Greek transliterations of names apparently similar to those found in the Safaitic inscriptions showed that the diphthongs /ai/ and /au/ had been monophthongized to [e:] and [o:] respectively (1943:xiii). However, by the Roman period, there were no appropriate diphthongs left in Koine Greek with which to transliterate any which may have existed in Safaitic, so the question must remain open.

As in all Semitic alphabets, doubled consonants are written singly in the Ancient North Arabian scripts (e.g. \**umm* “mother” appears as ' *m*). However, it has been suggested



that doubled /l/ and /n/ are occasionally expressed in writing. This is based mainly on the spelling *kll* “all” (cf. Classical Arabic *kull*) which is found in Dadanitic, Hismaic, and Safaitic (Littmann 1943:xiii). But it is perfectly possible that the word was pronounced with a short vowel between the two *l*s (e.g., \**kulil*). The other supposed examples of this feature are also capable of alternative explanations (see §4.2.1) and at present the hypothesis must be regarded as not proven.

### 3. PHONOLOGY

#### 3.1 Consonants

Given the nature of the sources, our knowledge of the phonology of the dialects of Ancient North Arabian is necessarily fragmentary. Most dialects appear to have had a consonantal phonemic repertoire of roughly twenty-eight sounds. Unless there is evidence to the contrary, these are usually assumed to have been similar, though not always identical, to their equivalents in Classical Arabic. They are presented in Table 16.1 using the Roman letters with which Ancient North Arabian texts are conventionally transliterated, rather than phonetic symbols, to emphasize that this is a purely hypothetical schema based partly on the traditional pronunciation of the cognate phonemes in Classical Arabic, as described by the early Arab grammarians (eighth century AD), and partly on reconstructions (see below).

The phonemes /b, /d/, /d̥/, /h/, /k/, /l/, /m/, /n/, /t/, /t̥/, /w/, /y/, /z/ were probably pronounced more or less like their equivalents in Classical Arabic. There is no way of telling whether certain phonemes had aspirated allophones (the so-called “bghadhkphath”), as, for example, in Masoretic Hebrew and Aramaic of the Christian era. The phoneme shown here as /f/, could have been pronounced [p] in some or all positions (as in Ugaritic, Hebrew, Aramaic, Akkadian, etc.) or as [f] throughout, as in Arabic. It is worth noting that in Safaitic (as also in early Arabic) the letter *f* is used to transliterate both Greek φ and π (e.g., *flfš*

**Table 16.1 The consonantal phonemes of Ancient North Arabian**

Manner of articulation	Place of articulation								
	Bilabial	Labio-dental	Inter-dental	Dental/Alveolar	Palato-alveolar	Palatal	Velar	Uvular	Pharyngeal/ Glottal
<i>Stop</i>									
Voiceless				t			k	q	ʔ
Emphatic				t̥					
Voiced	b			d			g (?)		
<i>Fricative</i>									
Voiceless		f	t̥	s <sup>3</sup>	s <sup>1</sup>	y		ħ	h
Emphatic			z̥	š					
Voiced	w		d̥	z			ġ		ʕ
Emphatic			d̥						
<i>Trill</i>				r					
<i>Lateral cont.</i>									
Voiceless				s <sup>2</sup>					
Voiced				l					
<i>Nasal</i>	m			n					

for Φίλιππος), the well-known confusion of [b] and [p] in Arabic being a much later phenomenon.

### 3.1.1 Stops

In Hismaic, there is a small amount of evidence for the occasional confusion of /d/ and /ḏ/, probably under the influence of the Aramaic used by the neighboring Nabataeans: for example, *d-s<sup>2</sup>ry* for the divine name *ḏ-s<sup>2</sup>ry*; *dkrt* for *ḏkrt*; and *d l* “he of the lineage of” for *ḏ l* (Macdonald 2004d). However, there is no evidence for the supposed alternation of /t/ and /ṭ/ in this dialect. On both these, see King 1990:69–70. However, in Dadanitic the numeral “three” is found as *tlṭṭ*, *ṭlt*, and *tlt* (see §4.4.1 and Table 16.2) which might suggest a weakening of the distinction between these two sounds in this dialect, though it may equally have been confined to the phonetic conditions of this particular word.

It is impossible to tell whether /g/ was pronounced [g], as in some Arabic dialects, or [j] as in Classical Arabic, or even [ž] as in some dialects of Syria and Southern Iraq. It is also impossible to determine whether /k/ had an allophone [č] in certain positions, as in many dialects in Syria, Iraq, Arabia and the Gulf Coast.

The phonemes /h/ and /ḡ/ were probably realised as [x] and [ɣ] respectively as in Arabic. The consonant transcribed /q/ in Table 16.1 may have been a uvular stop as in Classical Arabic, or, alternatively, an “emphatic” correlate of /k/ (i.e., /kʔ/), as in Hebrew and Aramaic. Whatever its exact pronunciation it appears generally to have remained distinct since only one instance has so far been identified in which it is confused with another phoneme. This is in an unpublished Safaitic text in which the author spells the word *qyz* “he spent the dry season” as *ʔyḏ* in an unequivocal context. This is the earliest attestation of a pronunciation in which the etymological phonemes /q/ and /z/ had fallen under /ʔ/ and /ḏ/ respectively, a feature of modern urban Arabic in such cities as Damascus, Jerusalem, and Cairo.

In the orthography of the Ancient North Arabian scripts, the letter ʾ represents a phonemic consonant in all contexts and never the equivalent of Classical Arabic *hamzat al-waṣl*, that is, a prosthetic glottal stop, the sole function of which is to carry an initial vowel and which disappears when the latter is assimilated to a preceding vowel. Thus *bn* (“son,” in all positions) as against Classical Arabic (ʾ)*ibn*. This contrasts with Old Arabic personal names found in Nabataean orthography (for instance in the Nabataean inscriptions of Sinai), where ʾ is regularly written in *ʔbn* (e.g., the name *ʔbn-l-qyny*). For a discussion of this phenomenon see Macdonald, forthcoming. There are a few personal names in Safaitic texts written with two successive ʾs, e.g., ʾʾs<sup>1</sup>d (cf. Classical Arabic *āsud* < \*ʾaʾsud; see Littmann 1943:xii–xiii), but as yet no examples in words have been identified, so we do not know whether this was a living feature of the language or merely a fossil inherited in particular names.

Very occasionally, ʾ is found unexpectedly in medial position and it has been suggested that this may represent a medial /a:/ (Winnett and Harding 1978:12). However, this is highly unlikely and the few examples cited are all capable of other explanations.

The ending which in Arabic appears as *-ah* in pause but *-at* before a vowel (i.e., *tāʾ marbūʿah*), is always written as *-t* in Ancient North Arabian, implying that it was pronounced \*-at in all contexts.

### 3.1.2 Fricatives

The voiceless nonemphatic sibilants in Ancient North Arabian, Ancient South Arabian, Old Arabic, and Classical Arabic up to the ninth century AD, present a complex problem (see

Beeston 1962). Proto-Semitic had a voiceless dental fricative \*/s/, a voiceless palato-alveolar fricative \*/š/, and a third sibilant, conventionally written \*/ṣ/, the exact nature of which is uncertain but which may have been a lateral dental fricative [ɬ]. While the Ancient (and Modern) South Arabian languages (in common with Hebrew and early Aramaic) retained all three, in Arabic and, with one possible exception, the Ancient North Arabian dialects they were reduced to two:

(1) The voiceless nonemphatic sibilants in Ancient North Arabian

<i>Proto-Semitic</i>		<i>Ancient North Arabian (except Taymanitic)</i>		<i>Proto-Semitic</i>		<i>Taymanitic</i>
* /š/	}	→ [š] (written <i>s</i> <sup>1</sup> )		* /š/	→	[š] (written <i>s</i> <sup>1</sup> )
* /s/				* /s/	→	[s] (written <i>s</i> <sup>3</sup> )
* /ṣ/				* /ṣ/	→	[ɬ] ? (written <i>s</i> <sup>2</sup> )

We know from the phonetic descriptions by the early Arab grammarian Sibawaihi (died c. AD 796) that in early Classical Arabic, س the reflex of Proto-Semitic \*/s/ + \*/š/, was pronounced something approaching [š], and that ش the reflex of Proto-Semitic \*/ṣ/, was pronounced something approaching [ɬ]. It was only subsequently that the pronunciation of س shifted to the [s] (*sīn*), and that of ش to the [š] (*šīn*) of later Arabic. This can be tabulated as follows:

(2) The voiceless nonemphatic sibilants in Arabic

<i>Proto-Semitic</i>		<i>Arabic before the 9th century AD</i>		<i>Arabic after the 9th century AD</i>
* /š/	}	→ [š] (written س)	→	[s] (written س)
* /s/				
* /ṣ/				[š] (written ش)

This means that Ancient North Arabian /s<sup>1</sup>/ (which is cognate with later Arabic س *sīn*) was actually pronounced like something approaching [š], while Ancient North Arabian /s<sup>2</sup>/ (which is cognate with later Arabic ش *šīn*) was probably pronounced something like Welsh -ll- [ɬ]. These findings are confirmed by the treatments of loans from Aramaic. Thus, for example, the Aramaic name of the great Syrian sky-god, Ba'al-Šamīn "lord of heaven," was borrowed into Dadanitic and Safaitic as *b'ls'mn*, that is, with Aramaic /š/ represented by Ancient North Arabian *s*<sup>1</sup>, not *s*<sup>2</sup>.

It follows from this that Ancient North Arabian (and Arabic before the ninth century AD) had no [s]. However, there is one possible exception. Taymanitic appears to have had a letter, graphically related to South Arabian *s*<sup>3</sup> (= [s]), which seems to represent [s] in transliterations of the name of the Egyptian god Osiris occurring in two personal names. Rather different forms of what is probably the same letter have been identified in two other Taymanitic texts (see Müller and Said 2001:114–116) and there is one further example on a seal of Babylonian design, but in a context which raises considerable difficulties. Since, at present, only a little over four hundred Taymanitic inscriptions are known, and few of them are more than twenty letters long, no firm conclusions can be drawn from this until more evidence appears. However, it seems unlikely that the Taymanitic alphabet would have employed a letter to represent a sound which did not exist in the Taymanitic dialect, and



so there is certainly a possibility that, at some stage in its history, Taymanitic used all three voiceless nonemphatic sibilants (see Macdonald 1991).

In Taymanitic, Thamudic D, and possibly Thamudic C, it seems that /d/ had probably merged with /z/ (as in Hebrew), since the *z* sign is used for both phonemes.

### 3.1.3 Emphatics

The etymological phonemes /s/, /t/, /d/, and /z/ are emphatics. In most Semitic languages /s/ is the emphatic correlate of [s]. However, since there was no [s] in Safaitic and Hismaic, *ṣ* is often used in these dialects to transliterate Greek *sigma* (e.g., *qsr* for καῖσαρ [“Caesar”]; *flṣ* for Φίλιππος [“Philip”]; etc.) and in the Hismaic abecetary *ṣ* is put in the position of Phoenico-Aramaic *samek* (= [s]). It is not certain whether this implies a weakening of the “emphatic” quality or whether it was simply felt to be the nearest equivalent to the foreign sound. The fact that in other transliterations the letter *s*<sup>1</sup> (approximately[š]) was used for Latin *s* (e.g., *tts*<sup>1</sup> for *Titus*) and Greek *sigma* (e.g., *grgs*<sup>1</sup> for Γεωργός [George]), points perhaps to the latter (see Macdonald 1992b).

The phoneme /t/ was almost certainly the emphatic correlate of /t/, and /d/ was, at least in origin, that of /d/. However, the Akkadian transliteration of the Ancient North Arabian divine name *rdw* as *Ruldaiu* points to a strongly lateralized pronunciation of /d/, at least in North Arabia in the seventh century BC. It has also been suggested that the god Ὀροτάλτ, who Herodotus says was worshiped by the Arabs in eastern Egypt in the fifth century BC, represents a garbled transliteration of a similar pronunciation of the divine name *rdw*, though this is more speculative. On the other hand, in the Roman period, Greek transcriptions of names which include /d/ always represented it by *sigma* (e.g., Σαίφηνος for *h-dfy*, “the Dayfite”, Macdonald 1993:306). In Nabataean, native Aramaic words show the cognate of North Arabian /d/ as /ʿ/ ([f]) (e.g., Nabataean ʿr “against Safaitic ʿr “earth, land”), as is normal from Imperial Aramaic onwards. However, in loanwords and transcriptions of names which are linguistically North Arabian, /d/ is consistently represented by *ṣ* (e.g., Nabataean *ṣryh* from Arabic *darīh* “trench, cist,” or the name *rṣwt* as against Safaitic *rdwt*). Kofler quotes examples of the confusion of /d/ and /s/ in early Arabic dialects and suggests that /d/ may have been pronounced more as a fricative than a stop (1940–1942:95–97). There is no example in Safaitic and Hismaic of a confusion of /d/ and /s/, so the two sounds seem to have remained distinct in these dialects. However, if /d/ was pronounced as the emphatic correlate of /d/ (rather than of /d/), i.e., as an emphatic interdental fricative, as it is in all modern Bedouin dialects, it would have shared its place of articulation, emphatization, and fricative release with /s/, and the two sounds would have been sufficiently similar for /d/ to be transcribed by /s/ in scripts such as Nabataean Aramaic which had no letter for /d/ (I owe this interesting observation to Professor Clive Holes).

The conventional symbol *z* (originally taken over from the Cairene pronunciation of Classical and Modern Standard Arabic) is unfortunate since the phoneme it is intended to represent was probably the emphatic correlate of an interdental (/t/, or perhaps /d/), and not a dental sibilant. The former would be more likely, at least in Hismaic and Safaitic, if, as suggested above, /d/ was pronounced as the emphatic correlate of /d/. In Dadanitic, Hismaic, and Safaitic, /z/ is clearly distinguished from other phonemes except in the one example of ʿyḏ for *qyz* mentioned above. It has been suggested that, in Dadanitic, /z/ might have fallen under /t/ (as in Aramaic), but no conclusive evidence has yet been presented for this shift and the two phonemes appear to be represented by distinct letter-forms. A sign for *z* has not yet been identified in Dumaitic, Taymanitic, Thamudic B, C, and D, or in Hasaitic, but since it is a relatively rare phoneme, it is, at present, impossible to determine whether this is significant.

### 3.1.4 The sounds /w/ and /y/

In Safaitic, there is considerable alternation of /w/ and /y/, which when represented in the Ancient North Arabian scripts are always consonants, not vowels (Robin 2001: 553 is incorrect on this point). This variation is found in all positions, e.g., *wrh/yrh* “month”; *ts<sup>2</sup>wq/ts<sup>2</sup>yq* (unpublished) “he longed for”; *s<sup>2</sup>ty/s<sup>2</sup>tw* (CSNS 324) “to winter.” In each case, the first item in these pairs is the common form and the second a much less frequent variant. Given the difficulty of dating most of the texts, it is impossible to say at present whether these variations represent chronological developments or synchronic dialectal differences.

However, forms with -w and -y are almost equally common in the divine name *rdw/rdy* in Safaitic inscriptions. This deity is also found in Dumaitic and Thamudic B texts, but there only as *rdw*. The Dumaitic, and at least some of the Thamudic B inscriptions, are considerably earlier than the Safaitic, and this might seem to suggest that the form *rdw* is the older and that the advent of *rdy* marks a change of pronunciation. However, the Akkadian transliteration *Ruldaiu*, which is securely dated to the early seventh century BC, implies a pronunciation \**ruḏayu* (i.e., *rdy*), and it therefore seems more likely that the two spellings represent dialectal (?) differences. It is not yet possible to tell whether the same is true of the other cases of w/y variation.

In a number of other cases, Safaitic and Hismaic have /-y/ where Classical Arabic has /-a:/ or /-a:ʔ/, thus Safaitic *s<sup>l</sup>my* “sky, clouds,” as against Arabic *samāʔ*, or Safaitic and Hismaic *bny* “he built” and *byt* “he spent the night,” as against Arabic *banā* and *bāta*. In some of these cases, there is evidence that Dumaitic and Thamudic B agreed with Arabic. Thus, the divine name *tr-s<sup>l</sup>m*, which occurs in Dumaitic and Thamudic B texts and in which *s<sup>l</sup>m* is the word for “heaven,” implies a pronunciation \**s<sup>l</sup>amā* (in which the /-a:/ would not appear in the consonantal script), as opposed to Safaitic *s<sup>l</sup>my* (\**s<sup>l</sup>umiyy* ?), see Macdonald *et al.* 1996:479–480.

Conversely, there are some words in which final /-a:/ is written with a -y in Arabic, but which in Ancient North Arabian did not end in consonantal /y/. These are most notably the prepositions which in Safaitic, Hismaic, and Thamudic B appear as *ʔ* (cf. Arabic *ʔilā*) “towards, for,” and *ʔl* (cf. Arabic *ʔalā*) “on, over, against.” In Dadanitic, both *ʔl* and *ʔly* are found, though the former is more common. This implies that the final sound may have been a diphthong -ay (/ai/), which would have been left unwritten in all the Ancient North Arabian scripts, except Dadanitic (see §2), where it would appear as -y (*pace* Drewes 1985, who believes diphthongs had been monophthongized in Dadanitic and that final -y represented [e:]). The forms without -y in Dadanitic may then represent either an uncertainty about writing diphthongs or a pronunciation with a final short vowel, as in some modern Arabic dialects (i.e., \*ai > \*ā (as in Classical Arabic) > \*a).

### 3.1.5 Nasal assimilation

As in Hebrew and Aramaic, but in contrast to Arabic, vowelless /n/ is frequently assimilated in most Ancient North Arabian dialects. This is particularly common in Safaitic and Hismaic where, for example, *mn* (cf. Arabic *min*) “from” and *mn* (cf. Arabic *man*) “whoever” are sporadically reduced to *m* (though curiously not in *mn ngd* “from high ground,” CSNS 381). Thus, the plural of *nfs<sup>l</sup>t* (“funerary monument”) sometimes appears as *ʔs<sup>l</sup>* (< \*ʔanfus<sup>l</sup>), and the verb \**intazar* (“to wait for”) always appears as *tʔr* (= \*ittazar?). Similarly, in Taymanitic, Thamudic B, Hismaic, and Safaitic (though rarely in Dadanitic), *bnt* (“daughter”) is occasionally spelled *bt*. However, this feature has not yet been identified in Hasaitic, where we find *bnt* (*passim*) and *ʔntt*, “wife” (CIH 984a) compare Dadanitic and Thamudic B *ʔt*, though



the corpus of Hasaitic texts is as yet so small that no firm conclusions can be drawn from this.

Assimilation of vowelless /n/ would also account for a feature characteristic of Taymanitic, that is the reduction of *bn* to *b* (“son of”) in genealogies, which contrasts with *bn* (= \**banī*, lit. “the sons of”) where the /n/ is followed by a vowel (Macdonald 1992a:31).

### 3.2 Vowels

Little of substance can be said about the vowels of Ancient North Arabian. The vowel inventory is assumed to have consisted of both short and long /a/, /i/, and /u/, but there is no evidence for or against this, except for final /a:/ and /u:/ in Dadanitic (see §2). Attempts to show that the diphthongs /au/ and /ai/ had been monophthongized to /o:/ and /e:/ respectively (as in many spoken Arabic dialects) are not convincing, though they cannot entirely be refuted either (see, again, §2).

## 4. MORPHOLOGY

Since Safaitic and Dadanitic are by far the best attested of the Ancient North Arabian dialects, the morphological descriptions below will concentrate on them, with information from the others when it is available.

It should be noted that several unusual forms have been attributed to Dadanitic on the basis of their apparent occurrence in JSLih 71 (= CLL 91). However, it is now recognized that, with the exception of the article *hn-* in the tribal name, the language of this text is Old Arabic, not Dadanitic. See Beeston *et al.* 1973:69–70 and Macdonald 2000:52–53 and forthcoming.

As in all Semitic languages, the morphology of the Ancient North Arabian dialects is based on the trilateral root, found in its simplest form in the third singular masculine of the suffix-conjugation (often known as the “perfect”).

The fact that, in most dialects of Ancient North Arabian, final -y is written in words such as *bny* “he built,” *s<sup>l</sup>my* “sky, clouds” and the gentilic ending (e.g., Safaitic *h-nbty* “the Nabataean” which in Arabic would be *al-nabaṭī*) suggests the presence of final short vowels, since without them the /-y/ would have become a long vowel [i:] or a diphthong [ai], and would not then have been represented in the orthography of any of the scripts, except in the case of the diphthong, that of Dadanitic. By contrast, the tiny amount of evidence available suggests that final short vowels may not have been present in the forms of Old Arabic represented in the documents so far identified (see Macdonald, forthcoming).

### 4.1 Nominal morphology

Nouns, adjectives, and pronouns will be discussed in this section. The purely consonantal Ancient North Arabian scripts must often conceal distinctions of number and possibly of case which would have been marked by changes in vowels. As in Arabic, the endings of nouns and adjectives can vary according to whether they stand alone (“in pause,” “pausal forms”) or are annexed to another noun or to an enclitic pronoun (“in construct”), see §5.1.3 below and Ch. 6, §3.3.2.1.

#### 4.1.1 Gender

The normal feminine singular ending in all Ancient North Arabian dialects is -t (even in pause; see §3.1.1): for example, *mr<sup>t</sup>* “woman,” Dadanitic (JSLih 64/2); *frs<sup>t</sup>* “mare,”

Thamudic B (e.g., HU 494); *bkr̥t* “young she-camel,” Safaitic (e.g., WH 344). Participles (see §4.2.6) are also marked for gender, and the feminine singular takes the *-t* ending of the nominals, as in *rg̥mt* (\**raḡīmat*) “humbled” (fem.), Safaitic (NST 2).

The word \**ym* “day” (attested only in the dual *ymn* and the plural *ʔym*) appears to have been treated as feminine in Dadanitic and Safaitic, as it is in Jibbālī and Mehri, though it is masculine in most other Semitic languages (see §4.4.1).

### 4.1.2 Number

Nominals in Ancient North Arabian have three numbers, singular (unmarked), dual and plural. On “external” (§4.1.2.2) and “internal” (§4.1.2.3) plurals in Semitic, see Chapter 6, §3.3.2.4.

#### 4.1.2.1 Dual

Clear evidence of the dual is found only in Dadanitic, Thamudic B, and Safaitic.

“In pause” (see §4.1), the normal ending of the dual is *-n* (cf. Classical Arabic *-āni*): for example, Dadanitic *h-mṭbr-n* “the two tomb-chambers” (JSLih 45/3); Thamudic B, *h-gml-n* “the two camels” (HU 296/2); Safaitic, *h-bkr̥t-n* “the two young she-camels” (e.g., WH 402, beside a drawing of them), *ym-n* “two days” (CSNS 796 and see p. iii).

A curious, and as yet unexplained, form of the dual in pause is found in one Safaitic text (LP 305), where *ḍll-y* “lost” (i.e., “dead”) refers to two people and is contrasted with *ḍll-n*, referring to three, in the same text (see §4.1.2.2). *Ḍll-y* is similar to the form of the dual which, in Classical Arabic, would be used in the oblique case “in construct” (see §4.1), namely *ḍalīlay*. However, in LP 305, while it would be in the oblique case (if this existed in Safaitic), it is clearly in pause and one would anyway not expect *y* to be used to represent a diphthong in the Safaitic script.

In Classical Arabic the *-n* of the dual is dropped in construct, leaving a long vowel (*-ā*), in the nominative, or a diphthong (*-ay*) in the oblique case. In Dadanitic, the only dialect with an orthography that represents some *final* long vowels and diphthongs, the ending seems to be a diphthong, represented by *-y*, regardless of case (if, indeed, this existed); thus, “nominative” *kbry s<sup>2</sup> t h-n/šj*, “the two kabirs of the company of H-NŠ” (JSLih 72/3–4; cf. Arabic *kabīrā*); “oblique” *b-ḥqwy kfr*, “on two sides of a tomb” (JSLih 77/7; cf. Arabic *ḥaqway*). As yet, there are not enough examples to assess the significance of this. Compare the situation in the modern spoken Arabic dialects where the dual ending in nouns is always *-ē(n)* (presumably <\**ay(n)*) regardless of whether the noun is grammatically in the “nominative” or “oblique” case. Again, this is a feature found in the early Arabic papyri (see Hopkins 1984:98–104).

When the second element of the construct was a pronominal suffix, the diphthong (\**-ay*) was considered to be *medial* and was therefore not represented in the Dadanitic script. The result is that the form *ḥw-hm* (JSLih 79/3) could represent either the dual “their two brothers” (\**ḥaway-hum*, cf. Classical Arabic *ḥawā-hum*, since the context requires it to be in the nominative) or the plural “their brothers” (cf. Classical Arabic *ḥuwwu-hum*).

A similar problem is found in Safaitic, where one of the few examples of the dual in construct yet identified is *ḥw-h* “his two brothers” (see LP 386, where the two persons are named). However, in C 657 *ḥw-h* is followed by the names of three persons, and in the other examples the numbers are not specified. It therefore appears that the form *ḥw* in Safaitic probably represents both the dual (\**ḥaway*) and the plural (\**ḥuwwu*) as in Dadanitic. The supposed plural *ḥwn* (in C 2534, 2779, 2955, cf. Arabic *ḥwān*) should almost certainly be read *ḥwl* (plural of *hl* “maternal uncle”).

The form *bny-h* in Safaitic has also been regarded as a possible dual (e.g., in C 3365, WH 1249, 3838, cf. Arabic *ibnay-hi* “his two sons,” oblique case). However, since Safaitic orthography does not show diphthongs, it is more likely that *bny-h* represents a diminutive (cf. Arabic *bunayyi-hi*, “his little son”), as it must do in C 4076, where it refers to only one person.

#### 4.1.2.2 External masculine plural

In pause this is formed by adding *-n* to the singular and is thus indistinguishable in the purely consonantal script from the regular form of the dual in pause. In construct the *-n* is dropped:

##### (3) A. Dadanitic

In pause      *ʿšdqn* “rightful heirs[?]” (CLL 65/2)

In construct    *bnw s<sup>l</sup>d<sup>l</sup>* “the sons of S<sup>l</sup>d<sup>l</sup>” (AH 1/2–3, see Sima 1999:35–36)

##### B. Safaitic

In pause      *zbyn* “male gazelles” (CSNS 550 beside a drawing of six, cf. Ar. *zabyān*)

Participles (see §4.2.6) are similarly marked: thus, *ḏll-n* “lost” (i.e., “dead” in LP 305, referring to three people, cf. Arabic, oblique case, *ḏalīlīn*).

#### 4.1.2.3 Internal masculine plural

In Arabic, this type of plural is often marked by changes in vowels within the word, and such changes would be invisible in the Ancient North Arabian consonantal scripts. Still, a few types have forms which show up even in the Ancient North Arabian orthographies, such as the following:

##### (4) Pattern Dadanitic

ʾaf ʾāl      *ʾym* (sg. \**ym*, “day,” e.g., JSLih 68/4, 349, cf. Ar. *ʾayyām*)

*ʾzll* (sg. *zll*, “zll-ceremony,” U 43, 115, etc. see Sima 1999: 95–96)

*ʾzl* (sg. *zll*, “zll-ceremony,” U 50/3)

ʾaf ʾilat      *ʾzlt* (sg. *zll*, “zll-ceremony,” U 32/3–4 and see Wright 1896–1898: i, 212)

fi ʾlat      *ʾzlt* (sg. *zll*, “zll-ceremony,” U 13/3, and see Stiehl 1971:6 and cf. Wright 1896–1898: i, 209, XII/4 for the form)

fu ʾāl      *hgg* (sg. \**hg*, “pilgrim,” JSLih 6/4, cf. Ar. *ḥuḡḡāḡ*)

Note also Dadanitic *ʾhw-hm* (“their brothers,” JSLih 79/3, \**ʾuḥuw* as in Safaitic, see §4.1.2.1).

##### Pattern Safaitic

ʾaf ʾāl      *ʾs<sup>2</sup>y<sup>c</sup>* (sg. \**s<sup>2</sup>c*, “companion,” cf. Ar. *ʾašyāʿ*)

*ʾhwl* (sg. *hl*, “maternal uncle,” e.g., HCH 71, cf. Ar. *ʾahwāl*)

fu ʾūl      *ḥtt* (sg. \**ht*, “line, carving,” cf. Ar. *ḥuṭūṭ*)

Note also Safaitic *ʾhw-h*, see §4.1.2.1.

#### 4.1.2.4 External feminine plural

This is *-t*, and so is identical in appearance to the singular (see §4.1.1), the change presumably lying in the vowel of the ending (cf. Arabic sg. *-ah/at*; pl. *-āt*); thus Safaitic *zbyt* “female gazelles” (WH 3373, the plural confirmed by the accompanying drawing); and Hismaic *nʾrt* “girls” (unpublished).

#### 4.1.2.5 Collective nouns

These are represented in Safaitic by *ʔbl* (“camels,” cf. Arabic *ʔibil*), and *mʕzy* (“goats,” cf. Arabic *miʕzan*). It is not clear whether they are grammatically feminine, as in Classical Arabic.

#### 4.1.3 Case

Since the Safaitic script shows no vowels, it is impossible to be certain whether case endings existed. However, by the same token, the spelling of such nouns as *mʕzy*, *ʔby*, and the gentilic (see §4.1.6) – for example, *h-yhdy*, “the Jew” (which in Arabic would be *al-yahūdī*) – imply that the final *-y* was pronounced with a short vowel, since, if it were not, it would itself become a long vowel and so would not be shown. Beyond this, little can be said with certainty at present. The same applies to Dadanitic.

#### 4.1.4 State

Caskel argued that the expression *h-ʔʃlmn* (CLL 19/3–4 = JSLih 62/3–4) indicates that, at an early period, a *determinate state*, marked by a suffixed *-n*, existed in Dadanitic, as in the Ancient South Arabian languages (1954:68). However, such an explanation would mean that the word was doubly defined (with a prefixed article *h-* and the suffixed *-n*), and Caskel’s attempt to explain the former as a demonstrative is unconvincing in view of the fact that elsewhere in Dadanitic the demonstrative adjective always follows the defined noun, thus *h-ʔʃlmn hḏh* (JSLih 82/1). It is much more likely that *ʔlmn* is a dual or an external plural, or perhaps a diminutive (see Brockelmann 1908–1913: i, 394), with a specialized meaning such as “statuette” as opposed to “statue” (cf. Aramaic *ʔlmnyṯ* which seems to mean “small female idols” in *Israel Exploration Journal* 29 (1979), p. 119).

#### 4.1.5 Determination

There is no visible mark of indetermination (comparable to *tanwīn* in Arabic), and had *tanwīn* been present it would have been represented in the Ancient North Arabian scripts. Determination is marked by the definite article (see §4.3.1) or annexation either to another noun or to a pronominal suffix.

#### 4.1.6 Diminutives

If diminutives were formed in Ancient North Arabian in the same way as in Arabic, by use of the *fuʕayl* form, they would be invisible in the Ancient North Arabian orthographies. Only exceptional forms such as *ʔhyt* (cf. Arabic *ʔuhayyat* “little sister”, C 893) and *bny* (cf. Arabic *bunayy* “little son”, WH1249) can be identified.

#### 4.1.7 Adjectives

These follow the noun and agree with it in gender, number, and determination: for example, in Safaitic *h-gs<sup>2</sup> h-rdf* (\**ha-gays<sup>2</sup> h-radīf*) “the rear guard” (LP 146); or *kll ʕ<sup>2</sup>r ʕdq* “every true kinsman” in Safaitic (HCH 191) and Hismaic (MNM 6).

As in Arabic, an adjective referring to a noun in the plural signifying nonsentient beings is put in the feminine singular, thus *rtg ʔqʔds<sup>l</sup>t* (cf. Arabic *rutuḡ qadīṣah*) “sacred portals” (CLL 85/3).

A gentile adjective (Arabic *nisbah*) is formed with *-y*: for example, *h-rmy*, “the Roman.” For demonstrative adjectives, see §4.1.8.4.

#### 4.1.8 Pronouns

Independent and enclitic personal pronouns are attested in Ancient North Arabian, as are relative and demonstrative pronouns.

##### 4.1.8.1 Independent personal pronouns

Only three independent personal pronouns are so far securely attested in Ancient North Arabian:

1. First singular *ʾn*: There is only one certain example in each of Safaitic (WH 1403b) and Dadanitic (JSLih 347/2). It is found occasionally in Hismaic (unpublished) and Thamudic D (e.g., JSTham 637), and is frequent in Thamudic B and C. It has not yet been found in Hasaitic.
2. Second singular *ʾt*: two possible examples are known so far, both in Thamudic B (HU 796 and 627?).
3. Third plural masculine *hm*: known from only one example in Dadanitic (JSLih 79/3).

##### 4.1.8.2 Enclitic personal pronouns

Enclitic personal pronouns can be attached to verbs representing the object (e.g., *qtl-h* “he killed him”) or to nouns indicating possession (e.g., *ʾb-h* “his father”) or to prepositions which govern them (e.g., *l-h* “for him”). Those so far attested on verbs in Ancient North Arabian are shown in 1 through 4.

1. First singular or plural *-n*: If the enclitic pronouns of the first persons singular and plural on verbs were similar to those in Classical Arabic (i.e., *-nī* = “me,” *-nā* = “us”) they would be indistinguishable in all Ancient North Arabian scripts except Dadanitic, where no certain example of either has yet been found. Thus, in Safaitic *ʾwd-n* “protect me/us” (unpublished); in Hismaic *dkrt-n lt* “may Lt be mindful of me/us” (unpublished); and in Thamudic B, where it is best attested, as in *flṭ-n* “deliver me/us” (LP 495).
2. Third singular masculine or feminine *-h*: This occurs in Dadanitic: for example, *rd-h w-sʾ d-h* “favor him and help him” (e.g., U 4/4); *rd-h w-ʾhrt-h w sʾ d-h* “favor her and her descendants and help her” (U 6/4–5). It is surprisingly rare in Taymanitic and Thamudic B, C, and D, but is found in both Safaitic – thus *yʾwr-h* “he will scratch it out” (e.g., LP 329), *qtl-h* “he killed him” (LP 385, etc.); and in Hismaic: for example, *hṭṭ-h* “he inscribed it” (JSTham 665).
3. Third dual *-hmy*: Several examples are found in Dadanitic, such as *sʾ d-hmy* “help both of them” (U 69/5–6). This presumably represents a diphthong \*-humay in contrast to Classical Arabic *-humā*.
4. Third plural *-hm*: This is found in Dadanitic *rd-hm* “favor them” (of four persons, AH 1/5 [see Sima 1999:35–36]).

On nouns and prepositions, the following enclitic personal pronouns are found:

5. First singular: If the enclitic pronoun of the first person singular was \*-ī on nouns and prepositions, as in Arabic and most Semitic languages, one would not expect it to show up in any of the Ancient North Arabian orthographies. However, there are a

handful of possible examples in Thamudic B: for example, *wdd-y* “my beloved” (HU 736), *s<sup>l</sup>m<sup>ʿ</sup> l-y* “listen to me” (HU 713). Since, the orthography of Thamudic B does not represent vowels in other cases, as far as we can tell, it would seem that the enclitic pronoun may have been pronounced \*īya or \*ayya, as when in Classical Arabic it is attached to a word ending in a long vowel, a diphthong, or *ʿalif maqṣūrah*.

6. Second singular *-k*: Safaitic *ʿwd-k* “your protection” (referring to one deity, unpublished) and Thamudic B *b-k* “in you” (e.g., HU 207, WTI 25, etc.) are attested. It is not yet identified in Dadanitic, Thamudic C and D, Hismaic, or Hasaitic.
7. Third singular masculine and feminine *-h*: This is common in Safaitic *ʿb-h* “his father” (e.g., WH 1275), *l-h* “for him” (e.g., WH 3420), “for her” (e.g., CSNS 412). The frequent omission of the definite article *h*- immediately after the third singular enclitic personal pronoun (e.g., *l-h rgm* “the cairn is his/hers,” as in the examples above) suggests that the suffix may have been pronounced \*-uh (masc.) / \*-āh (fem.), as in many Arabic dialects, rather than \*-hu (masc.) / \*-hā (fem.), as in Classical Arabic. The /h/ of the article may have been assimilated to that of the enclitic pronoun, leaving only its vowel and the possible reinforcement of the initial consonant of the following word (see §4.3.1), thus \*l-uh ha-(r)rugm > \*l-uh-a-(r)rugm “the cairn is his.” See also *s<sup>l</sup>d-h-rḏw* for \*s<sup>l</sup>ḏ-h h rḏw “help him O Rḏw” (CSNS 2), though this could also represent an optative perfect *s<sup>l</sup>d-h rḏw* “may Rḏw help him.” In Hismaic we find *kl-l-h* “all of it” (unpublished), *b-h* “in it” (unpublished); and in Dadanitic *ml-h* “his winter crop” (e.g., U 35/5), “her winter crop” (U 6/3). In Hasaitic there is *ʿht-h* “her sister” (Ja 1046). The nature of the texts in Taymanitic and Thamudic B, C, and D means that no certain examples of this suffix have yet been identified.
8. Second dual *-km*: In Safaitic there is *ʿwd-km*, “your protection” (referring to two deities, unpublished); compare Classical Arabic *-kumā*.
9. Third dual *-hmy*: This is found only in Dadanitic: *tmrt-hmy* “their fruit-trees” (U 69/4); compare Classical Arabic *-humā*.
10. Third dual *-hm*: In Dadanitic there are also examples of *-hm* being used to refer to two people. This could represent a difference in orthography or in pronunciation, or could simply be the use of the plural instead of the dual (see §5.2). Thus *ml-hm* “their winter crop” (referring to a man and a woman, following a verb in the dual U 19/5); *ml-hm* (referring to two men but following a verb in the 3rd pl. masc., U 36/4). In contrast to Dadanitic (cf. 9), this is the form which would be expected in the Thamudic B and Safaitic orthographies which show neither vowels nor diphthongs. There is one possible example in Thamudic B, *{h-/g}ml-n kl-hm* “both the camels” (HU 160) and one in Safaitic, *ʿl-hm* “on account of both of them” (HCH 34, referring to two persons).
11. First plural *-n*: Safaitic provides *ʿlh-n* “our god” (C 2526), *l-n* “for us” (C 2840). Hismaic has *ʿs<sup>2</sup>y<sup>ʿ</sup>-n* “our companions” (unpublished); *wq<sup>ʿ</sup>-n* “our inscription” (MNM 6).
12. Third plural masculine *-hm*: Examples include Dadanitic *ʿhrt-hm* “their descendants” (referring to three persons, U 90/5); Thamudic B: *kl-hm* (?) “all of them” (HU 160); Safaitic *ʿh-hm* “their brother” (LP 413); Hismaic *kl-hm*, “all of them” (unpublished).
13. Third plural feminine: At present there is no certain evidence for this, though Caskel sought unconvincingly to restore one, *-[h]n*, in CLL 69/1, 2.

#### 4.1.8.3 Relative pronouns

1. *mn/m* “who, whoever”: Compare Arabic *man*. In Safaitic this relative pronoun occurs in the very common curse *ʿwr m(n) y<sup>ʿ</sup>wr* “blind whoever scratches out [the writing],”

and in Hismaic in the expression *kl mn yqry* “anyone who may read” (MNM 6). No certain example of *mn* has yet been found in the other dialects. There is no example in Ancient North Arabian of *mn* or *m* used as an interrogative pronoun, but this is probably due to the nature of the texts.

2. *mh* “which, that which”: So far this has been found only in Dadanitic: for example, *m{h} ḥd* “that which has been taken” (CLL 82/2–3); and *m-l-hm* “that which [belongs] to them” (U 19/5, where the three elements are treated as one unit and the \**ā* of *mh* is not shown by a *mater lectionis* since it is no longer in final position).
3. *d* “who, whoever, which, that which”: Compare the relative pronoun *dū* which was particularly characteristic of the early Arabic dialect of the tribe of Ṭayyi’ (Wright 1896–1898:i, 272–273; Kofler 1940–1942:259–260; Rabin 1951:203–205). In Safaitic, this relative pronoun has so far been found only with reference to people, thus in the very common *ʿwr d yʿwr h-sʿlfr* “blind whoever scratches out the writing,” or *ʿyr m-d qtl-h* “recompense from him who killed him” (LP 385). In Dadanitic, however, *d*- is found referring to both people and things. Thus, *d-kn l-hm b-bdr* “that which [belongs] to them at Bdr” (U 73/4–5) which parallels *m-kn l-h b-dt* “that which [belongs] to him at Dt” (U 59/3–4). There are as yet no certain occurrences in the other dialects.
4. *d* followed by the name of a social group is the normal way of expressing group affiliation in Dadanitic (cf. 5), as in South Arabian (e.g., AH 1/1–3 [see Sima 1999: 35–36]; N w-N w-N w-N *bnw N d-N*.Trib., see also JSLih 197/2, 216/2).
5. *d ʿl*: This phrase is used as one of three ways of expressing affiliation to a social group in Safaitic and is the only method used in Hismaic and Hasaitic. There is no certain example of *d ʿl* in Dadanitic, where *d*- plus the ethnicon is the norm (cf. 4, the apparent example in AH 19/2 [= U 47/2] has been reread from the photograph as *d ʿlh* and interpreted as an error for *d ʿhl* (?) in Sima 1999:19, 84–85). It is not found at all in Taymanitic, where *ʿl* is simply placed after the last name in the genealogy (see Macdonald 1992a:31, 40, n. 74). There is also no certain example in any of the types of Thamudic. The phrase *d ʿl* is made up of a particle *d* + *ʿl*, a noun meaning any social group from immediate family to nation (cf. Arabic *ʿāl*). It is placed before the name of the group, thus *d ʿl ḥzy* “of the lineage of Ḥzy.” The masculine *d* seems to have been considered an inseparable particle, since in texts employing word-dividers it is always attached to *ʿl*, in contrast to the feminine *dʿt*, which is always separated from *ʿl*. The feminine, *dʿt ʿl*, is found in Safaitic (e.g., CSNS 412), Hismaic (unpublished), and Hasaitic (e.g., *Atlal* 6, 1982:139, lines 6–7). Here the *ʿ* is consonantal, in contrast to Classical Arabic *dāt* (perhaps < \**dāʿt* [?]; cf. the Hebrew feminine demonstrative *zōʿt* < \**zāʿt*?). A possible plural is found in Safaitic *dww ʿl yzr* “members of the ʿl Yzr” (C 2156); compare Classical Arabic *dawū*. Littmann (1943:xvi) compared this particle *d* to Classical Arabic *dū* “possessor of” (< “he of . . .”?). This is probably also the case with *d* (without *ʿl*) in Dadanitic (see 4). The exact relationship of this particle to the relative and demonstrative pronouns (§4.1.8.4) is not yet clear.

#### 4.1.8.4 Demonstrative pronouns

A demonstrative pronoun, *zn* (or perhaps *dñ*) is found in Thamudic D (*zn N*, “this is N”) and is used for both masculine and feminine: thus *zn ḡnm bn ʿbdmnt* “this is Ḡnm son of ʿbdmnt” (JSTham 584); and *zn rqs<sup>2</sup> bnt ʿbdmnt* “this is Rqs<sup>2</sup> daughter of ʿbdmnt” (JSTham 1, and another example in 219). It has been suggested that another demonstrative pronoun, *zt*, is attested in Thamudic C, but this is highly questionable. No demonstratives have yet been identified in Taymanitic or Thamudic B.



The only evidence at present for a demonstrative pronoun in Dadanitic is the adverb *b-dh* “here”, literally “in this”, (Jshih 279). Caskel (1954:64) suggested that some Dadanitic inscriptions begin with a demonstrative pronoun *d*, “this”: for example, *d / ms<sup>l</sup>lmh* “this is Ms<sup>l</sup>lmh” (CLL 102); *d lm jkl lt* “this is <sup>l</sup>lm priest of Lt” (CLL 104). However, the *d*-sign at the beginning of these graffiti is almost certainly an apotropaic sign (perhaps *d* for the deity *d-ḡbt*); see JSLih 284, where it occurs at the beginning and the end of the text and 297, where these signs are excluded from the cartouche around the name.

## 4.2 Verbal morphology

The different dialects of Ancient North Arabian contribute fragmentary evidence on verbal inflection for three persons (first, second, and third), three numbers (singular, dual, and plural) and two genders (masculine and feminine), at least in the third-person singular in which the vast majority of these inscriptions are couched. The various verb-stems (see §4.2.2) are inflected in two conjugations – one suffixed, the other prefixed (see §4.2.3). The verb appears in active and passive voice, though the morphology of the latter is difficult to identify, as discussed in §4.2.4. In a similar fashion, modal distinctions are obscured by the orthography; see §4.2.5.

A notable difference between Arabic and Ancient North Arabian lies in the treatment of verbs in which the third radical is /w/ or /y/. In Arabic, even in the pre-Islamic period, verbs of the form \*šatawa (“to pass the winter”) and \*banaya (“to build”) appear to have been contracted to \*šatā and \*banā respectively, since in purely consonantal scripts (e.g., Sabaic) they appear with no final radical (e.g., *bn* for \*banā in the ‘Igl bn Hf‘m inscription from Qaryat al-Faw, see Beeston 1979b:1–2) and in those which use *matres lectionis* (e.g., Nabataean) they appear with final -’ (= -ā). However, in Ancient North Arabian the third radical is always retained, thus *s<sup>2</sup>tw* (more commonly *s<sup>2</sup>ty*, see above) and *bny* (see Macdonald, forthcoming).

This feature is also found in verbs which have a middle radical /w/ or /y/. In Classical Arabic, this is commonly reduced to -ā- when between two short vowels: for example, \*ḥawara > ḥāra, and \*bayata > bāta. But in Safaitic, these verbs are written with the middle radical intact, both in the base stem (cf. Arabic Form I), for example *ḥwr* “he returned,” *byt* “he spent the night,” etc.; and in the ʔ-prefix stem (cf. Arabic Form IV), for example, ʔwr “he blinded in one eye” (MSTJ 11, cf. Arabic ʔaʔara but also ʔaʔwara). It has been suggested that verbs of this type are sometimes found in a contracted form in the base stem (e.g., Safaitic *šf* [supposedly representing \*šāfa] for *šyf* “he spent the early summer”), and that the forms with medial *w* or *y* represent the equivalent of the Arabic Forms II (faʔala) or III (fāʔala), where the middle radical has a consonantal value (for Dadanitic, Caskel 1954:67; for Safaitic, Littmann 1943:xvii–xviii). However, the only plausible case of such contraction yet identified in an Ancient North Arabian text is *kn* (cf. Arabic *kāna* “he/it exists”) in the Dadanitic phrase *d kn-l-h* “that which is to him” (i.e., “is his,” e.g., in U 85/3). In most cases, the sense requires the verb written with medial *w/y* to be the equivalent of Classical Arabic Form I rather than Forms II or III, though it should be noted that in most modern Arabic dialects forms I and II of many verbs are used interchangeably with little discernible difference in meaning (I am most grateful to Professor Clive Holes for this information).

There appears to be an interesting difference between Safaitic and Hismaic as regards verbs which (in Arabic) have ʔ as their third radical. Thus, *yqr* “he may read” (C 4803) in Safaitic (and Classical Arabic) as against *yqry* in Hismaic (MNM 6). On this root’s significance for



the etymology of Classical Arabic *qara'a* (meaning “to read”) in Ancient North Arabian, see Macdonald, forthcoming. See also Safaitic *ks<sup>l</sup>*’ “a track” (C 523, cf. Arabic *kus*’ “rear, behind”) as against Hismaic *ks<sup>y</sup>*, “pursuing” (unpublished, cf. Arabic *kas*’). It is also possible that this *y* contrast is sometimes found in medial position. In one Hismaic text (CTSS 3) we find *dyl* for *d<sup>y</sup>l*, the normal marker of affiliation to an ethnic or social group. However, this example is so far unique, and elsewhere in Hismaic we find *d<sup>y</sup>l*, as in Safaitic. All in all, there are at present too few examples of this apparent *y* contrast to be sure that it is really a dialectal feature.

In certain cases, Safaitic has a geminate verb where the equivalent in Classical Arabic has *w* or *y* as the third radical. Thus Safaitic *gzz* “to raid” as against Arabic *gazā* (root *g-z-w*, see Beeston 1979a:134).

#### 4.2.1 Verb patterns

Arabic grammar knows fifteen possible *forms* or *patterns* of the verb (conventionally illustrated by the verb *fa‘ala*), of which only the first ten are common. Several of these are distinguished by vowel lengthening or by doubling of the second or third radical. Since vowels and doubled consonants are not expressed in the Ancient North Arabian scripts (apart from some final long vowels in Dadanitic which are irrelevant in this case), it would be impossible to distinguish between the equivalents of Arabic Forms I (*fa‘ala*), II (*fa‘āla*), and III (*fā‘ala*), all of which would appear simply as *\*fʿl*, except possibly in the case of geminate verbs (see below). Similarly, V (*tafa‘ala*) and VI (*tafā‘ala*) would both appear as *\*tfʿl*. This means that there is no way of telling whether Ancient North Arabian had a structure of verbal Forms similar to that of Classical Arabic. It therefore seems more prudent to describe the stems simply by the ways in which they appear in the texts.

It might be thought that the geminate verbs would be an exception to the above, since one would expect the equivalent of the Arabic Form I to appear as *hl* (*\*ḥalla*), and the equivalent of the Arabic Form II to appear as *hll* (*\*ḥallala*). However, the *hl* form is rare in Safaitic and is always found in exactly the same contexts as *hll* with no apparent difference in sense between the two. Similarly, the verb *wdd* “he loved,” which is very common in Thamudic B, is rarely, if ever, found as *wd*. In Dadanitic, there is no clear example of the *hl* form in the base stem, though there is considerable variation in the *y*-prefix stem, namely: *ʔll* (U 14/2, etc.) as against *ʔl* (U 18/2, etc.); *ʔllt* (U 68/4, etc.) as against *ʔlt* (U 6/2, etc.); *ʔllw* (U 119/5, etc.) as against *ʔlw* (U 90/3, etc.) – where Arabic would have *ʔazalla*, *ʔazallat*, *ʔazallū*, respectively. Similarly, in Dadanitic, the active participle *ʔr* (HE 1) implies a pronunciation such as *\*ʔārir*, in contrast to Arabic *ʔarr*. This suggests that in most contexts the second and third radicals of geminate verbs were separated by a vowel in Ancient North Arabian (at least in the pronunciation of some speakers), thus *\*ḥālal*, *\*ʔārir*, *\*ʔāzāl*, and so forth, in contrast to Classical Arabic where they were not, thus *ḥalla*, *ʔarr*, *ʔazall*. These verbs cannot therefore be used as evidence of a *fa‘ala* (Form II) in Ancient North Arabian.

#### 4.2.2 Verb-stems

Before presenting the Ancient North Arabian verb-stems, three things must be noted. First, because in Arabic, verbs which contain one or more of the phonemes */ʔ/*, */w/*, or */y/* behave somewhat differently from those which do not, examples of such verbs in Ancient North Arabian are listed below with the form of the cognate verb in Classical Arabic given for comparison. Second, reconstructions of the vocalized and unassimilated forms of Ancient

North Arabian verbs are purely hypothetical and are based on the equivalent forms in Classical Arabic. They represent only one of several possible realizations of the forms found in the texts, and should not be taken as anything more than a working hypothesis. Finally, references to texts are usually given only for unique or unusual occurrences.

#### 4.2.2.1 Safaitic verb-stems

##### (5) Base Stem fʿl (cf. Arabic Forms I, II, and III)

Radicals	Safaitic	cf. Arabic
	<i>dbh</i> “he sacrificed”	<i>dabaḥa</i>
I = ʾ	<i>ḥd</i> “he took possession of”	<i>ʾahada</i>
I = ʾ, III = y	<i>ʾty</i> “he came”	<i>ʾatā</i>
I = w	<i>wgm</i> “he grieved”	<i>waḡama</i>
I = y, II = ʾ	<i>yʾs<sup>l</sup></i> “he despaired” (SIJ 118)	<i>yaʾisa</i>
II = w	<i>ḥwr</i> “he returned”	<i>ḥāra</i>
II = w, III = y	<i>nwy</i> “he migrated with the whole tribe”	<i>nawā</i>
II = y	<i>byt</i> “he spent the night”	<i>bāta</i>
III = ʾ	<i>dt<sup>ʾ</sup></i> “he spent the season of the later rains”	
III = w	<i>s<sup>2</sup>tw</i> “he spent the winter”	<i>ṣatā</i>
III = y	<i>bny</i> “he built”	<i>banā</i>
II = III	<i>ḥll</i> “he camped”	<i>ḥalla</i>

Three derived stems can be identified in Safaitic: (i) the ʾ-prefix (ʿfʿl) stem (cf. Arabic Form IV *ʾafʿala*); (ii) the *t*-prefix (tfʿl) stem (cf. Arabic Forms V *tafaʿala* and VI *tafāʿala*); and (iii) the *t*-infix (ftʿl) stem (cf. Arabic Form VIII *iftaʿala*). These are illustrated below.

##### (6) ʾ-prefix stem ʿfʿl (cf. Arabic Form IV)

Radical	Safaitic	cf. Arabic
	<i>s<sup>2</sup>rq</i> “he migrated to the inner desert”	<i>ʾašraqa</i>
I = y, II = ʾ	<i>yʾs<sup>l</sup></i> “it drove to despair” (root y-ʾ-s <sup>l</sup> , WH 1022)	<i>ʾayʾasa</i>
II = w	<i>ʾwr</i> “he blinded in one eye” (root ʿ-w-r, MSTJ 11)	<i>ʾaʿāra</i> / <i>ʾaʿwara</i>
III = y	<i>ʾly</i> “he raised up” (root ʿ-l-y, WH 1696)	<i>ʾaʿlā</i>

Note that *yʾs<sup>l</sup>* presents a rare occasion when a diphthong may have been expressed in Safaitic (\*ʾayʾasa), unless a short vowel or, more likely, a *shewā* was inserted to ease the transition to the secondʿ.

Safaitic *t*-prefix stems are illustrated by the following:

##### (7) *t*-prefix stem tfʿl (cf. Arabic Forms V and VI)

Radical	Safaitic	cf. Arabic
I = n	<i>tnẓr</i> “he looked out for” (root n-z-r, WH 3294)	<i>tanazzara</i>
II = w	<i>ts<sup>2</sup>wq</i> “he longed for” (root *s <sup>2</sup> -w-q)	<i>tašawwaqa</i>

(8) *t*-infix stem ft'l (cf. Arabic Forms VIII)

Radical	Safaitic	cf. Arabic
	<i>qttl</i> "he died mad" (root q-t-l, MHES p. 286)	<i>iqtatāla</i>
I = n	<i>tʒr</i> "he waited" (root n-ʒ-r)	<i>intāzara</i>
I = y, II = ʾ	<i>tʾsʾ</i> "he despaired" (root y-ʾ-sʾ, LP 679)	<i>ittaʾasa</i>

On the assimilation of \**ntʒr* to *tʒr*, see §3.1.5.

4.2.2.2 *Dadanitic verb-stems*

The Dadanitic base stem can be illustrated by *ndr* "he vowed" (U 10/2). Examples of base stems with ʾ, w and y radicals and with geminate radicals are presented in (9):

## (9) Base stem f'l (cf. Arabic Forms I, II and III)

Radical	Dadanitic	cf. Arabic
	<i>ndr</i> "he vowed" (U 10/2)	<i>naḍara</i>
I = ʾ	<i>ḥd</i> "he took possession of" (JSLih 45/3)	ʾ <i>ahada</i>
I = ʾ, III = w	<i>ḡw</i> "he made provision for, attended to" (?) (U 71/2), see Müller in Stiehl 1971:566	
I = w, III = y	<i>wdy</i> "he erected" (?) (JSLih 40/5)	
II = w	<i>kn</i> "it is" (e.g., U 73/4)	<i>kāna</i>
III = y	<i>bny</i> "he built" (CLL 74/1)	<i>banā</i>
II = III	<i>ʾrr</i> "he dishonored" (HE 1/4–6)	<i>ʾarra</i>

Regarding *ḡw*, note, however, that Sima (1999: 93–94) takes this as an ʾ-stem of a verb \**ngw* which he interprets as "to clear out [an underground water channel]."

Dadanitic is the only Ancient North Arabian dialect in which there is clear evidence of a *h*-prefix stem (10) and even here it coexists with the ʾ-prefix (11) which is the norm in Safaitic. There are insufficient clear examples of verbs in the other dialects to draw any conclusions:

(10) *h*-prefix stem hf'l

Radical	Dadanitic
	<i>hmt</i> meaning uncertain (* <i>hamtaʾa</i> , root * <i>m-t-ʿ</i> , JSLih 7/3)
I = w	<i>hdqt</i> "she offered" (* <i>hawdaqat</i> , root * <i>w-d-q</i> , JSLih 62/3)
	<i>hwdqw</i> "they offered" (* <i>hawdaqū</i> , 3rd pl., JSLih 49/5–6)

The retention of the initial *w* of the root in *hwdqw* may reflect uncertainty about representing diphthongs in the Dadanitic script.

## (11) ʾ-prefix stem ʾf'l (cf. Arabic Form IV)

Radical	Dadanitic	cf. Arabic
I = w	<i>ʾdq</i> "he offered" (root * <i>w-d-q</i> , CLL 62/3)	ʾ <i>awdaqā</i>
I = w, III = y	<i>ʾfy</i> "he accomplished" (root * <i>w-f-y</i> , U 4/2)	ʾ <i>awfā</i>
II = III	<i>ʾzll</i> "he performed the <i>zll</i> -ceremony" (root * <i>ʾ-z-l-l</i> , e.g., U14/2)	ʾ <i>azalla</i>
	<i>ʾzl</i> "he performed the <i>zll</i> -ceremony" (root * <i>ʾ-z-l-l</i> , e.g., U 18/2)	

It is possible that *tqt* (e.g., in JSLih103) represents a ***t*-infix stem (ft'l)** in Dadanitic. Caskel interpreted this as a metathesized *t*-infix stem of *qʾt*, thus \**iqtaʾta* > \**itqaʾta* (CLL

p. 64). However, this is improbable. More likely it represents the *t*-infix stem of a root \*wqt (\*ittaqaṭa), or of a root \*nqt (\*intaqaṭa which, with the expected nasal assimilation (§3.1.5), would become \*ittaqaṭa).

Caskel sought to identify one verb with an *n*-prefix (equivalent to the Arabic Form VII) and another with a *st*-prefix (equivalent to the Arabic Form X), but in both cases the interpretations are very uncertain (Caskel 1954:64–65).

### 4.2.3 Verb conjugations

Two conjugations are identifiable in Ancient North Arabian, one in which person, number and gender are indicated by suffixes and one in which these are indicated by prefixes (and in some persons suffixes as well). If two prefix-conjugations existed, as in some Semitic and Hamitic languages, the Ancient North Arabian writing system, which shows neither vowels nor doubled consonants, has rendered them indistinguishable. On the uses of the suffix- and prefix-conjugations see §§5.3.1 and 5.3.2.

#### 4.2.3.1 Safaitic verb conjugations

Examples of those forms which are attested for the suffix-conjugation in Safaitic are listed in (12).

#### (12) The suffix-conjugation in Safaitic

Base stem			
Person	Radical	Safaitic	cf. Arabic
3rd sg. masc.		<i>dbḥ</i> “he sacrificed”	<i>ḏabaḥa</i>
	I = ’III = <i>y</i>	<i>’ty</i> “he came” (e.g., NST 3)	<i>’atā</i>
	II = <i>y</i>	<i>myt</i> “he died” (e.g., WH 387)	<i>māta</i>
	III = <i>y</i>	<i>r’y</i> “he pastured”	<i>ra’ā</i>
	II = III	<i>ḥl</i> “he camped” (Form I)	<i>ḥalla</i>
3rd sg. fem.		<i>ḥll</i> “he camped” (Form II)	<i>ḥallala</i>
		<i>gls<sup>1</sup> t</i> “she stopped briefly” (SIAM i 30)	<i>ḡalasat</i>
	II = <i>y</i>	<i>mtt</i> “she died” (NST 2)	<i>mātat</i>
2nd sg. fem.		<i>whbt</i> “may you give” (C 4037, optative §5.3.1)	<i>wahabti</i>
’-prefix stem			
Person	Radical	Safaitic	cf. Arabic
3rd sg. masc.		<i>’s<sup>2</sup>rq</i> “he migrated to the inner desert”	<i>’ašraqa</i>
<i>t</i> -prefix stem			
Person	Radical	Safaitic	cf. Arabic
3rd sg. masc.	II = <i>w</i>	<i>ts<sup>2</sup>wq</i> “he longed for”	<i>tašawwaqa</i>
3rd sg. fem.		<i>ts<sup>2</sup>wqt</i> “she longed for”	<i>tašawwaqat</i>
<i>t</i> -infix stem			
Person	Radical	Safaitic	cf. Arabic
3rd sg. masc.	I = <i>n</i>	<i>tṣr</i> “he waited”	<i>intazara</i>

The terminations of the dual, if it existed (cf. Dadanitic and Classical Arabic *-ā*) and the plural (cf. Dadanitic and Classical Arabic *-ū*) of the suffix conjugation are not visible in Safaitic orthography.

Examples of those forms which are attested for the prefix-conjugation in Safaitic are listed in (13).

(13) The prefix-conjugation in Safaitic

Base stem			
Person	Radical	Safaitic	cf. Arabic
3rd sg. masc.		<i>yḥbl</i> “he may damage”	<i>yahḥbalu</i>
	I = <i>w</i>	<i>yʿwr</i> “he may scratch out”	<i>yaʿūru</i>
			<i>yuʿawwiru</i>
	III = ʾ	<i>yqrʾ</i> “he may read” (C 4803)	<i>yaqraʾu</i>
	III = <i>y</i>	<i>yqry</i> “he may read” (Hismaic, MNM 6)	
3rd pl. masc.	II = III	<i>yrb̄b</i> “he is training” (C 1186)	<i>yurabbibu</i>
	II = <i>w</i>	<i>yʿwrn</i> “they may scratch out” (WH 2112)	<i>yaʿūrūna</i>
			<i>yuʿawwirūna</i>
1st pl.	III = <i>y</i>	<i>nngy</i> “may we escape” (WH 135)	<i>nanḡū</i>
	II = III = <i>y</i>	<i>nḥyy</i> “may we live prosperously” (Thamudic B, LP 495)	<i>naḥyā</i>
ʾ-prefix stem			
Person	Radical	Safaitic	cf. Arabic
3rd sg. masc.		<i>ys<sup>2</sup>rq</i> (in <i>l-ys<sup>2</sup>rq</i> “in order to go into the inner desert”, LP 180)	<i>yušriq</i> (Jussive)
<i>t</i> -prefix stem			
Person	Radical	Safaitic	cf. Arabic
3rd sg. masc.	II = <i>n</i>	<i>ytzr</i> “he will wait for” (?) (WH 3929)	<i>yantaziru</i>

4.2.3.2 Dadanitic verb conjugations

(14) The suffix-conjugation in Dadanitic

Base stem			
Person	Radical	Dadanitic	cf. Arabic
3rd sg. masc.	I = ʾ	<i>ʾḥd</i> “he took possession of” (e.g., JSLih 45/3)	<i>ʾaḥada</i>
	I = ʾ, III = <i>w</i>	<i>ʾgw</i> “he made provision for” (?) (U 71/2) (see Müller in Stiehl 1971:566)	
	III = <i>y</i>	<i>bny</i> “he built” (CLL 74/1)	<i>banā</i>
	II = III	<i>ʾrr</i> “may he dishonor” (HE 1/4, see §5.3.1)	<i>ʾarra</i>
3rd sg. fem.		<i>nḍrt</i> “she vowed” (JSLih 73/4–5)	<i>naḍarat</i>
3rd pl. masc.	III = <i>y</i>	<i>bnt</i> “she built” (root b-n-y, CLL 90/3)	<i>banat</i>
	I = ʾ	<i>ʾḥdw</i> “they took possession of” (JSLih 79/2)	<i>ʾaḥadū</i>
	III = <i>y</i>	<i>bnyw</i> “they built” (CLL 26/2)	<i>banaw</i>

On this last, *bnyw*, compare the form *binyaw* (instead of Classical Arabic *banaw*) in some “old sedentary dialects” of eastern Arabia and many others in Saudi Arabia [Clive Holes].

**ʾ-prefix stem**

Person	Radical	Dadanitic	cf. Arabic
<b>3rd sg. masc.</b>	I = <i>w</i>	<i>ʾdq</i> “he offered” (?) (root w-d-q, CLL 62/3)	<i>ʾawdaqqa</i>
<b>3rd sg. fem.</b>	I = <i>w</i> , III = <i>y</i> II = III	<i>ʾft</i> “she accomplished” (root w-f-y, U 5/2) <i>ʾzllt</i> “she performed the ʒll-ceremony” (U 68/4) <i>ʾzlt</i> “she performed the ʒll-ceremony” (U 6/2)	<i>ʾawfat</i> <i>ʾazallat</i>
<b>3rd du. masc.</b>	II = III	<i>ʾzlh</i> “they two performed the ʒll-ceremony” (U 19/3, but see §5.2)	<i>ʾazallā</i>
<b>3rd pl. masc.</b>	II = III	<i>ʾzllw</i> “they performed the ʒll-ceremony” of four persons (AH 1/3-4, see Sima 1999:35–36)	<i>ʾazallū</i>

**h-prefix stem**

Person	Radical	Dadanitic
<b>3rd sg. masc.</b>		<i>hmtʿ</i> meaning uncertain (*hamtaʿa, root m-t-ʿ, CLL 39/3)
<b>3rd sg. fem.</b>	I = <i>w</i>	<i>hdqt</i> “she offered” (?) (*hawdaqat, root w-d-q, JSLih 62/3)
<b>3rd pl. masc.</b>	I = <i>w</i>	<i>hwdqw</i> “they offered” (?) (*hawdaqū, JSLih 49/5-6)

**t-infix stem**

Person	Radical	Dadanitic
<b>3rd sg. masc.</b>	I = <i>n</i> or <i>w</i>	<i>tqt</i> (*ittaqaṭa ? root n-q-ṭ or w-q-ṭ, e.g., CLL 6, JSLih 103)

## (15) The prefix-conjugation in Dadanitic

**Base stem**

Person	Radical	Dadanitic	cf. Arabic
<b>3rd sg. masc.</b>		<i>yqʿd</i> “it will remain” (?) (JSLih 40/4)	<i>yaqʿudu</i>

**4.2.4 Voice**

Since no short vowels are expressed in the Arabian consonantal scripts, it is impossible to tell whether the Ancient North Arabian verbal system had a fully operational passive voice, indicated by changes of internal short vowels, as in Arabic. Thus, *s<sup>l</sup>nt qtl mʿn* (LP 297) presumably means “the year Mʿn was killed,” but it is not clear whether *qtl* here is a verb in the passive of the suffix-conjugation (equivalent to Arabic *qutla*), or a *maṣdar*, or verbal noun (equivalent to Arabic *qutl*, i.e., “the year of Mʿn’s being killed”), or even a passive participle (cf. Arabic and Aramaic *qatīl*) acting as a verb to produce a virtual relative (i.e. “the year [in which] Mʿn [was] killed”), as, for example, in Nabataean (Cantineau 1930–1932:i, 108); see §5.4.

In Dadanitic, a verb in the passive can occasionally be identified. Thus, for instance, the context in CLL 82/3 requires *ʾhd* to be a third singular masculine passive of the suffix-conjugation in *m{h} ʾhd ʿl-hmy* “that which has been acquired on behalf of both of them.” A possible example of the passive of the prefix-conjugation is *lh yʿd* “he will not be threatened” (root \*w-ʿ-d, CLL 31/6, cf. Arabic *lā yūʿadu*).

**4.2.5 Mood**

Similarly, the fact that no short vowels are indicated in the scripts makes it impossible to tell whether there were indicative, subjunctive, and jussive moods in the prefix-conjugation, distinguished by final short vowels (or lack of them) as in Classical Arabic.

The absence of short vowels in the scripts also means that the imperative can only be identified from context, and there is no visible distinction between the masculine and

feminine forms. Thus, in Safaitic, for example, *flṭ* “deliver!” occurs in some contexts where it must be masculine (cf. Arabic *ifliṭ* [masc.]) and others where it must be feminine (cf. Arabic *ifliṭi* [fem.]); similarly with *ʿwr* “blind!” (masc. and fem.; cf. Arabic *ʿawwir* [masc.], *ʿawwiri* [fem.]).

In Dadanitic, many inscriptions end with invocatory formulas consisting of a series of verbs in the imperative or in the suffix-conjugation with an optative sense (see §5.3.1). The most common of these formulas is *f-rḏ-h w-s<sup>l</sup> ʿd-h w-ḥrt-h* “and so favour him and help him and his descendants” (see JSLih 8, where the deity is mentioned, and U 14/5–6, etc., where it is not; see Sima 1999:105 for the variants of this formula at al-ʿUḏayb). Here *rḏ* is the masculine imperative of *rḏy* “to favor” (equivalent to Arabic *irḏa*) whereas *s<sup>l</sup> ʿd* can be compared with the Arabic Form III imperative *sāʿid*.

In the case of verbs whose first radical is *w* there seems to be a distinction between Safaitic and Thamudic B, though the small number of examples is restricted to the verb *whb*, which in Classical Arabic is exceptional in this respect (see Wright 1896–1898:i, 78–79). We cannot therefore be certain how widespread a phenomenon this was. In Safaitic (in all but two examples), the initial *w* of *whb* is retained in the imperative, whereas in Thamudic B it seems to be dropped (as in Classical Arabic). Thus, in Safaitic we find *w-whb l-h nqmt* “and give to him booty” (C 1808, cf. Classical Arabic *hab*); and *h rḏw whb l-h* . . . “O Rḏw give to him . . .” (WH 190). On the other hand, there are two Safaitic texts in which the imperative appears as *hb*: *h rḏw hb l-ʿbdʿl nqmt* “O Rḏw give to ʿbdʿl booty” (LP 460) and *h ʿlt flṭ l-bgʿ w-hb l-h nʿm* “O ʿlt [grant] deliverance to Bgʿ and give to him prosperity” (LP 504), though in both cases this could be due to haplography, as it could be in the Thamudic B text *h rḏw hb s<sup>2</sup>km* “O Rḏw give a gift” (unpublished).

## 4.2.6 Participles

As a verbal noun, the participle in Ancient North Arabian was inflected according to gender, number, and voice. On the uses of the participle see §5.4.

### 4.2.6.1 Active participle

#### Base stem

	sg. masc.	<i>qtl</i> (cf. Ar. “ <i>qātil</i> ”): Safaitic, in <i>ḫr mn qtl-h</i> “revenge on his killer” (CSNS 1004);
	pl.	<i>qbrn</i> (cf. Ar. <i>qābirūna</i> ): Safaitic, in <i>qbrn ḏw ʿl yṣr</i> “members of the ʿl Yṣr having performed the burial” (C 2156), see §5.4;
II = <i>w</i>	sg. masc.	<i>mʿwr</i> (cf. Ar. <i>muʿawwir</i> ): Safaitic, in <i>ʿwr l-m ʿwr</i> “blindness to a scratcher-out” (WH 408, etc.)
II = <i>w</i> , III = <i>y</i>	sg. masc.	<i>nwy</i> (cf. Ar. <i>nāwin</i> ): Safaitic, in <i>rʿy h-nhl nwy</i> “he pastured this valley while on migration” (C 3181)
III = <i>y</i>	sg. masc.	<i>rʿy</i> (cf. Ar. <i>rāʿin</i> ): Safaitic, in <i>ṣyr rʿy ḥrt</i> “he was on his way to permanent water pasturing the <i>ḥarra</i> [basalt desert]” (C 3131)
II = III	sg. masc.	<i>ʿrr</i> (* <i>ʿarir</i> , cf. Ar. <i>ʿarr</i> ): Dadanitic, in <i>ʿrr ḏgbt ʿrr h-s<sup>1</sup> fr ḏh</i> “may <i>ḏgbt</i> [the chief deity of Liḥyān] dishonor him who dishonors [lit. ‘the dishonor of] this inscription” (HE 1)

#### 4.2.6.2 *Passive participle*

There appear to be two morphological types of passive participle in the base stem – the *faʿīl*-type and the *mafʿūl*-type. Safaitic singular and plural examples of each follow:

1. *The faʿīl*-type: Singular masculine *qtl* “killed” (e.g., LP 658; see §4.2.4); singular feminine (i.e., of the form *faʿīlat*) *trḥt* “untimely dead” (e.g., NST 2); plural masculine *ḥrbn* “plundered and left destitute” (C 657, *pace* ed.; cf. Arabic *ḥarībīn*, oblique case).
2. *The mafʿūl*-type: Singular masculine *mqtl* “killed, murdered” (e.g., HCH 76; cf. Arabic *maqṭūl*); plural masculine *mḥrbn* “plundered and left destitute” (HCH 71; cf. Arabic *maḥrūbīn*, oblique case).

In Dadanitic, the only clear participial form, *h-mqtl* (JSLih 40/9), is in a damaged context and could represent either an active participle (cf. Arabic *muqattil* “mass killer”) or a passive (cf. Arabic *maqṭūl* “killed, murder victim”). There are no certain cases in the other dialects.

The feminine, dual, and external masculine plural forms of participles are similar to those of other nouns; see §4.1.2.

### 4.3 *Particles*

#### 4.3.1 *The definite article*

The most obvious difference between the two branches of North Arabian lies in the form of the definite article. In Old and Classical Arabic and the majority of the vernaculars, it is *ʾal-*, while in Ancient North Arabian it is either *h-* (*hn-*) or in some dialects possibly zero. The earliest evidence for both comes from the fifth century BC in the epithet of a goddess which Herodotus (3.8) quotes in its Old Arabic form, Ἀλῖλάτ (\**ʾal-ʾilat*), and which occurs in its Ancient North Arabian form, *hn-ʾlt*, in a number of Aramaic inscriptions on silver bowls found at Tell al-Maskhūṭah in northeastern Egypt (Rabinowitz 1956). In both cases, it means literally “the goddess.”

A definite article has not yet been identified in Hasaitic (except in names) or in Thamudic C and D, and there are doubts whether Hismaic employed one at all (see below). In Taymanitic, Thamudic B, and Safaitic, it is *h-* in all contexts. Since the script shows neither vowels nor the doubling of consonants, it is impossible to tell how this *h-* was vocalized and whether it was followed by systematic strengthening or doubling of the following consonant (as, for instance, in Hebrew, but in contrast to Arabic; see Macdonald, forthcoming, *contra* Ullendorff 1965). In Dadanitic (and in some names spread over a wide geographical and chronological range) it has the form *hn-* before *ʾ* and *ʿ*. In an inscription in the Safaitic script, the gentilic *hn-ḥwly* (a tribe apparently from the region of Dadan) attests to the use of this form before *ḥ* (Macdonald 1993:308). There are as yet no examples of the article before a word beginning with *h*, but it is possible that it was *hn-* here as well.

Traditionally, it has been assumed that this *hn-* in Dadanitic was the survivor of the original form of the article before all phonemes, in all Ancient North Arabian dialects. However, had this been so, we would expect to find scattered examples of this form in other dialects (which so far we have not) and in front of other phonemes in Dadanitic (see Macdonald 2000:41–42). At present, therefore, it seems more likely that this was a development peculiar to Dadanitic and that, even there, it was simply a euphonic or dissimilatory phenomenon before glottal and pharyngeal consonants.

It was once thought that a definite article *hl-* existed in Dadanitic. However, the only examples were in two texts, one of which has now been identified as being an abecedary in the South Semitic order (JSLih 158, see Müller 1982:22); while the other is not in the



Dadanitic language but in Old Arabic written in the Dadanitic script, where *h-l-* represents a preposed demonstrative, *h-*, plus the Old Arabic definite article (ʾ)l- (JSLih 71/8, see Beeston *et al.* 1973:69–70 and Macdonald 2000:70, n. 90 and forthcoming). Compare the situation in many modern Arabic dialects, where an invariant demonstrative *ha-* with a relatively weak demonstrative force is placed before the article (e.g., *ha-l-bēt* “this house,” *ha-s-sana* “this year”; Holes 1995:152–153).

In Safaitic, the distinction between the definite article and the nearer demonstrative (“this”) is not always clear and it is possible that the article had a mild demonstrative implication (e.g., *h-dr* “the/this place,” *h-s<sup>l</sup>nt* “this year”), as it can have in Arabic (e.g., *ʾal-yawm* “the/this day,” i.e., “today”). This, of course, is different from the case in JSLih 71/8 and the modern Arabic dialects mentioned in the previous paragraph, where the demonstratives *h-* and *ha-* respectively are prefixed to the article. In Hismaic, on the other hand, *h-* is relatively rare in contexts where it would appear to represent the definite article. Thus, for instance, there is, as yet, no example in Hismaic of affiliation to a social group being expressed by the *nisbah* (see §4.1.6), in contrast to Safaitic where it is common (e.g., *h-gdly* “the Gdlyte”), while in “signatures” to rock drawings *lN bkrt* alternates with *lN h-bkrt*, “by N is the young she-camel,” where in Safaitic only the latter is found. The few possible examples of *h-* as definite article in Hismaic could equally well represent the nearer demonstrative “this” and there is, as yet, no case where it could not. It is therefore an open question whether Hismaic employed a form of determination which does not show up in the script (e.g., a final vowel, as in the Aramaic “determined state”), or had no definite article (as, in effect, in Syriac).

### 4.3.2 Demonstrative adjectives

In Dadanitic and Hismaic demonstrative adjectives are formed with *ḍ* and follow a noun defined by the article or a pronominal suffix.

In Dadanitic the masculine demonstrative adjective is *ḍh* (probably \**ḍā*), for example *h-s<sup>l</sup>fr ḍh* “this writing” (HE 1) and the feminine is *ḍt* (probably \**ḍāt*), for example *h-ṣfht ḍt* “this section of cliff” (JSLih 66/2). The demonstrative adjective *hḍh* (probably \**hādā*) is found in *h-{ṣ/}lmn hḍh* “this statuette (?)” (JSLih 82/1, cf. Arabic *hādā*).

In Hismaic, a demonstrative adjective *ḍʾ* is attested only once, in *wq<sup>ʿ</sup>-n ḍʾ* “this our inscription” (MNM 6, *pace* ed. who reads *ḍh*, though ʾ is clear on the photograph). This is a curious form since it would be highly unusual for the ʾ to represent a vowel in Hismaic. If the ʾ represents a consonant, perhaps compare *ḍʾt* in §4.1.8.3, 5. It seems possible that in the relatively rare cases in Hismaic where *h-* is prefixed to a noun with no other visible form of definition, that this represents a demonstrative adjective rather than the definite article. See the discussion in §4.3.1.

In Safaitic, the prefixed *h-* is the only form of demonstrative so far clearly attested (see §4.3.1).

### 4.3.3 Introductory particles

Most of the Ancient North Arabian graffiti and the majority of the Dadanitic monumental inscriptions begin with the name of the “author” (see §5.1.1). In the Taymanitic, Thamudic B, C, and D, Safaitic, and some Hismaic graffiti, the name is usually introduced by a particle. In Taymanitic, this is often *l* (known as the *lām auctoris*), which is probably the preposition “for, of” (see §4.3.4) which in this context means “by” in the sense of authorship, as it can in Arabic. However, a particle *lm* is also used, apparently with the same meaning (perhaps cf. Hebrew *lʾmô*, found only in the Book of Job, the language of which is thought to exhibit many

North Arabian features). This particle is characteristic of Taymanitic (Winnett 1980:135–136). What is possibly a dialectal variant of this, *nm*, is found as an introductory particle in Thamudic B, while Thamudic D texts often begin *zn* “this is . . .” In Safaitic, all but a handful of texts begin with the *lām auctoris*, while in Hismaic the author’s name can be introduced by the *lām auctoris*, or by the conjunctions *w* or *f* (see §4.3.6). In Dadanitic, no introductory particle is used (except possibly in JSLih 128). Since most of the Hasaitic inscriptions are gravestones they begin *wgr w-qbr* “tomb-chamber and grave” (see Livingstone 1984:102) or *nfs<sup>1</sup> w-qbr* “memorial and grave.”

#### 4.3.4 Vocative particles

The vocative particle is *h* in Dumaitic, Dadanitic (JSLih 8), Thamudic B, Safaitic, Hismaic, and Hasaitic (sole example unpublished). None has yet been identified in Taymanitic and Thamudic C and D. Given the nature of these texts it is not surprising that it has been found only in prayers (e.g., *h rḏw s<sup>1</sup> ʿd N*, “O Rḏw help N”; *h lt s<sup>1</sup> lm*, “O Lt [grant] security”). In origin, it was probably a sound used to attract attention (\*hā), and can be paralleled in Arabic by the *hā* which forms the initial part of a number of interjections and of the demonstrative *hādā* “this” (Wright 1896–1898:i, 268, Brockelmann 1908–1913:i, 503).

It has been suggested that in Safaitic the forms *hlyt* “O Lt” (or “O Ylt”) and so forth represent a variant vocative particle, *hy*, equivalent to Arabic *hayā* (Winnett and Harding 1978:47) or *ʾayyuhā* (Littmann 1943:21), though other explanations for this are possible. In fact, the particle *ʾyh* (\*ʾayyuhā) occurs in the invocation *w-ʾyh lt* “and O Lt” in a Safaitic inscription (unpublished) recently found in southern Syria.

In some Hismaic texts an *-m* is suffixed to the divine names *Lh* and *Lt* in invocations, thus *h lh-m*, *h lt-m* (King 1990:80). This is probably an asseverative particle which may be compared with the *-mma* in Arabic *allāhumma* (sometimes *yā allāhumma*), and possibly the *-m* in such names as *ʾabīmāʾēl* (Genesis 10:28), and *ʾbm ʿttr*, and others from Haram and its environs on the northern borders of Yemen, where the local form of Sabaic may have included a number of North Arabian features (Müller 1992:20).

#### 4.3.5 Prepositions

1. *ʾl* “towards” (cf. Arabic *ʾilā*), “for” (after the verb *ts<sup>2</sup>wq* “to yearn”): Safaitic and Hismaic.
2. *ḏky* “up to”: Dadanitic (JSLih 72/6, see Müller 1982:33 and Beeston 1979a:4).
3. *ʾl* “over, on, for, against” (cf. Arabic *ʾalā*): Safaitic and Hismaic; in Dadanitic it is usually found as *ʾly* with nouns (e.g., JSLih 81/4, 5) but as *ʾl* with pronominal suffixes (e.g., JSLih 77/3). This suggests that the final sound was a diphthong, which would not be represented in the Safaitic and Hismaic scripts. Since Dadanitic orthography only shows diphthongs in final position, the *-y* was not written when followed by a pronominal suffix. However, there are also a few examples in Dadanitic of the form *ʾl* without a pronominal suffix (e.g., U 73/4) which may indicate a pronunciation with final *-ī* or simply an uncertainty about the representation of diphthongs.

4. *ʾn pace* Caskel (1954:72), there is no clear evidence in Ancient North Arabian for a preposition *ʾn* “from” (cf. Arabic *ʾan*).

5. *b* “in, at, with, by” (cf. Arabic *bi-*): Taymanitic, Dadanitic, Thamudic B, Safaitic, and Hismaic.

6. *bʿd* “after” (cf. Arabic *baʿda*): Safaitic (e.g., SIJ 787).

The preposition occurs in Dadanitic with the meaning “for the sake of” (e.g., U 5/4, etc.). Compare Hebrew *baʿad* which is used in this sense and in a very similar context in Ezekiel

45:22 and Job 42:8 (see Stiehl 1971:9). Clive Holes informs me that in eastern Arabia a woman will plead with a loved one *yā ba'ad rūh-ī! yā ba'ad ʿen-ī! yā ba'ad ʿibd-ī!*, which is usually explained as “O you who are [the dearest thing to me] *after* my spirit/eyes/liver,” but may in fact mean “please, O X, for the sake of my spirit/eyes/liver” (personal communication). Note that Sima (1999:99–105) interprets *bʿd* as “in the direction of” in the same Dadanitic texts.

7. *bn* “between” (cf. Arabic *bayna*): Safaitic, in *h lt whbt s<sup>2</sup>n<sup>2</sup>-h bn yd-h* “O Lt may you give his enemy into his hands” (C 4037). In Arabic, the expression *bayna yaday-hi*, “between his hands,” has come to mean “in front of,” but in Safaitic it seems to retain its literal sense. In the phrase *s<sup>1</sup>nt ws<sup>1</sup>q bn rm nbṭ*, which appears to mean “the year of the conflict between the Romans and the Nabataeans” (C 4866), either the connective *w* (see §4.3.5) was not considered necessary between the two nouns (as it would be in Arabic), or it was accidentally omitted by the author or the copyist.

8. *dn* “without” (cf. Arabic *dūna*): Hismaic (unpublished).

9. *f* *pace* Winnett and Harding (1978:643) and Caskel (1954:72), there is no clear evidence in Ancient North Arabian for a preposition *f* “in” (cf. Arabic *fī*).

10. *hlḥ* “after, behind” (cf. Arabic *ḥalḥa*): Dadanitic (JSLih 70/4).

11. *l* “to, for, on behalf of” (cf. Arabic *li-*): Taymanitic (*nṣr l-ṣlm*, “he gave help to Ṣlm,” e.g., WTay 15), Dadanitic, Thamudic B, Safaitic, Hismaic. The preposition is attested in several additional uses:

- A. *Indicating possession*: Safaitic (e.g., *l-N bn N h-rgm* “the cairn is N son of N’s”, WH 329); Dadanitic (e.g., *l-N bn N h-qbr ḏh* “this grave is N son of N’s”, JSLih 312).
  - B. *In dating formulas*: Dadanitic (e.g., *s<sup>1</sup>nt ḥms<sup>1</sup> l-hn<sup>1</sup>s<sup>1</sup> bn tlmy mlk lhyn* “year five of Hn<sup>1</sup>s<sup>1</sup> son of Tlmy, king of Lhyn”, JSLih 75/5–7).
  - C. *Indicating motion*: Safaitic (e.g., *l-mdbr* “to the inner desert”, LP 180).
  - D. *Indicating purpose*: Safaitic, used with verbs in the prefix-conjugation (e.g., *l-ys<sup>2</sup>rq* “in order to migrate to the inner desert”, LP 180).
12. *ldy* “to, up to” (cf. Arabic *ladā*): Dadanitic (JSLih 77/3).
13. *m<sup>c</sup>* “in company with” (cf. Arabic *maʿa / maʿ*): Safaitic (e.g., LP 325); Dadanitic (JSLih 52/3).
14. *mn/m* “from” (cf. Arabic *min*): Thamudic B, Dadanitic, Safaitic, Hismaic *passim*. In Safaitic also with the sense “on account of” (e.g. SIAM:34).
15. *qbl* “before” (temporal, cf. Arabic *qabla*): Dadanitic (CLL 80/4).
16. *tḥt* “below” (cf. Arabic *taḥta*): Dadanitic (JSLih 50/4).

### 4.3.6 Conjunctions

Two conjunctions, *w* “and” and *f* “and (so)” “and (then)”, are attested in Ancient North Arabian. The former is found in all dialects, the latter so for only in Dadanitic, Safaitic, and Hismaic (see the discussion in Sima 1999:110–114).

### 4.3.7 Other particles

- 1. *ḏh* “when” (cf. Arabic *ḏā*): Dadanitic (JSLih 55/2).
- 2. *ʿn* “that” (cf. Arabic *ʿan*): Safaitic, in *s<sup>1</sup>m<sup>c</sup> ʿn myt flḥ* “he heard that Philip had died” (MHES p. 286).
- 3. *ʿn* “if” (cf. Arabic *ʿin*): Dadanitic (JSLih 40/6, in a very damaged context).
- 4. *ʿn* “verily” (cf. Arabic *ʿinna*): Dadanitic (JSLih 40/7, in a very damaged context).

5. *lh* negative particle (cf. Arabic *lā*): Dadanitic, *f-lh y'd*, “and so he will not be threatened” (?) in a very damaged context (JSLih 40/6).
6. *lm* negative particle followed by the prefix-conjugation (cf. *lam* plus the jussive in Classical Arabic): Safaitic (unpublished). This particle, which is characteristic of North Arabian, is also found in some of the texts from Haram on the northern borders of Yemen which are in Sabaic with some North Arabian features (see Macdonald 2000:49–50, 55–56).

## 4.4 Numerals

### 4.4.1 Cardinal numbers

These are attested in Dadanitic, Safaitic, and Hasaitic.

#### 4.4.1.1 Cardinal numbers in Dadanitic

The Dadanitic cardinal numbers are presented in Table 16.2.

The final entry in the table is so read by Sima, though the first and last words are more or less invisible on the published photograph and these lines were not copied by Abū al-Ḥasan.

It will be seen from Table 16.2 that there are some interesting similarities and differences between the treatments of numerals in Dadanitic and in Classical Arabic.

1. As far as we can tell on present evidence, numerals precede the nouns to which they refer; nouns following the numbers three to ten are in the plural, while those following eleven and upwards are in the singular, as in Classical Arabic. However, the situation is obscured by the fact that, in Dadanitic, the vast majority of the examples of numerals are in dates, where the noun (*s'nt*) precedes the number and is, by definition, singular.

2. The principles of agreement in gender with the preceding or following noun appear to be roughly the same as in Classical Arabic, namely that numerals of a feminine form refer to nouns which (in the singular) are masculine and vice versa. Since *'ym* “days” follows the forms of numerals referring to a feminine noun in both Dadanitic (*'s<sup>2</sup>r 'ym*) and Safaitic (*s<sup>1</sup>t 'ym*), it seems probable that the word *\*ym* “day” must have been regarded as feminine in these dialects (see §4.1.1).

3. If this is correct, it is probable that the final *t* in *tlt* (*tlt 'ym*) is part of the root (*tlt* < *\*tlt*) rather than the equivalent of Arabic *tā' marbūṭah* (see §3.1.1). Unfortunately, the word *m'n* in *tlt m'n* has not yet been satisfactorily interpreted and so we cannot be certain whether or not it is the plural of a feminine noun and therefore whether the second *t* in *tlt* should be explained in the same way. However, it should be noted that the development *\*/t/ > /t/* is not typical of Dadanitic and so far appears to be peculiar to this word. The Dadanitic form, *tltt*, used with masculine nouns and Safaitic *tltt/tlt* are identical to the Classical Arabic forms.

4. In compound numbers, the units continue to take the opposite gender to the noun, but from twenty upwards the tens are (probably) of common gender, again as in Classical Arabic. However, an interesting difference is observable in the numbers thirteen through nineteen, where in Classical Arabic (and Safaitic, see §4.4.1.2) the ten takes the same gender as the noun and the unit the opposite. In the only Dadanitic example available so far, *s'nt 's<sup>2</sup>r w-s<sup>1</sup>b'* (where Classical Arabic would have *sanat sab'a 'ašrata*), either the ten was regarded as of common gender (like twenty, etc.) or it behaved in the same way as the units, taking the opposite gender to the noun.

5. In the compound numerals, the larger unit is generally placed before the smaller, contrary to the practice in Classical Arabic. This occurs both in the numbers from thirteen through nineteen (e.g., *'s<sup>2</sup>r w-s<sup>1</sup>b'* “seventeen,” cf. Classical Arabic *sab'a 'ašrata* and Safaitic

**Table 16.2 The cardinal numerals in Dadanitic**

Masculine	Common	Feminine
1		<i>s<sup>1</sup> nt ḥdy</i> “year one” (CLL 26/4)
2		<i>s<sup>1</sup> nt ttn</i> “year two” (JSLih 45/3)
3	<i>tltt ḥzl</i> “three ḥll ceremonies” (U 32/3–4) <i>tltt ḥl</i> “three ḥll ceremonies” (U 50/2–3)	<i>tlṭ ḥym</i> “three days” (JSLih 68/4) notes 2,3  <i>tlṭ m<sup>c</sup> n</i> “three . . .” (? , JSLih 47/2) note 3
5		<i>s<sup>1</sup> nt ḥms<sup>1</sup></i> “year five” (JSLih 75/5; Scagliarini 1996:96–97)
10	<i>ṣ<sup>2</sup> rt mnḥ{ḥ}</i> “ten canals” (JSLih 177/1)	<i>ṣ<sup>2</sup> r ḥym</i> “ten days” (CLL 86/3) note 2
17		<i>s<sup>1</sup> nt {ṣ<sup>2</sup>} r w-s<sup>1</sup> b<sup>c</sup></i> “year seventeen” (U 8/4–5)
20	<i>s<sup>1</sup> nt ṣ<sup>2</sup> rn</i> “year twenty” (JSLih 68/2–3; AH 63/5, 64/7–8? see Sima 1999:38)	
22		<i>s<sup>1</sup> nt ṣ<sup>2</sup> rn {w}-ttn</i> “year twenty-two” (JSLih 77/11)
29		<i>s<sup>1</sup> nt ṣ<sup>2</sup> rn w-ts<sup>1</sup></i> “year twenty-nine” (CLL 86/2–3; JSLih 83/6)
35		<i>s<sup>1</sup> nt tltn w ḥms<sup>1</sup></i> “year thirty-five” (JSLih 82/3–4)
40	<i>ṛb<sup>c</sup> n s<sup>1</sup> lṭ</i> “forty drachmas” (JSLih 177/2)	
120	<i>m<sup>2</sup> t w-ṣ<sup>2</sup> rn</i> . . . (JSLih 77/5)	
140	<i>m<sup>2</sup> t w-ṛb<sup>c</sup> n</i> . . . (CLL 33/2)	
145	<i>m<sup>2</sup> t w-ṛb<sup>c</sup> n w-ḥms<sup>1</sup> nḥl?</i> “one hundred and forty-five palm trees” (U 23/4–5 = AH 41)	

*tmn ṣ<sup>2</sup> rt*, see §4.4.1.2), and from twenty onwards (e.g., *tltn w-ḥms<sup>1</sup>*, cf. Classical Arabic *ḥamsun wa-ṭalāṭūna*). Note also that, in the teens, unit and ten are connected by *w-* in Dadanitic but not in Arabic. See the discussion in Sima 1999:119, but note that the supposed examples of *s<sup>1</sup> tt ṣ<sup>2</sup> r* and *s<sup>1</sup> t ṣ<sup>2</sup> r* are very doubtful and that the restoration *ṣ<sup>2</sup> r w-t[s<sup>1</sup>]/{<sup>c</sup>}* in AH 81/6 (n. 28) looks unlikely on the published copy.

6. The form *tltn* may have resulted from an original \**ṭintān* (i.e., without a prosthetic initial vowel, cf. Classical Arabic *ṭintāni* beside *itnatāni*, also *ṭintēn* in modern dialects of central and eastern Arabia) with the assimilation of vowelless /n/ characteristic of Dadanitic and other Ancient North Arabian dialects (see §3.1.5).

#### 4.4.1.2 Cardinal numbers in Safaitic

In Safaitic no example of the numeral “one” has yet been found, though a verb *whd* “he was alone” is well attested. The dual is used for “two”. The other Cardinal numbers attested in

Safaitic are as follows:

(16)	<i>Masculine</i>	<i>Common</i>	<i>Feminine</i>
3	<i>tl̥tt̥ s̥²hr</i> “three months” (WH 3792a)		<i>tl̥t̥ s̥¹nn</i> “three years” (AZNG)
4			<i>ʾrb̥ s̥¹nn</i> “four years” (WH 3094)
5	<i>hms̥¹t̥ mny</i> “five minas” [a coin] (C 3916)		<i>hms̥¹ ws̥¹q</i> “five herds of camels” (C 2088)
6			<i>s̥¹t̥ ʾym</i> “six days” (unpublished)
18			<i>s̥¹nt̥ t̥mn̥ s̥²rt̥</i> “year eighteen” (LP 1064)
100		<i>m̥ʾt̥ fr̥s̥¹</i> “a hundred horsemen” (WH 1849)	

In contrast to Dadanitic, the rules of agreement in gender and number between a numeral and the noun to which it refers appear to be the same in Safaitic as in Classical Arabic, except in the case of *s̥¹t̥ ʾym* (see note 2 above). Similarly, the form of the single example of a compound number in *s̥¹nt̥ t̥mn̥ s̥²rt̥* is paralleled almost exactly by Classical Arabic *sanat tamāniya ʿašrata*.

#### 4.4.1.3 Cardinal numbers in Hasaitic

The following cardinal numbers, all feminine, are attested in Hasaitic:

- (17) 6 *s̥¹nt̥ s̥¹t̥* (unpublished)  
 27 *s̥¹nt̥ s̥²{rn} w s̥¹b{ʿ}* (Robin-Mulayḥa 1, *contra* ed.)  
 34 *ʾrb̥ w-tl̥tn̥ s̥¹nt̥* giving a person’s age (Livingstone 1984:100)

#### 4.4.2 Ordinal numbers

No ordinal numbers have yet been identified.

#### 4.4.3 Totality

The notion of totality is expressed in Safaitic, Hismaic, and Dadanitic by *kll* (\*kulil (?), cf. Arabic *kull*). As in Arabic, when *kll* is followed by an undefined entity it means “each, every”: for example, *kll s̥²r̥ ṣdq* “every true kinsman” (HCH 191, Safaitic; MNM 6, Hismaic). When it is followed by a defined entity (so far only pronominal suffixes are attested), it means “all” or “the whole”: for example, in Dadanitic *h-mq̣ḍ kll-h* “the whole sitting-place” (HE 1); Safaitic *s̥²ỵh kll-hm* “all his companions” (LP 243).

## 5. SYNTAX

Given the fragmentary and formulaic nature of the available documents, no coherent description of Ancient North Arabian syntax can yet be attempted. The following notes represent some miscellaneous features which can be gleaned from the Dadanitic and Safaitic texts.

## 5.1 Word order

### 5.1.1 Word order in verbal sentences

#### 5.1.1.1 Dadanitic

The majority of Dadanitic inscriptions begin with the subject followed by the verb followed by the object (i.e., they are SVO) and then adverbial or prepositional phrases:

- (18) 1. N *bn* NN *qrb h-šlm l-dğbt*  
 “N son of NN offered the statue to Dğbt” (JSLih 41/1–3)  
 2. N *ktb-h b-dh*  
 “N wrote it here” (JSLih 279)  
 3. N<sub>1</sub> *w*-N<sub>2</sub> [SUBJECTS] *ʔllh* [verb] *h-ʔll* [OBJECT] *l-dğbt b-khl b'd ml-hm b-bdr*  
 [PREPOSITIONAL PHRASES]  
 “N<sub>1</sub> and N<sub>2</sub> have performed the ʔll ceremony **for Dğbt in**  
**Khl for the sake of** their winter crops **in Bdr**” (U 19/1–6)

This order may not reflect normal practice but rather the nature of the texts, which are mainly dedications, records of the performance of religious rites, and graffiti, in which the name of the “author” was inevitably given prominence.

By contrast, the VSO (or VOS) order, which is the norm in Classical Arabic, is very rarely attested in the Dadanitic inscriptions:

- (19) *hls<sup>1</sup>* N<sub>1</sub> *bn* N<sub>2</sub>  
 died N<sub>1</sub> son of N<sub>2</sub>  
 “N<sub>1</sub> son of N<sub>2</sub> died” (literally “was carried off,” CLL 78, 79, 80)

#### 5.1.1.2 Safaitic

Unlike the Dadanitic inscriptions, the Safaitic graffiti usually begin with the *lām auctoris* (see §4.3.2) followed by the author's name and part of his genealogy. Any statement is then linked to the genealogy by the connective *w* “and.” This permits a natural word order within the statement, in contrast to the Dadanitic texts where it may have been distorted by the need to begin the first sentence with the author's name for the sake of emphasis.

The usual word order in Safaitic is VSO or VOS, as in Classical Arabic. Even if they existed, case endings, being short vowels, would not show up in Safaitic orthography and it is therefore sometimes impossible to decide which is the subject and which the object in a sentence. Thus:

- (20) 1. *s<sup>1</sup>nt hrbt ʔ wḏ ʔ šbh,*  
 “the year the ʔ wḏ made war on [or “plundered”] the ʔ šbh,” or *vice versa*  
 (SIJ 59, see also C 2577)  
 2. *s<sup>1</sup>nt s<sup>1</sup>lm ʔ b'd ʔ wḏ,*  
 “the year the ʔ B'd made peace with the ʔ wḏ,” or *vice versa*  
 (C 4394, wrongly transliterated in C)

The indirect object can also precede the direct object:

- (21) 1. *ngy b-h-bqr h-nhl,*  
 he fled with-the-cows the-valley  
 “and he fled the valley with the cows” (LP 90)



2. *bny*      *l-s<sup>1</sup>‘d*      *h-rgm*,  
 he built    for-S<sup>1</sup>‘d    the-cairn  
 “he built the cairn for S<sup>1</sup>‘d” (WH 421)

Verbs in Safaitic can take multiple direct objects: for example, *r’y h-’bl h-nhl bql*, “he pastured the camels (*h-’bl*) [in] the valley (*h-nhl*) [on] spring herbage (*bql*)” (C 2670). Compare *r’y h-nhl bql n’m-hm*, “he pastured their small cattle (*n’m-hm*) [in] the valley [on] spring-herbage” (C 1534).

### 5.1.2 Word order in nominal sentences

In common with Arabic and other Semitic languages, the Ancient North Arabian dialects used nominal sentences instead of employing the verb “to be” as a copula.

Thus in Dadanitic: *w-’n N bn N*, “and I [am] N son of N” (CLL 57/2; also in Thamudic D e.g., JSTham 637, and Hismaic e.g., King 1990: KCJ 646)

*l-N h-mtbr* (literally “to/for N [is] the grave-chamber”), i.e., “the grave-chamber belongs to N” (JSLih 366/1)

There are numerous examples in Safaitic. Thus

*l-N h-htt*, “By N [are] the carvings” (e.g., WH 368)

*l-N w-h-htt*, ‘By N and the carving [is by him?]’ (WH 353)

*l-N w-h-rgm*, “For N and the cairn [is his]” (HCH 1, 2), where we know from other texts that this person was the occupant of the grave under the cairn.

*l-N w-l-h h-bkrt*, “By N and the young she-camel [is] his [or “is by him”]” (WH 2833b)

*l-N w-l-h-rgm*, “For N and for him/her [is the] cairn” (WH 3420, etc.); for the assimilation of the article to the preceding enclitic personal pronoun, see §4.1.8.2, 7.

*w-b’s<sup>1</sup>l-h*, literally “and distress [was] to him”, i.e. “he was in distress” (CSNS 779)

*l N h-dr*, literally “by/for N the place”. This is a very common expression in the Safaitic inscriptions. It is unlikely to be a claim to personal real estate, something which is impractical in the nomadic life. Instead, it almost certainly means simply “N was here”.

Note also the word order in the nominal phrase

*l-N b-ms<sup>1</sup>rt ’l ’mrt frs<sup>1</sup>*, “by N, a horseman (*frs<sup>1</sup>*) in the unit (*ms<sup>1</sup>rt*) of the ’l ’mrt” (Macdonald 1993: 374).

### 5.1.3 Annexation

Annexation (the *idāfa* of the Arab grammarians) is a fundamental feature of Semitic grammar (see Ch. 10) in which two or more elements are bound together to form a grammatical and semantic unit. Nothing is allowed to intervene between the elements (except in certain very specific circumstances of which we have no examples in Ancient North Arabian) and thus items such as adjectives (including demonstrative adjectives) follow the final element even if they refer to the first. The unit as a whole is defined or undefined according to the nature of the final element even if one of the preceding elements would otherwise normally take the definite article (see under Safaitic, below).

Examples of annexation in Dadanitic are:

**Undefined** *b-ḥqwy kfr* (\*ḥaqway) ‘on two sides of a tomb’ (JSLih 75/3)

**Defined 3-element annexation** *kbry s<sup>2c</sup>t h-nš* “the two kabīrs of the association of H-NŠ” (CLL 77/3-4)

**Defined + a demonstrative** *’rr h-s<sup>1</sup>fr dh* “the dishonorer of this inscription” (HE 1/5-6).

Examples of annexation in Safaitic are:



**Undefined + adjective** *kl l s<sup>2</sup>r šdq* “every true friend” (HCH 191, also in Hismaic MNM 6)  
**Defined by the article** (1) *m<sup>1</sup>wr h-s<sup>1</sup>fr* “the scratcher-out of the writing” (e.g., WH 1679),  
 (2) *nmrt h-s<sup>1</sup>ltn* “Namārah of the government” (LP 540). When not annexed, the place-name is *h-nmrt* (e.g., LP 330, cf. the modern name, *al-Namārah*)  
**Defined by a name** *hrb nbṭ* “the war of the Nabataeans” (C 3680).

### 5.1.4 Demonstrative Adjectives

When the modified noun is part of a noun phrase, two constructions are possible: (i) *h-zll dh l-dḡbt* (U 33/2-3) or (ii) *h-zll l-dḡbt dh* (U 4/3), both of which mean “this zll-ceremony for Dḡbt.” The second construction is bizarre and may be an error on the part of the engraver.

## 5.2 Agreement

In Ancient North Arabian verbs agree with their subjects in gender and number, regardless of their position in the sentence (in contrast to Classical Arabic, Wright 1896–1898: ii, 289–290).

In Dadanitic, the only dialect in which it is identifiable, the use of the dual in verbal agreement is erratic. Thus, it is used after two subjects in some texts:

- (22) *N<sub>1</sub> w-N<sub>2</sub> ṣlh h-zll*  
 “N<sub>1</sub> and N<sub>2</sub> have performed the zll-ceremony” (U 19/1–4)

whereas in others the plural verb is used:

- (23) A. *N<sub>1</sub> w-N<sub>2</sub> wdyw*  
 “N<sub>1</sub> and N<sub>2</sub> have erected (?)” (JSLih 77/2)  
 B. *kbry s<sup>2</sup>t h-nṣ ḥdw*  
 “The two kabīrs of the association of H-NṢ have taken possession” (CLL 77/3–4)

The same variation can be seen in the use of enclitic personal pronouns (§4.1.8.2). Thus, in U 19 the two subjects are followed by a verb in the dual (ṣlh), but are later referred to by the plural enclitic personal pronoun *-hm* (lines 5–7). By contrast, in U 69, the two subjects are followed by a verb in the plural (ṣllw), but are referred to later by the dual pronominal suffix *-hmy*. See Sima 1999:117–118 for tables showing the variations in agreement in the inscriptions from al-ʿUdayb. Compare the situation in the modern spoken Arabic dialects, where the dual is in general use on nouns, but requires plural concord in the verb, adjectives, and pronouns (Clive Holes). This is a very old feature in the dialects which can already be seen in the earliest Arabic papyri (see Hopkins 1984:94–98).

## 5.3 Verb conjugations

The suffix- and prefix-conjugations are each associated with particular usages.

### 5.3.1 The suffix-conjugation

In Dadanitic, the suffix-conjugation is used of completed acts, e.g., *N ḥd h-mqbr*, “N has taken possession of the tomb” (JSLih 306), and for the optative: *rr dḡbt*, “may Dḡbt dishonor” (HE 1/4–5); or *rdy-h*, “may he [the deity] favor him” (U 18/4–5) in contrast to the imperative, *rd-h*, “favor him,” which is more common in this formula.

In Safaitic, the suffix-conjugation has four distinct functions. First, it is used for completed acts and, in particular, acts which preceded the author's present state or actions (where Classical Arabic would have the perfect, or *kāna* + the perfect, or *qad* + the perfect): for example, *nfr mn rm* "he had fled from Roman territory" (e.g., C 3721); *wgd ʾtr ʾm-h f-ngʿ*, "he had found the inscription of his grandfather and so he was grieving" (e.g., C 793); *wgm ʾl N mqt l qtl-h ʾl ḥwlt*, "he was mourning for N, a murder-victim, whom the ʾl Ḥwlt had killed" (lit. "... killed the ʾl Ḥwlt killed him," HCH 126); *sʾmʿ n myt flfš* "he heard that Philip had died" (MHES p. 286).

Second, the suffix-conjugation is used for descriptions of the author's state, or acts which were not complete, at the time of writing: *dʾt* "he is spending the season of later rains"; *rʿy* "he is pasturing"; *wgm* "he is grieving"; *ḥrš* "he is keeping watch" (where Arabic would use the imperfect).

Third, in Safaitic, as in Classical Arabic, it is used for the optative: *f-h lt whbt sʾnʿ-h bn yd-h* "and so, O Lt, may you give his enemy into (lit. between) his hands" (C 4037). This construction is also frequent in Hismaic: for example, in *ḍkrt lt*, "may Lt be mindful of" (e.g., TIJ 58, etc.).

Fourth, the suffix-conjugation can be used as a virtual subjunctive: *sʾlm l-ḍ sʾr w-ʿwr l-ḍ ʿwr h-sʾfr*, "security to whoever leaves (i.e., "may leave") intact and blindness to whoever scratches out (i.e., "may scratch out") the inscription" (e.g., LP 361). Compare the same formula using the prefix-conjugation in §5.3.2.

### 5.3.2 The prefix-conjugation

The handful of Dadanitic examples of the prefix-conjugation are all in damaged or doubtful contexts.

However, four distinct uses of the prefix-conjugation can be identified for Safaitic. First, it is used in clauses expressing purpose: *l-ysʿrq* "in order to migrate to the inner desert" (LP 180).

Second, the Safaitic prefix-conjugation occurs with a jussive implication: *nngy* "may we escape" (WH 135). Note also *nhyy* "may we live prosperously" in Thamudic B (LP 495).

Third, after the negative particle *lm* the prefix-conjugation has a perfect implication as in Classical Arabic (in an unpublished text).

Finally, the prefix-conjugation is used with a subjunctive implication: *sʾlm l-ḍ sʾr w-ʿwr l-ḍ yʿwr*, "security to whoever leaves (i.e., "may leave") intact and blindness [cf. Arabic *ʿawar*] to whoever scratches out (i.e., "may scratch out")" (e.g., LP 391). There seems to be no difference in meaning between invocations which use the suffix-conjugation (see §5.3.1) and those which use the prefix-conjugation.

### 5.4 Participles

Several different uses of participles are attested in Safaitic. An active participle can function as a finite verb with a perfective sense: for example, *w-wgd ʾtr gsʿ-h qbrn ḍwʾl yzr* "he found the traces of his raiding party, members of the ʾl Yzr having performed the burial" (C 2156); *wlh ʾl ʾsʿyʿ-h ḥrbn ʾl tʿyʿ* "he grieved for his companions [who were] raiding [\*ḥāribīn] the tribe of ʾy" (C 2795). In addition, active participles often form a circumstance clause (in Arabic grammar, a *ḥāl*): for example, *w-whd ḡzz* "and he was alone on a raid" (WH 128), where *ḡzz* is an active participle (\*ḡāziz); *ḥll h-dr šyr m-mḍbr* "he camped at this place while returning to permanent water [šyr] from the inner desert" (C 2590), where *šyr* is an active participle (\*šāyir).

Participles can be used as virtual relative clauses (see §5.5). The active participle can take a direct object, as in C 2795 above, while a passive participle can be used either on its own (e.g., *wgm ʿl sʿyd mqtł* “he mourned for Sʿyd who had been killed”; CSNS 1004), or in construct with another word (e.g., N *mqtł ty* “N victim of [i.e., who had been killed by] Ty”; CSNS 1011). This is probably the explanation of the passive participles which often follow the names of those for whom an author mourns: thus N *trḥ* (\*tarīḥ) “N who is untimely dead”; N *rḡm mny* (\*raḡīm manāyā) “N who has been humbled by (lit. “of”) the Fates.”

## 5.5 Relative clauses

In Safaitic, relative clauses can be formed with the relative pronoun *d* (see §4.1.8.3, 3).

- (24) *h lt ʿyr m-d qtl-h*  
 O Lt recompense from-who killed-him  
 “O Lt [grant] recompense from [him] who killed him” (LP 385)

and with the relative *mn* (\**man*; see §4.1.8.3, 1):

- (25) *ʿwr l-mn yʿwr h-sʿl fr*  
 blindness to-whomever scratches out the writing  
 “And blindness to whoever scratches out the writing” (SIJ 284)

Relative clauses can also be formed without a relative pronoun simply by using the prefix-conjugation with an implied or explicit reference back to the antecedent. This type of relative clause can be used in Safaitic even after a defined antecedent, contrary to the practice in Classical Arabic, though it is found at earlier stages of the language (cf. Beeston 1970:50, n.1):

- (26) *l-h h-mhrt yrbh-h*  
 to-him [is] the-filly he is training-it  
 “His is the filly which he is training” (C 1186)

Such a relative clause can also be constructed using the suffix-conjugation, and again can be employed even after a defined antecedent:

- (27) *wgm... ʿl ʿnʿm qtl-h ʿl šbh*  
 he mourned... for-ʿnʿm killed-him ʿl Šbh  
 “He mourned... for ʿnʿm whom the ʿl Šbh had killed” (C 4443)

## 5.6 Invocations

In Safaitic, invocations can be expressed in three different ways: (i) by the vocative particle *h* + divine name + imperative + predicate (e.g., *h lt ʿwr d yʿwr h-sʿl fr* “O Lt blind whoever scratches out the writing”); (ii) by the vocative particle *h* + divine name + an understood verb + noun (e.g., *h lt ḡnmt* “O Lt [grant] booty”; cf. Arabic *ḥanānayka yā rabbi* “O Lord have mercy on me” for *taḥannan ʿalayya ḥanānan*, Wright 1896–1898:ii, 73); and (iii) by a verb in the suffix-conjugation with an optative implication + divine name + predicate. This is particularly common in Hismaic: for example, *ḏkrt lt N.*, “may Lt be mindful of N.”

## 6. LEXICON

Since Ancient North Arabian is known only from inscriptions, 98 percent of which are graffiti, there is a vast disproportion between the size of the recorded onomasticon and

the surviving lexicon. The former is huge, perhaps the largest collection of personal names in any group of Ancient Near Eastern texts. Indeed, in reality it is even larger than it appears, since no vowels or doubled letters are shown and in many cases the same group of consonants must have covered several different names distinguished only by their vocalizations or by consonant doubling (e.g., *S<sup>l</sup>lm* could represent \*S<sup>l</sup>alm, \*S<sup>l</sup>ālim, \*S<sup>l</sup>alim, etc.).

By contrast, the lexicon that has survived is tiny and is severely limited in range by the subject matter of the texts. This is particularly true of Dadanitic, where the vast majority of the monumental inscriptions are dedications, or record the performance of religious duties, whereas the graffiti consist almost entirely of names. Similarly, since the Hasaitic inscriptions found so far are virtually all gravestones, they have yielded a very limited vocabulary. On the other hand, the Safaitic (and, to a lesser extent, the Hismaic) graffiti deal with a wide range of subjects, albeit very laconically.

In the past, the main resource for interpreting the Ancient North Arabian lexicon has been Classical Arabic. However, Modern Arabic dialects are being used increasingly to help explain features in Ancient North Arabian (particularly Safaitic) which do not occur in the Classical language. For instance, the word *ʿs<sup>2</sup>rq* (found in Safaitic) has traditionally been translated “he went east,” based on Classical Arabic *šarraqa*. However, it is clear from the texts that their authors used *ʿs<sup>2</sup>rq* in the same way as the modern bedouins of the same area use *šarraqa*, in the sense of “he migrated to the inner desert,” regardless of whether that meant traveling north, south, east, or west. There are also a number of words where the meaning has not been preserved in Arabic, but can be found in the cognate in another Semitic language, for example the word *n<sup>h</sup>l* in Safaitic which means “a valley” (cf. Hebrew and Aramaic *naḥal*), as opposed to Arabic *naḥl* “a palm tree.” Similarly, the word *ʒ<sup>l</sup>* in Taymanitic and possibly Lihyanite is probably to be interpreted as “leader” on the basis of Sabaic (see Macdonald 1992a:30–31).

However, there are also a number of words for which etymology does not seem to provide an appropriate meaning and which therefore, at present, have to be explained from their context: for example, *h<sup>r</sup>s* in Safaitic which appears to mean “he kept watch,” or *wgm*, which seems to be one of the numerous words for “to mourn” in that dialect. Sima argues that the key words in the Dadanitic vocabulary of the inscriptions from al-ʿUdayb (a side-valley near al-ʿUlā) relate to the maintenance of the irrigation system (1999:90–105), but this is often difficult to justify philologically, and the context usually seems to point to the performance of a religious ceremony.

Given the nature of the material, a complete description of Ancient North Arabian will never be possible. However, large numbers of new, well-recorded texts are becoming available (particularly in Safaitic) and much careful analysis is being undertaken. It may therefore not be too long before it will be possible to present a rather more detailed description than that offered here.

## 7. READING LIST

In Macdonald 2000, I have discussed the languages of pre-Islamic Arabia (i.e., not just Ancient North Arabian) at a more general level and explained the terminology. For a masterly brief discussion of Ancient North Arabian (with some different views from those expressed here) see Müller 1982. Sass 1991 presents a detailed analysis of the dispersed ONA texts though for a brief critique of his use of paleography see Macdonald 2004a. Caskel 1954 is

still the most recent published overall description of Dadanitic (Lihyanite), though a number of unpublished doctoral theses have been devoted to the subject. Caskel's work is marred by many strained interpretations of the texts and an attempt to force the language into the mold of Classical Arabic. However, Sima 1999 presents an excellent edition and analysis of an important group of Dadanitic texts and, although some of his conclusions are disputed, this marks a significant advance in our knowledge of the language. For a brief general outline of the present state of Thamudic studies (plus Taymanitic and Hismaic), see Macdonald and King 1999 and references there. For a similarly brief outline of Safaitic, see Müller 1980 and Macdonald 1995. Readings of the full corpus of the Hasaitic inscriptions (though regrettably without photographs) together with an excellent study can be found in Sima 2002. Finally, it should be noted that readings and interpretations of Ancient North Arabian texts published by A. Jamme and A. van den Branden should be treated with great caution.

### Abbreviations

AH	Dadanitic inscriptions originally published in Abū al-Ḥasan 1997 and republished in Sima 1999
AZNG	Safaitic inscription in Abbadi and Zayadine 1996
C	Safaitic inscriptions in <i>Corpus Inscriptionum Semiticarum. Pars V</i> . Paris, 1950–1951
CIH	South Arabian and Hasaitic inscriptions in <i>Corpus Inscriptionum Semiticarum. Pars IV</i> . Paris, 1889–1932
CLL	Dadanitic inscriptions in Caskel 1954
CSNS	Safaitic inscriptions in Clark 1979
CTSS	Hismaic inscriptions in Clark 1980
HCH	Safaitic inscriptions in Harding 1953
HE	Dadanitic and Taymanitic inscriptions in Harding 1971b
HU	Taymanitic, Hismaic, and Thamudic B, C, and D inscriptions copied by C. Huber and renumbered in van den Branden 1950
Ja 1046	Hasaitic inscription in Jamme 1966:72–73
JSLih	Dadanitic inscriptions in Jaussen and Savignac 1909–1922
JSTham	Taymanitic, Hismaic, and Thamudic B, C, and D inscriptions in Jaussen and Savignac 1909–1922
LP	Safaitic and Thamudic B inscriptions in Littmann 1943
MHES	Safaitic inscriptions in Macdonald 1995b
MNM	Hismaic inscriptions in Milik 1958–1959
MSTJ	Safaitic inscriptions in Macdonald and Harding 1976
NST	Safaitic inscriptions in Harding 1951
Ph	Taymanitic, Hismaic, and Thamudic B, C, and D inscriptions copied by H. St.J. B. Philby and published in van den Branden 1956
Robin-Mulayḥa 1	Hasaitic inscription in Robin 1994:80–81
SIAM i	Safaitic inscriptions in Macdonald 1979
SIJ	Safaitic inscriptions in Winnett 1957
TIJ	Hismaic inscriptions in Harding and Littmann 1952
U	Dadanitic inscriptions from al-ʿUḏayb published (and republished) in Sima 1999
WH	Safaitic inscriptions in Winnett and Harding 1978
WTay	Taymanitic inscriptions in Winnett and Reed 1970
WTI	Dumaitic, Hismaic, and Thamudic B, C, and D inscriptions in Winnett and Reed 1970

## Bibliography

In the bibliography, works are listed alphabetically by author, but each one is marked with one or more of the following letters which give an indication of the subject matter:

- D** Dadanitic  
**G** General  
**H** Hismaic  
**Ha** Hasaitic  
**OA** Old Arabic  
**ONA** Oasis North Arabian  
**S** Safaitic  
**T** Taymanitic  
**Th** Thamudic B, C, D, Southern Thamudic

- Abbadī, S. and F. Zayādīn. 1996. "Nepos the governor of the Provincia Arabia in a Safaitic inscription?" *Semitica* 46:155–164. **S**
- Abū al-Ḥasan, H. 1997. *Qirāʾa li-kitābāt lihyāniyya min ḡabal ʾakma bi-miṭṭaqaṭ al-ʿulā*. Al-Riyāḍ: Maktabat al-malik fahd al-waṭaniyya. [Note that the inscriptions in this work were published in facsimile without any photographs and the author's readings must therefore be regarded as unverified. However, most of these texts, together with photographs of almost eighty of them, have now been republished in Sima 1999. References to "AH [= Abū al-Ḥasan] + number" are therefore to Sima's edition not to the *editio princeps*.] **D**
- . 2002. *Nuqūṣ lihyāniyah min miṭṭaqaṭ al-ʿulā*. (*Dirāsah taḥlīliyah muqāranah*). Riyadh: Wizārat al-maʿārif wakālat al-wizārah li-l-āṭār wa-l-matāḥif. **D**
- Beeston, A. 1962. "Arabian sibilants." *Journal of Semitic Studies* 7:222–233. **G**
- . 1970. *The Arabic Language Today*. London: Hutchinson. **G**
- . 1979a. "Nemara and Faw." *Bulletin of the School of Oriental and African Studies* 42:1–6. **OA**
- . 1979b. Review of Winnett and Harding 1978. *Antiquaries Journal* 59:133–134. **S**
- Beeston, A. et al. 1973. "The inscription Jaussen-Savignac 71." *Proceedings of the Seminar for Arabian Studies* 3:69–72. **OA**
- Brockelmann, C. 1908–1913. *Grundriß der vergleichenden Grammatik der semitischen Sprachen*. Berlin. Reprint 1966, Hildesheim: Olms. **G**
- Cantineau, J. 1930–1932. *Le Nabatéen*. (2 vols.). Paris: Ernest Leroux. **G**
- Caskel, W. 1954. *Lihyan und Lihyanisch*. Arbeitsgemeinschaft für Forschung des Landes Nordrhein-Westfalen, Geisteswissenschaften 4. Cologne: Westdeutscher Verlag. **D**
- Clark, V. 1979. *A Study of New Safaitic Inscriptions from Jordan*. Ph.D. thesis, University of Melbourne. Ann Arbor, MI: University Microfilms (1983). **S**
- . 1980. "Three Safaitic stones from Jordan." *Annual of the Department of Antiquities of Jordan* 24:125–128. **H**
- Drewes, A. 1985. "The phonemes of Lihyanite." In C. Robin (ed.), *Mélanges linguistiques offerts à Maxime Rodinson par ses élèves, ses collègues et ses amis*, pp. 165–173. GLECS Supplément 12. Paris: Paul Geuthner. **D**
- Eskoubi, Kh. 1999. *An Analytical and Comparative Study of Inscriptions from "Rum" Region, South West of Tayma* (in Arabic). Riyadh: Ministry of Education, Deputy Ministry of Antiquities and Museums. [This book reached me too late to permit the data from these new inscriptions to be included in this survey. While the photographs of each inscription make this an important publication, it should be remembered that it is an M.A. thesis and that the author's interpretations of many of these difficult texts are disputed.] **T, Th**
- Faber, A. 1997. "Genetic subgrouping of the Semitic languages." In R. Hetzron (ed.), *The Semitic Languages*, pp. 3–15. London: Routledge. **G**
- Harding, G. 1951. "New Safaitic texts." *Annual of the Department of Antiquities of Jordan* 1:25–29. **S**
- . 1953. "The cairn of Hani." *Annual of the Department of Antiquities of Jordan* 2:8–56. **S**
- . 1971a. *An Index and Concordance of Pre-Islamic Arabian Names and Inscriptions*. Near and Middle East Series 8. Toronto: University of Toronto Press. **G**
- . 1971b. "The Thamudic and Lihyanite texts." In P. Parr, G. Harding, and J. Dayton (eds.), "Preliminary Survey in NW Arabia, 1968. Part II: Epigraphy," pp. 36–52, 61. *Bulletin of the Institute of Archaeology, University of London* 10:36–61. **D, T, Th**



- Harding, G. and E. Littmann. 1952. *Some Thamudic Inscriptions from the Hashemite Kingdom of Jordan*. Leiden: Brill. **H**
- Holes, C. 1995. *Modern Arabic. Structures, Functions and Varieties*. London: Longman. **G**
- Hopkins S. 1984. *Studies in the Grammar of Early Arabic. Based upon Papyri Datable to before AH 300 / AD 912*. London Oriental Series 37. Oxford: Oxford University Press. **OA**
- Jamme, A. 1966. *Sabaeen and Ḥasaeen Inscriptions from Saudi Arabia*. Studi semitici 23. Rome: Università di Roma. **Ha**
- Jaussen, A. and M. Savignac. 1909–1922. *Mission archéologique en Arabie* (6 vols.). Paris: Leroux/Ernest Paul Geuthner. Reprint 1997, Cairo: Institut Français d'Archéologie Orientale. **D, T, Th**
- King, G. 1990. "Early North Arabian Thamudic E. A Preliminary Description Based on a New Corpus of Inscriptions from the Ḥismā desert of Southern Jordan and Published Material." Ph.D. thesis, University of London [unpublished]. **H**
- Kofler, H. 1940–1942. "Reste altarabischer Dialekte." *Wiener Zeitschrift für die Kunde des Morgenlandes* 47 (1940):61–130, 232–262; 48 (1941):52–88, 247–274; and 49 (1942):15–30, 234–256. **OA**
- Littmann, E. 1943. *Safaitic Inscriptions*. Syria. Publications of the Princeton University Archaeological Expeditions to Syria in 1904–1905 and 1909. Division IV, Section C. Leiden: Brill. **S**
- Livingstone, A. 1984. "A linguistic, tribal and onomastical study of the Hasaeen inscriptions." In M. Gazdar, D. Potts, and A. Livingstone, "Excavations at Thaj," pp. 86–108. *Atlat* 8:55–108. **Ha**
- Macdonald, M. 1979. "Safaitic inscriptions in the Amman Museum and other collections I." *Annual of the Department of Antiquities of Jordan* 23:101–119. **S**
- . 1986. "ABCs and letter order in Ancient North Arabian." *Proceedings of the Seminar for Arabian Studies* 16:101–168. **D, H, S**
- . 1991. "HU 501 and the use of *s*<sup>3</sup> in Taymanite." *Journal of Semitic Studies* 36:11–36. **T**
- . 1992a. "North Arabian epigraphic notes I." *Arabian Archaeology and Epigraphy* 3:23–43. **S, T**
- . 1992b. "On the placing of *ṣ* in the Maghribi *abjad* and the Khirbet al-Samrā' ABC." *Journal of Semitic Studies* 37:155–166. **H, S**
- . 1993. "Nomads and the Ḥawrān in the late Hellenistic and Roman periods: a reassessment of the epigraphic evidence." *Syria* 70:303–413. **S**
- . 1995a. "Safaitic." In *Encyclopaedia of Islam* (revised edition), vol. VIII, pp. 760–762. Leiden: Brill. **S**
- . 1995b. "Herodian echoes in the Syrian desert." In S. Bourke and J. P. Descœudres (eds.), *Trade, Contact, and the Movement of Peoples in the Eastern Mediterranean. Studies in Honour of J. Basil Hennessy*, pp. 285–290. Supplement to *Mediterranean Archaeology* 3. Sydney. **S**
- . 2000. "Reflections on the linguistic map of Pre-Islamic Arabia." *Arabian Archaeology and Epigraphy* 11:28–79. **D, H, Ha, OA, ONA, S, T, Th**
- . 2004a. "On the uses of writing in ancient Arabia and the role of palaeography in Studying them." *Arabian Archaeology and Epigraphy* 15. **D, G, H, S, Th**
- . 2004b. "A preliminary re-assessment of the scripts used in pre-Islamic Dedan." In S. Weninger (ed.), *Epigraphik und Archäologie des antiken Südarabien*. Wiesbaden: Otto Harrassowitz. **D**
- . 2004c. "Literacy in an Oral Environment." In P. Bienkowski, C. B. Mee and E. A. Slater (eds.), *Writing and Ancient Near Eastern Society*. British Academy Monographs in Archaeology. Oxford: Oxford University Press. **D, G, H, OA, S, Th**
- . 2004d. "From Dedān to Iran in four Ancient North Arabian inscriptions." In D. F. Graf and S. G. Schmidt (eds.), *Fawzi Zayadine Festschrift*. Amman: Department of Antiquities of Jordan. **D, H**
- . Forthcoming. *Old Arabic and its Rivals in the Age of Ignorance. Six Studies on the Emergence of Arabic as a Written Language*. **OA**
- Macdonald, M. and G. Harding. 1976. "More Safaitic texts from Jordan." *Annual of the Department of Antiquities of Jordan* 21:119–133. **S**
- Macdonald, M. and G. King. 1999. "Thamudic." In *Encyclopaedia of Islam* (revised edition), vol. X, pp. 436–438. Leiden: Brill. **H, T, Th**
- Macdonald, M., M. Al Mu'azzin, and L. Nehmé. 1996. "Les inscriptions safaitiques de Syrie, cent quarante ans après leur découverte." *Comptes rendus de l'Académie des Inscriptions et Belles-Lettres*, pp. 435–494. **S, Th**

- Milik, J. 1958–1959. “Nouvelles inscriptions sémitiques et grecques du pays de Moab.” *Liber Annuus* 9:330–358. **H**
- Müller, W. 1980. “Some remarks on the Safaitic inscriptions.” *Proceedings of the Seminar for Arabian Studies* 10:67–74. **S**
- . 1982. “Das Altarabische und das klassische Arabisch.” In W. Fischer (ed.), *Grundriß der Arabischen Philologie*, vol. I. *Sprachwissenschaft*, pp. 17–36. Wiesbaden: Dr. Ludwig Reichert. **D, H, Ha, OA, ONA, S, T, Th**
- . 1992. “Abimael.” In *Anchor Bible Dictionary*, vol. I, p. 20. New York: Doubleday.
- Müller, W. and S. Said. (2001). “Der babylonische König Nabonid in taymanischen Inschriften.” *Biblische Notizen* 107/108:109–119. **T**
- Rabin, C. 1951. *Ancient West-Arabian*. London: Taylor’s. **OA**
- Rabinowitz, I. 1956. “Aramaic inscriptions of the fifth century BCE from a North-Arab shrine in Egypt.” *Journal of Near Eastern Studies* 15:1–9. **G**
- Robin, C. 1974. “Monnaies provenant de l’Arabie du nord-est.” *Semitica* 24:83–125 (esp. pp. 112–118). **Ha**
- . 1994. “Documents de l’Arabie antique III.” *Raydān* 6:69–90. **Ha**
- . 2001. “Les inscriptions de l’Arabie antique et les études arabes.” *Arabica* 48: 509–577. **G, OA**
- Ryckmans, G. (ed.). 1950–1951. *Corpus Inscriptionum Semiticarum. Pars V*. Paris: Imprimerie Nationale. **S**
- Ryckmans, J. 1956. “Aspects nouveaux du problème thamoudéen.” *Studia Islamica* 5:5–17. **Th**
- Sass, B. 1991. *Studia Alphabetica. On the Origin and Early History of the Northwest Semitic, South Semitic, and Greek Alphabets*. Orbis Biblicus et Orientalis 102. Freiburg Schweiz: Universitätsverlag. **ONA**
- Scagliarini, F. 1996. “Šahr figlio di Han-Aws: il nome di un nuovo sovrano in un testo Liḥyanitico inedito.” *Studi epigrafici e linguistici sul vicino oriente antico* 13:91–97. **D**
- . 1999. “The Dadanitic inscriptions from Ġabal ‘Ikma in north-western Hejaz.” *Proceedings of the Seminar for Arabian Studies* 29:143–150. **D**
- Sima, A. 1999. *Die lihyanischen Inschriften von al-‘Udayb (Saudi-Arabien)*. Epigraphische Forschungen auf der Arabischen Halbinsel 1. Rahden/Westf.: Leidorf. [Note, I am not convinced by Sima’s interpretation of these inscriptions as referring to the maintenance of the underground water system of al-‘Ulā, hence my translations of the vocabulary of these texts frequently differ from his.] **D**
- . 2000. “Zum antiken Namen Dedan.” *Biblische Notizen* 104:42–47. **D**
- . 2002. “Die hasaitischen Inschriften.” In N. Nebes (ed.), *Neue Beiträge zur Semitistik*, pp. 167–200. Jenaer Beiträge zum Vorderen Orient 5. Wiesbaden: Harrassowitz. **Ha**
- Stiehl, R. 1971. “Neue liḥyānische Inschriften aus al-‘Uḏaib I,” mit einem Nachtrag M. Höfners. In F. Altheim and R. Stiehl (eds.), *Christentum am Roten Meer*, vol. I, pp. 3–40, 565–566, 569–594. Berlin: de Gruyter. **D**
- Ullendorff, E. 1965. “The form of the definite article in Arabic and other Semitic languages.” In G. Makdisi (ed.), *Arabic and Islamic Studies in Honor of H. A. R. Gibb*, pp. 631–637. Leiden: Brill. **G**
- Van den Branden, A. 1950. *Les inscriptions thamoudéennes*. Bibliothèque du Muséon 25. Louvain: Institut Orientaliste. **H, T, Th**
- . 1956. *Les textes thamoudéens de Philby*. Bibliothèque du Muséon 39 and 41. Louvain: Institut Orientaliste. **H, T, Th**
- Winnett, F. 1957. *Safaitic Inscriptions from Jordan*. Near and Middle East Series 2. Toronto: University of Toronto Press. **S**
- . 1980. “A reconsideration of some inscriptions from the Tayma area.” *Proceedings of the Seminar for Arabian Studies* 10:133–140. **T**
- Winnett, F. and G. Harding. 1978. *Inscriptions from Fifty Safaitic Cairns*. Near and Middle East Series 9. Toronto: University of Toronto Press. **S**
- Winnett, F. and W. Reed. 1970. *Ancient Records from North Arabia*. Near and Middle East Series 6. Toronto: University of Toronto Press. **D, H, T, Th S**
- . 1973. “An archaeological-epigraphical survey of the Ḥā’il area of northern Sa‘udi Arabia.” *Berytus* 22:53–113. **H, S, Th**
- Wright, W. 1896–1898. *A Grammar of the Arabic Language* (3rd edition, revised by W. Robertson Smith and M. J. de Goeje). Cambridge: Cambridge University Press [constantly reprinted]. **G**



# Indo-European

HENRY M. HOENIGSWALD AND

ROGER D. WOODARD

with a discussion of syntax by

JAMES P. T. CLACKSON

## 1. THEORETICAL AND HISTORICAL CONTEXTS

### 1.1 The comparative method

The parent language of the Indo-European linguistic family is an “ancient” language in a special sense: it is a *protolanguage*, not attested but reconstructed. Since a protolanguage is, broadly speaking, the collection of all *retentions* in the daughter languages, the ability to segregate *innovation* from retention in the latter is crucial for the reconstruction of the former. The “*comparative*” *method* (in the narrow, phonological sense of the term) accomplishes that segregation to a large extent (on the comparative method of historical linguistics, see also Ch. 45). Those innovations which we classify as *sound-changes* are capable of producing homophony among morphs; they are phonemic mergers, with the algebraic form  $/a/ > /m/$ ,  $/b/ > /m/$  (further elaboration is needed for conditioned sound-changes). Owing to the “Polivanov” property of sound changes (“no split without merger”), which follows from their definition as replacements statable in purely phonological terms (without reference, that is, to particular morphs), it is the case that if one phoneme, or one phonemic component (distinctive feature specification), or one phoneme combination (diphthong, cluster, syllable, etc.) in *language A* corresponds to one phoneme or phonemic component or phoneme combination in a related *language B* in one set of morphs, and to some other phoneme (etc.) in another set of morphs, then language A has in this detail innovated. As regards other details the converse may be the case, and language B may be the innovator. If A is found to have innovated in all details and B in none, A is a descendant (or later stage) of B and B the ancestor (or earlier stage) of A. In this case, language B may be predicted to have occurred in time before language A.

The comparative method aims at the recovery of the phonological shape of morphs. When it comes to morphemics – obsolescence, neologism, semantic change, borrowing, analogic change, and so forth – what is sometimes also called the comparative method is in reality something quite different (hence the preponderance of phonological subject matter in comparative work). The methods available for morphemic retrieval are much more akin to “comparison” in the everyday meaning of the word. They tend to rely on grammatical and lexical consensus and on resemblances and differences that do not by themselves, typological considerations aside, carry any clear-cut chronological implications. Extensive use, however, is made of internal reconstruction which operates not only with phonological alternations which result from conditioned sound-changes, but also with semantic isolation of forms in morphological and syntactic paradigms and the like.

## 1.2 Scholarly tradition

The conceit of related languages having their descent from a no longer spoken “parent language” is old (Metcalf 1974:251). For the Indo-European languages it was memorably voiced in 1786 by Sir William Jones (1746–1794), the justly admired and influential British jurist and scholar who served in India. Though without a marked intellectual interest in language as such, Jones was riding the crest of the new-found wave of enthusiasm (an enthusiasm in the creation of which he was himself a leading spirit) about things Indic. In matters of language he argued in traditional fashion from the “perfection” of Sanskrit, Greek, and Latin. Of the sober efforts directed at the Finno-Ugric languages by Strahlenberg (1676–1647) in 1730 and Sajnovics (1733–1785) in 1770 he was unaware.

The nineteenth and twentieth centuries witnessed the unfolding of the great work of filling old metaphors with a new technical content, not necessarily acknowledged in the abstract but abundantly clear from substantive, especially polemical, endeavor. Since the days of Wilhelm von Humboldt (1767–1835) proof of “relationship” in the form of carrying out convincing reconstructions has been provided for language families as diverse as Austronesian, Afro-Asiatic (including Semitic), Dravidian, Algonquian, among others.

## 1.3 Internal and external relations of Indo-European

Proto-Indo-European (PIE) may well have been spoken somewhere in the Black Sea area before the middle of the fifth millennium BC. At that time the speech community began to break up in a complex, long-drawn-out, and only partly recoverable process. The main branches which survived into historical times are (to list them in the chronological order of their first documentation): Anatolian (now extinct; see Chs. 18–23), Indo-Iranian (Chs. 26–30), Greek (Chs. 24–25), Italic (Chs. 32–33), Celtic (Ch. 35), Germanic (Chs. 36–37), Armenian (Ch. 38), Tocharian (extinct), Balto-Slavic, and Albanian (the three last-named being too recently attested for inclusion in the present volume). Additional Indo-European languages are attested in antiquity which do not clearly belong to any of these ten subfamilies, or whose membership is debated, such as Phrygian (Ch. 31), Venetic (Ch. 34) and Messapic (Ch. 1).

Once severed from one another, each branch went through changes that were largely but not entirely independent. Subgroupings based on the principle of shared innovation in the manner of the well-known *family tree* (German *Stammbaum*), or some other topological or geometrical scheme, will in general be discussed in the later chapters (noted above) which deal with the comparative evidence, that is, with the changes that define the descendant languages.

Proto-Indo-European is certain to have had outside connections of two kinds: (i) common descent from an anterior pre-protolanguage, and (ii) contacts recognizable from membership in areal typologies. Efforts to identify either kind have remained inconclusive.

# 2. PHONOLOGY

## 2.1 Consonants

The reconstructed consonantal inventory of Proto-Indo-European is comprised of obstruents (stops and fricatives), nasals, and sonorants (liquids and glides), as well as the so-called laryngeal consonants.

### 2.1.1 Obstruents

The stop phonemes of Proto-Indo-European, identified following established practice, are produced at five articulatory positions: (i) bilabial; (ii) dental; (iii) palatal; (iv) (pure) velar; (v) labiovelar. For each position, a (i) voiceless, (ii) voiced, and (iii) voiced aspirated stop is reconstructed:

(1)	<i>bilabial</i>	<i>dental</i>	<i>palatal</i>	<i>velar</i>	<i>labiovelar</i>
<i>voiceless</i>	p	t	ḱ	k	k <sup>w</sup>
<i>voiced</i>	b	d	ǵ	g	g <sup>w</sup>
<i>voiced asp.</i>	b <sup>h</sup>	d <sup>h</sup>	ǵ <sup>h</sup>	g <sup>h</sup>	g <sup>w</sup> <sup>h</sup>

The voiced bilabial *\*b* occurs only rarely. In the recently advocated “glottalic” view, the values of traditional *\*p*, *\*b*, and *\*b<sup>h</sup>* (etc.) are *\*p<sup>(h)</sup>* (aspirated, with unaspirated allophones), *\*p’* (voiceless glottalized), and *\*b<sup>(h)</sup>* (etc.) respectively; see Gamkrelidze and Ivanov 1995 and, for an evaluation, Watkins 1998:38.

The labiovelar phonemes *\*k<sup>w</sup>*, *\*g<sup>w</sup>*, and *\*g<sup>w</sup><sup>h</sup>* are distinct from the sequences *\*ḱw*, *\*ǵw*, *\*ǵ<sup>h</sup>w* (palatal stop + labiovelar glide) as well as perhaps from the sequences *\*kw*, *\*gw*, *\*g<sup>h</sup>w* (velar stop + labiovelar glide). Still, it is remarkable that something like the geminate prohibition (see §2.3) neutralizes labiovelars and velars before [u], with outcomes that are those of the velars. This is especially visible in post-Mycenaean Greek where *\*k<sup>w</sup>a* gives Attic [pa] but *\*k<sup>w</sup>u* yields [ku] (see Ch. 24, §3.7.1).

On the basis of the evolutionary outcome of the Proto-Indo-European palatal, velar, and labiovelar stops, Indo-Europeanists have traditionally divided the Indo-European daughter languages into two major groups, labeled *centum* (Latin for “100”) and *satem* (after Avestan *satəm* “100”; both forms from PIE *\*ḱmtom*). The general case is that western Indo-European (centum) languages merge the palatal and velar stops, whereas in the eastern (satem) dialects, the palatal stops exhibit distinct reflexes while the velars and labiovelars fall together (see Melchert 1987). The conspicuous exception to this distributional pattern is provided by Tocharian. Spoken far to the east in antiquity (with documentary remains surviving in the deserts of Chinese Turkestan or Xinjiang Uygur), Tocharian shows the centum treatment of back consonants.

Proto-Indo-European possessed the dental sibilant *\*s*, presumably with allophones [s] and [z], the latter occurring before plain voiced and voiced aspirated obstruents. The occurrence of an interdental fricative /þ/ has long been proposed to account for that stop/fricative correspondence seen in cognates such as, for example, Greek ἄρκτος (*árktos*) and Sanskrit *īkṣa-*, “bear,” but this remains problematic as another, more sophisticated solution has been proposed.

### 2.1.2 Sonorants

The Proto-Indo-European sonorant phonemes occur as both nonsyllabic and syllabic allophonic variants (see §2.1.4):

(2)	<i>nasals</i>	<i>liquids</i>	<i>glides</i>
	n/ṇ	r/ṛ	y/i
	m/ṁ	l/ḷ	w/u

### 2.1.3 Laryngeals

Those consonantal sounds identified as “laryngeal” likewise occur in nonsyllabic and syllabic forms:

(3)  $h_1/\text{ə}_1$   $h_2/\text{ə}_2$   $h_3/\text{ə}_3$

(for other notations and other views, see Watkins 1998:40). Phonetically, these are, to judge from their comportment in conditioned sound change in the descendant languages, neutral ( $h_1$ ), *a*-colored ( $h_2$ ), and *o*-colored ( $h_3$ ), respectively. The nonsyllabic allophones of the first two laryngeals seem to be voiceless; that of the third, voiced.

### 2.1.4 Vocalic versus consonantal

The full-grade vowels (see §2.2), the long vowels, the syllabic allophones of glides and laryngeals, and the diphthongs will henceforth be referred to, when convenient, as *vocalics*; nonvocalics are *consonantals*.

### 2.1.5 Nonsyllabic versus syllabic

In certain respects the three laryngeals resemble the sonorants. The resemblance is weakened and tends to disappear in the descendants. Very roughly, the following holds:

1. After a full-grade vowel (see §2.2) and preceding a consonant, both the sonorants and the laryngeals appear in their *nonsyllabic* shapes, the sonorant combinations forming diphthongs and the laryngeal combinations merging in the descendants (if not earlier), for the most part, with the long vowels. Similarly, syllabic  $*\bar{i}$  and  $*\bar{u}$  with a following laryngeal generate  $*\bar{i}$  and  $*\bar{u}$ . The syllabic allophones of the liquids and nasals lead to different results in the descendants.
2. Unless following a full-grade vowel, sonorants and laryngeals preceding a consonantal appear in their *syllabic* shapes. However, special provisions require certain sonorants and certain laryngeals to appear word-initially in nonsyllabic form when followed by certain nonsyllabic sonorants which are followed in turn by vowels, so as to form an initial sonorant cluster (e.g.,  $*\#[wr-]$ ). In the descendant languages, laryngeals in their syllabic shapes end up merged with the full-grade vowels and their outcomes ( $*h_1 = e$ ,  $*h_2 = a$ ,  $*h_3 = o$ ) – once again a process that may have commenced in Proto-Indo-European.
3. Word-medially when occurring after the sequence short vowel + one consonant and before a vocalic (VC \_\_\_\_ [+vocalic]), sonorants appear in their *nonsyllabic* shape (algebraically,  $..et[y]e..$ ). When occurring after the sequence vowel + two consonants, or long vowel + one consonant, and before a vocalic ( $\{VCC \text{ or } VVC\} \text{ ____ } [+vocalic]$ ), sonorants appear in their *syllabic* shape (i.e.,  $..ekt[i]e..$ , *Sievers' Law*). After a single word-initial consonant, syllabic and nonsyllabic shapes both occur – generalized from occurrences after a preceding word-final vowel or word-final consonant respectively (i.e.,  $..\#t[y]e..$ ,  $.. \#t[i]e..$ ).

## 2.2 Vowels

The Proto-Indo-European vowel inventory consisted of the “full-grade” *short* vowels  $*e$ ,  $*o$ , and  $*a$ , as well as  $*\bar{i}$  and  $*\bar{u}$ , the syllabic allophones of the glides  $*y$  and  $*w$  (see §2.1.2); and

the “lengthened-grade” long vowels \*ē, \*ō, and \*ā, plus long \*ī and \*ū. The resulting vowel systems, short and long, were thus:

(4)

	short		long	
	front	back	front	back
high	i	u	ī	ū
mid	e	o	ē	ō
low	a		ā	

Moreover, there occurred the automatic syllabic outcropping [ɐ] between obstruents, known as *schwa secundum* (see §2.4, 3.1).

## 2.3 Phonotaxis

Various phonotactic constraints limit the permissible sequences of sounds in Proto-Indo-European (see also §3.3):

1. There are no geminates. Geminate clusters arising across morpheme boundaries were simplified: for example, \**h<sub>1</sub>és-si* “you (sg.) are” yields \**h<sub>1</sub>ési*, as in Sanskrit *ási* (though a marginal process of gemination creates hypocoristic by-forms of personal names and the like; see Watkins 1998:40). The sequence \*..*t-t*.. was, however, analogically restored.
2. There are no clusters (*hiatus*) of full-grade vowels, both like and unlike. Where such sequences arise at morpheme boundaries, the vowels are contracted into long vowels bearing a distinctive accent in some descendant languages (the “long diphthongs,” where they are not contraction products [as in, e.g., the thematic dative singular ending, see §3.5.3], pose difficult problems).
3. Obstruent (and *s*) sequences are entirely voiced or entirely voiceless. If a voiceless and a voiced or voiced aspirated obstruent abut at a morpheme boundary, regressive dissimilation will take place. It is likely, by the same token, that the distinction between the three manners of articulation was neutralized, phonetically in favor of voicelessness, before a word boundary (see §2.5). The word-final sequences \*-*ms*# and \*-*ns*# are likewise neutralized (Leumann 1977:415); this is relevant for the animate accusative plural ending; see n. 36.
4. *Bartholomae’s Law* specifies that “if the first member of an obstruent cluster is ... aspirated, the assimilation is progressive” (see Watkins 1998:40–41).

## 2.4 Syllabicity

There are hints of an overarching principle governing syllabicity. This principle is accessible only in a schematically simplified and chronologically flat form which fails to convey the sliding nature of the scale along which developments took place, and which stretches from a remote past well into the era of the descendant languages. While most of the evidence is Indo-Iranian and Greek, it testifies nevertheless to a state of affairs that is essentially Indo-European. It is likely that syllabicity largely falls out in such a way as to preclude the accumulation of more than two consonantals in the flow of speech (with a *word boundary* as well as the sibilant *s* playing an uncertain role; see Beekes 1982:110) – hence, before vowels, *Sievers’ Law* (..*et*[*y*]e, ..*et*[*r*]e but ..*ekt*[*i*]e, ..*ekt*[*r*]e; see §2.1.5) as modified by *Lindeman’s Law* (which regulates word-initial obstruent + sonorant clusters; see Lindeman 1965).

There could well have existed an *Extension of Sievers' Law* before consonantals and before a pause if—as the surviving difference between Greek (Attic-Ionic) ὅφρα (óph<sup>h</sup>ra) “in order that” (with one short vowel and one consonant preceding) and ἥπαρ (hêpar) “liver” (with one long vowel and one consonant preceding) suggests—the allophonic notation \*[ɾ] stands for two quite distinguishable allophonic entities: \*[ɾ<sub>e</sub>], with (it may be imagined) increasingly prominent syllabicity (in ὅφρα); and \*[ɾ], with syllabicity decreasing to the rightward (in ἥπαρ).

In word-initial syllables where the determining environment is not built in, one would expect vacillation between [ra] and [ar], with a potential for mutual analogic exchanges and generalizations. This is indeed what one finds: for example, in Homeric κραδίη (*kradiē*) beside καρδία (*kardía*) “heart.”

In Greek, as in other descendant languages, this \*[ɾ<sub>e</sub>] adjacent to liquids and nasals became phonemic by merging with some existing vowel (in Attic-Ionic with [a]). In the case of Indo-European \*[y/i] and \*[w/u] (these from the oldest period), and (much later) Indo-Iranian \*[r/ɾ] (Sanskrit .../r [but \*[ɾ] > Sanskrit *ir* before vocalics (...*aktira*...) under Sievers' Law proper; i.e., not the Sievers' Law “Extension”), Avestan .../ərə), the three pairs of two positional variants are transformed into one segment each, perhaps of steady (i.e., neither increasing nor decreasing) vowel-like quality. Under similar circumstances [ɾ<sub>e</sub>] in the vicinity of obstruents can end up phonemic in the descendant languages by merging with one of the existing vowels, though here the data remain shadowy. As a result of all of this, overlong syllables (short vowels with more than two consonants, or long vowels or diphthongs with two consonants before the next vowel) are rare, for example in Vedic and in Greek, until sound changes create new overlengths (see Hoenigswald 1994 for the details; lengthened grade [see §3.2] in certain formations is [still?] extremely rare in Sanskrit before consonant clusters; see Debrunner 1954:61).

The phenomena treated above militate in their own typological way in favor of the retentive nature of pitch accent and quantitative meter; see §§2.6, 2.7.

## 2.5 Word boundaries

Word boundaries (i.e., seams between so-called minimum free forms; see Hoenigswald 1992) loom large as conditioning factors in sound changes. So far from indicating, however, that all word boundaries are phonologically marked and contrast with Ø in *word-interior* position (note §2.4 on phonetic conditioning across a word boundary), word boundary is best considered an analogical development made possible by the circumstance that *pause* (the absence of sound which contrasts with the presence of sound, a universal condition) is an option at word boundaries. Post-pausal and ante-pausal allophony was generalized and turned into apparent word-initial and word-final phonology, each contrasting with word-interior phonology. The descendant languages differ somewhat in the extent to which this analogic change is carried through. Where analogic generalization is complete, utterances may indeed be treated as “composed of” (rather than “analyzed into”) words in *external sandhi* (some of the sandhi phenomena of Insular Celtic may be relevant survivals – see Russell 1995; sandhi phenomena were, however, created again and again in the separate branches).

## 2.6 Accent

The fragmentary character of the scripts in which the texts of the descendant languages are recorded, combined with the neglect of relevant phenomena despite their syntactic

centrality, have prevented deciding whether *phonemic stresses* forming stressed morphs existed, let alone reconstructing them. For some daughter languages metrical indications are available but have scarcely been exploited. Such a determination would be of paramount importance for syntax. Much of syntax is customarily discussed, *faute de mieux*, in terms of *word order*. In many languages, however, word-order phenomena (recognizable in the texts) are correlated with, or even dependent on, stress phenomena (ignored in the texts); see Hoenigswald 1980.

A lexical *word accent* (/') – likely a *pitch* accent – contrasted with the absence of accentuation. Such an accent may be reconstructed from Vedic Sanskrit, Greek, Anatolian, Balto-Slavic, and from the effect it had in Germanic (*Verner's Law*; see Ch. 36, §3.6.2). Clitics were unaccented, enclitics occupying the second place in a clause (*Wackernagel's Law*; see Szemerényi 1996:81–82, with references).

Little is known about sentence intonations. It is possible, though unlikely, that the fixed high pitch of the question pronoun in Greek τίς, τί (*tís, tí*) represents the survival of an Indo-European interrogative intonation.

## 2.7 Meter

It is uncertain whether Proto-Indo-European meter is quantitative in nature and based on the characteristics of syllables, as it is in Sanskrit and in Greek, or whether these two daughter languages have innovated (so Watkins 1995:21). The absence of any metrical function for word accent in these two branches is often associated with quantitative meter, whether retained or innovated. Verner's Law in Germanic (see §2.6) as well as the dependence of the ablaut zero-grade (see §3.2) upon lack of accent seem to point to an original strongly “dynamic” character for word accent; see Lehmann 1952:109.

# 3. MORPHOLOGY

## 3.1 Word formation

The morphology of nouns/adjectives (including pronouns) and verbs, comprises derivation, inflection, and compounding. A single *root*, minimal or extended (see §3.3), precedes a derivational *suffix* or suffix sequence (or accommodates the *\*-n/ne-* infix) which, in tandem with syntactic function, define the resulting “word” (marked, as often as not, by the incidence of accent) as a *noun* or *verb*. The resulting root + affix complex is a *stem*, though in some instances the root alone can function as a stem. In compounding (always binary), noun stems combine to form more complex noun stems. Verbs are not in that sense capable of compounding. Stems in turn are followed by a single nominal or verbal *inflectional ending* which likewise contributes to syntactic identification. The paradigms that result in this synthetic structure are close-knit and, especially insofar as the endings are concerned, characterized by well-recognizable and clear-cut allomorphies.

## 3.2 Ablaut

Proto-Indo-European ablaut, or *apophony*, originally depended on word accent (see §2.6) in ways which are only in part transparent. The phenomenon is a pervasive, nonautomatic, morphologically conditioned alternation of the vowels of (5):



(5)	<i>ablaut vowel</i>	<i>designation</i>
	<i>e</i> (and infrequently <i>a</i> )	full-grade, or simply <i>e</i> -grade
	<i>o</i>	<i>o</i> -grade
	Ø	zero-grade
	<i>ē, ō</i> (and infrequently <i>ā</i> )	lengthened-grade

In the case of the zero-grade, accumulations of obstruents tend to be relieved by [e], the so-called *schwa secundum*. Processes such as, perhaps, the internally reconstructed sound change *\*..ers# > \*..ēr#* produce the lengthened-grade vowels; see Szemerényi 1996:115–116. If the derivative process known in Sanskrit as *vr̥ddhi* (see Ch. 26, §3.4.3) goes back to the Proto-Indo-European period, it is another source of lengthened-grade vowels.

### 3.3 Root structure

Minimal roots consist of two consonants (i.e., phonemes other than full-grade and lengthened-grade vowels):  $C_1 \dots C_2$ . Minimal roots may also be extended to form structures of three and four consonants:  $C_1 \dots C_2 \dots C_3 (\dots C_4)$ , always subject to phonological constraints in accordance with the the sonority of their components. Taken together with ablaut, and observing the rule that full-grade vowels (here represented by *e*) can occur only once within a root, the following varieties exist: (i) for  $C_1 C_2$ :  $C_1 e C_2$ ; (ii) for  $C_1 C_2 C_3$ : (a)  $C_1 e C_2 C_3$ , (b)  $C_1 C_2 e C_3$ ; (iii) for  $C_1 C_2 C_3 C_4$ :  $C_1 C_2 e C_3 C_4$  (see Watkins 1998:53, following Benveniste 1935 *passim*; there may be a few roots with initial full-grade vowels, but many roots which appear to fall into this category are in fact to be reconstructed with an initial laryngeal). In a given root,  $C_1$  may freely alternate with  $sC_1$  (*s mobile*) devoid of semantic function.

In addition, the initial and the final obstruents of roots with or without extensions are subject to a set of highly compact *compatibility rules* or *root constraints*. With insignificant exceptions, the initial and the final phoneme of a root must not be the same (note that this prevents the zero-grade from creating a geminate cluster [see §2.3]; in the case of minimal roots, not even the places of articulation of  $C_1$  and  $C_2$  are permitted to be the same): thus, roots of the form *\*nēn*, *\*tēt*, *\*tērt*, *\*d<sup>h</sup>ed<sup>h</sup>*, *\*d<sup>h</sup>ed* are excluded. Voiced obstruents do not occur with one another; neither do voiceless obstruents occur with voiced aspirated obstruents (*\*bed*, *\*b<sup>h</sup>et*, *\*ped<sup>h</sup>*, *\*perd<sup>h</sup>*, etc.). In contrast, (i) voiceless obstruents can co-occur, (ii) as can voiced aspirates, (iii) and voiceless obstruents can occur with voiced obstruents, (iv) and voiced obstruents with voiced aspirated: thus, *\*pet*, *\*ped*, *\*bet*, *\*b<sup>h</sup>ed<sup>h</sup>*, *\*b<sup>h</sup>ed*, *\*b<sup>h</sup>erd*, *\*b<sup>h</sup>end*, and so forth (but not *\*ted*). For an organization of these constraints, see Hoenigswald 1954:469, n. 2.

### 3.4 Athematic versus thematic

Noun/adjective and verb morphology show a thoroughgoing parallelism between *athematic* and *thematic* formation. The latter exhibits a stem suffix *e ~ o* (*o* before endings with *-m ...*) preceding the inflectional ending, whereas the former has no such vowel. Athematic formations frequently exhibit a play of ablaut in root, suffixation, and ending (associated with accent; for a critique of the classificatory schemes proposed to deal with accent in inflectional noun paradigms, see Watkins 1998:62), while the thematic vowel tends to freeze accent and ablaut.



### 3.5 Nominal morphology

Under this heading can be treated both nouns and adjectives, as well as pronouns. As one goes back in history, the difference between noun and adjective tends to lessen. A noun has one gender as an inherent characteristic. A given adjective, on the other hand, aside from its syntactic and semantic standing as attribute or predicate and as a counter for the rules of grammatical agreement, is defined, in most of the descendant languages at least, by the fact that it occurs in all three genders. For example, derivative suffixation, as it serves to create feminines (once these are established) from some masculines, becomes a part of the paradigm for any adjective.

#### 3.5.1 Derivation

Nominal (noun/adjective) derivation by means of suffixes (see §3.1), including simply the thematic vowel itself, is either *primary* (directly from the root) or *secondary* (from a stem). Nominal suffixes range from (i) athematic (including -Ø-, in the case of root nouns, with inflectional endings attached directly to the root, which thus serves as the stem); to (ii) thematic suffixes (i.e., suffixes ending in the thematic vowel; see §3.4); to (iii) the suffix *\*-eh<sub>2</sub>* (and the ablauting *\*-yeh<sub>2</sub>* [e-grade], *\*-ih<sub>2</sub>* [zero-grade]) which became completely recast as the sign of feminines and collectives in the descendants. Stems formed with athematic suffixes have been traditionally classified by the final segment of the suffix, for example:

(6)	<i>stem-class</i>	<i>nominative singular</i>	<i>genitive singular</i>	
	<i>t-stems</i>	*nók <sup>w</sup> -t-s	*nék <sup>w</sup> -t-s	“night”
	<i>r-stems</i>	*ph <sub>2</sub> -tér	*ph <sub>2</sub> -tr-és	“father”
	<i>n-stems</i>	*tér-mṇ	*tér-mṇ-s	“boundary”
	<i>i-stems</i>	*mén-ti-s	*mṇ-teí-s	“mind”
	<i>u-stems</i>	*pér-tu-s	*pr-teú-s	“a crossing over”

For a full discussion of derivational suffixes, see Watkins 1998:62–65.

There are two processes that compete with suffixation. One is *accent shift*; the contrast between Sanskrit *bráhman-* (neuter), the religious concept, and *brahmán-* (masculine) “singer, etc.” seems to be old. The other is *compounding*.

Both compounds and secondary derivation by suffix are, on the whole, exocentric rather than simply determinative. In compounds, while the first stem may indeed be said to modify the second, the compounding itself has a derivational function: Sanskrit *bahu-* means “much” and *vrihi-* “rice,” but *bahu-vrihi-* is not simply “much rice” but “having much rice” (see Ch. 26, §4.4.2.3). In consequence, certain secondary suffixes indicating “having” and the compound construction are complementary to each other. In Greek terms, θεός (*t<sup>h</sup>eós*) “god,” suffixed θε-ϊ-ος (*t<sup>h</sup>e-ï-os*) “divine,” but compounded θεο-ειδής (*t<sup>h</sup>eo-eidés*) “having a god’s appearance,” and **not** \*θε-ι-ο-ειδής (*t<sup>h</sup>e-i-o-eidés*), on a par with πολῦ-μητις ([polū-me:ti:s]) “of many counsels” (cf. Skt. *bahu-vrihi*), even though both θεῖος (*t<sup>h</sup>eíos*) and πολὺς (*polús*) are attributive adjectives.

In secondary derivation by suffix, too, mere modification of meaning, as in diminutives, pejoratives, augmentatives, and so forth, is very rare. To continue the preceding example, Greek θεῖος (*t<sup>h</sup>eíos*) is, in fact, typical: it refers not to some sort of “god” but to an outside person or object characterized by gods.

This relationship extends to the process of *internal derivation* by a rightward shift of word accent, which turns some athematic nouns into possessive adjectives. For example, *\*krétu-* “strength” yields *krtú-* “strong”; see Watkins 1998:62 and Schindler in Nussbaum 1998:14.

The secondary comparative in *\*-tero-*, going back to a primary suffix to express opposing attributes (“other,” etc.; cf. Latin *alter* “the other of two”), which later, in some descendants, competes uneasily with primary formations, is a notable exception to the foregoing generalization.

### 3.5.2 The Caland System

Recognition of “Caland” suffixation represents an insight of an unusual kind. A set of suffixes is distributed in such a way that the presence of one (in one semantic function) implies, almost to the point of predictability, the existence of some or all other members of the set (in other semantic functions). Thus, in Greek, adjectives in *-(e)rós-* (e.g., *κῦδ-ρός* (*kud-rós*) “famous”; *κρᾶτ-ερός* (*krat-erós*) “powerful”) or *-ús* (e.g., *κρᾶτύς* (*kratús*) “strong”) go together with neuter nouns in *-es/-os* (*κάρτ-ος* (*kárt-os*) “strength”; *κῦδ-ος* (*kúd-os*) “fame,” etc.); with the primary comparatives; with first compound members in *-i-* (*κῦδι-άνειρα* (*kudi-áneira*) “of famed men” fem.); and so forth. On the Caland System, see Risch 1974:65–97; 208.

### 3.5.3 Nominal endings

Noun/adjective stems are followed by declensional endings in which the categories of (i) *number* (singular, dual, plural) and (ii) *gender* (once animate and neuter; then masculine, feminine, and neuter) – these two being really derivational – as well as (iii) *case* (eight in number; see Table 17.1) are fused, with few or no hints at a more agglutinating prehistory (the animate accusative plural ending, *\*-ns*, perhaps was built from accusative *\*-m* [as in the singular] plus the plural *\*-s*). These endings, insofar as they can be retrieved with any assurance, are presented in Table 17.1 (cf. Watkins 1998:66):

**Table 17.1 Proto-Indo-European nominal endings**

	Athematic	Thematic
<i>Nominative</i>	*-s	*-o-s
<i>Vocative</i>	*-Ø	*-e
<i>Accusative</i>	*-m	*-o-m
<i>Nom./Acc. neuter</i>	*-Ø	*-o-m
<i>Genitive</i>	*-es/-os/-s	*-o-s/-o-s(y)o
<i>Ablative</i>	*-es/-os/-s	*-o-h <sub>2</sub> ed
<i>Dative</i>	*-ei	(*-o-ei>) *-ōi
<i>Locative</i>	*-i; *-Ø	*-e/o-i
Dual		
<i>Nom./Acc.</i>	*-h <sub>1</sub>	*-o-h <sub>1</sub>
Plural		
<i>Nom./Voc.</i>	*-es	*-os
<i>Accusative</i>	*-ms	*-o-ms
<i>Nom./Acc. neuter</i>	*-h <sub>2</sub>	*-e-h <sub>2</sub>
<i>Genitive</i>	*-om	(*-o-om>) *-ōm
<i>Dat./Abl.</i>	*-b <sup>h</sup> (y)os; *-mos	*-o-b <sup>h</sup> (y)os; *-o-mos
<i>Locative</i>	*-su	*-oisu
<i>Instrumental</i>	*-b <sup>h</sup> is; *-mis	*-ōis

### 3.5.4 Pronouns

Pronouns may be classified superficially into (i) personal pronouns and (ii) the various pronominal adjectives and adverbs that form well-integrated derivational and inflectional paradigms.

Among the *personal pronouns* it seems possible to reconstruct these nominatives:

- |     |  |             |
|-----|--|-------------|
| (7) | $*(h_1)e\dot{g}oh_2$ , $*(h_1)e\dot{g}h_2om$ | “I”         |
|     | $*tuh_2$                                     | “you” (sg.) |
|     | $*weis$ , $*h_1\eta smes$                    | “we”        |
|     | $*yutts$ , $*h_1usmes$                       | “you” (pl.) |

The other cases have each an orthotone and an enclitic variant. There is also the much remarked-on suppletion in the first-person singular paradigm between the nominative stem and the oblique case forms with initial  $*m-$ . The reconstruction of all these forms is complex and problematic; see Rix 1976:177–180, Szemerényi 1996:216–218.

A *reflexive* stem  $*s(w)e/o-$  is used for all three persons.

*Possessive* pronouns are thematic derivations based upon the personal pronouns. *Demonstratives* are a mixture of indeclinable particles and adjective-like paradigms built on the latter. Limiting this presentation again to the nominative (singular) forms, the conglomerate particle  $*so$  “and he” (maintained as such in Hittite) and the neuter  $*to-d$  (with the characteristic neuter singular ending that distinguishes pronouns from ordinary adjectives; cf. Latin neuter *aliud* “other”) combined in the non-Anatolian descendants to form a suppletive thematic paradigm: masculine  $*so$  (feminine  $*seh_2$ ), neuter  $*tod$ , preserved, for instance, as the Attic Greek “definite article,”  $\delta$  (*ho* – without a nominative ending!; feminine  $\eta$  (*hē*)),  $\tau\acute{o}$  (*tó*). The *interrogative* stems are  $*k^wo-$  and  $*k^wi-$  (it is a characteristic of pronominal inflection that thematic stems and *i*-stems can exist side by side); when enclitic, these serve as indefinites. In the *relative* function,  $*k^wo-/k^wi-$  competes with  $*(h_1)yo-$  which is possibly derived from the demonstrative  $*h_1i-$  (as in Latin *is* “that one”).

## 3.6 Verbal morphology

### 3.6.1 Derivation

Verb-stems carry derivational affixes – often governed by *principles* which duplicate the corresponding processes in noun formation (see §3.5.1; also §3.1). Affixes utilized in verb-stem formation include: (i) athematic and thematic ( $*-e/o-$ ) suffixes; (ii) both denominative and nondenominative  $*-ye/o-$ ; (iii) the nasal infix  $*-n/ne-$ ; (iv) the  $*-s-$  of the “sigmatic aorist”; (v) the iterative suffix  $*-ske/o-$ ; (vi) the thematic vowel itself as sign of the *subjunctive* mood; (vii) the *optative* suffix  $*-yeh_1/i\eta_1-$  (placed immediately before the ending; thus in athematic paradigm after the thematic vowel: 3rd sg. pres. act.  $*b^h\acute{e}r-o-y\eta_1-t >$  Gk.  $\phi\acute{\epsilon}ροι$  (*phéroī*) “may (s)he carry,” matching the indicative  $\phi\acute{\epsilon}ρει$  (*phérei*)); (viii) the thematic  $*-se/o-$  of some futures (a doubtful case for the parent Indo-European language, but so used among daughters); (ix) as well as reduplication; and (x), in athematic subparadigms, the play of ablaut.

These affixations are distributed over the *voices* (active and middle), *tenses* (non-perfect and perfect), *moods* (indicative, subjunctive, optative, injunctive, imperative), and *persons* (first, second, and third; with *numbers*, singular, dual, plural) of finite verbs in complicated but well-delineated patterns. In some of the more conservative descendants a given verb appears with paradigmatically predictable forms in (nearly) all the intersections of the categories named (e.g., “2nd-person plural, subjunctive, present, middle...”). The protolanguage is not like that. Seen from that more familiar standpoint, only certain particular

portions of the paradigm seem filled – in ways, however, that lend themselves to coherent and convincing internal reconstruction.

### 3.6.2 Verb endings

Verbs are inflected for the categories named above. In main clauses verbs are enclitic; in dependent clauses and under certain other conditions they are orthotone. Some of the *active* personal endings (personal endings being what makes these constructs “finite” forms, as distinct from participles – infinitives developing only in the descendant languages) are given in (8)–(10), for singular and plural only, and with the added category of *secondary* (unmarked) versus *primary*, the latter perhaps with an added morph, *\*-i*, the so-called *hic et nunc* particle (see Watkins 1998:60–62; “secondary” and “primary” endings are to be distinguished from secondary and primary affixation in noun derivation [see §3.5.1]; the homonymy is unfortunate). More loosely attached is the so-called augment *\*h<sub>1</sub>e-*, optionally prefixed to past tense indicatives, which survives in a number of descendants:

(8)		<i>Athematic</i>		<i>Thematic</i>	
		<i>Primary</i>	<i>Secondary</i>	<i>Primary</i>	<i>Secondary</i>
<i>Singular</i>	1.	-mi	-m	-o-h <sub>2</sub> ei	-o-m
	2.	-si	-s	-e-si	-e-s
	3.	-ti	-t	-e-ti	-e-t
<i>Plural</i>	1.	-me	-me	-o-me	-o-me
	2.	-te	-te	-e-te	-e-te
	3.	-enti	-ent	-o-nti	-o-nt

The athematic inflection appears to have exerted a strong influence on the thematic. A first-person singular primary thematic *\*-o-mi* can also be reconstructed for a common Indo-European stage. In addition, for the thematic inflection, earlier second- and third-person singular forms have been reconstructed:

(9)		<i>Primary</i>	<i>Secondary</i>
<i>Singular</i>	2.	-e-(th <sub>2</sub> e)i	-e-(th <sub>2</sub> e)
	3.	-e-i	-e

Distinct endings for the active *imperative* are reconstructed as follows:

(10)		<i>Athematic</i>	<i>Thematic</i>
<i>Singular</i>	2.	Ø, -d <sup>h</sup> i	-e-Ø
	3.	-tu	-e-tu
<i>Plural</i>	3.	-entu	-o-ntu

A similar array may be assembled for the *middle* voice, though there is considerable uncertainty regarding the forms of the first and second plural in the protolanguage:

(11)		<i>Athematic</i>		<i>Thematic</i>	
		<i>Primary</i>	<i>Secondary</i>	<i>Primary</i>	<i>Secondary</i>
<i>Singular</i>	1.	-h <sub>2</sub> ei	-h <sub>2</sub> e	-o-h <sub>2</sub> ei	-o-h <sub>2</sub> e
	2.	-th <sub>2</sub> ei, -soi	-th <sub>2</sub> e, -so	-o-soi	-o-th <sub>2</sub> e, e-so
	3.	-oi, -toi	-o, -to	-o-i	-o, -e-to
<i>Plural</i>	3.	-ontoi	-onto	-o-ntoi	-o-nto

The *perfect* has no distinction of voice. It is largely reduplicated; its endings, insofar as they can be clearly reconstructed, are as follows:

(12) <i>Singular</i>	1.	-h <sub>2</sub> e
	2.	-th <sub>2</sub> e
	3.	-e
<i>Plural</i>	1.	-me
	2.	-e
	3.	-r

Two examples must suffice to illustrate some of the inflectional processes at work.

1. The verb “to go” is an athematic root present (i.e., the root itself serves as the present tense stem, without a suffix attached; see §3.6.1) with ablaut. Its constructs for the singular and plural of the indicative are *\*(h<sub>1</sub>)éi-mi*, *\*(h<sub>1</sub>)éi-si*, *\*(h<sub>1</sub>)éi-ti*; *\*(h<sub>1</sub>)i-més*, *\*(h<sub>1</sub>)i-té*, *\*(h<sub>1</sub>)i-énti*.
2. The verbs *\*l(e)ik<sup>w</sup>*- “leave” and *\*p(e)uh<sub>2</sub>*- “purify” form an indicative present from their zero grade with the ablauting nasal infix *\*-n(é)-*: thus, third singular active *\*li-né-k<sup>w</sup>-ti*, *\*pu-né-h<sub>2</sub>-ti*; third plural *\*li-n-k<sup>w</sup>-énti*, *\*pu-n-h<sub>2</sub>-énti* (giving Vedic Sanskrit *riṇákti*, *punáti*; *riñcánti*, *punánti*; see Watkins 1998:57).

### 3.6.3 Participles

There are four participles or participle-like verbal adjectives: one mostly primary, formed in *\*-tó-* (generally middle in meaning; e.g., Gk. κλυ-τό- (*klu-tó-*) “famous”), and three mostly secondary: (i) active, formed in *\*-nt-* (e.g., Gk. δό-ντ- (*dó-nt-*) “giving,” φέρ-ο-ντ- (*phér-o-nt-*) “carrying”); (ii) middle, in *\*-mh<sub>1</sub>n-o-* (*\*[mḁ<sub>1</sub>no-]*, *\*[-ṁh<sub>1</sub>no-]*; e.g., Gk. φερ-ό-μεν-ο- (*phér-ó-men-o-*) “being carried”); and (iii) perfect, in *\*-w(o)s-* (e.g., Gk. nom. masc. πε-ποιθ-ώς (*pe-poit<sup>h</sup>-ós* < earlier *\*pe-poit<sup>h</sup>-wós*), fem. πε-ποιθ-ῶϊα (*pe-poit<sup>h</sup>-uî-a* < earlier *\*pe-poit<sup>h</sup>-us-ya*) “trusting”).

## 3.7 Adverbs

Adverbs may be primary, even unanalyzable, or else derived – most typically from adjective stems. The forms more often known from some descendant languages in their function as prepositions or postpositions were adverbs that occurred in close syntactic construction with nouns/adjectives and verbs. They enter into compounds – *bahuvrīhi* compounds (see §3.5.1) – as first members, very much on a par with noun stems. A bit of derivational paradigm from Greek will illustrate not only their formal and semantic properties but also those of a number of prefixes such as the negative *\*h<sub>1</sub>n-* (Gk. ἄ- (*a-*), ἀν- (*an-*)), zero-grade of the sentence negation *\*h<sub>1</sub>né-*: πολύ-θεος (*polú-t<sup>h</sup>eos*) “belonging to many gods”; ἔν-θεος (*én-t<sup>h</sup>eos*) “having the god within, inspired”; ἄ-θεος (*á-t<sup>h</sup>eos*) “without a god.”

## 4. SYNTAX

The twentieth century saw a fundamental revision of the reconstructed phonology and morphology of Proto-Indo-European, but much of the nineteenth-century scholarship on reconstructed syntax, notably Delbrück (1893–1900) and Wackernagel (1926), is still standardly cited in books and articles, including this one, and their work is the starting point

for much current research – witness the volumes edited by Eichner and Rix (1990) and Crespo and García Ramón (1997). Although some writers take the resilience of Delbrück and Wackernagel's work as an indictment of more recent, and more transient, scholarship, it rather shows widespread agreement over many of the fundamentals of reconstructed syntax. Much of what we know about Indo-European syntax is tacitly assumed in morphological reconstruction: there were three numbers – singular, dual, and plural (on the “collective” see further below); adjectives show concord in number, gender, and case with their head noun; subject pronouns are not obligatorily present, but are encoded in the verbal inflections; case inflections marked both grammatical roles and local relations; and verbs are marked for mood and voice as well as tense (with certain restrictions, see §3.6 above).

Indeed, the reconstruction of any morphological category makes tacit assumptions about the syntax. Thus, the postulation of a nominative-accusative case system entails the reconstruction of nominative-accusative syntax. Since the end of the last century, many scholars have wondered whether the Proto-Indo-European verb might not in fact have had ergative syntax and have consequently relabeled the reconstructed nominative case “ergative” and the accusative “absolutive” (see the bibliography in Szemerényi 1996:331–332). The principal argument in support of this hypothesis is the syncretism of nominative and accusative in all numbers of neuter nouns, anomalous in terms of accusative syntax, but explainable if neuter nouns originally only occurred in the absolutive. However, despite a number of ingenious morphological arguments, there is no widely agreed route by which the ergative syntax and morphology could have given the nominative-accusative morphology as reconstructed in §3.5, and if Proto-Indo-European did have an “ergative phase,” it may have been earlier than we can reach using the standard methods of reconstruction.

Much as anomalous morphological reconstructions have led to theories of Proto-Indo-European syntax, so anomalous syntactic constructions in Indo-European languages have led to revisions in the morphology. A striking case in point is an apparent breach of the concord rules of subject noun and verb. In Greek prose, neuter plural subjects take a singular verb:

- (13) τὰ ζῶα τρέχει  
 the-NEUT.PL. animal-NEUT.PL. run-PRES.3RD.SG.  
 “The animals run”

The same rule applies in Hittite and Gathic Avestan. The agreement of such an unusual syntactic rule across three of the earliest attested Indo-European languages can only represent the survival of an archaism. However, it is now generally accepted that the apparent concord of a plural subject and singular plural is a reflection of the fact that the neuter plural was originally a collective, formed with a suffix *\*-h₂*, which was later incorporated into a full paradigm. Consequently, we cannot set up a special syntactic rule of concord for Proto-Indo-European, but have rather to reconstruct a new morphological category – the collective.

Since Delbrück, the major work on reconstructing syntax has been done in two broad areas: word order studies and hypotaxis, particularly the syntax of relative clauses. Any account of Proto-Indo-European word order must begin with a statement of *Wackernagel's Law*, already mentioned in §2.6: enclitics occupy second position in the clause.

The case for the validity of Wackernagel's Law as an Indo-European phenomenon has been supported by the decipherments of Hittite and Mycenaean Greek, which show more rigorous applications of the law than Homeric Greek or Vedic Sanskrit. However, in recent years scholars have paid closer attention to the law's shortcomings (see especially Hale 1987, Krisch 1990, Adams 1994). In Wackernagel's original article on the law (Wackernagel 1892), he envisaged “enclitics” to cover three separate categories of unaccented words: (i) sentence

particles (these may be further categorized, see Hale 1987:19–20); (ii) enclitic forms of personal pronouns; and (iii) accentless verbal forms. Although difficulties of script and interpretation mean that we do not always have a clear idea of which words were truly clitics in early Indo-European languages, it appears that Wackernagel's Law is best observed (given certain modifications) with enclitics of class (i), while pronouns also show a tendency to associate with the verb phrase. The behavior of accentless verb forms is more complicated. In Vedic Sanskrit, verbs are usually accented in subordinate clauses but unaccented in main clauses, and Wackernagel saw an exact parallel to this in the Modern German verb-second order of main clauses, but verb-final order in subordinate clauses (1892:427). However, this correspondence appears to be fortuitous, and since Delbrück (1900:82), scholars have argued that only the copula verb was truly an enclitic.

It seems likely that Proto-Indo-European did not have fixed word order, and the attempt to fit Proto-Indo-European syntax into the straitjacket of typological universals has now largely been superseded by more nuanced assessments of word placement (see in particular the criticisms of Lehmann 1974 in Watkins 1976). The unmarked order appears to have been head-final, although pragmatic and prosodic factors may have played an important role. Note, for example, that Vedic Sanskrit, Greek, and Hittite all allow constituents to be fronted to a TOPIC position to the left of the sentence proper (Hale 1987:14f.).

The reconstruction of subordination and embedding for Proto-Indo-European continues to provoke debate. Even the reconstruction of relative clauses is controversial. Most of the Indo-European languages mark relative clauses with the reflex of either *\*yo-* (Greek, Sanskrit, Celtic, etc.) or *\*k<sup>wo</sup>-*/*\*k<sup>wi</sup>-* (Hittite, Latin, Tocharian, etc.). Although some scholars have argued that the use of two different markers shows that Proto-Indo-European did not have relative clauses of any type, others reconstruct both relative pronouns for the parent language, with an original distinction between *\*k<sup>wo</sup>-*/*\*k<sup>wi</sup>-*, functioning as a restrictive or defining relative, and *\*yo-* as an appositional or descriptive relative (see Hettrich 1988 for discussion).

Those who deny the existence of any relative pronouns in Proto-Indo-European envisage a development of relatives, and other subordinate clause types, in the daughter languages from earlier paratactic structures. Indeed, Kiparsky (1995) argues that the difficulty of reconstructing any complementizers for Proto-Indo-European implies that there was no complementation at all. However, the reconstruction of participles (§3.6.3), and compounding (§3.5.1), suggests that some forms of syntactic embedding were possible, and further research in this area is needed.

## 5. READING LIST

Fundamental and classic works on Proto-Indo-European grammar include Brugmann 1930, and the shorter Brugmann 1902–1904; Hirt 1921–1937; and Meillet 1964. On the Proto-Indo-European lexicon, an invaluable, if somewhat outdated, source is Pokorny 1973. A recent reworking of the lexicon is Rix 2001. For a valuable and up-to-date treatment of the Proto-Indo-European roots of English vocabulary, see Watkins 2000. More recent presentations of Proto-Indo-European phonology and morphology include Meier-Brügger 2002, Szemerényi 1996, Beekes 1995 (each with helpful bibliography), Cowgill and Mayrhofer 1986, Watkins 1969, and Kuryłowicz 1968. Surveys of various Indo-European daughter languages can be found in Bader 1994, Ramat and Ramat 1998, and Baldi 1983. A survey of Indo-European linguistic laws is presented in Collinge 1985.



The authors wish to express their indebtedness to the many scholars cited herein, as well as to Sara Kimball and Jochem Schindler. Most especially we are indebted to Calvert Watkins.

## Bibliography

- Adams, J. 1994. *Wackernagel's Law and the Placement of the Copula esse in Classical Latin*. Cambridge: Cambridge Philological Society.
- Bader, F. 1994. *Langues indo-européennes*. Paris: CNRS Editions.
- Baldi, P. 1983. *An Introduction to the Indo-European Languages*. Carbondale, IL: Southern Illinois University Press.
- Beekes, R. 1982. Review of M. Peters, *Untersuchungen zur Vertretung der indogermanischen Laryngale im Griechischen*. *Kratylos* 26:6–15.
- . 1995. *Comparative Indo-European Linguistics*. Amsterdam: John Benjamins.
- Benveniste, E. 1935. *Origines de la formation des noms en indo-européen*. Paris: Adrien-Maisonneuve.
- Brugmann, K. 1902–1904. *Kurze vergleichenden Grammatik der indogermanischen Sprachen*. Strasbourg: Trübner.
- . 1930. *Grundriss der vergleichenden Grammatik der indogermanischen Sprachen* (2nd edition). Berlin and New York: de Gruyter.
- Collinge, N. 1985. *The Laws of Indo-European*. Amsterdam: John Benjamins.
- Comrie, B. 1998. "The Indo-European language family: genetic and typological perspectives." In Ramat and Ramat 1998, pp. 74–97.
- Cowgill, W. and M. Mayrhofer. 1986. *Indogermanische Grammatik I/1–2. Lautlehre*. Heidelberg: Carl Winter.
- Crespo, E. and J. García Ramón. 1997. *Berthold Delbrück y la sintaxis indoeuropea hoy: Actas del Coloquio de la Indogermanische Gesellschaft, Madrid, 21–24 de septiembre de 1994*. Madrid: Ediciones de la Universidad Autónoma de Madrid.
- Debrunner, A. 1954. *Die Nominalsuffixe*. (= J. Wackernagel, *Altindische Grammatik*, vol. II, part 2). Göttingen: Vandenhoeck and Ruprecht.
- Delbrück, B. 1893–1895. *Vergleichende Syntax der indogermanischen Sprachen*. (= K. Brugmann and B. Delbrück, *Grundriss der vergleichenden Grammatik der indogermanischen Sprachen*). Strasbourg: Trübner.
- Eichner, H. and H. Rix. 1990. *Sprachwissenschaft und Philologie: Jacob Wackernagel und die Indogermanistik heute: Kolloquium der Indogermanischen Gesellschaft vom 13. bis 15. Oktober 1988 in Basel*. Wiesbaden: Dr. Ludwig Reichert.
- Gamkrelidze, T. and V. Ivanov. 1995. *Indo-European and the Indo-Europeans*. Berlin: Mouton de Gruyter.
- Hale, M. 1987. "Studies in the comparative syntax of the oldest Indo-Iranian languages." Ph.D. dissertation, Harvard University.
- Hettrich, H. 1988. *Untersuchungen zur Hypotaxe im Vedischen*. Berlin: de Gruyter.
- Hirt, H. 1921–1937. *Indogermanische Grammatik*. Heidelberg: Carl Winter.
- Hoenigswald, H. 1954. Review of *Proto-Indo-European Phonology*, by W. Lehmann. *Language* 30:468–474.
- . 1980. "Notes on reconstruction, word order and stress." In P. Ramat (ed.), *Linguistic Reconstruction and Indo-European*, pp. 69–87. Amsterdam: John Benjamins.
- . 1992. "Minimum freedom and the sentence." In S. Hwang and R. Merrifield (eds.), *Language in Context: Essays for Robert E. Longacre*, pp. 531–536. Arlington, TX: Summer Institute of Linguistics.
- . 1994. "Meter and phonology: the chronological interpretation of idealized reconstructions." In G. Dunkel, G. Meyer, S. Scarlata, et al. (eds.), *Früh-, Mittel-, Spätindogermanisch – Akten der IX. Fachtagung der Indogermanischen Gesellschaft vom 5. bis 9. Oktober 1992 in Zürich*, pp. 135–148. Wiesbaden: Dr. Ludwig Reichert.
- Kiparsky, P. 1995. "Indo-European origins of Germanic syntax." In A. Battye and I. Roberts (eds.), *Clause Structure and Language Change*, pp. 140–169. New York/Oxford: Oxford University Press.

- Krisch, T. 1990. "Das Wackernagelsche Gesetz aus heutiger Sicht." In Eichner and Rix 1990, pp. 64–81.
- Kurylowicz, J. 1968. *Indogermanische Grammatik ii. Akzent, Ablaut*. Heidelberg: Carl Winter.
- Lehmann, W. 1952. *Proto-Indo-European Phonology*. Austin: University of Texas Press.
- . 1974. *Proto-Indo-European Syntax*. Austin: University of Texas Press.
- Leumann, M. 1977. *Lateinische Laut- und Formenlehre*. Handbuch der Altertumswissenschaft. Munich: Beck.
- Lindeman, F. 1965. "La loi de Sievers et le début de mot en indo-européen." *Norsk Tidsskrift for Sprogvidenskap* 20:38–108.
- . 1997. *Introduction to the Laryngeal Theory*. Innsbruck: Institut für Sprachwissenschaft der Universität Innsbruck.
- Meier-Brügger, M. 2002. *Indogermanische Sprachwissenschaft* (8th edition). Berlin/New York: de Gruyter.
- Meillet, A. 1964. *Introduction à l'étude comparative des langues indo-européennes* (8th edition). Tuscaloosa: University of Alabama Press.
- Melchert, C. 1987. "Proto-Indo-European velars in Luvian." In C. Watkins (ed.), *Studies in Memory of Warren Cowgill*, pp. 182–204. Berlin/New York: de Gruyter.
- Metcalf, G. 1974. "The Indo-European hypothesis in the sixteenth and seventeenth centuries." In D. Hymes (ed.), *Studies in the History of Linguistics. Traditions and Paradigms*, pp. 233–257. Bloomington: Indiana University Press.
- Nussbaum, A. 1998. *Two Studies in Greek and Homeric Linguistics. Hypomnemata 120*. Göttingen: Vandenhoeck and Ruprecht.
- Pokorny, J. 1973. *Indogermanisches etymologisches Wörterbuch*. Berlin/New York: de Gruyter.
- Ramat, A. and P. Ramat (eds.). 1998. *The Indo-European Languages*. London/New York: Routledge.
- Risch, E. 1974. *Wortbildung der homerischen Sprache* (2nd edition). Berlin/New York: Walter de Gruyter.
- Rix, H. 1976. *Historische Grammatik des Griechischen*. Darmstadt: Wissenschaftliche Buchgesellschaft.
- . 2001. *Lexikon der indogermanischen Verben: Die Wurzeln und ihre Primärstammbildungen*. Wiesbaden: Dr. Ludwig Reichert.
- Russell, P. 1995. *An Introduction to the Celtic Languages*. Harlow: Longman.
- Szemerényi, O. 1996. *Introduction to Indo-European Linguistics*. Oxford: Clarendon Press.
- Wackernagel, J. 1892. "Über ein Gesetz der indogermanischen Wortstellung." *Indogermanische Forschungen* 1:333–436.
- . 1926. *Vorlesungen über Syntax*. Basel: Birkhäuser. Reprint 1950.
- Watkins, C. 1969. *Indogermanische Grammatik III/1. Verbalflexion*. Heidelberg: Carl Winter.
- . 1976. "Towards Proto-Indo-European syntax: problems and pseudo-problems." In S. Steever, C. Walker, and S. Mufwene (eds.), *Papers from the Parasession on Diachronic Syntax*, April 22, 1976, pp. 305–326. Chicago: Chicago Linguistic Society.
- . 1995. *How to Kill a Dragon*. New York/Oxford: Oxford University Press.
- . 1998. "Proto-Indo-European: comparison and reconstruction." In Ramat and Ramat 1998, pp. 25–73.
- . 2000. *The American Heritage Dictionary of Indo-European Roots* (2nd edition). Boston: Houghton Mifflin.

# Hittite

CALVERT WATKINS

## 1. HISTORICAL AND CULTURAL CONTEXTS

Hittite is a member of the Anatolian branch of the Indo-European family, and the earliest attested Indo-European language. Anatolian is generally regarded as the first branch to have separated from the other Indo-European languages. Aside from Hittite it includes Luvian (Cuneiform and Hieroglyphic) and Palaic, all from the second millennium BC, and Hieroglyphic Luvian, Lycian, Lydian, and the scantily attested Carian, Pisidian, and Sidetic in the first millennium BC.

The speakers of Hittite were in place in Central Anatolia by the nineteenth–eighteenth century BC, since a few words of the language (notably *išhiul-* “contract”) appear in Old Assyrian documents from the merchant colonies like *Kārum Kaneš*, Hittite *Nešaš*, modern Kültepe. As an Old Hittite origin legend shows (Otten 1973), the Hittites regarded this city as their original home; it is the base of their designation of their own language, <sup>URU</sup>*nišili*, *nešumnili* “in Hittite,” literally “in the language of (the inhabitants of) Nešaš.” With the beginning of our documentation of the language proper we distinguish Old Hittite (seventeenth or early sixteenth century–c. 1500), Middle Hittite (c. 1500–c. 1375), and Neo-Hittite (c. 1375–c. 1200). Adherents of the “short chronology” would lower these dates somewhat, particularly at the upper end.

Speakers of what was to be the Anatolian branch of the Indo-European family apparently migrated into Asia Minor, probably from the Balkans across the Bosphorus, in the course of the third millennium BC. It is not unlikely, though not susceptible of proof, that these immigrating future Anatolians were already dialectally differentiated into (at least) Pre-Hittites, Pre-Palaites, and Pre-Luvians. On the Central Anatolian plateau Pre-Hittites came in contact with the autochthonous Anatolian Hattic speakers, from whose self-designation (KUR <sup>URU</sup>*Hatti* “land of Hatti,” cf. *hattili* “in Hattic”) the Hittites took their name, as well as many aspects of their culture and religion.

The earliest Hittite history is one of warring petty kingdoms, described in our earliest Hittite text, that of Anittas (Neu 1974), eighteenth/seventeenth century BC. These city-states were subsequently united to form the Old Kingdom under Hattusilis I and his adopted son Mursilis I (seventeenth/sixteenth century), a period of rapid Hittite expansion into Syria, Hurrian Mittani, and Western Anatolia, “making the sea the boundaries.” Internal dissension and pressure from the hostile nomadic Kaska people to the north brought about retraction of Hittite hegemony during the succeeding Middle Kingdom, c. 1500–1375. The New Kingdom or Empire was founded by Suppiluliumas I, c. 1375 (he spoke late Middle Hittite; his son Mursilis III spoke classical Neo-Hittite). This was the period of greatest expansion of the Hittites and their role on the international scene. The Hittite Empire

came to an abrupt end shortly after 1200, during the reign of Suppiluliumas II, with the destruction of Hattusas by an unknown people, in all likelihood part of the general upheaval in the Eastern Mediterranean area caused by the “Peoples of the Sea,” the “*Šikalayū* who live in ships,” the people from the land of *Šikila*, as Suppiluliumas II referred to them in a letter to a prefect of Ugarit (Dietrich and Loretz 1978).

The Hittite language is preserved for us on clay tablets written in a cuneiform syllabary, the archives of the palace or central authority in the capital city of Hattusas (Boğazköy, now Boğazkale), and a few other urban centers like Maşat, Ortaköy, and Kuşaklı, the tablets themselves written over the period from the seventeenth/sixteenth to the end of the thirteenth century. One of the important functions of the Hittite “state” was to assure the regular performance of ritual, and the correct preservation of the appropriate words and actions of ritual procedure. The great majority of our texts deal with religion and the administration of cult, festivals, and both public and private rituals, as well as magic, oracles, and divination. Our texts also include the Hittite political archives, treaties, political and some personal correspondence, land grants, as well as historical texts and annals (by regnal year) of individual rulers (see Beckman 1996). We find also “instructions” for religious and secular administrations and military personnel, all – like the treaties with foreign powers – regarded as engagements of personal fealty and labeled simply *išḫiul*–“contract.” We have a highly original law “code” composed and written down originally in the Old Kingdom, together with later copies (Hoffner 1997), but only a few documents dealing with the administration of public or private justice. Literary texts are primarily mythological (Hoffner 1990) in character, and both native compositions and translations from Hattic, Hurrian, and Sumero-Akkadian sources. The archives also include foreign-language cultic material, sometimes with Hittite translation, in Hattic, Hurrian, Sumerian and Akkadian, Cuneiform Luvian, and Palaic, attesting the significant cultural influence of all of these. For a catalogue of the Hittite texts then known see Laroche 1971 and supplement.

Hittite was clearly the language of the ruling classes, of public and private administration, and of the army, as our texts show. The changes over the four or five hundred years of our documentation of Hittite are entirely consistent with the development of a spoken language. At the same time, the extensive Luvian elements in Hittite personal names, the practice in the later empire of setting up large public inscriptions in Hieroglyphic script and in the Luvian language, and the frequency of Luvian loanwords in Hittite texts, often marked as foreign by the prefixation of the *Glossenkeil* (ε), would point to widespread use of Luvian and bilingualism.

Dialectal variation is virtually nonexistent in Hittite, not surprisingly since our texts are probably all produced in the same tradition of professional scribes. One or two texts like KUB 48.69 point to genuine dialect variation, but by and large they are remarkably homogeneous, as is to be expected in a literary language.

## 2. WRITING SYSTEMS

Our preserved Hittite texts were written by professional scribes on clay tablets, impressed with a stylus and then baked (plus one bronze tablet with signs hammered in). The writing system is the Mesopotamian cuneiform syllabary of the second millennium, borrowed probably in Northern Syria from a Peripheral Akkadian (see Ch. 8 §1.1) scribal school source, in the seventeenth century at the beginning of the Old Kingdom period. The signs in use in Boğazköy most closely resemble the Old Babylonian variants (Labat 1976).

The Old Assyrian variety of the merchant colonies in central Anatolia at an earlier period left no trace on Hittite literacy.

In addition to the cuneiform written by professional scribes on clay tablets, the Hittites also made use of another syllabary, the hieroglyphic. This syllabary, which made extensive use of logograms as well, was used for monumental carved rock inscriptions in the Luvian language in the empire (and continued in Southeastern Anatolia and Syria to c. 750 BC), and from the time of the Old Kingdom on, for names and titles on seals. The latter were doubtless logographic and not “in” any language, but read in Hittite context as Hittite, like numerals in modern scripts. The same may have been true for the monumental public inscriptions in the Empire, and for the wooden tablets inscribed with hieroglyphs the existence of which is evidenced in text references. For discussion of the hieroglyphic script see Chapter 19.

The cuneiform syllabary notes syllables of the structure V (the vowels *a, e, i, u*), CV (i.e., consonant + vowel), VC, and some CVC. The sets of CV and VC signs are incomplete for inherent *e*, and CVC signs distinguish only the vowels *a, i, u*, and these not always. For the cuneiform script, see Appendix 2.

The writing system also makes use of a number of logograms from Sumerian (*Sumerograms*) and Akkadian (*Akkadograms*, written syllabically). The Hittitological convention is to transliterate syllables, writing Hittite in lower case, Sumerograms in roman capitals, and Akkadograms in italic capitals: *at-ta-aš* “father,” *e-eš-zi* “is,” *LÚ* “man,” *LUGAL* “king,” *BI-IB-RU* “rhyton,” *QA-TAM-MA* “as follows.” *Narrow transcription* separates each sign of a word by a hyphen, as in *at-ta-aš, e-eš-zi*; *broad transcription* (with greater phonetic accuracy) erases the hyphens and deletes one of the identical vowels of CV–VC sign sequence, as in *attaš*, and if two vowels remain, marks a macron, as in *ēšzi*.

Akkadograms and Sumerograms sometimes alternate with syllabic Hittite spellings in duplicate texts, which shows that they functioned as rebus writing, purely graphic variants of the Hittite words actually pronounced, just as the Sumerograms were read and pronounced as Akkadian in the source script of the Hittite writing system. The same conclusion is indicated by the common practice of following a Sumerogram with a phonetic complement which may serve to indicate grammatical endings. Thus, for example, the Sumerogram DINGIR “god” may be followed by the Akkadian phonetic complement *LIM*, conventionally transliterated superscript DINGIR<sup>LIM</sup>, to write the (Old) Akkadian genitive singular *ILIM*. So read in Akkadian, the whole in Hittite may receive a further phonetic complement written syllabically, DINGIR<sup>LIM</sup>-*na-aš*, to write the genitive singular of the word for “god” in Hittite, *šunaš*.

A Sumerian scribal practice, continued as graphic convention in Akkadian and then in Hittite, is the use of *determiners* prefixed to words and names to classify them by semantic category. These are conventionally transliterated superscript, and were doubtless not pronounced in Hittite (or Akkadian). They indicate categories like male person (<sup>m</sup> or <sup>l</sup>), female person (<sup>f</sup>), god/goddess (<sup>D</sup> abbreviated for <sup>DINGIR</sup>), city (<sup>URU</sup>), stone (<sup>NA<sub>4</sub></sup>), wooden object (<sup>GIŠ</sup>), and the like.

A further, specifically Hittite graphic convention is to mark grammatical cases of nouns or names written as logograms by preceding them with an Akkadogram. Thus, *ŠA* (Akk. “the one of”) marks genitive; *I-NA* (Akk. “in”) indicates dative-locative, and “allative” with inanimates; *A-NA* (Akk. “to”) indicates dative with animates; *IŠ-TU* (Akk. “from, by”) marks both ablative and instrumental. Proper names preceded by determiner or Akkadographic case-marker are frequently, though not always, unmarked for case and thus function by graphic convention as quasi-logograms.

We may illustrate these spelling conventions with Figure 18.1 (Bo 91/1314), a seal of the founder of the empire (from Otten 1995). The outer and inner ring legends are cuneiform:



**Figure 18.1** The seal of Suppiluliumas I  
<sup>NA4</sup> KIŠIB <sup>m</sup>Šu-up-pi-lu-li-u-ma LUGAL GAL UR.SAG  
 DUMU <sup>m</sup>Du-ut-ta-li-ya LUGAL GAL UR.SAG  
 Seal (of) Suppiluliumas, great king, hero,  
 Son (of) Duthaliyas, great king, hero.

The inner field in hieroglyphs shows the royal emblem of the winged sun, corresponding to the title <sup>D</sup>UTU <sup>ŠI</sup> for <sup>D</sup>ŠAMŠI “my sun,” over the signs right and left MAGNUS REX “great king” (hieroglyphs are conventionally transcribed in Latin) flanking the three signs of the name: PURUS.FONS-*ma/i* for Suppi-luli-(u)ma (PURUS = Hittite *šuppi-* “pure, sacred,” FONS = Hittite *luli-* “pond, spring,” with phonetic complement). Under the name as space-filler is the (cuneiform) Sumerogram TI “life,” upside down.

In the four to five hundred years of its documented history the Hittite cuneiform writing system and scribal practices did not undergo any massive or dramatic changes. But small changes in the shapes of certain signs and the general appearance of the tablets and their *ductus* over this period have enabled scholars to date the tablets fairly precisely to the early or late Old, Middle, and Neo-Hittite periods respectively. The original impetus was given by the discovery in the early 1950s of a tablet fragment (the Zuckraši-text, Laroche CTH 15) in a stratigraphically certain Old Kingdom archeological context; its characteristic *ductus* was found to recur on many of the tablets already unearthed from the palace archives. Those tablets exhibiting the old *ductus* were then seen to preserve certain characteristic features of language and orthography which could be identified as archaic. The periodization of our corpus of texts and the attendant conclusions about the history of the Hittite language have been the subject of intense investigation by philologists and linguists in the latter part of the twentieth century, and the results are by now generally accepted. We can distinguish paleographically Old, Middle, and New Script (OS, MS, NS); original compositions from these periods are in Old, Middle, and Neo-Hittite (OH, MH, NH). Documents were often recopied later than their composition, such that we can classify the tablets, following the



convention of *The Chicago Hittite Dictionary*, as early or late OH/OS, OH/MS, OH/NS, MH/MS, MH/NS, NH/NS.

At the time of the German archeological excavations at Boğazköy under Hugo Winckler beginning in 1906, which unearthed the initial collection of tablets, the Akkadian cuneiform writing system had already been deciphered. The Hittite tablets could therefore be “read,” i.e., transliterated, but not understood. The actual decipherment of the language and its identification as Indo-European was the work of a young Czech Assyriologist, Bedřich (Frédéric) Hrozný, during World War I. His first-hand account of his decipherment can be found in the article “Hittite language” of *The Encyclopaedia Britannica* (14th edition).

### 3. PHONOLOGY

#### 3.1 Graphic considerations

Any discussion of the phonological system of Hittite must begin with consideration of the distinctions made by the cuneiform writing system. The phonological structure of Hittite was clearly different from that of the Semitic language from which the cuneiform was first borrowed. For the details of what follows see Melchert 1994.

Using the symbols V = vowel, C = consonant, we may state that the cuneiform syllabary had signs of the structure V, CV, VC, and CVC (see the above discussion of the Hittite writing systems). The vowels were *a, e, i, u*, and the consonants of the CV series *p, t, k, q, b, d, g, ħ, š, s, z* (an affricate *tš*, Semitic *š* and *z*), *m, n, r, l, w, y*. CV signs with inherent vowel distinguish *a, i, u*, but not all possibilities with inherent *e* are present: thus *ta, da, ti, di, tu, du*, and *te* but not *\*de*, and only *ya, wa* (and secondarily acrophonic *wi<sub>5</sub>* after GEŠTIN “wine,” Hitt. *wiyana-*). The VC series made fewer distinctions, merging voice (*at = ad*) and often ignoring inherent *e* (*iš* vs. *eš*, but only *im*, for example), and the CVC series was less systematic (e.g., *šap* but no *\*šak*). For writing, Hittite *ka* and the rarer *qa* (*ka<sub>4</sub>*) are treated as equivalent, and with few exceptions *š* (*a*, etc.) is used exclusively for writing the single Hittite sibilant, to the exclusion of *s* (*a*, etc.).

The Hittites did not utilize the Semitic orthographic opposition of voiced : voiceless (*da* : *ta*, *ga* : *ka*, etc.), but rather, most clearly in intervocalic position, opposed simple versus geminate (double) consonants, thus *a-ta* (or *a-da*) versus *at-ta* (or *ad-da*), *a-ħa* versus *aħ-ħa*, etc., probably pointing to a phonological contrast of lax : tense (lenis : fortis) respectively. In the case of *š* and the liquids and nasals simple versus double consonants likewise contrasted: *a-na* versus *an-na*, *a-ša* versus *aš-ša*. In initial position the same word could in principle be written with either the voiced or the voiceless sign, the choice governed by scribal convention, for example, third singular *da-a-i* “puts” but third plural *ti-ya-an-zi* “they put.” Previously regarded as arbitrary, this fact has now been explained as indicating a merger of inherited voiced : voiceless (lax : tense) stops in initial position, with generalization of the voiceless or tense stop. Word-finally, the voiced or lax stops were generalized, as is clear from spellings with simple stop before enclitic: *pait = as* [paydas] “went he,” *natid = a* [nadida] “but with an arrow.”

#### 3.2 Consonants

The Hittite inventory of phonemic consonantal segments distinguishes four places of articulation (labial, apico-dental, velar, and labiovelar, the last-named usually written with the *ku* sign before a vowel or consonant, but occasionally *uk* before a consonant), five



manners of articulation (stop, affricate, fricative, nasal, liquid, and glide), and two glottal modes (tense/voiceless and lax/voiced). Here and below, the symbols <> enclose spelling (orthographic) forms.

### (1) Hittite consonantal phonemes

p	t	k	k <sup>w</sup>
b	d	g	g <sup>w</sup>
	ʦs <z>		
	s	H <-ḫḫ->	
		h <-ḫ->	
m	n		
	l	r	
w		y	

## 3.3 Vowels

The inventory of vowels has four members and a correlation of length. Long vowels are noted (inconsistently) by so-called *scriptio plena* or plene-writing, *Ca-a* versus *Ca*, *Ca-a-aC* versus *Ca-aC*, i.e., [Cā] versus [Ca], [CāC] versus [CaC]:

### (2) Hittite vowel phonemes

i	u	ī	ū
e	a	ē	ā

Diphthongal combination like that of *ā* and the glides *w* and *y*, noted *(a-)a-i*, *(a-)a-u*, are also permitted.

## 3.4 Phonological variation

Morphophonemic variants are not numerous. A *w* adjacent to *u* is replaced by *m*. This involves the sequence *uwV*, in part generated from *wV* after a heavy syllable by the inherited feature known as Sievers' Law, in part from *u+wV* across morpheme boundary, as well as the mirror image *Vw (+)u*. Compare first plural *tar-weni* "we say" but *šarr-umēni* "we break," *tepnū-mēni* "we belittle," or nominative plural *idālaw-eš* "bad" but accusative plural *idālam-uš*.

The original inherited sequence *\*VnsV* became in Hittite *VššV*, as in *\*densu-* > *daššu-* "massive." This treatment was generalized across morpheme boundary in *accusative singular + enclitic possessive*, for example, *annan+šan* > *annaššan* "his/her mother."

The enclitic conjunction *-a* "and" (cf. Luvian *-ḫa* "and") causes gemination of a preceding consonant – *ūk* "I," *ūkka* (*ūgga*) "and I" – and thus can be distinguished from enclitic *-a* "but, however": *ūka* (*ūga*) "I, however."

Hittite, like other Anatolian languages, shows the effects of correlation of vowel length (see §3.6) and the inherited Indo-European accent (see §3.7). In particular, unaccented long vowels were shortened. Short vowels were lengthened (originally, at least, allophonically) in accented open syllables, and the mid and low vowels *e* and *a* in accented closed syllables as well: *\*pēdom* (cf. Greek πέδον) > *pe-e-da-an* [pédan] "place," *\*h<sub>1</sub>ésti* (cf. Greek ἔστι) > *e-eš-zi* [éšzi] "is." To what degree these are synchronic rules in Hittite is controversial; see Melchert 1994 for discussion.

### 3.5 Consonant clusters

The cuneiform syllabary does not permit the unambiguous notation of clusters of two or more consonants in word-initial or word-final position, nor clusters of more than two consonants word-medially. Spelling variation indicates that at least some consonant clusters were real, and involved an “empty” vowel, e.g., *ma-li-it-tu-* and *mi-li-it-tu-* “sweet” for [mlitu-]. Inherited initial \*sT- clusters (where T = stop) are usually noted iš-TV-; whether the prothetic vowel is real or not is much debated. A number of examples point to the existence of real anaptyctic vowels breaking up clusters, like *akkiš* “died,” *lakkiš* “knocked over” < \*aks, \*laks with voicing assimilation, from etymological \*og-s(-), \*log<sup>h</sup>-s(-). Their interpretation remains controversial.

### 3.6 Vowel length

Hittite inherited the Common Anatolian and Indo-European opposition of long and short vowels. The subsequent lengthening of accented short vowels in open and in some cases closed syllables, and the shortening of unaccented long vowels (see §3.4), affected the distribution of long and short vowels but not the opposition per se. The correlation of stress and vowel length is very uncommon in Indo-European languages of this antiquity (but compare the similar phenomenon in Middle English), and the lengthening of accented vowels in closed as well as open syllables is typologically rare cross-linguistically.

### 3.7 Accent

Hittite likewise inherited from Common Anatolian the Indo-European *accent*, traditionally described in terms of pitch but clearly including a stress component as well. The secondary effects of the Hittite accent or its absence, lengthening and shortening of vowels respectively (see §3.4), are those typical of a stress accent cross-linguistically.

Hittite normally preserved the place of the Indo-European accent, including mobile accent in some paradigms: compare 3rd sg. *ēszi*, 3rd pl. *ašanzi* “is, are,” ppl. *ašānt-*, probably from earlier \*ēs-ti, (a)s-énti, (a)s-ónt-; or *tēkan* “earth” [tégan], loc. sg. *taknī* [tagní]. In some cases, the position of the accent has shifted: nom.-acc. pl. *widār* “waters” [widár], PIE \*wédōr, gen. sg. *kūnaš* “dog” [kúnas], PIE *kúnós* (cf. Greek κυνός but Vedic *śúnas*). The question awaits a systematic solution.

### 3.8 Diachronic developments

#### 3.8.1 Stops

The Proto-Indo-European stop system is usually reconstructed as follows (Cowgill and Mayrhofer 1986):

(3)	p	t	ḱ	k	k <sup>w</sup>
	b	d	ḡ	g	g <sup>w</sup>
	b <sup>h</sup>	d <sup>h</sup>	ḡ <sup>h</sup>	g <sup>h</sup>	g <sup>wh</sup>

The plain voiced and voiced aspirate series merged in Proto-Anatolian, yielding

(4)	p	t	ḱ	k	k <sup>w</sup>
	b	d	ḡ	g	g <sup>w</sup>

The phonological contrast of voiceless : voiced was probably revalued to tense : lax, with the tense member longer in duration than the lax. For the need to recognize three dorsal points of articulation – palatal, plain velar, labiovelar – see Melchert 1994 with earlier literature. Common Anatolian further affricated  $*t$  before  $*\gamma$  to  $*[tʃ']$ , originally allophonically. In the dialect(s) ancestral to Luvian and Lycian, Proto-Anatolian  $*\hat{k}$  apparently merged with this  $*[tʃ']$  (<z>) and gave it phonemic status (see Ch. 19, §3.1). In the dialect ancestral to Hittite,  $*\hat{k}$  and  $*\hat{g}$  merged completely with  $*k$  and  $*g$ , while the affrication of  $*t$  to  $*[tʃ]$  before  $*i$  and further developments led to its phonemic status as <z> =  $[tʃ]$ . As a result, both dialects of late Proto-Anatolian showed the same inventory,

- (5) tense (long):    p    t    z    k    k<sup>w</sup>  
                       lax (short):    b    d            g    g<sup>w</sup>

but with differing distribution. The distribution is further altered by the “lenition” rules in Proto-Anatolian, by which tense (long, i.e., inherited voiceless) stops become the corresponding lax (short, i.e., inherited voiced) stop after accented long vowel or diphthong and between unaccented vowels. For these rules see Eichner 1973:79ff. and 100, fn. 86 and (more clearly) Morpurgo Davies 1982–1983, especially for Luvian and Lycian, as well as Melchert 1994:60 *et passim*. The effects of this rule are most palpable in the endings originally beginning with dental in the Luvian languages, where, for example, depending on accent and quantity the third singular is active  $-(t)ti$  or  $-di$ , middle  $-(t)ta-$  or  $-da-$ . In Hittite the effects of the rule have been largely leveled out (Melchert 1994:61), save for a handful of isolated instances, and it is unclear how the effects of the rule were eliminated in this language.

Typologically, the Anatolian reduction of the Indo-European stop system to a tense : lax opposition, and that only in medial position, with neutralization to [+tense] in initial position, [–tense] finally, is unique in the Indo-European family. It seems to be an areal feature in second-millennium Anatolia. The neutralization to [+tense] in initial position is controversial but plausible for Hittite and the other second-millennium cuneiform languages; it is certain for the alphabetically written languages of the first millennium, Lycian and Lydian.

The above consonant treatments as well as the prosodic developments discussed earlier (see §3.4) began as synchronic developments, and to a certain extent may still be so analyzed. They represent challenging problems for linguistic typology.

### 3.8.2 Laryngeals

Hittite and the Anatolian family are noted for preserving two of the three Indo-European “laryngeal” consonants in initial position, the “a-coloring”  $h_2$  and the “o-coloring”  $h_3$ :  $harki-$  “white,” PIE  $*h_2ar\hat{g}i-$  <  $h_2er\hat{g}-i-$  (cf. Latin *arg-entum* “silver”);  $palhi-$  “broad,” PIE  $*p_lh_2-i-$ ;  $\check{s}alli-$  “large,” PIE  $*solh_2-i-$ ;  $h\ddot{a}ppar$  “transaction,” PIE  $*h_3op-$  <  $*h_3ep-$  (cf. Latin *ops* “wealth”);  $h\ddot{a}ran-$  “eagle,” PIE  $*h_3or-n-$  <  $*h_3er-n-$  (cf. Greek ὄρνις, English *erne*);  $harra-$  “crush,” PIE  $*h_2arh_3-o-$  <  $*h_2erh_3-o-$  (cf. Greek ἄρο-τρον “plough”).

### 3.8.3 Sonorants

The difference in preforms between “broad” ( $*p_lh_2-i-$ ) and “large” ( $*solh_2-i-$ ) or “crush” ( $*h_2arh_3-o-$ ) shows that Proto-Anatolian still preserved the Indo-European syllabic sonorants  $*j$ ,  $*l$ ,  $*\eta$ ,  $*n$ , and their replacement by *ar*, *al*, *un*, *an*, occurred not long before the historical period. The special reflex *un* of word-final  $*\eta$  is controversial, but plausible.

## 4. MORPHOLOGY

Hittite as the earliest attested member of the Indo-European family of languages shows the familiar Indo-European pattern of morphological type known as *fusional*: a single inflexional morpheme regularly expresses a combination of grammatical categories, for example, *-s* marks nominative case, singular number, animate gender. The language shows a fairly rich inflexion of nominal, pronominal, and verbal categories.

### 4.1 Word formation

Words in Hittite are either inflected or uninflected. The *structure* of the Hittite inflected *word* is R(oot) + S(uffix or suffixes) + E(nding). The *root* gives the basic lexical semantic content, and the *suffix* or suffixes add derivational and grammatical meaning, as well as specifying the part of speech. Root + Suffix(es) together are termed the *stem*, and constitute a lexical or dictionary entry, an inflected word in the language. Thus, the noun *kartimmiyatt-* “anger” is built by the nominalizing action noun suffix *-att-* on the verb-stem *kartimmiya-*, with denominative verbal suffix *-iya-*, itself formed from a probable nominal stem *\*kartim(m)a-* with suffix *-(i)ma-*. The double *m* is probably just due to the usual spelling with the sign *tim* (*kar-tim-mi-*), Akkadian also *tì, di<sub>11</sub>*, without final *m*. The stem *\*kartima-* in turn is built (via a probable denominative verbal stem *\*kart-ailiya-*, cf. *šallakartāi-* “behave arrogantly toward”) on the noun stem *kart-* = *kard-* of the body part “heart,” PIE *\*k<sub>1</sub>rd-*.

Uninflected words are either frozen inflectible (R+S+E) stems, for example, the adverb *karū* “formerly” (with suffix and zero ending), or they are particles (on which see §5, Syntax). Though the evidence is only indirect, Hittite probably inherited from Indo-European the property that the numerals 1 to 4 were inflected adjectives while 5 to 10 were uninflected “particles.”

### 4.2 Nominal morphology

The Hittite nominal system includes the substantive, the adjective, and the lower numbers. Its *inflexional categories* are gender, number, and case.

#### 4.2.1 Gender

Hittite has two genders, animate (frequently termed common) and inanimate (frequently termed neuter). Comparative evidence, notably Lycian, shows that Proto-Anatolian had the traditional Indo-European three-gender system of masculine (Lycian nom. *-e*, acc. *-ē* < *\*-os*, *\*-om*), feminine (Lycian nom. *-a*, acc. *ā* < *\*-ā*, *\*-ām* < *\*-ah<sub>2</sub>*, and underlying *\*-ah<sub>2</sub>m* by Stang’s Law), and neuter (see Ch. 21, §4.1). Compare also the Luvian abstract suffix *-ah<sub>2</sub>-id-* from *\*-ah<sub>2</sub>-*, with the same suffix as the Greek abstract type *τομή* “cutting.” Hittite as well as the other cuneiform Anatolian languages of the second millennium (see §1) has innovated by eliminating the feminine gender by merger, as a consequence of certain phonological developments. Thus, Indo-European feminine *\*-ah<sub>2</sub>* (underlying *\*-eh<sub>2</sub>*) lost its final laryngeal by rule, and the undercharacterized nominative *\*-a*, like the *\*-a* resulting from unstressed nominative *\*-ō* of the *n*-stems, was further marked by nominative *-s*, and the resulting *-aš* rendered identical to *-aš* from masculine thematic *\*-os*. Thus, Hittite

nominative *ḫāššaš* “hearth” (cf. Latin *āra*, fem.) like *ḫāraš* “eagle” (cf. Old High German *aro* < \**h<sub>3</sub>órō*).

#### 4.2.2 Number

Hittite has two numbers, singular and plural. Some scholars have seen a trace of the Indo-European dual in such forms as *šākuwa* “eyes” (and Luvian *tawa* “eyes,” *iššara* “hands,” *GÌR<sup>MEŠ</sup>-ta* = *pāta* “feet”), comparing either Vedic dual *pādā* or Mycenaean Greek (*tiri*)*pode*, but the Anatolian ending is indistinguishable from the neuter plural. The latter is frequently used to form a collective plural opposed to an individual (count) plural of animate nouns: *alpāš* “cloud,” individual accusative plural *alpuš*, collective *alpa<sup>HLA</sup>*.

#### 4.2.3 Case

The Old Hittite noun shows nine cases. These are nominative, vocative, accusative, genitive, dative-locative, directive (also termed allative), ablative, instrumental, and ergative. The function of most of these cases is the one that is familiar in an older Indo-European language, largely self-explanatory. The dative-locative marks both location and the indirect object, and may represent a syncretism of two earlier distinct cases. Directive (allative) and ablative mark motion to or from. Hittite and the other Anatolian languages show a *split-ergative* system, in which neuter nouns functioning as agents, subjects of transitive verbs, are marked by a special ergative case ending (see Garrett 1990, 1996). The development of the system of split ergativity is an important common innovation of the Anatolian branch of the Indo-European family. It is closely connected with another important morpho-syntactic innovation of Common Anatolian, the development of enclitic subject pronouns with “unaccusative” intransitive verbs (Garrett, *ibid.*). See further the sections on voice, on the pronominal system, and on diachronic syntax.

The distribution of the Old Hittite cases between the two numbers, with their usual formal exponents, is as follows (commas separate variants):

Table 18.1 Old Hittite noun inflection				
	Singular		Plural	
	Animate	Inanimate	Animate	Inanimate
Nominative	-aš, -š, -Ø	-an, -Ø	-eš	-a
Vocative	-i, -Ø	-an, -Ø	-eš	-a
Accusative	-an	-an, -Ø	-uš	-a
Genitive		-aš		-an
Dative-locative		-i, -Ø		-aš
Directive		-a		-aš
Ablative		-az		-az
Instrumental		-it		-it
Ergative		-anz(a)		-anteš

By Neo-Hittite this system had undergone a number of changes. The separate vocative is disused, and the directive is lost by merger with the dative-locative; the genitive plural

merges with the dative-locative, and ablative and instrumental become noncontrastive, as do nominative and accusative animate plurals. The result is as follows:

**Table 18.2 Neo-Hittite noun inflection**

	Singular		Plural	
	Animate	Inanimate	Animate	Inanimate
<i>Nominative</i>	-aš, -š	-an, -Ø	-eš, -uš	-a
<i>Accusative</i>	-an	-an, -Ø	-eš, -uš	-a
<i>Genitive</i>	-aš		-aš	
<i>Dative-locative</i>	-i		-aš	
<i>Abl.-instr.</i>	-az, -it		-az, -it	
<i>Ergative</i>	-anz(a)		-anteš	

#### 4.2.4 Adjectives

Hittite adjectives show agreement in gender and number with nouns. The endings are the same as for the nouns. Adjectives are not inflected for degrees of comparison; comparative and superlative are expressed by syntactic means alone, positive plus dative-locative or ablative, and positive plus genitive plural (dative-locative plural?) respectively: *iškiši šalli* “big to the (other’s) back” = “bigger than the (other’s) back,” *šallayaš=kan* DINGIR<sup>MES</sup>-*aš kuiš šallis* “who of the great gods (is the) great(est).” This syntactic pattern is found marginally in other ancient Indo-European languages as well, like Vedic *yé devānām yajñtīyā yajñīyānam* “who of the worshipworthy gods is (the most) worshipworthy,” or Homeric Greek *δῖα γυναικῶν* “(the most) divine of women.”

#### 4.2.5 Nominal stem-classes

The stem-classes or declensions of the Hittite nominal are as follows; the case endings themselves have been given above. We distinguish first *athematic* and *thematic* formations, which differ only in the nominative singular: athematic animate -š (combining with stem-final dental to yield -z, spelled -za; *ašānt-* “being, real, true,” nom. sg. *ašānza*), Ø (OH *keššar* “hand”); inanimate -Ø (*milit* “honey”). Compare thematic animate -aš (*ḫartaggaš* “bear”), inanimate -an (*yugan* “yoke”). Athematic stems are *consonant stems* (see below); it is convenient, on the basis of the nominative singular, to term *vocalic stems* both the thematic stems (nom. -a-š, *ḫartagg-a-š* “bear”; -a-n, *pēd-a-n* “place”) and *i-* and *u-* stems (nom. -i-š, -u-š). The latter show the ending -Ø for the inanimate nominative-accusative singular (*ēšri* “form,” *gēnu* “knee”). Diphthongal stems in -ai-, -au-, -e(i)- are also found, again with the inanimate nominative-accusative singular -Ø. The thematic stem is invariant; *i-* and *u-* stems may show ablaut of the predesinential element: -a[y]/-/i-, -aw-/u-, e(i)/-/i-. The -u- and -i- before *a* appear as -uw-, -iy-. Intervocalic *y* in -a[y]- is subject to deletion, with coalescence of like vowels, but is sometimes analogically restored. Ablaut is characteristic of adjectival stems (*āššu-/āššaw-* “good” vs. *āššu-/āššu-* “good(s)”) but many substantives show it as well. We may illustrate typical vocalic stems (stem-vowel + case ending) of animate nouns and adjectives; it will be sufficient to give nominative, accusative, and genitive since the remaining case endings are added to the stem as it appears in the genitive:

(6)	Thematic stem	i-stem	i-stem adjective	u-stem
Nom.	-aš	-iš	-iš	-uš
Acc.	-an	-in	-in	-un
Gen.	-aš	-iyaš	-a[y]aš	-uwaš
	Diphthongal i-stem	Diphthongal u-stem	u-stem adjective	
Nom.	-aiš	-auš	-uš	
Acc.	-ain	-aun	-un	
Gen.	-iyaš	-uwaš	-awaš	

Note also the inanimate nom.-acc. *utnē* “land” (underlyingly *-ēi*), gen. *utniyaš*.

Consonantal stems of both genders are found ending in obstruents (anim. *kašt-* “hunger,” nom. sg. *kašza*; inan. *šēppitt-* a cereal, “wheat”?, nom. sg. *šēppit*; inan. *nēpiš-* “heaven”) and sonorants (anim. *hašter-* “star,” nom. sg. *hašterza*; inan. *lāman-* “name”). Many show paradigmatic ablaut, often with accent shift: nom.-acc. *tēkan* “earth,” dat.-loc. and directive (allative) *taknī*, *taknā*, suffixless *dagān*; nom. *keššar* “hand,” acc. *kiššeran*, gen. *kišraš*, dat.-loc. *kiš(ša)rī*, instr. (OH) *kiššarta*, *kiššarat* ([kés(s)ard] or [kis(s)árd]). Very common in Hittite and Anatolian, though residual elsewhere in Indo-European, are *r/n*-stem inanimates with nominative-accusative singular and plural in *-r* and remaining cases in *-n*: *haršar* “head,” gen. *haršanaš*, pl. nom.-acc. *haršār*. Compare Old Avestan *aīiarə* “day,” stem *aīian-*, nom.-acc. pl. *aīiārə*.

### 4.3 Pronouns

The Hittite pronominal system includes the personal pronouns, the demonstratives, and the interrogative-relative-indefinite pronouns. These differ in inflection from the nominal system in a number of ways, as they do in other old Indo-European languages. The *personal pronouns* distinguish stressed (tonic) and enclitic forms. Hittite is a “PRO-drop” language, incorporating the subject into finite verb forms, and the stressed pronouns of the first and second persons both subject and other, oblique arguments are used only for emphasis or contrast. The normal expression of pronominal objects is by enclitics. The usual Old Hittite forms are as follows; note that direct and indirect object (accusative, dative-locative) in the personal pronoun proper (first and second persons) are not distinguished, and the instrumental is not found:

**Table 18.3 First- and second-person pronouns**

	Singular			
	First person		Second person	
	Tonic	Enclitic	Tonic	Enclitic
Nom.	ūk		zīk	
Obl.	ammuk	=mu	tuk	=ta (=du=za)
Gen.	ammēl		tuēl	
Abl.	ammēdaz		tuēdaz	
	Plural			
	First person		Second person	
	Tonic	Enclitic	Tonic	Enclitic
Nom.	wēš		šumēš	
Obl.	anzāš	=naš	šumāš	=šmaš
Gen.	anzēl		šumenzan	
Abl.	anzēdaz		šumēdaz	



For the third person, only enclitic forms occur, in three cases: nominative, accusative, dative-locative. The third-person nominative (subject) pronouns are found, as noted above, only with the “unaccusative” subset of intransitive verbs. The Old Hittite forms are:

(7)	<i>Singular</i>		<i>Plural</i>	
	<i>Animate</i>	<i>Inanimate</i>	<i>Animate</i>	<i>Inanimate</i>
	<i>Nom.</i>	-aš -at	-e (NH -at)	-e (NH -at)
	<i>Acc.</i>	-an -at	-uš (NH -aš)	-e (NH -at)
	<i>Dat.-loc.</i>	-ši	-šmaš	

If more than one third-person object enclitic is present, accusative precedes dative-locative; third person usually precedes other persons, but first and second plural dative-locative precedes third singular accusative (Friedrich 1960, §288).

Old Hittite marks possession by a set of enclitic pronouns of all three persons singular and plural, suffixed directly to the possessed noun, and agreeing with it in gender. They show the stem-vowels *-i/-e-* for the nominative animate and inanimate before the pronominal endings *-š* and *-t*, otherwise the thematic vowel *-a-*:

(8)		<i>First</i>		<i>Second</i>		<i>Third</i>	
	<i>Singular</i>						
	<i>Nom.</i>	-miš	-met	-tiš	-tet	-šiš	-šet
	<i>Acc.</i>	-man		-tan		šan	
	<i>Gen.</i>	-maš		-taš		-šaš	
	<i>Dat.-loc.</i>	-mi		etc.		etc.	
	<i>Dir.</i>	-ma					
	<i>Abl.-instr.</i>	-mit					
	<i>Plural</i>						
	<i>Nom.</i>	-meš	-met				
	<i>Acc.</i>	-muš					
	<i>Gen.</i>	-man					

A possessed noun may appear anywhere in the sentence, but if it comes first, any other enclitics present follow the possessive suffix. Old Hittite also commonly employs the pleonastic possessive construction NOUN/PRONOUN<sub>gen</sub> NOUN + poss. suff., “of X its Y” = “the Y of X.”

Hittite has two demonstrative pronouns of “here” and “there” deixis, *kāš* (inan. *kī*) “this” and *apāš* (inan. *apāt*) “that,” which outside the nominative singular inflect alike: acc. *kūn* (*apūn*, etc.), gen *kēl*, dat.-loc. *kēdani*, abl. *kēz*, instr. *kēt*. The stems are respectively \**kó-* and \**ob<sup>h</sup>ó-*; while the former has numerous cognates elsewhere in Indo-European (like the Germanic family of English *he*, *him*, *her*, dialectal *hit*), the latter is apparently found only in the Anatolian branch (Luv. *apā-* Lyc. *ebe-* “this”).

The interrogative and relative “WH” pronoun is *kuiš*, inan. *kuit*, gen. *kuēl*, dat.-loc. *kuēdani*, abl. *kuēz*. The indefinite pronoun “someone” is *kuiški*, inan. *kuitki*, gen. *kuēlka*, with suffixed particle. Another suffixed particle, geminating *-a* “and,” appears in *kuišša* “each,” inan. *kuitta*; compare Latin *quisque* “each,” with suffixed particle *-que* “and.”

#### 4.4 Verbal morphology

The inflectional categories of the Hittite finite verb are person, number, voice, tense, aspect, and mood.

#### 4.4.1 Person

The persons are the familiar Indo-European first [+ personal, + subjective], second [+ personal, – subjective], third [– personal]: the third person is the zero-person.

#### 4.4.2 Number

As in the noun, only two *numbers* are recognized: singular and plural. The Hittite (and Common Anatolian) first-person plural endings, however, with their characteristic *-w-* (*-weni*, *-wen*) resemble Indo-European first-person dual endings, like Vedic *-vas*, *-va*, Lithuanian *-va*, rather than the first-person plural endings in *-m-* like Vedic *-mas(i)*, *-ma*, Greek *-μεν*, Lithuanian *-me*. Anatolian may thus have originally had a dual in the verb, which was generalized for the first-person plural, on the basis of the discourse-prominent first dual = “you (sg.) and I.”

#### 4.4.3 Voice

Indo-European languages characteristically show a semantic opposition between active and middle; the latter, the marked member, indicates the subject as “internal to” the action. Similar semantics are exhibited by some reflexive verbs in many modern Romance, Germanic, and Slavic languages. Hittite distinguishes active and middle endings in the verb, with the latter also marking the syntactic category of passive as well as subject-internality, reciprocity, and impersonal-hood, as in active *akkiškizzi* “(s)he is dying” versus middle *akkiškittari* “people are dying.” Most verbs in Hittite are inflected as either active or middle only.

The expression of reflexivity and its relation to voice in Hittite is complex. The language has a particle *-za/-az* [-ts], Common Anatolian *\*-ti* of unknown origin, commonly termed “reflexive,” though it has other functions as well. With some transitive active verbs *-za* can express benefit of the subject: *-za . . . dāi* “takes for himself.” For some others it appears to mark a real reflexive object: *nu-za apez arri* “he washes (*arri* active transitive) himself (*-za*) with this.” But some verbs in the language also show an intransitive “middle reflexive” (Garrett 1996) with middle endings, enclitic subject pronoun, and the reflexive particle *-za*: *n=aš=za ārškitta* “he (=aš) is washing (imperfective middle third singular) himself (=za)”; *ŠA KASKAL-NI=za A-az ārrahhut* “wash (imperative middle second singular) yourself (=za) with water of the road!”

#### 4.4.4 Ergativity

As noted earlier, the semantic category of *voice* in the Hittite verb is complicated by its interaction with the syntactic and semantic category of *transitivity*. Neuter nouns functioning as agents, subjects of transitive verbs, must go into the ergative case. The counterpart of this is that the class of third-person enclitic subject pronouns – a class which has no counterpart in any other older Indo-European language – occur only with intransitive verbs, but not with all of these. Specifically, subject clitics occur only with one of the two types of intransitive verb recognized in current syntactic theory: “unaccusative” verbs, with subjects that are less “agentive” and are notionally equivalent to the object of their corresponding transitive counterparts. The other type of intransitive verb is the “unergative,” which has subjects that are more “agentive,” and are notionally identical to the subject of their corresponding transitive counterparts. The repartition is lexically conditioned: in Hittite *šarra-* “break,” “burn,” “hide,” “die,” “go” are unaccusative and take subject clitics, while *tuwarni-* “break,”

“look,” “open,” “speak,” “clean” are unergative and do not. For the contrast between the two intransitive types in the selection of auxiliary, “be” and “have” respectively, in the formation of the periphrastic perfect see further below. For all these questions see Garrett 1990 and 1996, with earlier literature.

#### 4.4.5 Tense-aspect

The Hittite tense-aspect system is relatively simple by comparison with that of Greek or Indo-Iranian. The fundamental tense opposition, expressed by the endings (primary : secondary), is *past* (preterite), the marked member : *non-past* (present, also functioning as future, prospective, and historical present in past narrative), the unmarked member. The stem is the same: past *kuen-ta* “he killed,” non-past *kuen-zi* present “he kills,” future “he will kill.” For the prospective, compare *kuenzi=ma-an* LUGAL-*uš* *ḫuišnuzi=ya=an* LUGAL-*uš* “The king can kill him [or] the king can let him live.” For the narrative present in past time compare: “The Queen thereupon gave birth (pret. *ḫāšta*) to 30 daughters and she raised (pret. *šallanušket*) them herself. (Meanwhile) the sons were going back (pres. *āppa yanzi*) to Nesa and driving (pres. *nanniyanzi*) a donkey. When they arrived (pret. *arer*) in Tamarmara, they said (pres. *taršikanzi*) . . .”

The fundamental aspectual opposition in Hittite is *imperfective*, the marked member, versus the *nonimperfective* base form, root or stem. The primary exponent of the imperfective, usually termed “iterative,” is the suffix *-ške/a-*; sporadic instances of suffixes *-anna/i-* and *-išš(a)-* in similar function are found sometimes marking a particular mode of action or *Aktionsart*. Virtually all Hittite verbs except *eš-* “be” form an imperfective. The imperfective is inflected for tense like the base verb. The tense/aspect opposition can be illustrated by the third singular of the derived (causative) verb *parkunu-* “cleanse, purify”:

- |     |   |  |
|-----|---|--|
| (9) | pres. <i>parkunuzzi</i> “purifies”                | pret. <i>parkunut</i> “purified”         |
|     | impftv. pres. <i>parkunuškizzi</i> “is purifying” | pret. <i>parkunuškit</i> “was purifying” |

Hittite further shows a periphrastic verbal formation usually termed “perfect,” with the past participle and the verbs “have, hold” *ḫar(k)-* and “be” *eš-*. Transitive and unergative intransitive verbs select “have,” and unaccusative intransitives select “be” in the *perfect active*; with “have,” the participle is invariant nominative-accusative neuter, with “be” it agrees with the surface subject: tr. *piyan ḫarta* “had given,” intr. *ḫarkanza ēšta* “had perished.” The value is that of an attained state: *tarahḫan ḫarta* “held conquered.” Transitive verbs select “be” for the *perfect passive*: *piyanteš ešer* “had been given,” *parkunanteš ešer* “had been purified.” The transitive can also form an impersonal, subjectless construction with a direct object: *išḫēniuš=šmaš=kan dān ēšdu* “hairs [acc.]=to them=part. let it be taken,” in other words, “let their hair have been cut.”

#### 4.4.6 Mood

Of the traditional moods the Hittite verb has only indicative and imperative. The Indo-European modal categories of subjunctive and optative, with their respective morphemes *\*-e/o-* and *\*-yeh<sub>1</sub>-/-ih<sub>1</sub>-*, are simply not present. Contrafactual, volitional, and other notions are expressed by the use of the particles *mān*, *man*, with the past or present indicative tense, or by other syntactic means.

The imperative usually shows the bare stem in the fundamental second singular, with traces of the Proto-Indo-European particle *\*-d<sup>hi</sup>i* in *ūt* “go!” = Greek *ἴθι*, as well as a particle *\*h<sub>2</sub>u* with full grade of the same root in the quasi-interjection *eḫu* “come!” Both particles

are suffixed to form the imperative middle second singular: *ārraḥḥut* “wash yourself!” The third-person imperatives replace indicative *-i* with *-u*, agreeing with Vedic Sanskrit: *ēštu* = Vedic *āstu* “let him/her/it be.” The first person expresses volition, the wish of the subject: *ēšlit* “I’d like to be,” “let me be,” with a particle of obscure origin. It has variants *ēšlut*, but *-lit* recurs in one other verb, *talit* “I’d like to take, let me take.” A first singular imperative ending *-allu*, of somewhat different shape and perhaps origin, is also found: *ašallu* “may I be,” *akkallu* “may I die.”

#### 4.4.7 Verb conjugation

The conjugation of the Hittite finite verb is dominated by two sets of endings in the active singular, with no functional difference; they are termed after the first singular present the *mi*-conjugation and the *hi*-conjugation. The basic endings are as follows:

(10)	Present		Preterite	
Sg. 1	-mi	-(ḥ)ḥi (OH -ḥḥe)	-(n)un	-(ḥ)ḥun
2	-ši	-ti	-š (-ta)	-(š)ta
3	-zi	-i (OH -e)	-t(a)	-š
Pl. 1	-weni		-wen	
2	-teni		-ten	
3	-anzi		-er/-ir	

We find a single set of endings of the middle voice, save that some verbs show a third singular in *-a* while others show *-ta*:

(11)	Present	Preterite
Sg. 1	-ḥḥa(ḥa)	-ḥḥa(ḥa)t
2	-ta	-tat
3	-a or -ta	-at or -tat
Pl. 1	-wašta	-waštat
2	-dduma	-ddumat
3	-anta	-antat

The endings of the present may show a further suffixed optional particle *-ri*; those of the preterite may end in *-ti* rather than the usual (apocopated) *-t*.

Middle verbs show then a present third singular in *-a(ri)* or in *-ta(ri)*; the endings are not correlated with *hi*- or *mi*-conjugation actives if the latter are present (most primary middles are inflected in that voice only, and show no active forms): compare *ki-tta(ri)* “lies,” *kīš-a(ri)* “occurs.” Some verbs show *scriptio plena* (repeating the vowel of a CV or VC sign with the matching V sign) in the third singular ending *-āri*, and here the particle *-ri* is obligatory: *tukkāri* “is prescribed, important.” The ending *-ta(ri)* never shows *scriptio plena*. Originally *-ór* → *-ār+i* but unaccented *-(t)or* → *-(t)a* by phonological rule, whence analogical *-(t)a+ri*, which spreads during attested Hittite history (see Yoshida 1990). Secondary thematic middles show only the ending *-ta(ri)*, not *-a(ri)*: *-ietta(ri)*, *-iyatta(ri)*, *-škitta(ri)*.

The special endings of the imperative were given above in section 4.4.6.

#### 4.4.8 Verbal stem-classes

A number of different stem-classes of the Hittite verb may be recognized; to distinguish all or even most of them would exceed the limits of this presentation. Important variables

include stems in final consonant (*ēdmi* “I eat,” *ārḫi* “I arrive”) and in final vowel (*ḫariemi* “I bury,” *tepnumi* “I belittle,” *ḫalziḫḫi* “I call,” *munnaḫḫi* “I conceal”), as well as stems with various types of ablaut (*kuenzi* “kills,” 3rd pl. *kunanzi*; *ēpzi* “takes,” 3rd pl. *appanzi*; *wēkzi* “asks for,” 3rd pl. *wekkanzi*; *ārḫi* “I arrive” 3rd pl. *aranzi*; *dākki* “fits, corresponds,” 3rd pl. *takkanzi*; *sākki* “knows,” 3rd pl. *šekkanzi*), and reduplicated stems (*nanakkušzi* “gets dark,” *lelḫuw(a)i* “pours”). A complete descriptive analysis according to the chronological strata of the language remains a desideratum; the best to date is Oettinger 1979, supplement 1992.

It is noteworthy that while inherited primary athematic *mi*-verbs are common in Hittite, the Indo-European thematic conjugation is found only in active and middle secondary, derived verbs (*-ške/a-* < \**skelo-*, *-ieliya-* < \**ye/o-*). The primary thematic types of Latin *agō*, Greek ἄγω, Vedic *ájāmi*, Latin *uehō*, Greek dial. 3rd sg. impv.  $\text{ἔ}\chi\epsilon\tau\omega$ , Vedic *váhāmi* are not represented at all, and the Hittite thematic first singular active is *-škimi*, *-ieliyami* rather than the ending of Latin *-scō*, *-iō*, etc. The fact is significant; see Jasanoff 1994.

Historically, within the *mi*-conjugation, we have a number of inherited primary formations, derived from the root: athematic presents with ablaut *é* :  $\emptyset$  (*kuen-* : *kun-*, remade in *ēš-* : *aš-*, *ēp-* : *app-*); acrostatic (“Narten”) presents with ablaut *é* : *é* (*ēdmi* : *edwani*, remade in *adweni*); nasal-infix presents (*ḫarni(n)k-* “destroy” beside *ḫark-* “perish”) with probably innovated transitivity value. Of secondary formations, derived from synchronically coexisting stems, we have imperfectives in *-ške/a-* (\**-šké/ó-*); deverbative causatives in *-nu-* and in \**-éyelo-* (Hittite *-ela-*, *waššezi* “dresses (someone)”; deverbative and denominative \**-yelo-* (*karpīya-*, *karpizzi* “lifts” beside root present *karapzi*; *lamniya-* from the noun *lāman-* “name”); statives in *-ē-* (\**-eh<sub>1</sub>-*) and inchoatives in *-ēš-* (\**-eh<sub>1</sub>-s-*), for example, *marše-*, *maršeš-* “be, become false” from the adjective *maršant-* “false,” and the very common derivatives in *-ai-/ā-* from \**-ah<sub>2</sub>-yelo-*, for example, *par(a)šnāizzi* “squats” (cf. Latin *perna* “ham”).

Stem-classes of the Hittite *ḫi*-conjugation are numerous and varied. Primary formations show stems in both final consonant (*ār-ḫi* “I arrive,” 3rd sg. *ār-i*; reduplicated *wewakk-i* “requests”) and final vowel (*tarna-ḫḫi* “I leave,” 3rd sg. *tarna-i*, also *tarn-i*; reduplicated *mimma-i* “refuses” < \**mi-mnV-*). Several old monosyllabic long vowel or diphthongal stems are found: *dāḫḫi* “I take,” 3rd sg. *dāi*; *teḫḫi* (< \**daiḫḫi*) “I place,” 3rd sg. *dāi*; *nehḫi* “I lead,” 3rd sg. *nāi*; *peḫḫi* “I give,” 3rd sg. *pāi*. Secondary *ḫi*-conjugation classes (built on existing stems) are considerably less frequent than *mi*-forms. Note the factitives in *-aḫḫ-* built on adjectives (*šuppiy-aḫḫ-i* “makes pure” from *šuppi-*); the iterative-imperfectives in *-šš(a)-* (*ḫalzi-šša-i* “calls” from *ḫalzi/a-*; *išša-i* “does” but athematic 2nd pl. *ištēni* from *iē-/iya-*); and the “duratives” in *-anna-i* (*iyanna-i* “starts walking” from *ie-/iya-* “walk”).

#### 4.4.9 Origin of the *ḫi*-conjugation

The origins of the *ḫi*-conjugation are surely the most hotly debated in the whole Hittite verb. The endings of the singular are basically those of the classical Indo-European perfect: Greek  $-\alpha$ ,  $-\theta\alpha$ ,  $-\epsilon$ ; compare Latin *-ī* < *-ai*, *-(is)tī* < \**-tai*, *-it* < *-īt* < \**-ei(t)*. But while a very few *ḫi*-verbs agree in meaning but not in form with some Indo-European perfects (*šākki* “knows” like Greek οἶδε, Vedic *veda*), and while a very few look formally like Indo-European perfects (reduplicated *wewakk-i* “requests” beside *wēk-zi* in the same meaning) it has proven impossible to derive the whole *ḫi*-class from such an origin. It is likelier that the *ḫi*-conjugation of Hittite (and the other Anatolian languages) is a reflex of a distinct

present type in Proto-Indo-European originally with affinities to the (proto-)middle voice and singular endings *\*-h<sub>2</sub>e*, *\*-th<sub>2</sub>e*, *\*-e* (with Jasanoff 1994). In Anatolian, this formation then developed into the active *hi*-conjugation, and subsequently in most of the other branches into both the “classical Indo-European” perfect and in part the “classical Indo-European” simple thematic present. This explanation remains controversial, however plausible; for other earlier views compare Cowgill 1979, Kuryłowicz 1979, Eichner 1975, Oettinger 1979 and 1992.

#### 4.4.10 Nonfinite verbals

The nonfinite forms of the Hittite verb include a single adjective or participle, with the suffix *-ant-*, the function of which is to mark the accomplishment of the semantic notion of the verb. With transitive verbs the value is past passive: *ēp-zi* “takes,” *app-ant-* “taken, captive”; with intransitives it denotes an attained state: *ak-i* “dies,” *akk-ant-* “dead.” The suffix is commonly written plene, *ap-pa-a-an-t* = *appānt-*.

Hittite has an infinitive, which functions as a complement of another verb. The infinitive has two forms. Infinitive I *-anna* (var. *-ānna*), to the weak grade of ablauting *-mi*-verbs: *ēp-zi* “takes,” *app-ānna*; Infinitive II *-wanzi* (*-manzi* after stem in *-u-*), to all other verbs: *išḫamai* “sings,” *išḫamiya-wanzi zinnizzi* “stops singing.”

In addition, the imperfective in *-ške-* forms a *supine*, as it is conventionally termed, functioning as complement of the verbs *dai-* “set” and *tiya-* “step, proceed” in the meaning “begin X-ing”: *akkiškiwan dāir* “they began dying.”

The verb can be nominalized to form a neuter verbal noun, in *-war*, with genitive *-waš*: *ganeš-zi* “recognizes,” *kanešš-uwar* “recognition.” Some verbs, including but not limited to ablauting *mi*-verbs, form a verbal noun in *-atar* (*-ātar*), genitive *-annaš*: *app-atar* “taking, seizure,” *akk-atar* “dying, plague.”

The verbal noun *-war*, genitive *-waš* reflects an Indo-European heteroclit *\*-w<sub>ṛ</sub>*, *\*-wen-s* (with “closed” inflection), and the infinitive *-wanzi* is a frozen case form (ablative or instrumental) of the same suffix. That in *-atar*, genitive *-annaš* is from *\*(-ā)-t<sub>ṛ</sub>*, *\*-tn-os*, and the infinitive in *-anna* must be a case form (directive) from the same suffix.

Two isolated instances of a *gerundive* in *-la* are found in a single text: *dalugnula* and *parganula*, “to be lengthened” and “to be made high.”

### 4.5 Derivational morphology

The wealth of secondary *verbal derivational* processes, both inherited and innovated, may be illustrated with forms made from the root/stem *luk(k)-* of athematic *luk-ta* “it grows light” (PIE *\*leuk-/louk-*): *lukk-izzi*, 3rd pl. *lukk-anzi* “set on fire” (*\*louk-éye/o-* in Vedic *rocáyati*, Old Latin *lūcent*), *lukkeš-* “become bright” (cf. Latin *lūcēscere* “grow light”), reduplicated *lalukke-* “be(come) bright, luminous,” *lalukkeš-* “become bright, luminous,” with causative *lalukkešnu-* “give light, illuminate,” and its imperfective *lalukkešnuške/a-* “keep shining.” *Nominal derivatives* from the same root include *lalukkima-* “source of light, radiance,” *lalukkiwant-* “resplendent.” Compare also the set of derivatives underlying *kartimmiyatt-* “anger” given in §4.1.

Other illustrative sets are (from PIE *\*leg<sup>h</sup>-*, English *lie, lay*) active *hi*-verb *lāki* “knocks out (tooth); turns (ear),” middle *lag-āri* “totters,” reduplicated active *hi*-verb *lilakk-i* “causes (a tree) to fall,” and *n*-stem neuter noun *lag-an* “bent, disposition” in *aliyaš laganaš* “of

the disposition of a deer”; (from PIE *\*lah<sub>2</sub>-* in Greek λαῖ *ἄρως* “the people under arms”) *lahḫ-iy-āizzi* “goes on campaign, to war,” verbal noun *lahḫ-iy-war* in genitive *lahḫiyawaš išḫiul* “the obligation of going to war,” abstract *lahḫ-iy-atar* “campaign,” *lahḫ-e-ma* “errand” in *lahḫemuš ḫueškizzi* “he is always running errands” (*ḫuwai/iya-* “run”). Note the imperfective *lahḫ-e-ški-ši* “you go to war, too,” and the unique Neo-Hittite doubly specified iterative-durative creation *lahḫ-iy-anni-ška-weni* “we shall always go to war,” which shows how freely these morphemes could be manipulated.

## 4.6 Compounds

Hittite makes considerable use of *semantic compounding* of sentential adverb (“preverb”) and verb, while maintaining the phonological independence and separability of the two elements: *anda paizzi* “goes in,” *āppa paizzi* “goes back, returns,” *āppan paizzi* “goes after, behind,” always written with a space between the two. Two preverbs are frequent: *āppan arḫa paizzi* “passes behind,” *piran arḫa uizzi* “passes in front of.” The first preverb may be fronted and separated from the verb: *āppa=ma=aš kuwapi uizzi* “but when he returns.” Such semantic compounding occurs also in the nominalization of verb phrases of object and verb: *kurur ēpzi* “makes/begins hostility/hostile action,” whence *kurur appatar* “making hostility, declaring war.” But the phonological composition of two lexical elements to form a single phonological word is extremely rare in Hittite. The case of *šallakard-* “arrogant, arrogance” (*šalli-* “great,” *kard-* “heart”) underlying several verbal derivatives has been noted; the example of *pattarpalḫi* “kind of bird observed in divination” (*pattar* “wing,” *palḫi-* “broad”) was shown to be a loan-translation (calque) on Akkadian *kappu-rapaš* “id.” (*Chicago Hittite Dictionary* s.v., with references). Occasional geographic names like *ḫarašḫapaš* “Eagle River” (*ḫaraš* “eagle,” *ḫapaš* “river”) are juxtapositions, not true compounds with first member in stem form. Negative composition with the *-ant-* participle is found in *ām(m)iyant-* “immature” from *\*ḫ-mih<sub>1</sub>-ont-*, probably a (frozen?) archaism, cf. Vedic *āsant-* “untrue, false” from *\*ḫ-h<sub>1</sub>s-ent-*. Otherwise Hittite (with other Anatolian languages) has a very few negative compounds in *ni-* (*niwalla-* “weak”), apparently from the old sentence negation *\*ne* or *\*nei*. A unique numerical compound is *dā-yugaš* “two-year-old” (see §4.7).

## 4.7 Numerals

The numerals in Hittite texts are virtually always written in cuneiform ciphers, and almost never written out. We are left with inferences from a few forms and derivatives. See on all these Eichner 1992.

1. The very occasional writing of a stem *a-a-* (*ant-*), which may be the stem of “one,” *\*a[y]-ant-*, *\*oy-(ónt-)*, cf. Old Latin *oi-nos*, Vedic *éka-* < *\*ó-i-ko-*, Avestan *aēuua-* < *\*oi-wo-*. For the ordinal “first” *ḫantezzi(ya)š* is used, derivative of the adverb *ḫantī* “apart, in front,” from *ḫant-* “front, forehead.”
2. The numerical adverb *tān* (*dān*) “second(ly),” juxtaposed in such expressions as *dān pedaš* (gen. sg.) “of second place, rank” and prefixed (with loss of *n* before *y*) in the compound (?) *tā-yugaš* “two-year-old” (of animals) is apparently from PIE *\*dwo-yo-m*. The military term <sup>LÜ</sup>*duyanalliš* “man (officer) of the second rank” from *\*dwi-yo-* is probably from its phonology a Luvian loanword.
3. The numeral “three” is exceptionally written out in *teriyaš UD-aš* (gen. pl.) “of three days.” The Hittite stem *teri-* shows a real anaptyctic vowel, from IE *\*tri-*; its Luvian



counterpart *tarri-* (in <sup>LÜ</sup>*tarriyanalliš* “man of the third rank,” etc.) shows a special Luvian development of the same vowel.

4. Hittite and the Luvian languages, thus perhaps already Common Anatolian, apparently replaced the Indo-European word for “four” by a neologism of disputed inflection and origin, occasionally written out as *mēuw* (-*aš* dat. pl.), *meu-*, and in Luvian *māuw* (-*ati* abl.-instr.). It may go back to a *\*meyu-/myeu-* (simplified to *meu-*?) and originally have meant “little” hand (minus the thumb).

The remaining numerals are never written out, and can only be guessed at, with the exception of the ritual libation drink *šiptamiyal 7-miya* (beside *teriyalla/ 3-yalla*), which doubtless contains a reflex of PIE *\*septm̥* “seven” (beside *teri-* “three”) in cardinal, ordinal, or fractional function.

The formation of ordinals is not clear. For other suffixed forms, like the distributive *-anki* “*n*-times” see Eichner 1992, as well as the several dictionaries (under Sumerograms).

## 5. SYNTAX

### 5.1 Word order and clause structure

Hittite in its ordinary unmarked word order is by and large regularly verb-final (OV [Object–Verb] in the case of transitive verbs), with the possibility of emphatic initial position of the verb (VO in the case of transitive verbs) as special or marked order. In fact, we can distinguish a number of different syntactic constituents in the Hittite sentence which show a fairly fixed order relative to each other. These include the *sentence connective particles* (symbol N) *nu* (OH also *ta, šu*) which regularly begin most clauses. The virtually obligatory use of overt markers, sentence-initial or enclitic, to connect all but the first sentence in a discourse is one of the three defining syntactic isoglosses of Common Anatolian (Melchert 1994).

Sentence-initial particles or other words may be followed by one or more (up to six) *enclitics* (symbol E), which thus occupy second, *Wackernagel's Law*, position. The enclitic chain of particles and anaphoric pronouns is one of the most striking and salient features of Hittite syntax, and indeed another defining syntactic isogloss of Common Anatolian. The enclitic chain may include members of each of six ordered classes: (i) connectives =*a* (geminating), =*ya* “and,” =*a* (nongeminating) “but, however,” correlative focus =*ma*, weakly adversative adding new information; (ii) quotative particle =*wa(r)* marking direct speech; (iii) pronominal third-person object accusative (of transitive verbs) or subject nominative (cf. certain intransitive verbs); (iv) pronominal third-person dative-locative or first-/second-person oblique; (v) reflexive particle =*za(=z)*; (vi) local (/aspectual?) sentential particles =*kan*, =*šan*, =*ašta*, OH =(a)*p(a)*, =*an*. While usually no more than three of these are present, up to five are not uncommon, for example, =*ma=war=an=z=šan*.

Conjunctions like *mān* “when, if,” OH *takku* “if,” *māḥḥan* “as, how, when” occupy the C(omplementizer) slot, following optional connective and enclitic(s) (X indicates the rest of the sentence):

(12) ## (N) (E) C X ##

If no N is present, the C slot becomes sentence-initial, and E follows if present. E always occupies sentence second position. Thus:

- (13) A.  $n=aš\ māḥḥan \dots$  “when he . . .”  
           N   E   C  
       B.  $māḥḥann=a=kan \dots$  “and when . . .”  
           C   E<sub>1</sub> E<sub>2</sub>  
       C.  $n=ašta\ mān \dots$  “(then) if . . .”  
           N   E   C  
       D.  $mān=a=šta \dots$  “but if . . .”  
           C   E<sub>1</sub> E<sub>2</sub>

When conjunctions *mān* and *takku* are initial and followed by optional enclitics, the enclitic =*ma* is in Old Hittite and Middle Hittite delayed to the second word in its clause: *mān* <sup>URU</sup>*Ḥattuša=ma* “but when to Hattusas . . .,” *mān āppa=ma* <sup>URU</sup>*Nēsa* “but when back to Nesas . . .” This rule is no longer observed in Neo-Hittite (*mān=ma passim*); we have a diachronic syntactic change.

Any sentence element can be fronted, by moving into a TOP(icalization) Phrase to the left of C, thus

- (14) ## (N) (E) <sup>TOP</sup> (C) X ##  
        $n=ašta$  <sup>DIM</sup>*unni-ma mān* “and when to the Stormgod . . .”

If no N is present, as often for C and usually for TOP, we have

- (15) ## <sup>TOP</sup> (E) C X ##  
        $kinun=a=wa mān$  “but now how . . .?”

Coordinated clauses of the type “if X, (then) Y,” “when X, (then) Y,” “because X, (then) Y,” are almost always in that order (X,Y). Similarly in “subordinate” clauses the R(elative) C(lause) virtually always precedes the M(ain) or M(atrix) C(lause). The basic rule for *indefinite* RC (“whoever”) is “Move *kui-* (“*wh-*”) to C(omplementizer)”:

- (16) ## (N) (E) *kui-* X V ##

That for definite RC (“s/he who”) is the above rule, followed by a fronting rule, “Move something else to TOP(icalizer), to the left of *kui-* in C”:

- (17) ## (N) (E) \_\_\_\_ *kui-* X V ##

Compare, with connective, *nu kuiš MEŠEDI* “whichever guard” beside *nu 1 MEŠEDI kuiš* “The one guard who.” In the absence of connective and other elements save V, we have just two-word sentences: *kuiš paprezzi* “Whoever is unclean” beside *paprezzi kuiš* “He who is unclean.”

These movement rules are complex. In the following example, the interrogative *wh-* has been fronted around the RC, even though its domain is actually the MC:

- (18)  $kuwat=wa$     *URUDU-an*    *kuin*    *lāḥun*     $nu=wa=mu$     *āppa*    *ḥūrzakizi*  
       Why=quot.    copper            which    I cast    N=quot.=me    back    is cursing  
       “Why is the copper which I cast cursing me back?”

In the following example the whole of the RC has fronted to precede the *wh-* word “because” in C:

- (19) *nu* <sup>m</sup>*Aparruš* *LÚ* <sup>URU</sup>*Kalašma* *kūruriahta* *kuit*  
 N Aparrus the Kalasmean had begun hostilities because  
*nu=za* 3 LIM KASKAL *ninikta*  
 N=refl. 3,000 army he raised  
 “Because Aparrus the Kalasmean had begun hostilities he raised an army of 3,000”

Note also that the antecedent of the relative clause commonly appears in both the RC and the MC:

- (20) *ḥaššikkitten* *kuedani* *šiwatti* *nu=wa* *kāša* *apēl* *šiwattaš* *laleš*  
 you quarreled on which day N=quot. behold of that day the tongues  
 “Behold the tongues of the day on which you quarreled”

## 5.2 Agreement

In Hittite and other Anatolian languages agreement is generally like that of other old Indo-European languages: attributive adjectives, participles, and pronouns agree in gender, case, and number. In conformity to their origin as collectives, neuter plurals take a singular verb, as in some other old Indo-European languages. Agreement in predicates, nominal and verbal, frequently shows *constructio ad sensum*, especially in number: *n=an* GIM-*an* KUR-*eanza aušta* *n=at* *naḥšariyandari* “When the land (erg. sg.) saw him, they (nom. pl.) were afraid (3rd pl.).” In a single sentence the animate plural and collective (neuter plural or singular) may shift back and forth freely: *nu=mu* MUŠEN<sup>HLA</sup> *kue uppešta* *n=at* *arḥa ḥarranteš ešir* *n=aš edunn=a* *ŪL ūḥḥunn=a=aš* *ŪL man=at* SIG<sub>5</sub>-*anteš* *man=at* *ŪL* “the birds (neut. pl. acc. coll.) which (neut. pl. acc.) you sent, they (anim. pl. nom.) were spoiled (3rd pl.), and I neither ate them (anim. pl. acc.) nor did I see them (anim. pl. acc.) whether they (anim. pl. nom.) [were] good (anim. pl. nom.) [or] whether they (anim. pl. nom.) [were] not.” The sentence is a good example of Hittite complex clauses; notice the right dislocation of the two negatives, and the respective positions of the coordinated verbs *n=aš edunn=a* and fronted *ūḥḥunn=a=aš* vis-à-vis their enclitic objects. The underlying presence of enclitic *=a* “and” in the latter is guaranteed by the gemination of the final *n* of the first singular preterite *-un* in both verbs. Historically, *=a* “and” developed from *\*h<sub>2</sub>o* (Luv. =*ḥa*), and the gemination reflects generalization to all consonants of a phonetic rule *-VRHV- → -VRRV-*. Compare *kuišša* “each,” acc. *kuinna*, with Luvian *kuišḥa* “some/any(one),” acc. *kuinḥa*.

## 5.3 Syntactic innovation

The most interesting and striking syntactic innovations of Hittite and the other Anatolian languages are doubtless the system of split ergativity and the related development of enclitic subject pronouns with certain (“unaccusative”) intransitive verbs, both of which (with references) have been discussed above. Both are illustrated in the sentences cited in the preceding paragraph. Likewise striking is the enclitic chain of Hittite and the other Anatolian languages, also discussed above (see §§4.2.3, 4.4.4). As the incorporation into the chain of the innovated subject pronouns would suggest, the elaboration of these lengthy chains of enclitic elements, particles, and pronouns is itself a syntactic innovation built on much smaller inherited beginnings. Given the presence of such apparent chains in the unrelated Hattic, and to some extent Hurrian, languages it is likely that we are in the presence of an Anatolian areal feature.

## 6. LEXICON

The core grammatical structure of Hittite in nominal, pronominal, and verbal morphology is clearly Indo-European, with a few innovations like the loss of the categories of subjunctive and optative mood, the comparative *\*-yos-* (but not contrastive *\*-tero-*) and superlative, and the feminine gender. Features like the *hi*-conjugation are now seen to be archaisms, and the perfect and thematic conjugation of later Indo-European are innovations posterior to the separation of Anatolian. The former view, current early in this century, that the Hittite lexicon was largely of “foreign” non-Indo-European provenience, has proved incorrect. That false impression was due to the technical character of the vocabulary of a large number of texts dealing with public and private ritual, cult, augury, and the like, on the one hand, and details of social and military organization, on the other, all of which reflect the culture of Asia Minor and the Near Eastern world of the second millennium BC. On the level of basic core vocabulary, Hittite (and the rest of the Anatolian family) is firmly Indo-European. The existence of two large and ongoing etymological dictionaries or glossaries (Tischler 1977–, Puhvel 1984–) and a monumental Anatolian historical phonology (Melchert 1994) is ample testimony to the mine of information now available on the Indo-European origins of the Hittite lexicon. About half of the 230-odd Indo-European roots cited in Watkins 2000a, 2000b are represented in Hittite, and studies of selected semantic fields like body parts show a high percentage with Indo-European etymology. “Foot” (*pad-*) agrees with English and Greek, and “hand” (*keššar-*) with Greek; if “tooth” (*kaga-*) is cognate with English “hook,” we need only recall that the Slavic and Tocharian words for “tooth” are cognate with English “comb.” Sometimes the Hittite facts require revision of accepted semantic views: thus the usual Indo-European verb “drink” (*\*pō-*, *\*poh<sub>3</sub>-*) means “take a swallow” in Hittite (*pāš-*), and the usual Hittite verb “drink” (*eg<sup>w</sup>-*, *eku-*) has a cognate in Tocharian, and otherwise survives only residually in the Greek verb for “go without drink” and the Latin for “drunk” (*eb-rius*).

The Hittites settled in their homeland of central Anatolia when it was already populated by urbanized non-Indo-European Hattic speakers, and they borrowed or absorbed many features of Hattic culture, especially in the sphere of religion and cult. Our documents include many bilingual Hattic-Hittite texts, and the continued use of Hattic as a cultic language in the Old Kingdom accounts for numerous lexical and onomastic borrowings in this cultural area. The existence of a Hattic substratum of speakers having given up Hattic for Hittite (or dialects related to each) in the early centuries of the second millennium or earlier may also be ultimately responsible for such apparently areal syntactic features as the clitic chain in Hittite and other Indo-European languages, or such areal phonological features as the preservation as *h*, *h<sub>1</sub>h* of two of the three Proto-Indo-European laryngeals.

Already in the nineteenth century BC the Hittites in Kaneš (Nešaš) were in contact with the Semitic world, with the Old Assyrian merchant colonies. The cuneiform documents of the latter attest intermarriage and far-reaching cultural and economic contact between Assyrians and Anatolians, many of whom were Hittite-speaking. The use of writing in cuneiform on clay tablets came to the Hittites only later, from contacts with Peripheral Akkadian speakers and scribal schools in Northern Syria writing a form of Akkadian similar to Old Babylonian. Akkadian was, at the beginning of the historical period, the language of Hittite literary productions like the Siege of Uršu, and of bilinguals conceived in Hittite and then translated into Akkadian; it continued in use for ceremonial purposes in Middle Hittite, witness the inscription on an Aegean sword by Duthaliyaš II commemorating his destruction of the Aššuwa coalition, and throughout the Middle Kingdom and empire as the language of some treaties and international correspondence.

The Hittites were in contact with the non-Indo-European Hurrians from at least the time of the Old Kingdom on, and the early hostilities were succeeded in Middle Hittite times by a period of intense cultural symbiosis, particularly in religion and cult. The translation of some Hurrian texts and the composition of others on Hurrian models was a major factor in the flowering of Hittite culture, and the Hurrian linguistic legacy in the technical terminology of ritual as well as the onomastics of the new pantheon was immense.

The Hittites were also in continual contact since the Old Kingdom with other Indo-European languages of Anatolia. Palaic, the language of Palā to the northeast (classical Paphlagonia) was preserved as the language of local cults in a few tablets in Hattusas; the language appears to have died during the time of the Old Kingdom. Luvian, the language of Arzawa to the west and Kizzuwatna to the south, is attested in a number of rituals of Kizzuwatnan provenience in Hattusas from the Old Kingdom on. Both onomastics and prosopography attest a growing Luvian presence throughout Hittite history, and with the establishment of the empire and probably long before we may assume widespread Luvian–Hittite, bilingualism. Already at the end of the Old Kingdom or the beginning of Middle Hittite one text describes some orders to the palace guard to be given in Hittite, and others in Luvian. The use of the Hieroglyphic Luvian script and language for monumental and identificational (glyptic) purposes surely contributed to this linguistically cosmopolitan atmosphere which is so characteristic of Hittite culture. From the last recorded Hittite king, Suppiluliumas II, we have not only Hittite cuneiform texts, but two Hieroglyphic Luvian monuments setting forth his manly deeds, as well as the pathetic letter in Akkadian inquiring after the *šikalāyu*, the Sea Peoples who brought about his own and his empire's downfall.

## Bibliography and further reading

### References

- Beckman, G. 1996. *Hittite Diplomatic Texts*. SBL Writings from the Ancient World Series. Atlanta: Scholars Press.
- Cowgill, W. 1979. "Anatolian *hi*-conjugation and Indo-European perfect: Installment II." In E. Neu and W. Meid (eds.), *Hethitisch und Indogermanisch*. Innsbruck: Innsbrucker Beiträge zur Sprachwissenschaft.
- Cowgill, W. and M. Mayrhofer. 1986. *Indogermanische Grammatik*, vol. I. Heidelberg: Carl Winter.
- Dietrich, M. and O. Loretz. 1978. "Das 'seefahrende Volk' von Šikila (RS 34.129)." *Ugarit Forschungen* 10:53–56.
- Eichner, H. 1973. "Die Etymologie von heth. *mēhur*." *Münchener Studien zur Sprachwissenschaft* 31:53–107.
- . 1975. "Die Vorgeschichte des hethitischen Verbalsystems." In H. Rix (ed.), *Flexion und Wortbildung*, pp. 71–103. Wiesbaden: Dr. Ludwig Reichert.
- . 1992. "Anatolian." In J. Gvozdanović (ed.), *Indo-European Numerals*, pp. 29–96. Berlin/New York: Mouton de Gruyter.
- Garrett, A. 1990. "The origin of NP split ergativity." *Language* 66:261–296.
- . 1996. "Wackernagel's law and unaccusativity." In A. L. Halpern and A. M. Zwicky (eds.), *Approaching Second: Second Position Clitics and Related Phenomena*, pp. 85–133. Stanford: CSLI Publications.
- Hoffner, H. A., Jr. 1990. *Hittite Myths*. SBL Writings from the Ancient World Series. Atlanta: Scholars Press.
- . 1997. *The Laws of the Hittites. A Critical Edition*. Leiden: Brill.
- Jasanoff, J. 1994. "Aspects of the internal history of the PIE verbal system." In G. Dunkel, G. Meyer, S. Scarlata, et al. (eds.), *Früh-, Mittel-, und Spätindogermanisch*, pp. 149–167. Wiesbaden: Otto Harrassowitz.

- . Forthcoming. *Hittite and the Indo-European Verb*. Oxford: Oxford University Press.
- Kurylowicz, J. 1979. "Die hethitische *hi*-Konjugation." In E. Neu and W. Meid (eds.), *Hethitisch und Indogermanisch*, Innsbruck: Innsbrucker Beiträge zur Sprachwissenschaft.
- Labat, R. 1976. *Manuel d'épigraphie akkadienne* (5th edition). Paris: Imprimerie Nationale.
- Laroche, E. 1971. *Catalogue des textes hittites*. Supplements RHA 30, 33. Paris: Klincksieck.
- Melchert, H. C. 1994. *Anatolian Historical Phonology*. Leiden Studies in Indo-European 3. Amsterdam/Atlanta, GA: Rodopi.
- Morpurgo Davies, A. 1982–1983. "Dentals, rhotacism, and verbal endings in the Luwian languages." *KZ* 96:245–270.
- Neu, E. 1974. *Der Anitta-Text*. Studien zu den Boğazköy-Texten 18. Wiesbaden: Otto Harrassowitz.
- Oettinger, N. 1979. *Die Stammbildung des hethitischen Verbums*. Nürnberg: Carl.
- . 1992. "Die hethitischen Verbalstämme." In O. Carruba (ed.), *Per una grammatica ittita*, pp. 213–249. Pavia: Uculano.
- Otten, H. 1973. *Eine althethitische Erzählung um die Stadt Zalpa*. StBoT 17 Wiesbaden: Otto Harrassowitz.
- . 1995. *Die hethitischen Königssiegel der frühen Grossreichszeit*. Stuttgart: Steiner.
- Puhvel, J. 1984–. *Hittite Etymological Dictionary*. Berlin: Mouton de Gruyter.
- Rüster, C. and E. Neu. 1989. *Hethitisches Zeichenlexikon*. StBoT, Suppl. 2. Wiesbaden: Otto Harrassowitz.
- Tischler, J. 1977–. *Hethitisches etymologisches Glossar*. Innsbruck: Innsbrucker Beiträge zur Sprachwissenschaft.
- Watkins, C. 2000a. *Indo-European and the Indo-Europeans. The American Heritage Dictionary of Indo-European Roots* (2nd, revised edition), pp. vii–xxxv. Boston: Houghton Mifflin.
- . 2000b. *Indo-European and Indo-Europeans. The American Heritage Dictionary of the English Language* (4th, revised edition). Boston: Houghton Mifflin.
- Yoshida, K. 1990. *The Hittite Mediopassive Endings in -ri*. Berlin: de Gruyter.

### Further reading

#### Grammar:

Friedrich, J. 1960. *Hethitisches Elementarbuch I<sup>2</sup>*. Heidelberg: Carl Winter.

#### Dictionaries:

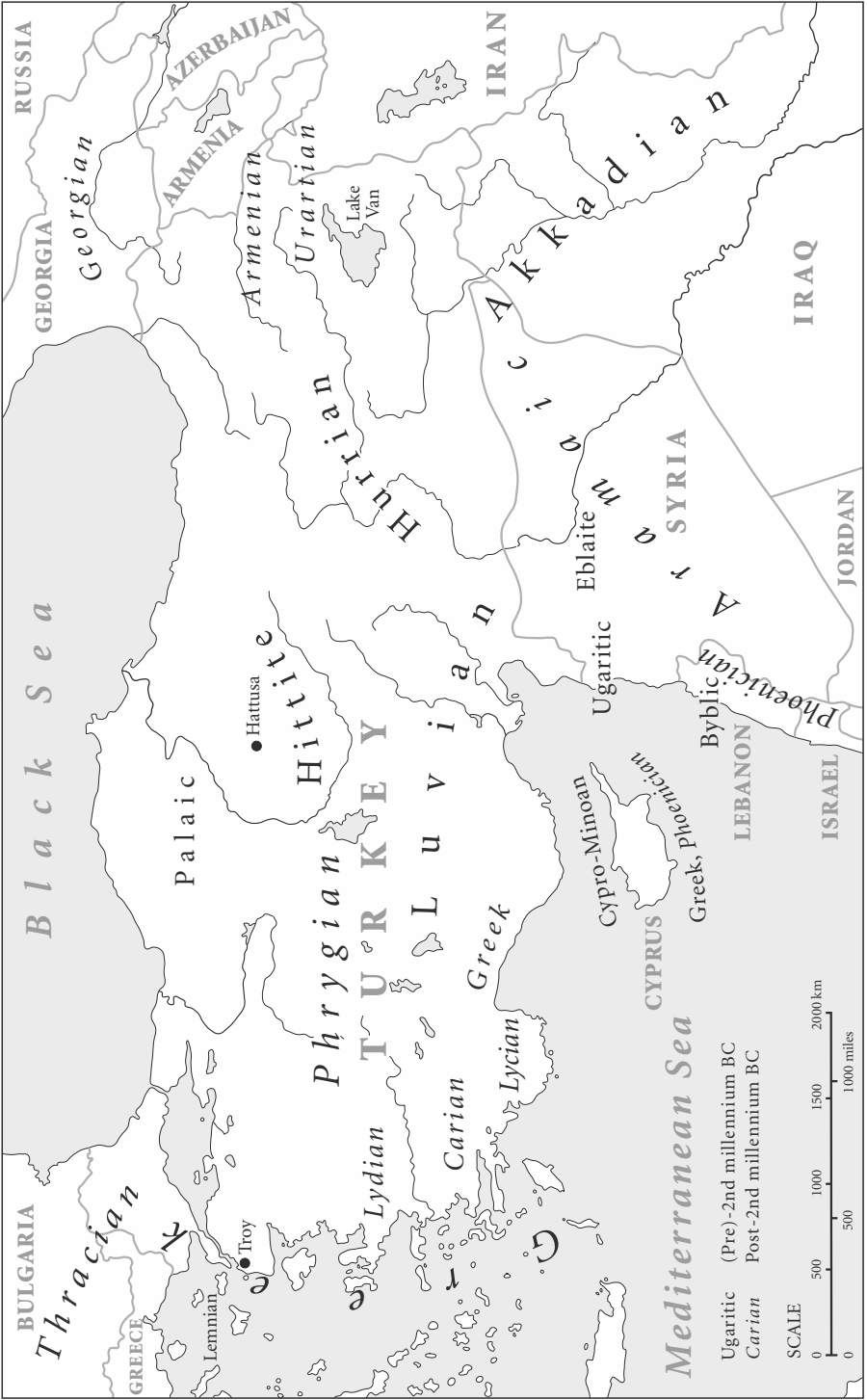
- Friedrich, J. 1952. *Hethitisches Wörterbuch, mit Ergänzungshefte* (3 vols.). Heidelberg: Carl Winter.
- Friedrich, J. and A. Kammenhuber. 1975–. *Hethitisches Wörterbuch* (2nd edition). Heidelberg: Carl Winter.
- Güterbock, H. G. and H. A. Hoffner, Jr. 1980–. *The Chicago Hittite Dictionary*. Chicago: Oriental Institute.

#### Texts:

See the two series Studien zu den Boğazköy-Texten (StBoT) and Texte der Hethiter. Representative texts are translated by A. Goetze in Prichard, James. 1955. *Ancient Near Eastern Texts relating to the Old Testament*. Princeton: Princeton University Press. See also Hoffner 1990 and Beckman 1996.







**Map 2.** The Ancient languages of Anatolia and surrounding regions (languages which are first attested in the first millennium BC or later are marked in italics)

# Luvian

H. CRAIG MELCHERT

## 1. HISTORICAL AND CULTURAL CONTEXTS

Luvian (or Luwian) was arguably the most widely spoken member of the Anatolian subgroup of Indo-European. Evidence for the language is twofold. First, the cuneiform archives of the Hittite capital Hattuša in central Anatolia contain a number of texts with passages in a language designated *luwili*; that is, of the land *Luwiya*, which the Old Hittite Laws list as one of three major divisions of the Hittite state.

Starke (1985) has shown in his excellent edition of the Cuneiform Luvian (CLuvian) corpus that the apparently extensive texts actually represent variations on scarcely a dozen distinct compositions (aside from a few fragments). With one or two exceptions, the texts are rituals, some of a private, therapeutic nature, others belonging to the state cult. The CLuvian manuscripts, like the Hittite, date from the sixteenth to thirteenth centuries BC, including a few from the Old Hittite period (see Ch. 18, §1). Beyond this highly restricted material, there are also many isolated Luvianisms scattered throughout the Hittite texts, both as foreign words and as genuine loanwords adapted to Hittite inflection. Starke (1990 and elsewhere) has demonstrated that Luvian influence on Hittite was both earlier (including prehistoric) and more extensive than previously acknowledged. However, the fact that the two languages are very closely related makes it difficult to distinguish with certainty Luvian loanwords into Hittite from native Hittite cognates of Luvian lexemes, and not all of Starke's claims are equally persuasive (see Melchert 1992).

The second source for Luvian consists of texts written in a system of Anatolian hieroglyphs. Aside from a few letters and economic documents on soft lead strips, the vast majority are monumental inscriptions on stone, both natural rock-faces and man-made structures. A few date from the fourteenth and thirteenth centuries BC, the later period of the Hittite Empire, and most of these are attributable to known Hittite kings. Most Hieroglyphic Luvian (HLuvian) texts, however, date from the tenth to seventh centuries, after the fall of the Hittite Empire itself, and describe the activities of local rulers and their subordinates in the various newly independent small states of southern Anatolia and northern Syria. These inscriptions on stone are generally dedicatory in content, but often contain lengthy historical sections.

Both references in the Hittite texts and the geographical distribution of the extant HLuvian inscriptions suggest that the Luvian "heartland" lay in southern and southwestern Anatolia, penetrating into what is now northern Syria. However, inscriptions have also been found in western and central Anatolia (including at Hattuša itself). The status of Luvian as a spoken language in the latter areas is quite unclear. The influence of Luvian on Hittite, particularly in the Late Empire period, has led to suggestions that by this time Luvian was the spoken

language in Hattuša, with Hittite preserved only as a written “chancellery” language. This possibility must be taken seriously, but it should be regarded as merely one of several hypotheses.

Also problematic is the precise relationship between CLuvian and HLuvian. There is nothing to recommend the view that CLuvian is in any sense the “Hattuša dialect” of Luvian. The highly restricted nature of the CLuvian evidence and limited understanding of the contemporary HLuvian inscriptions of the second millennium preclude any definitive statements at present. The prudent current consensus is to treat the two as closely related but independent coequal dialects of a single language with no further presumptions. A last complication to be mentioned is the chance that one set of CLuvian texts, the so-called “Istanuvian songs,” represents a dialect distinct from the rest of CLuvian (and HLuvian). The evidence is suggestive (see the references in Melchert 1994a:11), but the Istanuvian texts are too poorly understood to assert anything with confidence. Radical revision of the readings of many basic HLuvian signs by Hawkins, Morpurgo Davies, and Neumann (1974) has shown that differences between CLuvian and HLuvian are minimal. The single description which follows is meant to apply to both unless stated otherwise.

## 2. WRITING SYSTEMS

### 2.1 Cuneiform Luvian





















CLuvian was written by Hittite scribes, using the same version of the Old Babylonian syllabary that they used to write Hittite, and the writing conventions are mostly the same (see Ch. 18, §2). Word spacing and paragraph dividers are used consistently. Logograms are less frequent than in Hittite, but more common than in Palaic. One should certainly regard *scriptio plena* (repeating the V of a VC or CV sign with the matching V sign) as a marker of vowel length (Melchert 1994a:27, extending the claims of Kimball and Eichner for Hittite). The contrast of intervocalic single and geminate consonants is significant as in Hittite, however one interprets this in phonetic terms (see the lengthy discussion with references in Melchert 1994a:13ff.). The most striking spelling habit is the free use of word-initial *scriptio plena*, almost nonexistent in Hittite: CLuvian *i-i-ti* for /idi/ “goes,” *a-an-ta* for /ānda/ “in(to).” Since it is virtually certain that Luvian does not distinguish /i/ and /e/, the sporadic use of cuneiform signs with *e*-vocalism for /i/ is surely insignificant.

Emil Forrer in 1919 already established CLuvian as a distinct language with close affinities to Hittite. Further significant progress came following World War II with the publication of a large number of texts and analyses by Bernhard Rosenkranz, Heinrich Otten, and Emmanuel Laroche. A new era in CLuvian studies began with the publication of the radically reorganized corpus by Starke (1985).

### 2.2 Hieroglyphic Luvian

The Anatolian hieroglyphs are first attested on Hittite personal seals dating from the fifteenth and fourteenth centuries. These inscriptions, consisting of names, titles, and good luck signs, can hardly be said to represent texts in a given language. Except for a few isolated cases (Uartean glosses on pithoi, Hurrian divine names in the pantheon at Yazilikaya), all genuine texts in the hieroglyphs discovered thus far are in Luvian. That the system was invented for writing Luvian is supported by evidence from acrophony (secondary use of a logogram as a phonetic sign based on the initial sequence of the word represented): e.g., the

**Table 19.1** Examples of Hieroglyphic Luvian logograms

Character	Value	Character	Value	Character	Value
	"head"		"god"		"offer"
	"I"		"moon"		"put"
	"speak"		"sun"		"below"
	"anger"		"boundary"		"above"
	"king"		"scribe"		"after"
	"kingdom"		"Aleppo"		"before"
	"camp"		"give"		

sign *tara/i-* clearly is based on Luvian /tarri-/ "three" (vs. Hittite *tēri-*). The precise time and place of development of the hieroglyphs and the relationship of their use on seals to that for writing texts remain to be elucidated: see Hawkins 1986 for further discussion.

The HLuvian system is mixed logographic-syllabic. A word may be written as a logogram (particularly common in the second-millennium texts), entirely phonetically, or as a logogram with phonetic complements. The logogram may also stand before the complete phonetic word as a determinative (semantic marker). The nominative singular of "cow" may thus be written in four ways: (i) BOS (by a now widely accepted convention, logograms are with a few exceptions represented by capitalized Latin equivalents); (ii) *wa/i-wa/i-sa*; (iii) BOS-*wa/i-sa*; or (iv) (BOS)*wa/i-wa/i-sa* (where BOS is used as a determinative). The phonetic portion of the system includes only signs for V and CV sequences (and rarely CVCV). This fact means that neither word-final consonants nor any consonant clusters may be directly represented: see *wa/i-wa/i-sa* for /wawīs/ above. One should note in particular the failure to indicate preconsonantal nasals: the animate nominative plural ending /-intsi/ is spelled -*Ci-zi*. For a provisional list of logograms and phonetic signs see Hawkins 1975:153ff.

The system does not distinguish simple from geminate consonants nor a possible voicing contrast in stops. Repetition of the vowel of a CV sign by a V sign does **not** indicate vowel length, but is regulated by aesthetic principles (there is a strong tendency to fill available space evenly). The syllabary only gradually and imperfectly developed separate CV signs for /a/ and /i/ vocalism, hence the rather awkward use of spellings like *wa/i-wa/i-sa*. For more on the system see Melchert 1996.

It had long been surmised that the monumental inscriptions in Anatolian hieroglyphs were associated with the Hittite Empire, but it was only the discovery of the cuneiform Hittite documents at Hattuša/Boğazköy in the early twentieth century that permitted serious work on deciphering the hieroglyphs. Emil Forrer, Bedřich Hrozný, Piero Meriggi, and Ignace Gelb all made important pioneering contributions, and by 1940 a partial decipherment of the script and basic understanding of many texts had been achieved. It was also clear that "Hieroglyphic Hittite" was closely related to cuneiform Hittite and Luvian, but the precise relationship remained uncertain. The discovery of an extensive Hieroglyphic-Phoenician

bilingual at Karatepe in 1947, published in stages by Helmut Bossert, promised to revolutionize study of the hieroglyphs, but it was not until the mid-seventies that David Hawkins, Anna Morpurgo Davies, and Günter Neumann, building on work of Bossert, could demonstrate that the phonetic readings of a number of basic signs required radical revision. The major reassessment demanded by these changes confirmed the early claim of Meriggi that the language of the hieroglyphs is a form of Luvian, and indeed one very closely related to CLuvian.

The multiple ambiguities of the HLuvian syllabary mean that analysis of Luvian phonology is based primarily on CLuvian data. In compensation, the much more varied content of the HLuvian texts tends to give a broader picture of Luvian morphology.

### 3. PHONOLOGY

#### 3.1 Consonants

The Luvian phonemic inventory consists of at least the following consonants:

##### (1) Luwian consonantal phonemes

p	t	k
b	d	g
	ts	
	s	h
		ʃ
m	n	
	l	r
w		y

The absence of positive evidence for a unitary labiovelar /k<sup>w</sup>/ as in Hittite may be accidental, and words like *ku-i-* “who, which” from \*k<sup>w</sup>i- may still contain /k<sup>w</sup>/. It is certain that there is no corresponding voiced stop, because inherited \*g<sup>w</sup> merges with \*w: recall /wawī-/ “cow” from \*g<sup>w</sup>ow- above. As already noted, the orthography of stops in CLuvian follows the same principles as in Hittite (intervocalic contrast of simple vs. geminate; see Ch. 18, §3.1), although the specific distribution diverges due to different prehistoric changes. Interpretation of this orthography remains controversial. HLuvian obviously can render no assistance, but the restriction of rhotacism (see below) to the voiced dental stop confirms that some kind of phonemic contrast remained between inherited voiceless and voiced stops, whatever its precise synchronic realization.

The sound conventionally transliterated *z* represents sequences of /t/ + /s/, as well as the result of prehistoric assibilation of \*t before \*y and Proto-Indo-European palatal \*k̑ (for the last see Melchert 1987 and 1989). While there is no reason to assume more than one synchronic phoneme, it is quite possible that /ts/ includes a palatal or palatalized allophone. Despite the hesitation of Melchert (1994a:274), there is good reason to suspect that graphic *z* also in some cases represents a voiced dental fricative /z/ (cf. the same possibility for Lycian *z*). The transliteration of the voiceless coronal sibilant as *š* in CLuvian is merely conventional, as in Hittite, and there is no reason to suppose that the sound is anything other than a dental-alveolar /s/. As in the case of Hittite and Palaic, the characterization of the sounds spelled -*h*h- and -*h*- in CLuvian as pharyngeals is by no means assured, and velar fricatives /x/ and /ɣ/ are quite viable alternatives.

### 3.1.1 Diachronic variation

Two diachronic developments affecting Luvian consonants are worthy of mention. The first is Čop's Law, by which a prehistoric sequence \**ǵ.C<sub>1</sub>* becomes Luvian *aC<sub>1</sub>.C<sub>1</sub>*: for example, \**nǵb<sup>(h)</sup>es-* > CLuvian *tappaš-* "heaven"; \**mǵlid-* > CLuvian *mallit-* "honey" (see Čop 1970 and Melchert 1994b:305 for further details). The second is *rhotacism*, a sporadic change by which *d*, *l*, and rarely *n* become *r* in HLuvian: for example, HLuvian /*iri*/ beside /*īdi*/ "goes," /*wara-*/ beside /*wala-*/ "die" (see Morpurgo Davies 1982–1983 for details).

## 3.2 Vowels

Luvian has only three vowels, /*a*/, /*i*/, and /*u*/, in contrasting short and long varieties. While there are some underlying long vowels, most phonetic length is due to synchronic rules which lengthen underlying short vowels under the accent: contrast sentence-initial conjunction *pā* < /*pá*/ versus enclitic *-pa* < /-*pá*/, or adverb *ānnan* "under" < /*ánnan*/ versus *annān pātanza* "under the feet" < /*annán*/ with accent shift in a prepositional phrase (see Melchert 1994a:247 for further discussion). There are clearly falling diphthongs /*a:y*/ and /*a:w*/. Corresponding short /*ay*/ and /*aw*/ are likely, but difficult to prove.

Certain facts about the placement of the accent may be inferred from the prehistoric and synchronic rules cited in the preceding two paragraphs, but the evidence is limited, and the risk of circularity of argument is high.

## 3.3 Synchronic variation

In addition to the vowel-lengthening rules referred to above, synchronic rules include the loss of word-final /-*d*/ in certain noun paradigms and the insertion of /-*s*/ between dental stops (*aztūwari* "you (pl.) eat" < /*ad-tuwari*/), the latter rule inherited from Proto-Indo-European.

## 3.4 Phonotaxis

Phonotactic restrictions apply chiefly to initial and final consonants. Only /*s*/, /*l*/, /*r*/, and /*n*/ appear word-finally, with /-(*n*)ts/ the only final cluster. All consonants appear regularly word-initially except /*r*/, for which HLuvian shows a single example. For the possibility that only voiceless obstruents appear word-initially see Melchert 1994a:18ff. The very limited evidence regarding consonant clusters is summarized by Melchert 1994a:248ff. Vowels occur freely in all positions. There are no assured cases of hiatus.

# 4. MORPHOLOGY

## 4.1 Nominal morphology

Luvian is a typical older Indo-European language with a well-developed, almost exclusively suffixing derivational and inflectional morphology. The noun inflects for two numbers, singular and plural. Some animate nouns have a collective beside a count plural: *dušduma* "(set of) vouchers" beside unattested \**dušduminzi* "vouchers" to *dušduma/i-*. Reference to more than one collective set requires a special "individualizing" suffix *-ant-*: for example, /*tawa*/ (collective plural) "eyes" (of one person), but /*tawanta*/ "sets of eyes." There are

two genders, animate and inanimate. The former is in most stem-classes marked not only by a distinct set of endings, but also by an obligatory /-i-/ inserted between stem and ending just in the nominative and accusative cases (see Starke 1990:59ff.). The origin of this latter feature is a matter of lively debate. CLuvian has five cases: nominative, vocative, accusative, dative-locative, and ablative-instrumental. The vocative is rare and restricted to the singular. The inanimate gender expectedly has a single nominative and accusative, and the ablative-instrumental does not distinguish number. HLuvian merges the animate nominative and accusative plural. CLuvian replaces the genitive entirely with a “relational adjective” modifying the head noun: “divine favor” for “favor of the god(s).” HLuvian uses both the modifying adjective and a true nominal genitive case, sometimes combining them in remarkable ways (see the examples cited in Melchert 1990:202ff.). Both dialects permit use of the relational adjective even when a noun phrase includes a second genitive dependent on another (“the son of the lord of the country”), in which case both genitives are expressed by adjectives agreeing in case and number with the head noun (see Neumann 1982).

A disadvantage of the relational adjective is that it cannot express the number of the possessor. CLuvian has partially remedied this situation by developing a marker *-anz-* which is inserted between the stem and nondirect case endings in the relational adjective to mark the possessor as plural: *waššaraḫitati maššanaššanazati* “by the favor of the gods” (see Melchert 2000).

In HLuvian the inanimate nominative-accusative singular of nouns is obligatorily marked by a postposed particle *-sa/-za*. This marker is also frequent in CLuvian, where some trace of an original anaphoric or deictic function seems visible (see Arbeitman 1992:22ff. for discussion).

The suffixes marking number, gender, and case are mostly recognizable as inherited from Proto-Indo-European: for example, animate nominative singular /-s/ and accusative singular /-n/ < \*-m. However, Luvian has innovated significantly in the plural, building a new system apparently based on the old animate accusative plural \*-ons: animate nominative plural /-Vntsi/, animate accusative plural /-Vnts/, dative-locative plural /-ants/.

## 4.2 Pronouns

The personal pronouns, as far as attested, are recognizable as inherited from Proto-Indo-European, with the peculiar Anatolian *u*-vocalism in the first person singular: HLuvian *amu* “I, me.” Luvian also shows the characteristic Anatolian demonstratives *apā-* “that” and *za/i-* “this” (the latter equaling Hittite *ka/i-*) and the inherited relative-interrogative *kui-*. Inflection appears to follow that of the noun more closely than in Hittite, but evidence for the nondirect cases is sparse.

## 4.3 Verbal morphology

The verb has the expected three persons and two numbers. There are only two moods, indicative and imperative. Evidence for a medio-passive beside the active voice is limited, but use of the middle of *a(ya)-* “make, do” effectively as a passive (“be made, become”) suggests that the functions of the medio-passive are of the expected sort. Tenses are limited to a present-future and a preterite. The basic verb-stem is unmarked for aspect, but there is also an imperfective marked with the suffixes *-sa-* and *-za-*. Luvian shows the same division into *mi-* and *ḫi-* conjugations as Hittite (see Ch. 18, §§4.4.7, 4.4.9; Morpurgo Davies 1979). The verbal endings are mostly clearly inherited, but there may have been a noteworthy



replacement of the preterite third-person active endings by the medial endings (see Yoshida 1993).

Luvian has a single participle formed with the suffix *-a(i)mma-*, with a past passive value for transitive verbs and a stative one for intransitives, and an infinitive in *-una*.

#### 4.4 Derivational morphology

Luvian shows a rich set of derivational suffixes in both the noun and verb. Even the massive study of Starke (1990) covers systematically only the consonant stems in the noun.

#### 4.5 Compounds

There are no assured compounds among appellatives, but examples in personal names suggest that further analysis and additional evidence will reveal some.

#### 4.6 Numerals

Knowledge of numerals in Luvian is limited by their frequent spelling with logograms. See Eichner 1992 for what is known.

### 5. SYNTAX

#### 5.1 Word order and clause structure

The functionally unmarked word order is SOV (Subject–Object–Verb), but any major constituent may be fronted to initial position for emphasis, and elements may also be extraposed to the right of the verb for the same purpose. Relative clauses typically precede the main clause with a resumptive pronoun, but postposed relatives also occur. Luvian has several subordinating conjunctions marking temporal or conditional clauses. There are no coordinated clauses in the strict sense, but the enclitic conjunction *-ha* which conjoins noun phrases can be used to mean “also.” Like Hittite (see Ch. 18, §5.1), Luvian links sentences in narration with prosecutive conjunctions, *a-* (functionally = Hittite *nu-*) or *pā/-pa* (= Hittite *-ma*, marking change of topic, lightly adversative). Adjectives, including demonstratives and relational adjectives, typically precede their head noun. Luvian appears to have both prepositions and postpositions, as well as local adverbs which occur independently and as preverbs.

As in the other Indo-European Anatolian languages, anaphoric pronouns, conjunctions, and various particles regularly appear in Luvian as enclitics, attached to the first accented element in a clause by *Wackernagel's Law*. The conjunction *-ha* “and” which conjoins noun phrases is also an enclitic (like Latin *-que*), and note the particle *-sa/-za* cited above (§4.1).

#### 5.2 Syntactic miscellanea

HLuvian shows at least one example of the Anatolian construction in which the direct object of an infinitive is unexpectedly in the dative: *za-ti CASTRUM-si AEDIFICARE+MI-na* “this (dat.) fort (dat.) to build” = “to build this fort.” This usage is comparable to the “double dative” of Sanskrit. Examples with the expected accusative also occur in HLuvian.

The syntax of cardinal numerals is complex. They may occur as adjectives agreeing with nouns in number and case, but one also finds singular nouns with numbers above one (see Eichner 1992, *passim*).

## 6. LEXICON

Luvian core vocabulary appears to be for the most part inherited from Proto-Indo-European, but evidence is limited: of the fifty-one words from the Swadesh-Voegelin hundred-word list which are known, thirty-nine or 80 percent are of Proto-Indo-European origin. The only major source of loanwords is Hurrian, from which many terms in various technical fields such as divination passed into Luvian and then into Hittite.

## 7. READING LIST

Marazzi 1990 offers a thorough bibliography for HLuvian along with a grammatical sketch which is mostly valid also for CLuvian, and a partial lexicon. Werner 1991 is also useful and reliable. HLuvian text editions are currently scattered through secondary works. A complete new edition of the HLuvian texts of the first millennium is now available in Hawkins 2000. The older standard works by Meriggi and Laroche (cited in Marazzi) are now rendered almost useless by the outdated phonetic values of several crucial signs. Starke 1985 gives the available CLuvian texts in transliteration. Melchert 1993 offers a complete lexicon for Starke 1985 plus selected Luvianisms in Hittite contexts. The lexicon and grammatical sketch of Laroche 1959 are still useful, but must be read in conjunction with the works cited above.

### Bibliography

- Arbeitman, Y. 1992. "Luwian *za-* and *-sa* (/ *-za*): how I have changed my mind (with ruminations on Palaic)." *Linguistique balkanique* 35:3–58.
- Čop, B. 1970. "Eine luwische orthographisch-phonetische Regel." *Indogermanische Forschungen* 75:85–96.
- Eichner, H. 1992. "Anatolian." In J. Gvozdanović (ed.), *Indo-European Numerals*, pp. 29–96. Berlin/New York: Mouton de Gruyter.
- Hawkins, D. 1975. "The negatives in Hieroglyphic Luwian." *Anatolian Studies* 25:119–156.
- . 1986. "Writing in Anatolia: imported and indigenous systems." *World Archaeology* 17:364–376.
- . 2000. *Corpus of Hieroglyphic Luwian Inscriptions. Volume I. Inscriptions of the Iron Age*. Berlin/New York: de Gruyter.
- Hawkins, D., A. Morpurgo Davies, and G. Neumann. 1974. *Hittite Hieroglyphs and Luwian: New Evidence for the Connection*. Göttingen: Vandenhoeck and Ruprecht.
- Laroche, E. 1959. *Dictionnaire de la langue louvite*. Paris: Adrien-Maisonneuve.
- Marazzi, M. 1990. *Il geroglifico anatolico. Problemi di analisi e prospettive di ricerca*. Rome: Università "Sapienza."
- Melchert, H. C. 1987. "PIE velars in Luvian." In C. Watkins (ed.), *Studies in Memory of Warren Cowgill*, Berlin/New York: de Gruyter.
- . 1989. "New Luvo-Lycian isoglosses." *Historische Sprachforschung* 102:23–45.
- . 1990. "Adjective stems in *\*-iyo-* in Anatolian." *Historische Sprachforschung* 103:198–207.
- . 1992. Review of *Untersuchung zur Stammbildung des keilschrift-luwischen Nomens*, by Frank Starke. *Historische Sprachforschung* 105:309–312.
- . 1993. *Cuneiform Luvian Lexicon*. Chapel Hill: self-published.

- \_\_\_\_\_. 1994a. *Anatolian Historical Phonology*. Leiden Studies in Indo-European 3. Amsterdam: Rodopi.
- \_\_\_\_\_. 1994b. "Çop's Law' in Common Anatolian." In J. Rasmussen (ed.), *In honorem Holger Pedersen*, pp. 297–306. Wiesbaden: Dr. Ludwig Reichert.
- \_\_\_\_\_. 1996. "Anatolian Hieroglyphs." In P. Daniels and W. Bright (eds.), *The World's Writing Systems*, pp. 120–124. New York/Oxford: Oxford University Press.
- \_\_\_\_\_. 2000. "Aspects of Cuneiform Luvian nominal inflection." In Y. Arbeitman (ed.), *The Asia Minor Connexion: Studies on the Pre-Greek Languages in Memory of Charles Carter*, pp. 173–183. Louvain-la-Neuve: Peeters.
- Morpurgo Davies, A. 1979. "The Luwian languages and the Hittite *-hi* conjugation." In B. Brogyanyi (ed.), *Studies in Diachronic, Synchronic and Typological Linguistics. Festschrift for Oswald Szemerényi on the Occasion of his 65th Birthday*, pp. 577–610. Amsterdam: John Benjamins.
- \_\_\_\_\_. 1982–1983. "Dentals, rhotacism, and verbal endings in the Luwian languages." *Zeitschrift für Vergleichende Sprachforschung* 96:245–270.
- Neumann, G. 1982. "Die Konstruktionen mit Adiectiva genitivalia in den luwischen Sprachen." In E. Neu (ed.), *Investigationes philologicae et comparativae: Gedenkschrift für Heinz Kronasser*, pp. 149–161. Wiesbaden: Otto Harrassowitz.
- Starke, F. 1985. *Die keilschrift-luwischen Texte in Umschrift*. Studien zu den Boğazköy-Texten 30. Wiesbaden: Otto Harrassowitz.
- \_\_\_\_\_. 1990. *Untersuchung zur Stammbildung des keilschrift-luwischen Nomens*. Studien zu den Boğazköy-Texten 31. Wiesbaden: Otto Harrassowitz.
- Werner, R. 1991. *Kleine Einführung in das Hieroglyphen-Luwische*. Göttingen: Vandenhoeck and Ruprecht.
- Yoshida, K. 1993. "Notes on the prehistory of preterite verbal endings in Anatolian." *Historische Sprachforschung* 106:26–35.

# Palaic

H. CRAIG MELCHERT

## 1. HISTORICAL AND CULTURAL CONTEXTS

Palaic was once the spoken language of the land of Palā, generally agreed to have been located to the northwest of Hittite territory across the Halys River (modern Kızıl Irmak) in what is now north central Turkey. The country name is surely reflected in the later classical Blaëne and Paphlagonia. Palaic is attested in scarcely a dozen ritual fragments from the cuneiform archives of the ancient Hittite capital Hattuša (modern Boğazköy). The documents are contemporary with the Hittite (sixteenth to thirteenth centuries BC), including a couple of manuscripts from the Old Hittite period.

What little Palaic we have owes its preservation to liturgical use by the Hittites, chiefly for the cult of the Hattic god Za/ıparfa. Palā, mentioned in the Old Hittite Laws as one of the three divisions of the Hittite state along with Hatti and Luwiya, appears only rarely in later texts. Its decline in importance is sometimes attributed to the depredations of the Kaskians, a people of the northern mountains who caused serious problems for the Hittites throughout their history. It is likely that Palaic was extinct as a spoken language by the thirteenth century, and it may well have been so by the time of our earliest texts in the sixteenth. The extremely sparse documentation makes all aspects of the following description provisional. Palaic is interpreted largely in light of the much better attested Hittite so far as the facts permit. This is a reasonable and necessary procedure, but its obvious risks should constantly be borne in mind.

## 2. WRITING SYSTEM

Palaic was written by Hittite scribes, using the same version of the Old Babylonian cuneiform syllabary they employed for writing Hittite. The spelling conventions are the same as for Hittite (see Ch. 18, §2), with very few exceptions, the most notable being the use of special signs for a phoneme /f/ absent in Hittite (see below) and the near-total absence of logograms. The syllabary has V, CV, VC, and some CVC signs. It thus can indicate initial and final consonant clusters (and internal clusters of more than two) only by the use of “empty” vowels. Such sequences are interpreted largely on comparative and etymological grounds. There is no longer any reason to doubt that the use of *scriptio plena* (repeating the vowel of a CV or VC sign with the matching V sign) marks synchronic vowel length (see for Hittite, Kimball 1983 *passim*, et al.). The system uses both word spacing and paragraph-dividers.

Emil Forrer in 1919 already recognized Palaic as one of the eight distinct languages of the Boğazköy archives, and after a brief false start tentatively identified it as an Indo-European language closely related to Hittite. It was not until 1944, however, that Heinrich Otten was

able firmly to establish the status of Palaic on the basis of further documentation. Important further contributions to understanding the language were made by Kammenhuber (1959) and Carruba (1970). There have been no new textual finds since Carruba's work.

### 3. PHONOLOGY

#### 3.1 Consonants

The phonemic inventory includes at least the following consonants:

##### (1) Palaic consonantal phonemes

p	t	k
b	d	g
	ts	
f	s	ʒ
		ħ
		ʕ
m	n	
	l	r
w		y

The absence of positive evidence, as in Hittite, for unitary labiovelars may easily be accidental, and there is a good chance that we should also assume a voiceless labiovelar stop /kʷ/ in words such as *kui-* “who, which.” Owing to a prehistoric change, there likely is no corresponding voiced stop, but there may be a labialized /ɣʷ/ in cases like *aḫu-* “drink.”

##### 3.1.1 Stops

Characterization of the Palaic stop series as voiceless versus voiced is based on their etymological value. The synchronic phonetic status of the stops in the three cuneiform languages Hittite, Palaic, and Cuneiform Luvian is a vexing and controversial problem: see Melchert 1994:13–21 for an extensive discussion with references. What is clear is that etymological voiceless stops appear as graphic and probably linguistically real geminates in intervocalic position, while inherited voiced stops appear as single stops (so-called *Sturtevant's Law*): in Palaic contrast particle *-ppa* < \**-pe* (cf. Latin *nem-pe* “surely”) with *apā-* “that” < \**ob*<sup>(h)</sup>*ó-*. It is tolerably certain that voiced stops have been generalized in word-final position (*šarkut=at* “\_\_ed them,” with preterite third singular -t [d] < \**-t*), while it is likely but not assured that voiceless stops and fricatives have been generalized word-initially (see Melchert 1994:18–20, *et al.*). This partial neutralization of the voicing distinction may have contributed to a re-analysis of the stop contrast as one of fortis versus lenis, but this analysis cannot be regarded as proven.

##### 3.1.2 Fricatives

The phoneme /f/ appears in Hattic loanwords into Palaic such as *wu<sub>ii</sub>/pu-la-a-ši-na-* (a kind of bread). As the cited example shows, the fricative /f/ is indicated by a special series of cuneiform signs, consisting of *wa* plus a *mater lectionis* marking vowel quality, transliterated *wV<sub>V</sub>* (sometimes alternately with graphic *p*). It cannot be excluded that in some or all cases the fricative is a voiced /v/ rather than /f/.

The voiceless coronal fricative is spelled with the cuneiform series which indicates a palatal sibilant in Akkadian, whence the conventional transliteration as *š*, as in Hittite.

There is no reason, however, to think that the sound represented is anything other than a dental/alveolar /s/. The sound transliterated as *z* is in most cases a voiceless affricate /ts/, but it cannot be ruled out that in some instances it indicates a voiced sibilant /z/ instead.

The phoneme rendered above as /ʒ/ represents a weak palatal fricative, the result of a prehistoric sequence of *\*h<sub>2</sub>y* (the Proto-Indo-European “second laryngeal” plus *\*y*), spelled alternately with -g- and zero (see Watkins 1975:373 for the derivation and Carruba 1970:39 for the spelling). The phonetic definition of this sound obviously is merely an approximation, and one may entertain other possibilities.

Palaic shows both the regular and “lenited” reflexes of the Proto-Indo-European “second laryngeal” *\*h<sub>2</sub>*, spelled intervocally with geminate -*h<sub>2</sub>h<sub>2</sub>*- and simple -*h<sub>2</sub>*- respectively (for Proto-Anatolian “lenition” see Eichner 1973:79ff.). The characterization above as pharyngeals is nonbinding, and a pair of velar fricatives /x/ and /ɣ/ is equally possible.

### 3.2 Vowels

There are at least three vowel phonemes /a/, /i/, and /u/ and contrasting long /a:/, /i:/, and /u:/. It is very likely that there are also /e/ and /e:/, but the phonemic status of the latter is controversial (see Melchert 1994:198f., but also Carruba 1970:9, and Wallace 1983). While there are a few nonpredictable and thus contrastive long vowels, most surface vowel length is allophonic, due to synchronic rules of lengthening under the accent: for examples see the next paragraph. The vowel /a(:)/ combines with /y/ to form a falling diphthong /a(:)i/. The absence in our data of a corresponding /a(:)u/ is probably accidental.

### 3.3 Synchronic variation

There is limited but solid evidence for a synchronic rule in Palaic whereby the word accent shifts one syllable to the right with the addition of an enclitic (cf. the similar rule in Latin): underlying /*hāran*-/ “eagle” appears as regular [ḥá:ranas] in the genitive singular *hāranaš*, but compare nominative singular [ḥará:s] in the phrase *hārāš=kuwar*. The fact that the length of the *a* in both syllables of “eagle” depends on the accent argues that the vowel in each case is underlyingly short, with the long [a:] a conditioned allophone. There are actually two such synchronic lengthening rules, one applying to all accented vowels in open syllables, the other to /a/ and /e/ in accented closed syllables (see Melchert 1994:204f. for further discussion).

Word-final -*n* is sporadically assimilated to an initial labial of a following clitic: =*am=pi* beside =*an=pa=ti*. The sibilant /s/ appears occasionally as *z* next to a sonorant (= *kuwar=zi* for =*kuwar=ši*) and rarely other consonants. This may or may not represent voicing to [z].

### 3.4 Phonotaxis

Phonotactic restrictions are unremarkable, so far as the extremely limited evidence permits a judgment. Final consonants are highly restricted: voiced stops (only /d/ is actually attested), /s/, /ʃ/, /n/, /r/, and /l/. The only attested final cluster is /-(n)ts/. As indicated above, probably only voiceless obstruents are permitted word-initially, along with /m/, /n/, /l/, and /w/. The absence of examples of initial /y/ is surely accidental, but the lack of initial /r/ is systematic, as elsewhere in the ancient Anatolian languages. Initial clusters are mostly limited to biconsonantal sequences of rising sonority, but there may be some cases of fricative plus stop. Medial clusters are predictably more varied and complex: for an exhaustive list of examples, see Melchert 1994:206f. All vowels occur freely in initial, medial, and final position. Sequences with hiatus are rare, but *iu* “come!” (cf. Hittite *ehu*) certainly represents [i\_u] historically and probably also synchronically.

## 4. MORPHOLOGY

Palaic is a typical ancient Indo-European language in its morphological typology: that is, fusional, with a well-developed system of derivation and inflection, the latter exclusively suffixing, the former nearly so. However, it is also characteristically Anatolian in having a relatively limited set of inflectional categories in the nominal and verbal systems in comparison to Sanskrit or Ancient Greek.

### 4.1 Nominal morphology

The noun distinguishes two numbers, singular and plural, and two genders, animate and inanimate. There is no evidence for a separate dual or a feminine gender. There are at least six cases in the singular: nominative, vocative, accusative, genitive, dative, and locative. The last two cases are not distinguished in the plural, and as expected the nominative plural serves also for the vocative. It seems extremely likely that Palaic also has an ablative-instrumental corresponding to that of the other Anatolian languages, but no examples have yet been found. As usual in Indo-European, the nominative and accusative are not distinguished in the inanimate gender. In addition to the genitive case, Palaic also expresses possession by means of a relational adjective (“paternal house” for “father’s house”), but this usage does not seem to be as widely developed as in the western Anatolian languages. It is impossible to tell whether there is any functional distinction between the two constructions.

The case endings are mostly recognizable as Indo-European: animate nominative singular /-s/, vocative singular zero, animate accusative singular /-n/, inanimate nominative-accusative singular zero or /-an/, genitive singular /-as/, dative singular /-i/ or /-ai/, animate nominative plural /-es/ (or /-as/ < \*-ōs), inanimate nominative-accusative plural /-a/. The animate accusative plural is not securely attested. The locative ending /-a/ is cognate with the allative of Hittite continuing Proto-Indo-European \*-h<sub>2</sub>e and \*-oh<sub>2</sub> (cf. for the latter Latin *quō* “whither”). The dative-locative plural /-as/ matches the endings of Hittite and Lycian, reflecting Proto-Indo-European \*-os (cf. Latin *-bus*, etc. minus the initial labial).

### 4.2 Pronouns

The only reasonably well-attested pronominal stem is the relative-interrogative *kui-*, but the existence of the characteristic Anatolian demonstratives *kā-* “this” and *apā-* “that” is at least assured. For the few other extant pronominal forms see Carruba 1970:44.

### 4.3 Verbal morphology

The verb is inflected for singular and plural and the expected three persons. There are two moods, indicative and imperative, and two tenses, present (also used for the future) and preterite. Beside the active voice there is a medio-passive, surely with the usual range of functions, although the few attested examples happen to be *media tantum* with intransitive meaning (“lie” and “be warm”). The basic verbal stem may express various aspectual nuances according to context, but imperfective aspect may also be overtly marked by suffixes cognate with those which serve the same function in Hittite: *pī-ša* “give!” (distributive, with multiple objects) or *i-škā* “be!” (durative, in a construction indicating possession). The verbal endings formally are cognate with those of the other Anatolian languages, but the limited evidence suggests that the distribution in Palaic does not quite match that of Hittite or Luvian.



There is an infinitive in *-una* (e.g., *aḥuna* “to drink”) cognate with that in Luvian. Palaic interestingly appears to employ both *-ant-* and *-amma-* as suffixes to form participles (*takkuwānteš* and *patamman*), but the meaning of the latter example is quite uncertain, and it may be a lexicalized relic. It is likely that there is a single functional category expressing an attained state (passive for transitive verbs), as in Hittite (see Ch. 18, §4.4.5).

#### 4.4 Morphological miscellanea

Palaic has a range of nominal stem-classes (at least *-a-*, *-i-*, *-u-*, *-(n)t-*, *-n-*, and *-r-*) and probably two verbal conjugations corresponding to the *mi-* and *hi-*conjugations of Hittite (see Ch. 18, §§4.4.7; 4.4.9), although evidence for the latter is arguable. It is not clear to what extent Palaic shows the phenomenon of “i-mutation” so characteristic of the western Anatolian languages (see Starke 1990:71ff.). Several well-known Indo-European nominal and verbal derivational suffixes are attested, and further data would undoubtedly yield further examples.

#### 4.5 Compounds

There is one assured compound: *aš=kummawa-*, literally “mouth-pure,” i.e., “sacralized and fit for the gods to eat” (see Watkins 1987:399f., after Szemerényi). The absence of additional examples is undoubtedly due to chance.

### 5. SYNTAX

#### 5.1 Word order and clause structure

As an inflected language, Palaic predictably has rather free word order. The unmarked order is SOV (Subject–Object–Verb), but others are by no means rare. Essentially any major constituent may be placed in clause-initial position for emphasis: verb, direct or indirect object, adverb, and preverb are all attested besides subject.

There are no assured cases of coordinated clauses. Probably not by accident there are also no clear examples of coordinated noun phrases, while asyndeton is common. Palaic has the enclitic conjunctions *-ku* and *-ḥa*, but the latter certainly means “also,” and probably so does the former. Neither is a simple connective “and.” The only certain subordinating conjunction is *mān* “when(ever), if,” cognate with Hittite *mān*. Relative clauses preceding the main clause with a resumptive pronoun are assured (*kuiš=a . . . =apan* “whichever . . . that one”), and there likely is at least one example of a postposed relative clause.

#### 5.2 Agreement

Gender and number agreement is mostly of a standard sort, but Palaic does preserve the Proto-Indo-European construction whereby a neuter plural (actually an old collective) as subject takes a singular verb: *tilila ḥāri* “the t. (a food) are warm” (lit. “is warm”). As in other Anatolian languages, one also finds in Palaic a singular verb apparently agreeing with the first of multiple subjects: *lukīt=ku tabarnaš tawannannaš* “The king (and) queen have also distributed” (*lukīt* is preterite third singular).

### 5.3 Clitics

The Indo-European Anatolian languages are famous for their use of clitics, and Palaic is no exception. There is ample evidence for the appearance of subject and object pronouns and various sentence particles (often ill-defined) as enclitics to the first accented word in the clause (so-called *Wackernagel's Law*): *arūn=am=pi witeši* “tall=him=particle you shall build,” that is, “You shall build him tall.” In addition, however, Palaic shows sporadic cliticism of words which are normally accented: *nū=wašu* (sentence conjunction plus “good”) versus normal *wāšu* (note the loss of length on the first vowel of the noun, as per the rule mentioned above). The conditioning and function of this usage are unclear. There is also evidence for enclitic use of the demonstrative (see Melchert 1984:28ff.). The apparent restriction of this usage to the neuter singular seems strange, but is probably paralleled in Luvian.

## 6. LEXICON

The severely restricted corpus precludes definitive statements about the lexicon: only twenty-two of the words in the Swadesh-Voegelin hundred-word core vocabulary list are attested and identified. One cannot place too much weight on the fact that 87 percent of these are inherited. The facts of Hittite suggest, however, that the nature of our evidence presents a misleading picture. The apparent heavy influence of Hattic is probably due simply to the fact that our texts nearly all deal with the cult of the Hattic pantheon. Palaic has also borrowed at least the title for the Hittite king, *tabarna-*, from Luvian (not from Hattic), and this is likely for the queen's title, *tawananna-*, as well.

## 7. READING LIST

Carruba (1970) provides a convenient and excellent vade mecum: all texts in transliteration (but without translation), grammar, and lexicon, plus bibliography to that date.

### Bibliography

- Carruba, O. 1970. *Das Palaische. Texte, Grammatik, Lexikon*. Studien zu den Boğazköy-Texten 10. Wiesbaden: Otto Harrassowitz.
- Eichner, H. 1973. “Die Etymologie von heth. *mēhur*.” *Münchener Studien zur Sprachwissenschaft* 31:53–107.
- Kammenhuber, A. 1959. “Das Palaische: Texte und Wortschatz.” *Revue hittite et asianique* 14: 1–21.
- Kimball, S. 1983. “Hittite Plene Writing.” Ph.D. dissertation, University of Pennsylvania.
- Melchert, H. C. 1984. “Notes on Palaic.” *Zeitschrift für Vergleichende Sprachforschung* 97: 22–43.
- . 1994. *Anatolian Historical Phonology*. Amsterdam: Rodopi.
- Starke, F. 1990. *Untersuchung zur Stammbildung des keilschrift-luwischen Nomens*. Studien zu den Boğazköy-Texten, 31. Wiesbaden: Otto Harrassowitz.
- Wallace, R. 1983. “The development of PIE \*ē in Palaic.” *Die Sprache* 29:159–173.
- Watkins, C. 1975. “Die Vertretung der Laryngale in gewissen morphologischen Kategorien in den indogermanischen Sprachen Anatoliens.” In H. Rix (ed.), *Flexion und Wortbildung*, pp. 358–378. Wiesbaden: Dr. Ludwig Reichert.
- . 1987. “Two Anatolian forms: Palaic *aškumāuwa*, Cuneiform Luvian *wāša*.” In G. Cardona and N. H. Zide (eds.), *Festschrift for Henry Hoenigswald*, pp. 399–404. Tübingen: Narr.

# Lycian

H. CRAIG MELCHERT

## 1. HISTORICAL AND CULTURAL CONTEXTS

Lycian was the autochthonous language of the land of Lycia at least during the middle and late first millennium BC. Recent evidence from the Hieroglyphic Luvian inscription of Yalburt – specifically, forms of the place names for Tlos, Pinara, and Xanthos – has now proven that the “Lukka-Lands” of the second-millennium Hittite cuneiform texts do refer to historical Lycia, that is, roughly the mountainous peninsula on the southwest coast of Anatolia lying between the Gulf of Telmessos and the Bay of Attaleia (modern Gulf of Fethiye and Gulf of Antalya; see Poetto 1993). Obviously, without direct textual evidence from Lycia itself during the second millennium it is quite impossible to characterize with any precision the language of “Lukka” in that era.

Lycian shares a number of specific features, including innovations, with Luvian, and it is widely held that Lycian and Luvian form a subgroup within the Anatolian family; in other words, that they reflect a prehistoric “Proto-Luvian” language which had developed out of Proto-Anatolian along different lines from Hittite, Palaic, and Lydian, the other assured members of the Anatolian group (see, *inter alios*, Oettinger 1978). One may even read that Lycian is a later form of Luvian, though not necessarily of that form of Luvian which is directly attested in the second millennium. The shared features of Lycian and Luvian are undeniable, but several of these are also common to Lydian, while there are also crucial divergences between Lycian and Luvian (see Gusmani 1960 and Melchert 1992a). These divergences make it impossible to reconstruct a coherent Proto-Luvian language distinct from Proto-Anatolian. One should rather view the common features of Luvian and Lycian in terms of dialect geography. As the individual languages began to diverge in their development from Proto-Anatolian, they remained in contact, and innovations which arose in various places spread in the typical irregular fashion. Luvian, which occupied a geographically central position, unsurprisingly shares some isoglosses with Lycian (and to a lesser extent Lydian) to the west, and others with Hittite and Palaic to the east.

The extant Lycian corpus includes more than 150 inscriptions on stone, over 200 on coins (many not yet published), and a handful on other objects. The overwhelming majority of those on stone are sepulchral texts, with highly stereotyped content. Apart from several poorly preserved decrees, the most important exceptions are the inscribed stele of Xanthos, which describes the military exploits and building activities of a local dynasty, and the Lycian–Greek–Aramaic trilingual of the *Létôon*, which records the founding of a cult for the goddess Leto by the citizens of Xanthos at a temple a few miles south of the city. The latter text of some forty-one lines has predictably proven to be of immense importance in

advancing understanding of Lycian. Much of the text of the Xanthos stele remains opaque due to problems of vocabulary which result from the nearly unique subject matter.

Two of the Lycian texts (one of which is the last portion of the Xanthos stele) are written in a distinct dialect known either as *Lycian B* (vs. ordinary *Lycian A*) or as *Milyan*. The relationship of the two dialects is indeterminate. Milyan is more archaic than ordinary Lycian in certain features, and it is noteworthy that both Milyan texts are in verse (see Eichner 1993 with references). However, it would be dangerous to conclude from these limited facts that Milyan is merely an older stage of Lycian preserved for special literary purposes. This is only one of several viable possibilities: see Gusmani (1989–1990) for a useful discussion of the problem. Unless stated otherwise, the description which follows applies to both forms of Lycian, but the bulk of the evidence comes from Lycian (A). Extrapolation of the description to Milyan is often based on very limited evidence and should be viewed as highly provisional. Special features of Milyan will be explicitly noted where appropriate.

Thanks to the Létôon Trilingual and exploitation of the features shared with Luvian, understanding of Lycian has improved dramatically in the last two decades (with the notable exception of the Xanthos stele and Milyan). However, certain features of morphology and syntax cited below impose some quite serious limitations. One should regard the following description as intermediate in completeness and reliability between those for Palaic and Lydian on the one hand, and that for Luvian on the other.

## 2. WRITING SYSTEM

Lycian is written in an alphabet derived from or closely related to that of Greek. The details of the relationship remain unclear: for discussion see Carruba 1978a. The direction of writing is left to right. Use of word-dividers is frequent, but by no means absolutely consistent. This fact means that the status of certain morphemes as clitics is, strictly speaking, a matter of interpretation, which can be supported but not proven by the mode of writing. Problems involving individual letters will be dealt with below in the phonology.

## 3. PHONOLOGY

### 3.1 Consonants

The Lycian segmental inventory includes the following consonantal phonemes:

#### (1) Lycian consonantal phonemes

p	t	c	k<	k	k>	k <sup>w</sup>
	ts					
	θ					h
β	ð			γ		
m	n					
	l	r				
w		y				

Of the phonemes listed, /c/, /θ/, and /h/ occur only in Lycian (A), not in Milyan, due to different prehistoric sound changes. The sound very tentatively identified as /k<sup>w</sup>/ is attested only in Milyan and in personal names. Its absence in Lycian (A) may or may not be due to chance.

**Table 21.1 The Lycian alphabet**

Character	Transcription
𐤀	a
𐤁	e
𐤂 𐤃	b (/β/)
𐤄	β (/k <sup>w</sup> /?)
𐤅 𐤆	g (/γ/)
𐤇	d (/ð/)
𐤈	i
𐤉	w
𐤊	z (/t <sup>s</sup> /)
𐤋	θ
𐤌	y
𐤍	k (/k</)
𐤎	q (/k/)
𐤏	l
𐤐	m
𐤑	n
𐤒	m̃ (/m <sup>̃</sup> /)
𐤓	ñ (/n <sup>̃</sup> /)
𐤔	u
𐤕	p
𐤖	(≈ /k</?)
𐤗	r
𐤘	s
𐤙	t
𐤚	τ (/c/)
𐤛 𐤜 𐤝 𐤞 𐤟	ā
𐤠 𐤡 𐤢 𐤣 𐤤	ē
𐤥	h
𐤦 𐤧 𐤨 𐤩	x (/k>/)

### 3.1.1 Stops

The stop phonemes given here as /p/, /t/, /k</, /k/, and /k>/ are spelled respectively *p*, *t*, *k*, *q*, and *x* according to the current standard transliteration (but one must be prepared to find *c* for *k* and *k* for *x* respectively in older works). There is a consensus that these stop phonemes have voiceless and voiced allophones. The conditioning is also straightforward: the voiced allophones occur after nasals (including nasalized vowels), the voiceless allophones elsewhere. Note, for example, *trqqñt-* (name of the Storm-god) for [tərkənd-], rendered in Greek as Τροκονδος/Τερκανδας.

There is on the contrary a decided absence of agreement concerning the further features of the stops aside from labial /p/ and dental/alveolar /t/. The rare sound defined here as /c/ (transliterated as τ) alternates with /t/ in all cases. We know that prehistoric *\*k<sup>w</sup>* becomes Lycian (A) *t* before *i* (e.g., *ti-* < *\*k<sup>w</sup>i-* “who, which”), and several plausible, but not entirely

compelling etymologies have been adduced for  $t/\tau < *k^w$  before  $*e$  (see Carruba 1978b: 165ff.). If we accept this derivation, a palatal stop  $/c/$  seems a plausible transition sound, since the development includes fronting and delabialization (the value  $/t^w/$  suggested in Melchert 1994a:282 was an unfortunate lapsus). Note that in Milyan the result of a labiovelar before front vowel is  $k$  (*ki-* “who, which”), which will be argued below to be a front velar  $/k</math>. The development in Lycian (A) may be viewed as a further fronting to a palatal and eventually dental stop.$

The characterization of the dorsal stops  $k$ ,  $q$ , and  $x$  as front, mid, and back velar  $/k</math>,  $/k/$ , and  $/k>/$  respectively represents a personal point of view, and one should compare the in part very different opinions of Rasmussen (1974:53ff.), Laroche (1979:84), van den Hout (1995), and Hajnal (1995:26ff.). Evidence for a relatively front value for  $k$  (formerly transliterated  $c$ ) consists of its strong tendency to occur before (often between) front vowels and its rendering in Greek alternatively by *sigma* (*Tikeukepre-* =  $\tau\iota\sigma\epsilon\upsilon\sigma\epsilon\mu\beta\rho\alpha\nu$ ) and by *kappa* or *gamma* (*Sbikasa* =  $\Sigma\pi\iota\gamma\alpha\sigma\alpha$ ). The predilection of  $x$  (formerly  $k$ ) for appearing before back vowels suggests a relatively back consonant. The major point of dispute is whether it is an ordinary stop or instead an aspirated stop or even fricative. The only basis for the last assumption (hence the now standard transliteration  $x$ ) is etymological: Lycian  $x$  in most cases corresponds to a cuneiform  $h$ , both in names (*Xākbi* = *Ḫinduwa*) and in inherited words reflecting the Proto-Indo-European second laryngeal (preterite first singular ending  $-xa < *-h_2e$ ). There is, however, not a shred of evidence for anything but a plain stop synchronically: Greek rendering of Lycian  $x$  in names is consistently either with *kappa* or *qoppa*, never *chi* (the single exception  $\mu\omicron\sigma\chi\tilde{\alpha}\varsigma$  for *Musxxah* [cited by van den Hout 1995:134, correcting Melchert 1993:105] says nothing, since the aspirate may be a Greek phenomenon conditioned by the preceding  $s$ ).$

The question of whether  $q$  is an ordinary velar stop  $/k/$  as given here or is labialized depends on etymological considerations which cannot be treated here: see Melchert (1994a:306) for a discussion with references to other opinions. Even more problematic is the status of the sound represented by the rare letter  $\mathbb{W}$ . The Létôon Trilingual assures that it is some kind of dorsal stop (personal name *ArKKazuma* = Greek  $\text{Αρκεσιμ}\alpha$ ), but the tentative analysis as a labiovelar  $/k^w/$  is based on etymological and distributional arguments which are merely suggestive, not compelling (see Hajnal 1995:25f. and Eichner 1993:145, among others).

### 3.1.2 Affricate and fricatives

Lycian  $z$  in at least some cases represents a voiceless affricate  $/ts/$  (e.g., *hr-zze/i-* “upper” with suffix  $-zze-$  < Proto-Anatolian  $*-tsy-$  < PIE  $*-tyo-$ ). In other cases, however, a plausible case has been made for a voiced fricative  $/z/$ : see Melchert 1994a:314f. (with reference to Gusmani) and Hajnal 1995:21ff.

Lycian (A)  $\theta$  is clearly the reflex of prehistoric  $*d+h$ . Since  $*d$  is spirantized to voiced  $[\delta]$ , it seems reasonable to assume that the outcome of the sequence is a voiceless interdental fricative, and the Lycian version of a Persian name *Miθrapata-* appears to confirm this. Lycian (A)  $h$  is ignored in Greek renderings of personal names, suggesting that it is probably ordinary  $/h/$  (generally absent from Anatolian Greek). It reflects a conditioned change of  $*s > h$  in Lycian (A) which did not take place in Milyan.

There is near-universal agreement that the Lycian letters  $b$ ,  $d$ , and  $g$  stand for voiced fricatives. Evidence cited includes  $\Lambda\pi\alpha\rho\alpha\varsigma$  for *Dapara* and the Lycian rendering of Darius as *Ṃtarijeus-* (recall that voiceless stops are voiced after nasals). One may compare for the latter device Modern Greek. Neither of the cited spellings makes sense if Lycian  $d$  were a voiced stop  $[d]$ .

### 3.1.3 Sonorants

Cases such as *hrppi* “above” or *sñta* (a numeral) seem to indicate that liquids and nasals had syllabic allophones, and the standard view is that the special letters *ñ* and *ṁ* stand for syllabic nasals. This may have been true when the graphemes were invented, but this analysis cannot account for postvocalic occurrences such as *qāñti* “they slay.” The gemination in *hrppi* (see below) argues that at the phonetic level the pronunciation was [hərp.pi] with an anaptyctic vowel. If one makes the reasonable inference that the same is true for nasals (*sñta* = [sənta]), then one may make the generalization that *ñ* and *ṁ* occur only in syllable-final position. This distribution suggests that they are unreleased allophones of the nasal consonants.

The glides /w/ and /y/ are usually spelled with the letters transliterated *w* and *j*, but when they represent the second part of falling diphthongs they are spelled with the corresponding vowels: *ai*, *ei*, *au*, and so forth. Examples such as *ebeija* “these” (neut. nom.-acc. pl.) must apparently be interpreted as [eβej.ja]. Prehistoric \**w* appears as *b* after a consonant, suggesting that it has become a fricative in this position (e.g., *esbe*- “horse” < \**ekwo*-). Since this *b* never geminates after a consonant like ordinary /β/ (e.g., *erbbe*- “battle” or “defeat”), it should probably be treated as an allophone of /w/ synchronically.

### 3.1.4 Consonant gemination

One of the most striking and problematic features of Lycian consonantism is the widespread gemination of consonants (at least orthographically). No entirely satisfactory explanation has yet been presented: see for attempts Melchert 1994a:295f. and 316, and van den Hout 1995. Word-initial and some internal geminates probably reflect prehistoric processes (notably syncope) and must be synchronically analyzed as present in underlying structure: for example, *ttaraha*, adjective to *tetere/i*- “city”(?) (see Heubeck 1985 and Hajnal 1995:184ff.). However, the highly regular gemination of the second members of certain consonant clusters (versus its absence in others) is surely due to a synchronic rule in which syllable structure plays a crucial if not yet fully defined role: compare, for example, *hrppi* “above” (probably [hərp.pi]) versus *epre/i*- “back-, rear-” (probably [e.pre/i-]).

## 3.2 Vowels

Lycian has eight vowel phonemes: /i/, /u/, /e/, and /a/ and corresponding contrastive nasalized varieties of each. There are separate letters for /ā/ and /ē/, but not for the nasalized high vowels. Their likely existence is inferred from cases like *Ιῡβρος* for Lycian *Ipre*- ([ibre-]). The non-high vowels form several falling diphthongs with the glides: *ai*, *ei*, *āi*, *ei*, *au*, *eu*. There is no evidence that Lycian has synchronic contrastive vowel length.

### 3.2.1 Vowel assimilation

The most important process affecting Lycian vowels is a pervasive vowel assimilation rule which may be stated in its simplest form as: V [-high] > V [α back] / \_\_ C<sub>0</sub>V [α back]. The rule applies iteratively from right to left within the phonological word (including sequences with proclitics): for example, *tese*- “oath” but collective plural *tasa*; personal name \*/Armanani-/ attested as *Erṁmenēni*. There are many exceptions to the rule as just formulated: thus, dative singular *ladi* (not \**ledi*) to *lada*- “wife.” Some of these may be attributed to paradigmatic analogy, but it is not clear what such a description means in synchronic terms. Furthermore, Hajnal (1995:80ff.), in the most thorough discussion of the phenomenon to the present,



rightly points out that not all exceptions may be attributed to “analogy” in any case. The existence of Lycian umlaut is assured, but a rigorous account of its diachronic and synchronic status requires further study.

### 3.2.2 Syncope

Lycian shows widespread prehistoric syncope. For two independent attempts at a comprehensive description see Melchert 1994a:318ff. and Hajnal 1995:175ff. The broad agreement between the two accounts, despite differences in detail, suggests that their general thrust is correct. Nevertheless, since our knowledge of Lycian accent is indirect, being based almost entirely on the effects of the syncope, the risk of circularity of argument is high, and neither analysis should be taken as remotely definitive.

## 3.3 Phonotaxis

The most noteworthy features of Lycian phonotaxis are the restrictions on initial and final consonants. Inherited word-initial voiced stops were devoiced prehistorically, so that neither /β/ nor /γ/ occurs initially. Initial *dd-* (virtually always spelled as a geminate) does unexpectedly occur. Its source remains unknown. Initial /r-/ occurs in Lycian (A) only rarely, as the result of aphaeresis, and the few examples in Milyan probably should be attributed to the same process. The absence of initial /y-/ may be accidental or systematic. Aside from a handful of cases with unexplained final (unreleased) nasal, Lycian permits only /-s/ in word-final position. Milyan adds -z. Initial consonant clusters are common, including stop plus stop (at least at the phonemic level). The limited number of medial clusters probably is due merely to the very restricted attested lexicon (for a list see Melchert 1994a:297ff.). No final consonant clusters are permitted. Vowels occur freely in all positions in the word. There are no assured examples of heterosyllabic vowel sequences.

## 4. MORPHOLOGY

Lycian inflectional and derivational morphology seems upon first examination to be rather impoverished in comparison with that of other ancient Indo-European languages, Anatolian and non-Anatolian. Closer scrutiny shows that this probably is a misleading impression, an artifact of the relatively limited corpus and the crucial absence of distinct signs for nasalized high vowels.

### 4.1 Nominal morphology

The noun inflects for two numbers (singular and plural), and two genders (animate and inanimate). Animate nouns may have a collective plural beside a count plural (e.g., *wawa/luwa* “cattle” beside anim. acc. sg. *wawā* and unattested nom. and acc. pl. *wawāi\*/wawas\** “cows”). Synchronically, there is evidence only for two genders. However, the contrast between animate nouns with nominative singular *\*-e*, accusative singular *\*-ē < \*-os, \*-om* (respectively), animate nouns with nominative singular *-a*, accusative singular *-ā < \*-eh<sub>2</sub>, \*-eh<sub>2</sub>m* (respectively), and collective *pluralia tantum* in *-a < \*-eh<sub>2</sub>* argues that Lycian (and hence Proto-Anatolian) did inherit from Proto-Indo-European a feminine gender distinct from the masculine and neuter (see Melchert 1992a). There are at least five cases and perhaps six: nominative, accusative, genitive, dative-locative, and ablative-instrumental.

In some noun classes there may be a locative singular distinct from the dative (cf. *a*-stem *ladi* “for/to the wife” vs. *xupa* “in the grave”). The inanimate gender predictably has a single nominative-accusative, and the ablative-instrumental does not distinguish singular and plural, as elsewhere in Anatolian. While there is a genitive plural case, a corresponding genitive singular is found only in a handful of personal names. Possession is normally expressed by means of a relational adjective which agrees in number and case with the head noun and does not indicate the number of the possessor: *mahanahe/i* “divine; of the god(s).” This usage is inherited from Proto-Indo-European, but its nearly complete replacement of the genitive case is a characteristic feature of the western Anatolian languages. The Lycian case endings are inherited or built on inherited material, but the loss of nearly all final consonants (especially postvocalic \*-s) leads to a serious degree of homonymy between case forms.

An important feature of Lycian nominal inflection, shared at least with Luvian and Lydian, is *i*-Motion (better *i*-mutation), as established by Starke (1990:59ff.): many, indeed, most animate nouns and animate forms of adjectives obligatorily add a suffix *-i* to the stem just in the (animate) nominative and accusative, singular and plural. When the base stem ends in *-e*- (< PIE \*-o-), the suffix *-i*- replaces the stem-final *-e*-: for example, *hrzze* “upper” inflects as anim. nom. sg. *hrzzi*\*, anim. acc. sg. *hrzzi* [hərt.tsi], anim. nom. pl. *hrzzi*\*, anim. acc. pl. *hrzzis*\*, but inan. nom.-acc. sg. *hrzzē*, nom.-acc. pl. *hrzza*\*, dat.-loc. pl. *hrzze*\*. The origin of this phenomenon is a matter of serious debate (see Melchert 1994b and Oettinger 1987), but its existence as a synchronic feature of the western Anatolian languages is beyond doubt. The effective inflection of most Lycian nominal stems as *i*-stems in the nominative and accusative has very serious consequences for understanding the Lycian texts. The *i*-stems happen to have the most genuine homonymy of any stem-class: anim. nom. sg., dat. sg., and anim. nom. pl. *-i*. The spelling of anim. acc. sg. [-i] as *-i* as well completes the confusion.

There are clear reflexes of several Proto-Indo-European derivational suffixes, and absence of others is surely due to the restricted corpus.

## 4.2 Pronouns

Lycian attests typical Anatolian features in the first-person singular pronoun *e/amu* “I, me” with *u*-vocalism, in the demonstrative stem *ebe*- “this” (formally matching *apā*- “that” of Hittite, Palaic, and Luvian), and in the interrogative-relative *tī*- < \**k<sup>w</sup>i*-. The enclitic “reflexive” particle *-ti* also is clearly cognate with Luvian *-ti* and Hittite *-z(a)*, but the function of this morpheme in all these languages requires much further study. Evidence for the rest of the pronominal system is almost entirely lacking.

## 4.3 Verbal morphology

The very incomplete picture of the Lycian verb provided by the limited data agrees in most respects with that of the other Anatolian languages: the expected three persons, two numbers (singular and plural), two moods (indicative and imperative), two voices (active and mediopassive), and two tenses (present-future and preterite). There is very limited evidence for a *hi*-conjugation alongside the *mi*-conjugation, as in Hittite (see Ch. 18, §§4.4.7; 4.4.9). The inflectional endings, to the extent that they are known, are comparable to those of Hittite or Luvian, with the exception of medial endings with a nasal: for example, *sijēni* “lies” (see Melchert 1992b for the Lycian, but a convincing account of the prehistory is lacking). One unique feature of Lycian is the morphosyntactic alternation between nasalized and non-nasalized finite verbs: for example, *adeladē* “he/she did/made.” For a persuasive analysis of

this phenomenon see Garrett 1991. The most important of PIE verbal derivational suffixes are securely attested.

There is an infinitive in *-ne/a* which most likely is cognate with Luvian and Palaic *-una*, as per Laroche 1960:172f., contra Melchert 1992a:47, fn. 15. For the source of the final vowel alternation see Hajnal 1995:98. There is a single synchronic participle, with a past passive value for transitive verbs and a stative one for intransitives, as in the other Indo-European Anatolian languages. The suffix is *-Vime/i-*, matching Luvian *-V(i)mma/i-* < \*-(o)mno-. All examples of the suffix *-āt-/ēt(i)-* < \*-e/ont- are lexicalized relics: for example, *lāta-* “dead” (a noun).

## 4.4 Compounds

Attested compounds are not frequent, but they do occur. Neumann (1993:37f.) has convincingly explained *tidere/i-* as “collacteus” < “teat-companion”: compare Hittite *tēda-* “teat” and *arā-* “companion.”

# 5. SYNTAX

## 5.1 Word order and clause structure

Lycian is unique among the Indo-European Anatolian languages in its configurational syntax. There are good reasons to assume an unmarked VSO (Verb–Subject–Object) word order, but two preposing rules which affect the direct object as well as other constituents lead to a surface OVS order. The particular diction of much of the extant corpus happens to make the latter the most frequently attested order. See Garrett 1994 for a discussion of both the synchrony and diachrony of this phenomenon; the same article analyzes in detail the syntax of Lycian relative clauses. Demonstratives and most adjectives typically precede the noun they modify, but the order noun plus adjective is not infrequent and indeed seems to be regular for the relational adjective in *-ahe/i-*. Lycian has several prepositions, but no postpositions. Local adverbs occur both as independent elements and as preverbs.

Lycian is also unique in Anatolian in having true coordinated clauses, marked with *se* “and” (also used to conjoin noun phrases). The conjunction *me* marks prosecutive clauses. There are subordinating temporal and conditional conjunctions, but fronting is also used to mark conditions: *hrppi=ije me tadi* . . . “On-it conj. puts,” in other words, “If one puts thereon” versus *me=ije hrppi=tadi* “conj.-it on-puts,” that is “And (then) one puts thereon” (cf. English “Were I,” equivalent to “If I were”).

## 5.2 Clitics

Lycian employs enclitic pronouns chiefly in clitic doubling in conjunction with topicalization (see Garrett 1992). Conjunctions in Lycian are proclitic (*se* and *me*), not enclitic as in the related Anatolian languages. Lycian does have a few “local particles” which appear as enclitics to the first word in a clause, corresponding to those of Luvian or Hittite.

## 5.3 Syntactic miscellanea

Lycian has at least one example of the Anatolian construction with the direct object of an infinitive in the dative: *esedeñnewi epttehi ntepi=tane* “collateral descendance (dat. sg.)

their in-to put,” in other words, “to put in their collateral descendants.” Examples with the expected accusative also occur.

## 6. LEXICON

The peculiar nature of the extant corpus restricts the known lexicon to an extent which makes statistics about core vocabulary meaningless. However, there is no positive reason to think that the inherited portion of the lexicon is significantly less than the 75–80 percent demonstrated for Hittite. The few identifiable loanwords are predictably from Greek and Iranian and mostly in the expected spheres of government and “high culture”: for example, *sstala*- “stele,” *trijere*- “trireme,” and undoubtedly *sttrat*[ ] “general” from Greek; *xssadrapa*- “satrap” and *sixla*- “shekel” from Iranian (the last of these being ultimately a Semitic word). The only exception to this pattern known to me is *stta*- “stand, be placed standing,” the phonology of which argues that it is a Greek loanword rather than an inheritance.

## 7. READING LIST

The standard edition of Lycian texts discovered by the turn of the century is Kalinka 1901, but these are available in more convenient and often more accurate form in Friedrich 1932. More recent texts are found in Neumann 1979, Laroche 1979 – the *Létôon Trilingual* – and Bousquet 1992. For inscriptions on coins see Mørkholm-Neumann 1978, but many remain unpublished. The most thorough discussion of the alphabet is found in Carruba 1978a. The best description of the synchronic grammar remains that of Neumann 1969, although it is now dated in several respects. For all aspects of Lycian grammar, synchronic and diachronic, global reference should be made to Hajnal 1995. A complete lexicon is available in Melchert 1993. Bryce 1986 offers the best account of the historical and cultural setting.

## Bibliography

- Bousquet, J. 1992. “Les inscriptions gréco-lyciennes.” In H. Metzger (ed.), *Fouilles de Xanthos IX*, pp. 147–199. Paris: Klincksieck.
- Bryce, T. 1986. *The Lycians I. The Lycians in Literary and Epigraphic Sources*. Copenhagen: Museum Tusculanum Press.
- Carruba, O. 1978a. “La scrittura licia.” *Annali della scuola normale superiore di Pisa, classe di lettere e filosofia*, 3rd series, 8:849–867.
- . 1978b. “Il relativo e gli indefiniti in licio.” *Sprache* 24:163–179.
- Eichner, H. 1993. “Probleme von Vers und Metrum in epichorischer Dichtung Altkleinasiens.” In G. Dobesch and G. Rehrenböck (eds.), *Hundert Jahre Kleinasiatische Kommission der Österreichischen Akademie der Wissenschaften*, Vienna: Österreichische Akademie der Wissenschaften.
- Friedrich, J. 1932. *Kleinasiatische Sprachdenkmäler*. Berlin: de Gruyter.
- Garrett, A. 1991. “The Lycian nasalized preterite.” *Münchener Studien zur Sprachwissenschaft* 52:15–26.
- . 1992. “Topics in Lycian syntax.” *Historische Sprachforschung* 105:200–212.
- . 1994. “Relative clause syntax in Lycian and Hittite.” *Sprache* 36:29–69.
- Gusmani, R. 1960. “Concordanze e discordanze nella flessione nominale del licio e del luvio.” *Rendiconti del Istituto Lombardo di Scienze e Lettere. Classe di lettere e scienze morali e storiche* 94:497–512.
- . 1989–1990. “Lo stato delle ricerche sul miliaco.” *Incontri linguistici* 13:69–78.
- Hajnal, Ivo. 1995. *Der lykische Vokalismus*. Graz: Leykam.

- Heubeck, Alfred. 1985. "Konsonantische Geminaten im lykischen Wortanlaut." *Zeitschrift für Vergleichende Sprachforschung* 98:36–46.
- Kalinka, E. 1901. *Tituli Lyciae, lingua lycia conscripti*. Vienna: Hoelder-Pichler-Tempsky.
- Laroche, E. 1960. "Comparaison du louvite et du lycien." *Bulletin de la Société Linguistique de Paris* 55:155–185.
- . 1979. "L'inscription lycienne." In H. Metzger (ed.), *Fouilles de Xanthos VI*, pp. 49–127. Paris: Klincksieck.
- Melchert, H. C. 1992a. "Relative chronology and Anatolian: the vowel system." In R. Beekes, A. Lubotsky, J. Weitenberg, et al., *Rekonstruktion und relative Chronologie. Akten der VIII. Fachtagung der Indogermanischen Gesellschaft, Leiden, 31. August–4. September, 1987*, (eds.) pp. 41–53. Innsbruck: Institut für Sprachwissenschaft der Universität Innsbruck.
- . 1992b. "The middle voice in Lycian." *Historische Sprachforschung* 105:189–199.
- . 1993. *Lycian Lexicon* (2nd fully revised edition). Chapel Hill: self-published.
- . 1994a. *Anatolian Historical Phonology*. Amsterdam: Rodopi.
- . 1994b. "The feminine gender in Anatolian." In G. Dunkel, G. Meyer, S. Scarlata, et al. (eds.), *Früh-, Mittel-, Spätindogermanisch*, pp. 231–244. Wiesbaden: Dr. Ludwig Reichert.
- Mørkholm, O. and G. Neumann. 1978. *Lykische Münzlegenden*. Göttingen: Vandenhoeck and Ruprecht.
- Neumann, G. 1969. "Lykisch." In B. Spuler (ed.), *Altkeinasatische Sprachen*, pp. 358–396. *Handbuch der Orientalistik*. Leiden: Brill.
- . 1979. *Neufunde lykischer Inschriften seit 1901*. Vienna: Österreichische Akademie der Wissenschaften.
- . 1993. "Neue Erkenntnisse zur lykischen Sprache. Appositionen zu Personennamen." In J. Borchhardt and G. Dobesch (eds.), *Akten des II. Internationalen Lykien-Symposiums. Wien, 6.-12. Mai, 1990*, vol. I, pp. 35–38. Vienna: Österreichische Akademie der Wissenschaften.
- Oettinger, N. 1978. "Die Gliederung des anatolischen Sprachgebietes." *Zeitschrift für Vergleichende Sprachforschung* 92:74–92.
- . 1987. "Bemerkungen zur anatolischen *i*-motion und Genusfrage." *Zeitschrift für Vergleichende Sprachforschung* 100:35–43.
- Poetto, M. 1993. *L'iscrizione luvio-geroglifica di Yalburt*. *Studia Mediterranea* 8. Pavia: Iuculano.
- Rasmussen, J. 1974. *Haeretica Indogermanica*. Copenhagen: Munksgaard.
- Starke, Frank. 1990. *Untersuchung zur Stammbildung des Keilschriftluwischen Nomens*. Wiesbaden: Otto Harrassowitz.
- Van den Hout, T. 1995. "Lycian consonantal orthography and some of its consequences for Lycian phonology." In T. van den Hout and J. de Roos (eds.), *Studio Historiae Ardens*, pp. 105–141. Istanbul: Nederlands Historisch-Archaeologisch Instituut.

# Lydian

H. CRAIG MELCHERT

## 1. HISTORICAL AND CULTURAL CONTEXTS

The land called Lydia in Greek sources lay during the first millennium BC on the west central coast of Anatolia, centering on the River Hermos (modern Gediz), with its capital at Sardis (near modern Turgutlu). The indigenous language is attested in graffiti and on coins from the end of the eighth or beginning of the seventh century down to the third, but well-preserved inscriptions of significant length are presently limited to the fifth and fourth centuries, during the period of Persian domination. Lydian texts are thus effectively contemporaneous with those in Lycian.

Extant Lydian texts now number slightly over one hundred, but fewer than thirty of these consist of more than a few words and are reasonably complete. Aside from coins, graffiti, and very short inscriptions on various objects, the overwhelming majority of the inscriptions are on stone. The bulk of these are sepulchral in content, but several of the texts are decrees of one sort or another. Some half-dozen texts are in verse, with a stress-based meter and vowel assonance at line end (see Eichner 1986a and 1993:114ff., with references). All but a handful of the Lydian texts have been found in or near Sardis. For several isolated finds much farther afield see Gusmani 1995:9f. One short Lydian–Aramaic bilingual text helped establish the rudiments of Lydian grammar, but no extensive Lydian–Greek bilingual comparable to the Létôon Trilingual for Lycian (see Ch. 21, §1) has yet come to light.

Lydian shares several characterizing innovations with Hittite and related languages and definitely belongs to the Anatolian subgroup of Indo-European as narrowly defined (see Meriggi 1936, and Melchert 1994a:6f.). Lydian is thus to be derived from an intermediate prehistoric stage we may call Proto-Anatolian. Earlier suggestions that the resemblances between Lydian and the other Anatolian languages are due to convergence are no longer tenable. Nevertheless, the position of Lydian within the Anatolian group is unique and problematic, for at least two reasons.

First, understanding of Lydian remains very limited, comparable to that of Palaic and markedly inferior to that of Luvian or Lycian. The basic grammatical structure of most sentences is clear (aside from some in the verse texts, where unusual word order retards analysis). With rare exceptions, however, grasp of the semantic content ranges from approximate at best to zero at worst. All aspects of the following description should thus be viewed as representing mere hypotheses, of varying degrees of plausibility, not as established facts.

A second difficulty is that Lydian undeniably shows a number of features which are not shared by any other language of the Anatolian group. The limited evidence makes assessment of this fact difficult: are these unique features archaisms preserved only in Lydian, or do

they result from a series of peculiarly pre-Lybian developments? Until a more satisfactory answer to this question is available, the status of Lybian within Anatolian will remain a “special” one.

## 2. WRITING SYSTEM

The Lybian writing system, which is strictly alphabetic, is related to or derived from that of Greek. The exact relationship remains unclear (see Gusmani 1978 and 1995:12). The direction of writing in the older texts is either left to right or right to left. Later texts show exclusively the latter. Use of word-dividers is variable. Values of individual letters are discussed below in the phonology.

**Table 22.1 The Lybian alphabet**

Character	Transcription
Α	a
Β	b
Γ	g
Δ	d (/ð/)
Ε	e
Ϝ	v
Ι	i
Ϟ	y
Κ	k
Λ	l
Μ	m
Ν	n
Ο	o
Ρ	r
Σ Ϝ Ϟ	ś (/s/)
Τ	t
Υ	u
Φ	F
Ψ	q (/kʷ/)
Χ Ϝ	s (/ç/)
Ξ	τ (/tʰ/)
Μ	ā
Υ	ē
Υ	λ
Ϟ	v
↑	c (/dʒ/?)



### 3. PHONOLOGY

#### 3.1 Consonants

The Lydian phonemic inventory consists of the following consonants:

##### (1) Lydian consonantal phonemes

p	t	k	k <sup>w</sup>
	ts		
	dz		
f	s	ç	
v	ð		
m	n	v	
	l	r	ʎ

The nasal conventionally transliterated as *v* (Greek *nu*) is of indeterminate value (see §3.1.3).

##### 3.1.1 Stops

Lydian has a single set of stop phonemes which are probably underlyingly voiceless (for Lydian graphic *b* as /p/ see Gusmani 1965:204ff.). It is very likely that they are realized as voiced allophones in favorable environments, regularly so after nasals (including nasalized vowels), as in Lycian. The name *Alexander* appears as *Alīksāntru-*, while *\*ēndo* “in(to)” results in [ēd-] spelled *ēt-*. There is at least a strong tendency to voicing also next to /r/: note the names *Srkaštu-* and *Atrašta-* rendered in Greek as Σურγαστης and Αδραστης respectively. In rare cases the voiced allophone of the velar /k/ is spelled with a separate letter *g* (e.g., the hapax *qig* for normal *qik* “whatever”), but the allophonic variation is, as expected, not normally indicated in the spelling. Since there is no voicing contrast, there may well be some free variation (note the Greek equivalents of intervocalic /t/ in personal names cited by Gusmani 1988a:191ff.). The place of articulation of /p/ (letter *b*), /t/, and /k/ is undisputed. For arguments that *q* represents a synchronic as well as etymological labiovelar /k<sup>w</sup>/ (e.g., in *qi-* “who, which”) see Heubeck 1959:1–50 and especially Gusmani 1964:33f.

##### 3.1.2 Fricatives and affricates

The letter transliterated as τ (Greek tau) is certainly a voiceless coronal affricate: see Gusmani 1969 with references to Shevoroshkin and others. The definition as a dental alveolar is based on etymological considerations, and a palatal or palatalized articulation cannot be excluded. Characterization of the letter conventionally transliterated *c* as the corresponding voiced affricate is merely an educated guess, and almost any **voiced** coronal affricate or fricative is possible. The one assured source of the sound is an assibilated *\*d*: *civ-* “god” < *\*diw-*.

The synchronic status of the fricatives /f/, /v/, /s/, and /ç/ is clear. The last is the result of prehistoric palatalization of *\*s*, and obviously a palatalized [s] instead of an alveo-palatal or pure palatal is quite possible. By an unfortunate convention too long established to be changed, the sibilant transliterated as *ś* is the dental-alveolar /s/, while *s* is the palatal /ç/. There is strong evidence that the sound represented by the letter *d* is not a voiced stop: the borrowing of the name *Demeter* as *Lamētru-* and internal evidence suggest that Lydian had

no word-initial voiced stops, but *d*- occurs word-initially. The likeliest possibility is a voiced interdental fricative /ð/ (compare Lycian), for which see Melchert 1994c with references to prior works, but some other kind of voiced coronal fricative would also fit the current facts.

### 3.1.3 Sonorants

The nasals and liquids are mostly straightforward. Synchronic /ɰ/ (transliterated with Greek *lambda*) is the result of prehistoric palatalization of \**l*, and once again a palatalized instead of palatal articulation is possible. The value of the nasal transliterated as *ν* (Greek *nu*) is problematic. Its only clear source is **original** word-final nasal (both \**m* and \**n*): loss of word-final vowels makes it synchronically contrastive with the other nasals. This and other distributional facts point to some kind of weakly articulated nasal consonant, but a precise definition is elusive: see Gusmani 1978:842ff. and Melchert 1994a:339 for discussion. The presence of special letters for nasalized vowels (see below) makes it unlikely that the letter *ν* merely indicates nasalization of the preceding vowel. It seems reasonably certain that the sonorants could function as syllabic peaks in Lydian when occurring between other consonants (or alternatively that such sequences were realized phonetically with inserted [ə]): for /m/ note the sentence-initial sequence *fa=k=m=s=ad*, for /r/ *caqrla*- and so forth, and for /ɰ/ *bλtarvod*. Examples such as *kšbλta*- and *dctdid* suggest that even fricatives could form syllabic peaks, at least phonologically (see Eichner 1986a:8).

It is noteworthy that Lydian has no glides, unlike all the other ancient Indo-European Anatolian languages.

## 3.2 Vowels

Lydian probably has a standard five-vowel system /i/, /u/, /e/, /o/, and /a/ plus two contrastive nasalized vowels, though the precise place of articulation of these vowels is open to debate. The vowel transliterated as *γ* is in all likelihood merely an unstressed allophone of /i/ (see Gusmani 1983:57ff. and Melchert 1994a:342). Evidence of Greek transliterations of Lydian names and etymological considerations suggest that the mid vowels /e/ and /o/ were relatively high and long phonetically. It is unlikely that vowel length was synchronically contrastive in Lydian. While the status of occasional spellings in *aa* has not been fully clarified, the existence of other inconsistent attempts to indicate allophonic variation (note *g* and *γ* above) makes it likely that the length of the /a/ in such cases is likewise merely conditioned lengthening under the accent (see Eichner 1986b:215f., and below).

Eichner (1986b, especially 211, n. 21) has presented compelling arguments that the vowels transliterated as *ā* and *ē* represent nasalized vowels, confirming a long-held but occasionally doubted interpretation. The transliteration of *ē* is misleading, however, in that it alternates morphophonemically with /a/ (never with /e/): compare *cēqra*- and derivative *caqrla*-. Etymological considerations (see Melchert 1994a:343) point to a phonetic contrast in length (*ā* = /ā:/, *ē* = /ā/), but this is anything but assured, and one may entertain several other possibilities. As there are no glides, there are no diphthongs.

## 3.3 Accent

Eichner (1986a and 1986b) has convincingly established the essentials of the Lydian accent. While certain details of his analysis may require revision, the skepticism of Gusmani (1988b and elsewhere) of the overall scheme is wholly unjustified. As Eichner demonstrates, the Lydian vowels /e/, /o/, /ā/, and /ē/ regularly occur only under the accent. Using this and other

evidence (syncope and meter), he concludes that all simplex Lydian words are marked by a single accent, which is free to appear on any syllable. Aside from obscure and unanalyzable sequences, the only exceptions to this rule are univerbations of preverb and verb, and indeed, virtually all cases involve specifically the preverbs *ēn-* and *ēt-* “in(to).” Since metrical evidence suggests that in preverb–verb combinations the accent falls on the verb, the irregular appearance of *ē* in these cases is probably due to the influence of the associated free-standing adverb *ēn*.

### 3.4 Synchronic variation

Various cases of allophonic variation have already been cited above. Aside from “sandhi-rules” simplifying certain consonant clusters at morpheme boundary (see Melchert 1994a:351), the only known morphophonemic rule is that by which the nasalized vowels *ā* and *ē* become *a* when unaccented: note again *céqra-* versus *caqrlá-* cited above and see Eichner 1986b:211ff.

### 3.5 Phonotaxis

Lydian phonotactic restrictions differ markedly from those of the other Anatolian languages – in fact, this is superficially perhaps the most striking feature of the language from a comparative point of view. Prehistoric syncope at least as massive as that in Lycian plus regular apocope of original final short vowels combine to produce consonant clusters more typical of Caucasian languages than Indo-European: recall *dctdid* or *kśbłtok-*. For a very preliminary first attempt to describe the syncope see Melchert 1994a:373ff. All Lydian consonants occur word-initially except /*ɰ*/ and /*v*/ . Initial /*r*-/ is rare and surely secondary. Unlike its immediate Anatolian relatives, Lydian permits a wide range of final consonants, including several clusters. As the extreme examples cited above indicate, initial and medial clusters are frequent: for an exhaustive list see Melchert 1994a:352ff.

## 4. MORPHOLOGY

Lydian inflectional morphology is significantly reduced in comparison with other Anatolian languages or older Indo-European languages in general, but typologically it must still be regarded as belonging to the traditional *inflectional* class. The near-absence of demonstrable derivational morphology is surely also due to our limited understanding of the language.

### 4.1 Nominal morphology

The Lydian noun and adjective inflect for the expected two numbers (singular and plural) and two genders (animate and inanimate). Alleged examples of a separate feminine gender have conclusively been shown to be instead collective *pluralia tantum* (see Carruba 1969:44ff.). Assured cases include only nominative, accusative, and dative-locative. The inanimate gender naturally does not distinguish nominative and accusative in either singular or plural. One or two examples of the dative-locative plural (which formally represents the PIE genitive plural *\*-om*) appear to function as an adnominal genitive (*artimułibśimvav* “(to) Artemis of the Ephesians”), but this fact hardly justifies positing a distinct genitive case. Possession and appurtenance are regularly expressed in Lydian by a relational adjective which agrees with the head noun in gender, number, and case: for example, *siuvala/i-* “divine, of the

god(s).” This virtually complete replacement of the genitive case by a relational adjective is a characterizing feature of western Anatolian, shared with Luvian, Lycian, and surely also the poorly attested Carian, Pisidian, and Sidetic. The one or two alleged examples of an ablative-instrumental are almost surely false. The absence of evidence for such a case could be accidental, but one must seriously consider the possibility that Lydian expresses such relationships by the use of adpositions with the dative: note *artimuλdāv* probably “from Artemis.” The formal markers for number, gender, and case are mostly recognizable as inherited, with the notable exception of the dative-locative singular in *-λ*, the origin of which is disputed. Noteworthy is the spread of the Proto-Indo-European pronominal neuter nominative-accusative singular ending in *\*-d* to the noun and adjective: for example, *šfarvad* “oath.”

Lydian shares the feature of “*i*-mutation” described above for Luvian and Lycian (see §4.1), but the phenomenon does not appear to be as widespread: see Starke 1990:82ff. and Melchert 1994b:232ff.

## 4.2 Pronouns

Lydian *amu* “I, me” shows the peculiar Anatolian *u*-vocalism of the first-person singular pronoun. The only assured deictic pronoun is *es-* “this,” of uncertain origin. Decidedly less certain are *os-* “that” (see Eichner 1988) or *āna-* “this” and *ēna-* “that” (see Melchert 1991:137f.). The stem *bi-*, cognate with the deictic stem *apā-* “that” of Hittite, Palaic, and Luvian, functions in Lydian only as the stressed third-person pronoun “he, she, it, they.” Lydian also has as expected enclitic personal pronouns, some formally straightforward (*-av* “him, her, it,” acc. sg. < *\*-om*), others much less so (*-mλ* “to/for him, her,” dat. sg.).

## 4.3 Verbal morphology

The verb has the expected three persons, and two tenses (present-future and preterite). Evidence for a mediopassive beside the active is uncertain, as are possible examples of an imperative contrasting with the indicative. It is tolerably certain that there is a distinction between singular and plural (preterite first singular *-v* vs. preterite first plural *-vv*), but there clearly is no number distinction in the third person, either in the present or in the preterite (respectively *-t/d* and *-l*). An infinitive in *-l* seems reasonably assured, but the status of various proposed participial formations remains uncertain: see for all of this Gusmani 1964:42f.

# 5. SYNTAX

## 5.1 Word order and clause structure

The unmarked word order is SOV (Subject–Object–Verb), but fronting of the verb and other elements for emphasis is not uncommon, and one also finds extraposing of constituents to the right of the verb. These phenomena are by no means limited to the texts in verse. Relative clauses typically precede, with a resumptive pronoun in the main clause, but there are likely examples of postposed relative clauses. Adjectives, including demonstratives and relational adjectives, usually precede their head noun. At least one postposition, *dāv* “from,” seems assured, and others are likely. Lydian cognates of the local adverbs found in other Anatolian languages appear to be limited to use as preverbs, and indeed only unverbated with the verb.

The conjunction *ak-* apparently links Lydian clauses prosecutively (cf. Luvian *a-* and functionally Hittite *nu*). The disjunctive conjunction *buk* “or” conjoins both clauses and noun phrases, while *-k* “also, and” apparently links only noun phrases. Putative subordinating conjunctions are all uncertain.

## 5.2 Clitics

Lydian shows the typical Anatolian use of anaphoric pronouns and sentential “particles” as enclitics to the first accented word in the clause. The function of the various particles is poorly understood, but see Melchert 1991 for the reflexive *-š/is*.

## 5.3 Syntactic miscellanea

Lydian attests at least one example of the Anatolian usage of the dative for the direct object of an infinitive: *karola(v)=š šfēndav arvol*, literally, “of Karos (dat. pl.)-emphatic particle property (dat. pl.) to steal”; in other words, “to steal the property of **Karos**.” The expected accusative is also found.

# 6. LEXICON

For reasons cited in §1 above, it is impossible to say anything useful concerning the Lydian lexicon.

# 7. READING LIST

Gusmani 1964 with supplements (1980, 1982, 1986) furnishes grammar, texts in transliteration, and lexicon combined, along with extensive bibliography. The most thorough discussion of the writing system is Gusmani (1978).

## Bibliography

- Carruba, O. 1969. “Zur Grammatik des Lydischen.” *Athenaeum* NS 47:39–83.  
 Eichner, H. 1986a. “Die Akzentuation des Lydischen.” *Sprache* 32:7–21.  
 ———. 1986b. “Neue Wege im Lydischen I: Vokalnasalität vor Nasalkonsonanten.” *Zeitschrift für Vergleichende Sprachforschung* 99:203–219.  
 ———. 1988. “*os-*, eine sidetisch-lydische Wortgleichung?” *Kadmos* 27:44–56.  
 ———. 1993. *Probleme von Vers und Metrum in epichorischer Dichtung Altkleinasiens. Hundert Jahre Kleinasiatische Kommission der Österreichischen Akademie der Wissenschaften*, pp. 97–169. Vienna: Österreichische Akademie der Wissenschaften.  
 Gusmani, R. 1964. *Lydisches Wörterbuch*. Heidelberg: Carl Winter.  
 ———. 1965. “Sulle consonanti del lidio.” *Oriens Antiquus* 4:203–210.  
 ———. 1969. “Sul samekh lidio.” *Athenaeum* NS 47:136–143.  
 ———. 1978. “La scrittura lidia.” *Annali della Scuola Normale Superiore di Pisa, classe di lettere e filosofia*, 3rd series, 8:833–847.  
 ———. 1980, 1982, 1986. *Lydisches Wörterbuch, Ergänzungsband*, Heidelberg: Carl Winter.  
 ———. 1983. “Ein Weihrauchbrenner mit lydischer Inschrift im Metropolitan Museum.” *Kadmos* 22:56–60.  
 ———. 1988a. “Anthroponomie in den lydischen Inschriften.” In Y. Arbeitman (ed.), *A Linguistic Happening in Memory of Ben Schwartz*, pp. 179–196. Louvain-la-Neuve: Peeters.  
 ———. 1988b. “Zur lydischen Betonung.” *Historische Sprachforschung* 101:244–248.

- \_\_\_\_\_. 1995. "Zum Stand der Erforschung der lydischen Sprache." In E. Schwertheim (ed.), *Forschungen in Lydien*, pp. 9–19. Bonn: Dr. Rudolf Habelt.
- Heubeck, A. 1959. *Lydiaka*. Erlangen: Universitätsbund Erlangen.
- Melchert, C. 1991. "The Lydian emphasizing and reflexive particle *-ś/-is*." *Kadmos* 30:131–142.
- \_\_\_\_\_. 1994a. *Anatolian Historical Phonology*. Leiden Studies in Indo-European 3. Amsterdam: Rodopi.
- \_\_\_\_\_. 1994b. "The feminine gender in Anatolian." In G. Dunkel, G. Meyer, S. Scarlata, *et al.* (eds.), *Früh-, Mittel-, Spätindogermanisch*, pp. 231–244. Wiesbaden: Dr. Ludwig Reichert.
- \_\_\_\_\_. 1994c. "PIE \*y > Lydian d." In P. Vavroušek (ed.), *Iranian and Indo-European Studies. Memorial Volume of Otakar Klíma*, pp. 181–187. Prague: Enigma Press.
- Meriggi, P. 1936. "Der Indogermanische Charakter des Lydischen." In H. Arntz (ed.), *Germanen und Indogermanen: Volkstum, Sprache, Heimat, Kultur. Festschrift für Hermann Hirt*, vol. II, pp. 283–290. Heidelberg: Carl Winter.
- Starke, F. 1990. *Untersuchung zur Stammbildung des keilschrift-luwischen Nomens*. Studien zu den Boğazköy-Texten 31. Wiesbaden: Otto Harrassowitz.

# Carian

H. CRAIG MELCHERT

## 1. HISTORICAL AND CULTURAL CONTEXTS

The land of Caria lay during the first millennium BC in the southwest of Anatolia between Lydia and Lycia. A few dozen texts in the epichoric language, mostly very short or fragmentary, have been found in Caria itself or on objects likely to have originated there. These are dated very approximately to the fourth to third centuries BC. There is also a very fragmentary Carian–Greek bilingual from Athens, dated to the sixth century. By far the largest number of Carian texts consists of tomb inscriptions and graffiti left by Carian mercenaries in Egypt, dating from the seventh to fifth centuries BC. A new epoch in Carian studies has now begun with the dramatic discovery in 1996 of an extensive Carian–Greek bilingual by Turkish excavators in Kaunos and its remarkably swift publication by Frei and Marek (1997).

## 2. WRITING SYSTEM

The Carian script surely stands in some relationship to the Greek alphabet. The direction of writing is predominantly right to left in texts from Egypt, and left to right in those from Caria. *Scriptio continua* is frequent, and use of word-dividers is sporadic.

Decipherment of the Carian script has been a long and arduous task. Pioneering efforts by A. H. Sayce at the end of the nineteenth century were followed by several false steps based on the erroneous assumption of a syllabic or semisyllabic system and a long period of relative neglect. It was the merit of V. Shevoroshkin (1965) to have shown that the Carian script is an alphabet. However, the specific values he and others assigned to individual letters led to no breakthrough in our understanding of the language. Particularly striking was the virtually complete absence of any matches between Carian personal names, as attested in Greek sources, and putative examples in the native alphabet.

A new era began in 1981 when John Ray first successfully exploited the evidence of the Carian–Egyptian bilingual tomb inscriptions to establish radically new values for several Carian letters, as well as to confirm the values of others. Additional investigation, notably by Ray, Ignacio-Javier Adiego, and Diether Schürr, has led to further revisions and refinements of the new system. The basic validity of this approach was shown by its correct prediction of Carian personal names which have subsequently appeared in Greek sources. Nevertheless, many uncertainties and unsolved problems remained, and several reputable experts were skeptical of the new interpretation of the Carian alphabet. One can



conveniently gain a sense of the state of Carian studies prior to 1997 from Giannotta *et al.* 1994.

The new Carian–Greek bilingual from Kaunos has shown conclusively the essential validity of the Ray–Adiego–Schürr system, while also confirming the suspicion of local variation in the use of the Carian alphabet. While some rarer signs remain to be elucidated, the question of the Carian alphabet may be viewed as decided. The new bilingual has not led to immediate equally dramatic progress in our grasp of the language. One reason for this is that the Greek text of the Kaunos Bilingual is a formulaic proxenia decree, while the corresponding Carian is manifestly quite independent in its phrasing of what must be essentially the same contents. The Kaunos Bilingual has provided welcome confirmation of the view that Carian is an Indo-European Anatolian language, and indeed, of the western type of Luvian, Lycian, and Lydian. However, one cannot speak of a complete decipherment until there are generally accepted interpretations of a substantial body of texts – a stage not yet fully attained. This remark applies even to the new bilingual, as one can easily confirm by reading the competing linguistic analyses in Blümel, Frei, and Marek 1998. The following very sketchy description of the language must therefore be taken as highly provisional!

**Table 23.1 A subset of characters of the Carian alphabet**

Character	Transcription
Α	a
ϸ	d
Δ	l
Ϸ	ù
ϸ	r
Ι	λ
⊕	q
Γ	b
Λ	m
Ο	o
ϣ	t
ϣ	š
Μ	s
Υ	u
Χ	x
Υ	n
ΔΔ	p
Φ	ś
Θ	i
□	e
ϣ	w
▽	k
Π	ú
Η	í
↑	τ
ϣ	w

### 3. PHONOLOGY

#### 3.1 Consonants

##### 3.1.1 Obstruents

Carian certainly has a series of voiceless stops /p/, /t/, and /k/. There are actually three letters for dorsal stops: *k*, *q*, and *x*. It is quite unclear whether this orthographic distinction reflects a linguistic contrast, phonetic or phonemic, and if so, of what nature. There are also letters for *b* and *d* the basic value of which is assured by Greek renderings of Carian names. Whether these sounds are voiced stops or fricatives cannot yet be determined. Several indications point to the latter: the existence of separate signs for [mb] and [nd] (Schürr, 1991–1993: 169ff.); the absence or extreme rarity of a corresponding velar; and the apparent lack of voicing contrast in the velar stop(s) as suggested by Greek correspondences. Compare for the first and last points the situation in Lycian. One should, however, avoid premature conclusions.

There are three contrasting sibilant phonemes. Carian *š* is palatal or palatalized, based on Egyptian correspondences in personal names and etymological considerations (*šr*- “up(per)” or similar < \**ser*-; cf. Lydian *serli*- “supreme” likewise with palatal(ized) sibilant). Carian *ś* reflects Proto-Anatolian \*-*ss*- in the relational adjective suffix -*ś*-. The nature of the contrast with the third sibilant *s* remains to be defined. The Carian sound transliterated as *τ* is some kind of coronal obstruent, probably an affricate, but its source and hence its precise value is unknown.

##### 3.1.2 Sonorants

Carian sonorants include /m/, /n/, /r/, and /l/. There is a second lateral transliterated *λ*, which definitely contrasts with ordinary *l* (/l/). The former is rendered consistently in Greek as geminate *λλ* or *λδ* and probably continues prehistoric geminate \*-*ll*-. It is reasonably certain that the sonorants have syllabic allophones.

There are no certain distinct signs for glides, but there are undoubtedly nonsyllabic correspondents of the high vowels /i/ and /u/. Their phonemic status is indeterminate.

#### 3.2 Vowels

Carian appears to have a standard five-vowel system: /i/, /u/, /e/, /o/, and /a/. That the mid-vowels /e/ and /o/ are relatively close and long is suggested by both their likely historical sources and by Greek renderings (Adiego 1994:48ff.). A synchronic contrast in vowel length is unlikely. There is an apparent surfeit of letters for /u/ (*u*, *ú*, *ù*, *ü*, *w*), and additional linguistically real contrasts may eventually emerge, but the possibility of multiple graphemes for a single phoneme must also be taken seriously. Diphthongs /ai/ and /au/ seem assured.

### 4. MORPHOLOGY

#### 4.1 Nominal morphology

Only fragments of Carian morphology are as yet recoverable. In the noun one may identify an animate accusative singular ending /-n/ contrasting with animate nominative singular

ending zero (for the first see Melchert 1993: 79, and now the confirmatory evidence of the Kaunos Bilingual). Possession or appurtenance is indicated by a suffix which appears as *-ś* in the frequent patronymics and ethnica of the sepulchral inscriptions. Identification by Schürr (1992:138) of an animate accusative form in *-śñ* argues that the examples in *-ś* represent the corresponding animate nominative singular of a relational adjective (thus also Adiego 1994:54), rather than a nominal genitive case-marker. The equation of this relational adjective suffix *-ś* with that of Luvian and Lycian is one of the strongest arguments for the status of Carian as an Indo-European Anatolian language. The Kaunos Bilingual has now also given us *kbdun-* “Caunian,” where the *-un-* clearly is cognate with the Luvian suffix *-wann(i)-* which also derives adjectives from place names.

## 4.2 Pronouns

Adiego (1992:32f.) and Melchert (1993:79) have identified a demonstrative stem *s(a)n-* “this.” Hajnal (1997) has now fully confirmed the earlier suspicion that Carian enclitic *-xi* represents the Proto-Indo-European relative pronoun *\*k<sup>w</sup>i-*.

## 4.3 Verbal morphology

Melchert (1993:78f.) has argued that *wbt* represents a preterite third singular verb “has dedicated” (matching Lycian *ubete*). Janda (1994:178) proposes that the verb of the sentence in question is rather *pidl* “has given,” corresponding to Lydian *bil(l)* (< *\*bidl*). A choice between these alternatives depends on finding further convincing examples of one or the other. The lack thus far of any other persuasive identifications of finite verb forms, due in part to the nature of the available corpus, is the most serious obstacle to a complete decipherment of the language. The Kaunos Bilingual has not yet remedied this situation.

# 5. SYNTAX

Hajnal (1997) has compellingly analyzed the enclitic *-xi* as functioning in some instances as a relative pronoun, but in most cases as an invariant particle marking a definite noun phrase. Also noteworthy is the coordinating conjunction *sb* “and,” first correctly identified by Neumann (comparing Milyan *sebe*).

# 6. LEXICON

In addition to various lexemes cited above, one should note the recent identification of *ted* “father” and *en* “mother” (Schürr 1996). Important also is the stem *otr-* “oneself” of the Kaunos Bilingual, independently identified by several scholars with Lycian *atra-* “oneself.”

# 7. READING LIST

The most complete survey is Adiego 1993. One should also consult Giannotta *et al.* 1994, and Blümel, Frei, and Marek 1998.

## Bibliography

- Adiego, I.-J. 1992. "Recherches cariennes: essai d'amélioration du système de J. D. Ray." *Kadmos* 31:25–39.
- . 1993. *Studia Carica. Investigaciones sobre la escritura y lengua carias*. Barcelona: PPU.
- . 1994. "Les identifications onomastiques dans le déchiffrement du carien." In Giannotta *et al.*, pp. 27–63.
- Blümel, W., P. Frei, and C. Marek 1998. *Colloquium Caricum. Akten der Internationalen Tagung über die karisch-griechische Bilingue von Kaunos. 31.10–1.11.1997 in Feusisberg bei Zürich* (= *Kadmos* 37).
- Frei, P. and C. Marek 1997. "Die karisch-griechische Bilingue von Kaunos." *Kadmos* 36:1–89.
- Giannotta, M. E., R. Gusmani, L. Innocente, *et al.* 1994. *La decifrazione del cario. Atti del 1. Simposio Internazionale: Roma, 2–4 maggio 1993*. Rome: Consiglio Nazionale delle Ricerche.
- Hajnal, I. 1997. "Indogermanische Syntax in einer neu erschlossenen anatolischen Sprache: die karische Partikel -xi." In E. Crespo and J. L. García-Ramon (eds.), *Coloquio "Berthold Delbrück y la sintaxis indoeuropes hoy"*, pp. 193–217. Madrid/Wiesbaden: Dr. Ludwig Reichert.
- Janda, M. 1994. "Beiträge zum Karischen." In Giannotta *et al.*, 171–190.
- Melchert, H. C. 1993. "Some remarks on new readings in Carian." *Kadmos* 32:77–86.
- Ray, J. 1981. "An approach to the Carian script." *Kadmos* 20:150–162.
- Schürr, D. 1991–1993. "Imbr- in lykischer und karischer Schrift." *Sprache* 35:163–175.
- . 1992. "Zur Bestimmung der Lautwerte des karischen Alphabets." *Kadmos* 31:127–156.
- . 1996. "Karisch 'Mutter' und 'Vater'." *Sprache* 38:93–98.
- Shevoroshkin, V. 1965. *Issledovanija po dešifrovke karijskix nadpisej*. Moscow: Nauka.

# Attic Greek

ROGER D. WOODARD

## 1. HISTORICAL AND CULTURAL CONTEXTS

Though in this introductory section, and at certain other points as well, attention is given to the ancient Greek language as a whole, the central topic of this chapter will be that dialect called Attic, the spoken dialect of the region of Attica and the principal written dialect of Classical Greek literature. The many other dialects of Greek attested in antiquity will properly be the focus of Chapter 25.

Greek is a member of the Indo-European family of languages. It resides in that major subdivision of the family called centum (see Ch. 17), though its closest linguistic affinities are with the Indo-Iranian and Armenian languages, both members of the satem subset. The arrival in the Balkan peninsula of those Indo-Europeans who would in time be called the Greeks is most probably to be dated to c. 2100 or 1900 BC. One of the three earliest attested Indo-European languages, Greek is first documented on clay tablets recovered from the ruins of various Mycenaean palaces found on the Greek mainland and on the island of Crete, dating c. 1400–1200 BC; already during the Mycenaean period, the language displays dialectal variation. Ancient Greek is phonologically and morphologically quite conservative and has been a cornerstone in the reconstruction of Proto-Indo-European.

The history of the language has been traditionally divided into several chronological phases. Subsequent to the Mycenaean period, the Greeks fell into a prolonged period of illiteracy (though not in Cyprus, see Ch. 25). The language which reappears at the end of this Dark Age is called Archaic Greek, represented principally by the writings of Homer and Hesiod (eighth century BC). With the advent of the fifth-century BC Greek literati, the language is labeled Classical. Though numerous dialects of Greek are attested during the first millennium BC, in both literary and nonliterary sources (enumerated in Ch. 25), the principal dialect of classical literature is Attic. With the expansion of Hellenic culture under Philip of Macedon in the middle of the fourth century BC, the Attic dialect begins to spread geographically, developing into a Hellenistic Koine. This Hellenistic period of Greek continues until the fourth century AD. The final phase of Greek in antiquity is that of the Byzantine era, stretching from the fourth to the fourteenth century AD. All of the dialects of Modern Greek are descendants of Attic, aside from the dialect of Tsaconian, which traces its ancestry to the ancient Laconian dialect.

## 2. WRITING SYSTEMS

The earliest preserved Greek writing systems are syllabic scripts, the Linear B syllabary of the Mycenaeans and the distinct, though clearly related, Cypriot syllabary. Both are discussed in Chapter 25, §§2.1–2.2.

The third of the ancient Greek writing systems and the longest employed is the Greek alphabet. As in the case of the two syllabic scripts which preceded it, the alphabet was founded upon a writing system that the Greeks acquired from a non-Greek people, in this instance the Phoenicians. In typical Canaanite fashion, the segmental writing system of the Phoenicians was consonantal, containing no distinct vowel characters. As the Greek adapters of this Semitic script had no phonetic need for several of the Phoenician consonantal characters (representing consonants not occurring in the Greek language), the Greeks assigned vowel values to these characters, thus creating the first fully alphabetic writing system (i.e., a segmental system containing both distinct consonant and vowel graphemes; see Table 24.1). For example, to the Phoenician character *'aleph*, representing a glottal stop, the Greeks assigned the value of *a* (*alpha*); and to the Phoenician symbol for a voiced pharyngeal fricative, *'ayin*, the Greeks gave the value of *o* (*omicron*). To the end of the Phoenician script (terminating in *taw* (*t*)), additional characters were appended (not all at the same time) – symbols for vowels and for consonants, the latter showing some variation in value among the many local alphabets which arose in the Greek world. The Greek acquisition of the Phoenician script is most probably to be placed in Cyprus, likely in the ninth century BC, in the author's view, though numerous other ideas have been offered.

The numerous local or epichoric alphabets which developed as use of the script spread across the Greek-speaking world can be divided into certain fundamental alphabet-types. This classification is based chiefly, though not solely, on the presence and variety of the so-called “supplemental,” non-Phoenician consonantal characters. The alphabet of Athens and the surrounding region of Attica had belonged to the category of “light blue” alphabets (the color terms which are commonly applied to ancient Greek alphabets have their origin

**Table 24.1 The Greek alphabet**

Character	Phonetic value	Character	Phonetic value
A, α	a(:)	Ξ, ξ	k + s
B, β	b	Ο, ο	o
Γ, γ	g	Π, π	p
Δ, δ	d	Ϻ	s
Ε, ε	e	Ϙ	k
Ϝ	w	Ρ, ρ	r
Ζ, ζ	z + d	Σ, σ	s
Η, η	ɛ:	Τ, τ	t
Θ, θ	t <sup>h</sup>	Υ, υ	ü(:)
Ι, ι	i(:)	Φ, φ	p <sup>h</sup>
Κ, κ	k	Χ, χ	k <sup>h</sup>
Λ, λ	l	Ψ, ψ	p + s
Μ, μ	m	Ω, ω	ɔ:
Ν, ν	n		

in Kirchhoff 1887; see Ch. 25, §2). In 403–402 BC, however, Athens officially adopted the east Ionian alphabet (a “dark blue” script); and it is this form of the alphabetic Greek script which is most familiar to modern readers of Greek (see Table 24.1).

### 3. PHONOLOGY

#### 3.1 Consonants

The phonemic inventory of Attic Greek consonants is presented in Table 24.2.

As illustrated, Attic possesses a symmetrical system of nine oral stops: three manners of stops (voiceless unaspirated, voiceless aspirated, and voiced) produced at three distinct points of articulation (bilabial, dental, and velar; labiovelar stops /k<sup>w</sup>/, /k<sup>wh</sup>/, and /g<sup>w</sup>/ are attested in the second millennium BC dialect of Mycenaean Greek, on which see Ch. 25). Filling out the set of obstruents are two voiceless fricatives – the dental /s/ and the glottal /h/. The Classical Attic sonorant system consists of two nasals, bilabial /m/ and dental /n/ (on velar [ŋ] see below), and two dental liquids, /l/ and /r/. A labiovelar glide /w/ had existed at an earlier phase of Attic and has limited attestation in Attic’s sister dialect of Ionic (and various other dialects; see Ch. 25).

In addition to the bilabial and dental nasal phonemes /m/ and /n/, Attic also possessed a velar nasal [ŋ]. Velar [ŋ] is a positional variant which occurs in two contexts: the dental /n/ becomes [ŋ] when it precedes a velar stop (i.e., /n/ → [ŋ] / \_\_\_\_ {/k/, /g/, /k<sup>h</sup>/}); and the velar stop /g/ becomes [ŋ] when it occurs before the bilabial nasal [m] (i.e., /g/ → [ŋ] / \_\_\_\_ /m/) and perhaps before the dental /n/ as well. There is no distinct alphabetic symbol for the velar nasal; instead the sound is represented by the letter *gamma* (i.e., γκ, γγ, γχ, γμ). *Agma* is reported by Latin grammarians to be the name which the Greeks gave to *gamma* when used to spell [ŋ] (see Allen 1987:33–37).

In early Attic inscriptions, the alphabetic symbol *qoppa* (ϙ) was used to represent a /k/ which occurred next to a back vowel. Such spelling clearly suggests a backed allophone of the velar stop in this position.

#### 3.2 Vowels

Figure 24.1 illustrates the vowel phonemes of Classical Attic and their approximate relative arrangement.

**Table 24.2 The consonantal phonemes of Classical Attic Greek**

Manner of articulation	Place of articulation			
	Bilabial	Dental	Velar	Glottal
<i>Stops</i>				
<i>Voiceless unaspirated</i>	p	t	k	
<i>Voiceless aspirated</i>	p <sup>h</sup>	t <sup>h</sup>	k <sup>h</sup>	
<i>Voiced</i>	b	d	g	
<i>Fricatives</i>		s		h
<i>Nasals</i>	m	n		
<i>Liquids</i>				
<i>Lateral</i>		l		
<i>Nonlateral</i>		r		



	FRONT	CENTRAL	BACK
HIGH	i: / ü:		u:
	i / ü		
	ɛ:		
MID	e		o
	ɛ:		ɔ:
		a	
LOW		a:	

**Figure 24.1** The vowel phonemes of Classical Attic Greek

As can be seen, the vowel system of Classical Attic is markedly asymmetric, with front vowels outnumbering back vowels by more than two to one. Four high-front vowels occur, /i/ (ι), /i:/ (ι), /ü/ (υ), /ü:/ (υ), distinguished by vowel length and presence or absence of lip rounding. In the mid-front region there are three vowels: long tense /ɛ:/ (ει), long lax /ɛ:/ (η) and short /e/ (ε). Two vowels are produced in the low-central region: long /a:/ (α) and short /a/ (α). At the back of the mouth, only three vowels are articulated: long lax mid-back /ɔ:/ (ω), short mid-back /o/ (ο), and long high-back /u:/ (ου). As indicated, long and short vowels are distinguished orthographically only in the case of the mid vowels.

In addition to the monophthongs of Figure 24.1, Classical Attic is characterized by eleven diphthongs:

- (1) “Short” diphthongs
- |           |           |
|-----------|-----------|
| /ai/ (αι) | /au/ (αυ) |
|           | /eu/ (ευ) |
| /oi/ (οι) |           |
| /üi/ (υι) |           |
- “Long” diphthongs
- |                 |            |
|-----------------|------------|
| /a:i/ (ᾱι or ϝ) | /a:u/ (ᾱυ) |
| /ɛ:i/ (ηι or η) | /ɛ:u/ (ηυ) |
| /ɔ:i/ (ωι or ω) | /ɔ:u/ (ωυ) |

At an earlier time in the history of the Attic dialect (perhaps still in the early period of Classical Attic), the vowel sounds written ει and ου had also been diphthongs, /ei/ and /ou/ respectively. However, by the fourth century BC, ει had come to be regularly used to spell both the reflex of the inherited diphthong \*/ei/ and that of the long vowel \*/ɛ:/ (a long vowel which was the product of contraction and compensatory lengthening processes). Likewise, ου was utilized to represent both that sound which descended from the earlier diphthong \*/ou/ and that one which continued the long monophthong \*/ɔ:/ (likewise the outcome of contraction and compensatory lengthening). The orthographic merger of the two vowel sounds in each instance reveals a prior phonological merger: either the inherited diphthongs \*/ei/ and \*/ou/ had become monophthongs or the earlier long monophthongs \*/ɛ:/ and \*/ɔ:/ had undergone diphthongization. Throughout the history of the Greek language, monophthongization is attested recurringly, leaving little doubt that \*/ei/ and \*/ou/ became /ɛ:/ and /ɔ:/ respectively, and not vice versa. This monophthongization had

probably occurred by the fifth century BC. Hence Classical Attic  $\epsilon\iota$  and  $\omicron\upsilon$  are digraphic spellings of monophthongs; one often encounters the term “spurious diphthong” for these digraphs.

A second fundamental diachronic characteristic of Greek vocalic phonology is the fronting and raising of vowels, particularly long vowels, along the periphery of the vowel space. The mid-back vowel  $*/\text{ɔ}:/$  (which had arisen by contraction, compensatory lengthening and monophthongization as discussed above) was raised to become high-back  $*/\text{u}:/$  (probably by the fourth century BC). This raising process appears to have followed upon an earlier fronting of inherited  $*/\text{u}/$  and  $*/\text{u}:/$  to  $*/\text{ü}/$  and  $*/\text{ü}:/$  respectively (perhaps in the sixth century BC or earlier). Fronting and raising of the low-central vowel  $*/\text{a}:/$  perhaps produced an allophone  $*/\text{æ}:/$  which occurred in all contexts except after a preceding  $*/\text{e}/$ ,  $*/\text{i}/$ ,  $*/\text{i}:/$ , or  $*/\text{r}/$  and which would subsequently be further raised to merge with  $*/\text{e}:/$  (though it has also been argued that the raising affected all instances of  $*/\text{a}:/$  and a subsequent back-change of  $*/\text{æ}:/$  to  $*/\text{a}:/$  took place after  $*/\text{e}/$ ,  $*/\text{i}/$ ,  $*/\text{i}:/$ , or  $*/\text{r}/$ ).

### 3.3 Phonotaxis

Attic Greek permits consonants to cluster freely. Word-initially, a variety of biconsonantal clusters occurs ([s + stop]; [s + nasal]; [stop + stop]; [stop + s]; [stop + nasal]; [stop + liquid]; [nasal + nasal]; and at an earlier phase [glide + liquid]) as well as two triconsonantal sequences ([s + stop + liquid]; [s + stop + nasal]). Word-internally, the juxtaposition of syllable-final and syllable-initial consonant clusters generates yet additional permutations of consonants (though many earlier word-internal clusters had been simplified prior to the fifth century). In word-final position the set of possible consonant sequences is more limited: [l + s]; [(m +) p + s]; [( {η, r} +) k + s]. This phonotactic restriction on possible word-final clusters reflects that one which allows only three single word-final consonants in Greek – [r], [n], and [s] (except in the case of clitics).

### 3.4 Syllable structure

It is generally the case that in Attic as in other Greek dialects, word-internal consonant clusters are heterosyllabic. In the case of biconsonantal clusters, a syllable boundary simply falls between the two consonants, regardless of the consonants involved. If the cluster consists of three or more consonants, the boundary falls within the cluster, with its precise location being primarily a function of the relative sonority of the particular consonants which form the cluster. Classical Attic, however, provides a notable exception to the foregoing generalization, showing a certain propensity for open syllables followed by a complex onset in the following syllable. This behavior is observed in the case of a subset of [stop + liquid] and [stop + nasal] clusters (clusters traditionally designated *muta cum liquida*); thus, metrical patterns of Classical Attic verse reveal that at times words such as [kúpris] (Κύπρις “Cyprus”) and [tékmar] (τέκμαρ “token”) are syllabified [k ú | pris] and [té | kmar].

### 3.5 Vowel length

As indicated in Figure 24.1, vowel length is phonemic in ancient Greek. Since the time of Gottfried Hermann, Greek vowel duration has been described in terms of morae: a short vowel is said to consist of a single mora; a long vowel or diphthong of two morae. In antiquity vowel duration was defined in terms of an essentially identical unit, the *k<sup>h</sup>ronós prōtos* (χρονός πρῶτος “primary measure”; see Allen 1987:99–100). By the preceding criteria,

one might anticipate the so-called long diphthongs to consist of three morae; however, for purposes of accent placement, a phenomenon dependent upon the moric structure of a syllable, long diphthongs are treated like other diphthongs and long vowels, in other words as if they were bimoric. Long diphthongs, both those inherited from Proto-Indo-European and those which developed secondarily by contraction, were eliminated over time through shortening of the first vowel of the diphthong or through loss of the second. By the first century BC the spoken Greek language probably no longer possessed such sounds; though in some instances they continued as a part of Greek orthography into the Byzantine period (and hence remain part of the traditional orthography of ancient Greek), represented by the *iota*-subscript (herein transcribed by an *i* within parens).

### 3.6 Accent

Ancient Greek, like its Proto-Indo-European ancestor, was characterized by a pitch or tonal accent. In the traditional orthography of Attic, three different accentual markings are used: acute (´); grave (`) and circumflex (^). The acute and grave diacritics are allographic variants marking high pitch and occurring in complementary distribution: the grave is used on final syllables, unless the accented word occurs at sentence end or is followed by an enclitic, or the accented word is an interrogative; in these exceptional contexts and elsewhere the acute is used. High pitch marked by the acute/grave accent can occur on syllables containing one mora (those with a short vowel) and on syllables of two morae (those with a long vowel or diphthong). In the latter case, high pitch occurs on the rightmost mora of the syllable (i.e., ... |M M̃|Σ ...). In contrast, the circumflex can only occur on syllables containing two morae; within such syllables high pitch occurs on the leftmost mora and falling pitch on the ensuing (rightmost) mora (i.e., ... |M̃ M̃|Σ ...). In the case of the high pitch marked by the acute accent, falling pitch also follows, but in this instance the fall occurs across the succeeding syllable (rather than on the succeeding mora within the same syllable; Allen 1973:234).

While the pitch accent of Proto-Indo-European was free, that of Greek was fixed. The Greek accent can only occur on the final three syllables of a word: the ultima (final), penult (second to final), and antepenult (third to final). The accent of nouns tends to remain on the same syllable throughout the paradigm (subject to the aforementioned limitations), but that of verbs tends to be recessive, occurring as far from the end of the word as the limit of accentuation permits. No more than one mora is permitted to follow the pitch fall which ensues high pitch. The result is that the circumflex accent (... |M̃ M̃|Σ ...) is limited to the ultima and penult, and can only occur on the penult when the ultima contains a short vowel (i.e., only a single mora). The acute accent (i.e., ... |(M) M̃|Σ ...) can then occur on the ultima (in which case it is normally marked by the grave allograph), the penult, and the antepenult, but the antepenult can only bear the acute accent (i.e., have high pitch) if the vowel of the ultima is short.

Attic accent is further characterized by particular requirements. For example, by the so-called Final Trochee Rule of Attic, the occurrence of acute and circumflex accents on the penult is a matter of complementary distribution. If the vowel of the ultima is short and that of the penult is long, high pitch occurring on the rightmost mora of the penult (i.e., acute accent) is retracted to the leftmost mora (i.e., becomes circumflex); in other words [... |M M̃|M̃ #] → [... |M̃ M̃|M̃ #], compare Doric [gūnaïkes] (γυναικες “women”) and Attic [gūnaïkes] (γυναικες). Thus in Attic a penult with a long vowel bears the circumflex if the ultima is short, and the acute if the ultima is long (recall that a circumflex cannot occur on the penult if the ultima is long).

### 3.7 Diachronic developments

#### 3.7.1 Obstruents

Except where affected by conditioned sound changes, the stops of Proto-Indo-European (voiceless, voiced, and voiced aspirated) retain their integrity in Greek, though the voiced aspirates are devoiced:  $*b^h \rightarrow [p^h]$  ( $\phi$ ),  $*d^h \rightarrow [t^h]$  ( $\theta$ ), and so forth. In addition the palatal and velar stop phonemes of Proto-Indo-European merge as Greek velars; thus  $*k̑$  and  $*k \rightarrow [k]$  ( $\kappa$ ),  $*g̑$  and  $*g \rightarrow [g]$  ( $\gamma$ ), while  $*g^h$  and  $*g̑^h \rightarrow [k^h]$  ( $\chi$ ). A subset of the Proto-Indo-European voiced aspirated stops will emerge in historical Greek as plain voiceless stops, without aspiration, by the operation of *Grassman's Law*: within a word, the first of two (non-contiguous) aspirated consonants loses its aspiration (a dissimilatory change also occurring in Sanskrit). Thus, Proto-Greek (PG)  $*t^hrik^hos \rightarrow [trik^h\acute{o}s]$  ( $\tau\rho\iota\chi\acute{o}s$  “of hair”). Voiceless aspirated stops also lose their aspiration before the fricative  $s$ ; this deaspiration occurred prior to the Grassman's Law change, thus bleeding potential instances of such change. For example,  $*t^hrik^hs$ , the Proto-Greek nominative of  $[trik^h\acute{o}s]$ , becomes  $[t^hriks]$  ( $\theta\rho\iota\acute{\varsigma}$ ), removing the conditioning context for aspirate dissimilation and stranding the initial aspirated stop (irregularity so introduced into many paradigms was eliminated by analogy). The Grassman's Law deaspiration also affected instances of  $h$  which precede an aspirated stop; for example, PG  $*hek^h\bar{\rho} \rightarrow [ék^h\acute{\rho}:]$  ( $\varepsilon\chi\omega$  “I have”). Compare the future  $[hék^h\acute{\rho}:]$  ( $\varepsilon\chi\omega$ , in which the initial  $[h-]$  is preserved as a result of  $*k^h$  having previously lost its aspiration before  $[-s-]$ ).

The flagrant exception to the preservation of the integrity of Proto-Indo-European stops is provided by the reflexes of the labiovelar in Attic and other Greek dialects of the first millennium BC. Though the labiovelars are generally preserved in the second-millennium dialect of Mycenaean Greek (with loss of voicing in the case of  $*g^{wh}$ ), they have disappeared completely by the time of the earliest attestation of Attic. Bilabial reflexes emerge as the default development of the labiovelars; in other words, PIE  $*k^w$ ,  $*g^w$ ,  $*g^{wh} \rightarrow [p, b, p^h]$  ( $\pi, \beta, \phi$ ) respectively. Other developments are contextually conditioned. Before and after the high-back rounded vowel  $u$ , the labial element of the labiovelar is dissimilated, producing a velar reflex:  $*k^w, *g^w, *g^{wh} \rightarrow [k, g, k^h]$  ( $\kappa, \gamma, \chi$ ). For example, PIE  $*su-g^wih_3-\bar{e}s \rightarrow [hügíe:s]$  ( $\acute{\upsilon}\gamma\iota\acute{\varsigma}$  “healthy”). In Attic, the labiovelars developed into dental stops when found before the mid-front vowels: PIE  $*k^w, *g^w, *g^{wh} \rightarrow [t, d, t^h]$  ( $\tau, \delta, \theta$ ) respectively; for example,  $*g^welb^h-u- \rightarrow [delp^h\acute{u}s]$  ( $\delta\epsilon\lambda\phi\acute{u}s$  “womb”). Dental reflexes also arise before the high-front vowel  $i$ , but only in the case of the voiceless labiovelar  $*k^w$ ; voiced  $*g^w$  and aspirated  $*g^{wh}$  here give rise to the bilabial reflexes,  $[b]$  and  $[p^h]$  respectively. Thus,  $*k^wi-nu- \rightarrow [tín\acute{\rho}:]$  ( $\tau\acute{\iota}\nu\omega$  “I pay”), while  $*g^wih_3-o- \rightarrow [bíos]$  ( $\beta\acute{\iota}\acute{o}s$  “life”); compare  $[hügíe:s]$  from the same root.

An almost identical course of development is displayed by the Proto-Indo-European consonantal sequence of *palatal stop* + *labiovelar glide*, except that a geminate reflex is generated word-internally. For example, PIE  $*ek^wos \rightarrow [híp\acute{p}os]$  ( $\acute{\iota}\pi\pi\acute{o}s$  “horse”). Word-initially, the outcome is identical to the labiovelar stop development: PIE  $*g^hwēr \rightarrow [t^h\acute{e}:r]$  ( $\theta\acute{\eta}\rho$  “beast”).

Though involved in many particular contextual developments, the Proto-Indo-European fricative  $*s$  shows, broadly speaking, three principal reflexes in Greek:  $[s]$ ,  $[h]$ , and  $\emptyset$ . Word-initially,  $*s$ -becomes  $[h]$  when followed by either a vowel,  $[w]$ , a liquid, or a nasal; for example, PIE  $*septm \rightarrow [heptá]$  ( $\acute{\epsilon}\pi\tau\acute{\alpha}$  “seven”). When the ensuing consonant is  $[l]$  or a nasal, the  $[h]$  is subsequently lost (still preserved in early inscriptional Attic and in other dialects); thus, PIE  $*sla-m-g^w-o- \rightarrow [lambán\acute{\rho}:]$  ( $\lambda\alpha\mu\beta\acute{\alpha}\nu\omega$  “I take”). Intervocally,  $*s$ - likewise becomes

[-h-] and subsequently is lost (without attestation in the first millennium): *\*ĝenh<sub>1</sub>-es-os* → Homeric [géneos] (γένεος; and with vowel contraction) → Attic [génu:s] (γένους “of race”). The Proto-Indo-European fricative is preserved (i) word-initially when followed by a voiceless stop (e.g., *\*sth<sub>2</sub>-tos* → [statós] (στατός “placed”)); (ii) when flanked by a voiceless stop on one side and a vowel on the other (e.g., *\*h<sub>1</sub>esti* → [estí] (ἐστί “(s)he is”)); and (iii) word-finally (as in [génu:s]).

### 3.7.2 Sonorants

The Proto-Indo-European consonantal nasals, *\*m* and *\*n*, and liquids, *\*r* and *\*l*, are well preserved in Attic as in other Greek dialects; though like *\*s*, these consonants are affected by a number of changes which occur in combination with other consonants (see below). Also, Proto-Indo-European *\*m* regularly becomes Greek [-n] in word-final position: for example, *\*sem* → [hén] (ἓν “one”). On the other hand, the Proto-Indo-European syllabic nasals, *\*ṁ* and *\*ṇ*, and syllabic liquids, *\*ṛ* and *\*ḷ*, are both modified in all contexts. The nasals *\*ṁ* and *\*ṇ* become respectively the Greek sequences [am] and [an] before a vowel (optionally preceded by a laryngeal, on which see below) and before a glide; elsewhere they show the common reflex [a]. Thus, *\*deḱṁ* becomes [déka] (δέκα “ten”), while the negative prefix *\*h<sub>2</sub>-* shows up as [an-] in [án-üdros] (ἄν-υδρος “without water”). The syllabic liquids also show a bifurcation of reflexes in Attic, though with somewhat different results. PIE *\*ṛ* gives rise to either [ar] or [ra]. There is uncertainty regarding the precise regular distribution of these two reflexes, though [ar] may occur in approximately the same contexts as [am] and [an], as well as in word-final position. Thus, PIE *\*yēk<sup>w</sup>ṛ* → [hê:par] (ἥπαρ “liver”), while PIE *\*str<sub>1</sub>-to-* → [stratós] (στρατός “army”). The lateral syllabic liquid *\*ḷ* similarly becomes Attic [al] or [la], with perhaps the same distribution as [ar] and [ra], though without word-final reflexes; PIE *\*p<sub>1</sub>lth<sub>2</sub>-u-* → [platús] (πλατύς “wide, flat”).

The two remaining PIE sonorant consonants, *\*y* and *\*w*, are far less persistent in Greek. A palatal glide phoneme /y/ is never attested in ancient Attic, or in any other Greek dialect of the first millennium BC (a [y] offglide which occurs between [i] and an ensuing vowel is sometimes spelled in the syllabic writing system of the Cypriot Greeks and presumably existed in other dialects as well). Word-initially PIE *\*y* in some instances becomes Greek [h], as in [hê:par] (ἥπαρ “liver”), but in other, practically identical word-initial contexts, the Greek reflex is [zd]: PIE *\*yes-o-* → [zdéq:] (ζέω “I boil”). The factors conditioning this split remain unclear. Intervocalic *\*y* has disappeared from the Attic dialect; indirect evidence suggests that *\*[h]* was an intermediate reflex in this process. Thus, PIE *\*treyes* → *\*[trehes]* → *\*[trees]* → (by contraction and raising) [trê:s] (τρῆς “three”). The palatal glide is also involved in various changes in combination with other consonants.

While PIE *\*w* is preserved in many Greek dialects as late as the fourth century BC, its disappearance from Attic-Ionic is relatively early, being attested only in a very few Central and West Ionic inscriptions (in Attic spelling the alphabetic symbol for /w/, *F*, occurs at times, used to represent a [w-] on-glide preceding the vowel /u/). Somewhat like *\*y*, the labiovelar glide shows a developmental bifurcation at the beginning of the word: *\*w* becomes [h] word-initially when followed by [r]; further erosion to *ϕ* occurs when the ensuing sound is a vowel or [l] (though instances of an [h] reflex before a vowel do occur – perhaps conditioned by an [s] following the vowel). Thus, PIE *\*wreh<sub>1</sub>-* → [hrê:tra:] (ῥήτρα “verbal agreement”), while *\*woik-* → [oikos] (οἶκος “house”). Intervocalically, as with *\*y*, *\*w* disappears in Attic without a trace: PIE *\*h<sub>3</sub>ewi-* → [óis] (οἶς “sheep”). When occurring in consonantal sequences, *\*w* experiences yet additional developments.

### 3.7.3 Combinatory changes

In the preceding paragraph, and repeatedly in the foregoing discussion, reference has been made to phonological reflexes which arise when consonants are in contact with one another (so-called combinatory or syntagmatic changes). The following chart summarizes some of the more significant of these phonological developments in Attic:

#### (2) Combinatory phonological developments of Attic

- A. PG  $*p^{(h)}$   $y \rightarrow [pt]$
- B. PG  $*t^{(h)}$   $y \rightarrow [s]$
- C. PG  $*t^{(h)} + y \rightarrow [tt]$  (i.e., when a detectable intervening morpheme boundary occurs; on this complex matter, see Rix 1976:90–91; Lejeune 1982:103–104)
- D. PG  $*k^{(w)(h)}$   $y \rightarrow [t]$  word-initially (i.e., PG  $*k$ ,  $*k^h$ ,  $*k^w$ ,  $*k^{wh}$ )
- E. PG  $*k^{(w)(h)}$   $y \rightarrow [tt]$  elsewhere
- F. PG  $*dy \rightarrow [zd]$
- G. PG  $*g^{(w)}$   $y \rightarrow [zd]$
- H. PG  $*tw \rightarrow [s]$  word-initially
- I. PG  $*tw \rightarrow [tt]$  elsewhere
- J. PG  $*t^{(h)}$ ,  $d\}w \rightarrow \{[t^{(h)}], [d]\}$
- K. PG  $*dl \rightarrow [ll]$
- L. PG  $*bn \rightarrow [mn]$
- M. PG  $*p^{(h)}$ ,  $b\}m \rightarrow [mm]$
- N. PG  $*p^h$ ,  $b\}s \rightarrow [ps]$
- O. PG  $*k^h$ ,  $g\}s \rightarrow [ks]$
- P. PG  $*t^{(h)}$ ,  $d\}s \rightarrow [s]$
- Q. PG  $*ss \rightarrow [s]$
- R. PG  $*ti \rightarrow [si]$  however, the change does not occur if  $*ti$  is preceded by  $*s$
- S. PG  $*t^{(h)}$ ,  $d\}t^{(h)} \rightarrow [st^{(h)}]$
- T. PG  $*r$ ,  $n\}y \rightarrow [y\{r, n\}] / [\{a, o\}]$  —
- U. PG  $*r$ ,  $n\}y \rightarrow [\{r, n\}] / [\{e, i, u\}]$  — with compensatory lengthening of the preceding vowel
- V. PG  $*ly \rightarrow [ll]$
- W. PG  $*ln \rightarrow [l]$  with compensatory lengthening of a preceding vowel
- X. PG  $*r$ ,  $l$ ,  $n$ ,  $s\}w \rightarrow [\{r, l, n\}]$  where  $*s$  is of secondary origin (i.e., not inherited from Proto-Indo-European), without compensatory lengthening of a preceding vowel
- Y. PG  $*N \rightarrow \alpha$  place of articulation / —  $[\text{stop}]_\alpha$  place of articulation (where  $N = \text{nasal}$ )
- Z. PG  $*m\{y, s\} \rightarrow [n\{y, s\}]$
- AA. PG  $*ns \rightarrow [s]$  word-finally; with compensatory lengthening of a preceding vowel
- BB. PG  $*nsV \rightarrow [sV]$  where  $*s$  is of secondary origin (i.e., not inherited from Proto-Indo-European); with compensatory lengthening of a preceding vowel
- CC. PG  $*nsC \rightarrow [sC]$  without compensatory lengthening of a preceding vowel
- DD. PG  $*NsV \rightarrow [NV]$  where  $*s$  is inherited; with compensatory lengthening of a preceding vowel
- EE. PG  $*m\{r, l\} \rightarrow [b\{r, l\}]$  and  $*nr \rightarrow [dr]$  word-initially
- FF. PG  $*m\{r, l\} \rightarrow [mb\{r, l\}]$  and  $*nr \rightarrow [ndr]$  intervocalically
- GG. PG  $*t^{(h)}$ ,  $d\}sC \rightarrow [sC]$
- HH. PG  $*C_i s C_i \rightarrow [sC_i]$ 
  - II. PG  $*CsC \rightarrow [CC]$ , in the case of most remaining PG  $*CsC$  clusters
  - JJ. PG  $*Vs_w \rightarrow [Vw]$  where  $*s$  is inherited; with compensatory lengthening of the preceding vowel and subsequent loss of  $[w]$

KK. PG  $*Vs\{r, l, m, n\} \rightarrow [V\{r, l, m, n\}]$  with compensatory lengthening of the preceding vowel

LL. PG  $*rs \rightarrow [rr]$  where  $*s$  does not belong to the aorist suffix

MM. PG  $\{r, l\}s \rightarrow [s]$  where  $*s$  belongs to the aorist suffix; with compensatory lengthening of the preceding vowel (cf. DD)

### 3.7.4 Laryngeals

It is the Greek language which best preserves evidence of the Proto-Indo-European consonants conventionally called *laryngeal* ( $*h_1$ ,  $*h_2$ , and  $*h_3$ ). When these parent laryngeal sounds are sandwiched between two consonants, each shows a distinctive vowel reflex in Greek ([e], [a], and [o] respectively): for example, PIE  $*ph_2tēr$  gives Greek [patḗ:r] (πατήρ “father”). A laryngeal following the vowel  $*e$  results in a long vowel reflex, also distinctively colored (i.e.,  $*eh_1 \rightarrow [ē:]$ ;  $*eh_2 \rightarrow [ā:] \rightarrow [ē:]$  in Attic-Ionic;  $*eh_3 \rightarrow [ō:]$ ); thus, PIE  $*deh_3-$  yields, with reduplication, [dí-dō:-mi] (δί-δω-μι “I give”). If, on the other hand, the laryngeal precedes a vowel  $*e$ , it distinctively colors but does not lengthen the vowel (i.e.,  $*h_1e \rightarrow [e]$ ;  $*h_2e \rightarrow [a]$ ;  $*h_3e \rightarrow [o]$ ): for example, PIE  $*dh_3-ent-$  produces the aorist participial stem [dont-] (δοντ- “given”). For additional laryngeal developments in Greek, see Rix 1976:68–76.

### 3.7.5 Vowels

As indicated above, the reduction of consonant clusters in Attic is frequently accompanied by lengthening of a short vowel which precedes the cluster. In addition, long vowels were generated by contraction of short vowels which had become contiguous through loss of intervocalic  $*s$ ,  $*y$ , and  $*w$  (most commonly occurring singly, but sometimes in combination) and through morphological restructuring. Contraction is a relatively recent phenomenon in ancient Greek, as is reflected by variation in the outcome of contraction among the different first-millennium dialects. The general results of contraction in Attic are as follows:

- (3) A. Two identical short vowels contract to produce the corresponding long vowel, though the mid vowels [e] + [e] yield [ē:], and [o] + [o] produce  $*[ō:]$ , subsequently raised to [u:] (see §3.2)
- B. A short mid-back vowel contracts with a short mid-front or a low vowel to yield a long mid-back vowel: for example, [a] + [o] gives [ō:] and [e] + [o] gives  $*[ō:]$ , raised to [u:]
- C. While [a] + [e] produces [ā:], [e] + [a] yields [ē:]
- D. The high vowels [i] and  $*[u]$  (see §3.2) form *i*- and *u*-diphthongs with a preceding vowel

Conversely, in Attic, as in all dialects, long vowels become short in certain contexts. Proto-Greek long vowels (though not those arising later) were shortened when they preceded the sequence *sonorant* + *consonant*; thus PG  $*stāntes$  produces Attic [stántes] (στάντες “stood”) – the Greek expression of *Osthoff’s Law*. As a consequence, the first vowel of the so-called long diphthongs is shortened in most word-internal contexts (the second diphthongal element serving as a glide in the operation of this change). At times, long vowels in Attic and certain other dialects also undergo shortening when followed by another vowel: compare Homeric [basilē:ō:n] (βασιλήων) and Attic-Ionic [basiléō:n] (βασιλέων “of kings”). However, in the case of the sequences [ē:a] and [ē:o], concomitant with this shortening, the second vowel is sometimes lengthened (quantitative metathesis) in Ionic and, especially, Attic: thus, Homeric [basilē:os] (βασιλήος), but Attic [basiléō:s] (βασιλέως “of a king”).



## 4. MORPHOLOGY

### 4.1 Nominal morphology

The Greek nominal is morphologically marked for case, gender, and number. Five different grammatical cases are identified: vocative, nominative, accusative, genitive, and dative. In certain inflectional classes, each case-marker has a distinct morphological form. The functions of the Proto-Indo-European ablative have been absorbed by the Greek genitive, and the locative and instrumental by the Greek dative. Three nominal genders, feminine, masculine, and neuter, are distinguished; and nouns are inflected in three numbers: singular, dual, and plural. By the fifth century BC, however, the dual has become restricted in use, and by the Hellenistic period has disappeared except in a few frozen contexts.

#### 4.1.1 Noun classes

Within Greek grammatical tradition, nouns are divided into three declensional classes: the principally feminine first declension; the predominantly masculine and neuter second declension; and the third declension, of mixed gender. Each of the declensions has Proto-Indo-European ancestry. Within the parent Indo-European language, nominals, as well as verbals, are characterized by a tripartite structure; each word consists of a *root*, to which is optionally attached a *suffix*, followed in turn by an *ending* (R + (S) + E). Regarding morphological typology, Greek is predominantly a fusional language. This is clearly illustrated by the paradigm of (4) below, in which endings and suffixes freely combine and lose their morphological integrity.

##### 4.1.1.1 First declension

The majority of first declension feminine nouns of Greek are descended from Proto-Indo-European nouns formed with the suffix *\*-eh<sub>2</sub>-*. As noted above, by regular sound change PIE *\*-eh<sub>2</sub>-* becomes Greek [a:] (ᾱ), which in Attic, in most contexts, is raised and fronted to [ɛ:] (η). This characteristic η vowel is obscured in the plural of the first declension by contraction and morphological restructuring. As an example of first declension nouns of this type, consider the paradigm of *tīmē* (τίμη “honor”).

#### (4) The Attic first declension I

	<i>Singular</i>	<i>Dual</i>	<i>Plural</i>
<i>Nominative</i>	tīmē (τίμη)	tīmá (τίμᾱ)	tīmái (τίμαί)
<i>Vocative</i>	tīmē (τίμη)	tīmá (τίμᾱ)	tīmái (τίμαί)
<i>Accusative</i>	tīmēn (τίμην)	tīmá (τίμᾱ)	tīmás (τίμας)
<i>Genitive</i>	tīmēs (τίμης)	tīmáin (τίμαῖν)	tīmôn (τίμων)
<i>Dative</i>	tīmē(i) (τίμηι)	tīmáin (τίμαῖν)	tīmáis (τίμαῖς)

Early Attic attests a dative plural in which the η stem-vowel is still preserved, as in *díkēsi* (δίκησι “for penalties”). The long ᾱ of the nominative, vocative, and accusative dual is secondary.

When the noun root ends in [e, i, i:] or [r], the [a:] reflex of the PIE *\*-eh<sub>2</sub>-* suffix is preserved in Attic, thus producing a first declension singular of the type of *k<sup>h</sup>ōrā* (χώρᾱ “place”):



## (5) The Attic first declension II

	<i>Singular</i>
<i>Nominative</i>	k <sup>h</sup> ǫrā (χώρᾱ)
<i>Vocative</i>	k <sup>h</sup> ǫrā (χώρᾱ)
<i>Accusative</i>	k <sup>h</sup> ǫrān (χώρᾱν)
<i>Genitive</i>	k <sup>h</sup> ǫrās (χώρᾱς)
<i>Dative</i>	k <sup>h</sup> ǫrā(i) (χώρᾱι)

The dual and plural of this type are identical to those of the *tīmē* type.

Proto-Indo-European also formed nominals with an ablauting suffix *\*-yeh<sub>2</sub>-* (*e*-grade), *\*-ih<sub>2</sub>-* (*o*-grade). Developing the respective Proto-Greek reflexes *\*-yā* and *\*-ya*, Attic [-ē:] (ἡ) and [-a] (α), nouns of this type fall formally into the feminine first declension. This suffix is quite frequently attached to roots and stems ending in a consonant, which, in combination with the ensuing glide *\*-y*, is subject to sound change. Thus, the root *\*ped-* (“foot”) provides a noun *trápezda* (τράπεζα “table”; see (2F)), *\*glok<sup>h</sup>-* gives *glōtta* (γλῶττα “tongue”; see (2E)), *\*smor-* gives *moira* (μοῖρα “portion”; see (2S)), and so forth.

## (6) The Attic first declension III

	<i>Singular</i>	
<i>Nominative</i>	trápezda (τράπεζα)	} with the suffix <i>*-ih<sub>2</sub>-</i>
<i>Vocative</i>	trápezda (τράπεζα)	
<i>Accusative</i>	trápezdan (τράπεζαν)	
<i>Genitive</i>	trapézdēs (τραπέζης)	} with the suffix <i>*-yeh<sub>2</sub>-</i>
<i>Dative</i>	trapézdē(i) (τραπέζηι)	

The dual and plural are formed like that of *tīmē* and *khǫrā*. Thus, the so-called *ā*-feminine of the first declension differs from the other feminine nouns of this declension only in the nominative, accusative, and vocative of the singular.

Also derived from stems in *\*-eh<sub>2</sub>-* and placed within the Greek first declension is a group of masculine nouns having a nominative singular ending in *-ēs* (-ης):

## (7) The Attic first declension IV

	<i>Singular</i>
<i>Nominative</i>	polítēs (πολίτης)
<i>Vocative</i>	políta (πολίτα)
<i>Accusative</i>	polítēn (πολίτην)
<i>Genitive</i>	polítū (πολίτου)
<i>Dative</i>	polítē(i) (πολίτηι)

The nominative and genitive singular have been influenced by the masculine nouns of the second declension. Both the dual and plural are formed like those of the feminine nouns of the first declension.

## 4.1.1.2 Second declension

The nouns of the Greek second declension, continuing the thematic stems of Proto-Indo-European, are characterized by a suffix terminating in the vowel *o* or *e* (sometimes obscured by sound change). The inflection of the masculine nouns is here demonstrated with *lúkos* (λύκος “wolf”):

## (8) The Attic second declension I

	<i>Singular</i>	<i>Dual</i>	<i>Plural</i>
<i>Nominative</i>	lúkos (λύκος)	lúkō (λύκω)	lúkoī (λύκοι)
<i>Vocative</i>	lúke (λύκε)	lúkō (λύκω)	lúkoī (λύκοι)
<i>Accusative</i>	lúkon (λύκον)	lúkō (λύκω)	lúkūs (λύκους)
<i>Genitive</i>	lúkū (λύκου)	lúkoīn (λύκοιν)	lúkōn (λύκων)
<i>Dative</i>	lúkō(i) (λύκω)	lúkoīn (λύκοιν)	lúkoīs (λύκοις)

Early Attic preserves a dative plural ending in *-oisi* (-οισι). A very few nouns following the above inflectional pattern have feminine gender.

With the exception of the nominative, vocative, and accusative case forms, both singular and plural, neuter nouns of the second declension have the same inflection as the masculine nouns. Consider the paradigm of *zdiugón* (ζυγόν “yoke”):

## (9) The Attic second declension II

	<i>Singular</i>	<i>Dual</i>	<i>Plural</i>
<i>Nominative</i>	zdügón (ζυγόν)	zdügō (ζυγώ)	zdügá (ζυγά)
<i>Vocative</i>	zdügón (ζυγόν)	zdügō (ζυγώ)	zdügá (ζυγά)
<i>Accusative</i>	zdügón (ζυγόν)	zdügō (ζυγώ)	zdügá (ζυγά)
<i>Genitive</i>	zdügū (ζυγοῦ)	zdügoīn (ζυγοῖν)	zdügōn (ζυγῶν)
<i>Dative</i>	zdügō(i) (ζυγῶ)	zdügoīn (ζυγοῖν)	zdügoīs (ζυγοῖς)

Contraction of the thematic vowel with a preceding *-o-* or *-e-* gives rise to a set of second declension masculine and neuter nominals having a long vowel in the inflection of the nominative, accusative, and vocative singular: for example, nominative masculine singular *nûs* (νοῦς “mind”); accusative singular *nûn* (νοῦν); nominative, accusative neuter singular *ostûn* (ὀστοῦν “bone”). Contraction often also occurs in the nominative, accusative neuter plural, yielding a final long *-ā*, as in *ostâ* (ὀστᾶ).

Yet other sound changes, including quantitative metathesis, produce a distinctive second declension inflectional paradigm marked by the presence of the long vowel *-ō-* (-ω-), the so-called Attic declension. Consider the paradigm of Attic *neōs* (νεώς “temple”; Ionic *nēōs*, νηός) as an example:

## (10) The Attic second declension III

	<i>Singular</i>	<i>Dual</i>	<i>Plural</i>
<i>Nominative</i>	neōs (νεώς)	neō (νεώ)	neō(i) (νεῶ)
<i>Vocative</i>	neōs (νεώς)	neō (νεώ)	neō(i) (νεῶ)
<i>Accusative</i>	neōn (νεῶν)	neō (νεώ)	neōs (νεώς)
<i>Genitive</i>	neō (νεῶ)	neō(i)n (νεῶν)	neōn (νεῶν)
<i>Dative</i>	neō(i) (νεῶ)	neō(i)n (νεῶν)	neō(i)s (νεῶς)

## 4.1.1.3 Third declension

The Greek third declension is the historical, grammatical repository of a broad array of Proto-Indo-European athematic noun stems. These stems are *athematic* in that they end in a consonant or in the vowel *i* or *u* (in other words, in some sound other than the thematic vowel *o/e*). In Proto-Indo-European such stems were characterized by distinctive patterns of ablaut variation and accent placement. No fewer than four fundamental patterns have

been identified for the parent language (though this is a matter on which there is not full agreement among Indo-Europeanists): acrostatic (with two subtypes), amphikinetic, proterokinetic, and hystero-kinetic. The following table schematically summarizes ablaut gradation (*e*-grade/*o*-grade/*ø*-grade) and accent placement for each of these athematic noun-types of Proto-Indo-European:

**Table 24.3 Ablauting noun patterns of PIE**

	Strong stem			Weak stem		
	Root	Suffix	Ending	Root	Suffix	Ending
Acrostatic I	ó	ø	ø	é	ø	ø
Acrostatic II	ě	ø	ø	é	ø	ø
Amphikinetic	é	o	ø	ø	ø	é
Proterokinetic	é	ø	ø	ø	é	ø
Hystero-kinetic	ø	é	ø	ø	ø	é

In addition to these, Proto-Indo-European also possessed root nouns (athematic nouns having a root which serves as a stem without attachment of a suffix) displaying a distinct pattern of accent and ablaut variation between strong and weak stems. For masculine and feminine nouns, the strong stem is usually identified as that of the (a) nominative singular, dual, and plural; and (b) the accusative singular and dual. The strong stem of the neuter is that of the nominative and accusative plural. The stem of all other cases is weak.

Greek is one of the languages which best provides evidence of this Proto-Indo-European inflectional phenomenon. Even so, the ancestral patterns have often been obscured in Greek by processes of paradigm regularization; for example, within a given paradigm Greek has essentially limited ablaut variation to the suffix. Consequently, in a synchronic grammatical description of Greek, third declension noun stems are more appropriately and efficiently categorized by their final member than by their ancestral ablaut and accent pattern.

The endings which are attached to Greek nouns of the third declension are the following:

#### (11) The Attic third declension endings

	<i>Singular</i>	<i>Dual</i>	<i>Plural</i>
<i>Nominative</i>	-s (-ς) or ø	-e (-ε)	-es (-ες) or -a (-α)
<i>Vocative</i>	-s (-ς) or ø	-e (-ε)	-es (-ες) or -a (-α)
<i>Accusative</i>	-a (-α) or -n (-ν)	-e (-ε)	-as (-ας), -s (-ς) or -a (-α)
<i>Genitive</i>	-os (-ος)	-oin (-οιν)	-ōn (-ων)
<i>Dative</i>	-i (-ι)	-oin (-οιν)	-si (-σι)

The endings of the third declension and those of the first declension share a common Proto-Indo-European heritage – distinct from that set of endings utilized for inflecting thematic nouns (second declension). Sound changes will in some instances arise when the ending is attached to the stem, obscuring the phonetic shape of both ending and stem. Analogical remodeling of particular case forms also commonly occurs within third declension paradigms.

Each of the principal third declension stem-types is here illustrated using a partial paradigm (the illustration is not, however, exhaustive, as various distinct subcategories exist for most stem-types):

1. stop-stems (stems ending in a stop). (A) *p<sup>h</sup>lēb-* (“vein,” fem.): *p<sup>h</sup>lēp-s* (φλέψ, nom. sg.), *p<sup>h</sup>lēb-ós* (φλεβός, gen. sg.), *p<sup>h</sup>lēb-a* (φλέβα, acc. sg.); (B) *pod-* (“foot,” masc.): *pú-s* (πούς, nom. sg., the vowel is irregular, < \**pod-s*); *pod-ós* (ποδός, gen. sg.); *po-sí* (ποσί, dat. pl., < \**pod-si*).
2. *s*-stems. *genes-* (“race,” neut.): *gén-os* (γένος, nom./acc. sg., i.e., *gén-os-ø*), *gén-ūs* (γένους, gen. sg., < \**gen-e-os* < \**gen-es-os*), *gén-ē* (γένη, nom./acc. pl., < \**gen-e-a* < \**gen-es-a*).
3. *n*-stems. (A) *poimen-* (“shepherd,” masc.): *poi-mēn* (ποιμήν, nom. sg., i.e., *poi-mēn-ø*, lengthening of stem-vowel is of Proto-Indo-European date), *poi-mén-os* (ποιμένος, gen. sg.), *poi-mé-si* (ποιμέσι, dat. pl. < \**poi-mn-si* with *ø*-grade of the suffix; regular phonological reflex *-ma-* analogically modified to *-me-*); (B) *sōmat-* (“body,” neut.): *sō-ma* (σῶμα, nom./acc. sg., < \**sō-mn-ø*), *sō-mat-os* (σώματος, gen. sg., < \**sō-mn-t-os*, the source of the *-t-* is uncertain; it occurs throughout the paradigm of the neuter *n*-stems, other than in the nom./acc. sg., and is found also in other types of third declension paradigms).
4. *r*-stems. *pater-* (“father,” masc.): *pa-tēr* (πατήρ, nom. sg., i.e., *pa-tēr-ø*, lengthening of stem-vowel is of Proto-Indo-European date), *pa-tr-ós* (πατρός, gen. sg.), *pa-tér-as* (πατέρας, acc. pl.).
5. *r/n*-heteroclite stems (*r*-stem in the nom./acc. sg. and *n*-stem elsewhere). *hēpar-* (“liver,” neut.): *hēp-ar* (ἥπαρ, nom./acc. sg., i.e., *hēp-ar-ø*), *hēp-at-i* (ἥπατι, dat. sg., with *-t-* as in neuter *n*-stems), *hēp-a-si* (ἥπασι, dat. pl.).
6. *i*-stems. (A) *poli-* (“city,” fem., ablauting suffix): *pól-i-s* (πόλις, nom. sg.), *pól-e-ōs* (πόλεως, gen. sg., < *pól-ē-os* by quantitative metathesis), *pól-ēs* (πόλεις, nom. pl. < \**pol-ey-es*); (B) *oi-* (“sheep,” masc./fem., nonablauting suffix): *oi-s* (οἷς, nom. sg.), *oi-ós* (οἰός, gen. sg.), *oi-es* (οἷες, nom. pl.); see also Ch. 25, §4.1.1.3.
7. *u*-stems. (A) *pēk<sup>h</sup>ü-* (“forearm,” masc., ablauting *-ü-* suffix): *pēk<sup>h</sup>-ü-s* (πῆχυς, nom. sg.), *pēk<sup>h</sup>-ēs* (πήχεις, nom. pl. < \**pēk<sup>h</sup>-ew-es*); (B) *sü-* (“sow,” fem., nonablauting *-ü-* stem): *sü-s* (σῦς, nom. sg.), *sü-es* (σύες, nom. pl. < \**suw-es*).
8. diphthongal *u*-stems. *basileu-* (“king,” masc., *ēu*-stem): *basil-eú-s* (βασιλεύς, nom. sg., < \**basil-ēu-s*), *basil-é-ōs* (βασιλέως, gen. sg., < \**basil-ēw-os* by quantitative metathesis), *basil-é-ās* (βασιλεῶς, acc. pl., < \**basil-ēw-as* by quantitative metathesis).

#### 4.1.2 Adjectives

Greek adjectives are constructed by utilizing most of the nominal stem-types which were elaborated above. As adjectives agree with the nouns they modify in case, gender, and number, any single adjective, unlike most nouns, can be assigned multiple genders. The most commonly occurring adjectives are those which form the feminine, in Attic, using an *-ē-* stem (first declension) and form the masculine and neuter using a thematic stem (second declension): *agath-ós* (ἀγαθός “good,” masc.), *agath-ē* (ἀγαθή, fem.), *agath-ón* (ἀγαθόν, neut.). Some adjectives make no *morphological* distinction between masculine and feminine gender. A subset of these are thematic adjectives with the common nonneuter gender marked by masculine inflection; such adjectives commonly contain prefixes: *á-dik-os* (ἄδικος “unjust,” masc. and fem.), *á-dik-on* (ἄδικον, neut.). Certain adjectives of this type conform to the “Attic declension” discussed above. Similarly, consonant stem adjectives commonly have a single masculine/feminine form: for example, the *s*-stem *alēthēs* (ἀληθής,

“true,” masc. and fem.), *alē<sup>h</sup>ēs* (ἀληθής, neut.). In contrast, adjectives formed from *u*-stems (stems formed with a short *-u-* suffix as opposed to the long *-ū-* of most *u*-stem nouns) distinguish the three genders morphologically, forming the feminine by utilizing the short *-a-* morphology of the first declension (i.e., the PG suffix *\*-ya/ya-*, PIE *\*-ih<sub>2</sub>/yeh<sub>2</sub>-*): *hēdūs* (ἡδύς, “sweet” masc.), *hēdēa* (ἡδεῖα [from PG *\*swād-ew-ya*], fem.), *hēdū* (ἡδύ, neut.). Certain *n*-stem adjectives as well as adjectives formed with a suffix terminating in *-nt-* (compare the active participle below) also make a three-way morphological distinction, utilizing the *\*-ya/ya-* formant for the feminine.

Comparatives and superlatives are productively generated by attaching the suffixes *-tero-* and *-tato-* respectively to the adjective stem: *glūkūs* (γλυκύς “sweet”), *glūkū<sup>h</sup>-tero-s* (γλυκύ-τερο-ς “sweeter”), *glūkū<sup>h</sup>-tato-s* (γλυκύ-τατο-ς “sweetest”). Less commonly, Greek produces the comparative with a suffix *-iōn-* attached directly to the adjective root, in origin the *ø*-grade (*\*-is-*) of an ablauting *s*-stem suffix *\*-yes-* to which Greek appended a nasal formant: *hēd-ūs* (ἡδ-ύς “sweet”), *hēd-iōn* (ἡδ-ίων “sweeter”). The corresponding superlative marker is produced by attaching *-to-* to the *ø*-grade: *hēd-is-to-s* (ἡδ-ισ-το-ς “sweetest”).

### 4.1.3 Pronouns

Attic and the other dialects of ancient Greek possess a wealth of pronouns.

#### 4.1.3.1 Personal pronouns

Personal pronouns, enclitic and accented forms, occur in the singular, dual, and plural for each of the three persons, though by the period of Classical Attic, the third-person forms, aside from the dative singular and plural, are little used, and when they are used have a reflexive function (see 4.1.3.2):

#### (12) Attic personal pronouns

	<i>Singular</i>		
	<i>First</i>	<i>Second</i>	<i>Third</i>
<i>Nominative</i>	egō (ἐγώ)	sū (σύ)	—
<i>Genitive</i>	emū (ἐμοῦ)	sū (σοῦ)	hū (οῦ)
<i>Dative</i>	emoi (ἐμοί)	soi (σοί)	hoi (οἱ)
<i>Accusative</i>	emé (ἐμέ)	sé (σέ)	hé (ἐ)
	<i>Dual</i>		
	<i>First</i>	<i>Second</i>	
<i>Nom./Acc.</i>	nō (νώ)	sp <sup>h</sup> ō (σφώ)	
<i>Gen./Dat.</i>	nō(i)n (νῶν)	sp <sup>h</sup> ō(i)n (σφῶν)	
	<i>Plural</i>		
	<i>First</i>	<i>Second</i>	<i>Third</i>
<i>Nominative</i>	hēmēs (ἡμεῖς)	hūmēs (ὕμεῖς)	sp <sup>h</sup> ēs (σφεῖς)
<i>Genitive</i>	hēmōn (ἡμῶν)	hūmōn (ὕμῶν)	sp <sup>h</sup> ōn (σφῶν)
<i>Dative</i>	hēmin (ἡμῖν)	hūmin (ὕμῖν)	sp <sup>h</sup> isi (σφίσι)
<i>Accusative</i>	hēmās (ἡμᾶς)	hūmās (ὕμᾶς)	sp <sup>h</sup> ās (σφᾶς)

The oblique forms of the singular personal pronouns and the dative of the third-person plural also occur as enclitics, in which case the first-person pronouns lack the initial *e-* (i.e., *mū* (μου), etc.). Furthermore, the oblique cases of the first and second plural pronouns are found with accent on the initial syllable; such forms have been similarly designated as enclitic or, alternatively, as simply “unemphatic” (see Allen 1973:243).

Utilizing the stem of the personal pronouns, *possessive pronominal adjectives* were derived by attaching the thematic suffixes *-o-* and *-tero-*; feminine forms are constructed with the long *-ā-* morphology of the first declension. Nominatives of the first and second persons respectively are formed as follows: (i) *emós* (ἐμός masc.), *emē* (ἐμή fem.), *emón* (ἐμόν neut.); (ii) *sós* (σός masc.), *sē* (σή fem.), *són* (σόν neut.). Instead of the third-person possessive adjective – *hós* (ὅς masc.), *hē* (ἡ fem.), *hón* (ὅν neut.) – Classical Attic normally uses masculine/neuter *autū* (αὐτοῦ) and feminine *autēs* (αὐτῆς), genitives of the pronoun *autós* (αὐτός, etc., see below). First and second singular possessives are at times also used reflexively. Plural possessives of the first and second persons appear in the nominative masculine singular as *hēmēteros* (ἡμέτερος) and *hūmēteros* (ὕμέτερος) respectively. Attic normally uses *autōn* (αὐτῶν), the genitive plural of *autós*, for third-person possession. A third-person possessive *sp<sup>h</sup>ēteros* (σφέτερος), etc. is reflexive in use, normally accompanied by *autōn*; the first and second plural forms are commonly used as reflexive possessives also (usually in combination with *autōn*).

#### 4.1.3.2 Reflexive pronouns

The reflexive pronouns of Attic were formed from the personal pronouns used in combination with the pronoun *autós*. In the singular these have undergone univerbation (not yet having been joined in Homer) and only the second member shows inflection (occurring only in the oblique cases), with a thematic masculine/neuter and long *-ā-* feminine. The genitive singular is thus formed as follows: (i) first person *emautū* (ἐμαυτοῦ “myself,” masc.), *emautēs* (ἐμαυτῆς fem.); (ii) second person *s(e)autū* (σ(ε)αυτοῦ “yourself,” masc.), *s(e)autēs* (σ(ε)αυτῆς fem.); (iii) *h(e)autū* (ἐαυτοῦ or αὐτοῦ “himself, itself,” masc./neut.), *h(e)autēs* (ἐαυτῆς or αὐτῆς, “herself,” fem.). In contrast, the two elements of the plural reflexives remain independent; consider the genitive plural (note that the genitive plural is identical for all genders): (i) first person *hēmōn autōn* (ἡμῶν αὐτῶν “ourselves”); (ii) second person *hūmōn autōn* (ὕμῶν αὐτῶν “yourselves”); (iii) third person *sp<sup>h</sup>ōn autōn* (σφῶν αὐτῶν “themselves”). However, at an early period in Attic, the third singular reflexive is generalized to the third plural so that *h(e)autōn* and the other case forms eventually usurp the position of *sp<sup>h</sup>ōn autōn*, etc. (moreover, the *h(e)aut-* morpheme will in time be completely generalized, replacing the reflexive forms of the first and second persons, singular and plural). As pointed out above, Attic also uses the third-person pronouns (*hū*, *hoi*, *hé*, *sp<sup>h</sup>ōn*, *sp<sup>h</sup>isi*, *sp<sup>h</sup>ās*) reflexively. These function as the so-called “indirect” or “long-distance” reflexives, appearing in subordinate clauses and having an antecedent in a higher clause (though the *h(e)aut-* third-person reflexive frequently is also so used).

#### 4.1.3.3 Reciprocal pronoun

In addition to the reflexive, Greek possesses a reciprocal pronoun *allēlo-* (ἀλλήλο-), meaning “each other, one another.” It occurs in the oblique cases of the dual and plural. The accusative masculine, feminine, and neuter plural are offered as examples: *allēlūs* (ἀλλήλους), *allēlās* (ἀλλήλας), *allēla* (ἀλλήλα).

#### 4.1.3.4 Definite article

Under the heading of demonstrative pronouns can be treated the Greek definite article, which had its origin as a demonstrative and still functions as such in Homer. Like the reflexive and reciprocal pronouns, the demonstratives form a thematic masculine/neuter stem and a long *-ā-* feminine; however, the declension of these pronouns is not at all points identical to that of the corresponding nouns. Such differences are to be seen in the paradigm of the Attic article; note the nominative masculine singular and the nominative/accusative neuter singular:

#### (13) Attic definite article

	<i>Singular</i>		
	<i>Masculine</i>	<i>Feminine</i>	<i>Neuter</i>
<i>Nominative</i>	hō (ὁ)	hē (ἡ)	tó (τό)
<i>Genitive</i>	tū (τοῦ)	tēs (τῆς)	tū (τοῦ)
<i>Dative</i>	tō(i) (τῷ)	tē(i) (τῇ)	tō(i) (τῷ)
<i>Accusative</i>	tón (τόν)	tēn (τήν)	tó (τό)
	<i>Dual</i>		
	<i>Masculine</i>	<i>Feminine</i>	<i>Neuter</i>
<i>Nom./Acc.</i>	tō(τῶ)	tō(τῶ)	tō(τῶ)
<i>Gen./Dat.</i>	toîn (τοῖν)	toîn (τοῖν)	toîn (τοῖν)
	<i>Plural</i>		
	<i>Masculine</i>	<i>Feminine</i>	<i>Neuter</i>
<i>Nominative</i>	hoi (οἱ)	hai (αἱ)	tá (τά)
<i>Genitive</i>	tōn (τῶν)	tōn (τῶν)	tōn (τῶν)
<i>Dative</i>	tois (τοῖς)	tais (ταῖς)	tois (τοῖς)
<i>Accusative</i>	tūs (τούς)	tās (ταῖς)	tá (τά)

The nominative/accusative singular termination *-o* is from PIE *\*-od* and characterizes various demonstrative pronouns.

#### 4.1.3.5 Demonstrative pronouns

Attic has three principal demonstratives, one of which was formed from that early demonstrative which became the article, plus a particle *-de*: *hóde* (ὅδε), *hēde* (ἧδε), *tóde* (τόδε). The demonstrative *hūtos* (οὗτος masc.), *hautē* (αὕτη fem.), *tūto* (τοῦτο neut.) appears to trace its origin to the same source, constructed with a particle *-u-* and a formant *-to-*. Both *hóde* and *hūtos* function as near demonstratives the former is generally used to refer to some entity in nearer proximity to the speaker than the latter. The far demonstrative of Greek is *ekēnos* (ἐκεῖνος masc.), *ekēnē* (ἐκεῖνη fem.), *ekēno* (ἐκεῖνο neut.). Declined like *ekēnos* is the so-called emphatic pronoun *autós* (αὐτός masc.), *autē* (αὐτή fem.), *autó* (αὐτό neut.). As noted above, *autós* is utilized in reflexive constructions and serves in lieu of the third-person personal pronoun in the oblique cases; in addition *autós* is used in conjunction with a noun to express emphasis or sameness.

#### 4.1.3.6 Interrogative/indefinite pronoun

Greek inherited from Proto-Indo-European an interrogative/indefinite pronoun. The interrogative *tís*, *tí* (τίς, τί; “who, which, what”) is tonic, while the segmentally identical indefinite



*tis*, *ti* (“someone, something, etc.”) is enclitic. The interrogative is illustrated in (14); the nasal which appears in most of the oblique cases has been generalized from an inherited accusative singular *\*tin*; as with adjectives of two endings, a gender distinction occurs only in the nominative and accusative:

(14) Attic interrogative pronoun

	<i>Singular</i>	<i>Dual</i>	<i>Plural</i>
<i>Nom. masc./fem.</i>	tís (τίς)	tíne (τίνε)	tínes (τίνες)
<i>Acc. masc./fem.</i>	tína (τίνα)	tíne (τίνε)	tínas (τίνας)
<i>Nom./acc. neut.</i>	tí (τί)	tíne (τίνε)	tína (τίνα)
<i>Genitive</i>	tínos (τίνος)	tínoin (τίνοιν)	tínōn (τίνων)
<i>Dative</i>	tíni (τίνι)	tínoin (τίνοιν)	tísi (τίσι)

A thematic variant is preserved in various dialects, found in Attic in the genitive singular *tû* (τοῦ), from which a dative *tô* (τῷ, Homeric τέω) was created.

#### 4.1.3.7 Relative pronouns

The Greek relative pronoun developed from a Proto-Indo-European stem *\*yo-*, *\*yeh<sub>2</sub>-*; the inflection is that characteristic of *ekēnos* and *autós*: nominative *hós* (ὅς masc.), *hē* (ἡ fem.), *hó* (ὃ neut.); genitive *hû* (οὗ masc./neut.), *hês* (ἧς fem.), and so forth. In addition, Greek possesses an indefinite relative pronoun (“whoever, whatever, etc.”) composed of the relative and indefinite pronouns in combination, with both members inflected: for example, *hóstis* (ὅστις masc.), *hētis* (ἧτις fem.), *hóti* (ὅτι neut.); genitive *hústinos* (οὗστινος masc./neut.), *hêstinos* (ἧστινος fem.). In Attic there also exist variant forms of the genitive and dative, singular and plural, which consist of an uninflected first member *hó-* joined to a thematized second member: for example, *hótû* (ὁτου gen. sg.), *hótô(i)* (ὁτω dat. sg., both masc./neut.).

## 4.2 Verbal morphology

The verbal system of ancient Greek is quite complex. Greek verbs are marked for tense, voice, mood, person, and number. The so-called tenses of Greek require some discussion and are treated in the immediately following paragraphs. Verbs are inflected for three voices (active, middle, and passive), three persons (first, second, and third) and three numbers (singular, dual, and plural). Stems are marked for four moods: indicative (the mood of declaration, factual statement), subjunctive (future-oriented, the mood of will and probability), optative (the mood of wish and potentiality), and imperative (the mood of command).

The Greek verbal system is characterized by seven inflectionally distinct tenses: present, imperfect, aorist, perfect, pluperfect, future, and future perfect. Though these verbal categories have been traditionally labeled “tenses,” they possess independent temporal significance only in the indicative mood. Most fundamentally, the so-called tenses of Greek register aspectual differences.

### 4.2.1 Verbal aspect

At least three different verbal aspects can be identified in Greek: perfective, imperfective, and aoristic. The perfective aspect signifies action which the speaker views as complete, as a packaged whole, and the results of which continue to exist. This is the aspectual significance

of the Greek perfect “tense,” which Gildersleeve (1900:99) aptly and succinctly described in stating that it “looks at both ends of an action.” The pluperfect (which is limited to the indicative mood and so always has temporal significance in independent clauses) denotes complete action producing a result which continued into some referential moment in the past. Similarly the future perfect represents complete action producing a result which will continue into some referential moment in the future.

While the perfective aspect signifies complete action, the imperfective aspect represents action which is continuing, ongoing (and hence not complete). The present stem denotes the imperfective aspect and provides two distinct tenses in the indicative mood: the present and the imperfect. The latter is used of action taking place in the past (and only occurs in the indicative mood), the former of non-past action. Compare imperfect indicative *égraph<sup>h</sup>on*, “I was writing” (ἔγραφον) with perfect indicative *gégraph<sup>h</sup>a*, “I have written” (γέγραφα).

The aoristic aspect is conveyed by the aorist “tense” stem and signifies action which is reported simply as an occurrence, an event, without suggestion as to its completeness or continuance—hence the name of the tense: aorist (ἀόριστος “undefined, unlimited”). Within the indicative mood, the aorist has temporal significance and represents past action.

The aspectual distinctions outlined above are relatively discrete in the indicative mood even though verb-stems conveying particular aspectual notions in this mood also have temporal significance (i.e., actually have tense value). However, this aspectual distinctiveness begins to blur in the case of the present and future indicative. We have seen that the present stem is a carrier of the imperfective aspect and that this is the stem of both the imperfect indicative and present indicative. While the imperfect regularly signifies imperfective aspect and the present indicative often does so, in some instances the present indicative is aspectually aoristic, being used simply to record the occurrence of an action in present time without any notion of continuation. The future indicative is sometimes analyzed as fundamentally signifying aoristic aspect, and perhaps in a majority of instances the future does simply cite the occurrence of an action, in aoristic fashion. However, in other instances the future clearly is used in an imperfective sense to signify continuous action.

#### 4.2.2 Thematic present tense stems

In the parent Indo-European language, various means existed for forming the present tense stem, most of which survive in the grammar of Greek, at least vestigially. For the formation of thematic stems, the Attic dialect utilizes each of the following constructions:

1. The present tense stem can be formed by attaching the thematic vowel to the verb root. In Proto-Indo-European, present tense stems thus formed were of two types – those with accented *e*-grade of the root, and those with *ø*-grade of the root with accent on the thematic suffix (the so-called *tudáti* type). Reflexes of both types occur in Greek: *p<sup>h</sup>érō* (φέρω “I bear,” < \**b<sup>h</sup>ér-e/o-*) is of the former type, and *gráp<sup>h</sup>ō* (γράφω “I write,” < \**g<sup>r</sup>b<sup>h</sup>-é/ó-*) is of the latter. Reduplicated forms of the thematic present tense stem occur in Greek, as they did in Proto-Indo-European; the vowel used in constructing the reduplicated syllable is *-i-*, as in, for example, *tíktō* (τίκτω, “I bring forth,” < \**ti-tk-e/o-*).

2. In Greek, as in its Indo-European ancestor language, a highly productive suffix *-ye/yo-* was used to build verb-stems either by attaching the suffix directly to a verb root (primary suffix) or by adding the suffix to an already existing stem (secondary suffix), most commonly to noun stems (forming denominative verbs), but also to verb-stems (forming deverbative verbs). Primary formations are of two types – one with *e*-grade of the root, the other with *ø*-grade. Though a commonly utilized formant, the occurrence of *-ye/yo-* is opaque because

its addition results in numerous phonological modifications to stems. These modifications give rise to three of the traditionally identified classes of present tense stems: the *tau*-class, the *iota*-class, and the contract verbs.

**2A. The  $\tau$ -class:** The verbs assigned to the *tau*-class are characterized by the presence of the consonantal cluster *-pt-* ( $-\pi\tau-$ ), the reflex of an earlier sequence *\*bilabial + y*; for example, *sképtomai* (σκέπτομαι “I look carefully”) < PG *\*skēp-ye/o-*.

**2B. The  $\iota$ -class:** A heterogeneous set of verbs, the *iota*-class consists of several subtypes:

- (i) Verb-stems formed in *-tt-* ( $-\tau\tau-$ ) in Attic (but *-ss-* ( $-\sigma\sigma-$ ) in many dialects), from the earlier sequences *voiceless {dental, velar, labiovelar} stop + y*; for example, *péttō* (πέττω, “I cook”) < PG *\*pek<sup>w</sup>-ye/o-*.
- (ii) Verb-stems formed in *-zd-* ( $-\zeta-$ ), from the Proto-Greek sequences *voiced {dental, labial, labiovelar} stop + y*; for example, *nízdō* (νίζω “I wash”) from *\*nig<sup>w</sup>-ye/o-*.
- (iii) Verb-stems formed in *-ll-* ( $-\lambda\lambda-$ ) from the Proto-Greek sequence *\*-ly-*; for example, *stéllō* (στέλλω “I set in order”), from *\*stel-ye/o-*.
- (iv) Verb-stems in *-ain-* ( $-\alphaίν-$ ) and *-air-* ( $-\αίρ-$ ), from the earlier sequences *\*-any-*, *\*-amy-*, *\*-ary-*; for example, *bainō* (βαίνω, “I walk, go”), from PG *\*g<sup>w</sup>m<sup>w</sup>-ye/o-*.
- (v) Verb-stems in *-{ǵ, ǵ̃, ũ}n-* and *-{ǵ, ǵ̃, ũ}r-* from the Proto-Greek sequences *\*-{e, i, u}{n, r}y-*; for example, *ténō* (τείνω “I stretch”) from *\*ten-ye/o-*.
- (vi) Verb-stems in *-ai-* ( $-\alphaί-$ ) and *\*-ei-* ( $*-εί-$ ) from Proto-Greek sequences *\*{a, e}w-ye/o-*; for example, *kaiō* (καίω “I light,”) from *\*kaw-ye/o-*. In Attic and most other dialects, verbs ending in *\*-eiō* ( $*-είω$ ) were analogically modified to *-euō* ( $-\epsilonυω$ ), under the influence of nonpresent tenses and corresponding nouns in *-eu-s* ( $-\epsilonυ-ς$ ).

**2C. The contract verbs:** A large class of verbs built with the *-ye/o-* suffix is that of the so-called contract verbs. These are predominantly denominatives, constructed by the addition of *-ye/o-* to a stem ending in a vowel (sometimes as the result of consonant loss). With the loss of intervocalic *-y-*, the resulting adjacent vowels contracted, giving this class its defining characteristic. Contract presents are of three principal types: those in *-aō* ( $-\alphaω$ ), *-eō* ( $-\epsilonω$ ), and *-oō* ( $-\οω$ ).

- (i) Verbs ending in *-aō* ( $-\alphaω$ ). These are primarily denominative verbs formed from noun stems in *-ā-* (first declension nouns); for example, *tīmāō* (τιμάω “I honor”) from PG *\*tīm-a-ye/o-* (cf. *tīmē* [τιμή “honor”]).
- (ii) Verbs ending in *-eō* ( $-\epsilonω$ ). This somewhat heterogeneous class of verbs consists predominantly of denominative verbs made from thematic noun stems (second declension nouns) having *e*-grade of the thematic suffix; for example *oikēō* (οικέω “I inhabit”) from PG *\*woik-e-ye/o-* (cf. *oikos* [οἶκος “house”]). Among other Proto-Indo-European formations which contribute to this set are *s*-stems (e.g., *telēō* [τελέω, “I complete”] from PG *\*tel-es-ye/o-*); iterative/causatives formed with *o*-grade of the root and a suffix *-éye/o-* (e.g., *p<sup>h</sup>obéō* [φοβέω “I strike with fear”] from PG *\*p<sup>h</sup>og<sup>w</sup>-eye/o-*); and stems built with a stative formant *\*-eh<sub>1</sub>-* (e.g., *hrīgēō* [ῥιγέω “I shiver with cold”] from PIE *\*rig-eh<sub>1</sub>-*).
- (iii) Verbs in *-oō* ( $-\οω$ ). While the preceding two types of contract verbs have Proto-Indo-European antecedents, this third type, as a productive category, is original to Greek. Contract verbs of the *-aō* ( $-\alphaω$ ) type furnished the pattern for analogical creation of *-oō* ( $-\οω$ ) denominatives of second declension nominals. Such verbs are commonly factitive in sense; for example *dēlōō* (δηλόω “I make clear”) beside *dēlos* (δῆλος “clear”).

In Attic, contraction of the vowels which were juxtaposed subsequent to the loss of *-y-* adhered to the contraction patterns outlined above, thus producing present (active indicative) paradigms such as those of *timāō* (τιμάω), *oikēō* (οικέω), *dēlōō* (δηλόω):

(15) Attic contract verbs

1st sg.	tīmō	(τῖμῶ)	oikō	(οἰκῶ)	dēlō	(δηλῶ)
2nd sg.	tīmā(i)s	(τῖμαῖς)	oikēs	(οἰκεῖς)	dēlois	(δηλοῖς)
3rd sg.	tīmā(i)	(τῖμαῖ)	oikē	(οἰκεῖ)	dēloi	(δηλοῖ)
1st pl.	tīmōmen	(τῖμῶμεν)	oikūmen	(οἰκοῦμεν)	dēlūmen	(δηλοῦμεν)
2nd pl.	tīmāte	(τῖμαῖτε)	oikēte	(οἰκεῖτε)	dēlūte	(δηλοῦτε)
3rd pl.	tīmōsi	(τῖμῶσι)	oikūsi	(οἰκοῦσι)	dēlūsi	(δηλοῦσι)

3. The Greek thematic suffix *-ske/o-* is descended from PIE *\*-s(ĥ)e/o-*, originally used in the formation of iteratives. Among stem formations found are those with *o*-grade of the root, for example *bá-ske* (βά-σκε “go!”), in some instances with reduplication, as in *di-dá-skō* (δι-δά-σκω “I teach”).

4. A fourth present tense formation is that of the *nu*-class, a set of verb-stems having Proto-Indo-European antecedents, built with various formants containing *n*. In the parent language, nasal presents (originally iterative or inchoative in sense) were formed by insertion of an ablauting infix *\*-ne/n-* before the final consonant of the root; from roots ending in *\*-w* and *\*-h* were abstracted new suffixes *\*-neu-/nu-* and *\*-neh-/nh-*. The parent infix *\*-n-* is preserved in some Greek thematic verbs, but is used in conjunction with a suffix *-ane/o-*, itself derived originally from the Proto-Indo-European nasal infix: for example, *pū-n-t<sup>h</sup>-áno-mai* (πυ-ν-θ-άνο-μαι “I learn”). Still other Greek nasal presents are formed with this same suffix, but without the nasal infix: for example, *auks-ánō* (αὐξ-άνω “I increase”). A third thematic nasal present of Greek is built with a suffix *-ne/o-*: for example, *dák-nō* (δάκ-νω “I bite”); certain stems display a thematicized form of the above-mentioned Proto-Indo-European suffix *\*-nu-*, that is PG *\*-nwe/o-*: for example, Attic *tí-nō* (τί-νω “I pay”); cf. Ionic *tí-nō* (τί-νω), from Proto-Attic-Ionic *\*ti-nwō* (*\*ti-vFw*). From athematic *nu*-stems (see below) developed a thematized formant *-nuo/e-*: for example, *dēk-nūō* (δεῖκ-νύω, “I point out”).

### 4.2.3 Athematic present tense stems

Athematic present tense stems are of four basic types. Two of these involve Greek reflexes of the Proto-Indo-European nasal suffixes abstracted from nasal infixed roots ending in *\*-h* and *\*-w* (noted above):

1. In Attic, athematic present tense stems are formed with the suffix *-nē-* (*-nā-* outside of Attic and Ionic, from PIE *\*-neh<sub>2</sub>-*) or *-na-* (from PIE *\*-nh<sub>2</sub>-*). The former occurs in the active singular, the latter elsewhere: thus, active *pér-nē-mi* (πέρ-νη-μι “I sell”).
2. The Attic suffixes *-nū* (instead of *\*-neu-*, an analogical formation based on *\*-nā-*) and *-nū-* show the same distribution as *-nē-* and *-na-*: for example, *ár-nū-mai* (ἄρ-νυ-μαι “I win,” with short *ü*, being in the middle voice).
3. Root presents are formed by attaching athematic endings directly to the verb root; *e*-grade of the root occurs in the active indicative singular and in the subjunctive, elsewhere the zero-grade: for example, *p<sup>h</sup>ē-mí* (φη-μί “I say,” from PIE *\*b<sup>h</sup>eh<sub>2</sub>-*), *p<sup>h</sup>a-mén* (φα-μέν “we say,” from PIE *\*b<sup>h</sup>h<sub>2</sub>-*).

4. Athematic reduplicated presents are likewise formed utilizing a root stem, but with reduplication of the initial consonant of the root: for example, *dí-dō-mi* (δί-δω-μι “I give,” from PIE *\*-deh<sub>3</sub>-*), *dí-do-men* (δί-δο-μεν “we give,” from PIE *\*-dh<sub>3</sub>-*).

#### 4.2.4 Imperfect tense

As was noted above, the Greek imperfect is built with present tense stems. The imperfect differs from the present by the use of secondary, rather than primary, verb endings (see below), and by the presence of the temporal prefix *e-*, the *augment*. Thus, beside present *p<sup>h</sup>ér-ō* (φέρ-ω, “I bear”), there is formed an imperfect *é-p<sup>h</sup>er-on* (ἐ-φερ-ον, “I was bearing”). The *e*-augment is also used in the formation of the other “secondary” tenses – the aorist and pluperfect – and is also attested in Indo-Iranian, Phrygian, and Armenian. Its use is optional in early Greek, as in Vedic Sanskrit and Avestan, but in time becomes requisite.

#### 4.2.5 Future tense stems

The future tense of Greek is formed with a suffix *-se/o-*, descended from the Proto-Indo-European desiderative suffix *\*-s(y)e/o-*. Greek future stems are of two principal types, the sigmatic (or *s-*) future and the contract future. The former occurs with roots ending in a stop or *-s*, such as *dék-sō* (δείκ-σω “I will show”), and with certain roots (and stems) having a final vowel, in which case the intervocalic *-s-* of *-se/o-* (having been lost by regular sound change) has been restored analogically: *lú-sō* (λύ-σω “I will loose”). Contract futures have their origin in the future stems of Proto-Indo-European roots terminating in the laryngeals *\*h<sub>1</sub>* and *\*h<sub>2</sub>*, such as *ere-sō* (ἐρε-σω “I will speak”) from PIE *\*werh<sub>1</sub>-*; and *ela-sō* (ἐλα-σω “I will drive”) from PIE *\*h<sub>1</sub>elh<sub>2</sub>-*. Regular loss of the Proto-Greek intervocalic *-s-* yields contract verbs in *-eō* and *-aō*: thus *eréō* (ἐρέω) and *eláō* (ἐλάω). The *-eō* contract future was then generalized to almost all Greek verb formants ending in a liquid or a nasal. In Attic this future construction was extended to yet an even wider range of verbs, resulting in the inflection dubbed the “Attic future”: compare (with contraction) Attic *telō* (τελῶ “I will complete”; Ionic *teléō* (τελέω) and Homeric *teléssō* (τελέσσω), from the stem *teles-*.

Greek future tense verbs are not uncommonly inflected with middle endings, for example *pé-somai* (πεί-σομαι “I will suffer”), and in instances show reduplication as well, for example Homeric *de-dék-somai* (δεδέξομαι “I will receive”). Both of these characteristics likely have their origin in the morphology of the Proto-Indo-European desiderative. Though future middle inflection could also be used to convey passive voice, a new future passive construction was built utilizing the aorist passive suffixes (see below) *-t<sup>h</sup>ē-* (*-θη-*, the first future passive) and *-ē-* (*-η-*, the second future passive) to which was attached the future middle *-somai* (*-σομαι*), etc. The construction is little known in Homer but has become common by the period of Classical Attic.

#### 4.2.6 Aorist tense stems

The morphology of the aorist tense is of three basic types: athematic, thematic (which together comprise the traditional *second aorist* category), and sigmatic (*first aorist*), each with Proto-Indo-European ancestry. The class of Greek athematic aorists consists primarily

of nonablauting root verbs (though preserving traces of Proto-Indo-European ablaut): for example, *é-bē-n* (ἐ-βη-ν “I went,” from PIE \**g<sup>w</sup>eh<sub>2</sub>-*). In the case of a small subset of three verbs, the singular athematic aorist is formed with a *-k-* extension of the root, preserving vowel gradation: for example, *é-t<sup>h</sup>ē-k-a* (ἐ-θη-κ-α “I placed,” from PIE \**d<sup>h</sup>eh<sub>1</sub>-*; cf. Latin *fe-c-i* “I made”), displaying so-called *alpha-thematic* morphology (where *-a*, the regular reflex of the first singular ending \**-m*, and which arose regularly in the third plural, is extended through much of the paradigm [thus second singular *é-t<sup>h</sup>ē-k-a-s* (ἐ-θη-κ-α-ς)] – a morphology also characteristic of certain other root aorists).

Thematic aorists are formed predominantly with *ø*-grade of the root, originally accented on the thematic suffix: for example, *é-lip-on* (ἐ-λιπ-ον “I left”). As in Sanskrit, some display reduplication: *ēp-on* (εἶπον “I spoke,” from \**e-we-wk<sup>w</sup>-o-*).

The Greek sigmatic aorist is clearly inherited from Proto-Indo-European, though the origin of its characteristic *-s-* marker is disputed: *é-dēk-sa* (ἐδειξα, “I showed,” from \**e-deik-s-m*). The *-s-a(-)* reflex, regular in the first singular and the third plural, was analogically extended through most of the sigmatic aorist paradigm, i.e., the paradigm has become alpha-thematic.

The passive voice of the aorist could be expressed by middle inflection in early Greek, as in Sanskrit; however, a morphologically distinct aorist passive developed from intransitive aorist actives in *-ē-*, which formant is likely to be traced to a Proto-Indo-European stative suffix \**-eh<sub>1</sub>-/-h<sub>1</sub>-*: thus, *e-k<sup>h</sup>ár-ē-n* (ἐ-χάρ-η-ν, “I rejoiced; I was delighted”); with possible *o*-grade, survives only *he-ál-ō-n* (ἐ-ἄλ-ω-ν “I was taken”). The details of origin are uncertain, but alongside *-ē-* there developed an aorist passive marker *-t<sup>h</sup>ē-* (*second* and *first* aorist passives respectively), perhaps of greater utility for verb bases ending in a vowel, as in *e-lú-t<sup>h</sup>ē-n* (ἐ-λύ-θη-ν “I was released”).

#### 4.2.7 Perfect tense stems

The Greek perfect stem is formed in four principal manners and, in the active indicative, inflected with a set of *perfect endings*, continuing in part those of Proto-Indo-European. The archaic verb *oíd-a* (οἶδ-α “I know” [in origin “I have seen”], from PIE \**wid-*), one perfect type in and of itself, preserves the Proto-Indo-European pattern of *o*-grade of the root in the active singular, *ø*-grade in the plural (*ís-men* [ἴσ-μεν “we know”]), with endings attached to the root.

The so-called *first perfect* of Attic is the most commonly occurring perfect stem; its hallmark is a *-k-* formant which precedes the endings, probably to be linked to the *-k-* of the three athematic aorists mentioned above. Relatively late in origin and a uniquely Greek formation, the *k*-perfect began with verb roots ending in a long vowel, as in, for example, *bé-bē-k-a* (βέ-βη-κ-α “I have gone,” from PIE \**g<sup>w</sup>eh<sub>2</sub>-*). The construction first appeared in the singular, spreading subsequently to the plural and to verb roots of other shapes. As in the preceding example, perfect stems normally show an initial reduplicated syllable (to be found already in the parent Indo-European language), on which see immediately below.

Lacking the *-k-* formant of the first perfect, the Attic *second perfect* is characterized by an absence of root alternation in the active voice. Both this perfect stem and that of the *k*-perfect display alpha-thematic inflection in the active indicative (extended from the first singular and third plural).

The fourth perfect type, the aspirated perfect, is primarily an Attic-Ionic development, one which had its origin in the middle voice. The perfect middle is formed by attaching



endings directly to the verb root. Each of the perfect middle endings begins with a consonant except for the early third plural *-atai* (-αται). Through processes of assimilation, all root-final bilabial stops, whether *-p*, *-b*, or *-p<sup>h</sup>*, are modified by the attached consonant of the ending – and all undergo identical modification, so that the original quality of the bilabial stop (voiceless, voiced, or aspirated) is obscured. Root-final velar stops (*-k*, *-g*, *-k<sup>h</sup>*) are likewise neutralized. For example, from *trép-ō* (τρέπω “I turn”) is generated a middle second plural *té-trap<sup>h</sup>-t<sup>h</sup>e* (τέ-τραφ-θε “you have turned”). In the case of some roots with a final bilabial or velar, the aspirated reflex of the second plural spreads to the third plural, as in *te-tráp<sup>h</sup>-atai* (τε-τράφ-αται, rather than *\*te-tráp-atai*) – a stage which is preserved in Homer. From this starting point, the aspirate is then generalized through the perfect active: thus, *té-trop<sup>h</sup>-a* (τέ-τροφ-α “I have turned”) rather than *\*té-trop-a*.

#### 4.2.7.1 Perfect stem reduplication

Most commonly roots beginning #C<sub>1</sub>(C<sub>2</sub>)V- reduplicate as #C<sub>1</sub>e-C<sub>1</sub>(C<sub>2</sub>)V- (as in *bé-bē-k-a*), though a good number of root-initial #CC- sequences in Attic (e.g., *ps-*, *ks-*, *gn-*) “reduplicate” synchronically by prefixing the vowel *e-* (e.g., *é-psau-k-a* (ἐ-ψαυ-κ-α “I have touched”). The latter reduplication appears to have spread from perfects of verb roots with initial #sC- clusters: by regular sound change *\*se-sC-* yields *#he-sC-* (e.g., *hé-stē-k-a* (ἐ-στη-κ-α “I have stood”) from *\*se-stē-k-a*). The spread of unaspirated *e-* (rather than *he-*) was likely supported by the *e-* augment of the other preterite tenses, aorist and imperfect. Moreover, in some instances of initial #sC- clusters, regular dissimilatory processes of deaspiration produced an *e-* reduplication: thus, *é-lēp<sup>h</sup>-a* (ἐ-ληφ-α “I have taken”) from *\*he-l<sup>h</sup>āp<sup>h</sup>-a*, from *\*se-slāp<sup>h</sup>-a* (certain #s + sonorant clusters perhaps being particularly susceptible to this development).

Proto-Indo-European verb roots beginning with a laryngeal produce Greek perfect stems which synchronically appear to “reduplicate” by lengthening an initial vowel: for example, *ēg-mai* (ἤγ-μαι “I have led,” perfect of *ágō* (ἄγω)), from PG *\*āg-*, from PIE *\*h<sub>2</sub>e-h<sub>2</sub>ǵ-*. This synchronic pattern of producing the perfect stem by lengthening an initial vowel then spread to other vowel-initial roots.

Yet a distinct type of reduplication is exhibited by verb roots which begin with a vowel followed by a sonorant consonant; such roots form a perfect stem by reduplicating the *vowel* + *sonorant* sequence and lengthening the vowel of the root. The exact origin of the structure is a matter of disagreement, though again is likely to lie in the presence of an initial laryngeal: thus, *elēlūth<sup>h</sup>-a* (ἐλ-ήλυθ-α “I have come”), from *\*h<sub>1</sub>le-h<sub>1</sub>lud<sup>h</sup>-*. The pattern is extended to other verb roots beginning with a vowel and becomes especially common in Attic (and Ionic), thus being dubbed *Attic reduplication*.

#### 4.2.7.2 Pluperfect and future perfect

Before leaving perfect morphology, attention needs to be given to the pluperfect and future perfect. Both of these tenses are Greek innovations, not to be found in Proto-Indo-European. The Attic pluperfect is formed with the perfect stem, to which the augment is prefixed if the stem begins with a consonant (such is the general case at least). In the active voice, the Classical Attic pluperfect endings preserve a formant *-e-*, of uncertain origin (attested in Homer), which is followed in turn by the perfect endings in the singular and the secondary endings in the dual and plural (though the third plural appends *-san*). In the singular, Attic contracts the *-e-* and the ensuing morph: thus *\*e-le-lū-k-e-a* (\*ε-λε-λύ-κ-ε-α) yields *e-le-lū-k-ē* (ἐ-λε-λύ-κ-η “I had released”). In both Homer and later Attic, variant pluperfect active morphology occurs. The Attic middle is produced by adding the secondary



middle endings to the pluperfect stem (as described above): *e-le-lŭ-mēn* (ἐ-λε-λŭ-μην “I had ransomed”). Attachment of the sigmatic future *s* + *ending* complex to the perfect stem yields the future perfect.

## 4.2.8 Nonindicative moods

The above elaboration of the tense stems of Attic has focused upon stems as they occur in the indicative, the unmarked mood of Greek by the fifth century. A survey of the morphology of the nonindicative moods follows.

### 4.2.8.1 Subjunctive mood

In Proto-Indo-European the subjunctive is marked by an ablauting suffix *-e/o-* attached to the root. The Greek reflex of this construction, the so-called *short vowel subjunctive*, characterized athematic stems in early Greek and is preserved in Homer and elsewhere: for example *ēd-o-men* (ἐῖδ-ο-μεν “may we know”, from the *e*-grade of *\*wid-*, perfect subjunctive). The attachment of this suffix *-e/o-* to thematic stems yielded, by contraction with the thematic vowel, the Greek *long vowel subjunctive*: *lŭ-ō-men* (λŭ-ω-μεν “may we release,” present subjunctive). Extension of this long vowel morphology to the aforementioned athematic stems results in, for example, Attic *ēd-ō-men* (ἐῖδ-ŭ-μεν).

### 4.2.8.2 Optative mood

The optative mood in the parent Indo-European language was marked by the suffix *\*-yeh<sub>1</sub>/ih<sub>1</sub>-*, originally attached to the root, with *\*-ih<sub>1</sub>-* subsequently also affixed to thematic stems. The former is antecedent to the Greek athematic optative suffix, as in Attic *ēē-n* (earlier *\*eiē-n* [εῖη-ν “I would be”] from *\*h<sub>1</sub>s-yéh<sub>1</sub>-*) and *ē-men* (earlier *\*ei-men* [εῖ-μεν “we would be”] from *\*h<sub>1</sub>s-ih<sub>1</sub>-*). In the case of thematic and alpha-thematic stems, the Attic reflex is *-oi-* and *-ai-* respectively: *p<sup>h</sup>ér-oi-mi* (φέρ-οι-μι “I would bear”), with the primary athematic ending (on primary and secondary endings see below) extended to earlier *p<sup>h</sup>ér-oi-a* (φέρ-οι-α, from *\*b<sup>h</sup>ér-o-ih<sub>1</sub>-m*), though secondary endings are commonly preserved in the optative paradigm; *lŭ-s-ai-mi* (λŭ-σ-αι-μι “I would release,” aorist).

### 4.2.8.3 Imperative mood

A multiplicity of morphological markings characterizes the Greek imperative. As in Proto-Indo-European, the active second singular is formed with the bare stem alone (i.e., without an ending), or by attaching to the stem the particle *-t<sup>h</sup>i* (PIE *\*-d<sup>h</sup>i*); the former construction provides the most frequently occurring Greek imperative, the latter is limited to athematic stems: for example, *p<sup>h</sup>ér-e* (φέρ-ε “carry !”); *í-t<sup>h</sup>i* (ί-θι, “go !”). In addition, the second singular is formed in Attic by attachment of the word-final formants *-s* (the secondary ending) and *-i* (both occurring rarely), as well as *-on*, characterizing alpha-thematic aorist inflection. Proto-Indo-European filled out portions of the imperative paradigm utilizing the injunctive mood (like the indicative in form but with secondary, rather than primary, endings and expressing “timeless truths”). Injunctive morphology is preserved in the Greek second-person plural imperative (*p<sup>h</sup>érete* [φέρετε “carry!”]), looking like the Greek indicative (as does the second dual). The third-person singular imperative is marked by the appending of a particle *-tō* (-τω, PIE *\*-tōd*) to the verb-stem (*p<sup>h</sup>eré-tō*

[φέρέ-τω “let him/her carry”]), from which a third dual marker *-tōn* (-των) was created. The third-person plural takes several forms in Attic, building with the particle *-tō*, such as *p<sup>h</sup>érō-ntōn* (φέρó-ντων “let them carry”), where the bookend nasals are taken over from the primary (*\*-onti > -ūsi*) and secondary (*-on*) third plural endings. Middle imperative endings likewise continue injunctive morphology (e.g., second singular *\*p<sup>h</sup>ere-so* (\*φερε-σο), becoming *p<sup>h</sup>érū* (φέρου)) and display analogical reshaping (e.g., the third singular ending *-st<sup>h</sup>ō* (-σθω), after middle second plural *-st<sup>h</sup>e* (-σθε) and active third singular *-tō* (-τω)).

#### 4.2.9 Verb endings

The verb endings of Greek are traditionally classified as *primary* and *secondary*. In broad terms, the primary endings are used with non-past tenses, the secondary endings with past tenses and the optative mood. Endings are further differentiated as thematic (attached to a thematic stem) and (otherwise) athematic. The following charts illustrate Attic verb endings. In the case of thematic verbs, division is made between the root and thematic suffix; for athematic, division is marked before the ending. In (16) primary active thematic and athematic endings are illustrated by the present active indicative of *p<sup>h</sup>érō* (φέρω “I carry”) and *títēmi* (τίθημι “I place”) respectively:

##### (16) Attic verb endings I: primary active

		<i>Thematic</i>		<i>Athematic</i>	
<i>Singular</i>	1.	p <sup>h</sup> ér-ō	(φέρ-ω)	tít <sup>h</sup> ē -mi	(τίθη-μι)
	2.	p <sup>h</sup> ér-ēs	(φέρ-εις)	tít <sup>h</sup> ē -s	(τίθη-ς)
	3.	p <sup>h</sup> ér-ē	(φέρ-ει)	tít <sup>h</sup> ē -si	(τίθη-σι)
<i>Dual</i>	2.	p <sup>h</sup> ér-eton	(φέρ-ετον)	tít <sup>h</sup> e -ton	(τίθε-τον)
	3.	p <sup>h</sup> ér-eton	(φέρ-ετον)	tít <sup>h</sup> e -ton	(τίθε-τον)
<i>Plural</i>	1.	p <sup>h</sup> ér-omen	(φέρ-ομεν)	tít <sup>h</sup> e -men	(τίθε-μεν)
	2.	p <sup>h</sup> ér-ete	(φέρ-ετε)	tít <sup>h</sup> e -te	(τίθε-τε)
	3.	p <sup>h</sup> ér-ūsi	(φέρ-ουσι)	tít <sup>h</sup> e -āsi	(τίθε-ᾱσι)

Secondary active thematic and athematic endings are illustrated by the imperfect active indicative paradigms of *p<sup>h</sup>érō* and *hístēmi* (ίστημι “I stand”) respectively:

##### (17) Attic verb endings II: secondary active

		<i>Thematic</i>		<i>Athematic</i>	
<i>Singular</i>	1.	ép <sup>h</sup> er-on	(ἔφερ-ον)	hístē -n	(ίστη-ν)
	2.	ép <sup>h</sup> er-es	(ἔφερ-ες)	hístē -s	(ίστη-ς)
	3.	ép <sup>h</sup> er-e	(ἔφερ-ε)	hístē	(ίστη)
<i>Dual</i>	2.	ep <sup>h</sup> ér-eton	(ἔφερ-ετον)	hístā -ton	(ίστά-τον)
	3.	ep <sup>h</sup> ér-étēn	(ἔφερ-έτην)	hístā -tēn	(ίστά-την)
<i>Plural</i>	1.	ep <sup>h</sup> ér-omen	(ἔφερ-ομεν)	hístā -men	(ίστα-μεν)
	2.	ep <sup>h</sup> ér-ete	(ἔφερ-ετε)	hístā -te	(ίστα-τε)
	3.	ép <sup>h</sup> er-on	(ἔφερ-ον)	hístā -san	(ίστα-σαν)

Middle endings are used to express both middle and passive voice, as in Proto-Indo-European; though distinct passive inflection developed for particular tenses, as noted above.

In (18), the primary middle endings are illustrated with the present middle indicative paradigms of thematic *p<sup>h</sup>ērō* and athematic *tít<sup>h</sup>ēmi*:

(18) Attic verb endings III: primary middle

		<i>Thematic</i>		<i>Athematic</i>	
<i>Singular</i>	1.	p <sup>h</sup> ēr-omai	(φέρ-ομαι)	tít <sup>h</sup> e-mai	(τίθε-μαι)
	2.	p <sup>h</sup> ēr-ē(i)	(φέρ-η)	tít <sup>h</sup> e-sai	(τίθε-σαι)
	3.	p <sup>h</sup> ēr-etai	(φέρ-εται)	tít <sup>h</sup> e-tai	(τίθε-ται)
<i>Dual</i>	2.	p <sup>h</sup> ēr-est <sup>h</sup> on	(φέρ-εσθον)	tít <sup>h</sup> e-st <sup>h</sup> on	(τίθε-σθον)
	3.	p <sup>h</sup> ēr-est <sup>h</sup> on	(φέρ-εσθον)	tít <sup>h</sup> e-st <sup>h</sup> on	(τίθε-σθον)
<i>Plural</i>	1.	p <sup>h</sup> ēr-ómet <sup>h</sup> a	(φερ-όμεθα)	tít <sup>h</sup> e-met <sup>h</sup> a	(τίθε-μεθα)
	2.	p <sup>h</sup> ēr-est <sup>h</sup> e	(φέρ-εσθε)	tít <sup>h</sup> e-st <sup>h</sup> e	(τίθε-σθε)
	3.	p <sup>h</sup> ēr-ontai	(φέρ-ονται)	tít <sup>h</sup> e-ntai	(τίθε-νται)

(19) presents the secondary middle endings, utilizing the imperfect middle indicative paradigms of thematic *p<sup>h</sup>ērō* and athematic *tít<sup>h</sup>ēmi*:

(19) Attic verb endings IV: secondary middle

		<i>Thematic</i>		<i>Athematic</i>	
<i>Singular</i>	1.	ep <sup>h</sup> ēr-ómēn	(ἐφερ-όμεν)	etít <sup>h</sup> e-mēn	(ἐτίθε-μεν)
	2.	ep <sup>h</sup> ēr-ū	(ἐφέρ-ου)	etít <sup>h</sup> e-so	(ἐτίθε-σο)
	3.	ep <sup>h</sup> ēr-eto	(ἐφέρ-ετο)	etít <sup>h</sup> e-to	(ἐτίθε-το)
<i>Dual</i>	2.	ep <sup>h</sup> ēr-est <sup>h</sup> on	(ἐφέρ-εσθον)	etít <sup>h</sup> e-st <sup>h</sup> on	(ἐτίθε-σθον)
	3.	ep <sup>h</sup> ēr-ést <sup>h</sup> ēn	(ἐφερ-έσθην)	etít <sup>h</sup> e-st <sup>h</sup> ēn	(ἐτίθε-σθην)
<i>Plural</i>	1.	ep <sup>h</sup> ēr-ómet <sup>h</sup> a	(ἐφερ-όμεθα)	etít <sup>h</sup> e-met <sup>h</sup> a	(ἐτίθε-μεθα)
	2.	ep <sup>h</sup> ēr-est <sup>h</sup> e	(ἐφέρ-εσθε)	etít <sup>h</sup> e-st <sup>h</sup> e	(ἐτίθε-σθε)
	3.	ep <sup>h</sup> ēr-onto	(ἐφέρ-οντο)	etít <sup>h</sup> e-nto	(ἐτίθε-ντο)

In the singular active indicative, Greek preserves an inherited set of perfect endings, seen in the inflection of *oíd-a* (οἶδ-α “I know”). In the other perfect stem-types, represented below by *léloipa* (λέλοιπα “I have left”), the secondary second-person singular -s (-ς) has been invoked to replace inherited -t<sup>h</sup>a (-θα):

(20) Attic verb endings V: perfect

<i>Singular</i>	1.	oíd-a	(οἶδ-α)	léloip-a	(λέλοιπ-α)
	2.	oís-t <sup>h</sup> a	(οἶσ-θα)	léloip-as	(λέλοιπ-ας)
	3.	oíd-e	(οἶδ-ε)	léloip-e	(λέλοιπ-ε)

#### 4.2.10 Infinitives

Attic possesses active, middle, and passive (or middle-passive) infinitives in the present, future, aorist, and perfect tenses. While the origin of the Greek infinitives is a matter of some uncertainty, it appears likely that they developed from verbal nouns inflected for particular cases. Attic thematic stems produce an active infinitive which terminates in -ēn (-ειν, earlier \*-ein), apparently in origin an endingless locative of an *n*-stem (probably from \*-sen, with loss of \*-s- and contraction of the thematic vowel and the initial vowel of the remaining \*-en). Athematic verbs in Attic form the active infinitive in -(e)nai: thus, *tít<sup>h</sup>énai*

(τιθέναι “to place,” present); *dūnai* (δοῦναι “to give,” aorist); *ēdénai* (εἰδέναι “to know,” perfect). The origin of the formant *-(e)nai* is disputed – perhaps arising from a particle *\*-ai* appended to an *n*-stem, perhaps from a locative in *\*-eneh<sub>2</sub>-i*. The active infinitive of sigmatic aorists terminates in *-sai* (-σαι), which perhaps preserves the particle *\*-ai* mentioned above, or again is perhaps to be traced to a locative. The middle infinitives – present, future, aorist, and perfect; thematic and athematic – end in *-st<sup>h</sup>ai* (-σθαί), often conjectured to be related to Indo-Iranian infinitives in *\*-d<sup>h</sup>yai* (as the aorist active *-sai* has been conjectured to be so related).

#### 4.2.11 Participles

Active, middle, and passive (or middle-passive) participles occur in the present, future, aorist, and perfect tenses, and are inflected for all three genders. The active participle of the present, future, and aorist is formed with the suffix *-nt-* (-ντ-). When attached to a thematic stem, the stem bears the *o*-grade of the thematic vowel: for example, the present active participles *p<sup>h</sup>ér-o-nt-os* (φέρ-ο-ντ-ος “carrying,” gen. masc./neut. sg.); *p<sup>h</sup>ér-ōn* (φέρ-ων, nom. masc. sg. from *\*p<sup>h</sup>er-o-nt-s*, with irregular lengthening of the final vowel); *p<sup>h</sup>ér-ūs-a* (φέρ-ουσ-α, nom. fem. sg. from *\*p<sup>h</sup>er-o-nt-ya*). As the preceding examples illustrate, the masculine and neuter active participles have the expected consonant-stem inflection; feminines follow the inflection of (first declension) nouns of the *\*-ih<sub>2</sub>/yeh<sub>2</sub>-* type. Sigmatic aorists form the present active participle with a formant *-ant-* rather than *\*-at-* (as expected by regular sound change, from *\*-s-rt-*) under the influence of thematic stems: *lū́-s-ant-os* (λῦ-σ-αντ-ος, “releasing,” gen. masc./neut. sg.). The perfect active participle is formed with a suffix *\*-wos-* (prior to the disappearance of Attic *w*) in the masculine and neuter, zero-grade *-us-* in the feminine: *ēd-ōs* (εἰδ-ώς “knowing,” nom. masc. sg., from *\*weid-wōs*); *ēd-uia* (εἰδ-υῖα, nom. fem. sg., from *\*wid-us-ih<sub>2</sub>*). Middle participles are formed utilizing a thematic suffix *-meno-*.

#### 4.2.12 Verbal adjectives

In various daughter languages, including Greek, verbal adjectives developed from the Proto-Indo-European stem formant consisting of *φ*-grade of the root plus the suffix *\*-tó-*. While the original sense was passive, the Greek verbal adjective came to express active notions as well, and lacked the root constraint of the parent language: *kliu-tó-s* (κλυ-τό-ς, “heard of, famous”); *p<sup>h</sup>ilē-tó-s* (φιλη-τό-ς “to be loved”); *pis-tó-s* (πισ-τό-ς, “to be believed; believing”). This is perhaps the same suffix used in the formation of ordinals and superlatives. Adjectives indicating necessity are formed with a suffix *-téo-*, of disputed origin though frequently linked to Sanskrit *-tavya-*: *grap-téo-s* (γραπ-τέο-ς “must be written”).

### 4.3 Adverbs

Attic, like other Greek dialects, productively forms adverbs from adjectives utilizing a formant *-ōs*: *kakós* (κακός “bad”), *kakōs* (κακώς “ill”); *hēdús* (ἡδύς “sweet”), *hēdeōs* (ἡδεώς “sweetly”). For the comparative adverb, the accusative neuter singular of the comparative adjective is used, and for the superlative adverb, the accusative neuter plural of the superlative

adjective. In addition, Greek possesses many adverbs which are simply lexicalized nouns of various case forms (some no longer productive in Attic): for example, nominative (apparently) *hápaks* (ἅπαξ “once”); accusative *tēmeron* (τήμερον “today”); dative *koinē(i)* (κοινῇ “in common”); locative *oíkoι* (οἴκοι “at home”); instrumental *lát<sup>h</sup>rā* (λάθρᾱ “secretly”). Similarly some adverbs are lexicalized unverbated prepositional phrases: *ek-podōn* (ἐκποδῶν, “out of the way,” literally “away from the feet”). Numerous suffixes, of uncertain origin, are also used for adverb formation, such as *-t<sup>h</sup>en*, with ablative sense, in, for example, *én-t<sup>h</sup>en* (ἐν-θεν “thence”).

## 4.4 Compounds

Nominal compounding is a common phenomenon in Greek as it was in the parent Indo-European language. In Greek, nominal compounds are most frequently composed of two elements, infrequently more than two, and show inflection of the last member only. While Attic displays a wide variety of compound types, these can be conveniently, if not exhaustively, classified as endocentric and exocentric, invoking categories from traditional Indo-European grammar. The former can be subdivided into copulative and determinative; the principal representative of the exocentric type is the possessive compound.

Copulative compounds coordinate two (or more) members: for example, *núk<sup>h</sup>t<sup>h</sup>-ēmeron* (νυχθ-ήμερον “night and day”). Determinatives may be descriptive (the first member modifies the second adjectivally or adverbially) or dependent (the first member is grammatically dependent on the second, or occasionally vice versa): for example, *akró-polis* (ἀκρό-πολις “upper city”) and *Diós-kūroi* (Διός-κουροι “sons (boys) of Zeus”) respectively. Possessive compounds are similar in sense to determinatives, but are used adjectivally to indicate possession of a trait or quality: *argūrō-toksos* (ἀργυρό-τοξος “having a silver bow”). At times in Greek, as commonly in Sanskrit, possessive compounds are derived from determinatives by a shift in accent.

## 4.5 Numerals

Of the Attic cardinals 1 through 10, only the first four are declined, as in Proto-Indo-European:

### (21) The Attic cardinals

1	hēs, mía, hén	εἷς, μία, ἓν	(masc. fem., neut.)
2	duō	δύο	(declined as a dual)
3	trēs, tría	τρεις, τρία	(masc./fem., neut.)
4	téttares, téttara	τέτταρες, τέτταρα	(masc./fem., neut.)
5	pēnte	πέντε	
6	héks	ἕξ	
7	heptá	ἑπτὰ	
8	oktō	ὀκτώ	
9	ennéa	ἐννέα	
10	déka	δέκα	

From 11 through 199, the cardinals are indeclinable. Between 11 and 19, these are composed

of compounds with *déka*: for example, *dō-deka* (δῶ-δεκα “12”). The decades 20 to 90, composed of a form of the appropriate monad and a reflex of Proto-Indo-European *\*dēkm t-* or *o-grade \*dēkont-* (cf. *\*dēkm(t)* “ten”) are as follows:

(22) The Attic decades

20	ἑkosi	εἴκοσι
30	triákonta	τριάκοντα
40	tettarákonta	τετταράκοντα
50	pentēkonta	πεντήκοντα
60	heksēkonta	ἑξήκοντα
70	hebdomēkonta	ἑβδομήκοντα
80	ogdoēkonta	ὀγδοήκοντα
90	enenēkonta	ἐνενήκοντα

Hundreds are expressed by *-katon* (used for 100, PIE *\*k̑mtom*) and its inflected Attic derivative *-kósioi* preceded by a form of the appropriate monad; for example:

(23) The Attic decades

100	hekatón	ἑκατόν
200	diákósioi	διακόσιοι
300	triákósioi	τριακόσιοι
400	tetrakósioi	τετρακόσιοι

One thousand is *k<sup>h</sup>ílioi* (χίλιοι) and 10,000 is *mūrioí* (μύριοι).

Compound numbers are expressed in various ways. Where *x* is the smaller number and *Y* the larger, the typical formulae are: (i) *x kai Y* (where *kai* (καί) is the conjunction “and”); (ii) *Y (kai) x*. In the second, *kai* is optional; compare English “three and twenty blackbirds” and “twenty-three.” If the last digit of the compound is eight or nine the common practice is to express the number as the next highest decade minus two or one respectively: for example, *duoîn déontes pentēkonta* (δυσὶν δέοντες πενήκοντα “forty-eight,” literally “fifty lacking two”).

Ordinals are generally derived from the corresponding cardinals utilizing the suffix *-to-*. The ordinals “first” and “second” are exceptions regarding the cardinal base, and “seventh” and “eighth” show variation of the suffix. All ordinals are declined.

(24) The Attic ordinals

first	prōtos	πρῶτος
second	deúteros	δεύτερος
third	trítos	τρίτος
fourth	tétartos	τέταρτος
fifth	pémp tos	πέμπτος
sixth	héktos	ἕκτος
seventh	hébdomos	ἑβδόμος
eighth	ógdoos	ὀγδοος
ninth	énatos	ἐνατος
tenth	dékatos	δέκατος

## 5. SYNTAX

### 5.1 Word order

As is the case with many early Indo-European languages possessing well-developed systems of nominal and verbal morphological marking, the word order of Greek is identified as free. That is to say, the order of sentence constituents is highly variable, though not all possible permutations can actually occur. Various investigators conducting statistical examinations of Greek texts have noted a tendency in Classical Greek for the subject to precede the verb (SV) and likewise for the object to precede the verb (OV). The result is that SV and OV have been identified as “unmarked” orders, and variation in these and other basic constituent orders has been commonly attributed to stylistic, pragmatic, and even prosodic factors.

### 5.2 Clitics

In the preceding sections allusion has been made to clitic elements; Classical Attic, like the other dialects of ancient Greek, possessed numerous such clitics, divided into the two broad classes of *enclitics* and *proclitics*. Traditionally these are analyzed as unaccented (atonic) lexemes which form an accent unit with the preceding (enclitics) or following (proclitics) tonic form. Among the enclitics are included the oblique cases of the singular personal pronouns, the indefinite pronoun and adverbs, and various grammatical particles (of which Greek has many, both adverbial and conjunctive). Under the heading of proclitic have been listed monosyllabic forms of the article which begin with a vowel, and certain prepositions and conjunctions. It should be noted that a proclitic class was not a notion treated by the Greek grammarians and that the breadth of its membership and its prosodic nature have been debated by modern scholars (see Devine and Stephens 1994:356–361). The occurrence of clitics in the parent Indo-European language and their placement in Wackernagel’s position (after the first accented word of the sentence) is a well-established phenomenon, preserved particularly clearly in Anatolian (see Chs. 18–23).

### 5.3 Post- and prepositives

Classes of Greek lexemes can be further distinguished as *postpositive* and *prepositive*. Enclitics constitute roughly a large subset of the former and proclitics of the latter (for enumeration of class membership see Dover 1960:12–14). As formulated by Dover, postpositives (*q*) are generally not permitted in clause-initial position, while prepositives (*p*) are normally excluded from clause-final position. Words which are not so limited – most of the words of the language – can be labeled *mobile* (*M*), again following Dover (1960:12). In early Greek, postpositives tend to aggregate after the first mobile word of the sentence, but over time this tendency is progressively eroded. A familiar sentence-initial syntactic pattern of Classical Attic is  $\#pq_1Mq_2$  where  $q_1$  can only be a connecting particle,  $q_2$  can be any other postpositive (Dover 1960:16; Dover attributes the emergence of this pattern to an interaction of factors, including a partial coalescence of prepositives and mobile forms).

### 5.4 Coordination

Greek freely allows coordination and subordination. Coordination is commonly effected utilizing the enclitic conjunction *te* (τε) and the tonic *kai* (καί). Both can be used to conjoin



individual words, clauses and sentences and are frequently used in combination with one another and with still other conjunctions. The conjunctive particle *dé* (δέ) is frequently used to introduce clauses and occurs in second position. Often a clause so introduced is coupled with a second clause marked by the particle *mén* (μέν), the two existing in a contrastive relationship (“on the one hand”... “on the other hand”).

## 5.5 Subordination

With regard to syntax, the subordinate clauses of Greek are of three basic types, distinguished by the verb form – finite, infinitival, and participial. Within each type structural variation occurs. Subordinate clauses frequently contain a finite verb and are introduced by a complementizer, of which the language possesses several. For example, the complementizers *hína* (ἵνα), *hōs* (ὥς), and *hópōs* (ὥπως) are used to mark subordinate clauses containing a finite verb in the subjunctive or optative mood – subordinate constructions traditionally identified as *purpose* (or *final*) *clauses*. If the verb of the matrix clause is inflected in a so-called primary tense (present, future, perfect, future perfect), the subjunctive is used in the embedded clause; if the tense of the matrix verb is “secondary” (imperfect, aorist, pluperfect), the subordinate verb appears in the optative (or subjunctive) mood.

- (25) παιδεύω τὸ παιδίον ἵνα ἐκμάθῃ  
 paideúō tò paidíon hína ekmát<sup>h</sup>ē(i)  
 “I teach (PRESENT) the child in order that he may learn (SUBJUNCTIVE)”

After a verb expressing the notion of SAYING, a complement clause commonly is introduced by *hóti* (ὅτι) or *hōs* (ὥς); if the tense of matrix verb is primary, the mood of the subordinate verb is unaltered (i.e., the mood is retained which would have been present had the subordinate clause been independent), but may be changed to the optative if the matrix verb tense is secondary.

- (26) ἔλεξεν ὅτι Σωκράτης παιδεύει τὸ παιδίον  
 éleksen hótī Sōkrátēs paideúoi tò paidíon  
 “(S)he said (AORIST) that Socrates was teaching (OPTATIVE) the child”

The second fundamental type of subordinate clause construction is that in which the verb is infinitival. For example, this syntax is typical of clauses embedded in matrix sentences containing a verb of THINKING or, in some cases, a verb of SAYING. If the subject of the embedded clause is identical to that of the matrix clause, it is not expressed; if the two are different, the embedded subject appears in the accusative case.

- (27) νομίζει Σωκράτην παιδεύειν τὸ παιδίον  
 nomízē Sōkrátēn paideúēn tò paidíon  
 “(S)he thinks that Socrates<sub>ACCUSATIVE</sub> is teaching (INFINITIVE) the child”

Third and less commonly, a subordinate clause may be constructed with a participial verb. Certain verbs expressing PERCEPTION and KNOWING take subordinate clauses of this construction. If both matrix and embedded clause have the same subject, the participle stands in the nominative case. If the subjects are different, the subordinate subject and participle are inflected as accusatives (or, in certain instances, some other oblique case).

- (28) ἀκούω Σωκράτην παιδεύοντα τὸ παιδίον  
 akouō Sōkrátēn paideúonta tò paidíon  
 “I hear that Socrates (ACCUSATIVE) is teaching (PARTICIPLE) the child”

## 5.6 Conditional clauses

Attic possesses an elaborate syntactico-semantic system of conditional clauses. No fewer than eight distinct patterns can be identified, varying structurally by the verb tense and/or mood found in the protasis and in the apodosis (and the presence or absence of the particle *án* (ὅν)). The various conditional constructions differ in nuance by the partial intersection of three semantic factors: temporality (past, present, future); likelihood of fulfillment; and generality (or specificity) of the event to which reference is made. For example, the imperfect indicative in both protasis and apodosis signals a *present unreal* (or *contrary to fact*) conditional – a conditional relation which could, but does not in fact, exist:

- (29) εἰ Σωκράτης ἐπαίδευε τὸ παιδίον, ἂν ἐδύνατο γράφειν  
 ē Sōkrátēs epáideue tò paidíon, àn ed únato gráphēn  
 “If Socrates taught the child, he would be able to write (but Socrates does not teach the child)”

## 5.7 Agreement

Agreement is expressed between: (i) subject and verb in person and number; (ii) adjective and noun in case, gender, and number; (iii) a word and its appositive in case; and (iv) a relative pronoun and its antecedent in gender and number. The case of a relative pronoun is determined by its syntactic position in the relative clause; however, the relative pronoun frequently is inflected to agree with the case of its antecedent (case attraction). A notable exception to regular subject/verb agreement is of Proto-Indo-European origin: neuter plural subjects (collectives in origin) take singular subjects.

## 5.8 Long-distance anaphora

In the classical Attic dialect of the fifth century BC, there exists a well-developed system of reflexive pronouns. As described above, a distinct reflexive formant occurs for each of the three persons of the singular and plural, though the third singular form has begun to be utilized in lieu of the existing third plural. A reflexive pronoun is employed when it and its antecedent occur within the same clause. In the case of the third person, however, the reflexive can also appear in a subordinate clause when its antecedent is in a dominating clause. The *h(e)aut-* third-person form is sometimes utilized in this “long-distance” fashion. As discussed earlier (see §4.1.3.2), there is a morphologically distinct, so-called “indirect” reflexive which also functions in this manner – in origin the early personal pronouns of the third person, familiar from Homer.

# 6. LEXICON

The lexicon of a language is a mirror of its speakers’ culture and a footprint of its history. Ancient Greek is one of the grammatically most conservative of the attested Indo-European languages and not surprisingly preserves, at least within its core lexicon, many words of Proto-Indo-European pedigree (a number of which have been encountered above). These include kinship terms such as *patḗr* (πατήρ “father”), *mḗtēr* (μήτηρ “mother”), *tḥügátēr* (θυγάτηρ “daughter”); names of domesticated and wild animals, for example *híppos* (ἵππος “horse”), *taûros* (ταῦρος “bull”), *hús* (ὕς “pig”), *ópḥis* (ὄφις “snake”), *mûs* (μῦς “mouse”); names of body parts such as *kardía* (καρδία “heart”), *hēpar* (ἥπαρ “liver”), *ompḥalós* (ὀμφαλός “navel”).

The reader may consult the discussion of verb morphology in §4.2 for numerous examples of inherited Proto-Indo-European verb roots.

There are a great many words of the Greek language, however, which have no clear Indo-European etymology. When the Greeks arrived in the Balkan peninsula late in the third millennium, they came to a place which had an indigenous population, and from the language or, more likely, languages of this population the Greeks certainly acquired a part of their lexicon. Some scholars have attributed a subset of these borrowings to an unattested, broadly distributed “Mediterranean” or “Aegean” substratum language, as superficially similar forms crop up in numerous of the attested languages of the ancient Mediterranean. Under this rubric have been listed words such as *erébinthos* (ἐρέβινθος “chick-pea”), *mint<sup>h</sup>a* (μίνθα “mint”), *sûkon* (σῦκον “fig”), *hródon* (ρόδον “rose”), *hiúakint<sup>h</sup>os* (ὑάκινθος “hyacinth”), *mólubdos* (μόλυβδος “lead”).

Some scholars have held out the possibility that one or more Indo-European languages were already spoken in the Balkan peninsula at the time the Greeks arrived and that these languages similarly provided loans to the Greek lexicon. Thus, a so-called Pelasgian element of the Greek vocabulary has been proposed, with forms cited such as *tûmbos* (τύμβος “grave”) beside *táp<sup>h</sup>os* (τάφος, the regular Greek reflex of PIE \**d<sup>h</sup>mb<sup>h</sup>os*) and *pûrgos* (πύργος “tower”), compare Germanic \**burgs* (“hill-fort”). The Pelasgian hypothesis has not been widely received without reservation.

Among the attested languages of antiquity from which Greek unquestionably acquired vocabulary, Semitic occupies a prominent position. Securely identified Semitic loanwords include *déltos* (δέλτος “writing tablet”), *k<sup>h</sup>itōn* (χιτών a garment; of Sumerian origin), *k<sup>h</sup>rūsós* (χρῦσός “gold”), *krókos* (κρόκος “saffron,” though not of Semitic origin; perhaps originally from an Anatolian place name), *mált<sup>h</sup>ē* (μάλθη “wax”), and *sēsamon* (σήσαμον “sesame seed”). Hittite loans include *kúanos* (κύανος “dark blue enamel”; though itself likely of non-Hittite origin). Iranian appears to provide, among other forms, *kaunákēs* (καυνάκης a woolen robe).

## 7. READING LIST

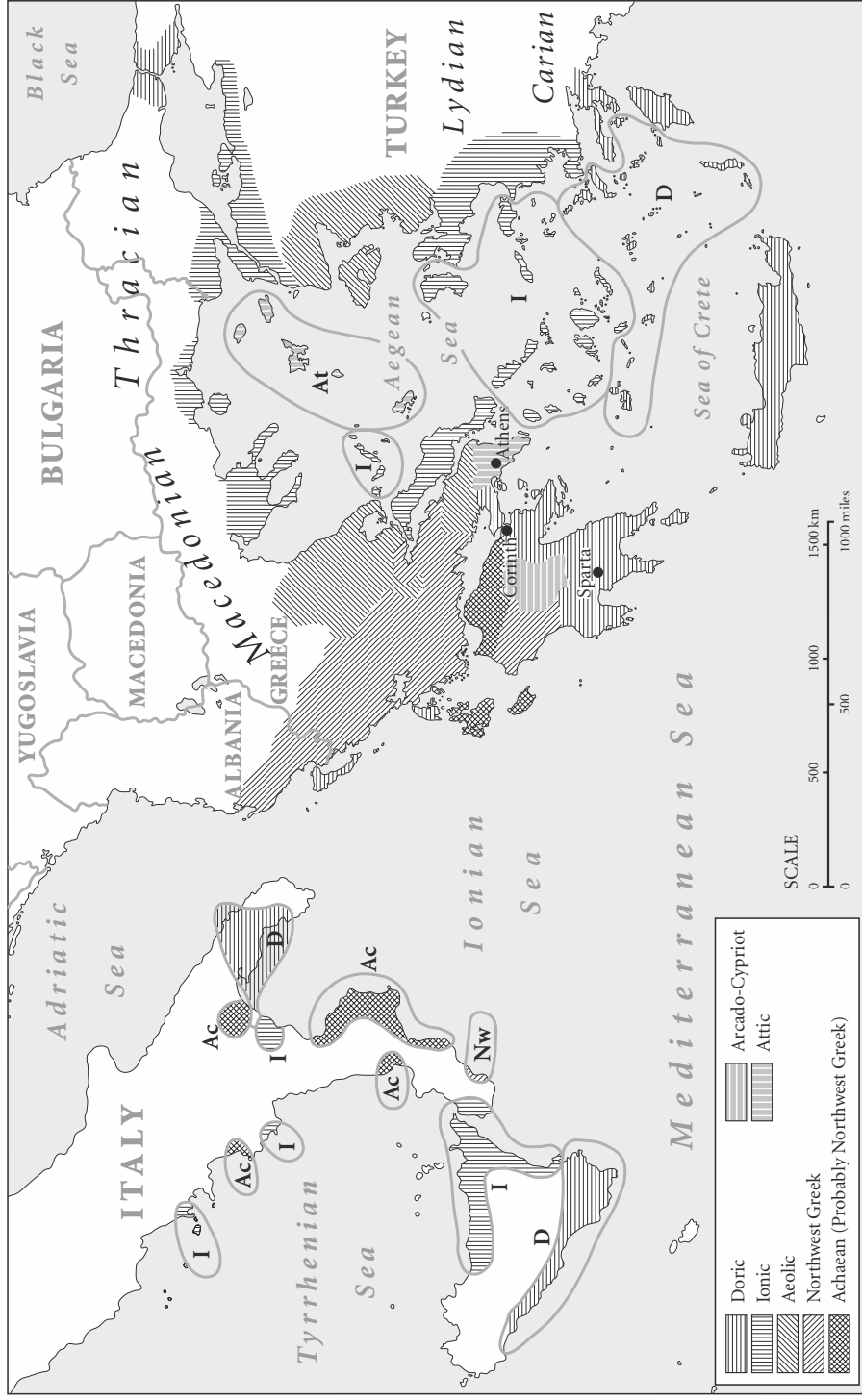
For a traditional grammatical treatment of classical Greek, Smyth 1956 is a standard and comprehensive work. Excellent linguistic overviews of Greek are to be found in Buck 1933 (updated and modified in Sihler 1995), Palmer 1980 and Rix 1976. Jeffery 1990 provides a valuable and detailed discussion of the Greek alphabets; on the alphabet and especially its origin, see also Woodard 1997. For phonetics and phonology, see the excellent treatments in Allen 1987, Devine and Stephens 1994 and Lejeune 1982. Chantraine 1984 provides a valuable survey of Greek morphology. Dover 1960 offers an insightful analysis of Greek word order. For the Greek lexicon, various etymological dictionaries are available; see particularly Chantraine 1968ff. An excellent overview of the development of Greek beyond the period examined herein is to be found in Browning 1983.

## Bibliography

- Allen, W. 1973. *Accent and Rhythm*. Cambridge: Cambridge University Press.  
 ———. 1987. *Vox Graeca* (3rd edition). Cambridge: Cambridge University Press.  
 Browning, R. 1983. *Medieval and Modern Greek* (2nd edition). Cambridge: Cambridge University Press.  
 Buck, C. 1933. *Comparative Grammar of Greek and Latin*. Chicago: University of Chicago Press.

- Buck, C. and W. Petersen. 1945. *A Reverse Index of Greek Nouns and Adjectives*. Chicago: University of Chicago Press.
- Chantraine, P. 1968ff. *Dictionnaire étymologique de la langue grecque: Histoire des mots*. Paris: Klincksieck.
- . 1984. *Morphologie historique du grec* (2nd edition). Paris: Klincksieck.
- Denniston, J. 1954. *The Greek Particles* (revised edition by K. Dover). Oxford: Oxford University Press.
- Devine, A. and L. Stephens. 1994. *The Prosody of Greek Speech*. Oxford: Oxford University Press.
- Dover, K. 1960. *Greek Word Order*. Cambridge: Cambridge University Press.
- Gildersleeve, B. 1900. *Greek Syntax*. Amsterdam: Gieben.
- Jeffery, L. 1990. *The Local Scripts of Archaic Greece* (revised edition). Supplement by A. Johnston. Oxford: Oxford University Press.
- Kirchhoff, A. 1887. *Studien zur Geschichte des griechischen Alphabets*. Gütersloh: Bertelsmann.
- Lejeune, M. 1982. *Phonétique historique du mycénien et du grec ancien*. Paris: Klincksieck.
- Masson, O. 1983. *Les inscriptions chypriotes syllabiques*. Paris: Edition E. de Boccard.
- Meier-Brügger, M. 1992. *Griechische Sprachwissenschaft* (2 vols.) Berlin: DeGruyter.
- Meillet, A. 1965. *Aperçu d'une histoire de la langue grecque* (7th edition). Paris: Klincksieck.
- Meillet, A. and J. Vendryes. 1979. *Traité de grammaire comparée des langues classiques* (5th edition). Paris: Champion.
- Nagy, G. 1974. *Comparative Studies in Greek and Indic Meter*. Cambridge, MA: Harvard University Press.
- Palmer, L. 1980. *The Greek Language*. Atlantic Highlands, NJ: Humanities Press.
- Rix, H. 1976. *Historische Grammatik des Griechischen*. Darmstadt: Wissenschaftliche Buchgesellschaft.
- Schwyzler, E. 1939ff. *Griechische Grammatik* (4 vols.). Munich: Beck.
- Sihler, A. 1995. *New Comparative Grammar of Greek and Latin*. Oxford: Oxford University Press.
- Smyth, H. 1956. *Greek Grammar* (revised edition by G. Messing). Cambridge, MA: Harvard University Press.
- Threatte, L. 1980. *The Grammar of Attic Inscriptions*, vol. I. Berlin: De Gruyter.
- Vendryes, J. 1916. *Traité d'accentuation grecque*. Paris: Klincksieck.
- Ventris, M. and J. Chadwick. 1973. *Documents in Mycenaean Greek* (2nd edition). Cambridge: Cambridge University Press.
- Woodard, R. 1997. *Greek Writing from Knossos to Homer*. Oxford: Oxford University Press.





# Greek dialects

ROGER D. WOODARD

## 1. HISTORICAL AND CULTURAL CONTEXTS

### 1.1 The dialects of the first millennium BC

The ancient Greeks themselves traced their ethnic and linguistic heritage to Hellen, the eponym of both Greece (*Hellas*, Ἑλλάς) and the Greeks (*Hellenes*, Ἕλληνες). Hellen was said to be a son of Deucalion, a son of Prometheus and survivor of the great primeval flood of Greek tradition. The self-recognized diversity of Greek culture and language was attributed to descent from Hellen's three sons, Dorus, Xanthus, and Aeolus, being the alleged progenitors of the Dorian, Ionian, and Aeolian Greeks respectively.

Modern scholars recognize a dialectal distinction which fundamentally parallels this ancient tripartite division. Prior to Michael Ventris' decipherment of the Linear B tablets of the Mycenaean Greeks (see §2.1) in 1952 (see Ventris and Chadwick 1973:3–27), the ancient Greek dialects (i.e., of the first millennium BC) were broadly separated into (i) Attic-Ionic; (ii) Arcado-Cypriot; (iii) Aeolic; (iv) Doric; and (v) Northwest Greek. Each of these, in turn, shows some lesser or greater degree of internal differentiation.

#### 1.1.1 Attic-Ionic

Attic is the dialect of Athens and the surrounding region of Attica (and is the focus of the linguistic description presented in Ch. 24). Its closely related sister dialect of Ionic is divided into three subdivisions, East, Central, and West Ionic. East Ionic is comprised of the dialects of the Ionian cities of western Anatolia (Hallicarnassus, Miletus, Smyrna, etc.) along with those of neighboring islands (such as Samos and Chios), and the Ionic of areas surrounding the Hellespont and of coastal regions along the Thracian Sea. Central Ionic is the language of the Ionian Cycladic Islands such as Naxos and Paros; while West Ionic was spoken in Euboea. The Ionic dialect contributes a significant portion to the literary language of Greek epic and is the dialect of the fifth-century historian Herodotus and the physician Hippocrates of Cos (where the native dialect was Doric), among still other Greek writers.

#### 1.1.2 Arcado-Cypriot

Arcado-Cypriot is the dialectal subdivision to which belong the geographically far-flung but remarkably homogeneous dialects of the island of Cyprus (see §2.2) and the mountainous region of Arcadia in the Peloponnese. Their similarity is chiefly the result of the preservation of archaic features of a common ancestor dialect in two linguistically isolated areas.



Somewhat similar is the dialect of Pamphylia in southern Anatolia. Pamphylian, however, also shows similarities to West Greek, and its proper position within the network of the Greek dialects is uncertain.

### 1.1.3 Aeolic

The Aeolic dialect is further divided into Lesbian (Anatolian Aeolic), Thessalian, and Boeotian (Balkan Aeolic). Lesbian is the dialect of the northwest Anatolian coast (lying northward of the East Ionic regions) and associated islands, chief of which is Lesbos. The poets Sappho and Alcaeus of Lesbos composed in a literary form of their native dialect. Boeotian and Thessalian are the dialects of the regions of Boeotia and Thessaly in northeast Greece. The latter has itself two subdialects, those of Pelasgiotis (spoken in cities such as Larisa) and Thessaliotis (known from Pharsalus and elsewhere). Like Ionic, Aeolic provides linguistic components to the literary dialect of Greek epic.

### 1.1.4 Doric

Doric is the dialect which is attested in the greatest variety of distinct local forms. Rhodian is the dialect of the island of Rhodes and of neighboring smaller islands and coastal towns of southwest Anatolia (south of the Ionic-speaking region). A distinct Doric form is found on the islands of Cos and Calymna (northwest of Rhodes), and another on the Cycladic islands of Thera and Melos. The dialect of Crete is Doric, and itself shows internal variation. On the Balkan Peninsula, several Doric dialects are identified: Megarian, Argolic, Corinthian, Messanian, and Laconian. In literary usage, Doric figures prominently in the language of Greek choral lyric.

### 1.1.5 Northwest Greek

The remaining dialect group is that of Northwest Greek, being clearly a close relative of Doric. The principal Northwest Greek dialects are three. Phocian is the dialect from the area of Delphi; East and West Locrian were spoken in Locris (along the northwest coast of the Gulf of Corinth); Elean is known chiefly from the city of Olympia (in the northwest Peloponnese). In addition, a Northwest Greek Koine is known – fundamentally a hybrid dialect of Attic and certain distinctively Northwest Greek (and Doric) linguistic features. Its use is chiefly associated with the Aetolian Confederacy (Rome's Greek allies against the Macedonians; later subjugated by Rome) and dates to the second and third centuries BC.

## 1.2 The dialects of the second millennium BC

With the decipherment of Linear B and the translation of the documents from Pylos, Knossos, Thebes, and still other Mycenaean sites, a Greek dialect came to light – a dialect of the second millennium BC – not identical to any of those known from the later, alphabetic period (described in §1.1). Moreover, continued study of the Linear B documents led to the realization that they preserve not one, but two different dialectal forms. These distinctions were first teased apart in print by Risch (1966), who assigned to them the names *Normal Mycenaean* (*mycénien normal*; the more commonly attested type) and *Special Mycenaean* (*mycénien spécial*). Further analysis of the variation was provided by Nagy 1968 and Woodard 1986. Herein the two dialects will be referred to as *Mycenaean I* (Normal) and *Mycenaean II* (Special).

The two Mycenaean dialects are distinguished by four morphological and phonological isoglosses. On the one hand, the following features characterize Mycenaean I:

1. The athematic dative singular ending is *-ei*
2. The Proto-Indo-European syllabic nasals *\*m̥* and *\*n̥* develop into the mid vowel *o* in the vicinity of a labial consonant
3. The mid vowel *\*e* is raised to *i* in the vicinity of a labial consonant
4. Before the high vowel *\*i*, the voiceless dental stop *\*t* becomes the fricative *s*

In contrast, Mycenaean II shows the following traits:

5. The athematic dative singular ending is *-i*
6. The syllabic nasals develop into the low vowel *a*
7. The mid vowel *\*e* is preserved in the vicinity of a labial consonant
8. The inherited sequence *\*ti* is preserved

Almost ironically, of the four isoglosses which are characteristic of the more commonly attested dialect, Mycenaean I, only a single one (4: the shift of *\*ti* to *si*) is attested among the known post-Mycenaean dialects (see §3.4.3).

Of the first-millennium dialects, it is Arcado-Cypriot to which Mycenaean Greek is most closely related. The Mycenaean language as attested in the Linear B tablets does not appear, however, to be the direct precursor of Arcado-Cypriot. More than that, a comparison of the Mycenaean dialects with those of the alphabetic period suggests a linguistic heterogeneity in the second millennium which goes considerably beyond the dialectal variation preserved in the Linear B tablets (see Cowgill 1966).

### 1.3 Dialect interrelations

Prior to 1955, the ancient Greek dialects were conventionally divided into two major groups: *West Greek*, composed of Doric and Northwest Greek; and *East Greek*, consisting of Aeolic, Arcado-Cypriot, and Attic-Ionic. With the decipherment of the Linear B tablets, Mycenaean was folded into East Greek. A serious challenge to this analysis, however, was put forward by Ernst Risch (1955) utilizing various linguistic methodologies (such as relative chronology of language change and dialect geography) and building upon then recently published work by Walter Porzig (1954).

Risch argued that the proper bifurcation of Greek dialects is one of North versus South. A *North Greek* phylum consists of Doric, Northwest Greek, and Aeolic; *South Greek* of Mycenaean, Arcado-Cypriot, and Attic-Ionic. It is the swing position of Aeolic, obviously, which distinguishes Risch's classification from the old East versus West analysis. The Balkan Aeolic dialects, Thessalian and Boeotian, show similarities to West Greek – similarities which had been attributed to West Greek influence in the former scheme. Risch, however, contends cogently that the traits which Thessalian and Boeotian share with West Greek are archaic, while the East Greek features of Lesbian (Anatolian Aeolic) are innovations which that dialect experienced under Ionic influence.

Owing to the highly complex nature of Greek dialect geography, it can hardly be said that there presently exists a consensus regarding the proper classification of Greek dialects – East versus West or North versus South. Risch's analysis is not without its uncertainties (see the comments of Cowgill 1966:80–81; see also Coleman 1963) but offers much to commend itself.

## 2. WRITING SYSTEMS

### 2.1 Linear B

Three separate writing systems were used for recording the Greek language in antiquity. The earliest of these is the syllabic script of the Mycenaeans called Linear B (see Table 25.1). Perhaps developed in the fifteenth century BC and based upon the Minoan Linear A script, Linear B consists almost entirely of V (vowel) and CV (consonant + vowel) characters. Owing to the common occurrence of consonant clusters in the Greek language, special strategies were of necessity devised for representing consonant sequences in the Linear B script. In some instances, the initial member of a cluster is simply deleted from the orthography, as in the spelling *pe-mo* for *spérmo* (σπέρμο “seed”). Alternatively, all members of a cluster may be spelled utilizing phonetically fictitious vowel graphemes: thus, *trípos* (τρίπος “tripod”) is spelled *ti-ri-po* (note that word-final consonants are not spelled). Linear B spelling does not distinguish voiced, voiceless, and voiceless aspirated consonants from one another, with the exception of the dental *d* which is distinguished from *t*<sup>(h)</sup>. Linear B script ceases to be attested after the downfall of Mycenaean society in the twelfth century BC.

### 2.2 The Cypriot syllabary

Consequent to the demise of Mycenaean society, large numbers of Greek émigrés settled on the island of Cyprus, where by at least the middle of the eleventh century a distinct syllabic script had been developed for writing Greek. The Cypriot syllabary appears to have been modeled graphically upon the Cypro-Minoan scripts of Cyprus, which are attested as early as the sixteenth century BC. The graphemic inventory of this the second of the Greek syllabaries was likewise composed predominantly of V and CV symbols (see Table 25.2). Fundamentally the scribal strategies utilized for spelling consonant sequences are the same as those found in Linear B practice, except that those clusters which Linear B scribes spelled with the omission strategy are now written with a phonetically fictitious vowel grapheme – one which is identical to the phonetic vowel which immediately *precedes* the cluster. Thus, *argúro* (ἀργύρω “of silver”) is spelled *a-ra-ku-ro*. Cypriot spelling practice also differs from the Mycenaean in that all word-initial clusters are spelled and word-final consonants are at times written (for both practices a fictitious-vowel strategy is employed). Much like their Mycenaean predecessors, Cypriot scribes fail to distinguish orthographically between voiced, voiceless, and voiceless aspirated consonants – including dentals in Cypriot practice. The Cypriot syllabary remained in use until the late third century BC.

### 2.3 The epichoric alphabets

For general discussion concerning the development of the Greek alphabet from the Phoenician consonantal script, see Chapter 24, §2.

As the Greek alphabet was carried across the Greek world in the eighth century BC, numerous local or epichoric alphabets developed. While many of these differ on the basis of variation in letter-shapes, ranging from subtle to radical, the various alphabets fall grossly into four or five fundamental groups according chiefly to the absence or presence (and form and arrangement) of the so-called “supplementals,” the non-Phoenician characters appended to the end of the Greek alphabet. While there is some correlation between Greek

**Table 25.1 The Linear B script**

Basic values													
A		E		I		O		U					
DA		DE		DI		DO		DU					
JA		JE				JO							
KA		KE		KI		KO		KU					
MA		ME		MI		MO		MU					
NA		NE		NI		NO		NU					
PA		PE		PI		PO		PU					
QA		QE		QI		QO							
RA		RE		RI		RO		RU					
SA		SE		SI		SO		SU					
TA		TE		TI		TO		TU					
WA		WE		WI		WO							
ZA		ZE				ZO							
Special values													
HA		AI		AU		DWE		DWO					
NWA		PTE		PHU		RYA		RAI					
RYO		TYA		TWE		TWO							
NUMERALS	1	10–100		1000		10,000							
WEIGHTS	12	1	4	1	30	1							
MEASURES	Dry	Wet	4	6	4	1	T; 10 T = 1 UNIT	6	4	1	T; 3 T = 1 UNIT		
Ideograms													
	MAN		WOMAN		RAM		EWE		BULL/OX		COW		
	WOOL		LINEN		CLOTH		OXHIDE		SHEEPSKIN		WHEAT		
	BARLEY		OLIVES		OLIVE OIL		FIGS		WINE				
	TRIPOD		JUG		AMPHORA		PAN		STIRRUP JAR				
	SWORD/DAGGER		CORSLET		CHARIOT		HORSE						
	GOLD		BRONZE		INGOT		FOOTSTOOL						

dialect and alphabet, such correlation is only partial. In certain cases, quite distinct dialects utilize alphabets of the same type; in others, conversely, closely related dialects are written with different alphabet-types.

Since the work of Kirchhoff 1887, the fundamental alphabet-types have been commonly referenced by color terms, following the color-coded map which Kirchhoff included at the end of the volume. Alphabets are green, blue, or red.

**Table 25.2 The Cypriot syllabary**

<i>a</i>	✱	<i>e</i>	✱	<i>i</i>	✱	<i>o</i>	≍	<i>u</i>	Υ
<i>ya</i>	ϙ					<i>yo</i>	ϙ		
<i>wa</i>	λλ	<i>we</i>	⊥	<i>wi</i>	Υλ	<i>wo</i>	⊥		
<i>ra</i>	Ω	<i>re</i>	⊥	<i>ri</i>	ϙ	<i>ro</i>	λ	<i>ru</i>	λλ
<i>la</i>	ϙ	<i>le</i>	8	<i>li</i>	≤	<i>lo</i>	+	<i>lu</i>	ω
<i>ma</i>	ϙ	<i>me</i>	✱	<i>mi</i>	ϙ	<i>mo</i>	ω	<i>mu</i>	✱
<i>na</i>	⊥	<i>ne</i>	λλ	<i>ni</i>	ϙ	<i>no</i>	λλ	<i>nu</i>	λλ
<i>pa</i>	⊥	<i>pe</i>	λ	<i>pi</i>	≍	<i>po</i>	λ	<i>pu</i>	ω
<i>ta</i>	⊥	<i>te</i>	✱	<i>ti</i>	⊥	<i>to</i>	⊥	<i>tu</i>	⊥
<i>ka</i>	⊥	<i>ke</i>	λ	<i>ki</i>	Υ	<i>ko</i>	λ	<i>ku</i>	✱
<i>sa</i>	ϙ	<i>se</i>	⊥	<i>si</i>	⊥	<i>so</i>	≍	<i>su</i>	λλ
<i>za</i>	λλ					<i>zo</i>	λλ		
<i>ksa</i>	λλ	<i>kse</i>	⊥						

### 2.3.1 Green alphabets

The green alphabets (or the “primitives”) are those of Crete and the neighboring islands of Thera and Melos (and are thus used by speakers of different dialects of Doric; see §1.1.4). This alphabet-type is characterized by the absence of the supplementals as well as by the absence of a character having the sequential value [k] + [s] (the existence of which is one of the hallmark idiosyncrasies of the Greek alphabet; see Woodard 1997:147–161).

### 2.3.2 Blue alphabets

The blue alphabets contain the non-Phoenician supplementals – or at least a subset thereof, as this group shows internal variants, distinguished as *dark blue* versus *light blue*. Both the dark blue and light blue alphabet-types have the supplementals Φ, representing [p<sup>h</sup>], and X, for [k<sup>h</sup>]. In addition, the dark blue type has the supplemental Ψ, a biconsonantal symbol representing the sequence [p] + [s]. The light blue type, however, lacks this symbol and spells the sequence [p] + [s] componentially with the two letters Φ + Σ. Furthermore, while dark blue alphabets have the letter Ξ, spelling [k] + [s], light blue scripts lack the character and spell the sequence with two letters, X + Σ (paralleling the spelling of [p] + [s]). The blue alphabet-types (particularly the dark blue) are far more widely distributed geographically than the green. A light blue alphabet was used early in Attica and on various of the Ionian Cycladic islands, for example. Dark blue alphabets occurred, among other places, in the northeastern Peloponnese and in Ionian cities of Anatolia. The Ionian dark blue script was adopted as the official alphabet of Athens at the end of the fifth century BC.

### 2.3.3 Red alphabets

Like the dark blue alphabet-type, the red alphabets are marked by the presence of the non-Phoenician supplementals. However, the value assigned to these symbols only partially agrees with their blue values. Red Φ represents [p<sup>h</sup>] (as in the blue alphabets), but X has the sequential value [k] + [s] (and not [k<sup>h</sup>]), and Ψ spells [k<sup>h</sup>] (and not [p] + [s]; for which there is no single red-alphabet character). Red alphabets were used widely throughout the

**Table 25.3 Epichoric Greek alphabets**

Letter-Name	Printed Letter	Local alphabets		Transcription
		Corinthian (blue-type)	Boeotian (red-type)	
<i>Alpha</i>	A, α	Α, Α	Α, Α	a
<i>Beta</i>	B, β	Β	Β, Β	b
<i>Gamma</i>	Γ, γ	Γ, Γ	Γ, Γ	g
<i>Delta</i>	Δ, δ	Δ	Δ, Δ	d
<i>Epsilon</i>	E, ε	Ε, Ε	Ε, Ε, Ε	e
<i>Digamma</i>	Ϝ	Ϝ, Ϝ	Ϝ, Ϝ, Ϝ	w
<i>Zeta</i>	Z, ζ	Ζ	Ζ	z + d
<i>Spiritus asper</i>	‘	Η, Η	Η, Η	h
<i>Theta</i>	Θ, θ	Θ, Θ, Θ	Θ, Θ, Θ	t <sup>h</sup>
<i>Iota</i>	I, ι	Ι	Ι	i
<i>Kappa</i>	K, κ	Κ	Κ	k
<i>Lambda</i>	Λ, λ	Λ	Λ	l
<i>Mu</i>	M, μ	Μ, Μ	Μ, Μ	m
<i>Nu</i>	N, ν	Ν, Ν	Ν, Ν	n
<i>Xi</i>	Ξ, ξ	Ξ	Ξ, Ξ, Ξ	k + s
<i>Omicron</i>	O, ο	Ο	Ο	o
<i>Pi</i>	Π, π	Π	Π, Π	p
<i>San</i>	Ϻ	Ϻ	—	s
<i>Qoppa</i>	Ϙ	Ϙ	Ϙ, Ϙ	q
<i>Rho</i>	Ρ, ρ	Ρ, Ρ, Ρ	Ρ, Ρ, Ρ	r
<i>Sigma</i>	Σ, σ	—	Σ, Σ	s
<i>Tau</i>	Τ, τ	Τ	Τ	t
<i>Upsilon</i>	Υ, υ	Υ, Υ, Υ	Υ, Υ, Υ	u
<i>Phi</i>	Φ, φ	Φ, Φ	Φ, Φ	p <sup>h</sup>
<i>Chi</i>	Χ, χ	Χ, Χ	Χ, Χ, Χ	k <sup>h</sup>
<i>Psi</i>	Ψ, ψ	Ψ, Ψ	Ψ, Ψ	p + s

Balkan Peninsula, thus blanketing numerous dialect boundaries. This alphabet-type was also in widespread use in Sicily and Magna Graecia (Greek Italy) and is the source of the Etruscan and Roman alphabets. On the possibility of distinguishing a “light red” from a “dark red” alphabet-type (paralleling the blue division) see Woodard 1997:215–216.

### 2.3.4 The Fayum alphabet

An additional Greek alphabet, one which does not fit into the preceding tripartite scheme, is known from four copper plaques, purported to have come from the Fayum in Egypt. Three of the four (two from the Schøyen collection in Oslo and one from the University of Würzburg Museum) have been examined carefully at the J. Paul Getty Museum in Los Angeles. An analysis of the physical remains reveals the plaques and the alphabets inscribed on them to be of great antiquity but does not permit an exact dating. The alphabet is epigraphically interesting in various ways, perhaps most interesting in that it ends in the letter *tau* (Τ), just as does the Phoenician precursor of the Greek alphabet. It is the only known Greek alphabet which matches the Semitic template in this manner, all others having the vowel

letter *upsilon* (Υ) added after *tau* (on the Greek creation of vowel characters, see Ch. 24, §2), and may represent the earliest form of the Greek alphabet (see Heubeck 1986, Scott, Woodard, McCarter, *et al.* 2003, Woodard 1997).

### 3. PHONOLOGY

In the remaining sections of this chapter, the discussion of Greek dialectal linguistic features closely follows the format of the treatment of Attic grammar presented in Chapter 24 and is dependent upon it. For background discussion of each section, the reader should consult the corresponding section in Chapter 24. Hereafter “dialects” should be construed to refer generally to all dialects other than Classical Attic, unless stated otherwise.

#### 3.1 Consonants

The inventory of consonant phonemes in the dialects is *grosso modo* the same as that of Attic. Variations do occur, however.

##### 3.1.1 Obstruents

As mentioned in Chapter 24, §3.1 the Proto-Greek labiovelar stop phonemes, /k<sup>w</sup>/, /k<sup>wh</sup>/, and /g<sup>w</sup>/, are preserved in Mycenaean Greek. The same dialect also appears to have possessed both a voiceless and a voiced palatalized stop (or perhaps affricate), sounds which developed from earlier sequences of \*[k<sup>(h)</sup>y], in the case of the voiceless, and \*[dy], \*[gy] as well as some instances of word-initial \*[y-], in the case of the voiced. Among the very few CCV characters occurring in the Linear B syllabary (see §2.1) are the symbols *twe*, *two*, *dwe*, *dwo*, *nwa*, *tya*, *rya*, and *ryo*. The existence of the signs may reveal the occurrence of palatalized and of labialized dental phonemes in the dialect at some time within the period of Mycenaean literacy and/or they may be relics of the phonological system of the non-Greek language for which the ancestor script of Linear B was designed.

The voiceless aspirated stops of Attic, /p<sup>h</sup>/, /t<sup>h</sup>/, and /k<sup>h</sup>/, would become the fricatives /f/, /θ/, and /x/ respectively in the post-Classical period – probably by the first or second century AD (perhaps earlier; see Allen 1987:20–23). However, there is evidence of a fricative reflex at a much earlier period among some of the dialects, such as the Doric dialect of Laconian. Thus, the later fifth-century authors Thucydides and Aristophanes, when reproducing Doric speech, use the letter σ (/s/) to spell the sound corresponding to Attic /t<sup>h</sup>/, suggesting an attempt to render a fricative pronunciation (i.e., /θ/). By the fourth century BC, a similar spelling practice is observed in Laconian inscriptions.

Like the voiceless aspirated stops, the voiced stops of Classical Attic have become voiced fricatives in Modern Greek: /b/, /d/, and /g/ yield /v/, /ð/, and /ɣ/ respectively. The date of the change is probably considerably later than that of the voiceless stops (as would be expected on typological grounds), though is difficult to pinpoint. In the ninth century AD, when Greek missionaries created a writing system for recording scripture translations in Old Bulgarian, the Greek letter β (Classical Attic /b/) provided a symbol for the Slavic voiced fricative /v/ (evidence from the earlier Greek-based Gothic and Armenian alphabets is inconclusive; see Allen 1987:28–30). A much earlier date (first century AD) for the shift of the voiced stops to fricatives in the Hellenistic Koine of Egypt, at least in some phonetic contexts, is suggested by spellings in nonliterary papyri (Allen 1987:154). Outside of Hellenistic Koine and its



descendant, just as with voiceless stops, the fricativization of voiced stops had a dialectal head start. In Laconian inscriptions dating as early as the fourth century BC,  $\beta$  is used at times in lieu of  $F$  (/w/), suggesting that the sound of  $\beta$  was, in some instances, a continuant (or that /w/ had become a labial fricative – or both; the same spelling variation is also found in Cretan). Similarly,  $\zeta$  (Attic /z/ + /d/) is at times used in place of  $\delta$  (/d/) in early Elean writing, and in Boeotian (as well as in Pamphylian) inscriptions  $\gamma$  (Attic /g/) is at times replaced by the vowel character  $\iota$  or deleted altogether.

Beyond the aforementioned early fricative reflexes of stops and the ubiquitous dental sibilant /s/, there is orthographic evidence of additional sibilant consonants occurring dialectally. In the alphabets of several Greek cities of Anatolia, there occurs a character  $\Pi$ , used to spell the common reflex of Proto-Greek  $*k^{(w)(h)}$ ,  $*t^{(h)}$  +  $y$  and  $*tw$ . The eventual reflex of the Proto-Greek consonantal sequences will be [-ss-] in the Ionic dialect of these Anatolian cities. In all likelihood the character  $\Pi$  represents an intermediate phonological stage – a strident sound which is distinct from the /s/ represented by  $\sigma$  (see Lejeune 1982:89, 101; Woodard 1997:178–179). More secure is the presence of distinct sibilants in Arcadian; the evidence is again orthographic. In the Arcadian alphabet a form of the letter *san*,  $\Pi$ , is used to spell the reflex of Proto-Greek  $*k^w$  occurring before front vowels. This sibilant reflex of Arcadian (found also in Cypriot, written syllabically, however, rather than alphabetically) must be distinct from dental /s/, spelled with *sigma* ( $\sigma$ ), and likely is to be identified as an affricate. The sound has a voiced counterpart in Arcadian (though not in Cypriot), with *zeta* ( $\zeta$ ) appropriated for its spelling (see Woodard 1997:178–184, 187–188).

The glottal fricative, /h/, of Attic, limited in native vocabulary to word-initial position, is shared by several other dialects. The presence or absence of this initial fricative (the *spiritus asper*, “rough breathing”) has served as a major isogloss in traditional Greek grammatical studies. Dialects which lack it, East Ionic, Lesbian, Cretan, and Elean, are called *psilotic*. In a few dialects, such as Laconian and Argolic, /h/ also occurs intervocalically; Cypriot is included in this number, though the presence of /h/ in this dialect is made less transparent by the syllabic script of Cyprus. Intervocalic /h/ may have also occurred in Mycenaean Greek, though the orthographic evidence is open to alternative interpretation.

### 3.1.2 Sonorants

In the Doric dialect of Cretan, when the lateral liquid (Attic /l/, spelled  $\lambda$ ) follows a vowel, the two are often spelled as a diphthong,  $V\upsilon$ : thus  $\acute{\alpha}\delta\epsilon\upsilon\pi\iota\alpha\acute{\iota}$  for Attic  $\acute{\alpha}\delta\epsilon\lambda\phi\alpha\acute{\iota}$  (*adelph<sup>h</sup>ai*) “sisters.” The use of *upsilon* ( $\upsilon$ ) for the liquid suggests some sort of back articulation, a velar-*l*, or perhaps a uvular approximant or fricative (see Bile 1988:120, who notes the occurrence of a velar-*l* in modern Cretan). Allen (1987:39) observes that Old Armenian transcriptions of Greek words may also suggest the presence of a velar-*l* in Asian Greek.

The use of the graphemic sequence  $\lambda\epsilon$  for /l/ in inscriptions from the island of Cos (also attested at neighboring Cnidos and on Melos and Thasos) may be an attempt to represent a lateral alveolar fricative [ɬ] (cf. Buck 1955:64).

The Proto-Greek labial glide  $*w$  appears in Mycenaean Greek and survives later in many of the first-millennium dialects than it had in Attic. The sound first disappears word-internally; in inscriptional spellings its grapheme ( $F$ ) continues word-initially until as late as the second century AD, though by this time its sound had perhaps become a fricative (see Buck 1955:46–48).

The palatal glide /y/ does not exist phonemically in any of the dialects of the first millennium BC. In Mycenaean Greek it occurs word-initially, as well as after vowels and sonorants, though in some of these contexts /y/ was perhaps in the process of evolving into /h/ during

the period from which Mycenaean documents survive (see Lejeune 1982:155–156, 162, 165, 167–169, 171).

### 3.2 Vowels

The short and long front rounded vowel phonemes of Attic, /ü/ and /ü:/, which arose by the fronting of Proto-Greek \**ǵ*, are not present in all dialects. The Aeolic dialect of Boeotian, for example, preserved the high-back position, /u(:)/; this is revealed by the use of the digraph ου (= Attic /u:/) in lieu of υ (= Attic /ü(:)/) when, in the middle of the fourth century BC, Boeotian speakers adopted the Attic alphabet: for example, Boeotian ἀργούριον (*argurion*) beside Attic ἀργύριον (“money, silver”). Boeotian, however, had developed its own front rounded vowel by the third century BC, through the fronting of the earlier diphthong /oi/ (οι) – the result being perhaps /ö/, then /ü/, ultimately /i/. Similarly, Boeotian /ai/ underwent monophthongization, becoming a long lower mid-front vowel, spelled with η after the acquisition of the Attic alphabet. Proto-Greek \**e*i, which had given rise to the long higher mid-front vowel /ē:/ in Attic, probably underwent a similar development in Boeotian; but by the fifth century BC, the Boeotian vowel had moved farther upward along the front periphery of the vowel track to merge with /i:/.

Throughout the history of Greek, the language has demonstrated a tendency for vowel monophthongization and movement forward and upward. The aforementioned vowel developments that characterized Boeotian at an early period occurred later in the Attic-based Hellenistic Koine, ultimate parent of Modern Greek (see Ch. 24, §1). The change of /ē:/ (ει) to /i:/ was already well underway by the third century BC. The lower mid-front /ē:/ (η) in response was raised (perhaps to /ē:/), eventually itself becoming /i:/, in some areas, perhaps by the second century AD. By about the beginning of that century, inscriptional spellings reveal that the diphthong /ai/ was undergoing monophthongization and raising to /ē:/ – in essence filling the gap created by the upward shift of earlier /ē:/ (η). The new lower mid-front vowel /ē:/ would, prior to the Byzantine era, merge with the vowel of ε. As earlier in Boeotian, so in Hellenistic Koine, the diphthong /oi/ shifted forward, developing into a front rounded monophthong, prior to the middle of the third century AD. Both this vowel and the already existing /ü(:)/ eventually unrounded, becoming Modern Greek /i/. For detailed discussion of these and related developments, see, *inter alia*, Allen 1987:74, with further page references.

### 3.3 Accent

Whatever accentual idiosyncrasies might have characterized the various dialects are for the most part unknown. Aeolic is notably different than Classical Attic in that the accent of all words (except conjunctions and prepositions) – and not only verbs – is recessive: for example, [pótamos] beside Attic [potamós] (ποταμός “river”); [basíleus] beside Attic [basileús] (βασιλεὺς “king”); [zdeús] beside Attic [zdeús] (Ζεὺς “Zeus”); and so forth (see Thumb-Scherer 1959:86–87; Allen 1973:238–239; 256–257). In the present work, all first-millennium dialect forms are conventionally marked with the appropriate Attic accent except in those cases in which it is known that the dialect accentuation differs from that of Attic.

The tonal accent of Classical Attic eventually gave way to a stress accent, present still in Modern Greek. Dating the shift from a pitch to a stress system is an uncertain affair, though it appears that the change was in progress in Attic by at least the first centuries AD, and perhaps much earlier. Evidence provided by Egyptian papyri suggests that among Egyptian

Greeks the change may have occurred by the last two centuries BC (see Allen 1987:119–120; Devine and Stephens 1994:215–223).

### 3.4 Diachronic developments

#### 3.4.1 Obstruents

As noted earlier (see §3.1.1), the Proto-Indo-European labiovelar stops,  $*k^w$ ,  $*g^w$ , and  $*g^{wh}$ , are generally preserved in Mycenaean Greek, except, of course, that voiced aspirated  $*g^{wh}$  has devoiced to  $k^{wh}$  (a Proto-Greek development), though the same set of CV graphemes is used for spelling all three types (voiceless, voiced, voiceless aspirated). The Mycenaean dialect exhibits a tendency to dissimilate one of two labiovelar consonants found within a word. Compare the spelling of the proper name *qe-re-qo-ta* (Pylos) beside *pe-re-qo-ta* (Pylos and Knossos); see Ventris and Chadwick 1973:82, 245, 399, 447).

The Aeolic dialects show labiovelar developments which are in part distinct from those of Classical Attic (see Ch. 24, §3.7.1) and other dialects of the first millennium BC (except Cypriot; see below). On the one hand, just as in Attic, bilabial stops constitute the unconditioned reflexes of these sounds in Aeolic: PIE  $*k^w$ ,  $*g^w$ ,  $*g^{wh}$  → [p, b, p<sup>h</sup>] (π, β, φ) respectively. Aeolic also agrees with other dialects in dissimilating the labiovelars before and after the high back rounded vowel *u*: PIE  $*k^w$ ,  $*g^w$ ,  $*g^{wh}$  → [k, g, k<sup>h</sup>] (κ, γ, χ). On the other hand, however, when occurring before mid-front vowels, labiovelars become bilabials in Aeolic (the default development), rather than dentals as in Attic: PIE  $*k^w$ ,  $*g^w$ ,  $*g^{wh}$  → Aeolic [p, b, p<sup>h</sup>] (π, β, φ) (cf. Attic [t, d, t<sup>h</sup>] (τ, δ, θ)). Thus, for example, while  $*g^w elb^h$ - gives Attic [delp<sup>h</sup>is] (δελφίς “dolphin”), the Lesbian reflex is [bél<sup>h</sup>is] (βέλφίς); and PIE  $*k^w etw̥r$ - becomes Attic [téttares] (τέτταρες “four”), but Thessalian and Boeotian [péttares] (πέτταρες). Note, however, that in the case of the enclitic conjunction  $*-k^we$ , the outcome is [-te] (-τε “and”) in all dialects (i.e., in a clitic context, Aeolic participates in a change, a palatalization process, found regularly in non-Aeolic dialects). Aeolic also agrees with Attic in the development of a voiceless dental reflex ([t]) before the high front vowel [i], but bilabial voiced and voiceless aspirated reflexes ([b] and [p<sup>h</sup>]) in this environment. On these labiovelar developments and their wave-like spread through dialect regions, see Stephens and Woodard 1986.

Distinct labiovelar developments also occur within Arcado-Cypriot. In part these developments constitute an isogloss distinguishing the dialectal group from all others; in part they divide the two members of the group. Within Arcado-Cypriot the palatalization of the labiovelars is carried a step beyond the [t] reflex seen elsewhere before the high front vowel. Both Arcadian and Cypriot develop a continuant reflex in this context (probably an affricate). The two sister dialects differ, however, in the extent of the development: in Arcadian the change is more pervasive, occurring before mid-front vowels also, and affecting both voiceless and voiced labiovelars (and likely the voiceless aspirated as well, though this is not yet attested). In Cypriot, the labiovelars develop bilabial reflexes before mid-front vowels, as in Aeolic. On these developments, see Woodard 1997:180–184.

#### 3.4.2 Sonorants

The vocalization of the Proto-Indo-European syllabic liquids,  $*r$  and  $*l$  shows dialectal variation. In Aeolic the reflex of  $*r$  is either [or] or [ro], rather than the [ar] or [ra] of Attic: for example, PIE  $*str̥-to-$  → Lesbian and Boeotian [strótos] (στρότος “army”) beside Attic [stratós] (στρατός). The same treatment is found in Arcado-Cypriot, as in Arcadian

[storpá] (στορπά “lightning”), as well as in Mycenaean: for example, *k<sup>w</sup>etro-* (“four”) beside Attic [tetra-] (τετρα- “four”); compare Thessalian [petro-] (πετρο-). The Proto-Indo-European lateral syllabic liquid *\*l̥* similarly gives [ol] or [lo] (Attic [al] or [la]): for example, from PIE *\*ǵ<sup>h</sup>l̥* - develops Lesbian [k<sup>h</sup>ólaisi] (χόλαισι “they loosen”), Attic [k<sup>h</sup>alô:si] (χαλῶσι).

On the *o*-reflex of the Proto-Indo-European syllabic nasals in Mycenaean (Mycenaean I), see §1.2. First-millennium dialects agree with Attic and the less commonly attested form of Mycenaean (Mycenaean II) in showing the *a*-reflex.

### 3.4.3 Combinatory changes

The chief phonological developments which occur in Attic when two or more phonetic segments come into contact are detailed in Chapter 24, §3.7.3. Among those changes, the following dialectal developments constitute significant isoglosses (these dialectal distributions should be viewed as tendencies rather than absolutes); reconstructed sequences are presented first, followed by their reflexes in the various dialects:

1. PG *\*t<sup>(h)</sup>γ*: (A) [tt] in Boeotian and Cretan; (B) *\*[ss]*, then becoming [s] in Attic, Ionic (though Homer has both [ss] and [s]) and Arcadian; (C) [ss] in other dialects; (D) however, following a consonant or long vowel, as well as word-initially, all dialects have [s].
2. PG *\*t<sup>(h)</sup> + γ* (i.e., when a detectable morpheme boundary separates the two consonants): (A) [tt] in Attic, Boeotian, and Cretan; (B) [ss] in other dialects (on this complex matter, see Rix 1976:90–91; Lejeune 1982:103–104).
3. PG *\*k<sup>(w)</sup>(h)γ*: (A) [tt] in Attic, Boeotian, and Cretan ([t] word-initially); (B) [ss] in other dialects ([s] word-initially).
4. PG *\*dγ* and *\*g<sup>(w)</sup>γ*: (A) [dd] in Boeotian, Thessalian, Laconian, Elean, and Cretan ([d] word-initially); (B) [zd] in other dialects.
5. PG *\*tw*: (A) [tt] in Attic, Boeotian, and Cretan; (B) [ss] in other dialects; (C) however, all dialects have [s] word-initially.
6. PG *\*t<sup>(h)</sup>, d}s*: (A) [tt] in Boeotian and Cretan; (B) *\*[ss]*, then becoming [s] in Attic, Ionic (though Homer has both [ss] and [s]) and Arcadian; (C) [ss] in other dialects; (D) however, following a long vowel, as well as word-finally, all dialects have [s].
7. PG *\*ss*: (A) [s] in Attic, Ionic (though Homer has both [ss] and [s]) and Arcadian; (B) [ss] preserved in other dialects; (C) however, following a long vowel all dialects have [s].
8. PG *\*ti*: (A) [si] in Attic, Ionic, Arcado-Cypriot, and Lesbian, as well as Mycenaean I (see §1.2); (B) [ti] remains in other dialects; (C) however, the change does not occur if *\*ti* is preceded by *\*s*; (D) and in the case of certain words [si] develops in all dialects, in the case of others [ti] is preserved in all dialects (see Buck 1955:57–58).
9. PG *\*{r, n}γ* after *\*{e, i, u}*: (A) geminate [{rr, nn}] in Lesbian and Thessalian; (B) [{r, n}] with compensatory lengthening of the preceding vowel in other dialects.
10. PG *\*ln*: (A) geminate [ll] in Lesbian and Thessalian; (B) [l] with compensatory lengthening of a preceding vowel in other dialects.
11. PG *\*{r, l, n, s}w*, where *\*s* is of secondary origin (i.e., not inherited from Proto-Indo-European): (A) [{r, l, n}] with compensatory lengthening of a preceding vowel in East and Central Ionic, and in several Doric dialects (Argolic, Cretan, Thera, and the dialects of Rhode and Cos); (B) [{r, l, n}] without compensatory lengthening of a preceding vowel in other dialects.

12. PG *\*ns#* (word-final): (A) [ns#] preserved in Argolic and central Cretan; (B) [s#] with no effect on the preceding vowel from the loss of [n] in Arcadian, Thessalian, and Thera Doric; (C) [s#] with *i*-diphthongization of the preceding vowel in Lesbian and Elean; (D) [s#] with compensatory lengthening of a preceding vowel in most other dialects (on original conditioning by the first sound of the ensuing word and the occurrence of doublets, see Buck 1955:68; Lejeune 1982:131–132).
13. PG *\*nsV*, where *\*s* is of secondary origin (i.e., not inherited from Proto-Indo-European): (A) [nsV] preserved in Arcadian, Thessalian, Argolic, and central Cretan (contrast 12 [A] and [B] for dialect distribution); (B) [sV] with *i*-diphthongization of the preceding vowel in Lesbian and Cyrenaean Doric; (C) [sV] with compensatory lengthening of a preceding vowel in other dialects.
14. PG *\*NsV*, where *\*s* is inherited: (A) geminate [NNV] in Lesbian and Thessalian; (B) [NV] with compensatory lengthening of a preceding vowel in other dialects.
15. PG *\*Vsw*, where *\*s* is inherited: (A) geminate [Vww] in Lesbian and Thessalian; (B) [Vw] with compensatory lengthening of the preceding vowel and eventual loss of [w] in other dialects.
16. PG *\*Vs{r, l, m, n}*: (A) geminate [V{rr, ll, mm, nn}] in Lesbian and Thessalian; (B) [V{r, l, m, n}] with compensatory lengthening of the preceding vowel in other dialects.
17. PG *\*rs*, where *\*s* does not belong to the aorist suffix: (A) geminate [rr] in Attic, West Ionic, Arcadian, Elean, and Thera Doric; (B) [rs] preserved in most other dialects.
18. PG *\*{r, l}s* where *\*s* belongs to the aorist suffix: (A) geminate [{rr, ll}] in Lesbian and Thessalian; (B) [s] with compensatory lengthening of the preceding vowel in other dialects (cf. 14).
19. PG *\*Vs#* (word-final): (A) [Vr#] in Elean (especially in later inscriptions) and in late Laconian; (B) preserved in other dialects, though Plato (Cratylus 434 C) has Socrates note that Eretrian speakers have a final *-r* in their pronunciation of the word *sklērotēs* (σκληρότης, “hardness”).
20. PG *\*VsV*: (A) [VrV] in early Eretrian, though not consistently attested; (B) [s] lost in other dialects but sometimes restored by analogy.
21. PG *\*sC<sub>+voice</sub>*: (A) [rC<sub>+voice</sub>] attested in Eretrian, Thessalian, Cretan, and Laconian; (B) [zC<sub>+voice</sub>] in most dialects.

### 3.4.4 Vowels

The change of [a:] to [ɛ:] which occurs in both Attic and Ionic (see Ch. 24, §3.7.4) is not identical in its distribution in these sister dialects. While the change is thoroughgoing in Ionic, it does not occur (or is reversed; see Szemerényi 1968) in Attic when [a:] is preceded by [e], [i], or [r]. The opposite change of [ɛ:] to [a:] appears to have occurred in Elean, though its attestation is inconsistent. In Northwest Greek generally, [e] is lowered to [a] when it occurs before [r]; while in Aeolic, high-front [i] is lowered to the mid vowel [e] when preceded by [r]. In Arcado-Cypriot, mid vowel [e] is raised to [i] when it occurs before the dental nasal [n]; at the back of the mouth, the same dialect raises mid vowel [o] to [u] in word-final position. In Cretan Doric [e] is raised to [i] when a vowel follows.

As in Attic-Ionic, the initial vowel in sequences of [ɛ:] + vowel commonly undergoes shortening in Doric and Northwest Greek, though without the quantitative metathesis found in Ionic and, especially, Attic (see Ch. 24, §3.7.5). In Arcado-Cypriot, Aeolic, and Elean, however, the initial vowel remains long.

The so-called spurious diphthongs of Attic, spelled εἰ and οὐ (actually long monophthongs written as digraphs; see Ch. 24, §3.2), are long vowels that arose secondarily by contraction

or compensatory lengthening, and vowels that are distinct from the long  $*\bar{e}$  and  $*\bar{o}$  vowels inherited from Proto-Indo-European. These vowels – [e:] and ( $*[o:] >$ ) [u:] in Attic – are found in numerous dialects. In other dialects, however, the long vowels which develop secondarily are identical to those inherited, as in the Aeolic dialects, Arcadian, Elean, Cretan, and Laconian.

## 4. MORPHOLOGY

### 4.1 Nominal morphology

For an overview of Greek nominal morphology, see Chapter 24, §4.1

#### 4.1.1 Noun classes

Greek nouns are traditionally divided into three declensional classes (first, second, and third); for a general discussion of these, see Chapter 24, §4.1.1.

##### 4.1.1.1 First declension

In dialects other than Attic-Ionic, the characteristic suffix of the first declension nouns remains [-a:] ( $-\bar{\alpha}$ , from PIE  $*-eh_2$ ), not having undergone the shift to [-e:] ( $-\eta$ ; see §3.4.4). The following Cretan (Doric) forms (see Bile 1988:188–190) exemplify the singular paradigm, where the greatest deviation from Attic-Ionic is found (cf. Ch. 24, §4.1.1.1 [4]–[5]):

(1)	<i>Singular</i>
<i>Nominative</i>	gâ (γᾱ̃, “earth”)
<i>Vocative</i>	t <sup>h</sup> eá (θεᾱ̃, “O goddess”)
<i>Accusative</i>	stégān (στέγα̃ν, “house”)
<i>Genitive</i>	tīmās (τῖμα̃ς, “of honor”)
<i>Dative</i>	stégāi (στέγα̃ι, “to [the] house”)

Among dialectal forms in the plural, the greatest variation occurs in the dative and, especially, the accusative (on which see below). While most dialects agree with Attic and have a dative in *-ais* ( $-\alpha\iota\varsigma$ ), Lesbian shows *-a(:)isi* ( $-\alpha\iota\sigma\iota/-\bar{\alpha}\iota\sigma\iota$ ) and Ionic commonly has  $-\tilde{\epsilon}\iota\sigma\iota$  ( $-\eta\iota\sigma\iota$ ). Cretan and other dialects sometimes attest the infrequent *-āsi* ( $-\bar{\alpha}\sigma\iota$ ).

Mycenaean Greek has a distinct suffix *-pi* ( $-\phi\iota$ ) marking the instrumental plural (as in Linear B *a-ni-ja-pi* “with reins”). The ending is also used with place names in apparently locative or ablative function.

The ancestral accusative plural  $*-āns$  is preserved in Argolic and Cretan (though *-ans* ( $-\alpha\nu\varsigma$ ) with vowel shortening by Osthoff’s Law; see Ch. 24, §3.7.5). With loss of [n] before word-final [s], diverging dialectal reflexes emerge. Most widely occurring is the *-ās* ( $-\bar{\alpha}\varsigma$ ) form found in Attic; a short vowel formant *-as* ( $-\alpha\varsigma$ ) characterizes Thessalian, Arcadian, and certain Doric dialects (including Cretan, also with *-ans*,  $-\alpha\nu\varsigma$ ). Elean and Lesbian have *-ais* ( $-\alpha\iota\varsigma$ ), with the former also showing a further development to *-air* ( $-\alpha\iota\rho$ ; see §3.4.3, 19).

Corresponding to the Attic-Ionic nominative singular  $-\tilde{\epsilon}\varsigma$  ( $-\eta\varsigma$ ) of the first declension masculine nouns (see Ch. 24, §4.1.1.1 [7]), most dialects show the formant *-ās* ( $-\bar{\alpha}\varsigma$ ). Under the influence of the second declension, the genitive singular of the masculine is commonly formed in *-ā* ( $-\bar{\alpha}$ ; from  $-\bar{\alpha}o$  ( $-\bar{\alpha}o$ )), giving Arcado-Cypriot *-āu* ( $-\bar{\alpha}u$ ; see §3.4.4).



### 4.1.1.2 Second declension

In the singular of the second declension (cf. Ch. 24, §4.1.1.2 (8)–(9)), both the genitive and dative show dialectal variants. Genitive *-oio* (-οιο), from PIE \*-*osyo*, is found, among other places, in Homer and is the source of Thessalian *-oi* (-οι). The Attic genitive formant *-ū* (-ου) is shared by Ionic and certain Doric and Northwest Greek dialects and arose by vowel contraction after the loss of the two intervening consonants. Showing a different long vowel reflex, other Doric, Northwest Greek, and Aeolic dialects, as well as Arcadian, are characterized by a genitive singular in *-ō* (-ω; on the dialectal distribution, see §3.4.4). Cypriot has both the expected *-ō* and an innovative *-ōn*.

As with the first declension, both the dative and accusative plural show dialectal variation. The dative formant *-oisi* (-οισι) of Early Attic also occurs, among other dialects, in Ionic, Lesbian, and Pamphylian. Most dialects agree with Classical Attic in having *-ois* (-οις).

The accusative plural distribution mirrors that of the first declension: archaic \*-*ons*, preserved in Argolic and Cretan (-ovς); Lesbian *-ois* (-οις; and Elean *-oir*, -οιρ; see §3.4.3, 19); in most dialects a long vowel reflex, with the quality of the vowel showing variation, (\*-*ōs* >) *-ūs* (-ους) or *-ōs* (-ως; see §3.4.4). Thessalian, Arcadian, and a subset of Doric dialects are again characterized by a short vowel form *-os* (-ος).

### 4.1.1.3 Third declension

Among third declension inflections, various dialectal forms occur. A widely distributed consonant stem variant is the dative plural in *-essi* (-εσσι; see Ch. 24, §4.1.1.3 (11)), found throughout the Aeolic branch and in scattered Doric dialects, as well as in Pamphylian. In Mycenaean Greek, the instrumental plural suffix *-pi* (-φι; see §4.1.1.1) also occurs on third declension nouns. Most dialects differ from Attic in preserving *s*-stem endings without contraction after loss of intervocalic \*-*s*-; thus, genitive *-e-os* (-εος), nominative-accusative neuter plural *-e-a* (-εα), and so forth (cf. Ch 24. §4.1.1.3, 2). Outside of Attic (and some varieties of Ionic), *i*-stems are uniformly of the type which preserve stem-vowel *-i-*. The difference between the Attic and a typical non-Attic type (i.e., between an ablauting and non-ablauting suffix) can be illustrated by the paradigms of the *i*-stem noun *pólis* (πόλις, “city”) of (2) and (3) respectively:

#### (2) Attic *i*-stem

	<i>Singular</i>	<i>Plural</i>
<i>Nominative</i>	pólis (πόλις)	pólēs (πόλεις)
<i>Vocative</i>	póli (πόλι)	pólēs (πόλεις)
<i>Accusative</i>	pólin (πόλιν)	pólēs (πόλεις)
<i>Genitive</i>	póleōs (πόλεως)	póleōn (πόλεων)
<i>Dative</i>	pólei (πόλει)	pólesi (πόλεσι)

#### (3) Non-Attic *i*-stem

	<i>Singular</i>	<i>Plural</i>
<i>Nominative</i>	pólis (πόλις)	pólies (πόλιες)
<i>Vocative</i>	póli (πόλι)	pólies (πόλιες)
<i>Accusative</i>	pólin (πόλιν)	pólīs (πόλῖς)
<i>Genitive</i>	pólios (πόλιος)	políōn (πολίων)
<i>Dative</i>	póli (πόλι)	pólisi (πόλισι)



With the type of (3), compare the Attic paradigm of *oīs* (οἷς, nom. sg.), *oiós* (οἰός, gen. sg.) “sheep” (see Ch. 24, §4.1.1.3, 6)

The diphthongal *ēu*-stems in Ionic, Doric, and Northwest Greek are unlike those of Attic (see Ch. 24, §4.1.1.3, 8) in that the initial vowel of the suffix is shortened before a vocalic ending (gen. *-eos* (-εος), etc.) rather than remaining long and triggering *quantitative metathesis* (see Ch. 24, §3.7.5). In some dialects, the second element of the diphthong is preserved intervocalically and spelled with digamma; when the second element is lost, contraction is common.

#### 4.1.2 Pronouns

For an overview of the pronominal system of Classical Greek, see Chapter 24, §4.1.3.

##### 4.1.2.1 Personal pronouns

Among the various dialectal differences in the personal pronouns, one of the most readily apparent is the form of the second-person stem in the singular. In some dialects the pronoun begins with *t*-, in others with *s*-; the dialectal distribution parallels that of *ti* versus *si* (see §3.4.3, 8). These pronouns together with a few additional forms suggest a more limited assibilation of PG *\*t* to *s* before *\*u*: Proto-Greek nominative *\*tú* “you” gives Attic (etc.) *sú* (σú), Doric (etc.) *tú* (τύ). Proto-Greek accusative *\*twé* produces Attic (etc.) *sé* (σέ; see §3.4.3, 5 – alternatively, the initial *s*- of the accusative could possibly be an analogical source for that of the nominative); the source of Doric (etc.) *té* (τέ) appears to have developed from a Proto-Greek variant *\*té* (see Lejeune 1982:66; Chantraine 1984:136–137). First-, second- and third-person singular pronouns from various dialects are presented in (4); compare those of Attic given in Chapter 24, §4.1.3.1:

(4)	<i>First</i>	<i>Second</i>	<i>Third</i>
<i>Nominative</i>	Boeotian <i>iō</i> (ἰώ)	Boeotian <i>toú</i> (τοῦ)	—
<i>Genitive</i>	Ionic <i>emeū</i> (ἐμεῦ)	Doric <i>teū</i> (τεῦ)	Aeolic <i>wét<sup>h</sup>en</i> (Fέθεν)
<i>Dative</i>	Doric <i>emín</i> (ἐμίν)	Doric <i>tín</i> (τίν)	Aeolic <i>woi</i> (Fοι)
<i>Accusative</i>	Ionic <i>emé</i> (ἐμέ)	Doric <i>té</i> (τέ)	Pamphylian <i>whe</i> (Fηε) Sicilian Doric <i>nín</i> (νίν)

The first- and second-person plural pronouns of Attic, *hēmês* (ἡμεῖς) and *hūmês* (ὅμεῖς) in the nominative, are formed from the Proto-Greek stems *\*h₂sme-* and *\*usme-* respectively. The Thessalian and Lesbian forms of these pronouns are thus marked by their characteristic geminate reflex of the cluster *\*sm* (see §3.4.3, 16). Selected dialect forms of the plural pronouns appear in (5); see, again, Chapter 24 for Attic equivalents:

(5)	<i>First</i>	<i>Second</i>
<i>Nominative</i>	Lesbian <i>ámmeš</i> (ἄμμεš)	Lesbian <i>úmmes</i> (ὕμμεš)
<i>Genitive</i>	Doric <i>hāmēōn</i> (ἡμέων)	Doric <i>hūméōn</i> (ὅμέων)
<i>Dative</i>	Lesbian <i>ámmi</i> ((ἄμμι)	Lesbian <i>úmmi</i> (ὕμμι)
<i>Accusative</i>	Thessalian <i>hammé</i> (ἄμμέ)	Doric <i>hūmé</i> (ὅμέ)

Dialect forms of the third-person plural pronoun are seen, for example, in the Doric dative *sp<sup>h</sup>in* (σφίν), Lesbian dative *ásp<sup>h</sup>i* (ἄσφι), Sicilian Doric *psín* (ψίν), Lesbian accusative *ásp<sup>h</sup>e* (ἄσφε), and Sicilian Doric *psé* (ψέ).

The formation of possessive adjectives in the various dialects is like that in Classical Attic (see Ch. 24, §4.1.3.1). In the nominative case, the first-person possessive appears, for example, as Doric *hāmós* (ἄμός, sg.), *hāméteros* (ἄμέτερος, pl.), Lesbian *ámmos* (ἄμμος, sg.), *amméteros* (ἄμμέτερος, pl.). The second-person singular shows the same *t*- ~ *s*- dialect alternation that occurs in the personal pronouns: thus, Lesbian and Doric *teós* (τεός). For the second plural, Doric has *hūméteros* (ὕμέτερος), Lesbian *umméteros* (ὕμμέτερος). Among third-person forms, nominative singular appears in Cretan as *wos* (Ϝος), in Doric as *heós* (ἑός). A third plural form *sp<sup>h</sup>ós* (σφός) occurs in both Doric and Lesbian.

#### 4.1.2.2 Reflexive pronouns

The dialects display a variety of constructions for the reflexive pronoun. The personal pronouns (see §4.1.2.1) alone are at times used as reflexives. Other formations involve the use of the pronoun *autós* (αὐτός) together with personal pronouns – either as a lexical pair or, as commonly in Classical Attic (see Ch. 24, §4.1.3.2), in a univerted (compound) form. The former type is seen, for example, in the Cretan third singular dative *wín autôî* (Ἔν αὐτῷ), the latter type in Cretan *wiautô* (Ἔιαυτῷ) third singular genitive. In some dialects, oblique forms of *autós* are used alone as reflexives; in some, *autós* is used in one of several reduplicated forms, such as Delphian *autosautón* (αὐτοσαυτόν), accusative singular.

#### 4.1.2.3 Definite article

The definite article of Classical Attic (see Ch. 24, §4.1.3.4) differs most conspicuously from dialectal forms in the nominative animate plural. In Northwest Greek and all Doric dialects except Cretan, as well as in the Aeolic dialects of Boeotian and, in part, Thessalian (and in Homer), the archaic masculine *toí* (τοί) and feminine *taí* (ταί) survive, in contrast to the innovative *hoi* (οἱ) and *hai* (αἱ) found elsewhere. The definite article does not occur in Mycenaean Greek; when the aforementioned formants appear, they function as demonstrative pronouns, as they do in the Homeric dialect (see Ch. 24, §4.1.3.4).

#### 4.1.2.4 Demonstrative pronouns

Dialectal variation occurs throughout the demonstrative pronoun paradigms (for Attic, see Ch. 24, §4.1.3.5). For example, the demonstratives *hóde* (ὅδε) and so forth of Classical Attic are formed with a particle other than *-de* in certain dialects: thus, Arcado-Cypriot has *ónu* (ὄνυ) beside Arcadian *oní* (ὀνί); Thessalian has *hóne* (ὄνε). The Attic near demonstratives *hûtos* (οὗτος) and *haútē* (αὕτη), masculine and feminine nominative singular respectively, appear in the nominative plural also with an initial *h*-; outside of the nominative (singular and plural), all members of the animate paradigm, as well as all neuter forms – including nominatives – have initial *t*-. In some Doric and Northwest Greek dialects, however, the initial *t*- of the animate nominative plural has been preserved, as in the paradigm of the article (see §4.1.2.3), thus masculine *tûtoi* (τοῦτοι) feminine *taûtai* (ταῦται). Boeotian, on the other hand, has generalized initial *h*- throughout the entire paradigm. The Attic far demonstrative, masculine *ekēnos* (ἐκεῖνος) and so forth, appears in Ionic, Lesbian, and certain Doric dialects without initial *e*-. In most Doric dialects, however, the far demonstrative takes the form *tēnos* (τῆνος) and so on.

#### 4.1.2.5 Interrogative/indefinite pronoun

The interrogative *tís, tí* (τίς, τί), indefinite *tis, ti* (τις, τι) of Classical Attic (see Ch. 24, §4.1.3.6) occurs in most dialects (from PIE *\*k<sup>w</sup>i-*). Showing the advanced stage of assibilation of the labiovelars, however, Cypriot has *si-se* (the syllabic Cypriot spelling) and Arcadian *hís* (see §§3.1.1, 3.4.1). The Thessalian pronoun takes the form *kís, kis* (κίς, κίς).

#### 4.1.2.6 Relative pronouns

The Classical Attic relative pronoun *hós* (ὅς masc.), *hē* (ἥ fem.), *hó* (ὃ neut.; see Ch. 24, §4.1.3.7) is found across the Greek dialect map. However, the definite article (see §4.1.2.3) is commonly used as a relative pronoun in Lesbian, Thessalian, and Arcado-Cypriot, and is attested in this use elsewhere as well.

### 4.2 Verbal morphology

The Classical Attic verb system described in Chapter 24, §4.2 is for the most part characteristic of all ancient Greek dialects. Particular differences are noted in the ensuing discussion.

#### 4.2.1 Present tense stems

Among the various types of thematic present tense stems of Classical Attic (see Ch. 24, §4.2.2), the most notable dialect variation occurs in the so-called contract verbs. In the Aeolic dialects of Lesbian and Thessalian as well as in Arcado-Cypriot, the contract verbs are inflected as *athematic* rather than thematic constructions, in that they take the endings of the *mi*-verbs (i.e., athematic verbs; for a full presentation of the forms of the endings, see Ch. 24, §4.2.9): for example, Lesbian *p<sup>h</sup>ilēmmi* (φίλημμι) for Attic *p<sup>h</sup>ilēō* (φιλέω). Conversely, in Ionic (and occasionally elsewhere, even in Attic), some *mi*-verbs are inflected as thematic contract verbs, as in *tithē* (τίθει) beside Classical Attic *tithēsi* (τίθησι). Certain Attic *-aō* (-αω) contract verbs appear in a variety of dialects as *-eō* (-εω) verbs.

#### 4.2.2 Future tense stems

In Doric and Northwest Greek, the future tense stem appears as a contract-verb construction formed in *-se-* (the so-called Doric future): thus, first singular Cretan *speusiō* (σπευσίω, with \**e* raised to *i* before a vowel; see §3.4.4) beside Attic *speúsō* (σπεύσω “I will hasten”; on the Attic future, see Ch. 24, §4.2.5).

These same dialects as well as Arcado-Cypriot and Balkan Aeolic (though only partially in Boeotian) show an innovative future morphology of verbs which have a present tense stem marked by *-zd-* (ζ). In a subset of such verb-stems, the cluster [zd] had developed historically from the consonantal sequence \**gy* (see §3.4.3, 4), where the verb root ends in a velar stop to which the thematic suffix *-ye/yo-* is attached (see Ch. 24, §4.2.2, 2). In the case of these verbs, the future stem would then be formed in *-ks-* (ξ, from a Proto-Greek velar stop followed by the *s*-formant of the future). In the aforementioned dialects, all future tense stems of *zd*-presents tend to be produced with a formant *-ks-* (rather than *-s-*), regardless of whether or not the root originally ended in a velar stop. For example, *erízdō* (ἐρίζω “I strive”) forms a Doric future *eríksō* (ἐρίξω) beside Attic *erísō* (ἐρίσω). In the present stem of this verb, the cluster *zd* arose from the Proto-Greek sequence \**dy* (see §3.4.3, 4) rather than \**gy*.

#### 4.2.3 Aorist tense stems

The sigmatic or *s*-aorist (see Ch. 24, §4.2.6) shows a dialect variation like that of the *s*-future stems described immediately above: in Doric, Northwest Greek, Thessalian, and partially in Boeotian and Arcado-Cypriot, the *s*-aorist of present stems terminating in *-zd-* tends to be formed in *-ks-*, regardless of the historical source of *zd*. Thus, Doric shows an aorist participle *kathíksas* (καθίξας) for the present *kathízdō* (καθίζω “I set, sit”), from the root \**sed-*.

#### 4.2.4 Perfect tense stems

Among dialectal peculiarities in perfect tense morphology, notable is the occurrence of thematic inflection of the perfect indicative which is attested in Sicilian Doric and in the Doric of Rhodes and neighboring regions. For example, the Syracusan author Theocritus uses perfects such as *dedoikō* (δεδοίκω “I fear”) and *pepónth<sup>h</sup>eis* (πεπόνθεις “you have suffered”). Compare – with perfect endings (see Ch. 24, §§4.2.7, 4.2.9) – the respective Attic forms *dēdoika* (δέδοικα) and *pépont<sup>h</sup>as* (πέπονθας). In Lesbian and some Doric and Northwest Greek dialects, the perfect infinitive is formed with the thematic formant *-ēn* (-ειν) rather than athematic *-enai* (-εναι). The Aeolic dialects form the perfect participle with thematic formants (see Ch. 24, §4.2.11).

#### 4.2.5 Nonindicative moods

Outside of the indicative mood, several dialectal variants can be noted.

##### 4.2.5.1 Subjunctive mood

In several dialects – such as Anatolian Ionic, Lesbian and Cretan – the *s*-aorist subjunctive is attested as a “short vowel subjunctive” (i.e., is formed with the short vowel suffix *-e/o-*; see Ch. 24, §4.2.8.1). Compare Ionic *poiĕ-se-i* (ποιήσει) and Attic *poiĕ-sē-i* (ποιήσῃ “may (s)he make”), a “long vowel subjunctive.”

##### 4.2.5.2 Imperative mood

Throughout the Greek dialects, there is extensive variation in the inflection of the third-person imperative. The Attic ending *-ntōn* (-ντων; see Ch. 24, §4.2.8.3) also occurs, among other dialects in Ionic and Cretan. Two other third plural endings are essentially substrating components of the formant *-ntōn*: *-ntō* (-ντω) occurs in Arcadian, Boeotian, and various Doric dialects; *-tōn* (-των) is found in Ionic. A short vowel variant *-nton* (-ντον) is used in Lesbian and also occurs in Pamphylian.

#### 4.2.6 Verb endings

For the verb endings of Attic, see Chapter 24, §4.2.9. In those Greek dialects which preserve *\*t* before *i* (see §3.4.3, 8), the athematic active third singular ending is *-ti* (-τι), rather than *-si* (-σι). In the same way, *\*-ti-* is preserved in the third plural: for example, Doric thematic *p<sup>h</sup>éronti* (φέροντι “they carry”), athematic *tí<sup>h</sup>enti* (τίθεντι “they place”).

In Doric and Northwest Greek the ending of the active first plural is *-mes* (-μες), rather than the *-men* (-μεν) of Attic and other dialects. This ending *-mes* finds cognates in Sankrit *-mas* and Latin *-mus*.

The middle third singular ending – *-tai* (-ται) in Attic and most other dialects – appears as *-toi* (-τοι) in Mycenaean and Arcadian.

#### 4.2.7 Infinitives

The Attic thematic active infinitive formant *-ēn* (-ειν; see Ch. 24, §4.2.10) or its variant *-ēn* (-ην, see §3.4.4) occurs in many other dialects as well – such as Ionic and certain Doric dialects (*-ēn*), and Lesbian, Laconian, and Elean (*-ēn*). A short vowel form *-en* (-εν) is found in Arcadian and various Doric dialects. The athematic infinitival formant *-nai* (-ναι) of Attic is also found in Ionic and Arcado-Cypriot; Lesbian uses *-menai* (-μεναι), while Boeotian,

Thessalian, Northwest Greek, and most Doric dialects have *-men* (which Boeotian and Thessalian also use with thematic verbs). Long vowel variants of the last-named occur in Cretan and Rhodian. On the perfect infinitive, see §4.2.4.

### 4.3 Numerals

Most, though not all, of the dialectal variations seen in the forms of numerals, vis-à-vis their Attic counterparts (see Ch. 24, §4.5), are the result of dialect sound changes. Selected examples are presented below:

(6)

- 1 Cretan éns (ἐνς, masc.), Aeolic ía (ἰα, fem.)
- 2 Laconian dúe (δύε), West Ionic dúwo (δύϜο)
- 3 Cretan trées (τρέες), Heracleian trís (τρίς)
- 4 Ionic tésseres (τέσσερες), Lesbian péssures (πέσσυρες), Boeotian péttares (πέτταρες)
- 5 Lesbian pémpre (πέμπε), Pamphylian péde (πέδε)
- 6 Cretan wéks (Ϝέξ)
- 7 Cretan ettá (ἐττά)
- 8 Lesbian ókto (ὄκτο), Elean optó (ὀπτό)
- 9 Heracleian hennéa (ἡεννέα)
- 10 Arcadian déko (δέκο)

## 5. SYNTAX

### 5.1 Dialectal syntactic features

The syntactic variation attested between dialects – to the extent that such variation can be or has been discerned – is quite minor and lexically specific. Most examples are so much so that they do not fall within the purview of the present work. A few prominent morphosyntactic isoglosses are noted below.

### 5.2 Coordination

In place of the pandialectal tonic conjunction *kai* (καί), Arcado-Cypriot uses the conjunction *kás* (κάς). Thessalian utilizes a particle *má* (μά) in lieu of *dé* (δέ).

### 5.3 Conditional clauses

The Attic conditional particle *án* (ἄν; see Ch. 24, §5.6) also occurs in Ionic as well as in Arcadian. In Lesbian, Thessalian, and Cypriot, a particle *ke* (κε) is used instead; while a form *ka* (κα) occurs in Boeotian, Doric, and Northwest Greek.

## 6. LEXICON

Making allowance for dialect-specific phonological and morphological variation, a great part of ancient Greek vocabulary is common to all dialects. Yet, with even a casual perusal of a comprehensive dictionary of ancient Greek, such as Liddell and Scott's *Greek-English*

*Lexicon* (1996), one cannot help but be impressed by how many words appear to be limited to a particular dialect or set of dialects. The lexical distribution that such an activity suggests is to some extent illusory, of course, owing to the haphazard nature of the survival of ancient Greek documents – had a greater, more evenly distributed body of material survived, many words would certainly be found to have a broader dialectal distribution. Apt testimony of this is provided by the Mycenaean vocabulary. Prior to the decipherment of Linear B, a number of lexemes which would emerge from the Bronze-Age Mycenaean tablets were only attested in relatively late, post-Classical Greek sources (see Ventris and Chadwick 1973:91). On the other hand, chances are that more extensive documentation would also reveal yet more dialect-limited vocabulary.

The Mycenaean lexicon also contains, expectedly, both vocabulary that is not otherwise attested in ancient Greek and vocabulary preserved in the archaic poetic language of Greek epic. The agent noun *to-ko-do-mo* (*toik<sup>h</sup>odomoi*) “builders,” for example, morphologically and semantically transparent, is not found elsewhere, though a denominative verb seemingly derived from it is attested in the fourth century BC. Among words which Mycenaean shares with Homeric epic are *pa-ka-na* (pl.), *p<sup>h</sup>ásganon* (φάσγανον “sword, dagger”); *e-ke-si* (dat. pl.), *ép<sup>h</sup>os* (ἐγχος “spear”); and *a-sa-mi-to*, *asáminthos* (ἀσάμινθος “bathing tub”).

The Arcado-Cypriot lexicon likewise contains numerous archaic words, shared with Mycenaean Greek and the language of Homer and poetry. In Cyprus, the Mycenaean word for “king,” *wanaks*, still survives as a royal title (Cypriot *wa-na-kse*), denoting the king’s sons and brothers; elsewhere in the first millennium BC, *ánaks* (ἄναξ) commonly means “lord” or “master of the house.” The notion “king” has come to be expressed by *basileús* (βασιλεὺς), which in Mycenaean (*g<sup>w</sup>asileus*) names, much more modestly, the “chief.” Among other words which Cypriot shares with epic are *e-le-i* (dative of *élos* “meadow”), Homeric ἔλος (cf. the Mycenaean place name *e-re-e/i*, dative); *-i-ja-te-ra-ne* (accusative of *ijātēr* “healer”); Homeric ἱητήρ (Mycenaean *i-ja-te*). Arcadian shares with the epic language, *inter alia*, *kéleut<sup>h</sup>os* (κέλευθος “path”) and *ámar* (ἄμαρ “day”), Homeric ἡμαρ (cf. Armenian *awr*). Archaic words shared by Arcadian and Cypriot include, among others, *euk<sup>h</sup>ōlā* (εὐχολᾶ “prayer”; Cypriot spelling *e-u-ko-la*), Homeric εὐχολή.

Interesting among dialect-specific lexemes are names of legal and religious officials. To cite but a few examples, Lesbian provides *dikáskopoi* (δικάσκοποι), the title of judges at Mytilene and Cyme (“inspectors of justice”); Thessalian has *tagós* (ταγός), the title of a magistrate at Larissa (the word more widely denotes “commander”). Among Northwest Greek dialects, Locrian shows *pentámeroi* (πεντάμεροι), officials who serve for five (*pénte*, πέντε) days (*améra*, ἄμέρα); compare the Phocian verb *pentamariteúō* (πενταμαριτεύω) “to hold office for five days” (*amára*).

In Laconian, the title of the office of overseer is *bíduoi* (βίδυοι) or *bídeoi* (βίδεοι), from the root *\*wid-* “to see,” evidencing the Spartan fricativization of the glide *\*w* (see §3.1.1). The regimentation of Spartan society with its grouping of boys and young men by age for military training and common life finds expression in the Laconian lexicon, producing words such as *pratopámpais* (πρατοπάμπαις), from *prato-* (πρατο-) “first,” *pân* (πᾶν) “all” and *país* (παῖς) “child, boy”; and *hatropámpais* (ἄτροπάμπαις), perhaps from *háteros* (ἄτερος) “another, second” – both denoting such a group of boys.

The Doric dialect of Coan preserves the title of a priest of Cos called the *gereaphóros* (γερεαφόρος), meaning approximately the “recipient of perks,” from *géras* (γέρας) “gift of honor, present.” Particularly intriguing is the title of a scribe preserved in a remarkable Cretan inscription, the *poinikastás* (ποινικαστάς; see Jeffery and Morpurgo Davies 1970; Thomas 1992:69–70); he (Spensithios is his name) is the scribe who writes with *p<sup>h</sup>oinikḗia* (φοινικήια) “Phoenician letters,” the term the Greeks use to denote the characters of their



alphabet, suggesting an awareness at the time of coining of another type of Greek character (see Burkert 1992:27). A natural further implication may be that the *poinikastás* is the scribe who wrote with Phoenician characters as opposed to a scribe using a different Greek script (see Woodard 2000) – perhaps a particularly archaic title preserved by the scribe's cultic affiliation (Spensithios is appointed to be scribe of both secular and sacred matters). Jeffery and Morpurgo Davies (1970:152) have drawn attention to the use of a similar scribal title *phoinikográphos* (φοινικογράφος) in the Aeolic dialect of Mytilene, attested alongside the common term for scribe *grammateús* (γραμματεús), suggesting that the former may be an old title preserved because of the scribe's affiliation with the cult of Hermes.

## 7. READING LIST

Though rendered somewhat out of date by the absence of Mycenaean Greek data, the most helpful English treatment of the Greek dialects remains Buck's excellent 1955 volume. Helpful summaries of the Greek dialects, including Mycenaean, appear in Palmer 1980. A detailed treatment of Greek dialects is provided by the two revised volumes of Thumb: Thumb–Kieckers 1932 and Thumb–Scherer 1959. Cowgill 1966 is an excellent summary of more recent work in Greek dialectology, including a review of the seminal studies by Porzig (1954) and Risch (1955). For Mycenaean Greek, see especially Ventris and Chadwick 1973. The standard English dictionary of ancient Greek is Liddell, Scott, Stuart Jones, and McKenzie 1996. On the Mycenaean lexicon see Aura Jorro 1985–1993.

## Bibliography

- Allen, W. 1973. *Accent and Rhythm*. Cambridge: Cambridge University Press.  
 ———. 1987. *Vox Graeca* (3rd edition). Cambridge: Cambridge University Press.  
 Aura Jorro, F. 1985–1993. *Diccionario micénico*. Madrid: Instituto Filología.  
 Bile, M. 1988. *Le dialecte crétois ancien*. Athens: Ecole Française.  
 Browning, R. 1983. *Medieval and Modern Greek* (2nd edition). Cambridge: Cambridge University Press.  
 Buck, C. 1955. *The Greek Dialects*. Chicago: University of Chicago Press.  
 Burkert, W. 1992. *The Orientalizing Revolution*. Translated by M. Pindar and W. Burkert. Cambridge, MA: Harvard University Press.  
 Chadwick, J. 1956. "The Greek dialects and Greek pre-history." *Greece and Rome* NS 3:38–50.  
 ———. 1976. "Who were the Dorians?" *Parola del passato*. 31:103–117.  
 Chantraine, P. 1984. *Morphologie historique du grec* (2nd edition). Paris: Klincksieck.  
 Coleman, R. 1963. "The dialect geography of ancient Greece." *Transactions of the Philological Society*, pp. 58–126.  
 Cowgill, W. 1966. "Ancient Greek dialectology in the light of Mycenaean." In H. Birnbaum and J. Puhvel (eds.), *Ancient Indo-European Dialects*, pp. 77–95. Berkeley/Los Angeles: University of California Press.  
 Devine, A. and L. Stephens. 1994. *The Prosody of Greek Speech*. Oxford: Oxford University Press.  
 Duhoux, Y. 1983. *Introduction aux dialectes grecs anciens*. Louvain-la-Neuve: Cabay.  
 Heubeck, A. 1986. "Die Würzburger Alphabettafel." *Würzburger Jahrbücher für die Altertumswissenschaft* NS 12:7–20.  
 Jeffery, L. 1990. *The Local Scripts of Archaic Greece* (revised edition). Supplement by A. Johnston. Oxford: Oxford University Press.  
 Jeffery, L. and A. Morpurgo Davies. 1970. "ΠΟΙΝΙΚΑΣΤΑΣ and ΠΟΙΝΙΚΑΖΕΝ: BM 1969. 4–2. 1, a new archaic inscription from Crete." *Kadmos* 9:118–156.  
 Kirchhoff, A. 1887. *Studien zur Geschichte des griechischen Alphabets*. Gütersloh: Bertelsmann.



- Lejeune, M. 1982. *Phonétique historique du mycénien et du grec ancien*. Paris: Klincksieck.
- Liddell, H. and R. Scott. 1996. *A Greek-English Lexicon with a Revised Supplement*. Revised by H. Stuart Jones and R. McKenzie. Oxford: Clarendon Press.
- Masson, O. 1983. *Les inscriptions chypriotes syllabiques*. Paris: Edition E. de Boccard.
- Nagy, G. 1968. "On dialectal anomalies in Pylia texts." In *Atti e memorie del 1° congresso internazionale di micenologia*, vol. II, pp. 663–679. Rome: Edizioni dell'Ateneo.
- . 1970. *Greek Dialects and the Transformation of an Indo-European Process*. Cambridge, MA: Harvard University Press.
- Palmer, L. 1980. *The Greek Language*. Atlantic Highlands, NJ: Humanities Press.
- Porzig, W. 1954. "Sprachgeographische Untersuchungen zu den altgriechischen Dialekten." *Indogermanische Forschungen* 61:147–169.
- Risch, E. 1955. "Die Gliederung der griechischen Dialekte in neuer Sicht." *Museum Helveticum* 12:61–76.
- . 1966. "Les différences dialectales dans le mycénien." In L. Palmer and J. Chadwick (eds.), *Proceedings of the Cambridge Colloquium on Mycenaean Studies*, pp. 150–157. Cambridge: Cambridge University Press.
- Rix, H. 1976. *Historische Grammatik des Griechischen*. Darmstadt: Wissenschaftliche Buchgesellschaft.
- Scott, D., R. Woodard, P. McCarter, et al. 2003. "Two ancient Greek copper plaques: Schøyen MS 108." In J. Braarvig (ed.), *Manuscripts in the Schøyen Collection (MSC)*, vol. IV. Oslo: Hermes Publications.
- Stephens, L. and R. Woodard. 1986. "The palatalization of the labiovelars in Greek: a reassessment in typological perspective." *Indogermanische Forschungen* 91:129–154.
- Szemerényi, O. 1968. "The Attic 'Rückverwandlung' or atomism and structuralism in action." In M. Mayrhofer (ed.), *Studien zur Sprachwissenschaft und Kulturkunde*, pp. 139–157. Innsbruck: Leopold-Franzens-Universität.
- Thomas, R. 1992. *Literacy and Orality in Ancient Greece*. Cambridge: Cambridge University Press.
- Threatte, L. 1980. *The Grammar of Attic Inscriptions*, vol. I. Berlin: De Gruyter.
- Thumb, A. and E. Kieckers. 1932. *Handbuch der griechischen Dialekte*, part 1. Heidelberg: Carl Winter.
- Thumb, A. and A. Scherer. 1959. *Handbuch der griechischen Dialekte*, part 2. Heidelberg: Carl Winter.
- Ventris, M. and J. Chadwick. 1973. *Documents in Mycenaean Greek* (2nd edition). Cambridge: Cambridge University Press.
- Woodard, R. 1986. "Dialectal differences at Knossos." *Kadmos* 25:49–74.
- . 1997. *Greek Writing from Knossos to Homer*. Oxford: Oxford University Press.
- . 2000. "Greek–Phoenician interaction and the origin of the alphabet." In A. Ovdiah (ed.), *Mediterranean Cultural Interaction*, pp. 33–51. Tel Aviv: Tel Aviv University Press.

# Sanskrit

STEPHANIE W. JAMISON

## 1. HISTORICAL AND CULTURAL CONTEXTS

Sanskrit is an Indo-European language, a member of the Indo-Aryan branch of the Indo-Iranian subgroup of that family. It is chronologically and in terms of linguistic development the “oldest” Indo-Aryan language and consequently often referred to as *Old Indic* (Altindisch) or *Old Indo-Aryan*; its descendants include a range of linguistic varieties classified under the rubric Middle Indic (or Prākṛit, see Ch. 27), as well as the Modern Indic (New Indo-Aryan) languages spoken today, such as Hindi, Gujarati, Bengali. It is not related genetically to the Dravidian languages of South India, such as Tamil and Telugu.

The oldest form of Sanskrit is so-called *Vedic Sanskrit*, the language of the four collections of liturgical texts known as the Vedas and of the early exegetical literature on these texts. The oldest Veda is the *Ṛgveda* (*Rig-veda*), a compilation of 1,028 hymns which took shape around 1500 BC in northwest India, though the composition and collection of hymns clearly occupied several centuries. In language, style, and phraseology the *Ṛgveda* resembles the earliest texts of its closest linguistic relative, the *Gāthās* attributed to the prophet Zarathustra, composed in Old Avestan (see Ch. 29).

Though the composition of Vedic texts can be dated with fair confidence to the period of c. 1500–500 BC, direct records of them are only found several millennia later. The “texts” were transmitted orally, with minimal alteration, and even after they were also committed to writing, the manuscripts were perishable and less reliable than the oral tradition.

Through the approximately thousand years of Vedic textual composition, the language shows gradual changes, especially in the loss of certain grammatical categories and the reduction of variant forms. Around 500 BC the Sanskrit then current among cultivated speakers received a magnificent description by the grammarian Pāṇini in his treatise, the *Aṣṭādhyāyī* (“[Work] consisting of eight chapters”), whose level of detail and theoretical sophistication has not been equaled to this day.

Pāṇini inadvertently froze the language in this particular form forever. What was composed as a *descriptive* grammar (though descriptive of a geographically and socio-culturally limited speech form, not the speech of the whole society) became a *prescriptive* grammar of a learned language. All subsequent Sanskrit follows, or attempts to follow, the rules of Pāṇini. Though there are systematic variations in later texts, these are essentially stylistic and distributed according to textual genre. The language of the great epics, the *Mahābhārata* and the *Rāmāyaṇa*, deviates somewhat from the Pāṇinian norm and is therefore sometimes distinguished as *Epic Sanskrit*; it displays some Middle Indic tendencies. Inscriptional Sanskrit also commonly shows nonsanctioned forms. Despite these minor exceptions, Sanskrit no longer had a history in the accepted linguistic sense of this term – even though the greater

part of its literature remained to be composed. The great flourishing of Sanskrit literary production lasted through the first millennium AD.

The language as fixed by Pāṇini is commonly known as *Classical Sanskrit*, or *Sanskrit* proper. Indeed, the term *saṃskṛta* means “perfected” and refers to the language generated according to Pāṇini’s rules, as opposed to the vernacular Prākṛits, from *prākṛta* “natural, unrefined.” Strictly speaking, the pre-Pāṇinian language of the Vedic texts is not “Sanskrit,” and is sometimes called simply *Vedic*, rather than Vedic Sanskrit. In this work, however, *Sanskrit* will denote all varieties of Old Indic.

## 2. WRITING SYSTEM

The earliest Sanskrit texts were composed and transmitted orally, not written down for centuries after their first “attestation.” Indeed, the first documentary evidence of Indo-Aryan languages in the Indian subcontinent comes not from Old Indic but Middle Indic: the inscriptions of the ruler Aśoka in the third century BC (see Ch. 27, §1.1) The first direct attestation of Sanskrit comes from around the beginning of the present era. The first extensive inscription is that of the ruler Rudradāman c. AD 150 at Girnar in western India; the first extant manuscripts, found in central Asia, date from about the same period.

The writing system found in most of the early inscriptions is *Brāhmī* (another, less widespread system, *Kharoṣṭhī*, an adaptation of Aramaic, is found in the northwest, already in the Aśokan edicts). *Brāhmī* seems to have been adapted from a Semitic writing system, though the exact details are unclear, as is the date of its introduction into India, a subject of much controversy. *Brāhmī* is the ancestor of most of the writing systems used in India.

Until the advent of printing and the regular publication of Sanskrit texts, Sanskrit manuscripts were written in various local scripts. Now Sanskrit is almost exclusively printed in a script known as *Nāgarī* or *Devanāgarī*, a medieval offshoot of *Brāhmī*, and perfectly adapted to the writing of Sanskrit, with a one-to-one correspondence between sound and symbol. The conventional transcription of *Devanāgarī* into Roman characters was established finally at the Tenth Congress of Orientalists, 1894. Transliterations in works published before often show deviations from the modern norm.

The system can be considered a modified or pseudo-syllabary in that each consonantal symbol represents a consonant with following short *a*-vowel (the commonest vowel in the language), for example, क = *ka*, ख = *kha*, ग = *ga*, घ = *gha* (not *k*, *kh*, *g*, *gh*); see Table 26.1. However, unlike “pure” syllabaries, a different symbol is not necessary to represent consonants followed by other vowels (e.g., *kā*, *ki*, *kī*, etc.). Instead, a set of universally applicable diacritics can be used to cancel the inherent short *a* and substitute a different following vowel: thus, का = *kā*, कि = *ki*, कु = *ku*, and so forth. There are also separate signs for independent vowels, for example, अ = *a*, ए = *e*.

Another drawback of some syllabaries, the inability to represent consonant clusters unambiguously, is overcome by the system of ligatures. Portions of each consonant in a cluster are combined into a single conventional sign, for example, त (ta) + क (ka) = त्क (tka). Final consonants can also be represented, by a stroke (*virāma*) under the sign, which cancels the short *a*: thus त = *ta*, but त् = *t*. Thus, the system combines the flexibility of an alphabet with some of the spatial economy of a syllabary.

*Devanāgarī* writing of Sanskrit lacks word divisions. Each linguistic string, regardless of morphosyntactic structure, is treated as a sequence of syllables (*akṣaras*) consisting of onset

**Table 26.1 The Devanāgarī script***Vowel symbols*

a	ā	i	ī	u	ū	ṛ	ṝ	ḷ
अ	आ	इ	ई	उ	ऊ	ऋ	ॠ	ऌ
e	ai	o	au					
ए	ऐ	ओ	औ					

*Consonant + vowel symbols*

ka	kha	ga	gha	ṇa
क	ख	ग	घ	ङ
ca	cha	ja	jha	ṇa
च	छ	ज	झ	ञ
ṭa	ṭha	ḍa	ḍha	ṇa
ट	ठ	ड	ढ	ण
ta	tha	da	dha	na
त	थ	द	ध	न
pa	pha	ba	bha	ma
प	फ	ब	भ	म
ya	ra	la	va	
य	र	ल	व	
śa	ṣa	sa	ha	
श	ष	स	ह	

*Sample vowel diacritics*

kā	ki	kī	ku	kū
का	कि	की	कु	कू
kṛ	ke	kai	ko	kau
कृ	के	कै	को	कौ

consonant(s) (if present) plus vowel. Thus, a string like *tad etad rūpam*, with word divisions as given in transliteration, would obligatorily appear in Devanāgarī as *ta de ta drū pa m* त दे त द्रूपम् (though without spaces between the characters).

### 3. PHONOLOGY

#### 3.1 Diachronic overview

From the point of view of reconstructed Proto-Indo-European, the most important phonological development in Sanskrit (and indeed in Indo-Iranian) is vowel-merger: short \*e, \*o, and \*a all merge as a; long \*ē, \*ō, \*ā (and short \*o under certain conditions) merge as ā. Since much of Proto-Indo-European morphology was based on alternations of vowels with \*e-timbre and those with \*o-timbre (qualitative ablaut), these mergers had major effects on the morphological system.

On the other hand, Sanskrit maintained the Proto-Indo-European consonantal system with some fidelity, only enlarging its inventory. The three series of stops – voiceless (T), voiced (D), and voiced aspirated (Dh) – traditionally reconstructed remain in Sanskrit, and

a fourth was added, voiceless aspirated (Th). As in other satem languages the labiovelars merged with the plain velars. There was secondary palatalization of the resulting segments, reflected in thoroughgoing synchronic alternations within Sanskrit (see §3.4.2.2). Otherwise, the inventory of places of articulation was increased by the creation of a series of retroflex dental stops. For the comparatist an especially important retention in Sanskrit is the preservation of \**ɣ*, \**w*, and \**s* intervocally, thus avoiding the loss of morphological clarity attendant on vowel contraction that bedevils the historical linguist in languages like Greek.

### 3.2 Vowels

The cardinal vowels *i*, *u*, *a* distinguish length; in addition, short *a* is a closer vowel than *ā*, equivalent to schwa. The mid vowels *ē* and *ō*, as monophthongizations of the Indo-Iranian diphthongs \**ai* and \**au* (preserved in Iranian), are inherently long and are so marked in the phonological sections of this work, though they are not usually so transcribed. The true diphthongs *āi* and *āu* (usually now transcribed simply *ai* and *au*) also count as long. The vocalic liquid *ṛ* represents a merger of PIE (Proto-Indo-European) \**r̥* and \**l̥*. However, long *ṝ* is an invention of the system and found in a few analogically generated morphological categories; PIE \**r̥* has different, biphonemic outcomes in Sanskrit, as we will see. Vocalic *ḷ* is even more limited, found in only one morpheme.

#### (1) Sanskrit vowel phonemes

monophthongs:	i / ī ē	u / ū ō	diphthongs:	āi	āu
	a ā		vocalic liquids:	ṛ / ṝ	ḷ

### 3.3 Consonants

The consonantal inventory of Sanskrit is presented in Table 26.2:

Table 26.2 The consonantal phonemes of Sanskrit						
Manner of articulation	Place of articulation					
	Labial	Dental	Retroflex	Palatal	Velar	Glottal
Stops and affricates						
Voiceless	p	t	ṭ	c	k	
Voiceless aspirated	ph	th	ṭh	ch	kh	
Voiced	b	d	ḍ	j	g	
Voiced aspirated	bh	dh	ḍh	jh	gh	
Nasals	m	n	ṇ	ñ	ṅ	
	ṁ anusvāra (see below)					
Fricatives						
Voiceless		s	ṣ	ś		h visarga
Voiced						h
Liquids		l	r			
Glides	v			y		

The apparent symmetry of this consonantal system conceals some failures of parallelism in distribution, often the results of historical changes:

1. The voiceless aspirated series is an addition to the system and significantly rarer than the other three. It is often found in etymologically obscure words.
2. The retroflex sibilant  $\mathring{s}$  is the automatic product of dental  $s$  following  $i$ ,  $u$ ,  $r$ , and  $k$  (mnemonically “ruki”), a process also found not only in Iranian but in part in Balto-Slavic.
3. The series of retroflex stops was a creation of Indic, in most cases as a conditioned result of regressive assimilation to ruki $\mathring{s}$  and therefore distributionally limited; in particular, initial retroflexes are almost never found. The retroflex nasal is ordinarily the automatic product of dental nasal when the word contains a preceding  $r$  (subject to some conditions). Thus, all the retroflexes are in origin conditioned alternants of dentals, though from the beginning of the language they have a qualified independence.
4. The palatals are affricates, not stops. In the palatal row the voiced aspirate  $jh$  is a new and extremely rare phoneme; the phoneme patterning with the palatals as the voiced aspirate for morphophonemic purposes is glottal  $h$  (see §3.4.2.1).
5. The palatal nasal is a conditioned variant of  $n$  occurring next to palatal obstruents; the velar nasal is also ordinarily a conditioned product of  $n$ , found before velar stops, but further phonological developments (loss of final or cluster-internal velar stop) can allow the velar nasal an independent if marginal existence. *Anusvāra* is a conditioned alternant of postvocalic nasals, under certain sandhi conditions.
6. *Visarga* is a word-final (sometimes morpheme-final) conditioned alternant of  $s$  and  $r$  under certain sandhi conditions.
7. The glides and liquids regularly alternate with vowels:  $i \approx y$ ;  $u \approx v$  ( $[w]$ );  $\mathring{r} \approx r$ ;  $\mathring{l} \approx l$  (under conditions discussed below).

### 3.4 Phonological alternations

Sanskrit is characterized by a pervasive series of phonological alternations occurring on several different linguistic levels and displaying varying degrees of transparency. We begin with the most transparent.

#### 3.4.1 External sandhi

The surface form of any linguistic string is subject to phonological rules of combination (*sandhi* or “putting together”). In other words, phenomena of the English *gonna* (from *going* + *to*) type apply to any two words in contact within a sentence, and even between sentences in a discourse. Most sandhi rules involve regressive assimilation, especially in voicing: for example, (with underlying *tad*) *tad bhavati* but *tat phalam*. Assimilation in manner of articulation is also met with (e.g., *tan manas*). Like vowels coalesce into a single long vowel (e.g., *vada* + *agne*  $\Rightarrow$  *vadāgne*), and unlike vowels undergo diphthongization or glide-formation (e.g., *vada oṣadhe*  $\Rightarrow$  *vadauṣadhe*; *asti agniḥ*  $\Rightarrow$  *asty agniḥ*). Despite the simplicity of the principles, the details of sandhi rules are sometimes opaque. For example, though the change of final *-as* to *-o* before voiced sounds historically involves regressive voicing assimilation, this process is not synchronically transparent. The rules of external sandhi ordinarily apply also at compound seams, and many but not all of the same rules at morpheme boundaries.

External sandhi in Vedic is more variable than in Classical Sanskrit, not only in the form of the rules but also in their application (or nonapplication). Sandhi in Middle Indic occurs only under conditions of close syntactic nexus. Given these facts, it seems likely that the pervasive system of obligatory sandhi characteristic of Classical Sanskrit involved an artificial imposition of an originally more flexible set of processes linking words within syntactically defined phrases.

### 3.4.2 Internal consonantal alternations

The rest of this section presupposes the concept of the *root* and the canonical structure of the Sanskrit word presented in §4.1.

#### 3.4.2.1 Voicing and aspiration

The voiceless, voiced, and voiced aspirated obstruents of a positional series regularly alternate with each other ( $p \approx b \approx bh$ ;  $t \approx d \approx dh$ , etc.; note, however,  $c \approx j \approx h$ ), such that, for example, a morpheme with an underlying voiced aspirate final may show alternants with all three stops under differing internal sandhi conditions: thus,  $\sqrt{buddh}$  “be aware” – *buddh-yate*, *bud-dha-*, *bhot-syate*.

Clusters containing unaspirated stops show regressive assimilation (e.g., *chit-ti-* from *\*chid + ti-*). But in those containing voiced aspirates the resulting cluster is both voiced and aspirated whatever the position of the aspirate in the underlying cluster (hence *buddha-* from *buddh-ta-*) – the change known as *Bartholomae’s Law*. In summary,

$$\begin{array}{lll}
 (2) \quad \text{A. } T + T & \text{B. } T + D & \text{C. } Dh + T \\
 & \Rightarrow T-T & \Rightarrow D-D \\
 D + T & D + D & Dh + D \Rightarrow D-Dh \\
 & & Dh + Dh
 \end{array}$$

Before *s* all stops become voiceless; hence *bhot-syate* above. This same form illustrates another, sporadic alternation: when roots with underlying final aspirates lose that aspiration, the initial consonant often acquires aspiration (hence *bhot-syate*, but *buddh-yate*). This represents a reconfiguring or reversal of a historical development – *Grassmann’s Law*, whereby di-aspirate roots dissimilated the first aspirated stop.

#### 3.4.2.2 Velars and palatals

The velar series (*k, g, gh*) regularly alternates with the palatal series (*c, j, h*). In particular, velar-initial roots reduplicate with palatals (e.g.,  $\sqrt{kṛ}$ : *ca-kāra*;  $\sqrt{gam}$ : *ja-gāma*); and preobstruent velars alternate with palatals in other phonological positions (e.g.,  $\sqrt{muc}$ : *muk-ta*, but *muc-yate*). This alternation is the historical result of a pan-Indo-Iranian palatalization of velars by following front vowels (and *y*), the conditioning of which was obscured by the subsequent merger of *\*e* with *\*a* and *\*o* noted above.

#### 3.4.2.3 Palatals and retroflexes

The structural position of the palatal series was further complicated by a different merger. Though Sanskrit *c* can only be the product of an old palatalized velar (*\*k<sup>(w)</sup>e*, etc.), both *j* and *h* have two sources: (i) not only palatalized velars (*\*g<sup>(w)</sup>e*, *\*g<sup>(w)h</sup>e*, etc.); (ii) but also PIE palatal stops (*\*ǵ* and *\*ǵ<sup>h</sup>*), whose voiceless equivalent (*\*ǵ̊*) yields Sanskrit *ś*. These underlying palatals enter into a set of synchronic alternations different from that of the old velars: palatals followed by dentals produce a retroflex cluster, for example,  $\sqrt{sṛj}$  “emit”: *sṛj + ta*  $\Rightarrow$  *sṛṣ-ṭa*. Thus, though the phonetic inventory of the language contains only a single palatal series, morphological alternations define two morphophonemically distinct series: (3A) *ś, j, h*; and (3B) *c, j, h*.



- (3) A. *palatals* ( $\approx$  *retroflexes*) B. *palatalized velars* ( $\approx$  *velars*)  
           ś (e.g., viś : viṣ-ṭa)                      c (e.g., muc : muk-ta)  
           j (e.g., sṛj : sṛṣ-ṭa)                      j (e.g., bhaj : bhak-ta)  
           h (e.g., ruh : rū-dha)                      h (e.g., snih : snig-dha)

(By Bartholomae's Law, a compensatorily lengthened vowel plus retroflex *ḍh* is the regular outcome in *rūḍha*.)

The distinction between these two series is neutralized before *s*, where both series (and all three manners) show *k*: for example, both *ruruk-ṣati* ( $\sqrt{\text{ruh}}$ ) and *sisnik-ṣati* ( $\sqrt{\text{snih}}$ ).

### 3.4.3 Internal vocalic alternations

The Sanskrit morphological system is pervaded by vocalic alternations, conveniently considered as the “strengthening” of an underlying vocalic element by two successive additions of the vowel *a* ( $V / a + V / a + aV$ ). The preconsonantal versions of these strengthenings are known by the indigenous terms *guṇa* and *vṛddhi*, but it is useful to consider these in conjunction with their prevocalic alternants as well. In terms familiar from Indo-European descriptive grammars, the unstrengthened state corresponds to *zero-grade*, *guṇa* to *full-* (or *normal-*) grade, and *vṛddhi* to *extended-* (or *lengthened-*) grade. Though Proto-Indo-European *qualitative* ablaut essentially disappeared in Indo-Iranian with the merger of *\*e* and *\*o*, *quantitative* ablaut is transparently continued by the Sanskrit system of vowel strengthening. Alternations between zero-grade and full-grade are prominent in the morphological system; *vṛddhi* is especially important in the derivation of adjectives of origin and appurtenance (*vṛddhi derivatives*).

The alternations between consonantal and vocalic versions of glides and liquids are also relevant here, and the system is in fact clearest with these segments, especially *ṛ*, where the successive additions of *a* are easily discerned (N.B. for ease of exposition, *ī* and *ū* are not included here, but will be discussed below):

(4)	<u>Zero-grade</u>		<u>Full-grade</u>		<u>Extended-grade</u>	
	PreC.	PreV.	PreC. ( <i>guṇa</i> )	PreV.	PreC. ( <i>vṛddhi</i> )	PreV.
ṛ	ṛ	r	ar	ar	ār	ār
i	i	y	ē	ay	āi	āy
u	u	v	ō	av	āu	āv
a	a		a		ā	
ā	ā		ā		ā	

As can be seen, with the simple vowels *a* and *ā*, the progressive addition of *a* is not so clear; moreover, prevocalic position is subject to complications.

Though Sanskrit does not have surface syllabic nasals (*\*ṁ* and *\*ṇ*) as reconstructed for Proto-Indo-European, the parallelism of morphological alternations compels us to posit such underlying vowels, which fit into the vowel gradation system as follows:

(5)	<u>Zero-grade</u>		<u>Full-grade</u>		<u>Extended-grade</u>	
	PreC.	PreV.	PreC. ( <i>guṇa</i> )	PreV.	PreC. ( <i>vṛddhi</i> )	PreV.
<i>*ṁ</i>	a	m	am	am	ām	ām
<i>*ṇ</i>	a	n	an	an	ān	ān

The following chart gives an example of a root with each vocalism (save for the *a*-vowels), with representative forms from the various categories:

(6)		<i>r̥</i> kṛ “do, make”	<i>i</i> ji “conquer”	<i>u</i> su “press”	* <i>m̐</i> gam “go”	* <i>ṇ̐</i> han “smash”
Zero:	PreC.	kṛ-ta	ji-ta	su-ta	ga-ta	ha-ta
	PreV.	cakr-ur	jigy-ur	suṣv-āna	jagm-ur	jaghn-ur
Full:	PreC.	kar-tum	jē-tum	sō-tum	gan-tum	han-tum
	PreV.	akar-am	jay-ati	asūṣav-at	agam-am	ahan-am
Ext.:	PreC.	kār-ya	ajāi-s	asāu-ṣīt	—	—
	PreV.	cakār-a	ajāy-i	asāv-i	jagām-a	jaghān-a

### 3.5 Syllable structure and phonotaxis

There are few constraints on syllable structure. Syllables may both begin and end with vowels, single consonants, or consonant clusters; and internal vowels may be of any weight, even before coda consonants. In Vedic, however, some traces of phonological processes (*Sievers–Edgerton Law*) seemingly function to avoid overlong syllables: some suffixes containing *y* or *v* must be read as *iy* and *uv* after heavy syllables, but *y* and *v* after light syllables. But this is a morphologically limited phenomenon, not a pervasive phonological rule.

There are constraints on word-final consonants, which apply before external sandhi rules operate. Final clusters are not allowed (though monomorphemic *r*+ obstruent is rarely retained), and certain classes of sounds, such as aspirates and palatals, are not permitted finally.

### 3.6 Accent

Vedic Sanskrit has a pitch accent system, described also by Pāṇini, but accent has disappeared in Classical Sanskrit. The Vedic accent can fall anywhere in the word and, as it is not phonologically predictable, the position of the accent often conveys morphological and syntactic information.

Most Vedic words possess one accent. A few loosely bound compounds keep accent on both members, and a number of linguistic forms lack accent: some particles, some pronouns, and, most interestingly, noninitial vocatives and noninitial finite verbs in main (but not subordinate) clauses.

For ease of exposition, accent will in many instances not be marked in the ensuing discussions.

### 3.7 Diachronic developments

As in most early Indo-European languages, the loss of the so-called *laryngeal* consonants (cover-symbol \**H*) of Proto-Indo-European had major effects on Sanskrit phonology and morphology. The phonological alternations originally caused by these segments have been morphologized in various ways, especially visible in the variant forms of roots.

1. *seṭ* vs. *aniṭ* roots: In many obstruent-final roots, an *i* (from vocalized \**H*) surfaces in preconsonantal position, with no counterpart prevocally. Such roots are known as *seṭ* (“with an *i*”), and contrast with apparently parallel *aniṭ* (“without an *i*”) roots. Compare examples of identical morphological categories:

(7)	<i>set:</i> pat “fly”	<i>aniṭ:</i> cit “think”
<i>PreC.</i>	pati-ta	cit-ta
<i>PreV.</i>	pat-ati	cet-ati

Because the distinction is neutralized in prevocalic position and because the interposition of the *i* helps to avoid the often awkward sandhi of consonant clusters, this *i* spreads beyond its original historical boundaries. Indeed, many suffixes and endings are reinterpreted as having an initial *i* (or at least an alternate form with initial *i*).

2. *Roots in \*ṛH*: Sonorants (or resonants; i.e., *i*, *u*; *ṛ*, (*ḷ*); \**ṛ*, \**ṛ*) followed by a laryngeal in Proto-Indo-European yield so-called *long sonorants*, having root-final alternation patterns as follows:

(8)	<u>Zero-grade</u>		<u>Full-grade</u>		<u>Extended-grade</u>	
	<i>PreC.</i>		<i>PreC.</i>		<i>PreC.</i>	
<i>Vowel</i>	<i>PreC.</i>	<i>PreV.</i>	( <i>guṇa</i> )	<i>PreV.</i>	( <i>vṛddhi</i> )	<i>PreV.</i>
* <i>ṛ</i> H	īṛ/ūr	ir/ur	ari	ar	āri	ār
* <i>i</i> H	ī	(i)y	ayi (> ē)	ay	āyi	āy
* <i>u</i> H	ū	(u)v	avi	av	āvi	āv
* <i>m</i> H	ām	(a)m	ami	am	āmi	ām
* <i>n</i> H	ā	(a)n	ani	an	āni	ān

Consider the following examples:

(9)		* <i>i</i> H tṛ “cross”	* <i>i</i> H nī “lead”	* <i>u</i> H bhū “become”	* <i>m</i> H kram “stride”	* <i>n</i> H jan “be born”
<i>Zero:</i>	<i>PreC.</i>	tīr-ṇa	nī-ta	bhū-ta	krām-ta	jā-ta
	<i>PreV.</i>	tir-ati	nin(i)y-ur	bhuv-āni	cakram-ur	jajñ-ur
<i>Full:</i>	<i>PreC.</i>	tari-ṣyati	nayi-tum	bhavi-tum	krami-ṣyati	jani-tum
	<i>PreV.</i>	tar-ati	nay-ati	bhav-ati	kram-ate	jan-ati

The distribution of *īṛ/ūr* and *ir/ur* forms in \**ṛ**h* roots was originally conditioned by the quality of the preceding consonant, with *u*-forms following labials (e.g., √*pṛ* “fill,” with *pūrṇa*).

3. *Roots in ā*: Such roots show an extremely anomalous set of alternations in comparison with the patterns set by other root types. As was first recognized by F. de Saussure in the 1870s, the anomalies can be explained by positing the same structure and alternations as in *set* roots; in other words, by rewriting (in modern terms) *ā* as \**VH* and its unstrengthened form as \**H*, yielding *i* before consonant and zero before vowel:

(10)	<u>Zero-grade</u>		<u>Full-grade</u>
	<i>PreC.</i>		<i>PreC.</i>
	<i>PreV.</i>		( <i>guṇa</i> )
(* <i>VH</i> ) > ā	(* <i>H</i> ) > i	ø	ā
for example,			
sthā “stand”	sthi-ta	tasth-ur	asthā-t

## 4. MORPHOLOGY

### 4.1 Word formation

The basis of Sanskrit morphology is the *root*, a morpheme bearing lexical meaning. Through the vowel-gradation processes described above and through the addition of affixes, verbal and nominal *stems* are derived from this root. The grammatical and syntactic identity of a stem in context is then fixed by the addition of an *ending*. In other words, the three major formal elements of the morphology are (i) root, (ii) affix, and (iii) ending; and they are roughly responsible for (i) lexical meaning, (ii) derivation, and (iii) inflection respectively. A (noncompound) word ordinarily contains only one root and one ending, but may have a theoretically unlimited number of affixes. Both ending and affix may also be represented as zero. The canonical structure of a Sanskrit word is thus:

#### (11) Root – Affix $0_n$ – Ending $0_1$

Numerous examples of roots and their alternants were given above. There are some phonological constraints on root structure, the most important being that no root can end in short *a*, though affixes and endings commonly do, and all roots are monosyllabic (not counting the *i* of *seṭ* roots; see §3.7). There are also some restrictions on co-occurrence of consonants: for example, roots do not contain two aspirates (the historical result of Grassmann's Law) or stops from the same positional series in onset and coda.

Affixes are almost entirely suffixes. There is one infix, alternating *-na/n-* found in a single verbal present class, and one clear prefix, the so-called *augment*, an *a-* prefixed to past tense verb forms in the imperfect and aorist tenses. In addition, the class of *preverbs* mimic prefixes, because they precede a verb (and its nominal derivatives) and modify it semantically (e.g., *ud* “up,” *pra* “forth”). In the earliest language, however, the status of these elements is not clear, or rather it fluctuates, as both their position and their accentuation show. In the *Rig-veda* preverbs regularly occur in *tmesis*, in other words, separated from the finite verb. Even when immediately preceding the verb, they maintain their own accent, except in subordinate clauses. This last context is the only one in which they clearly form a part of the phonological word of the finite verb. Preverbs always precede the one undeniable prefix, the augment. With nonfinite forms of the verb and with nominal derivatives thereof, preverbs show much clearer univerbation in Vedic, both by position and by accent, and by Classical Sanskrit *tmesis* is no longer possible even with finite forms.

In nominal morphology three elements, *a(n)-* “un,” *su-* “well,” and *dus-* “ill,” function like prefixes, though technically forming compounds, both determinative and possessive.

Besides these few exceptions, suffixes are the rule in affixation. Though there are few absolute phonological constraints on suffixes, most are monosyllabic (though sometimes with the old laryngeal *i* attached, see above) and have relatively simple structure: CV is a common shape. The same is true of endings.

Reduplication is a common morphological process in the verbal system. Although the details cannot be examined, several of the phonological alternation processes discussed above are exemplified in reduplication: dissimilation of aspirates ( $\sqrt{dhā}$ : *da-dhā-*), alternation of palatals and velars ( $\sqrt{kr}$ : *ca-kr-*).

Some words do not conform to the canonical structure. A few forms lack both inflection and root and do not ordinarily serve as derivational bases: for example, the negatives *ná* and *mā́*, particles of various functions like *sú* and *hí*, and conjunctions like *ca* and *vā* (some are tonic, some not). Preverbs can be classified here at least originally.

Moreover, a much larger number of words are inflected (and can enter into derivation) but lack a recognizable root. These include many terms of basic vocabulary – kinship terms (e.g., *mātar-* “mother”), body parts (e.g., *nas-* “nose”), flora and fauna (e.g., *śvan-* “dog”) – but are not limited to such semantic categories. Pronouns might be usefully classified here. Numerals also lack roots; some are inflected, some not.

Sanskrit morphology is conveniently divided into two fundamental categories, namely nominal forms and verbal forms, formally distinguished by the types of endings they take and the grammatical categories these endings mark. Adjectives and participles derived from verbs are not formally distinct from nouns; pronouns share the same grammatical categories with nouns, though they may deviate somewhat in inflection. “Adverbs” are usually frozen case forms of adjectives, and nonfinite verbal forms such as infinitives and gerunds also clearly show frozen nominal case endings.

Before discussing nominal and verbal forms separately, we should note certain features and processes they share. Perhaps the most important is the distinction in each between *thematic* and *athematic* inflection. Any *stem*, nominal or verbal, that ends in short *a* (i.e., ends with a suffix consisting of or containing short *a* as final vowel) is thematic. All thematic stems show fixed form throughout their inflection, modified only by the addition of endings. There are no stem alternants and there is no accent shift in the paradigm. Any stem not ending in short *a* is athematic and ordinarily will show stem alternants (as generated by the vowel strengthening patterns discussed above) and often movable accent. For example, the noun stem *deva-* “god” is thematic and maintains this form throughout, whereas *rājan-* “king” is athematic, with the following stem alternants: “strong” *rājān-* (*/rājā-*), “middle” *rājan-*, “weak” *rājñ-* (*/rāja-*). Similarly in verbs, a nonalternating thematic present stem like *bhava-* “become” contrasts with athematic *kṛnó-* / *kṛñu-* (with accent shifted to the ending) “make.” Given the relative simplicity of the former and the frequent morphophonemic complications of the latter, thematic inflection spreads at the expense of athematic inflection during the history of Sanskrit.

Two of the facts noted above – that affixes can be athematic (and alternating) as well as thematic, and that Sanskrit words can contain more than one affix – interact with each other. With very rare exceptions, only one element in any Sanskrit word will alternate within a single paradigm; all the rest will remain frozen in a nonalternating, usually weak form. Whenever a suffix (thematic or athematic) is added to a stem, all preceding elements become frozen. For example, the root  $\sqrt{kṛ}$  alternates within its root aorist paradigm: *ákar-am* “I have made” versus *ákr-an* “they have made.” However, when the present-stem alternating suffix *-nó/nu-* is added, the root syllable *kṛ* is fixed in zero-grade: *kṛ-nó-/kṛ-ñu-*. In turn, with the optative suffix *-yá/ī-* added to that, the present stem is frozen in weak form: *kṛñu-yá-/kṛñv-ī-*.

## 4.2 Nominal morphology

The grammatical categories of Sanskrit nominal forms are gender, number, and case.

### 4.2.1 Gender

Three genders exist: masculine, neuter, and feminine. Nouns have inherent gender; personal pronouns have no gender, though demonstrative and anaphoric pronouns do. The formal expression is not parallel among the three genders. The feminine is primarily expressed by *derivation*: there are two important feminine-forming suffixes, *-ā-* and *-ī-*. By contrast, the difference between masculine and neuter is primarily *inflectional*. For the most part the

same suffixes form both masculine and neuter nouns, and different case endings signal the different genders. Most stems formally encode masculine versus neuter only in nominative and accusative. A few stem-types (especially *i*-stems) form feminines as well as masculines and neuters, where the feminine is distinguished by different case endings and by the form of modifying adjectives.

#### 4.2.2 Number

Three numbers occur: singular, dual, and plural. The dual is a fully functioning category, used not merely for naturally paired objects, like eyes, but for any collection of two. Notable in Vedic is the “elliptical” dual, with a noun in the dual signalling a conventional paired opposition: for example, *dyāvā*, literally “the two heavens,” for “heaven and earth”; *mātarā*, literally “the two mothers,” for “mother and father.” Number is entirely an inflectional category, except in the personal pronoun.

#### 4.2.3 Case

Sanskrit has eight cases: nominative, accusative, instrumental, dative, ablative, genitive, locative, vocative, though no stems make all eight distinctions in all three numbers. In all stems the dual shows only three distinctions: (i) nominative, accusative, and vocative merge; as do (ii) instrumental, dative, ablative; and (iii) genitive, locative. In all nominal stems the plural collapses nominative and vocative, as well as dative and ablative; only the personal pronouns distinguish dative and ablative in the plural. Even in the singular most stems conflate ablative and genitive; only one nominal stem-type (though the most common, the short *a*-stem) and the pronouns distinguish ablative and genitive singular. Thus, since pronouns lack vocatives, only one stem-type (*a*-stem) has eight distinct case forms in any number. Case function is discussed in §5.

Case is marked inflectionally, by endings, and by stem-form alternations. In alternating paradigms some cases regularly pattern together, in other words, show the same stem alternants. Normally (i) nominative/accusative singular, (ii) nominative/vocative plural and (iii) nominative/accusative/vocative dual (the so-called *strong* cases) operate in opposition to the other, *weak cases* (the terms *direct* versus *oblique* have almost the same range of reference, but are syntactic not formal designations; moreover, the accusative plural is also a direct case).

#### 4.2.4 Nominal stem-classes

Unlike a language such as Latin or Greek, Sanskrit has no closed set of conventionally denoted *noun declensions*. Instead, there is a fairly large set of stem-types, some of which share features of patterning, as well as a sizable group of exceptional stems (not treated here). The first major division is between *root nouns* and *derived nouns*. As the name implies, root nouns combine the bare root, without suffixes, with endings, while derived nouns interpose suffix(es) between root and ending.

##### 4.2.4.1 Vowel stems

The major division in derived nouns is between vowel stems and consonant stems, distinguished by the patterning of stem alternants and to some extent by endings. Among vowel stems we can differentiate three types:

1. The short *a* thematic type, the commonest stem-type in the language, forming masculines (e.g., *deva-* “god”) and neuters (e.g., *phala-* “fruit”). Besides its invariant stem, it is distinguished by somewhat aberrant endings and by the fact that it alone has eight distinct forms in the singular.
2. The *ā* and *ī* feminine stems (e.g., *senā-* “army,” *devī-* “goddess”). In addition to their gender, these stems share a distinctive set of endings in the singular oblique cases.
3. The stems in short *i* and *u*, forming nouns of all three genders (e.g., masc. *agni-* “fire,” fem. *mati-* “thought,” neut. *vāri-* “water”; masc. *paśu-* “cow,” fem. *dhenu-* “milk-cow,” neut. *vasu-* “wealth”). In early Vedic the inflection of all three genders is essentially the same (save for the neuter endings of the direct cases), with weak forms of the stem in the singular direct cases (*agni-*) and strong forms in the singular oblique (*agnay-*). Gradually all three genders develop separate singular oblique forms. The feminine stems become more like the stems of type 2.

#### 4.2.4.2 Consonant stems

A number of varieties occur (*an-*, *ar-*, *ant-*, *vas-*, and *as-* stems, among others), forming primarily masculine and neuter nouns. Most consonant stems share a general patterning tendency: strong forms of the stem occur in the “strong” cases, weak in the “weak” cases (e.g., *rājān-* vs. *rājñ-*; *kartār-* vs. *kartr-*; *sánt-* vs. *sat-*, etc.), in direct opposition to the patterning of the short vowel stems just discussed. A few stem-types show no significant stem alternation (*in-* stems, neuter *s-* stems). Note also that *ar-* stems are often classified as vowel stems (i.e., as *r-* stems), and several of their cases have indeed adopted vowel-stem forms (especially acc. pl., gen. pl.). But the patterning of their stem alternants clearly classifies them with consonant stems, especially *an-* stems.

#### 4.2.5 Endings

Though no scheme of endings is applicable to all stems and all periods of the language, the following chart gives the most common patterns. When there are significant differences, both consonant and vowel-stem endings are given, as well as some feminine alternants.

(12)	Singular				Dual		Plural		
	Cons.	Vow.	Fem.	Neut.	Cons.	Neut.	Cons.	Vow.	Neut.
Nom.	∅	-s		∅	-au	-ī	-as		-Ṽñi
Acc.	-am	-m		[=nom.]	[=nom.]		-as	-Ṽñ	[=nom.]
Instr.	-ā	-nā	-ā		-bhyaṁ			-bhis	
Dat.	-ē	-ē	-āi		[=instr.]			-bhyas	
Abl.	-as	-s	-ās		[=instr.]			[=dat.]	
Gen.		[=abl.]			-os		-ām	-Ṽñām	
Loc.	-i	var.	-ām		[=gen.]			-su	
Voc.	∅	var.	var.		[=nom.]			[=nom.]	

#### 4.2.6 Comparison of adjectives

There are two different patterns for producing comparatives and superlatives, one primary, that is, by direct attachment to the root, not to a derived adjective (comp. *-īyas-*, splv. *-iṣṭha-*); the other secondary, by attachment to an existing adjective (*-tara-*, *-tama-*). An example of each follows:



- (13) *primary*    urú- “wide”    várīyas- “wider”    váriṣṭha- “widest”  
*secondary*    priyá- “dear”    priyátara- “dearer”    priyátama- “dearest”

In Vedic the secondary suffixes are used rather freely, for example, in compounds like *somapātama-* “most soma-drinking” (i.e., “best drinker of soma”); *vṛtrahantama-* “most Vṛtra-smashing” (i.e., “best smasher of Vṛtra”).

#### 4.2.7 Pronouns

The major division within this category is between (i) the personal pronouns of the first and second persons, unmarked for gender, and (ii) a larger number of gender-distinguishing demonstrative/deictic/anaphoric pronouns and adjectives.

##### 4.2.7.1 Personal pronouns

The cases of these pronouns were noted above, as was the occurrence of a different stem in each number. The number of stems is in fact still greater, in that the first singular and plural and the second plural use a different form for the nominative than for the rest of the paradigm:

(14)	<i>1st sg.</i>	<i>1st pl.</i>	<i>2nd pl.</i>
<i>Nom.</i>	ahám	vayám	yūyám
<i>Elsewhere</i>	m-	asm-	yuṣm-

The other stem formants are 1st dual *āv-*, 2nd. sg. *tu-*, 2nd dual *yuv-*. There also exist enclitic oblique forms, often with yet a different stem (e.g., 1st. pl. *nas*, 2nd pl. *vas*). The endings of the personal pronouns are in part unique to them.

##### 4.2.7.2 Gender-marking pronouns

Such pronouns are characterized by a number of different paradigms and partial paradigms, with different functions sometimes changing over time. Most can be used both as pronouns proper and as demonstrative adjectives. We will mention only the most important and widespread stems, beginning with the strong deictics, nearer *ayám* “this here,” farther *asáu* “that yonder.” Both have rather aberrant inflection, with an assortment of stems collected from different sources.

The most common pronominal stem is *sá/tám*, with a wide range of uses. While it serves as the anaphoric pronominal par excellence, it also shows traces in early Vedic of deictic usage. Moreover, it is the closest element Sanskrit possesses to both a third-person pronoun and to a definite article. It is also sometimes used with both second- and first-person reference. Its inflection shows archaic inherited features, with initial *s-* in nominative singular (masc. *sá* and fem. *sā́*), versus *t-* elsewhere (replicated by Greek masc. *ho*, fem. *hē* [with *h-* < \**s-*] but neut. *tó*; see Ch. 24, §4.1.3.4), and with an endless nominative singular masculine (under certain sandhi conditions).

This stem also shows some peculiarities of inflection, some of which are found also in the stems of the interrogative (*ká-*), the relative (*yá-*), and a class of “pronominal adjectives” such as “other” (*anyá-*), “all” (*víśva-*, replaced by *sárva-*), “one/some” (*éka-*).

### 4.3 Verbal morphology

Like nouns, verbs are either thematic or athematic. Athematic verbs regularly alternate strong (guṇa) forms in the active singular, weak in the rest of the inflection.

The grammatical categories of finite verbs are person, number, voice, tense/aspect, and mood. In general, person/number/voice are expressed by a portmanteau morpheme, the endings; tense/aspect by suffixes, morphological processes directly affecting the root, and/or endings; and mood by suffixes (or endings) following the tense/aspect markers. The canonical shape of a verb is thus:

(15) Root – (Tense/Aspect suffix) – (Mood suffix) – Per./Num./Voice ending

#### 4.3.1 Person and number

These categories index the subject of the verb. There are three persons, first, second, and third (in Western grammatical terminology); and three numbers, singular, dual, and plural. As in the noun, the dual is fully functioning, not limited to subjects naturally occurring in pairs. The nine-member grid defined by these two parameters is the basic building block of the Sanskrit verbal system, the paradigm. Each person/number pair is marked by a separate ending.

#### 4.3.2 Voice

The approach to this topic will differ depending on whether formal or functional aspects are emphasized. Formally, many Sanskrit nine-member paradigms come in matched pairs, in two different voices – with identical stems but different endings. The two voices are *active* and *middle* (or *mediopassive*), or, in the more perspicuous Sanskrit terms, *parasmaipada* “word for another” and *ātmanepada* “word for oneself.” A typical formal configuration, the endings of the present, active, and middle, is given below:

	Active			Middle		
	Singular	Dual	Plural	Singular	Dual	Plural
1st	-mi	-vas	-mas	-e	-vahe	-mahe
2nd	-si	-thas	-tha	-se	-āthe	-dhve
3rd	-ti	-tas	-anti	-te	-āte	-ante

The function of the separate voices is harder to define. Though there exist contrasting pairs such as act. *yajati* “sacrifices (on another’s behalf)” : mid. *yajate* “sacrifices (for one’s own benefit),” which illustrate the Sanskrit terminology, there are other active : middle functional relations: for example, transitive : intransitive, act. *vardhati* “increases X” : mid. *vardhate* “X increases.” Some middles are simply passive in value, though lacking overt passive suffix, and an even greater number have no obvious functional correlate: for example, the numerous *deponents* (to use the Latin term) inflected only in the middle (e.g., *āste* “sits”). The distinction between active and middle is, in the main, a purely formal one synchronically; not surprisingly, the distinction becomes attenuated in the development of the language.

There is, however, an important functional distinction in voice, with various formal encodings: that between active and passive. As just noted, the formal middle sometimes functions as a passive. One particular present-stem type, the suffix-accented *-yá-*present with middle endings, also becomes specialized as a passive (e.g., *ucyate* “is spoken”); and the aorist system contains a third singular of peculiar formation (heavy root syllable and mysterious ending *-i*; type *avāci* “was spoken”), the so-called aorist passive. Passive value is also expressed by several verbal adjectives, the gerundive (“future passive participle”) in *-ya-* and *-tavya-*, and especially the past passive participle in *-ta-* (*-na-*). The latter often substitutes for a finite verb as sentential predicate.

### 4.3.3 Tense-aspect

The backbone of the tense-aspect system is the three-way contrast between the *present* system, the *aorist* system, and the *perfect* system. Each of these stems produces one or more tenses, as well as (in the early language) moods and participles. The present system has two tenses, the present and the imperfect. In post-Rig-vedic Sanskrit both the aorist and the perfect have only one, though in the *Rig-veda* there is a marginal pluperfect beside the perfect. All three systems can be inflected in either voice:

(17) Stem	Tense
present	present
	imperfect
aorist	aorist
perfect	perfect
	(pluperfect)

Like voice, the tense-aspect system is an elaborate formal edifice whose functional motivations have essentially broken down. Though the system inherited from Proto-Indo-European was an aspectual one, aspect is no longer a clear category even in early Vedic, and only relics of the inherited system can be discerned in the *Rig-veda*. From the Sanskrit point of view, the salient functional distinction is *tense*: present (expressed by the present tense) versus past (expressed by three competing preterital forms, imperfect, aorist, and perfect, as well as by certain nonfinite forms used predicatively).

The old perfect was a stative present functionally; a few Vedic perfects maintain this function, but most already express simple past. The original distinction between the present and aorist systems was probably durative versus punctual, but this can no longer be discerned. Insofar as the aorist can be distinguished from the imperfect in Sanskrit, it expresses immediate past time. The loss of functional distinction among the three past tenses set the stage for the loss of those formal categories in later Indo-Aryan.

### 4.3.4 Perfect stem morphology

Formally, the perfect is characterized by special endings and, except for one widespread old form (*veda* “knows”), by reduplication. It is built directly to the root, without affixes, and shows ordinary strong/weak stem alternation (type *cakār-a/cakr-úr*). There is only one type of perfect stem formation (except for the “periphrastic perfect” of derivative presents; see §4.3.6).

### 4.3.5 Primary and secondary endings

The formal distinction between present and aorist systems is less well marked. The endings of the imperfect tense and the aorist are identical (the so-called *secondary* endings), and the endings of the present tense (the *primary* endings) closely resemble these. Compare, for example, the primary and secondary endings of the active singular, and contrast them with the corresponding perfect endings:

(18)	Primary	Secondary	Perfect
1st	-mi	-m	-a
2nd	-si	-s	-tha
3rd	-ti	-t	-a

Unlike the perfect both imperfect and aorist prefix the augment, regularly in Classical Sanskrit and optionally (but commonly) in Vedic. Moreover, several types of stem formation are common to both present and aorist.

#### 4.3.6 Present stem morphology

The indigenous grammarians distinguish ten present classes, which can be conveniently divided into thematic and athematic types. Four thematic classes occur, with the following suffixes added to the root: -a- (Class I); -á- (VI); -ya- (IV); and -áya- (X). The six athematic classes are as follows: simple root presents (endings added directly to the alternating root, Class II); reduplicated presents (III); and four classes continuing (directly or indirectly) nasal affixes – nasal infix (VII), and suffixed -nó/nu- (V), -ó/u- (VIII), and -ná/nī- (IX). Examples of each follow; thematic forms (with nonalternating stems) are given in the third singular active present, athematic forms in both third singular and third plural active, to display both stem alternants:

#### (19) Sanskrit present tense classes

I	simple thematic	√bhū “become”	bháva-ti
II	root	√as “be”	ás-ti, s-ánti
III	reduplicated	√hu “pour”	juhó-ti, juhv-ánti
IV	-ya-	√paś “see”	páśya-ti
V	-nó/nu-	√su “press”	sunó-ti, sunv-ánti
VI	-á-	√viś “enter”	viśá-ti
VII	nasal-infix	√yuj “yoke”	yunák-ti, yuñj-ánti
VIII	-ó/u-	√tan “spread”	tanó-ti, tanv-ánti
IX	-ná/nī-	√krī “buy”	krīṇá-ti, krīṇ-ánti
X	-áya-	√cint “think”	cintáya-ti

There is no longer any clear distinction in function among these various present classes, though again traces of prehistoric distinctions can occasionally be discerned.

Besides the above ten classes, several other formations are formally presents, but are classified separately because they have clear functional correlates.

The *future* is formed with the thematic suffix -syá- (or -iśyá- originally proper to *seṭ* roots) (e.g., *kariṣyáti* “will do”: √kr). There is also a periphrastic future, formed from a noun stem with the -tar- agent suffix.

The so-called secondary conjugations:

1. *Passive*, formed with accented -yá- and middle endings, for example, *nīyáte* “is led”: √nī “lead.” In Classical Sanskrit with the loss of accent the passive cannot be formally distinguished from a middle Class IV present.
2. *Intensive*, formed with heavy reduplication (sometimes disyllabic) and, in later Sanskrit, a -yá- suffix with middle endings. The intensive expresses repeated or intensively performed action, for example, *mármarj-*, *marmṛjyáte* “wipe repeatedly, groom”: √mrj “wipe.”

3. *Desiderative*, formed with reduplication in *-i-* and a *-sa-* suffix. The desiderative expresses action desired, intended, or about to take place, for example, *pīpāsati* “desires, intends, is about to drink”:  $\sqrt{pā}$  “drink.”
4. *Causative*, formed with a heavy root syllable and a suffix *-āya-*. Formally not distinguishable from Class X presents, except sometimes in the shape of the root syllable. In the earlier language the causative is ordinarily formed only to intransitive verbs, for example, *pādāyati* “cause to fall”:  $\sqrt{pad}$  “fall.”

In addition to present stems built to verbal roots, nouns and adjectives can form *denominative* presents by the addition of the suffix *-yá-*, for example, *ásva-* “horse”: *ásvayāti* “seek horses.”

The above derivative present stems can form a secondary periphrastic perfect, with a feminine accusative singular generated to the present stem, plus the perfect of  $\sqrt{kṛ}$  (in the earlier language),  $\sqrt{bhū}$  or  $\sqrt{as}$  (in the later language), of the type *pādayāṁ cakāra/āsa* “caused to fall.” The periphrastic perfect is especially common with causatives.

### 4.3.7 Aorist stem morphology

The aorist shares certain stem-types with the present system. The root aorist (e.g., *ābhūt*:  $\sqrt{bhū}$  “become”) and thematic aorist (*āvidat*:  $\sqrt{vid}$  “find”) resemble Class II and VI presents. Class III presents somewhat resemble the reduplicated aorist, though the aorist has certain formal characteristics (heavy *ī-* reduplication, thematic vowel) and a functional connection with the causative (type *āpīpadat* “caused to fall,” parallel to *pādāyati* “causes to fall”) that set it apart.

Proper to the aorist, however, are a variety of sigmatic formations. The *s*-aorist and *iṣ*-aorist were originally identically built, with *s*-suffix, to *aniṭ* and *seṭ* roots respectively. Especially notable in these formations is the consistent *vṛddhi* of the root in the entire active voice, an unusual distribution of grades (e.g., *s*-aor. *ājai-s-* “he conquered”:  $\sqrt{ji}$  “conquer”; *āpāvi-ṣ-* “purified”:  $\sqrt{pū}$  “purify”). Analogic extensions of these two aorist types led to the creation of the marginal types, *siṣ*-aorist and *sa*-aorist.

The passive aorist was noted in §4.3.2.

### 4.3.8 Mood

There are four clear moods in early Sanskrit: indicative, imperative, optative, and subjunctive. In addition, the so-called injunctive of early Vedic is considered a mood by some, and the precative, a subtype of the optative, develops in the course of Vedic. This system is reduced by Classical Sanskrit. One global change is the virtual restriction of nonindicative moods to the present stem; in Vedic, aorists and perfects displayed broader modality. Furthermore, the subjunctive is effectively lost, and the injunctive, insofar as it is a mood, becomes restricted in usage.

#### 4.3.8.1 Indicative

The indicative is the unmarked mood, used for statements, questions, etc.

#### 4.3.8.2 Imperative

The imperative expresses command and is marked by special endings on the appropriate tense stem. In Vedic the imperative has a defective paradigm, being found only in second and third persons, but as the subjunctive is lost as a functional category, its first-person

forms are incorporated into the imperative. The negative imperative (i.e., prohibitive) is expressed not by the formal imperative mood, but by the injunctive with a special form of the negative, namely *mā* (not *ná*).

There is also a rare second imperative formation, the so-called *future imperative*, made by adding *-tāt* to the tense stem, expressing a command to be executed after the action of an intervening verb. Its value is usually second singular.

#### 4.3.8.3 Optative

The optative expresses possibility (“might,” “could”), necessity (“should,” “ought to”), and will/desire (“would”), and is marked by a suffix added to the tense stem. For athematic stems, the suffix is *-yā-* in the active, *-ī-* in the middle, added to the weak stem form (e.g., *s-yā-* to root pres. *as-ti*, *s-anti*:  $\sqrt{as}$ ; *kṛṇu-yā-*, *kṛṇv-ī-* to *kṛṇōti*:  $\sqrt{kṛ}$ ). For thematic stems, *-ē-* is substituted for the thematic vowel *-a-* throughout (e.g., *bhāvē-* to thematic pres. *bháva-*:  $\sqrt{bhū}$ ). Both suffixes take secondary endings, with some special details.

The precative is a supercharged optative, primarily expressing desire. It is formed by interposing an *-s-* between the optative suffix and the ending. Thus, the ordinary athematic optative first singular ends in *-yām*, that of the precative in *-yāsam*; that of the first plural optative in *-yāma*, the precative in *-yāsma*.

#### 4.3.8.4 Subjunctive

This mood has disappeared (except for its formal representatives in the imperative) by Classical Sanskrit. It is formed by adding a suffix *-a-* (identical to the thematic vowel) to the tense stem; in thematic verbs this produces a contracted suffix *-ā-* (e.g., *bhāvā-* to *bháva-ti*). Athematic verbs add the *-a-* to their strong forms (e.g., *ás-a-* to *ás-ti*; *kṛṇáv-a-* to *kṛṇó-ti*). The subjunctive stem can take either primary or secondary endings (*ásati*, *ásat*, etc.); in addition, the typical final vowel of primary middle endings, *-e*, is usually strengthened to *-āi* after the Rig-vedic period.

The function of the subjunctive is difficult to define. It often seems to express the future, or volitional future, rather than the more strictly *modal* value its Western name implies. This interpretation fits well with the fact that the future tense is quite rare in early Vedic in finite forms; their place seems to be filled by the subjunctive.

#### 4.3.8.5 Injunctive

Formally the term *injunctive* simply refers to unaugmented preterite forms (i.e., imperfects and aorists). Such forms are quite common in the *Rig-veda* in a variety of contexts, but only one usage persists into later Vedic and Classical Sanskrit: the conjoining of aorist injunctive and the particle *mā* to express prohibitions. Despite the best efforts of numerous distinguished scholars, a common functional core cannot be discerned in the other Rig-vedic contexts, and it seems best to regard these forms as not belonging to a unified modal category, but rather representing a period when the prefixation of the augment was still optional in the preterite.

### 4.3.9 Nonfinite verbals

Sanskrit possesses a large number of verbal nouns and verbal adjectives, of common occurrence. These ordinarily show verbal syntax (objects in the accusative, for example), and many can stand as the main verb in a clause. Some are built directly to the root, some to tense stems.

#### 4.3.9.1 Infinitive

Classical Sanskrit has a single infinitive, built with the suffix *-tum* added directly to the root in guṇa form (type *kar-tum*: √*kṛ*), which is much rarer in textual usage than the infinitives of other early Indo-European languages. It continues the frozen accusative singular of a nominal stem with a *tu*-suffix, and indeed in Vedic other case forms of this stem appear in infinitival usage: dative *-tave* (*/-tavāi*), ablative-genitive *-tos*. In addition, other stem-types form infinitives or quasi-infinitives in Vedic, for example, datives to *as*-stems in *-ase*. The line between an infinitive and a simple noun can be difficult to draw in the early language.

Infinitives appear as complements to verbs such as √*śak* “be able” and are used to express purpose. They are neutral as to voice and can express either active (“to X”) or passive (“to be Xed”) value, usually depending on the voice of the form to which they are complement.

#### 4.3.9.2 Gerund

These frozen instrumentals, common in Sanskrit of all periods, are used to express an action prior to (or just simultaneous with) that of the main verb. Standard Classical Sanskrit has two formations, formally distributed: *-tvā* (also made to the *tu*-stem noted under the infinitive, §4.3.9.1) built to an uncompounded root; and *-(t)ya* built to preverb + root (thus the type *kṛ-tvā* vs. *pra-kṛ-tya*). This formal distribution is not always adhered to in the earlier language, and several other related suffixes are also employed.

#### 4.3.9.3 Tense-stem participles

As with the moods, participles tend to become restricted to the present stem in later Sanskrit, although Vedic allows participles to be built to all three tense-aspect stems. Tense-stem participles distinguish voice. The active participle suffix for present and aorist is *-ant-*; the middle suffix for all three tense-aspect stems is *-āna-* for athematic verbs, *-māna-* for thematic. The active perfect participle is made with the suffix *-vas-*, of curious inflection. Though most nonpresent participles disappear by Classical Sanskrit, the perfect participle to *vēda* “knows,” *vid-vās-*, survives as an adjective meaning “knowing, wise.”

#### 4.3.9.4 Past passive participle

This is an extremely common form, both as an attributive adjective and as a predicative verb substitute. It is built directly to the unstrengthened root with the suffixes *-tá-*, *-itá-* (originating in *seṭ* roots and still largely found there), *-ná-*, and, rarely, *-vá-*: types *kṛ-ta-* “made, done”: √*kṛ* “make, do”; *muṣitá-* “stolen”: √*muṣ(i)* “steal”; *san-ná* “seated”: √*sad* “sit”; *pakvā-* “cooked, ripe”: √*pac* “cook.” Competing with the three finite past tense forms discussed above, the past passive participle is often the successful contestant, and is responsible for the preterites in a number of later Indo-Aryan languages.

#### 4.3.9.5 Past active participle

Derived from the past passive participle by the addition of the possessive suffix *-vant-* (type *kṛtāvanta-* to *kṛtá-*), it is far less successful than its base.

#### 4.3.9.6 Gerundive (or future passive participle)

The gerundive is another form with passive value, but with the additional component of obligation or necessity (“to be X-ed”), often the equivalent of a passive optative (type *kartavya-* “to be done”). It is formed directly to the root by the addition of one of several suffixes, the most common being *-tavya-* and *-ya-*.



## 4.4 Compounds

Sanskrit has an extremely well-developed system of nominal compounding; verbal compounding hardly exists. In Vedic, though all types of nominal compounds occur and are frequently encountered, individual compounds are usually limited to two or three members. In Classical Sanskrit, compounds of dozens of members are not infrequent, especially in philosophical texts: the compounding process comes to take the place of the independent syntactic arrangement of inflected words.

### 4.4.1 Verbal compounds

The verb shows two types of quasi-compounding: (i) the gradual incorporation of preverbs (and functionally equivalent elements) into a verbal complex (type  $\sqrt{\text{gam}}$  “go”:  $\tilde{a}\sqrt{\text{gam}}$  “come”); (ii) the so-called *cvī* construction, which combines nouns and adjectives with both finite and nonfinite forms of the roots  $\sqrt{\text{kṛ}}$  “make” and  $\sqrt{\text{bhū}}$  “become” (meaning “make/become X”). In such cases, the nominal first member substitutes invariant *-ī-* for a stem-vowel *-a-* or *-i-*, *-ū-* for *-u-* (e.g., *stambhī-bhavati* “becomes a post”: *stambha-* “post”).

### 4.4.2 Nominal compounds

Formally, nominal compounding ordinarily involves the concatenation of uninflected words (i.e., stems), resulting in a unit with a single ending and a single accent. The stems may include nouns, adjectives (including participles), adverbs, and pronouns. Both the single ending and the single accent have exceptions in the early language. Inflected case forms may appear in prior compound members, as in *rathe-ṣṭhā-* “standing on a chariot” (with the first member in the locative case). And *paral* compounds (dual dvandvas; see §4.4.2.1) with both members in the dual *and* both accented (e.g., *mitrā-vāruṇā* “Mitra [and] Varuṇa”) are a well-attested feature of Rīg-vedic discourse.

There are three major types of nominal compounds: copulative, determinative, and possessive, known familiarly by their Sanskrit names as *dvandva*, *tatpuruṣa*, and *bahuvrīhi* respectively.

#### 4.4.2.1 Dvandvas

These copulative compounds conjoin two or more stems as parallel members of a series:  $X + Y + Z \dots$  (the “lions and tigers and bears” type). Formally the compound may either take the gender of its final member and be inflected as dual or plural (as appropriate), or be treated as a neuter singular collective. In either case the final member is accented (in accented texts). On the Rīg-vedic dual dvandvas, with double inflection and double accent, see §4.4.2.

#### 4.4.2.2 Tatpuruṣa

The prior member of this determinative compound limits the following member in some way. Two major subtypes can be distinguished according to the underlying case relations of the members: dependent (*tatpuruṣa* proper) and descriptive (*karmadhāraya*). In the former the prior compound member would be in a different case from that one which follows. A typical relation is genitive + head, as in *nṛ-pāti-*, literally “man-lord,” that is “lord of men”; but other relations are common, especially the limiting of a final past passive participle by an underlying instrumental agent – type *agni-taptā-*, literally “fire-heated,” that is, “heated by fire.” In *karmadhārayas* the prior member is either a qualifier in the same underlying

case as that member which it limits (typically an adjective, i.e., the “black-bird” type) or an adverbial element (*su-* “well,” *dus-* “ill,” and *a(n)-* “un-” are especially common). The accentual facts of determinatives are complex, but in general the accent falls on the final syllable or the final member.

#### 4.4.2.3 Bahuvrīhi

This possessive compound may be based on any of the preceding types, but adds to the concatenation the semantic feature of possession: the formal sequence  $X + Y$  means not simply “X-Y” but “possessing X-Y.” English has similar compounds; compare *red-head* and *Bluebeard*.

An important formal consequence of the addition of this semantic feature is that the compound, whose final member is a noun, must be transformed into an adjective, capable of inflection in all genders (hence the common designation “secondary adjective compound”). Sometimes the gender switch can be accomplished silently, as it were, as when neuter nouns in *-a-* simply take masculine endings in the nominative and accusative. Sometimes the adjustment simply requires lengthening or shortening the stem-vowel, as when masculine or neuter nouns in *-a-* become feminized as *ā-* stems or, vice versa, a feminine long *ā-* or *ī-* stem is inflected as a short *a-* or *i-* stem in the masculine or neuter. At other times more complex processes must be employed. These possessive adjectives are then often resubstantivized; bahuvrīhis are a rich source for proper names in Indic and other Indo-European languages (as *Bluebeard* demonstrates).

As with determinatives, the accentual facts are complex, but the accent generally falls on the first member. In accented texts it is thus easy to distinguish determinative compounds from bahuvrīhis, but in later Sanskrit this is not formally possible unless the bahuvrīhi has undergone gender shift.

We might note here that Sanskrit nominal morphology engages in a kind of conspiracy to express the semantic feature “possessing.” When a bahuvrīhi cannot be formed, because the notion being expressed is not a compound, a variety of suffixes may be utilized, especially *-vant-* (*-mant-*) and *-in-*, and in early Vedic simple accent shift is possible (e.g., *brāhman-* “formulation” gives *brahmán-* “possessing a formulation,” “priest”).

## 4.5 Numerals

The cardinals from 1 to 10, 20, 100, and 1,000 are:

(20)	1	<i>éka-</i>
	2	<i>dvá-</i>
	3	<i>trí-</i>
	4	<i>catúr-</i>
	5	<i>páñca</i>
	6	<i>ṣáṣ</i>
	7	<i>saptá</i>
	8	<i>aṣṭá</i>
	9	<i>náva</i>
	10	<i>dáśa</i>
	20	<i>vimśatí</i>
	100	<i>śatá</i>
	1,000	<i>sahásra</i>

The relation of most of these to numerals in other Indo-European languages should be obvious.

There are some unusual inflectional details. *Dvá-* “two” is inflected regularly as a dual in all three genders (masc. nom./acc. *dváu*; fem., neut. *dvé*, etc.). Both *trí-* “three” and *catúr-* “four” display some archaic inflectional features, especially the feminine formant *-sr-* between stem and ending; thus nom./acc. pl. *tisrás* (with dissimilation < \**tri-sr-as*), *cátasras*.

Ordinals are derived from cardinals with the suffixes *-ma-* (e.g., *pañcama-* “fifth”) and, rarely, *-tha-* (e.g., *ṣaṣṭha-* “sixth”). Irregular forms include

- (21) first      *prathama-*  
 second      *dvitīya-*  
 third      *tṛtīya-*  
 fourth      *turīya-* Vedic (< \**ktur-*), also *caturtha-*

## 5. SYNTAX

Because of its elaborate morphology many traditionally “syntactic” phenomena take place on the level of morphosyntax in Sanskrit. In particular the case system allows the syntactic roles of nominals to be encoded without recourse to rigid word order or obligatory adpositions. Both prepositions and postpositions are rare in early Sanskrit; they become more common later, developing from old preverbs and from frozen case forms of nouns.

### 5.1 Case usage

Sanskrit cases and their uses are typical of an early Indo-European language: vocative (address); nominative (subject); accusative (direct object; goal of motion; a number of adverbial uses, notably duration of time); instrumental (accompaniment; instrument; agent of the passive; adverbial uses); dative (purpose; indirect object, though the genitive is more commonly used for the latter); ablative (source; cause; comparison); genitive (found in all varieties of adnominal usage; a genitive absolute is also occasionally found, cf. locative absolute); locative (location in both space and time; goal of motion). The locative is also the normal “absolute” case: a noun and modifying participle in the locative can express the time or attendant circumstances under which the action of the main clause occurs: for example, “(on) the sun having risen,” “(on) the enemy fleeing.”

### 5.2 Word order

Although the case system obviates the need for rigid word order, the order of elements in a Sanskrit sentence is not entirely free. Ordinary prose is SOV (Subject–Object–Verb), with many of the standard typological features of this ordering, such as genitives preceding heads. Poetry and artful prose, however, exploit the opportunities that the syntactic clarity of the morphological system affords, by thoroughly scrambling the order of elements for expressive or discourse purposes. Even in the most extreme examples, however, it is usually possible to formulate principles of movement from a putative underlying order parallel to simple prose.

Overt marking of the subject is not necessary; the bare verb, with person/number markings, is sufficient. First- and second-person subject pronouns are used in addition to the verb only for emphatic or contrastive value. The third-person “pronoun” *sá* is more frequent with third-person verbs, but it ordinarily serves discourse functions: anaphoric to a noun previous in the discourse or coreferent with the relative pronoun in a subordinate clause.

Not only finite verb forms but also participles, especially the past passive participle, can fill the slot V. In this case the copula normally appears only in the first and second persons, and even in those circumstances the personal pronoun can serve instead:

- |      |                                |    |                                  |                    |
|------|--------------------------------|----|----------------------------------|--------------------|
| (22) | <i>gató 'smi</i> (with copula) | or | <i>ahám gatáh</i> (with pronoun) | "I went"           |
|      | <i>gató 'si</i> (with copula)  | or | <i>tvám gatáh</i> (with pronoun) | "you went"         |
|      | but                            |    | <i>gataḥ</i>                     | "he [she/it] went" |

Also common are nominal sentences – that is, the predication of a noun (or adjective) to a noun (or pronoun) without an overt copula.

### 5.3 Cliticization

As in other early Indo-European languages, sentences frequently begin with a chain of clitics attached to the initial, accented word of the sentence, occupying "Wackernagel's position." Such a chain of clitics (and pseudo-clitics – some carry accent) consists of sentential particles (often several to the sentence), conjunctions, and pronouns fronted from their underlying position in the clause; their order is determined by both syntactic class and phonological shape. Word-level conjunctions and pronominal clitics may also appear elsewhere in the clause, the latter ordinarily attached to their head. In such positions the pronoun may either precede or follow the word it is attached to, but clause-initial proclitics are not permitted: all clauses (and their metrical equivalent, verse lines) begin with an accented word. Especially common initial hosts include coordinating and subordinating conjunctions, preverbs in tmesis, and tonic demonstrative and anaphoric pronouns. Much recent work on Sanskrit syntax has concentrated on the constituents of this initial chain and their functions.

### 5.4 Subordination

A fully inflected relative pronoun *yá-* and a number of subordinating conjunctions built to this stem (*yadā* "when," *yádi* "if," etc.) mark subordinate clauses. These elements are normally fronted (*wh-movement*) from wherever they originate in the clause, but as other elements (including entire constituents) can be topicalized around them, the fronting is sometimes not superficially obvious. Relative clauses either precede or follow the main clause (the former is more usual except in the case of relative clauses of purpose); there is almost no embedding.

In early Vedic, subordinate clauses are sometimes marked only by verbal accentuation, not by a subordinating conjunction; and some particles (notably *hi*) also induce verbal accentuation, presumably a mark of subordination.

Indirect discourse is quite rare, especially in Classical Sanskrit; such clauses are usually expressed by direct discourse marked by the clause-final quotative particle *iti*. For example, "he thought that he would go" would be expressed as "he thought, 'I will go.'"

Other, nonclausal types of subordination are quite common. For example, a series of gerunds with nominal complements is often completed by a single finite verb (type "having come, having asked the king for permission, having received it, he went away"). A notable feature of the syntax of the gerund is that its subject is the logical *agent* of the main clause, not necessarily the overt grammatical subject (type "having smashed [ger.] with a cudgel, the tiger [nom.] was killed by the man [instr.]", where the subject of the gerund is "the man" in the instrumental).

Participles and possessive compounds often correspond to relative clauses in other languages. Noteworthy is the use of the present participle of the verb “to be” (*sánt-*) as a concessive marker (“*although* being X, ...”). Bahuvrīhis often serve as nonrestrictive relative clauses (type “Indra, [lit.] possessing slain Vṛtra”, i.e., “who had slain Vṛtra”).

Unlike some other early Indo-European languages, Sanskrit has no elaborate rules governing the succession of moods and tenses in conditional sentences.

## 5.5 Agreement

The usual agreement rules of early Indo-European languages hold for Sanskrit: subjects agree with their verbs in person and number; adjectives with the nouns they modify in number, gender, and case; relative pronouns with their antecedents in number and gender.

There are a few interesting exceptions. The well-known Ancient Greek rule, whereby a neuter plural subject takes a singular verb, is preserved only in a few Vedic relics; ordinarily a plural verb is used. Vedic prose has developed a subtype of defining relative clause (type: “... the X, which is Y”) in which the relative marker is always neuter singular *yád*, whatever the gender and number of X and Y. This usage is reminiscent of the Iranian *izafe* marker, which has developed from the same form, but it is not clear if the two constructions are directly related. In some other equational nominal clauses, by contrast, an anaphoric pronoun is attracted to the number and gender of its antecedent.

Though conjoined nominals ordinarily agree in case, an apparently inherited exception in Vedic involves the conjoining of vocatives by *ca* “and,” where the second underlying vocative appears instead in the nominative. This phenomenon is denominated the *vāyav indraś ca* construction after one of its principal examples (“o Vāyu [voc.] and Indra [nom.]”).

## 5.6 Stylistic syntactic developments

One may consider the history of Sanskrit a history of style, and style in turn is linked to textual genre. Although neither the grammar nor the syntax of Sanskrit shows any significant changes after the fixation of the language by the early grammarians, the usage of these fixed elements significantly alters its balance in the Classical period. The emphasis falls heavily on the nominal system, and the complex verbal system outlined above is exploited far less. We have already noted some of the features of this change in emphasis – the efflorescence of the compounding system, the employment of nominal formations built to verbal roots in preference to finite verbs, the expansion of the adnominal case, the genitive. Sanskrit works of “high” style, court literature and philosophical discourse, take these tendencies to remarkable extremes, while technical treatises, with an eye to verbal economy, arrive at a similar nominal style from a somewhat different angle.

## 6. LEXICON

A very large proportion of the Sanskrit vocabulary in all periods consists of transparent Indo-European inheritances. Examples need hardly be given; but the numerals given above (§4.5), as well as kinship terms like *pitar* “father,” *māter*, “mother,” *sūnu* “son,” *duhitar* “daughter” can serve as illustrations. Not surprisingly, however, even earliest Vedic has words without clear Indo-European correspondences. While some of these may nonetheless still continue Proto-Indo-European etyma, others doubtless were borrowed from languages with which

the Sanskrit speakers came in contact. The difficulty is determining the source languages, given the fact that we have no records of likely languages from remotely the same era. Though Sanskrit speakers no doubt encountered speakers of Dravidian language(s), no Dravidian language is attested until around the beginning of the present era, and then only in South India. We do not know what a northern Dravidian language would have looked like in the second millennium BC. Our knowledge of the Muṇḍa languages (belonging to the Austro-Asiatic family) comes only from the modern era. Many scholars have proposed Dravidian and Muṇḍa sources for Sanskrit words (and indeed phonemes, syntactic constructions, and so on). It is reasonable to accept the principle, but difficult to judge the plausibility of any particular suggestion. Even when a single etymon clearly reveals itself in Sanskrit and one or more Dravidian languages, for example, borrowing may have gone in the other direction, or both families may have borrowed from a third source. Later (i.e., post-Vedic) Dravidian borrowings into Sanskrit are less controversial.

In addition to borrowing from non-Indo-Aryan languages, Sanskrit also sometimes re-incorporates vocabulary showing Middle Indic phonological developments, often with some phonological hypercorrection.

## 7. READING LIST

The standard synchronic grammar of Sanskrit in English is Whitney 1889, which, along with its supplement, Whitney 1885, is invaluable. The standard historical grammar is the multivolume but still unfinished (lacking the verb) Wackernagel and Debrunner 1896–. The first volume, reissued in 1957, has a detailed general introduction to the language by L. Renou. Many of Renou's other works can be consulted with profit, including his short but elegant history of the language (1956). The classic work on syntax (but only of the Vedic period) is Delbrück 1888. Speijer 1886 treats the Classical language. The standard etymological dictionary is Mayrhofer 1956–1976, currently updated and significantly expanded in Mayrhofer 1986–. A general discussion of the language, though with personal views, is found in Burrow 1955. A short survey, along the same lines as this, is found in Cardona 1987. Both Bloch 1965 and Masica 1991, though concentrating on later Indo-Aryan, nonetheless treat many aspects of Sanskrit as starting points for later developments.

## Bibliography

- Bloch, J. 1965. *Indo-Aryan from the Vedas to Modern Times*. Translated and revised by A. Master. Paris: Adrien-Maisonneuve.
- Burrow, T. 1955. *The Sanskrit Language*. London: Faber and Faber.
- Cardona, G. 1987. "Sanskrit" In B. Comrie (ed.), *The World's Major Languages*, pp. 448–469. New York: Oxford University Press.
- Delbrück, B. 1888. *Altindische Syntax*. Reprint 1968, Darmstadt: Wissenschaftliche Buchgesellschaft.
- Masica, C. P. 1991. *The Indo-Aryan Languages*. Cambridge: Cambridge University Press.
- Mayrhofer, M. 1956–1976. *Kurzgefasstes etymologisches Wörterbuch des Altindischen*. Heidelberg: Carl Winter.
- . 1986–EWA. = *Etymologisches Wörterbuch des Altindischen*. Heidelberg: Carl Winter.
- Renou, L. 1956. *Histoire de la langue sanskrite*. Lyon: Editions IAC.
- . 1957. *Introduction générale, Altindische Grammatik*, vol. I [see Wackernagel 1896].
- Speijer, J. S. 1886. *Sanskrit Syntax*. Leiden: Brill.
- Wackernagel, J. and A. Debrunner, *Altindische Grammatik*. Göttingen: Vandenhoeck and Ruprecht. Vol. I 1896. *Lautlehre* (J. Wackernagel). Reprint 1957, with intro. by L. Renou.

- 
- Vol. II.1 1905. *Einleitung zur Wortlehre. Nominalkomposition* (J. Wackernagel). Reprint 1957.  
Vol. II.2 1954. *Die Nominalsuffixe* (A. Debrunner).  
Vol. III 1930. *Nominalflexion – Zahlwort – Pronomen* (A. Debrunner and J. Wackernagel).  
Whitney, W. D. 1885. *The Roots, Verb-Forms and Primary Derivatives of the Sanskrit Language*.  
Leipzig: Breitkopf and Härtel.  
———. 1889. *Sanskrit Grammar* (2nd edition). Cambridge, MA: Harvard University Press.



# Middle Indic

STEPHANIE W. JAMISON

## 1. HISTORICAL AND CULTURAL CONTEXTS

*Middle Indic* (or *Prākṛit*) is the designation for a range of Indo-Aryan languages displaying characteristic phonological and grammatical developments from Old Indic (i.e., Sanskrit, see Ch. 26). Like Sanskrit they belong to the Indo-Iranian branch of the Indo-European family and are directly attested beginning in the latter part of the first millennium BC and through the first millennium AD. *Middle Indic* is not strictly a chronological term, but rather refers to logical stages of linguistic development. Some of the defining characteristics of Middle Indic phonology are found already in lexical items in the oldest Sanskrit text, the *Rig-veda*, and Middle Indic languages are attested alongside Sanskrit for all of their history. The alternate designation, *Prākṛit*, means “natural, unrefined,” hence “vernacular,” as opposed to *saṃskṛta-* “perfected,” applied to the prescriptive, rule-governed Classical Sanskrit of the grammarians. As well as sometimes designating all Middle Indic speech forms, *Prākṛit* is often used in the narrow sense to refer to a subset of these languages. This latter usage will be followed here.

In the following sections, we will enumerate the various Middle Indic languages and describe the evidence for them.

### 1.1 Inscriptions

Though forms showing characteristic Middle Indic sound changes are found in our earliest Vedic Sanskrit text, no Middle Indic languages are directly attested until the third century BC. At this time appear the earliest inscriptional records of any Indo-Aryan languages, the inscriptions of the Buddhist Mauryan emperor Aśoka (Aś.). These consist of a number of proclamations (fourteen rock inscriptions, seven pillar edicts, etc.), each with identical texts composed in several different local dialects and distributed throughout India (there are also inscriptions in Greek and Aramaic, which fall outside the scope of this chapter). The language of the texts may be termed *Early Middle Indic*, and the local dialect features displayed allow the separate versions to be used as the basis for studying later dialect development. Most of the inscriptions are written in Brāhmī script (see Ch. 26, §2), except for those in the extreme Northwest, written in Kharoṣṭhī (both scripts were deciphered in the 1830s). After this spectacular beginning, inscriptions in Middle Indic were produced for more than a half-millennium in various parts of India, continuing to show local dialect features.

## 1.2 Pāli (Pā.)

The Buddha is said to have preached in the vernacular, not Sanskrit, and although we have no direct records from the time of the Buddha, early Buddhist documents are in Middle Indic. The most extensive and linguistically conservative records are found in the canon of the Theravāda school, composed in the language known as Pāli. Though the texts were preserved in Śrī Lanka, they were clearly brought originally from the mainland and seem to represent a Western dialect, with some admixture of Eastern features (the Buddha himself lived in the East). The redaction of the canon probably occurred around the beginning of the present era, though the texts doubtless continue older oral traditions.

## 1.3 Gāndhārī Prākṛit

Post-Aśokan Kharoṣṭhī inscriptions of the Northwest find, to some degree, their linguistic continuation in a large cache of third-century AD documents on wood, paper, and leather, discovered in Niya in Central Asia (the so-called Niya Documents), and in a fragmentary manuscript of the Dharmapada also found in Central Asia. The language of these texts has been denominated *Gāndhārī*, and recently announced finds of Buddhist texts appear to document its use at an earlier date than heretofore known. Not surprisingly, being a geographically marginal Indic language, it shows a number of aberrant features (and the influence of other Central Asian languages) not found in the “standard,” geographically more central Middle Indic Prākṛits.



**Figure 27.1** Part of scroll manuscript of the Anuvatapta-gatha or “Songs of Lake Anavatapta” in Gandhari Prakrit, c. first century AD

## 1.4 Prākṛits “proper” (Pkt.)

This term designates a number of different linguistic systems, which no doubt began as local dialects (as their separate names imply), but which have become geographically deracinated, stylized, and deemed appropriate to different genres and expressive functions. We have two major types of textual sources for these standard Prākṛits: (i) literary texts, including (a) epic and lyric poetry, and (b) the speech of most women and lower-born men in Classical Sanskrit dramas; and (ii) Jain religious texts. In addition, there is a tradition of Prākṛit

grammarians, paralleling that of the Sanskrit grammarians. The most common Prākṛits and their usages are as follows:

1. *Māhārāṣṭrī* (M.): poetic and literary Prākṛit, used also for songs in drama. Generally treated as the standard Prākṛit by grammarians.
2. *Śaurasenī* (Ś.): standard dramatic prose Prākṛit, spoken by high-born women and the *Vidūṣaka* (the king's buffoon, Brahmin by birth).
3. *Māgadhī* (Mg.): dramatic Prākṛit, spoken by low-born men. An Eastern dialect.
4. *Ardha-Māgadhī* (AMg.): Jain Prākṛit, language of the oldest parts of the canon.
5. *Jain-Śaurasenī* (JŚ.): language of the canon of the Digambara Jains.
6. *Jain-Mahārāṣṭrī* (JM.): noncanonical texts of the Śvetāmbara Jains.

In addition, snatches of other Prākṛits are found in the dramas. In the earliest dramas attested, the fragments of the Buddhist author Aśvaghoṣa found in Central Asia and dated to the first–second century BC, as well as the plays attributed to Bhāsa, the distribution and linguistic form of the Prākṛits are somewhat different from that encountered in Kālidāsa and later authors.

### 1.5 “Buddhist Hybrid Sanskrit”

Though early Buddhism was propagated in the vernacular, not in Sanskrit, as time passed various schools introduced Sanskrit or Sanskritized Prākṛit. Attested texts show different degrees of this Sanskritization; the term “Buddhist Hybrid Sanskrit” is especially appropriate to certain texts of the Mahāsāṃghika-Lokottaravāda school in North India, which (to oversimplify) show Sanskritic or hyper-Sanskritic phonology, but Middle Indic morphological traits. The extent to which the language of these texts was an actual spoken medium, rather than a result of textual hypercorrection, remains a topic of discussion.

### 1.6 Late Prākṛit (Apabhraṃśa)

Late Prākṛit falls beyond the chronological limits of this volume.

While it is customary to refer to the various forms of Middle Indic as “dialects,” the present treatment will ordinarily use “languages,” or the more neutral but awkward “speech forms.” Without appealing to mutual intelligibility or other such criteria, it seems condescending to apply the trivializing term “dialect” to linguistic systems used by such different social groups for such different purposes over a range of space and time.

## 2. WRITING SYSTEMS

The various Middle Indic languages are recorded using different writing systems. The inscriptions are for the most part in Brāhmī, except for those of the Northwest, where Kharoṣṭhī is found. The assemblage of speech forms denominated by the term *Gāndhārī* is also in Kharoṣṭhī. The other types of Middle Indic use various writing systems ultimately derived from Brāhmī. As noted in Chapter 26, §2, literary and religious texts were recorded in the appropriate local script, but nowadays will usually be printed in Devanāgarī, an offshoot of Northern Brāhmī.

### 3. PHONOLOGY

Since we are discussing not a single language, but a range of speech forms, all descended from Old Indic, we will give an overview of the characteristic Middle Indic developments from the Old Indic phonological system, rather than describing the synchronic phonological system of one (or more) Middle Indic languages. Accordingly, the Sanskrit phonological system described in Chapter 26, §3 is presupposed here.

#### 3.1 Vowels

1. All Middle Indic languages lose *r̥*, which is replaced by *a*, *i*, or *u* (e.g., Skt. *kr̥ta-* > Pā. *kata-*). There is evidence for this change already in the *Rig-veda*, which has, for example, *śithirā-* “loose” for etymological \**śṛthirā-*.
2. The long diphthongs *āi* and *āu* develop to *ē* and *ō* respectively. This change lends obscurity to the system of *vr̥ddhi* derivation (see Ch. 26, §3.4.3).
3. The sequences *-aya-* (*/-ayi-*) and *-ava-* (*/-avi-*) develop to *-ē-* and *-ō-* respectively. There is evidence for this change also in the *Rig-veda*, where the well-attested imperative *bo-dhí* “become” was remade from an original *bháva*.
4. One rhythmic rule had far-reaching effects on the grammatical system, the so-called *Zwei-Moren-Gesetz* (“Two-Mora-Rule”), whereby a long vowel before two consonants is not permitted. This blanket prohibition has variant manifestations: (i) the vowel can be shortened and the consonant cluster retained; (ii) the cluster can be simplified and the long vowel retained; (iii) an anaptyctic vowel may break up the cluster and the preceding long vowel be retained. The development of Skt. *dīrgha-* to Pkt. *diggha-* and *dīha-* illustrates (i) and (ii) respectively; that of Skt. *rājñā* to Pā. *raññā* and Áś. *lājñā* illustrates (i) and (iii). This rule has introduced two new phonemes into the language: the vowels *e* and *o*, which are always long in Sanskrit, develop short versions, originally in the position before two consonants. The resulting vowel system is thus more symmetrical than that of Sanskrit:

- (1) 

<i>i, ī</i>	<i>u, ū</i>
<i>e, ē</i>	<i>o, ō</i>
<i>a, ā</i>	

However, the backbone of the Sanskrit morphological system, the pattern of vowel gradation (see Ch. 26, §3.4.3), is seriously disturbed by these changes.

#### 3.2 Consonants

Middle Indic developments involve both individual segments and consonant clusters (compare Ch. 26, §3.3).

1. *Sibilants*: In most Middle Indic languages the three sibilants of Sanskrit (*s*, *ṣ*, and *ś*) merge, with the usual product *s* in the West and *ś* in the East (e.g., Skt. *śata* > Pā *sataṃ*, Mg. *śada*). Gāndhārī, however, keeps the three distinct.

2. *Liquids*: The two liquids merge, with *l* the usual Eastern product, *r* the Western: compare in the West (Girnar) Áś. *rājā* with *lājā* in the East (Jaugada).

3. *Glides*: Varying developments of *y* correspond roughly to chronological layers, though there are a number of aberrant changes. It is preserved in Áśoka and in Pāli, but in the

Prākṛits *y* is ordinarily lost between vowels (e.g., Skt. *priya* > Pā. *piya*-, Pkt. *pia*-) and often becomes *j* initially (Skt. *yadi* > Pā. *yadi*, Pkt. *jadi*, *jai*). The change to *j* is also sometimes found between vowels, with a *Verschärfung* to *-yy-* and then *-jj-*; for example, some Prākṛits have optatives in *-jja-* (from *\*-yā-*).

4. *Intervocalic single stops*: Pāli faithfully preserves voiceless, voiced, and voiced aspirated stops as distinct. In the Prākṛits, plain voiced stops are ordinarily lost between vowels (Skt. *hṛdaya*- > *hīaa*-). The more conservative Prākṛits (such as Śaurasenī) preserve some old intervocalic voiceless stops as voiced (e.g., Skt. *hita*- > Ś. *hida*-), while the more innovative (such as Māhārāṣṭrī) usually lose these too (M. *hia*-).

Old aspirates often lose their occlusion intervocalically in Prākṛit (Skt. *sakhī* > Pkt. *sahī*), and in some Prākṛits initially as well (Skt. *bhavati* > M. *hoi*). Again some conservative dialects voice *th* to *dh* between vowels (Skt. *atithi*- > Ś. *adidhi*-). The loss of occlusion in voiced aspirates is, of course, a sporadic feature of Sanskrit from the beginning (e.g., *hita*- past passive participle to *√dhā*).

As with many developments in the Prākṛits proper, it is difficult to formulate consistent rules within a single dialect, and even within a single text, because of dialect mixture, analogy, hypercharacterization on Sanskrit models, and scribal transmission. However, the “feel” of a particular Prākṛit and its stereotypical employment in literature is much affected by its treatment of intervocalic consonants. For example, the regular loss of most intervocalic consonants in Māhārāṣṭrī and the preservation of the resulting hiatus makes this speech form well adapted for its use in songs, but less suited for dialogue, due to the numerous homonyms created by these phonological changes (e.g., *maa* < Skt. *mata*, *mada*, *mṛta*, *mṛga*). The more conservative Śaurasenī is better fit for conveying meaning in a less ambiguous function, and so for dialogue.

5. *Final consonants*: All final consonants are lost, except *anusvāra* (*ṁ*, nasal offglide; see Ch. 26, §3.3), though in formulaic close nexus, consonants may sometimes be retained (e.g., Skt. *yad asti* > AMg. *jad atthi*). In some Prākṛits, final vowels (either original or produced by final-consonant loss) frequently acquire a nonetymological *anusvāra* (e.g., instr. sg. *-ena* > *-eṇaṁ*, instr. pl. *-ebhiḥ* > *-ehiṁ*).

6. *Final -as*: The outcome of this extremely common Sanskrit final is a shibboleth in Middle Indic: most dialects have *-o*, but Eastern Middle Indic has *-e* in the nominative singular (Aś. [Kālsī, Jaugaḍa], Mg., AMg.).

### 3.2.1 Consonant clusters

Probably the most conspicuous set of phonological changes spanning the Middle Indic languages involves the thoroughgoing assimilation in clusters, which significantly diminishes the transparency of the morphological system. An occasional alternative to assimilation is the insertion of an anaptyctic vowel. The assimilation rules involve a rough hierarchy of segment sonority: stops, nasals, sibilants, and (the sonorants) *l*, *v*, *y*, *r*— with a lower segment assimilating to a higher. When two segments belong to the same class, the first assimilates to the second.

1. *Stop + stop*: Total regressive assimilation occurs: for example, *-kt-* > *-tt-* (Skt. *mukta*- > *mutta*-), *-pt-* > *-tt-* (Skt. *sapta*- > *satta*-), *-dg-* > *-gg-* (Skt. *mudga*- > *mugga*), etc.
2. *Nasal + stop*: The nasal remains or becomes *anusvāra*.
3. *Stop + nasal*: The nasal assimilates (e.g., Skt. *agni*- > *aggi*-).

4. *Sibilant + stop*: The sibilant assimilates, but adds aspiration to the cluster (e.g., Skt. *asti* > *atthi*).
5. *Stop + sibilant*: The cluster *-ts-* ordinarily gives *-cch-* (e.g., Skt. *vatsa* > *vaccha*); *-kṣ-* gives either *-kkh-* or *-cch-* (e.g., Skt. *akṣi* > *akkhi/acchi*), originally distributed dialectally. A reflection of this change is probably already to be found in the *Rig-veda* in the form *akhkhalī(-kṛtya)* (sic), from *akṣara-*.
6. *Sonorant + stop / stop + sonorant*: Total assimilation of the sonorant to the stop occurs (e.g., *r*: Skt. *artha-* > *attha-*, *cakra-* > *cakka-*; *v*: *pakva-* > *pakka-*; *y*: *vākyā-* > *vakka-*. In the common combination of dental + *y*, there is palatalization of the cluster (e.g., *satya-* > *sacca-*, *adya-* > *ajja-*).
7. *Sibilant + nasal*: Ordinarily the outcome is nasal + *h*, with the aspiration also characteristic of sibilant + stop clusters (e.g., Skt. *grīṣma-* > *gimha-*).
8. *Sonorant + nasal / nasal + sonorant*: Total assimilation of the sonorant, as with stops (e.g., Skt. *anya-* > Pā. *añña-*, Pkt. *aṇṇa-*; Skt. *dharma-* > *dhamma-*).
9. *Sibilant + sonorant / sonorant + sibilant*: The sonorant assimilates to the sibilant (e.g., Skt. *aśva-* > *assa-*, *tasya-* > *tassa-*, *sahasra-* > *sahassa-*, *varṣa-* > *vassa-*).
10. *Sonorant + sonorant*: In general the lower sonorant assimilates to the higher, although there are a number of exceptions and special developments. In the hierarchy *l* prevails over the other sonorants (e.g., Skt. *durlabha-* > *dulla(b)ha-*). Next in strength is *v* (the labial glide), but clusters with *v* show some special developments. In Pāli *-rv-/ -vr-* and *-vy-* become *-bb-*, as opposed to the *-vv-* prevailing in the Prākritis (e.g., Skt. *sarva-* > Pā. *sabba-*, Pkt. *savva-*; *tīvra-* > *tibba-* / *tivva-*; *-tavya-* > *-tabba-* / *-tavva-*). Finally, *r* ordinarily submits to *y*, though again with some special developments: anaptyxis is fairly common (producing *-riy-*), and *-yy-* tends to develop to *-jj-* (e.g., Skt. *arya-* > Pā. *ayya-*, but Pkt. *ajja-*).

### 3.3 Phonotaxis

In most Middle Indic languages only one consonant is permitted initially, two intervocalically. Cluster simplification occurs after the assimilation processes described in §3.2.1. In the case Skt. *martya-*, for example, both *r* and *y* assimilate to the medial stop, which itself has undergone palatalization before *y*. A triconsonantal *\*maccca-*, the expected product of assimilation, simplifies to *macca-*. Similarly, Skt. *strī* is susceptible to both *r*-assimilation and *s*-assimilation with aspiration; the resulting *\*ttthī* simplifies to *thī* (and, with prothetic vowel, *itthī*). As this example shows, simplified single initial consonants can alternate with intervocalic geminates; thus, when simplex verbs are compounded with preverbs, the geminate cluster resurfaces (e.g., Skt. *kramati* > *kamati*, but *upa-kramati* > *upa-kkamati*).

### 3.4 Accent

Just as it had disappeared in the evolution of Sanskrit, so the Vedic pitch accent system did not survive in the Middle Indic languages. For the Vedic accent, see Chapter 26, §3.6.

### 3.5 The Middle Indic phonological system

Despite the global nature of the changes discussed in the preceding sections, they had essentially no effect on the phonological inventory of Middle Indic. All (or almost all) of the segments occurring in the Middle Indic languages already existed in Old Indic, and very

few segments or contrasts were eliminated (elimination occurred in the case of (i) the velar nasal  $\tilde{n}$ ; (ii) contrast between sibilants; (iii) contrast between liquids; (iv) in part, contrast between nasals; (v) the changes in the vowel system discussed above). However, the distribution of and phonotactic relations among segments have drastically changed in every Middle Indic language, and the syllable structure was entirely altered by the effects of the Two-Mora-Rule and the restrictions on length and types of clusters already discussed.

These changes also had dramatic repercussions in the morphological system. The elaborate but orderly morphophonemic alternations that pervade Sanskrit morphology (see Ch. 26, §3.4) were significantly obscured by the numerous consonantal assimilations, as well as by the loss of  $r$  and the shortening of  $v\ddot{r}$ ddhi vowels before two consonants. Some of the transparent variants of the root  $\sqrt{kṛ}$  in various morphological categories in Sanskrit and their Middle Indic equivalents provide a telling example:

(2)	<i>Sanskrit</i>	<i>Middle Indic</i>
	$kṛta-$	kata- (kada-, kida-, kaa-)
	$kṛṇoti$	kuṇadi
	$kṛtya$	kicca
	karma-	kamma-
	kartum	kātum (< expected *kattum)
	kārya-	kajja- (/kayya-)
	akārṣīt	akāsī
	kuryāt	kujjā

## 4. MORPHOLOGY

### 4.1 Word formation

As with phonology, the development of Middle Indic from Sanskrit involves the redistribution and reduction of existing categories, rather than the creation of new ones. Again, however, the result is superficially very distinct from its source. The structure of the word is theoretically as in Sanskrit: Root–Suffix(es)–Ending (see Ch. 26, §4.1). But phonological changes have conspired to make the root less consistently recognizable (as the above example of  $\sqrt{kṛ}$  demonstrates), and the morpheme-boundaries less delimitive than in Sanskrit. Because of its loss of salience, the root plays a far less prominent role in Middle Indic morphology than in Sanskrit, and in verbal forms the *present stem* is often the base of derivation.

It is again more convenient to describe the various developments of the Middle Indic languages from Sanskrit along a continuum, rather than producing a synchronic description of one Middle Indic speech form. Hence the following discussion presupposes the description of Sanskrit morphology found in Chapter 26, §4.

We noted the pervasive distinction between *thematic* inflection (invariant stems ending in *a*) and *athematic* inflection (stems, often alternating, ending in consonants or vowels other than short *a*) in Sanskrit (see Ch. 26, §4.1). With the loss of final consonants (as well as the general obscuring of morphophonemic relations), this distinction has become irrelevant in Middle Indic. Consonant-final noun stems are found only as marginal relics, and consonant-final verb-stems are essentially nonexistent. Almost all stems are nonalternating (again, except for relics), and endings are more uniform across stem-types. The formal



features that set thematic forms apart from athematic ones were generalized by all Middle Indic languages, and the formal inventory consists of different types of vocalic stems.

Middle Indic closely resembles Sanskrit with respect to affixation. Suffixation remains the dominant mode. The single true prefix of Sanskrit, the verb augment (*-a*), is still found in the past tense, but is not obligatory. The same inventory of preverbs is found as in Sanskrit with the same functions; indeed, tmesis is still possible in older Pāli. The single infix in Sanskrit (*-na/n-*) is only marginally present, frozen in the verbal inflection of a set of roots originally containing the infix.

Reduplication is no longer a functioning morphological process in Middle Indic.

## 4.2 Nominal morphology

The grammatical categories of Middle Indic nominal forms are gender, number, and case.

### 4.2.1 Gender

Three genders exist – masculine, neuter, and feminine. The differences are expressed, as in Sanskrit, primarily inflectionally between masculine and neuter, derivationally between masculine/neuter and feminine; that is to say, masculine and neuter nominals employ the same stem, but different endings (which differ only in the nominative/accusative), while the feminine is built to a different stem. In fact, this tendency has become more pronounced than in Sanskrit, in that the old short *i* and *u* feminines tend to fall together inflectionally with their long *ī* and *ū* counterparts, and become distinct from the short *i* and *u* masculine and neuter stems.

### 4.2.2 Number

There are only two numbers, Middle Indic having lost the dual. The plural takes its place.

### 4.2.3 Case

There are formally eight cases in older Middle Indic, as in Sanskrit: nominative, accusative, instrumental, dative, ablative, genitive, locative, vocative. However, the apparent identity of case systems masks a significant loss of strength in the distinctions among the cases.

As a general rule, the dative has been lost or almost lost in most Middle Indic languages. Formal expression of the dative is found only in the *a*-stem and only in Aśokan, Pāli, and a few Prākṛits (notably Māhārāṣṭrī and Ardha-Māgadhī), its function primarily restricted to expressing purpose. It is otherwise replaced by the genitive. This functional restriction and replacement by genitive we also noted in Sanskrit (see Ch. 26, §5).

The conflation of cases in many stems, also noted for Sanskrit (see Ch. 26, §4.2.3), has progressed much further in Middle Indic. For example, the instrumental and ablative plural of most stems coincide; for the most part, the feminine singular oblique cases employ only a single form, and so forth. Thus, the progressively developing tendency in Middle Indic is to distinguish the direct cases (nominative and accusative) from a less differentiated oblique inflection, though nowhere does this development come to completion. There is one major countercurrent: the creation of new ablative singulars with a *-tas* adverbial suffix, found also in Sanskrit.

Case is marked primarily by endings, since stem-form alternations have generally disappeared.

#### 4.2.4 Nominal stem-classes

The Sanskrit distinction between vowel and consonant stems has been almost entirely lost.

##### 4.2.4.1 Consonant stems

Consonant stems exist marginally, in relic forms and incomplete declensions. Most older consonant-stem nouns have been reformed as vowel stems. This is another legacy of the loss of final consonants; once this happened, old consonant-final stems could easily fit into the appropriate vocalic category (e.g., Skt. *caṣus-* becomes a neuter *u*-stem *cakkhu-*; *sarpis-* > *sappi-*, etc.). This reinterpretation was also favored by the ambiguity of certain common case forms. Thus, the masculine accusative singular of both consonant stems and *a*-stems ends in *-am*, allowing the abstraction of a new *a*-stem; for example, the Sanskrit *n*-stem *mūrdhan-*, with accusative singular *mūrdhānam* (= MI *muddhāṇam*), can backform a Prākṛit nominative singular *muddhāṇo* next to older *muddhā*.

##### 4.2.4.2 Vowel stems

These changes leave Middle Indic with a limited set of fully functioning vowel-final stems: masculine and neuter *a*-stems; feminine *ā-* and *ī-*stems; masculine and neuter *i-* and *u-*stems, continuing their Sanskrit counterparts. As in Sanskrit, the endings of the *a*-stems are somewhat aberrant; the feminine singular has a unique set of oblique endings; and the masculine and neuter are distinguished only in nominative/accusative.

#### 4.2.5 Endings

Because of the wide variation in endings both between and within Middle Indic languages, a generalized scheme of endings cannot be given here. Wholesale analogies and adaptations from other stems have operated in the generation of paradigms, especially when phonological change would have made the inherited ending inconvenient. A few examples show the types of rearrangements that occur:

1. The instrumental plural ending of the *a*-stems, *-ehi(ṃ)*, continues the Vedic Sanskrit alternant *-ebhis*, rather than Classical *-ais*, which would have become undercharacterized *\*-e*. In addition, the Vedic *a*-stem alternative nominative plural *-āsas* is attested in Pāli verse as *-āse*.
2. The accusative plural of masculine vowel stems would have fallen together with the accusative singular by regular sound change: *-am* and *-ān* > *-aṃ* (and probably *-im* / *-īn* > *-iṃ*, *-um* / *-ūn* > *uṃ*, though this is disputed). The *a*-stem has accusative plural *-e*, apparently adapted from the stem-vowel of the old oblique plurals (*-ebhis*, *-ebhyas*), though some scholars explain it as borrowed from the nominative plural of the demonstrative pronoun.
3. Short *i-* and *u-*stems show traces of stem alternation in the Pāli: nominative plural *aggayo* (< Skt. *agnayas*), stem *aggi-*; *bhikkhavo* (< Skt. *bhikṣavas*), stem *bhikkhu-*. But in the Prākṛits this has largely been replaced by forms of the type *aggiṇo* (from the *in*-stems) and *aggī* (from the accusative plural, itself probably borrowed from the feminine *i*-stems).

Relics of consonant-stem inflection are most conspicuous in conservative speech forms like Pāli, and there especially in the older layers of the language. Even so, only *r*-stems (both kinship terms and agent nouns), *n*-stems (esp. *rājan*- “king,” *ātman*- “self”), and *nt*-stems (pres. act. part.) preserve anything resembling a paradigm; the old *s*-stems offer relic instrumentals in *-asā*, but little else.

#### 4.2.6 Comparison of adjectives

The two forms of comparison found in Sanskrit (see Ch. 26, §4.2.6) are still utilized in Middle Indic. The simple adjective can also function as the comparative, and the comparative in *-tara-* tends to stand in for the superlative-*tama-*, at least in Pāli.

#### 4.2.7 Pronouns

As in Sanskrit, we can distinguish the nongender-marked personal pronouns of the first and second person and the gender-marking demonstrative/anaphoric types.

##### 4.2.7.1 Personal pronouns

Many of the idiosyncrasies of these pronouns in Sanskrit have been maintained and indeed built upon, producing an efflorescence of analogic confections across the range of Middle Indic languages. The distinction between nominative and oblique stem-forms is ordinarily kept in the first-person singular (*aham*, or forms based on it, vs. *m-*); in the plural, the oblique stem. *amh-* has generally spread to the nominative, though Pāli has *mayam*, built to old *vayam* (nom. pl.) with the initial of the singular oblique. In the second person, the plural has adopted the initial of the singular throughout, but otherwise keeps the plural stem (Skt. *yusm-* replaced by *tumh-*, after *t(u)vam*, etc.). The separate enclitic forms are also generally preserved in one form or another.

##### 4.2.7.2 Gender-marking pronouns

The Sanskrit anaphoric/demonstrative *sa*, *sā*, *ta-* also preserves many peculiarities of inflection through much of Middle Indic, including the distinctive distribution of *s-* and *t-*stems. The *ayam* and *asau* pronominal paradigms also remain alive, and the relative and interrogative stems, *ya-* and *ka-*, and the “pronominal adjectives” continue the Sanskrit forms.

### 4.3 Verbal morphology

The Middle Indic verbal system experienced a dramatic reduction in the formal categories which characterized the Sanskrit verb, and a consolidation of the functions of those which do remain, rather than the production of new categories. The motivations for these losses are often already to be seen in Sanskrit, where, as we noted, a number of separate formal categories were insufficiently distinguished functionally. The older Middle Indic languages often preserve relics of otherwise lost categories.

Formally alternating stems within paradigms have been virtually eliminated, and often the present stem has become the basis for all other categories.

The grammatical categories of finite verbal forms are person, number, tense, mood, and voice. There are also formal relics of distinct aspect stems in older Middle Indic, but without functional value. Generally, person and number are expressed by a portmanteau morpheme,

the ending; tense by suffixes and/or endings; mood also by suffixes and/or endings; voice by a suffix.

#### 4.3.1 Person and number

As in the noun, the dual was eliminated. The basic building block of the verbal system is thus the six-member paradigm: first, second, third person; singular and plural.

#### 4.3.2 Voice

As discussed in Chapter 26 (§4.3.2), Old Indic had a formal voice distinction, active versus middle (*parasmaipada* vs. *ātmanepada*), the functional differentiation of which was not always clear. This distinction has essentially been lost in Middle Indic; though some Middle Indic languages can employ old middle endings, they are used indifferently with the active, and active endings far outnumber them. Only in Aśoka are there some possible traces of a distinct *middle* use of middle endings.

Sanskrit also had a functional voice distinction, active versus passive, encoded in various ways. This distinction is maintained in Middle Indic. In addition, Middle Indic ordinarily opposes a causative stem to the simple present, and so it is useful to consider the grammar as displaying a three-way contrast: active, passive, causative.

The old Sanskrit *-ya-* passive suffix, often extended to *-īya-* or *-iyya-* (*-ijja-*), is added to either the root or the present stem. Unlike the Sanskrit passive stem, that of Middle Indic ordinarily takes active endings.

Since the Sanskrit *-aya-* causative falls together phonologically with other representatives of the expanding *ē*-stem present, a new causative suffix spreads in Middle Indic, based on the Sanskrit *p*-causative to roots ending in *ā*, Pā. *-(ā)paya-*, *-(ā)pē-*, Pkt. *-(ā)vē-*.

#### 4.3.3 Tense

As we saw in Chapter 26 (§4.3.3), Old Indic had a complex set of formal tense/aspect categories, some with low functional load. In particular, there were three competing past tenses – imperfect, aorist, perfect – as well as a possible nominal expression of past tense. From this elaborate system there emerges in Middle Indic a simple three-way expression of tense: present, future, and past. The present continues the old present and the future the old finite future. The past ordinarily continues the aorist in older Middle Indic, but the past passive participle in younger speech forms.

#### 4.3.4 Stem formation

In verbs, as in nouns, consonant-final stems have essentially disappeared, as have alternating stems within a paradigm. Therefore, a distinction between thematic (*a-*) stems and others is less relevant than a division into different varieties of vowel stems.

#### 4.3.5 Present stem morphology

Middle Indic present stems continue many of the Sanskrit types (see Ch. 26, §4.3.6), though transformed into invariant, vowel-final forms. However, the verb “to be” (Skt. *√as*) preserves consonant-final forms and even traces of stem alternation in many Middle Indic languages.

Compare the following Pāli forms, which differ from their Sanskrit equivalents only by regular phonological developments:

(3)	<i>Sanskrit</i>	<i>Pāli</i>
1st sg.	asmi	amhi
2nd sg.	asi	asi
3rd sg.	asti	atthi
3rd pl.	santi	santi

Other old root-class (Sanskrit class II) verbs are preserved only in traces in older Middle Indic.

The most numerous present type by far is the *a*-stem (incorporating Sanskrit classes I, IV, and VI), into which were attracted many old athematic presents (e.g., root *hanti* > *hanati*; reduplicated *dadhāti* > Pāli *dahati*; nasal infix *yunakti* > Pāli *yuñjati*). Sanskrit class X (*-aya-* verbs) and causatives and denominatives with the phonologically identical suffix produce a different stem-type in contracted *-ē-*, which is also widespread in Middle Indic and attracts old athematic presents to its inflection. Other old vowel-final stems are much rarer: for example, class VIII *karoti*, class IX *jānāti*.

Rather than listing numerous old Sanskrit presents and their outcomes, it is more instructive to examine the fate of a single present in various Middle Indic languages, that of the root  $\sqrt{kṛ}$  “make, do.” In early Vedic this formed a class V present, *kṛnoti*, *kṛnute*, soon replaced by the prevailing Sanskrit class VIII *karoti*, *kurute*. Traces of both these present types are found in Middle Indic: *kṛṇo-/ṇu-* in Pkt. *kuṇa-* (reformed according to the *a*-class); the strong stem *karo-* in Pā. *karoti*, the weak stem *kuru-* in Pā., Pkt. *kubba-/kuvva-* (< prevocalic *kurv-*, with *a*-stem inflection). In addition, Pāli and Prākṛit regularly inflect the simple root as an *a*-stem (*kara-*), and Prākṛit also as an *ē*-stem (*karē-*). Here and in many other cases, Middle Indic absorbed the leftovers of the formally diverse Sanskrit present system and redistributed them in a very few classes.

The productive future is built to the present stem, with the continuation of the Sanskrit future suffix which is proper to *seṭ* roots, Skt. *-īśya-* > *-issa-*, though in older Middle Indic we find some examples of old *-śya-* added directly to roots.

#### 4.3.6 Preterite stem morphology

The preterite in older Middle Indic (esp. Aśoka, Pāli, also more rarely in Ardha-Māgadhī, etc.) ordinarily continues an aorist form, though there are occasional traces of the other competing Sanskrit preterites. The old imperfect is preserved in *āsī* (Pāli and some Prākṛits, < *āsīt*) to the root  $\sqrt{as}$  “be” (which did not form an aorist in Sanskrit); and a few relics of the perfect remain, notably Aś., Pā., Pkt *āhu(ṃ)* to the root  $\sqrt{ah}$  “say,” Pā. *vidu(ṃ)* to  $\sqrt{vid}$  “know.”

Otherwise, forms of the old sigmatic aorists prevail, though reformed, redistributed, and usually attached to the present stem. Present stems ending in short vowels add *-i*, derived from the *-īś-*aorist (Pā. *pucchatī* “asks”: *pucchi* “asked”), while those in long vowels add *-sī*, from the *s*-aorist (Pā. *kathēti* “tells”: *kathēsi* “told”). Relics of other aorist types are also found (e.g., root aorist, Skt. *adāt* > Pā. *adā*).

The Sanskrit augment, which marked both imperfect and aorist (optionally in Vedic, obligatorily in Classical Sanskrit), is also found optionally in Middle Indic and seems especially utilized in shorter forms. The preterite is characterized by endings ultimately derived from the old Sanskrit secondary endings.

Finite forms of the preterite survive only as relics in Prākṛit, where the standard means of expressing past tense is the predicated past passive participle (with or without the copula), a periphrastic method commonly found already in Sanskrit and older Middle Indic.

### 4.3.7 Mood

Middle Indic attests three moods: indicative, imperative, and optative, all built to the present stem. These continue their Sanskrit counterparts both formally and functionally. The subjunctive has been lost (as in Classical Sanskrit), except for a few possible relics in early Middle Indic.

#### 4.3.7.1 Imperative

The imperative is marked entirely by special endings on the present stem. As in Sanskrit, the negative imperative is expressed with the special negation *mā* and the unaugmented preterite in older Middle Indic speech forms.

#### 4.3.7.2 Optative

The Middle Indic optative is ordinarily built to the present stem with the suffix *-eyy(a)-* (Pkt. *-ejja-*), most likely derived from the Sanskrit thematic optative (in prevocalic position: e.g., 1st. sg. *bhavey-am*). Traces of athematic optatives (note esp. Aś., Pā., AMg. *siyā* < Skt. *s(i)yāt*, √as “be”) and of preconsonantal thematic forms are also found.

### 4.3.8 Verb endings

The entire range of endings found in Middle Indic cannot be treated here, but a few general facts can be noted. The primary active endings are the major set retained from Old Indic. These are used for the present (including the passive and causative) and the future and have been preserved with remarkable fidelity:

(4)	<i>Sanskrit</i>	<i>Pāli</i>	<i>Prākṛit</i>
1st sg.	-mi	-mi	-mi
2nd sg.	-si	-si	-si
3rd sg.	-ti	-ti	-di (Ś.), -i (M.)
1st pl.	-mas	-ma	-mo
2nd pl.	-tha	-tha	-dha (Ś.), -ha (M.)
3rd pl.	-(a)nti	-(a)nti	-(a)nti

As noted before, middle (primary) endings are encountered sporadically, without distinctive function. The Sanskrit secondary endings (of imperfect and aorist) are continued, though less transparently and systematically, in the endings of the aorist and optative (the fact that they ended in consonants for the most part contributed to their transformation). The endings of the Sanskrit imperative are also rather well preserved, especially the distinctive third person *-(n)tu*.

### 4.3.9 Nonfinite verbals

Like Sanskrit, Middle Indic deploys a number of verbal nouns and adjectives, some built directly to the root, some to tense stems. They display verbal case-syntax and often can serve

as predicates. In the case of the verbal nouns, the formal connection between the formants of these frozen nominals and synchronic case endings is not as clear as in Sanskrit.

#### 4.3.9.1 *Infinitive*

Middle Indic knows continuators of *-tum* and Vedic *-tave*. The former can be built to the present stem as well as to the root.

#### 4.3.9.2 *Gerund*

This is a very well-developed formation in Middle Indic, with a number of suffixes, some continuing Sanskrit *-tvā* and *-(t)ya*. Again, these can be built to present stems as well as roots. The Classical Sanskrit rules for the distribution of *-tvā* and *-(t)ya* do not rigidly hold in Middle Indic.

#### 4.3.9.3 *Tense-stem participles*

The active present participle suffix of Sanskrit, consonant stem *-ant-*, is common in Middle Indic to all types of present stems (including the passive). It is regularly thematized as an *a*-stem (yielding *-anta-*), though abundant relics of the old athematic paradigm are found in older Middle Indic. The Sanskrit middle present participle suffix *-māna-* is also widely attested, built to originally active stems and with an “active” meaning. The old perfect participle *vid-vas-* “wise” (weak stem *vid-us-*) survives in the Pāli *u*-stem adjective *vidū-*.

#### 4.3.9.4 *Past passive participle*

As in Sanskrit, this is an extremely common form and, as noted, the basis for the preterite tense in younger Middle Indic. Its suffixes continue *-ta-*, *-ita-*, and *-na-*, and it is the part of the verbal system that most successfully resists the tendency to substitute the present stem for the root, though past passive participles built on tense stems are also quite common. Numerous Middle Indic past passive participles are historically identical to their Sanskrit counterparts, with regular phonological developments.

#### 4.3.9.5 *Past active participle*

This verbal (Skt. *-tavant-*) is also found in Middle Indic, though less commonly.

#### 4.3.9.6 *Gerundive*

The Middle Indic gerundive occurs commonly, with continuators of Sanskrit *-tavya-*, *-anīya-*, and *-ya-*. It too can be formed to the present stem as well as to the root.

### 4.4 Compounds

Middle Indic displays the same varieties of compounds as Sanskrit and employs them regularly and productively.

As in Sanskrit, verbs incorporate preverbs into both finite and nonfinite forms. In addition the *cvī*-formation (compounding of noun or adjective with forms of  $\sqrt{kr}$  “make” or  $\sqrt{bhū}$  “become”) continues (e.g., Pā. *udakī-bhū-* “consist of water”: *udaka-* “water”).

The three major types of Sanskrit nominal compounds – copulative, determinative, and possessive (dvandva, tatpuruṣa, and bahuvrīhi) – are present in Middle Indic, though the baroque exuberance of Classical Sanskrit multiple compounding is restrained.



## 4.5 Numerals

The numerals ordinarily continue the Sanskrit forms, with appropriate sound changes:

### (5) The Pāli cardinals

1	<i>eka-</i>
2	<i>d(v)i-</i>
3	<i>ti-</i>
4	<i>catur-</i>
5	<i>pañca</i>
6	<i>cha</i>
7	<i>satta</i>
8	<i>aṭṭha</i>
9	<i>nava</i>
10	<i>dasa</i>
20	<i>vīsati</i>
100	<i>sata-</i>
1,000	<i>sahassa-</i>

Prākṛit cardinals correspond to the above for the most part.

In Pāli both *ti-* (“three”) and *catur-* (“four”) maintain the archaic inflectional features found in Sanskrit, including distinction between strong and weak cases (Pā. masc. nom./acc. pl. *tayo*, gen. *tiṇṇaṃ*; *cattāro*, *catunnaṃ*) and the unusual feminine formant *-ss-* between stem and ending (thus Pā. fem. nom./acc. pl. *tisso*, *catasso*). Continuators of these forms are found in the Prākṛits, although not usually distributed systematically in a paradigm.

Ordinals are derived from cardinals as in Sanskrit, with the irregular forms preserved:

### (6) The Pāli ordinals

1st	<i>paṭhama-</i>
2nd	<i>dutiya-</i>
3rd	<i>titiya-</i>

## 5. SYNTAX

Most observations made regarding Sanskrit syntax are equally applicable to Middle Indic, especially in its older layers. Certain stylistic devices are especially well developed; for example, the piling up of numerous nonfinite clauses, each typically containing a final gerund, completed by a final verb having the same subject as the preceding clauses, is especially characteristic of Middle Indic prose style. On the other hand, the excessively nominal style of later Classical Sanskrit is usually avoided, except in genres founded directly on Sanskrit models.

In the younger layers of Middle Indic the most important syntactic development is the replacement of the old preterite(s) with forms of the past passive participle. Since the agent of the preterite is then expressed in the instrumental case and the patient in the nominative (as opposed to the active syntax – nominative agent, accusative patient – of the present and future tenses), the stage is set for the *split ergative* systems that arise in many modern Indo-Aryan languages. It also goes without saying that this originally participial form agrees with the grammatical subject in number and gender and takes nominal rather than

verbal endings. As in Sanskrit, the copula with the third person is optional and, in fact, rare.

## 6. LEXICON

The majority of the Middle Indic lexicon is derived from Sanskrit. In addition to words inherited from Sanskrit, and therefore undergoing regular phonological development, other words were borrowed directly from Sanskrit, which has served as a cultural word-hoard for all Indian languages (Dravidian as well as Indo-Aryan) through their histories. This distinction between actual inheritances and internal borrowings is similar, though not identical, to that made by the Prākṛit grammarians between *tadbhavas* and *tatsamas*. The latter (“same as that”) refers to words that are identical to their Sanskrit counterparts; this includes not only internal borrowings, but also inherited words, like *bhara*, *nāma*, etc., that have the same form in both Sanskrit and Middle Indic because no sound laws have applied. *Tadbhavas* (“arisen from that”) are Middle Indic words that show Sanskrit elements modified by normal sound change.

A third class of words is identified by the grammarians as *deśya* (*/deśī*), “belonging to the country, provincial” – words that do not have obvious Sanskrit equivalents. Etymologically these no doubt include numerous inherited Indo-Aryan words that lack Old Indic counterparts, as well as words borrowed from non-Indo-Aryan languages. Here, as with Sanskrit, identifying the source in a particular case is often difficult; though, especially in Northwest Prākṛit loanwords from Iranian, Greek, and Central Asian languages are identifiable.

The Middle Indic languages used primarily for religious texts, Buddhist or Jain, also have large stocks of technical religious terms. Though many of these terms are Sanskrit in origin, they have developed senses not found in their Sanskrit sources. At the opposite extreme are the dramatic Prākṛits, which can often be rendered, morpheme by morpheme, into intelligible Sanskrit simply by reversing the sound laws. In fact, most dramas have been provided with such translations, known as *chāyās* (“shadows”), much used by students.

## 7. READING LIST

Invaluable is the up-to-date, comprehensive, and detailed historical survey of von Hinüber a 2001. Also useful are the surveys of Bloch (1965) and Masica (1991), though both are more concerned with modern Indo-Aryan. Cardona 1987 gives a necessarily brief sketch of Middle Indic developments. There is no general dictionary of Middle Indic, but Turner 1966–1969 is often useful.

For individual languages, the standard work on Pāli is Geiger 1916, revised and edited by K.R. Norman (1994) and a useful introduction is provided by Warder 1963. For the Prākṛits, Pischel 1900 is standard, but difficult to use. The English translation (1981) provides a much needed *index verborum* (but omits the table of contents). A brief but well-organized introduction is provided by Woolner 1928. For Buddhist Hybrid Sanskrit the standard work is Edgerton 1953.

## Bibliography

Bloch, J. 1965. *Indo-Aryan from the Vedas to Modern Times*. Translated and revised by A. Master. Paris: Adrien-Maisonneuve.

- Cardona, G. 1987. "Indo-Aryan languages." In B. Comrie (ed.), *The World's Major Languages*, pp. 440–447. New York: Oxford University Press.
- Edgerton, F. 1953. *Buddhist Hybrid Sanskrit Grammar and Dictionary*. New Haven: Yale University Press.
- Geiger, W. 1916. *Pāli Literatur und Sprache*. Strasburg: Trübner. Translated into English by B. Ghosh, 1943; revised and edited by K. R. Norman, 1994 as *Pāli Grammar*. Oxford: Pali Text Society.
- Hinüber, O. Von 2002. *Das ältere Mittelindisch im Überblick* (2nd expanded edition). Vienna: Österreichische Akademie des Wissenschaft.
- Masica, C. P. 1991. *The Indo-Aryan Languages*. Cambridge: Cambridge University Press.
- Pischel, R. 1900. *Grammatik der Prakrit-Sprachen*. Strasburg: Trübner. Translated into English by S. Jha, Delhi: Motilal Banarsidass, 1981.
- Turner, R. L. 1966–1969. *A Comparative Dictionary of the Indo-Aryan Languages*. London: Oxford University Press.
- Warder, A. K. 1963. *Introduction to Pali*. Oxford: Pali Text Society.
- Woolner, A. C. 1928. *Introduction to Prakrit* (2nd edition). Reprint, Delhi: Motilal Banarsidass.

# Old Persian

RÜDIGER SCHMITT

## 1. HISTORICAL AND CULTURAL CONTEXTS

Old Persian is one of two Old Iranian languages which are attested in the Achaemenid royal inscriptions (see below), members of that branch of the Indo-European language family called Indo-Iranian, or Aryan (the Persians designate themselves and their language by the term *ariya-*). The Iranian languages began to take shape when the ancestors of the Indo-Aryans left the common homeland in the steppes of Central Asia in the first half of the second millennium BC. The Western Iranian peoples, the Medes who settled in Media and the Persians in Fārs (speaking a Northwestern and Southwestern Iranian dialect respectively), step into the light of history in the ninth century BC, when Median names are first attested in Assyrian documents.

While “Old Persian” was certainly the language of Fārs, the variety which is attested in the Achaemenid inscriptions appears to be a rather artificial idiom, peppered with dialectal and archaic words, unlike any dialect actually spoken (characteristics of a distinct spoken Old Persian may be discerned from certain spontaneous phonetic developments, and from Old Persian words and names as rendered in other languages). The language called Old Persian was thus restricted to royal usage (as was the cuneiform script in which Old Persian was recorded). Even so, Old Persian was neither the *lingua franca* nor the administrative language of the Achaemenid Empire, roles fulfilled by Aramaic and, to a limited extent, various regional languages spoken within the empire. As a consequence, the linguistic situation of the empire was a quite complex one; and epigraphical Old Persian was itself influenced by these other languages, particularly in its vocabulary and even syntax (e.g., in the occurrence of a postpositive genitive, as in *xšāyaθiya xšāyaθiyānām* “king of kings” or *vašnā Auramazdāha* “by the favor of Auramazdā”).

The language of the Old Persian inscriptions is dialectologically homogeneous in principle. Only some lexical items (technical terms, etc.) prove to be borrowed from other Iranian languages, mainly the Northwestern Iranian dialect of the Medes (see §6), the political predecessors of the Persian Achaemenids.

The only direct and authentic sources available for the Old Persian language are the cuneiform inscriptions on durable objects (rock, stone, metal, rarely clay tablets) ranging over the period from Darius I (522–486 BC) to Artaxerxes III (359/8–338/7 BC), but dating in the main from the reigns of Darius I and Xerxes I (486–465 BC). In this short period the inscriptions, for the most part, are trilingual (in Old Persian, Elamite, and Babylonian), but even the oldest text, the one of the Bisutūn monument of Darius I (see below), has sections which are only in Old Persian, or in Old Persian and Elamite. With Artaxerxes I

(465–425/4 BC) the number, size, and significance of the texts begin to decrease rapidly, and they consist almost exclusively of stereotyped formulae, which, in part, seem to have been poorly understood at the time of composition. On the other hand, however, apart from their trilingualism, it is just this monotonous stereotyped style of the texts, along with the great number of parallel texts with their often-repeated invocations of the supreme god and with the regularly quoted royal titles, that has facilitated an understanding of the language and texts and which has allowed reconstruction of fragmentary texts. The abbreviatory system of citing texts is presented at the end of the chapter.

The decreasing number of Old Persian texts after the reign of Xerxes I may be attributed to a loss of fluency with the royal language. By that period, spoken Persian had evolved into a somewhat different form, so discrepancies between everyday speech and the traditional language of inscriptions had arisen. Only upon that basis can the serious grammatical faults which appear in the texts of later Achaemenid kings (mainly of Artaxerxes II and III) be understood.

Most of those “corrupt” forms (incorrect endings, hybrid genitive forms, etc.) can be found in the monolingual inscription A<sup>3</sup>Pa of Artaxerxes III; but they also occur in most of the inscriptions of Artaxerxes II and in the monolingual texts claiming to have been composed by Ariaramnes and Arsames in the sixth century BC (that these texts were produced under Artaxerxes III instead, is suggested by the fact that among the later Achaemenids it is only this king who derives his lineage from Arsames, and not only from Darius’ father Hystaspes). The use of a form like *būmām* in lieu of the expected accusative singular feminine *būmīm* “earth” can best be explained by positing an actually spoken monosyllabic [bu:m] (like Middle Persian *būm*) and a scribal attempt to “transform” the spoken form into an Old Persian one (an attempt which was rendered detectable by its lack of success, as it used the *ā*-stems as the normal class of feminine nouns). A similar archaizing process is seen in the pseudo-Old Persian accusative singular *šāyatām* for expected *šiyātīm* “happiness,” where the later form *šāt* has been changed into *šāyat-* by reversing the regular sound change of Old Persian *āya* to Middle Persian *ā* (though being inappropriate here) and adding again the ending *-ām* of the feminine *ā*-stems.

## 2. WRITING SYSTEM

### 2.1 Graphemic shape and inventory

Old Persian texts are recorded only in a cuneiform script. This script does not, however, directly continue the Mesopotamian cuneiform tradition (see Ch. 8, §2), being similar to the other cuneiform systems only in the employment of “wedge-shaped” characters. In other words, the Old Persian script is not the result of an evolution of the Mesopotamian system, but a deliberate creation of the sixth century BC. It remains unclear why the Persians did not take over the Mesopotamian system in earlier times, as the Elamites and other peoples of the Near East had, and, for that matter, why the Persians did not adopt the Aramaic consonantal script (Aramaic being the lingua franca of the Persian Empire; see §1).

Old Persian cuneiform was used only by the Achaemenid kings for two centuries and only for their own language – that is, the rather artificial literary language of their royal inscriptions. The use of this script was thus in effect a royal privilege. It was a splendid and imposing script best suited for hard surfaces, and apparently used neither for poetic texts nor for administrative nor historical writings.

**Table 28.1 The Old Persian cuneiform script***Syllabic symbols*

𐎠	𐎡	𐎢								
a	i	u								
𐎣	𐎤	𐎥	𐎦	𐎧	𐎨	𐎩	𐎪	𐎫	𐎬	𐎭
b <sup>(a)</sup>	c <sup>(a)</sup>	ç <sup>(a)</sup>	d <sup>(a)</sup>	f <sup>(a)</sup>	g <sup>(a)</sup>	h <sup>(a)</sup>	j <sup>(a)</sup>	k <sup>(a)</sup>	l <sup>(a)</sup>	m <sup>(a)</sup>
𐎮	𐎯	𐎰	𐎱	𐎲	𐎳	𐎴	𐎵	𐎶	𐎷	𐎸
n <sup>(a)</sup>	p <sup>(a)</sup>	r <sup>(a)</sup>	s <sup>(a)</sup>	š <sup>(a)</sup>	t <sup>(a)</sup>	θ <sup>(a)</sup>	v <sup>(a)</sup>	x <sup>(a)</sup>	y <sup>(a)</sup>	z <sup>(a)</sup>
𐎹	𐎺	𐎻	𐎼							
d <sup>i</sup>	j <sup>i</sup>	m <sup>i</sup>	v <sup>i</sup>							
𐎽	𐎾	𐎿	𐏀	𐏁	𐏂	𐏃	𐏄			
d <sup>u</sup>	g <sup>u</sup>	k <sup>u</sup>	m <sup>u</sup>	n <sup>u</sup>	r <sup>u</sup>	t <sup>u</sup>				

*Logograms*

𐎠𐎣	𐎠𐎤	𐎠𐎥	𐎠𐎦	𐎠𐎧
XŠ	DH <sub>1</sub>	DH <sub>2</sub>	BG	BU
xšāyaθiya-	dahyu-	dahyu-	baga-	būmī-
“king”	“land”	“land”	“god”	“earth”
𐎠𐎨	𐎠𐎩	𐎠𐎪		
AM <sub>1</sub>	AM <sub>2</sub>	AMha		
Auramazdā	Auramazdā	Auramazdāha		
		(genitive singular)		

The total number of phonetic characters (which consist of two to five single elements) is thirty-six. These are naturally divided into four groups:

- (1) A. Three pure vowel (V) characters: *a*, *i*, *u*
- B. Twenty-two syllabic characters whose vowel component is *a* (C<sup>a</sup>), but which can also be used to represent a consonant occurring before another consonant or in word-final position (C): *b<sup>(a)</sup>*, *c<sup>(a)</sup>*, *ç<sup>(a)</sup>*, *d<sup>(a)</sup>*, *f<sup>(a)</sup>*, *g<sup>(a)</sup>*, *h<sup>(a)</sup>*, *j<sup>(a)</sup>*, *k<sup>(a)</sup>*, *l<sup>(a)</sup>*, *m<sup>(a)</sup>*, *n<sup>(a)</sup>*, *p<sup>(a)</sup>*, *r<sup>(a)</sup>*, *s<sup>(a)</sup>*, *š<sup>(a)</sup>*, *t<sup>(a)</sup>*, *θ<sup>(a)</sup>*, *v<sup>(a)</sup>*, *x<sup>(a)</sup>*, *y<sup>(a)</sup>*, *z<sup>(a)</sup>*
- C. Four syllabic characters with inherent *i* vowel (C<sup>i</sup>): *d<sup>i</sup>*, *j<sup>i</sup>*, *m<sup>i</sup>*, *v<sup>i</sup>*
- D. Seven syllabic characters with inherent *u* vowel (C<sup>u</sup>): *d<sup>u</sup>*, *g<sup>u</sup>*, *k<sup>u</sup>*, *m<sup>u</sup>*, *n<sup>u</sup>*, *r<sup>u</sup>*, *t<sup>u</sup>*

In addition, there are eight logograms for commonly used words such as “king,” “god” or “land”; these are not obligatory and are not used consistently. The logograms are of a more complex shape, contain up to twelve elements and even show angles placed above angles (as is the case with the numerals). Further, a word-divider is used as well as number symbols (vertical wedges for the units, angles for the tens, and a special symbol for 100 (found in a single inscription).

One of the remarkable stylistic features of Old Persian cuneiform is that the wedges and angles which make up the cuneiform symbols never cross. The attested characters (excluding the numerals and the word-divider) are presented in Table 28.1.

Within the relatively short period of its use this writing system shows a few changes in character shapes – an attempted standardization of the height of those wedges which at first

(i.e., in the Bisutūn text) took up only half the height of the line. However, the mechanics of the writing system (see below), with all its “imperfections,” remain unchanged.

## 2.2 Orthographic conventions

As the set of CV characters with inherent *i* or *u* vowel shows, the inventory as a whole is inconsistent and asymmetric in its structure, for no ascertained reason (phonetic or otherwise):

(2)	d <sup>a</sup>	g <sup>a</sup>	j <sup>a</sup>	k <sup>a</sup>	m <sup>a</sup>	n <sup>a</sup>	r <sup>a</sup>	t <sup>a</sup>	v <sup>a</sup>
	d <sup>i</sup>		j <sup>i</sup>		m <sup>i</sup>				v <sup>i</sup>
	d <sup>u</sup>	g <sup>u</sup>		k <sup>u</sup>	m <sup>u</sup>	n <sup>u</sup>	r <sup>u</sup>	t <sup>u</sup>	

Beyond this, there are no C<sup>i</sup> and C<sup>u</sup> characters of the form *b<sup>i/u</sup>*, *c<sup>i/u</sup>*, *ç<sup>i/u</sup>*, *f<sup>i/u</sup>*, *h<sup>i/u</sup>*, *l<sup>i/u</sup>*, *p<sup>i/u</sup>*, *s<sup>i/u</sup>*, *š<sup>i/u</sup>*, *ṣ<sup>i/u</sup>*, *x<sup>i/u</sup>*, *y<sup>i/u</sup>*, *z<sup>i/u</sup>*. Even if the writing system were not plagued by such omissions, the ambiguity of many spellings would not be eliminated; the entire group of C<sup>a</sup> graphemes has its own affiliated spelling difficulties, which reveal that this writing system is neither phonemic nor phonetic.

As a consequence of the preceding graphemic problems, a number of *orthographic conventions* had to be employed when particular phonemic sequences are written. The most important of these “rules” (to the extent that they can be identified with certainty) are the following:

1. Long vowels are not distinguished from short ones except for *ā* in medial position.
2. Proto-Iranian final *\*-a* is written with an additional <a> (i.e., as <-C<sup>a</sup>-a>), though in all probability this indicates an actual lengthening of the vowel.
3. The vowels *ī* and *ū* are written with the vocalic characters <i> and <u>, and medially with an additional preceding <C<sup>i</sup>> or <C<sup>u</sup>> sign (when available, otherwise <C<sup>a</sup>> is used).
4. Final *-ī* and *-ū* are written with an additional semivowel as <-i-y> and <-u-v> respectively.
5. The “short” diphthongs *ai* and *au* are written <-C<sup>a</sup>-i->, <-C<sup>a</sup>-u-> (in final position extended by <-y>, <-v>) and therefore can be only partially distinguished from simple vowels (namely, <d<sup>a</sup>-i> = *dai*, but <d<sup>i</sup>-i> = *dī* or *dī̄*, whereas <t<sup>a</sup>-i> = *tai* and *tī* or *tī̄*).
6. The so-called “long” diphthongs *āi* and *āu* are written <-C<sup>a</sup>-a-i->, <-C<sup>a</sup>-a-u-> and are thus unambiguous (except in initial position according to 1).
7. Syllabic *ṛ*, which in all probability was pronounced as [ər], is written with consonantal <r> as <C<sup>a</sup>-r-C<sup>x</sup>> (= C<sub>ṛ</sub>C) in medial position, and as <a-r-> (= *ṛ*-) word-initially (where it cannot be distinguished from *ar*- and *ār*-).
8. The nasal consonants *m* and *n* are written before consonants only in special cases, like *mn* in <k<sup>a</sup>-m-n<sup>a</sup>-> = *kamna*- “few”; otherwise they are not written, so that <b<sup>a</sup>-r<sup>a</sup>-t-i-y> spells *baranti* “they bear” as well as *barati* “(s)he bears.”
9. In word-final position the only consonants which appear are *-m*, *-r*, and *-š*. Thus, while final *-m* is commonly written, as in <a-b<sup>a</sup>-r<sup>a</sup>-m> = *abaram* “I brought,” final *-n* (from Proto-Iranian *\*-n* and ultimately from *\*-nt*) is omitted: <a-b<sup>a</sup>-r<sup>a</sup>> = *abaran* “they brought.”
10. The postconsonantal glides *y* and *w* are usually written <-i-y> and <-u-v> (with <-C<sup>i/a</sup>-i-y-> spelling [Ciy]).



11. Early Iranian \**h* (from Indo-Iranian \**s*) is omitted in writing before Old Persian *š*, *m*, and *r* (cf. <a-u-r- > = *Aura-*, equivalent to Avestan *ahura-* “lord”), apparently reflecting its phonetic status in the particular Old Persian dialect, on which the inscriptional language is based.
12. The Early Iranian cluster \**hw* is likewise spelled as Old Persian <u-v> (by 10 and 11).
13. The vowel *ī* is commonly omitted after the *h* sign, though not without exception, as in <h-i-d<sup>u</sup>-u-š> = *Hinduš* “Indus.”

Given the cumbersome nature of the writing system, clear, one-to-one correspondences between graphs and phonemes do not exist. Some of the above spelling rules result in critical morphology being hidden, particularly rule 5 (e.g., the absence of a distinction between *tai* and *tī* means that third singular, indicative present endings, active *-ti* and mediopassive *-tai* cannot be distinguished) and rule 8 (the omission of preconsonantal *n* blurs, for example, the distinction between the third-person singular and plural endings *-ti*, *-tu* and *-nti*, *-ntu*).

The ambiguous nature of Old Persian spelling means that there is normally some set of possible interpretations of a word. In any particular case then a correct reading is dependent upon careful philological and linguistic (in particular, etymological) analysis – chiefly by comparison with cognate languages (Avestan, Vedic, etc.) or with later Persian developments. In the case of names and technical terms, the forms which they take in Elamite and Babylonian versions of an Old Persian inscription plays a decisive role. For example, the Old Persian spelling <a-s<sup>(a)</sup>-t<sup>(a)</sup>-i-y<sup>(a)</sup>> “is” has, according to the above rules, seventy-two possible readings. Only from Avestan *asti*, Vedic *ásti*, Middle and Modern Persian *ast*, and so forth, does it become clear that the correct interpretation of this sequence is *a-s-t-i-y*, that is, *asti*. That the geographical name spelled <k<sup>(a)</sup>-p<sup>(a)</sup>-d<sup>(a)</sup>> is to be read *Kampanda* (with two nasals omitted in the spelling by rule 8 above) can be ascertained by the Elamite rendering *Ka-um-pan-taš*. Things are not, however, always so simple; a great number of uncertain readings remain unresolved, among them, for example, the second syllable of King Cambyses’ Persian name.

It is important to distinguish sharply between graphic and phonemic (and eventually phonetic) units in the publication of Old Persian inscriptions and discussion of lexical or grammatical problems. Most of the existing manuals (text editions, grammars, etc.) use a “normalizing” interpretation – a kind of blend of the graphic and the phonemic which often is determined by the views about Old Persian held by the particular scholar, her/his scholarly tradition, or her/his time.

### 2.3 Origin of the script

The problems of the origin of the Old Persian cuneiform script, of the date and process of its introduction, have been treated again and again without general agreement having yet been reached concerning the controversial issues. There are several factors that one must take into account:

1. The passage DB IV 88–92, in which a new “form of writing” (Old Persian *dipiciṣam*) is mentioned that Darius has made and is said to be *ariyā* “in Aryan.”
2. A number of archeological and stylistic observations regarding the Bisutūn monument, by which several subsequent stages in its genesis may be established.
3. Those Old Persian inscriptions that are supposed or claimed to predate Darius I.
4. The structural analysis of the script itself.

Though the oldest attested inscriptions in Old Persian language are the Bīsutūn texts (first the minor captions, then the major inscription), the creation of a new type of writing for recording the king's mother tongue seems to have begun already under Cyrus II. This assumption is based not least on the observation that the characters *k*<sup>u</sup> and *r*<sup>u</sup> needed for writing the royal name *Kuruš* must belong to some initial set of characters, for their shapes have a quite simple pattern, even though the phonemic sequences expressed by them are not very common. A similar observation reveals that this writing system was created for the Old Persian language and not for some other Iranian dialect like Median: the fricative *ç*, which is the Old Persian reflex of Proto-Iranian \**ʧr* and which was foreign to Median, likewise is represented by one of the simplest characters, which must have been among the earliest of signs created.

A number of striking features appear to suggest that the *invention* of the script indeed began under Cyrus, but that Darius was the first to employ it. An original strategy seems to have aimed at a consistent and unambiguous system of marking short and long vowels and diphthongs by means of a complete set of three CV characters – for each consonant – used in conjunction with three V signs; for example:

- (3)    \* <b<sup>a</sup>> = *ba*                      \* <b<sup>i</sup>> = *bi*                      \* <b<sup>u</sup>> = *bu*  
          \* <b<sup>a</sup>-a> = *bā*                      \* <b<sup>i</sup>-i> = *bī*                      \* <b<sup>u</sup>-u> = *bū*  
          \* <b<sup>a</sup>-i> = *bai*  
          \* <b<sup>a</sup>-a-i> = *bāi*  
          \* <b<sup>a</sup>-u> = *bau*  
          \* <b<sup>a</sup>-a-u> = *bāu*

But this concept (which would have required a total of sixty-nine symbols) must have been abandoned at some point in favor of the attested system with its many ambiguities. As can be seen from the system's inconsistent structure (see [2]), the reorganization of the original system must have been regulated by extralinguistic (formal and stylistic) considerations – for example, the tendency to avoid complex signs with crossed wedges or with more than five elements. In any event, the principle of “Occam's razor” was not employed in devising the Old Persian spelling practices to the extent that many spellings are quite uneconomical (e.g., that of final *-i*, *-u*, etc.).

It is the history and genesis of the Bīsutūn monument itself which strongly suggests that the Old Persian script was introduced in connection with these texts. The Old Persian captions of the figures represented in the relief and likewise the Old Persian text of the major inscription do not belong to the original design of the monument, but were added only later to the Elamite and Babylonian versions. That the mother tongue of the kings had been at first neglected on this monument certainly suggests that the Old Persian language had not been previously set to writing.

## 2.4 Decipherment

Because Old Persian cuneiform fell into disuse with the fall of the Achaemenid Empire, and thus knowledge of that script and of the values of its individual characters was lost already in antiquity, this writing system had to be deciphered in the modern era. Old Persian texts first came to the attention of the West during the seventeenth century. A solid basis for the decipherment was laid by C. Niebuhr, who in 1778 published the first precise copies of Achaemenid trilingual texts and who recognized that the first and most simple system was written from left to right. Following the identification of the word-divider and the attribution of the texts to the Achaemenids, G. F. Grotefend, in 1802, began the process

of decipherment. By assuming that the inscriptions were records of the ancient Persians and might therefore contain the names, titles, and genealogies of some of their kings, he succeeded in determining the approximate phonetic values of about ten signs.

From this starting point, other scholars, progressing step by step, brought the decipherment to its conclusion. In 1826 R. Rask identified the  $n^{(a)}$  and  $m^{(a)}$  signs in the genitive plural ending *-ānām* (corresponding to Avestan-*anām*) and thus produced the first evidence for a close relationship with the Avestan language. In 1836 E. Burnouf and C. Lassen undertook a more systematic comparison with Avestan. Lassen, in 1845, made the very important discovery that the consonant characters of the Old Persian script could have an inherent vowel, as in the ancient Indian scripts. The work was completed in 1846/1847 by H. C. Rawlinson with his publication, translation, and interpretation of the entire DB text. A final touch was added in 1851 by J. Oppert, who established the value of the last (and most rarely used) of the phonetic signs,  $l^{(a)}$ , which even now is attested only in four foreign names for the marginal phoneme /l/ (not belonging to Old Persian proper).

### 3. PHONOLOGY

#### 3.1 Phonemic inventory

Identifying the complete system of Old Persian phonemes is a rather difficult task, since only a minimal set of phonemes is revealed by the attested graphemes. In order to advance beyond that set, the data must be analyzed and evaluated on a language-internal basis and by methods of historical-comparative linguistic analysis.

##### 3.1.1 Consonants

The following consonantal phonemes can be confidently identified for Old Persian:

(4)	<i>Bilabial</i>	<i>Labiodental</i>	<i>Interdental</i>	<i>Dental</i>	<i>Velar</i>
<i>Stop</i>					
<i>Voiceless</i>	p			t	k
<i>Voiced</i>	b			d	g
<i>Fricative</i>		f	θ		x
<i>Nasal</i>	m			n	

(the velar nasal [ŋ] is only a positional variant with allophonic status). In addition, Old Persian possesses two so-called “palatal” affricates *c* and *j*, which in all probability were palato-alveolar /č/ and /ǰ/. There also occur six fricatives – /s/, /z/, /ç/, /š/, /ž/, and /h/, the liquids /r/ and /l/, and the glides /y/ and /w/.

The actual pronunciation of those phonemes is not as secure as is suggested by the conventional representation. Thus, regarding the voiced stops /b, d, g/, it has been hypothesized that they were – at least in intervocalic position (if not more generally) – voiced fricatives [β, ð, ɣ]. The sibilant /ž/, which is not represented graphically by a separate character, but is written with the *j* sign, must be postulated for reasons of historical phonology: DB II 64 *n-i-j-a-y-m* = [niž-āyam] “I departed, went off” presents evidence for the Proto-Aryan verbal root \**ay* + prefix \**niš-/niž-* (with a *j* sign denoting the reflex not of Proto-Aryan \**j*, but of \**ž*, the voiced counterpart of \**š* in the position before a voiced sound). For the time being, however, the question of whether *ž* and *ǰ* are two distinct phonemes or only allophones of one and the same archiphoneme remains unresolved.

The fricative phoneme identified as the palatal /ç/ is the Old Persian reflex of the Proto-Iranian cluster \*ʋr (which is preserved in [nearly] all other Old Iranian dialects). Its phonetic realization remains unclear, however. It can be said with certainty only that the sound was pronounced as a voiceless sibilant (certainly not as a palato-alveolar sibilant [ʃ] and not as an affricate [č]); in Middle Persian its reflex has merged with that of Old Persian /s/.

Old Persian has a syllabic [ɾ], which is only a contextually conditioned allophone of the liquid /r/ (between stops), however, and not an independent phoneme. The lateral /l/ has a marginal position in the phonemic inventory of Old Persian, since it is attested only in four foreign names.

### 3.1.2 Vowels

Old Persian possesses three short and three long vowel phonemes, presented in Figure 28.1:

	FRONT	CENTRAL	BACK
HIGH	ī / i		ū / u
LOW		ā / a	

**Figure 28.1** Old Persian vowels

Whether the long vowels are somewhat lower than the short ones cannot be established. In addition, there are two “short” and two “long” diphthongs, which are not phonemes, but only biphonemic combinations of the short or long low-central vowel with a subsequent short high-front or back vowel; since the first is the syllable nucleus, those diphthongs result in

(5)	<u>Short diphthongs</u>	<u>Long diphthongs</u>
	ai	āi
	au	āu

Those four diphthongs, inherited from Proto-Iranian, are preserved in Old Persian as such at the time of the origin of the Old Persian cuneiform script and during the reign of Darius I and Xerxes I, as can be deduced from their regular orthographic representation (see §2). From a later period, there is evidence of a monophthongization of *ai* and *au* to *ē* and *ō* respectively – seen in the development from Old to Middle Persian and revealed by transcriptions of Persian words in other languages (the “collateral” tradition; see §6). The only transcription evidence of any linguistic weight for Old Persian proper is provided by the Elamite language, which has no diphthongs itself (see Ch. 3, §3.2). The Elamite script therefore lacks a regular means of spelling such sounds and so offers little possibility of documenting an early (pre-460 BC) monophthongization. Even so there are, in fact, unmistakable Elamite attempts to render Old Persian diphthongs: for example, *ti-ig-ra-ka-u-da* for Old Persian *tigra-xauda*– “with pointed caps.”

It should be noted that not every graphic sequence seemingly pointing to *ai* and *au* actually records a diphthong. Spellings like *a-i-š-t-t-a* “he stood” (from Proto-Iranian \**a-hišta*<sup>o</sup>), the theonym *a-u-r-m-z-d-a* (from Proto-Iranian \**Ahura Mazdā*) or the country name *h-r-u-v-t-i-š* (from Eastern Iranian \**Harahwatī*– “Arachosia”) record sequences of two syllables, [-a\$ɪ-] and [-a\$u-] (i.e., *A-uramazdā*, not *Au-ramazdā*, etc.).

### 3.2 Phonotaxis

Vowels and diphthongs are not subject to any phonotactic restrictions, and likewise all single consonants appear in initial and intervocalic position. For the final position, however, only single consonants (neither geminate consonants nor any other consonant clusters) are found, and only *-m*, *-r*, and *-š* are written. Those final consonants which are omitted in writing were perhaps still pronounced but in some manner phonetically reduced. Note that original Proto-Iranian *\*-a* is written as Old Persian <-a> (i.e., [-a:]), but original *\*-an* or *\*-ad* is written as <-C<sup>a</sup>> (i.e., [-a]).

Even if Old Persian shows a certain preference for open syllables (see §3.3; suggested also by historical developments like that of the Proto-Iranian clusters *\*Cy*, *\*Cw* to *Ciy*, *Cuw*), consonant clusters appear in great number, especially biconsonantal clusters, and particularly in word-internal position. More complex clusters with three (*xšn-*, *-xšn-*, *-xtr-*, *-ršn-*, *-nst-*) or even four elements (only non-native *-xštr-*) are rare. Because of the very limited corpus of Old Persian texts, only a small subset of all clusters possible is actually attested. The most commonly occurring of the attested clusters are (i) those of the form *Cr* and *rC*; (ii) those having an initial sibilant (*sk*, *st*, *zd*, *zb*, *zm*, *šk*, *št*, etc.); and (iii) those having an initial nasal (though not written; *nk*, *ng*, *nt*, *nd*, *mp*, *mb*, etc.).

### 3.3 Syllable structure

It is difficult to make specific observations about the syllable structure of Old Persian. Most syllables appear to be open: [ $\$(C)V$ ]; more rarely [ $\$C_1C_2V\$$ ] (e.g., *xša-ça-* “kingdom”) or even [ $\$C_1C_2C_3V\$$ ] (e.g., *xšnā-sā-ti* “he may know”). In the case of consonant clusters the syllable boundary may fall within the cluster or before it; the position of the boundary may depend on various criteria: the relative sonority of the particular elements of the cluster; the presence and position of a morpheme boundary; whether or not the cluster concerned is permissible in word-initial position; and so forth. Syllables also occur with the structure [ $\$VC\$$ ], [ $\$CVC\$$ ], and [ $\$C_1C_2VC\$$ ] (e.g., *u-fraš-ta-* “well punished”), and perhaps also those with two consonants following the syllabic nucleus (e.g., *θans-ta-nai* “to say”).

### 3.4 Accent

Accent is not marked in the Old Persian writing system; consequently both the nature and the position of the accent are quite uncertain. In the development from Old to Middle Persian, final syllables disappear, suggesting that the accent was fixed in the manner of Classical Latin or later Old Indo-Aryan. There may be (indirect) evidence for the hypothesis that the inherited free accent (perhaps a pitch or tonal accent), of which there are traces in Avestan and in modern Iranian languages (especially Pashto), survived until the reign of Darius I.

### 3.5 Diachronic developments

In this section, only the most interesting and significant diachronic phonological developments will be presented (and only vis-à-vis Proto-Iranian).

#### 3.5.1 Consonants

Among consonantal developments, the most distinctive concerns the Old Persian reflexes of the Proto-Iranian continuants (presumably affricates *\*tʰ* and *\*dʰ*), which are themselves

reflexes of the Proto-Indo-European palatals  $*\hat{k}$ ,  $*\hat{g}$ ,  $*\hat{g}^h$ : in contrast to the other Iranian languages Old Persian shows  $\vartheta$  in, for example, *við-* “house, royal house” = Avestan *vīs-* = Vedic *vis-* from Proto-Aryan  $*wíc-$ , and *d* (if not  $[\delta]$ ; see §3.1.1) both in, for example, *yad-* “to worship” = Avestan *yaz-* = Vedic *yaj-* from Proto-Aryan  $*yaǵ-$ , and in *adam* “I” = Avestan *azəm* = Vedic *ahám* from Proto-Aryan  $*aǵ^hám$ .

There are also certain distinctive Old Persian consonantal changes of a conditioned or syntagmatic type. These changes show an Old Persian development which has progressed beyond that seen in the other Old Iranian languages. Thus, the Proto-Iranian cluster  $*\vartheta r$  develops into Old Persian  $\varsigma$  in, for example, *puça-* “son” = Avestan *puθra-* = Vedic *putrá-*. That this change is of a rather late date is suggested by the fact that Proto-Persian  $*\vartheta r$ , where  $\vartheta$  is a reflex of Proto-Indo-European  $*\hat{k}$ , Proto-Iranian  $*t^s$ , has also undergone the change: thus, one finds Old Persian *ni-çāraya-* “to restore” = Avestan *ni-srāraia-* from Proto-Aryan  $*çrai-$  and Proto-Indo-European  $*\hat{k}lei-$ .

Before  $*n$  or  $*y$  Proto-Iranian  $*\vartheta$  became Old Persian  $\varsigma$ : for example, *a-r-š-n-i-* ([arašni-]) “cubit” from Proto-Iranian  $*aravni-$  = Vedic *aratní-*; *h-š-i-y-* ([hašiya-]) “true” = Avestan *ha’θiia-* from Proto-Iranian  $*havθya-$  = Vedic *satyá-*.

Old Persian *šiy* develops from Proto-Iranian  $*čy$  (i.e., from a Proto-Indo-European  $*k^w$  that was palatalized before  $*y$ ): for example, *š-i-y-a-t-i-* ([šiya:ti-]) “happiness” = Avestan *šā’iti-* from Proto-Aryan  $*čyāti-$  = Latin *quiēti-*, nominative *quiēs*.

A completely independent development of Old Persian, setting it apart from all the other Iranian languages (and thus one of its chief innovative characteristics), is the simplification of the Proto-Iranian clusters  $*t^sv$  and  $*d^zv$ , producing Old Persian *s* and *z* (not *sp* and *zb*): for example, *a-s-* ([asa-]) “horse” = Avestan *aspa-* = Vedic *ásva-*; *v<sup>i</sup>-i-s-* ([visa-]) “all” = Avestan *vīspa-* = Vedic *visva-*; *h-z-a-n-m* (acc. sg. [hiza:nam]) “tongue” (for the spelling *h-z-* see §2.2, 13), evolving from Proto-Iranian  $*hid^zvā^o$  as do Avestan *hizuuā-* or Parthian *‘zb n* ([izβa:n]) from earlier  $*hizbān^o$ .

### 3.5.2 Vowels

The vowels and diphthongs of Proto-Iranian remained unchanged in Old Persian at least until the period of Darius I and Xerxes I (on the later monophthongization of the short diphthongs see §3.1.2). The reflex of Proto-Iranian word-final short  $*-a$  is usually written as  $<-C^a-a> = -ā$ , as in *u-t-a* ([uta:]) “and” (Avestan *uta*, Vedic *utá*); it appears probable that this lengthening was a linguistic reality and not only a graphic phenomenon. Vowel contraction seems to play a minor role in Old Persian. The most obvious example is that of  $*-iya-$  producing  $-ī-$ , as in *n-i-š-a-d-y-m* ([ni:ša:dayam]) from uncontracted  $*ni-a-šādayam$  “I have put down” (cf. the alternative form *n-i-y-š-a-d-y-m*), and in *m-r-i-k-* ([mari:ka-]) “young man” from  $*mariyaka-$  (with a secondary  $-Ciya-$  from  $*-Cya-$ , from Proto-Aryan  $*maryaka-$  (= Vedic *maryaká-*).

Proto-Iranian sonorants,  $*m$ ,  $*n$ ,  $*y$ ,  $*w$ , and  $*r$  (including Proto-Iranian  $*ar$  from Proto-Aryan  $*ǵH$  as in *darga-* “long” = Old Avestan *dar<sup>2</sup>ga-* = Vedic *dīrghá-*, etc.), remain unchanged in Old Persian. Proto-Aryan  $*Cy$  and  $*Cw$  developed into Old Persian *Ciy* and *Cuw* respectively, regularly written as  $<C^{i/a}-i-y>$  and  $<C^{u/a}-u-v>$ : for example, *a-n-i-y-* ([aniya-]) “other” = Avestan *a’niia-* = Vedic *anyá-*; *h-r<sup>u</sup>-u-v-* ([haruva-]) “all” = Avestan *ha<sup>u</sup>ruua-* = Vedic *sárva-*.

Syllabic  $*r$  as an allophone of consonantal  $*r$  occurring between consonants (C\_\_C) and word-initially before a consonant (#\_\_C) likewise is preserved in Old Persian and probably

was pronounced as [ər]. Since in Old Persian orthography this [ər] can be rendered only in a makeshift fashion (like the sequence [ar]) by <(C)a-r-C>, other unambiguous evidence is required to confirm the value [ər] – either morphological (e.g., *k-r-t-* “made, done” = [kārta-] with the zero-grade of the root like Avestan *kər̥ta-* and Vedic *kṛtā-*), or etymological (e.g., *a-r-š-t-i-* “spear” = [əršti-], revealed by Vedic *ṛṣṭi-*). A special case is the development of Proto-Iranian \**r* to Old Persian *u* in the present and aorist stems of the root *kar* “to do” (e.g., *k<sup>u</sup>-u-n-u-t-i-y* [kunauti] “he does” = Avestan *kər̥naoiti* = Vedic *kṛṇōti*); these are usually explained as allegro forms originating in (and spreading from) the imperative.

Two phonetic phenomena, which have given such a strange appearance to many Avestan words (see Ch. 29, §§3.3; 3.4.2; 3.4.10), are without significance for Old Persian. Epenthesis (i.e., the insertion of *i* or *u* into an existing syllable) is completely foreign to Old Persian, and anaptyxis (i.e., the development of a vowel between two consonants) is nearly unknown. The Avestan epenthesis, which is triggered by an ensuing *i/y* or *u/w* (as in Avestan *ha<sup>i</sup>θ iia-* “true” from \**haθya-*, see §3.5.1), is not attested in Old Persian inscriptions (transcription of Old Persian words in other languages may reveal that a late process of this sort characterized colloquial Old Persian). Anaptyxis is found only in the case of the clusters *dr* and *gd* when followed by *u*: for example, one finds *d<sup>u</sup>-u-r<sup>u</sup>-u-v-* ([duruva-]) “firm” = Avestan *druua-* ([druwa-]) = Vedic *dhruvā-*; present tense stem *d<sup>u</sup>-u-r<sup>u</sup>-u-j<sup>i</sup>-i-y-* ([durujiya-]) “to lie” = Vedic *drúhya-*; *s-u-g<sup>u</sup>-u-d-* ([Suguda-]), as well as *s-u-g-d-* ([Sugda-]), “Sogdiana.”

## 4. MORPHOLOGY

### 4.1 Morphological type

Typical of ancient Indo-European, Old Persian is an inflectional language with synthetic morphological patterns. Owing to lack of evidence, both the nominal and pronominal and, still more, the verbal paradigms are known only partially in most instances. Therefore it is not possible to give a fully formed account of the formation, function, and actual use of nominal, pronominal, and verbal forms. The same is true, by and large, with regard to nominal and verbal stem formation.

### 4.2 Nominal morphology

The grammatical categories marked on the Old Persian noun are case (seven), gender (three), and number (three). Whereas the three genders (masculine, feminine, and neuter) and the three numbers (singular, dual, and plural) inherited from Proto-Indo-European have preserved their usual significance and function, the case system has been reduced by one in Old Persian. Likewise gender and number show the expected and customary grammatical agreement (see §5.6), though there are some instances in which two singular subjects occur not (as would be expected) with a dual, but with a plural form of the verb.

The seven attested nominal cases are the following: (i) nominative (for subject); (ii) vocative (for direct address); (iii) accusative (for direct object and direction); (iv) genitive (used as possessive, subjective, objective, and partitive genitive); (v) locative (for indication of place or goal); (vi) instrumental (for indication of means, cause, and extension); and (vii) ablative (only combined with prepositions). The functions of the Proto-Indo-European dative (as the case of the indirect object) have been absorbed by the Old Persian genitive (e.g., *haya siyātim adā martiyahyā* “who created happiness for man”). Moreover, the case



system has also been reduced and simplified by abandoning formal distinctions; thus, for example, there are only three separate forms in the singular of the *ā*-stems: nom., voc. *-ā*; acc. *-ām*; gen.(*-dat.*), abl., loc., instr. *-āyā*.

#### 4.2.1 Stem formation

Old Persian has inherited from Proto-Indo-European its two chief means of nominal stem formation: (i) *derivation* (by means of primary or secondary suffixes attached to the underlying [verbal] root itself or to an already derived nominal stem), and (ii) *composition* of two word stems (with or without a particular [compositional] suffix). Also playing a role in stem formation are *ablaut* (see Ch. 17, §3.2) and, for derivation, the vowel-lengthening process known as *vrddhi*. Only some subset of the numerous inherited nominal suffixes of Old Persian can be treated here, since the scanty evidence available does not allow one to judge whether some particular formation is only a traditional relic within Old Persian or actually remains a living and productive process.

One of the productive suffixes is undoubtedly the “locative” suffix *-iya-*, forming adjectives, especially ethnics such as *Armin-iya-* “Armenian” (from *Armina-*), *Ūj-iya-* “Elamite” (from *Ūja-*), *Mac-iya-* “inhabitant of Makrān” (from *Maka-*), and so forth. The Proto-Iranian suffix *\*-hwa-/ \*-šwa-* forming fractions (see §4.6) seems to be similarly productive.

A distinctive phenomenon of derivation which Old Persian has inherited and which, as several indisputable examples show, is still productive in this language, is the lengthening of the first vowel of a word, a process traditionally called *vrddhi* (a term coined by the ancient Indian grammarians). The clearest examples attested are the ethnic *Mārgava-* “inhabitant of Margiana,” derived from *Margu-* “Marv, Margiana”; and the month name *Bāgayādi-*, based on *\*baga-yāda-* “worship of the gods.” Other apparent cases are not without problems: for example, the month name *Θāigraci-*; a form which – could *vrddhi* be confirmed – would be essential for settling the question of whether Old Persian derivatives of words with *i* or *u* vowels have the *vrddhi* form *āi* and *āu* like Old Indo-Aryan or the short diphthong *ai*, and *au*, as it is found in Avestan.

#### 4.2.2 Nominal declension

Old Persian nouns have been traditionally grouped into declensional classes, though with regard to the origin of the nominal system at an earlier stage of the Indo-European parent language, a number of other criteria are of relevance, chiefly accent placement and ablaut variation and their distribution over the root, the (optional) suffix, and the ending (see Ch. 24, §4.1.1.3). Old Persian evidence is available for stems ending in *-a-*, *-ā-*, *-i-*, *-ī-*, *-ī/yā-*, *-u-*, *-ū-*, *-h-* or *-š-*, *-r-*, *-n-* and in several stops and fricatives. The only productive stems, however, are those ending in vowels, and in particular those of the *a*-class, as those lexemes suggest which show forms of different declensions side by side: most clearly *tunuvant-* “strong” (in nom. sg. *tunuvā*) versus *tunuvanta-* (in gen. sg. *tunuvantahyā*); compare the “bridge” accusative singular *tunuvantam*.

The only paradigms which are known somewhat extensively are those of the stems in *a-* and *ā-*; their singular and plural forms may be given in (6) and (7) (for the dual see below); all other case forms and declensional patterns are presented only in the larger summary of (8) and (9):

(6) The Old Persian *a*-stems

	Singular		Plural	
	Example	Ending	Example	Ending
<i>Animate</i>				
<i>Nom.</i>	martiya “man”	-∅ < *-s	martiyā bagāha “god”	-ā < *-ās -āha < *-āsas
<i>Voc.</i>	martiyā	*-∅	—	
<i>Acc.</i>	martiyam	-m	martiyā	-ā < *-āns
<i>Gen.</i>	martiyahyā	-hyā	martiyānām	-ānām
<i>Abl.</i>	Pārsā	-ā < *-āt	Sakaibiš	=instr.
<i>Instr.</i>	kārā “army”	-ā	martiyaibiš	-aibiš
<i>Loc.</i>	Pārsai dastay-ā “hand”	-i -i + -ā	Mādaišuvā	-aišu + -ā
<i>Neuter</i>				
<i>Nom.-acc.</i>	xšačam “kingdom”	-m	āyadanā “place of worship”	-ā < *-ā

(7) The Old Persian *ā*-stems

	Singular		Plural	
	Example	Ending	Example	Ending
<i>Animate</i>				
<i>Nom.</i>	taumā “family”	-∅	stūnā “column”	-ā < *-ās
<i>Voc.</i>	—		—	
<i>Acc.</i>	taumām	-m	[hamiçi]yā “rebellious”	-ā < *-āns
<i>Gen.</i>	taumāyā	-yā < *-yās	°zanānām “with . . . races”	-ānām
<i>Abl.</i>	Same as genitive		—	
<i>Instr.</i>	framānāyā “order”	-yā	—	
<i>Loc.</i>	Aθurāyā	-i + ā	maškāuvā “skin”	-u < *-su + -ā

The set of case endings attested in Old Persian may be summarized in (8) and (9) without differentiating them by declensional class and without a detailed historical-comparative interpretation:

## (8) Summary of Old Persian singular case endings

<i>Animate</i>	
<i>Nom.</i>	-∅, -š from *-s; -∅ from *-∅
<i>Voc.</i>	-∅ from *-∅
<i>Acc.</i>	-m, -am from *-m, -m̥
<i>Gen.</i>	-a from *-as; -∅, -š from *-s; -hyā from *-sya; -yā from *-yās
<i>Abl.</i>	-ā from *-āt; -∅ from *-t; or identical to the genitive
<i>Instr.</i>	-ā from *-ā; -yā from *-yā
<i>Loc.</i>	-i from *-i; -∅ from *-∅, both with or without postpositive -ā
<i>Neuter</i>	
<i>Nom.-acc.</i>	-m from *-m; -∅ from *-∅

## (9) Summary of Old Persian plural case endings

*Animate*

<i>Nom.</i>	-a from *-as; -ā from *-ās; -āha from *-āsas
<i>Voc.</i>	Identical to the nominative, but not attested
<i>Acc.</i>	-ā from *-āns; -ø, -š from *-ns
<i>Gen.</i>	-ānām, -ūnām from *-Vnām
<i>Abl.</i>	Identical to the instrumental
<i>Instr.</i>	-biš, -aibiš from *-biš
<i>Loc.</i>	-aišuvā, -šuvā from *-šw-ā; -uvā from *-sw-ā, attested only with postpositive -ā

*Neuter*

<i>Nom.-acc.</i>	-ā from *-ā
------------------	-------------

Several dual forms are securely attested in Old Persian texts, such as nom. *u-b-a* ([uba:]) “both”; acc. *g-u-š-a* ([gauša:]) “both ears”; gen. *g-u-š-a-y-a* ([gauša:ya:]); instr. *d-s-t-i-b-i-y-a* ([dastaibiya:]) “with both hands,” all belonging to stems in *-a-*. In addition, the following occur: nom. *u-š-i-y* ([uši:]), as well as *u-š-i-y-a* ([ušiya:]), three times each, and instr. *u-š-i-b-i-y-a* ([uši:biya:]), from neuter *uši-* “intelligence” (literally “ear” and therefore in dual number).

Adjectives behave like the nouns with regard to stem formation and declension. The comparative is formed by means of the Proto-Indo-European suffix *\*-yes-/ -yos-* and the superlative by *\*-is-to-*. As examples, consider Old Persian nom. masc. sg. *t-u-v<sup>i</sup>-i-y-a* ([taviya:]), from *\*tau-yah-* “stronger,” and *m-θ-i-š-t* ([maθišta]) “greatest.”

## 4.3 Pronominal morphology

A variety of pronouns is attested in Old Persian: (i) personal pronouns (including the so-called anaphoric pronoun); (ii) several demonstrative pronouns; (iii) relative; and (iv) interrogative-indefinite pronouns.

## 4.3.1 Personal pronouns

The personal pronouns are characterized (i) by an absence of grammatical gender; (ii) by a remarkable heteroclisis between the nominative and oblique cases; and (iii) by the existence of frequently used enclitic forms. All these characteristics have Proto-Indo-European ancestry. The following personal pronouns are attested in Old Persian:

(10)

	<i>Accented forms</i>		
	<i>First</i>	<i>Second</i>	<i>First Plural</i>
<i>Nominative</i>	adam	tuvam	vayam
<i>Accusative</i>	mām	θuvām	—
<i>Genitive</i>	manā	—	amāxam
<i>Ablative</i>	-ma	—	—
	<i>Enclitic forms</i>		
<i>Accusative</i>	-mā	—	—
<i>Genitive</i>	-mai	-tai	—

The dual forms are not attested at all; the genitive has taken over the function of the dative. Ablative *-ma*, though being attested only in combination with the preposition “by,” *h-c-a-m* ([*hacā-ma*]) “by me,” is not enclitic (demonstrated by accented Vedic *mát*).

The anaphoric pronouns “he, she, it” share the characteristic features of the personal pronouns, though there are no nominative forms and no heteroclisis. Old Persian exhibits enclitic forms built from the stems *-ša-/-ši-* and *-di-*: acc. sg. *-šim* “him,” gen. *-šai* “his,” acc. pl. *-šiš* “them,” gen. *-šām* “their”; acc. sg. *-dim* “him” and acc. pl. *-diš* “them.”

#### 4.3.2 Demonstrative pronouns

Other pronominal stems exhibit grammatical gender distinctions and, in part, are characterized by a declension differing from that of nominal stems in *-a-* and *-ā-*. Included in this group are three demonstrative pronouns. The pronoun *iyam* (nom. sg. masc./fem.) “this” combines forms of the stems *i-*, *ima-*, and *a-*: for example, *ima* (nom.-acc. sg. neut.), *anā* (instr. sg. masc.), *ahyāyā* (loc. sg. fem.). The remaining two are *aita-* “this here” (more emphatic), and *hau-* (nom. sg. masc./fem.) “that”; the paradigm of the latter is supplemented in the oblique cases by the stem *ava-*: for example, *ava* (nom.-acc. sg. neut.), *avai* (nom.-acc. pl. masc.), *avaišām* (gen. pl. masc.), *av[ā]* (nom. dual masc.).

#### 4.3.3 Relative and interrogative pronouns

The relative pronoun, which has also acquired the function of an article (see §5.5), is an Old Persian innovation. Its stems *haya-* (nom. sg. masc./fem.) and *taya-* (elsewhere) “who, which” emerged from the fusion of the Proto-Aryan correlating demonstrative and relative pronouns *\*sá-/tá- + \*yá-* “the one, who.” The interrogative pronoun is not attested in Old Persian texts and can be recovered only from the indefinite pronouns *kaš-ci* (nom. sg. masc.) “somebody,” *ciš-ci* (neut.) “something,” which are derived by means of the generalizing particle *-ci*, as in *ya-ci* (nom.-acc. sg. neut.) “whatever.”

#### 4.3.4 Pronominal adjectives

The declension of certain adjectives, which are semantically close to the pronouns, shares also the special declensional forms of pronouns. Old Persian attests only *aniya-* “other” (e.g., nom.-acc. sg. neut. *aniya*, abl. sg. masc. *aniyanā*); *haruva-* “all” (e.g., loc. sg. fem. *haruvahyāyā*); and *hama-* “the same” (in gen. sg. fem. *hamahyāyā*).

### 4.4 Verbal morphology

The grammatical categories of the Old Persian verbal system were inherited from Proto-Aryan, the consequent and consistent structure of which can still plainly be observed in the earliest Vedic texts. But with regard to both function and form, a great number of fundamental innovations and reorganizations have occurred which leave the distinct impression that Old Persian, like Young Avestan (see Ch. 29, §1), has begun to part company with the Proto-Aryan system and already represents a kind of transitional stage from Old to Middle Iranian. This is revealed by phonetic developments and innovations in nominal morphology, but especially by changes in the system of verbal morphology: (i) the aspectual opposition of aorist versus imperfect has been lost; (ii) aorist and perfect tense forms are attested only rarely; (iii) a periphrastic “neo-perfect” has emerged (see §4.4.6); and (iv) present stems in *-aya-* begin to gain prominence.

Old Persian verbal forms are marked for tense (originally aspect), voice, mood, and the usual three persons and three numbers. The Old Persian evidence is, however, rather

unbalanced, owing to the nature of the contents of the inscriptions: thus, for example, the only dual form found in the texts is the third plural imperfect active *ajīvatam* “they both (still) lived.” Together with the three persons and numbers, two of the three voices (i.e., active and middle) find expression in two sets of personal endings: the so-called *primary* endings in the present indicative (which alone denotes a real present time) and subjunctive (which may do the same, at least in the speaker’s view), and the *secondary* endings otherwise, apart from the imperative, which has distinctive endings.

#### 4.4.1 Voice

The voices usually have their customary functions (inherited from the Indo-European parent language). A particularly striking exception is provided by certain third plural middle forms which lack middle function and are to be interpreted as having arisen only to avoid ambiguity. Passive morphology is more innovative, with the following attested: (i) forms built from the passive stem in *-ya-* (e.g., imperfect *a-ṽanh-ya* “it has been said”), common to Indo-Iranian for the present stem; (ii) middle forms like *a-naya-tā* “he was led”; and (iii) phrases consisting of a verbal adjective in *-ta-* plus the copula (which usually is omitted, however, in the third person: see §4.4.6).

#### 4.4.2 Mood

The five moods attested in Old Persian are indicative, subjunctive, optative, imperative, and, as an Indo-European relic, injunctive (see below). Typical of Iranian is both the use of the perfect optative for the irrealis of the past, and (even more so) the use of the present optative with the temporal augment *a-* (thus looking like an imperfect optative) to express a repeated action of the past (e.g., *avājaniyā* from \**ava-a-jan-yā-t* “he used to slay”).

The Old Persian moods exhibit the same functions as their counterparts in Young Avestan. The *indicative* is used to express factual statements – present indicative (formed with the primary endings) for those in present time, and imperfect indicative (the augment *a-* and secondary endings being added to the present stem) for those in past time. The *subjunctive* expresses the eventual or potential realization of actions in the present or future; the present subjunctive is formed with primary endings, which are added to the present stem enlarged by *-a-* (e.g., *ah-a-ti* “it may be”). The *optative* is used for wishes and prayers and is formed with a stem in *-iyā-* (in the athematic singular) or *-ī-* (otherwise) – suffixes descended from Proto-Indo-European \**-yeh<sub>1</sub>-/\*-ih<sub>1</sub>-*; the optative takes secondary endings (e.g., 2nd sg. mid. *yadaišā* “you may worship”). The *imperative* is the mood of command and prayer and makes use of distinctive imperative endings which are added to the present or aorist stem.

The *injunctive* (with secondary endings) is found in Old Persian only in prohibitive constructions introduced by the particle *mā* “not!” but even in preventive clauses never combined with forms of the aorist tense stem. Together with the loss of the aorist (see §4.4.3) Old Persian obviously has lost the inherited distinction between the inhibitive present injunctive and the preventive aorist injunctive. Moreover, if combined with the optative present, the prohibitive particle *mā* denotes a corrective notion with regard to a present action: for example, *daivā mā yadiyaiša* “the Daivas shall not be worshiped any longer!”

#### 4.4.3 Tense

The tenses find expression in stem formations which had originally been used to distinguish aspect (imperfective vs. perfective) and still did so in Proto-Aryan and Proto-Iranian. Several doublets of such forms make it clear, however, that the imperfect (which is built on the present stem and thus expressed the imperfective aspect of a past action) and the aorist (being the counterpart in the perfective aspect) are used in Old Persian without any obvious difference

in function, suggesting that aspectual distinctions were no longer being productively made. The “sigmatic” aorist *adarši* “I took possession of” (1st sg. indic. aor. middle of the root *dar-*) alone seems to point to a living use of the aorist indicative (i.e., for conveying the perfective aspect of an action). The one perfect form attested is an optative expressing past irrealis, *caxriyā* “he might have done.” Regarding perfect morphology, therefore, all that can be said is that Old Persian inherited stem reduplication (*ca-xr-* from Proto-Aryan \**ča-kr-* and Proto-Indo-European \**kʷe-kʷr-*), but nothing can be discerned about the particular endings of the perfect indicative active.

#### 4.4.4 Verbal stems

The stem formations occurring in Old Persian are essentially those inherited from Proto-Aryan and in the end often from Proto-Indo-European. This includes the inherited distinction between the *thematic* and the *athematic* stems marked by the presence or absence of the thematic vowel *-a-* (from Proto-Indo-European \**-e/o-*; see Ch. 17, §3.4) preceding the personal endings (e.g., athematic *as-ti* “he is,” but thematic *bav-a-ti* “he becomes”). The present and aorist stems (and likewise the only perfect stem attested; see §4.4.3) are formed either from the verbal root to which one of a set of suffixes is attached, or from the unsuffixed root itself (root presents and root aorists). Most numerous and to a certain degree productive are the present stems in *-aya-* like *tāvaya-* “to be able,” *mānaya-* “to wait, expect,” and so forth. Ancestral formations of Proto-Indo-European origin are the stems in *-sa-* (= Avestan *-sa-*) like *pṛsa-* “to ask, interrogate” (= Avestan *pṛśsa-*), *tṛsa-* “to be afraid” (= Avestan *tərśsa-*), *xšnāsa-* “to know.”

#### 4.4.5 Verbal endings

The various sets of verbal endings are only partially attested in Old Persian; these are presented in (11)–(16) together with their Proto-Aryan preforms:

##### (11) The Old Persian primary endings: active

	<i>Singular</i>
<i>First</i>	-mi from * <i>-mi</i> (also in the thematic verbs); -ni from * <i>-ni</i> (subjunctive)
<i>Second</i>	-hi from * <i>-si</i> (attested only in subjunctive)
<i>Third</i>	-ti from * <i>-ti</i>
	<i>Plural</i>
<i>First</i>	-mahi from * <i>-masi</i>
<i>Second</i>	—
<i>Third</i>	-nti from * <i>-nti</i>

##### (12) The Old Persian primary endings: middle

	<i>Singular</i>
<i>First</i>	-ai from * <i>-ai</i> ; -nai from Proto-Iranian * <i>-nai</i> (subjunctive)
<i>Second</i>	-hai from * <i>-sai</i>
<i>Third</i>	-tai from * <i>-tai</i>
	<i>Plural</i>
	Not attested

## (13) The Old Persian secondary endings: active

<i>Singular</i>	
<i>First</i>	-m from *-m; -am (athematic) from Proto-Aryan *-am replacing Proto-Indo-European *-m̥
<i>Second</i>	-ø from *-s
<i>Third</i>	-ø from *-t; -š after ai, au (in imperfect and optative forms like <i>akunauš</i> “he did” = Avestan <i>akər<sup>o</sup>naot</i> )
<i>Dual</i>	
<i>Third</i>	-tam = Avestan -təm (see §4.4)
<i>Plural</i>	
<i>First</i>	-mā from *-ma
<i>Second</i>	—
<i>Third</i>	-ø from *-nt; -h after a and -š after ai (in imperfect and optative forms like <i>abaraha</i> “they brought” or <i>yadiyaiša</i> “they shall not be worshiped”) from *-s

## (14) The Old Persian secondary endings: middle

<i>Singular</i>	
<i>First</i>	-i from *-i
<i>Second</i>	-šā from *-sa
<i>Third</i>	-tā from *-ta
<i>Plural</i>	
<i>First</i>	—
<i>Second</i>	—
<i>Third</i>	-ntā from *-nta

## (15) The Old Persian imperative endings: active

<i>Singular</i>	
<i>Second</i>	-ā from *-a (thematic) and -di from *-d <sup>h</sup> i (athematic)
<i>Third</i>	-tu from *-tu
<i>Plural</i>	
<i>Second</i>	-tā from *-ta
<i>Third</i>	-ntu from *-ntu

## (16) The Old Persian imperative endings: middle

<i>Singular</i>	
<i>Second</i>	-uvā and -šuvā from *-swa
<i>Third</i>	-tām from *-tām
<i>Plural</i>	
Not attested	

## 4.4.6 Nonfinite verbal forms

Old Persian exhibits only one type of infinitive: a construction with the formant *-t-n-i-y* ([*-tanai*] or [*-tani*]?), being an oblique case, dative (or locative) singular, of an action noun in *-tan-*, and built on the full-grade verb root: for example, *cartanai* “to do”; *bartanai* “to bear;”



*ʔanstanai* “to say.” In the case of *kantanai* “to dig” and *nipaištanai* “to engrave, write,” the passive interpretation “to be dug,” “to be engraved” cannot be ruled out.

The only reliably attested active participles are *tunuvant-* “strong” (literally “being able”; nom. sg. masc. *tunuvā*, from \**-wānt-s*) and *yaudant-* “being in turmoil” (only acc. sg. fem. *y-u-d-[t-i]-m* ([*yaudant-i(:)m*])). Present middle participles are formed by means of the suffix *-mna-* = Avestan *-mna-*, as in *xšaya-mna-* “being in control of.”

The commonly occurring verbal adjective or perfect passive participle in *-ta-* is inherited from the Proto-Indo-European formation in \**-to-*, which usually is added to the zero-grade verbal root: for example, *kṛta-* “done, made”; *jata-* “slain”; *pāta-* “protected”; but also *basta-* “bound” like Young Avestan *basta-* (in contrast to Vedic *baddhá-*) and the like. In addition, there are also some formations in *-ata-* (like *ʔak-ata-* “passed” or *han-gm-ata-* “assembled”; cf. Avestan *gmata-*) which go back to Proto-Indo-European \**-eto-*.

The verbal adjective in *-ta-* is used in Old Persian particularly for creating the new periphrastic perfect of the type *manā kṛtam* “(it was) done by me” (cf. Middle Persian *man kard*) replacing the inherited Proto-Aryan active perfect for expressing an accomplished action and/or a situation achieved by it. In origin this “neo-perfect” was formed by combining the copula “to be” with the *-ta-* adjective, though the third singular *asti* “she/he/it is” normally has been deleted. Moreover, the agent of transitive verbs is expressed in the genitive case (though the sense of the construction is not a possessive). Examples include the following: *ima, taya manā kṛtam* “this [is], what [has been] done by me”; *taya Brdiya avajata* “that Smerdis [had been] slain”; *yadi kāra Pārša pāta ahati* “if the Persian people shall be protected.”

## 4.5 Compounds

In principle, Old Persian exhibits all the types of compounds known from the other ancient Aryan languages (see Ch. 26, §4.4.2) and inherited from Proto-Indo-European (see Ch. 17, §3.5.1). Compounds contain two elements, the last of which is inflected. Attested are *determinative* and *possessive* compounds (including those which have an inseparable prefix like *a(n)-* “without, un-”; *u-* “well-”; or *duš-* “mis-, dis-” as first element), but no *copulative* compounds are attested as yet. Especially remarkable are the compounds having a verbal stem as the first element; Old Persian exhibits a number of such formations in anthroponomastics: for example, the throne names of Darius and Xerxes, *Dāraya-vauš* “holding the good” and *Xšaya-ṛšan-* “having command of heroes.” These forms reveal that Old Persian does not share in the Aryan recasting of the first element as a participial form in *-at-*, as one finds in Avestan and Old Indo-Aryan (cf. Avestan *Dāraiaṭ.raṭa-* “holding the chariot,” *xšaiiaṭ.vac-* “having (a good) command of speech”; Vedic *dhārayāt-kṣiti-* “sustaining the creatures,” *kṣayād-vīra-* “having command of heroes”).

## 4.6 Numerals

Since the cardinals are normally indicated by numeral signs and not written phonetically, hardly anything can be said about them. The number 1 is *aiwa-*, which like Avestan *aēuua-* goes back to Proto-Indo-European \**oi-wo-* “one, alone” (= Greek *oī(w)os* (οἷ(F)ος)). One hundred must have been \**ʔata-* (= Avestan *satəm* = Vedic *śatām*) and in all probability is attested in the name of the province Sattagydia, *ʔata-gu-*. Other cardinals are reflected in the “collateral” linguistic traditions (see §6), especially in Elamite garb, in compounded titles like \**daṭa-pati-* (Elamite *da-sa-bat-ti-iš*) “chief of ten, decurion” or \**ʔata-pati-* (Elamite *sa-ad-da-bat-ti-iš*) “chief of hundred, centurion.”

Of the ordinals there are attested in the Old Persian inscriptions: *fratama*- “first” = Avestan *fratəma*-; *duvitiya*- “second” = Old Avestan *d<sup>ai</sup>bitiia*-, Young Avestan *bitiia*- (= Vedic *dvitīya*-); *çitiya*- “third” = Avestan *θritiia*-; *navama*- “ninth” = Avestan *naoma*- (from *\*nawəma*-).

A quite interesting Iranian innovation is found in the fractions formed by addition of the Proto-Iranian suffix *\*-swa*- (realized as Avestan *-huua*- or *-šuua*-). The Old Persian reflexes are attested in Elamite renderings only and can be reconstructed as *\*çišuva*- “one-third” (Elamite *ši-iš-maš*; cf. Avestan *θrišuua*-); *\*çaçušuva*- and (with haplology) *\*çaçuva*- “one-quarter” (Elamite *za-aš-maš*, *za-iš-šu-maš*, *za-iš-šu-iš-maš*; cf. Avestan *caθrušuua*-); *\*pancauva*- “one-fifth” (Elamite *pan-su-ma-iš*; cf. Avestan *pañtaghuua*-); *\*aštauva*- “one-eighth” (Elamite *aš-du-maš*; cf. Avestan *aštahuua*-); *\*navauva*- “one-ninth” (Elamite *nu-ma-u-maš*); *\*daθauva*- “one-tenth” (Elamite *da-sa-maš*) and *\*vīstauva*- “one-twentieth” (Elamite *mi-iš-du-ma-kaš*, with an additional *ka*-suffix).

## 5. SYNTAX

### 5.1 Word order

The word order found in the Old Persian inscriptions is on the whole rather free, as is common among the ancient Indo-Iranian languages. The “unmarked” order, however, is Subject–Object–Verb (SOV):

- (17) Auramazdā-mai upastām abara  
 Auramazdā-me aid he brought  
 “Auramazdā brought me aid”

For enclitic *-mai*, see §5.3. Other complements, especially those indicating place, may follow the verb. There are attested, however, a number of cases showing varying order of the sentence constituents: for example, (i) of copula and predicate noun (cf. DNb 42f. *θanuvaniya uθanuvaniya ami* “as a bowman I am a good bowman” vs. DNb 44 *ṛštika ami uvṛštika* “as a spearman I am a good spearman”); or (ii) of two coordinated constituents (DB IV 72f. *yadi imām dipim vaināhi imaivā patikarā* “if you shall look at this inscription or these sculptures” vs. DB IV 77 *yadi imām dipim imaivā patikarā vaināhi*).

Nevertheless some peculiarities of word order must be noted, mainly “marked” sentence-initial or sentence-final position of words for reasons of emphasis. Here belong, for example, the initial position of the object (OSV) when expressed by a deictic pronoun

- (18) ima hadiš adam akunavam  
 this palace I I have built  
 “I have built this palace”

or the nonfinal (medial) position of verbs expressing an urgent plea. Notable is also the uncommon initial position of the verb in the formulaic expression *θāti NN xšāyaθiya* “proclaims NN, the king.”

When two or more coordinated elements form the subject or the object of a sentence, only the first element is placed before the verb, and the remaining elements follow, for example:

- (19) mām Auramazdā pātu utamai xšačam  
 me Auramazdā may he protect and my kingdom  
 “May Auramazdā protect me and my kingdom!”

Within phrases the word order is more fixed. A noun or pronoun (in the genitive case) which is dependent upon a noun precedes that noun: for example, *Kurauš puça* “son of Cyrus”; *manā pitā* “my father.” Exceptions which are attested in royal titles (cf. *xšāyaθiya xšāyaθiyānām* “king of kings” in contrast to Middle Persian *šāhān šāh*) or religious formulae (*vašnā Auramazdāha* “by the favor of Auramazdā”) are caused by foreign influence.

## 5.2 Topicalization

A striking feature of Old Persian syntax and stylistics is the frequent use of a sentence-initial (so-called) *casus pendens* (usually an absolute nominative), which is resumed by a demonstrative pronoun (20A) or adverb (20B):

- (20) A. Vištāspa    manā    pitā,    hau    Parθavai    āha  
          Hystaspes    my    father    that one    in Parthia    he was  
          “Hystaspes my father, he was in Parthia”  
       B. Pṛga    nāma    kaufa,    avadā...  
          Pṛga    by name    mountain    there  
          “There is a mountain, Pṛga by name, there ...”

This phenomenon is often combined with another stylistic peculiarity found in the Old Persian inscriptions, the origin of which must be sought, as Vedic parallels in prose texts show convincingly, in colloquial Proto-Aryan and not, as has been previously presumed, in Aramaic influence. This concerns parenthetical (more exactly, prosthothetical) constructions taking the form of nominal (i.e., verbless) clauses which introduce less common personal or geographical names: for example, *Dādr̥šiš nāma Arminiya, manā bandaka, avam...* “[There is] an Armenian, Dādr̥ši by name, my vassal, him...”

It should be noted that nominal sentences are very frequently used in Old Persian, mainly because the third singular form of the copula is normally omitted; consider DB I 27:

- (21) ima,    taya    manā    kṛtam  
          this    what    by me    done  
          “This [is], what [has been] done by me”

with relevant examples in both the main and relative clauses.

## 5.3 Clitics

Old Persian attests a number of enclitics (atonic lexemes which in Old Persian form a graphic unity with the preceding word); chiefly the following: (i) the oblique cases of the personal pronouns (including the anaphoric pronoun); (ii) the copulative and disjunctive conjunctions (*-cā* “and,” *-vā* “or”); and (iii) various emphatic particles. According to *Wackernagel’s Law* the enclitics are attached to the first accented word of the sentence or clause in Old Persian, as in Proto-Aryan and, still earlier, in Proto-Indo-European. This becomes particularly clear from examples like (17), *Auramazdā-mai upastām abara* “Auramazdā brought me aid,” when contrasted with

- (22) pasāva-mai    Auramazdā    upastām    abara  
          afterwards-me    Auramazdā    aid    he brought  
          “Afterwards Auramazdā brought me aid”

Enclitics which are construed with single words only and not with an entire sentence do

not follow Wackernagel's Law, but are attached to that particular word: for example, *yaθā paruvam-ci* "just as [it was] previously." For a special treatment of enclisis see Schmitt 1995.

## 5.4 Coordination and subordination

In the Old Persian inscriptions both coordination and subordination are used for expressing complex statements. It is not uncommon to find short simple sentences following one another, either accompanied by a connector (a coordinating conjunction like *utā* "and" or a temporal adverb like *pasāva* "afterwards, then"), or without such (asyndeton). In other cases (and, in part, in closely parallel passages), subordinate clauses occur introduced by a relative pronoun or by some appropriate conjunction. Most conjunctions used in Old Persian are derived from the (original) stem of the relative pronoun (as is the case in the cognate languages, too): for example, *yaθā* (often correlated with *avaθā* "thus") "when, after, so that" (introducing temporal, modal, and consecutive clauses); *yadi* "if" (normally with a subjunctive verb), "when" (with an indicative; introducing temporal and conditional clauses). While both of these are inherited, *yātā* "until, when, as long as" is a new formation, as is *taya* "that, so that" (acc. sg. neut. of the relative pronoun) which introduces causal, explicative clauses, indirectly reported speech, and so forth. Relative clauses are commonly attested, positioned both before and after the main clause.

There are also some passages that show a subordinate infinitive. Typical is that construction after a main clause containing verbs like "to order," "to be able," "to dare" (e.g., *adam nīštāyam imām dipim nipaīštanai* "I ordered to engrave this inscription"); another likewise typical use of an infinitive construction is that expressing purpose after verbs like "to go," "to send" (e.g., *paraitā patiš Dādṛšim hamaranam cartanai* "went forth against Dādṛši to fight a battle").

## 5.5 Relative constructions

The relative pronoun *haya-/taya-* functions as a definite article in expressions indicating various attributive complements to nouns, with case attraction if appropriate; for example:

- (23) A. Gaumāta haya maguš (nominative)  
Gaumātam tayam magum (accusative)  
"Gaumāta the magus"  
B. kāram tayam Mādam (accusative)  
"The Median army"  
C. viθam tayām amāxam (genitive plural)  
"Our [royal] house"  
D. xšačam taya Bābirau (locative)  
"The kingship in Babylonia"

Those constructions have similar counterparts in Avestan, but have spread considerably in Middle Persian and are ultimately the source of the Modern Persian *izāfat* construction.

## 5.6 Agreement

Grammatical agreement in Old Persian is of the sort common to the older Indo-European languages: (i) appositive and attributive adjectives and nouns agree in gender, number, and case; (ii) predicate nouns and adjectives agree at least in case, but now and then there are particular conditions for gender and number; (iii) relative, resumptive, and anaphoric

pronouns agree in gender and number, whereas their case is dependent upon their syntactic use (examples of case attraction not being attested); (iv) verbs agree with their subject in person and number. The existence in Old Persian of the Proto-Indo-European use of a singular verb with a neuter plural subject cannot be demonstrated, both for lack of evidence and for orthographic reasons. The only evidence is found in the usual dating formulae (see §6), and there the copula *āha* (with *ṭakatā* nom. pl. neut.) may be third-person singular as well as plural.

## 5.7 Stylistics

A comprehensive and systematic study of the stylistic features that may be detected in the Old Persian inscriptions (which show clear traces of stylization), is an urgent desideratum. There is found evidence for the stylistic figures of the asyndeton, of chiasmus, parallelism, and so forth; see the discussion in Kent 1953 (pp. 99f. §§ 316–317 in the relevant paragraphs). Some additional stylistic features can be briefly noted here. Epiphora (repetition of the same words at the end of each of a set of sentences) occurs several times: for example, in DPd 22 and 24 *hadā visaibiš багаibiš* “with all the gods.” Examples of personification are attested: for example, with *dahyu-* “land” (which “does not fear anybody else”) or *dušiyāra-* “crop failure” (which “may not come”). But attempts to demonstrate rhyming phrases in Old Persian texts or to detect metrical passages (especially in DB) are not convincing in this author’s view.

## 6. LEXICON

The Old Persian vocabulary is known only in part owing to the limited corpus of the texts and to their stereotyped character. On the whole it corresponds closely to the vocabulary of the other attested ancient Aryan languages, Avestan and Old Indo-Aryan (especially Vedic). A striking characteristic feature of Old Persian is the considerable quantity of foreign words and names which it uses. Such foreign influences, however, are only to be expected in such a multinational state as that of the Persian Empire. Among those foreign elements, borrowings from the Median language take a special place, and they can be justified historically without difficulty. The fact that particular terms are of Median origin can sometimes be established by phonetic criteria, even if the non-Persian phonetic developments observed are not unique to the Median language, but also belong to other Old Iranian dialects. Medisms occur more frequently among royal titles and among terms of the chancellery, military, and judicial affairs (*vazyka-* “great,” *zūra-* “evil,” *zūrakara-* “evil-doer,” etc.); they are found not least in the official characterizations of the empire and its countries (*uvaspa-* “with good horses,” *vispazana-* “with all races,” etc.).

From a dialectological perspective, one notes some peculiar developments. Particularly striking is the case of the verb “to say, speak”; Old Persian continues neither Proto-Iranian *\*wač-* nor *\*mrau-*, both of which are attested in Avestan, but has *gaub-*. A similar case is found with “to hear”: Old Persian has lost Proto-Iranian *\*srau-* (Avestan *srauu-*), and has instead the root *ā-xšnau-* (literally “to grasp, understand”).

In addition to the shared isogloss of Old Persian *gaub-* “to say, speak” and Sogdian *γωβ-* ([ʁo:β-]) “to praise,” there are a number of remarkable features common to Old Persian (Southwest Iranian) and Sogdian (East Iranian). For example, to both belong *\*kun-* “to do” (from Proto-Iranian *\*kar-*, pres. *\*krynau-*) in Old Persian *kunau-* = Sogdian *kwn-* ([kun-]). Both share the meaning “to have” for the Iranian root *\*dar-* “to hold, keep” (Old Persian *dar-*, pres. *dāraya-*), and the dating formulae of the type Old Persian *NN māhyā X raucabiš ṭakatā*

*āha* “in the month NN X days had passed” and Sogdian *pr ’tδrtyk YRH’ pr 10 syth* “in the third month at/after ten passed [days].”

In other cases, borrowings from some East Iranian language have been assumed: for example, *kāsaka*- “semiprecious stone.” In addition, the influence of the other languages spoken by the indigenous peoples of the Ancient Near East can be detected in the Old Persian lexicon. Thus, the Persians seem to have acquired *dipi*- “inscription” from Elamite, *maškā*- “[raft of] skin” from some Semitic language, and *pīru*- “ivory” likewise from some Near Eastern source.

A considerable portion of the Old Persian lexicon has simply not survived (because of the nature of the texts). However, the possibility exists of reconstructing Old Persian lexemes, provided they are inherited from Proto-Aryan (and from Proto-Indo-European), by comparing the Proto-Aryan vocabulary (which can be reconstructed from the very rich records available in Old Indo-Aryan) with Middle and Modern Persian words, since such later attested lexemes necessarily must have passed through an Old Persian stage.

In addition, a great many Old Persian lexemes, including proper names, are preserved in a borrowed form in non-Persian languages – the so-called “collateral” tradition of Old Persian (within or outside the Achaemenid Empire). The main sources of that tradition are Elamite (especially the Persepolis tablets), Late Babylonian (with numerous administrative texts), Aramaic (as the lingua franca of the official imperial administration), Hebrew, Egyptian, and Greek authors (from Aeschylus and Herodotus) and inscriptions. It must be borne in mind, however, that not every purported Old Iranian form attested in this manner is an actual lexeme of Old Persian. Thus, for example, the title “satrap,” best known in its Greek form σατράπης, in fact mirrors Median *\*xšaθra-pā*-, whereas the first element of the Old Persian form was *xšaça*- and the form attested epigraphically is *xšaça-pā-van*-. A collection of the complete material attested in the various branches of the collateral tradition is not available; Hinz 1975 offers the most comprehensive collection, though is far from being complete (e.g., by omitting even Median *\*xšaθra-pā*-) and is often unreliable.

## 7. READING LIST

The most comprehensive treatment of Old Persian (containing a full descriptive as well as historical grammar, the transcribed texts with English translation, and a lexicon with full references) is found in Kent 1953; for a traditional grammar see also Meillet and Benveniste 1931. A more structured outline of morphology and an etymological lexicon (including, in part, the collateral tradition) is presented by Mayrhofer in Brandenstein and Mayrhofer 1964 (pp. 55–82 and 99–157). Mayrhofer 1979: II (pp. 11–32) provides a special treatment of the personal names attested in the inscriptions. A brief account of the Old Persian language (with the most essential bibliography) is also presented in Schmitt 1989.

A complete corpus of all Old Persian Achaemenid inscriptions is not available; there are only partial collections outdated by later discoveries or limited to certain groups or types of texts. The Old Persian texts alone can be found in Kent 1953: 107–157 (with an English translation); this has been supplemented by Mayrhofer 1978, who also provides a full inventory list of the Old Persian texts (pp. 37–47); though even this list is not up to date.

## Abbreviations

The most important Old Persian texts are listed below. Texts are usually cited utilizing a system of abbreviations, in which the king's name normally appears first (D = Darius I, X = Xerxes I, A<sup>1–3</sup> = Artaxerxes I–III, etc.), followed by the place of origin (B = Bīsutūn,

P = Persepolis, N = Naqš-i Rostam, S = Susa, etc.). Several texts by the same king at the same place are distinguished by additional small letters:

- DB: the major inscription of Darius I at the rock of Mt. Bisutūn, the most extensive and most important trilingual inscription, with five columns and 414 lines of Old Persian text (newly edited by Schmitt 1991).
- DNa, DNb: two major trilingual inscriptions at the tomb of Darius I at Naqš-i Rostam, the lower text DNb being some kind of guide for the ideal ruler (new edition by Schmitt 2000:23–44).
- DPd, DPe: two monolingual Old Persian inscriptions which form part of an ensemble of texts at the southern wall of the Persepolis terrace and in all probability are the oldest Persepolitan inscriptions (new edition by Schmitt 2000:56–62).
- DSab: the trilingual cuneiform text on the Egyptian-made statue of Darius I excavated in Susa in 1972.
- DSe, DSf: two major trilingual building inscriptions from the palace of Susa, which are preserved, however, only in a great number of fragments.
- DZc: the longest of the cuneiform inscriptions from the Suez Canal.
- XPf: a bilingual (Old Persian and Babylonian) foundation document of Xerxes from Persepolis, which is of special historical importance owing to some details reported about the king's succession.
- XPh: the trilingual, so-called Daiva-inscription describing a revolt and praising the cult of Auramazdā (rather than the Daivas).
- XPl: an Old Persian text on a stone tablet, which is essentially parallel to DNb, but associated with the name of Xerxes I.

## Bibliography

- Brandenstein, W. and M. Mayrhofer. 1964. *Handbuch des Altpersischen*. Wiesbaden: Otto Harrassowitz.
- Hinz, W. 1975. *Altiranisches Sprachgut der Nebenüberlieferungen*. Wiesbaden: Otto Harrassowitz.
- Kent, R. G. 1953. *Old Persian: Grammar, Texts, Lexicon* (2nd edition). New Haven: American Oriental Society.
- Mayrhofer, M. 1978. *Supplement zur Sammlung der altpersischen Inschriften*. Vienna: Österreichische Akademie der Wissenschaften.
- . 1979. *Iranisches Personennamenbuch*. Vol. I. Vienna: Österreichische Akademie der Wissenschaften.
- Meillet, A. and E. Benveniste. 1931. *Grammaire du vieux-perse* (2nd edition). Paris: Édouard Champion.
- Schmitt, R. 1989. "Altpersisch." In R. Schmitt (ed.), *Compendium Linguarum Iranicarum*, pp. 56–85. Wiesbaden: Dr. Ludwig Reichert.
- . 1991. *The Bisutun Inscriptions of Darius the Great: Old Persian Text*. London: School of Oriental and African Studies.
- . 1995. "Zur Enklise im Altpersischen." In H. Hettrich, W. Hock, P. Mumm, et al. (eds.), *Verba et Structurae: Festschrift für Klaus Strunk*, pp. 285–301. Innsbruck: Institut für Sprachwissenschaft der Universität Innsbruck.
- . 2000. *The Old Persian Inscriptions of Naqsh-i Rostam and Persepolis*. London: School of Oriental and African Studies.



# Avestan

MARK HALE

## 1. HISTORICAL AND CULTURAL CONTEXTS

Avestan is a member of the Indo-European language family. It is the most richly attested ancient member of the Iranian branch of the Indo-Iranian subgroup of that family. As such, it is closely related to the Sanskrit language (which represents the most archaic member of the Indic subgroup of Indo-Iranian; see Ch. 26). For those who hold that the centum–satem division (fundamentally an east/west bifurcation in the Indo-European language family) is a matter of subgrouping, Indo-Iranian (and therefore Avestan) is a member of the satem group (indeed, it is the Avestan word for “100,” *satəm*, which gives that group its name).

There are uncertainties regarding both the dating and the geographical provenance of the surviving Avestan texts. The oldest manuscript is quite young (manuscript K7a, dating from AD 1278) and therefore of little assistance in resolving these matters. The issue of chronology is usually linked to the problems surrounding the dates of the founder of Zoroastrianism, the prophet Zarathuštra. Current scholarly consensus places his life considerably earlier than the traditional Zoroastrian sources are thought to, favoring a birth date before 1000 BC. Since the *Gāthās* are recognized as being the work of Zarathuštra, these Old Avestan texts appear to date from around that time. Precise dating of the Young Avestan texts, many of which appear to have a long oral transmission history, is in most cases impossible. Regarding geography, the Avestan language itself is now widely believed to be an Eastern Iranian language, though it cannot be directly connected to any known group of ancient Iranian speakers, thus greater geographical precision is not at this time possible. For the most recent and more coherent consideration of these complex issues, the interested reader is referred to the introduction to the first volume of Humbach *et al.* (1991). The Avestan texts continue to be used in ritual and other hieratic contexts in Zoroastrian communities.

Although Avestan is quite conservative in several crucial respects both phonologically and morphologically, many (though not all) of its archaisms are also found in the better-attested, better-preserved, and generally more widely studied Sanskrit language, leading to a certain degree of neglect of Avestan in Western scholarship. This has been rectified to some extent in the postwar era of Indo-European studies, during which the type of philological problems posed by the Avestan records have captured the attention of many prominent Indo-Europeanists.

It is traditional to refer to the two major dialects of Avestan as *Old* (or *Gāthic*) *Avestan* and *Young Avestan*. Nevertheless, it is apparent that the relationship between the two dialects is not strictly a chronological one (i.e., Young Avestan is not a direct descendant of Old Avestan). These labels may accurately reflect the relative chronology of the respective corpora, although the matter is complicated by the fact, noted above, that many of the Young Avestan texts

appear to be originally oral compositions with a potentially long transmission history before becoming fixed canonical texts.

The Avestan language is transmitted almost exclusively through the surviving text of the *Avesta*, a collection of Zoroastrian religious and legal texts. Unfortunately, the transmission history of these texts involves several serious disruptions, leading to loss of a large number of texts (the contents of which can be in part gleaned from a surviving Pahlavi, i.e., Middle Persian, summary) and challenging philological problems for those texts which do survive. Excluding a number of minor texts, there are three major sections of the surviving Avesta: (i) the *Yasnas* (Y.), containing prayers, hymns, and liturgical works; (ii) the *Yašts* (Yt.), containing invocations of specific holy figures and concepts; (iii) the *Vidēvdāt* (V.), containing “legal” texts, broadly construed. All of the Old Avestan texts are contained within the *Yasnas*. These texts include the *Gāθās* (metrical hymns the composition of which is attributed to the prophet Zarathuštra), the prose liturgy of the *Yasna Haptaŋhāiti*, and a set of short prayers, the most sacred in Zoroastrianism.





















Many of the *Yašts* are rather poorly preserved, or were not originally native-speaker compositions. One usually distinguishes between the best-transmitted *Yašts* – the so-called *Great Yašts* – and the lesser works. The *Great Yašts* represent the high points of Young Avestan literature. Included among them are Yt. 5 (in honor of Arəduuī, the personification of a mythic river); Yt. 8 (in honor of Tištriia, the personification of the star Sirius); Yt. 10 (in honor of Miθra, the personification of the contract); Yt. 13 (in honor of the Frauuašis – protective spirits of the faithful); Yt. 14 (in honor of Vərəθraϥna, the personification of victory); Yt. 17 (in honor of Aši Vərəuuhī, the personification of the reward of the pious); Yt. 19 (in honor of Xʼarənah, the personification of royal power/glory); and two *Yašts* preserved in the *Yasna* section of the Avesta: Y.9–Y.11.8 (in honor of Haoma, the Avestan cognate of Sanskrit *soma*, a ritualistic intoxicant) and Y.57 (in honor of Sraoša, the personification of obedience to divine will).

Finally, the *Vidēvdāt*, while containing some significant mythological material, focuses the bulk of its attention on matters of purity and pollution, of crime and of punishment. It is of great significance for our understanding of the history of Zoroastrian doctrine and practice.

As noted above in the discussion of the chronology of Avestan, the two major dialects are in part chronological and in part almost certainly geographical variants of one another. They are sufficiently distinct – although the bulk of the identified contrasts are in the phonological domain – that I have chosen to focus on the more extensively transmitted variant, that of Young Avestan, in what follows. I will not, however, hesitate to cite Gāthic forms where appropriate or necessary, noting the forms as such. Young Avestan itself does not appear to have been uniform, though the study of its variants faces a number of philological difficulties. The differences between Young Avestan dialects are, at any rate, too minor to be of concern in a survey of this type.

The texts themselves show clear evidence of indigenous scholarly redaction, much like the *pada*-texts of the Vedic Sanskrit tradition. For example, in the transmitted text of the Avesta, sandhi – phonological variation conditioned by the context in which a word is placed – has been for the most part eliminated through the generalization of a single sandhi variant for each final sequence. Clear evidence of redactorial intervention in the text can be seen in the orthographic repetition, in Gāthic Avestan, of preverbs which are separated from their verbs (i.e., in *tnesis*, much like German separable prefixes) in a position immediately preceding the verb itself. Thus, *Yasna* 32.14 transmits *nī . . . nī.dadaṭ* “they put down,” where the meter assures us that the intended reading is *nī . . . dadaṭ*. The doubling of the “preverb” *nī* before the verb *dadaṭ* appears to represent an indigenous analytical hypothesis about the syntactic

## 2. WRITING SYSTEM

1  a	2  ā	3  ä	4  ä̇	5  a	6  ä	7  ə	8  ä̇				
		9  e	10  ē	11  o	12  ō	13  i	14  ī	15  u	16  ū		
17  k	18  x	19  ẋ	20  ẋv	21  g	22  ġ	23  γ					
24  c	25  j										
26  t	27  θ	28  d	29  ḋ	30  ż							
31  p	32  f	33  b	34  β								
35  θ̇	36  θ̇	37  θ̇v	38  n	39  ñ	40  ñ̇	41  m	42  ṁ				
43  j̇	44  γ̇	45  v̇	46  ṙ								
47  s	48  z	49  ṡ	50  ż	51  ṡ̇	52  ż̇						
53  h											

1	2	3	4	5	6	7	8
ا	ب	ت	ث	ج	ح	ز	ح
ā	ā	ā	ā	ā	ā	ā	ā
		9	10	11	12	13	14
		و	و	و	و	و	و
		e	e	o	o	i	i
17	18	19	20	21	22	23	
ق	ك	ع	خ	غ	ج	ح	
k	x	ā	x <sup>v</sup>	g	g	γ	
24	25						
ف	ج						
c	j						
26	27	28	29	30			
ط	ث	د	ذ	ر			
t	θ	d	ð	ɾ			
31	32	33	34				
ز	ح	ج	ح				
p	f	b	β				
35	36	37	38	39	40	41	42
س	ش	ص	ض	ط	ظ	ع	غ
s	ʃ	ʂ	z	ʈ	ʈ	m	ŋ
43	44	45	46				
ق	ك	و	و				
q	q	v	r				
47	48	49	50	51	52		
ن	ف	و	و	و	و		
n	f	z	z	ʃ	ʃ		
53	54	55	56	57	58	59	60
هـ	و	و	و	و	و	و	و
h	h	h	h	h	h	h	h

The transliteration given in Table 29.1 for each character is now the standard, but differs in some details from prominent earlier work on Avestan (such as Reichelt's 1909 grammar and Bartholomae's 1904 dictionary). The principal differences are as follows:

- (1)  $\text{ā}$  (3) was formerly transliterated  $\text{ā̇}$ , thus identically to (4)
- $\text{ā}$  (6) was formerly transliterated  $\text{ā̇}$ , thus identically to (5)
- $\text{x}$  (19) was formerly transliterated  $\text{h}$
- $\text{g}$  (22) was formerly transliterated  $\text{g}$ , thus identically to (21)
- $\text{c}$  (24) was formerly transliterated  $\text{č}$
- $\text{j}$  (25) was formerly transliterated  $\text{j̇}$
- $\text{β}$  (34) was formerly transliterated  $\text{w}$
- $\text{η}^v$  (37) was generally transliterated  $\text{ηu}$
- $\text{n}$  (39) and  $\text{ṇ}$  (40) were formerly transliterated  $\text{n}$ , thus identically to (38)
- $\text{ṁ}$  (42) was formerly transliterated  $\text{hm}$
- $\text{y}$  (43), as well as  $\text{ii}$  sequences, were formerly transliterated  $\text{y}$ , thus identically to (44)
- $\text{uu}$  sequences were formerly transliterated  $\text{v}$ , thus identically to (45)
- $\text{ś}$  (51) and  $\text{ṣ}$  (52) were formerly transliterated  $\text{š}$ , thus identically to (49)

In many of these cases the underdifferentiation of characters extends to Western books printed in Avestan characters, including Geldner's (1886–1896) extensive critical edition of the bulk of the Avestan corpus. Character 6 ( $\text{ā̇}$ ), for example, is generally not used in Geldner's edition, even in the critical apparatus. Moreover, some of these distinctions are lacking in certain Avestan manuscripts or manuscript traditions (for example, the  $\text{y} : \text{j̇}$  contrast is generally, though not universally, absent from Indian manuscripts).

The phonetic value of some of these characters, especially some of the “minor” ones which were earlier not distinguished, is not particularly clear, though there is published speculation on virtually all of them. In general, however, we can be fairly confident about the values assigned to the vast majority of symbols.

### 3. PHONOLOGY

#### 3.1 Phonemic status

Determining the precise phonemic inventory of Avestan is problematic, though further research may allow us to resolve some or all of the outstanding issues. The writing system, designed to capture the nuances of hieratic recitation, is closer to the phonetic level. The principal difficulties arise from the fact that some relevant aspects of the sound system of Avestan are not explicitly indicated in the writing system. For example, there are no direct encodings of the position of stress (though some aspects of stress placement can probably be safely inferred), nor of syllable boundaries (which appear to be relevant to the determination of the phonemic status of some segments). In addition, as pointed out above (see §1), the final sandhi variants which were certainly present in the language (as indicated by their rare preservation in fixed phrases, for example) have been for the most part leveled out in the transmitted text.

#### 3.2 Consonants

The approximate phonetic values of the consonant symbols are generally not in dispute. The uncontroversial stops, fricatives, affricates, and nasals of Avestan are presented in Table 29.2.

**Table 29.2 Consonantal sounds of Avestan**

Manner of articulation	Place of articulation							
	Labial	Inter-dental	Dental/Alveolar	Palato-alveolar	Palatal	Retro-flex	Front velar	Round velar
<i>Stop</i>								
<i>Voiceless</i>	p		t					k
<i>Voiced</i>	b		d					g
<i>Fricative</i>								
<i>Voiceless</i>	f	θ	s	š	ś	ṣ	χ	x
<i>Voiced</i>	β	ð	z	ž			xʷ	ɣ
<i>Affricate</i>								
<i>Voiceless</i>					c			
<i>Voiced</i>					j			
<i>Nasal</i>								
<i>Voiceless</i>	ɱ							
<i>Voiced</i>	m		n		ɲ		ŋ	ŋʷ

In addition, it is generally recognized that the symbol *t̥* represents an “unreleased” voiceless dental stop – it is extremely limited in distribution, being regularly found only in word-final position and before certain obstruents. The phonetic nature of *ɲ* is taken by Hoffmann to be a “postuvular nasal” without oral occlusion of any type.

The values of the symbols which represent liquids and glides present only minor difficulties of detail. There appears to be a voicing contrast in the liquids between *r* and the digraph *hr*, the latter being voiceless. The symbol *v* appears to differ from *β* by the former being round, the latter not. While the symbol *h* is uncontroversially held to be a glottal approximant, there is some speculation that the symbol transliterated as *y* originally represented *ž*, contrasting therefore with the voiced palatal glide (which was represented by the symbol *ȳ*). As noted, the contrast is not observed in all manuscripts nor by the earlier Western scholarly tradition. A detailed study of the distribution of these two symbols in manuscripts which use both of them remains a desideratum.

### 3.3 Vowels

The confidently identified vowel symbols may be approximately distributed in the vowel space as in Figure 29.1:

	FRONT	CENTRAL	BACK
HIGH	ī / i		ū / u
MID	ē / e	ə / ə	ō / o
LOW		ā / a	ā

**Figure 29.1** Avestan vowels

A macron indicates vowel length; however, it seems likely, and is now generally accepted, that the original length contrast has become a qualitative one, either as well as or instead

of a purely quantitative one. It must be noted in this regard that the manuscripts do not, in general, do a particularly good job of distinguishing length contrasts in the high vowels *ī* and *ū*. The vowel *q* represents nasalized *a* as well as nasalized *ā*. In addition to these simple vowels, Avestan has a number of diphthongs, including the so-called short diphthongs *aē*, *ōi*, and *ao* and the long diphthongs *āi* and *āu* (see §3.4.9).

One of the most salient differences between Gāthic and Young Avestan concerns vowel quantities in absolute word-final position. In Gāthic Avestan all such vowels are long, whereas in Young Avestan final vowels are long only in monosyllables (discounting a few sandhi forms, on which more below). The fact that monosyllables are treated differently in this regard than polysyllables in Young Avestan allows one to determine certain otherwise somewhat obscure facts about the syllabification of Young Avestan word forms. For example, the instrumental singular of the word for “earth” (*zam-*) is transmitted as *zamā*, which must, given the rule just stated concerning final vowel quantities, represent a monosyllable. The epenthesis is thus phonologically irrelevant (either postdating the rule regulating final vowel quantities or too low-level phonetic to be of concern, or both). It is, as it turns out, also metrically irrelevant, the phonological facts thus supporting the analysis of the meter nicely. This case can be contrasted with that of the nominative singular of the word for “bowstring,” *jūia*, Sanskrit *jyā*, which must be disyllabic in Avestan given its short final vowel, as it originally was in Sanskrit.

### 3.4 Diachronic developments

#### 3.4.1 Proto-Indo-Iranian

Avestan, being an Indo-Iranian language, shares with Sanskrit the phonological developments of Proto-Indo-Iranian (PIIr.). The most salient of these are (i) the merger of the labiovelar and velar stop series (definitional of satem languages); (ii) the development of the syllabic nasals to PIIr. *\*a*; (iii) the RUKI-inducing backing of PIE *\*s* (in Avestan to *š*); (iv) the merger of PIE *\*e*, *\*a*, and *\*o* into PIIr. *\*a*, and that of PIE *\*ē*, *\*ā*, and *\*ō* into PIIr. *\*ā*. In keeping with *Brugmann’s Law*, short *\*o* in open syllables shows up in Avestan as *ā*, rather than the expected *ā* (examples include *dāuru* “wood” < PIE *\*doru* and *srāuuaiia-*, the causative stem of *sru* “to hear” < PIE *\*kloweye-*). The palatalization of the Proto-Indo-European velars (and the Proto-Indo-European labiovelars which had fallen together with this set of stops) before front vowels and *\*y* preceded the merger of the vowels.

Avestan provides key evidence for the status of PIE *\*T<sup>s</sup>T* (< *\*TT*, where *T* is any dental stop) in Proto-Indo-Iranian: whereas Sanskrit shows *TT* as the outcome of this sequence (*vittā-* “found” < PIE *\*vit<sup>s</sup> tó-*, morphologically *\*vid* + *\*-tó*), Avestan has *ST* (thus *vista-* “found”). The evidence of these two major branches of Indo-Iranian points to preservation in Proto-Indo-Iranian of PIE *\*T<sup>s</sup>T*, thus suggesting the reconstruction of an affricate-formation rule for Proto-Indo-European phonology.

#### 3.4.2 Indo-European laryngeals

In the matter of the laryngeals of Proto-Indo-European (see Ch. 17, §2.1.3), Avestan provides only limited direct phonological evidence. In virtually all positions, the laryngeals have disappeared without a trace. There are, however, two exceptions to this statement. First, in Old Avestan the hiatus left by intervocalic laryngeal loss is generally preserved, as indicated by the syllable-counting meter of the Gāthas. Thus, the apparently disyllabic *zrazdā*, the nominative plural of *zrazdā-* “having faith,” the second compound member of which comes

from PIE *\*d<sup>h</sup>eH<sub>1</sub>es*, scans as a trisyllable in Old Avestan (*\*d<sup>h</sup>eH<sub>1</sub>es* having become PIIr. *\*d<sup>h</sup>aHas* and then pre-Avestan *\*daas*, with two syllables). Unfortunately, our lack of a firm understanding of Young Avestan meter, coupled with, and in part deriving from, the flawed transmission of the relevant metrical texts, does not permit us to determine conclusively whether such scansions were also attested in this language.

Secondly, in a few instances interconsonantal laryngeals appear to have “vocalized” to *i* in Avestan, much as in Sanskrit. This is particularly clear in the paradigm of “father,” PIE *\*pH<sub>2</sub>ter-*, which shows laryngeal vocalization in the nominative singular (*pitā*, Skt. *pitā́*), accusative singular (*pitārem*, Skt. *pitāram*), and dative singular (*pit<sup>re</sup>*, Skt. *pit<sup>rē</sup>*). This development must be seen as dialectal, since Avestan also shows forms of this paradigm without traces of the vocalized laryngeal, including Old Avestan nominative singulars *ptā* and *tā*, Old Avestan dative singular *fəðrōi*, and the Young Avestan accusative plural *fəðrō* (the schwas in the last two forms are the result of a late epenthesis process – they do not count for purposes of the meter, and thus were apparently not there at the time of composition; such epenthetic schwas will not be explicitly pointed out in the discussion which follows). One may also contrast Avestan *duyðar-* “daughter” both with Sanskrit *duhitār-* (where the *i* represents the vocalized laryngeal) and with Greek *t<sup>h</sup>ugátēr* (where the laryngeal is represented by *a*), all three from PIE *\*d<sup>h</sup>ugH<sub>2</sub>ter-*. These forms make it impossible to see laryngeal vocalization to *i* as a property of Proto-Indo-Iranian itself in spite of the fact that only Sanskrit and, in some instances, Avestan, appear to show such a development within the Indo-European family.

Indirect evidence of the prior presence of the laryngeals is, by contrast, quite easy to come by. The sequence of syllabic nasal + laryngeal yields Avestan *ā*, giving rise to alternations of the type *zan-* “give birth” (< PIIr. *\*jan* < PIE *\*genH<sub>1</sub>-*, Skt. *jan*) : *zāta-* “born” (< PIIr. *\*jāta-* < PIE *\*gyH<sub>1</sub>to-*, Skt. *jātā-*). More interesting is the divergence between Avestan and Sanskrit in the treatment of pre-laryngeal syllabic liquids (PIE *\*l<sup>h</sup>H* and *\*l<sup>h</sup>H*). Whereas Sanskrit regularly shows *īr* from such sequences, the Avestan reflex is *ar*: for example, *darəya-* “long” < PIE *\*d<sup>l</sup>H<sub>1</sub>g<sup>h</sup>o-* (Skt. *dīrghā-*); *starəta-* “strewn” < PIE *\*st<sup>l</sup>H<sub>3</sub>to-* (cf. Skt. *stīrṇā-*), *tarō* “across” < PIE *\*t<sup>l</sup>H<sub>2</sub>es* (Skt. *tirāḥ*). The best reconstruction for the Proto-Indo-Iranian reflex of these sequences is not at all clear given the Avestan and Sanskrit developments.

### 3.4.3 Stops

A number of distinctive phonological developments in the consonant system give Avestan a quite different “look” from that of Sanskrit. Quite salient among these is the development of the Proto-Indo-European palatal stops (*\*k̑*, *\*g̑*, and *\*g̑<sup>h</sup>*). In the first instance, these stops develop into palatal fricatives in Proto-Indo-Iranian, usually designated *\*ç*, *\*j*, and *\*jh*, respectively (and thus distinguished from the outcome of the palatalization of the Proto-Indo-European plain and rounded velar stops, which became the affricates *\*ç̑*, *\*j̑*, and *\*jh̑*). The place of articulation of these fricatives then shifts to the dental region, and we find *s* as the regular reflex of *\*ç*, and *z* as the regular outcome of both *\*j* and *\*jh* (with the regular Avestan loss of distinctive aspiration of the voiced aspirates). Examples include *satəm* “100” < PIIr. *\*catam* < PIE *\*k̑ntom* (Skt. *śatām*); *zan-* “beget” < PIIr. *\*jan-* < PIE *\*genH<sub>1</sub>-* (Skt. *jan-*); *zari-* “yellow” < PIIr. *\*j<sup>h</sup>ali-* < PIE *\*g̑<sup>h</sup>eli-* (Skt. *hāri-*).

The voiceless unaspirated stops of Proto-Indo-Iranian have been generally preserved. However, they have developed into voiceless fricatives preconsonantly (excepting *\*p* before *\*t*, which remains unchanged): for example, Av. *xratu-* “insight” < PIIr. *\*kratu-* (Skt. *krātu-*); Av. *friia-* “beloved” < PIIr. *\*priHa* (Skt. *priyā-*). Contrast Av. *hapta* “seven” < PIIr. *\*sapta*



(Skt. *saptā*). This development does not take place if the stop in question is preceded by PIIr. \*s (or its RUKI-variant, \*š): thus, *vastra*- “clothing” < PIIr. \**wastra*- (Skt. *vāstra*-); *uštra*- “camel” (cf. Skt. *úštra*-), with preserved \*t.

The voiceless aspirated stops of Proto-Indo-European have become corresponding voiceless fricatives: for example, Av. *haxāii*- “companion” < PIIr. \**sak<sup>h</sup>āy*- (Skt. *sákhāy*-); Av. *kafa*- “foam” < PIIr. \**kaph<sup>h</sup>a*- (Skt. *kapha*-). Avestan preserves better than Sanskrit the paradigm-internal effects of aspiration arising from an ensuing \*H<sub>2</sub> in the Proto-Indo-Iranian word for “path,” which has the nominative singular *paṇtā* < PIIr. \**pantaH<sub>2</sub>-s* (contrast Skt. *pánthāḥ*, with generalized aspiration), genitive singular *paθō* < \**pṇtH<sub>2</sub>as* (Skt. *pañthāḥ*).

The voiced aspirated stops of Indo-European (and Indo-Iranian) have merged, via loss of aspiration, with corresponding simple voiced stops in Avestan (and Iranian generally). The resulting voiced stops are generally preserved as such in Old Avestan, but have lenited (or “weakened”) to voiced fricatives in all but a few positions in Young Avestan. They are generally preserved as stops only in word-initial position (except in a few word-initial consonant clusters) and after nasals and fricatives. These developments can be seen in the following examples, sorted by place of articulation:

1. Iranian \*b (< PIE \*b, \*b<sup>h</sup>) >
  - (i) Avestan *b*: *brātā* “brother” < PIIr. \**b<sup>h</sup>rātā* (Skt. *bhrātā*); Avestan *xumba*- “pot” < PIIr. \**k<sup>h</sup>umb<sup>h</sup>a*- (cf. Skt. *kumbhá*-).
  - (ii) Avestan *β*: *aīβi* “toward” (Old Avestan *aibī*) < PIIr. \**ab<sup>h</sup>i* (Skt. *abhí*).
2. Iranian \*d (< PIE \*d, \*d<sup>h</sup>) >
  - (i) Avestan *d*: *dasa* “ten” < PIIr. \**daśa* (Skt. *dáśa*); *viṇdənti* “they find” < PIIr. \**windanti* (Skt. *vindánti*).
  - (ii) Avestan *ð*: *maða*- “intoxicating drink” (Old Avestan *mada*-) < PIIr. \**mada*- (Skt. *máda*-).
3. Iranian \*g (< PIE unpalatalized \*g, \*g<sup>w</sup>, \*g<sup>h</sup>, \*g<sup>wh</sup>) >
  - (i) Avestan *g*: *garəma*- “warm” < PIIr. \**g<sup>h</sup>arma*- (Skt. *gharmá*-); *zəŋga*- “ankle” < PIIr. \**jang<sup>h</sup>a*- (cf. Skt. *jāṅghā*- “shin”); *mazga*- “marrow” (cf. Skt. *majján*-).
  - (ii) Avestan *γ*: *darəya*- “long” (Old Avestan *darəga*-) < PIIr. \**d<sup>l</sup>H<sub>1</sub>g<sup>h</sup>a*- (Skt. *dīrghá*-); *uyra*- “strong” (Old Avestan *ugra*-) < PIIr. \**ugra*- (Skt. *ugrá*-).
4. Iranian \*j (PIE palatalized \*g, \*g<sup>w</sup>, \*g<sup>h</sup>, \*g<sup>wh</sup>) >
  - (i) Avestan *j*: *jani*- “woman” < PIIr. \**janī*- (Skt. *jánī*-); *rəŋja*- “move quickly” < PIIr. \**ranj<sup>h</sup>a*- (Skt. *raṁha*- “run”).
  - (ii) Avestan *ž*: *azī*- “serpent” < PIIr. \**až<sup>h</sup>i*- (Skt. *áhi*-); *dažaiti* “he burns” (transitive) < PIIr. \**daž<sup>h</sup>ati* (Skt. *dáhati*).

Exceptions to the Young Avestan lenition processes evidenced above are attested. While some exceptional forms appear to represent the borrowing of religious vocabulary from the Gāthic dialect, others seem to require the assumption of dialectal developments within Young Avestan itself. Finally, in a number of cases, analogical restructuring appears to be at work. For example, in a reduplicated form such as *dadāθa* “you give,” built to the verbal root *dā*, the transparency of the reduplicative morphology has allowed the medial *d* to avoid lenition (or, more likely, to be remade to *d* after undergoing lenition). Similarly, in a number of transparent compounds the first member of which ends in a vowel and the second member of which begins with a voiced stop (e.g., *hu-baoði*- “having a good fragrance”), lenition of the morpheme-initial voiced stop is lacking. Analogy to the uncompounded form (*baoði*- “fragrance”) is clearly at work. Note that, in the example cited, the presence of lenition on the dental stop of *hubaoði*- makes a dialectal explanation for the lack of lenition on the labial stop unlikely.

In spite of the general loss of aspiration on voiced stops treated above, Avestan does preserve some morphological traces of the original aspiration through the workings of *Bartholomae's Law*. This law states that the direction of voicing assimilation in obstruent clusters (usually regressive) is reversed just in the cases in which the first obstruent is a voiced aspirate. In addition, the aspiration originally present on the first obstruent is shifted to the second. A Sanskrit example will make this clear: when the *-ta-* participial suffix is added to the verbal root *ṽṛdh* “grow,” Bartholomae’s Law triggers the following development: *ṽṛdh* + *ta-* > *ṽṛddha-* “grown.” The corresponding Avestan form, with the *ST* treatment of the dental cluster, as expected, is *ṽṛəzda-* “grown.” Although the aspiration is no longer present, its earlier existence is reflected in the rightward spread of voicing. The effects of Bartholomae’s Law are well preserved in Old Avestan, but frequently Young Avestan has analogically recreated the forms, applying the much more general regressive voicing assimilation to the cluster created in the remaking. Thus, corresponding to the Old Avestan third singular *aogədā* “he spoke,” from the verbal root *aog* (cf. Skt. *ohate*) + *-ta*, the ending of the third singular middle, Young Avestan generally has *aoxta*.

### 3.4.4 Fricatives

Avestan shares with Greek (though independently, of course) the development of Proto-Indo-European prenasorant *\*s* to *h*. This Proto-Iranian *\*h* underwent a number of conditioned changes in Avestan, of which the principal ones are as follows:

1. *\*h* > *ḡh* between low vowels (Av. *aḡhaiti* “he would be” < PIIr. *\*asati*, Skt. *asati*) – contrast the preservation of *\*h* before non-low vowels (Av. *ahi* “you are” < PIIr. *\*asi*, Skt. *asi*). Correspondingly, *\*āhwā* > *āḡhā*, *\*āhyā* > *āḡhā* (*pərəsaḡhā* “ask for your own benefit” < *\*pr̥cśaswa*, Skt. *pr̥cchasva*; *vaḡhō* “better” [nom. sg. neut.] < PIIr. *\*wasyas*, Skt. *váśyaḥ*)
2. Initial *\*hw-* > *xʷ-* (Av. *xʷafna-* “sleep” < PIIr. *\*swapna-*, Skt. *svápna-*).
3. Initial *\*hm-* > *m-* (Av. *mahi* “we are” < PIIr. *\*smasi*, Skt. *smasi*), contrast preservation of this sequence word-internally (Av. *ahmi* “I am” < PIIr. *\*asmi*, Skt. *asmi*).
4. Final *\*-ah* > *ō* (*-ō* nom. sg. masc. ending of thematic nouns < PIIr. *\*-as* < PIE *\*-os*), compare the Sanskrit sandhi of final *-ah* > *-o* before voiced segments.
5. Final *\*-āh* > *-ā* (*m ā* nom. sg. masc. “moon” < PIIr. *\*maas* < PIE *\*meHys*).

An exception to the development of Proto-Indo-Iranian *\*s* to Avestan *h* is provided by so-called RUKI contexts (i.e., when the *\*s* immediately followed any type of *r*, *ṽ*, *ṣ*, velar stop, or palatal affricate). In such a context, PIIr. *\*s* and *\*z* show up as Avestan *š* and *ž*, respectively. Examples include: *viša-* “poison” (Skt. *viṣā-*), *mižda-* “payment” (Skt. *mīḍhá-*). Interestingly, in Avestan (though not in Sanskrit), we find the same development after labials: *drafša-* “banner” < PIIr. *\*drapsa-* (Skt. *drapsā-*), *vaβža-ka-* “wasp” < PIIr. *\*wabzha-* < PIE *\*wobh-so-*.

### 3.4.5 Liquids

Proto-Indo-European *\*l* and *\*r* have merged as Avestan *r*, which is generally preserved as such. Interestingly, however, an *r* following a low vowel in the coda of a stressed syllable is devoiced before a following voiceless stop, the voiceless *r* being indicated by the digraph <*hr*>. In the case of *p* and *k*, nothing further befalls these segments: thus, *vəhrka-* “wolf” < PIIr. *\*wřka-* (Skt. *vřka-*); *kəhrpa-* “body” < PIIr. *\*křpa-* (Skt. *křpa-*). When the following voiceless stop was *t*, however, the sequence *hrt* became *š*: *mašīia-* “man” < PIIr.

\**mártiya-* (Skt. *mártya-*); *pəṣṣanā-* “battle” < PIIr. \**pṛtanā-* (Skt. *pṛtanā-*). Particularly instructive is the pair *mərəta-* “dead” (< PIIr. \**mṛtá-*, Skt. *mṛtá-*), *aməša-* “immortal” (< PIIr. \**a-mṛta-*, Skt. *amṛta-*).

### 3.4.6 Nasals

The nasals have developed into *ŋ* before stops and affricates: for example, *aŋtarə* “beside” (Skt. *antár*); *paŋca* “five” (Skt. *pañca*). The sequence *an* becomes *ɔ* before fricatives: *məθra-* “mantra” < \**mantra-* (Skt. *mántra-*). In most other positions PIIr nasals have been preserved.

It is worth pointing out that although in general the syllabic nasals have developed into PIIr. \**a*, before glides we find instead that PIE \**ŋ* > *an* (Av. *janiiāt* “he would smite” [with an analogical initial palatal] < PIIr. \**g<sup>h</sup>anyāt* < PIE \**g<sup>wh</sup>ŋ-yeH-t*, Skt. *hanyāt*) and PIE \**ŋ* > *am* (*jamiia* “you would go” [also with an analogical palatal] < PIIr. \**gamyās* < PIE \**g<sup>w</sup>ŋ-yeH-s*, Skt. *gamyāḥ*).

### 3.4.7 Glides

The glide \**w* shows a number of conditioned developments in Avestan. After the Proto-Indo-European palatal stops, this glide becomes a labial stop (voiceless after the Proto-Indo-European voiceless palatal stop, voiced after the Proto-Indo-European voiced and voiced aspirated palatal stops): for example, *aspa-* “horse” < PIE \**Heḱwo-* (Skt. *ásva-*); *zbaia-* present stem of “call” < PIIr. \**j<sup>h</sup>waya-* (Skt. *hváya-*). After the dental stops, it becomes a voiced labial fricative: *θβam* “you” (acc. sg.) < PIIr. \**twām* (Skt. *tvám*); *caθβārō* < PIE \**k<sup>w</sup>etwores* (Skt. *catvárah*); *aðβan-* “way” (Skt. *ádhan-*).

### 3.4.8 Vowels

The vowels of Avestan have in general undergone fewer modifications than the consonants, the exception being the short low vowel *a*. This vowel shows a number of conditioned changes, some of them apparently dialectal (and thus “sporadic” in our text), some of them quite regular. One of the more significant of the regular changes, because of its interaction with other phonological rules of Avestan, is the raising of *a* to *ə* before word-final nasals (and, dialectally, before word-internal nasals as well). The effects of this process are seen in nearly every line of the Avesta, producing forms such as the accusative singular of *a*-stems in *-əm* (thus *narəm* “man” [acc. sg.], Sanskrit *náram*) as well as forms such as *satəm* “100” (Sanskrit *śatám*).

This schwa is itself subject to further raising to *i* under the influence of a preceding palatal (*y*, *c*, *j*, or *ž*). Thus, the accusative singular masculine of the relative pronoun, corresponding precisely to Sanskrit *yám*, has undergone the following stages of development: \**yam* > \**yəm* > *yim*. Similarly, the accusative singular of the word for “deceit,” *druj-*, corresponding to Sanskrit *drúham*, is *drujim* (< earlier \**drujəm*).

Moreover, when the prenasal raising to schwa took place in the environment of a preceding consonant + glide sequence, the development went even further, with *-Cyə-* sequences becoming *-Cž-*, and *-Cwə-* sequences becoming *-Cũ-* (the lack of clarity about high vowel quantity is the result of the general problem of the transmission of quantities in the case of these vowels alluded to above). Examples include *haiθīm* “truth” < \**haθyəm* < \**satyam* (Skt. *satyám*) and *haurum* “whole” < \**harwəm* < \**sarwam* (Skt. *sárvam*).

The sequences *-ayə-* and *-awə-* show a corresponding assimilation of the *ə* to the preceding glide with subsequent loss of the glide. The end result of this development is the appropriate Avestan diphthong (see §3.4.9), as can be seen in the following examples: (i) *\*wayam* “we” (Skt. *vayám*) > *\*vayəm* > *\*vayim* > *\*vaim* > *vaēm*; (ii) *\*awam* “that” > *\*awəm* > *\*awum* > *\*aum* > *aom*. Glide + *ə* sequences show precisely parallel developments after *ā*, giving rise to the “long” diphthongs *āi* and *āu*.

### 3.4.9 Diphthongs

The development of the Proto-Indo-Iranian diphthong *\*ai* is dependent upon both position in the word (nonfinal vs. final) and syllable structure (open vs. closed) in Avestan. Turning first to the development of nonfinal *\*ai*, we find development to Avestan *aē* in open syllables – *aēiti* “he goes” < PIIr. *\*aiti* (Skt. *éti*) – but to Avestan *ōi* in closed syllables: *kauuōiš* “of the singer” (gen. sg. of *kauui-* “singer”) < PIIr. *\*kawaiš* (Skt. *kaveḥ*).

In word-final position, the usual development of PIIr. *\*ai* is to Young Avestan *-ē* (the length determined by syllable count, as always in Young Avestan), Old Avestan *-ōi*: for example, *naire* “man” (dat. sg.) < PIIr. *\*narai* (Skt. *nāre*, compare Gāthic *narōi*). After glides, however, the development is different, PIIr. *\*-wai* becoming *-uiie* (*aghuiie* “life,” dat. sg. of *ahu-*, < *\*ahwai*), PIIr. *\*-yai* becoming *-iiōi* (*maiōiiōi* “in the middle,” loc. sg. of *maiōiia-* < PIIr. *\*mad<sup>h</sup>yai*, Skt. *mādhye*; *yōi*, nom. pl. of the relative pronoun *ya-*, < *\*yai*, Skt. *yé*).

Proto-Indo-Iranian *\*au* does not show such a syllable-structure set of developments in Young Avestan, becoming *ao* in nonfinal position across the board: thus, *aojah-* “strength” < PIIr. *\*aujas-* (Skt. *ójas*); *gaoš* “of the cow” (gen. sg. of *gauu-*) < *\*gauš* (Skt. *góḥ*).

In final position, PIIr. *\*au* becomes Avestan *-uuō* (compare the *-iie* development of *\*ai* after glides): for example, *huuō* “that” < *\*sau* (Old Persian *hauv*, cf. Sanskrit *asáu*); *ərəzuuō* “O righteous one” (voc. sg. of *ərəzu-* “straight, correct, righteous”) < *\*rjvau*.

The so-called long diphthongs of Indo-Iranian, *\*āi* and *\*āu*, become Avestan *āi* and *āu*, respectively. Examples include the following: the dative singular of Avestan *a*-stems (PIE *o*-stems) such as (unattested) *aspāi*, dative singular of *aspa-* “horse” < PIE *\*Heḱwōi* (compare Greek *-ōi*, but contrast Skt. *-āya*); the nominative singular of the word *gauu-* “cow,” which has the form *gāuš* < PIIr. *\*gāus* (Skt. *gāuḥ*).

#### 3.4.10 Epenthesis

Avestan shows the effects of a relatively recent process of *i*-epenthesis. It is important to note that this epenthesis has no metrical effects and thus may postdate the time of the composition of the texts. There are two distinct versions of *i*-epenthesis – one word-initial, the other word-internal. The word-initial version is quite restricted, affecting only initial *\*ri-* and *\*θy-* (itself from *\*ty-*), as seen in *irišta-* “damaged” (Skt. *riṣṭa-*) and *iθiiejah-* “abandonment” (Skt. *tyájas-*). Both of these forms are disyllabic in Avestan. The word-internal version is much more general, occurring before dental and labial stops and fricatives as well as before *n*, *nt*, *r*, and *rm* if a front vowel or palatal glide follows. The phenomenon is quite common and can be seen in examples such as *barāiti* “he carries” (Sanskrit *bhāratī*) and *aīβi* “towards” (Sanskrit *abhi*).

Interestingly, this epenthesis appears to be an ongoing synchronic process. As such, it tells us something significant about the accentual system of Avestan at the stage during which *i*-epenthesis took place. The addition of the enclitic conjunction *-ca* “and” regularly undoes the effects of *i*-epenthesis in penultimate syllables (i.e., penultimate before the cliticization

of *-ca*). Thus, we find *baēšaziatica* “and he heals” for what would appear without *-ca* as *baēšaziiti*, or *varəðatica* “and it increases” next to *varəðaiti*. The standard explanation for such alternations is that the cliticization of *-ca* gave rise to an accent shift from the original penult to the syllable immediately preceding the *-ca*. Such shifts are characteristic of stress-based, rather than pitch-accent type accentual systems, indicating that unlike Sanskrit, the Avestan accentual system was of the former type. The *-ca* induced alternation also indicates that internal *i*-epenthesis should be expected only in stressed syllables.

Somewhat parallel to *i*-epenthesis, though much more restricted, is the phenomenon of *u*-epenthesis. Like *i*-epenthesis, the latter process is metrically irrelevant and thus would appear to be rather late. The phenomenon of *u*-epenthesis is essentially restricted to *ru* and *ruu* sequences. Standard examples include *uruuata-* “duty” (< \**rwata-*, which shows an Avestan metathesis of the initial cluster when compared to Skt. *vratá*) and *hauruua-* “whole” (< PIIr. \**sarwa-*, Skt. *sárva-*). Further evidence that this is a late process can be seen from the fact that in cases in which, dialectally, Young Avestan *β* has become *uu* (i.e., /w/) after *r*, the *u*-epenthesis is still triggered – thus *gəuruuāia-* “seize” (< PIIr. \**grb<sup>h</sup>āya-*).

## 4. MORPHOLOGY

### 4.1 Morphological type

Avestan is a highly inflected language, much like other Indo-European languages of very early attestation, making use of a rich set of derivational suffixes and inflectional endings in both the nominal and verbal systems.

### 4.2 Nominal morphology

The standard Indo-European cases, genders, and numbers are preserved in Avestan, where they serve to inflect nouns and adjectives, as well as pronouns. There are eight cases (nominative, accusative, instrumental, dative, ablative, genitive, locative, and vocative – generally cited in this order). There are three genders (masculine, feminine, and neuter), distributed in the usual archaic Indo-European manner (i.e., the masculine and neuter differ only in the nominative, vocative, and accusative, which are not distinguished from one another in the neuter). Finally, there are three numbers (singular, dual, and plural). Adjectives agree with their head nouns in case, number, and gender. The nominal inflection system appears quite robust throughout the period of attestation, although some breakdown in the understanding of the case system is evident in very late compositions.

The nominal paradigms may be roughly divided between *vocalic stems*, the descendants of PIE \*-*o* and \*-*eH<sub>2</sub>* stems, and *consonant stems*, continuing Proto-Indo-European consonant stems (see Ch. 17, §3.5). The latter frequently show ablaut variations in their suffixal (or occasionally root) syllables (on Indo-European ablaut, see Ch. 17, §3.2). For ablauting stems, it is often useful to distinguish between the so-called *strong cases* (nominative, accusative, locative, and vocative singular; nominative and accusative dual; and nominative plural) – characterized by full- or lengthened-grade ablaut before the ending – and *weak cases*, which show by contrast zero-grade ablaut. The paradigm of Indo-European thematic (*o*-stem) nouns, generally masculine or neuter, shows up in Avestan as follows (using \**Heḱwo-* “horse” > Avestan *aspa-* as an example, unattested cases of this particular lexeme being marked with an asterisk):

(2)	<i>Singular</i>	<i>Dual</i>	<i>Plural</i>
<i>Nominative</i>	aspō	aspa*	aspāñhō / aspa
<i>Accusative</i>	aspəm	aspa*	aspā*
<i>Instrumental</i>	aspa*	aspaēibiia	aspāiš*
<i>Dative</i>	aspāi*	aspaēibiia	aspaēibiiō*
<i>Ablative</i>	aspāt*	aspaēibiia	aspaēibiiō*
<i>Genitive</i>	aspahe	aspaiiā*	aspanam
<i>Locative</i>	aspe*	aspaiiō*	aspaēšu
<i>Vocative</i>	aspa*	aspa*	aspāñhō / aspa*

Indo-European stems in \*-eH<sub>2</sub>, generally feminines, inflect like Avestan *daēnā*- “religion”:

(3)	<i>Singular</i>	<i>Dual</i>	<i>Plural</i>
<i>Nominative</i>	daēna	daēne*	daēnā
<i>Accusative</i>	daēnam	daēnā	daēnā
<i>Instrumental</i>	daēna	daēnāibiia*	daēnābiš
<i>Dative</i>	daēnaiia / daēna	daēnāibiia*	aspaēibiiō
<i>Ablative</i>	daēnaiiāt	daēnāibiia*	daēnāibiiō
<i>Genitive</i>	daēnaiiā	daēnaiiā*	daēnanam*
<i>Locative</i>	daēnaiia*	daēnaiiā*	daēnāhu
<i>Vocative</i>	daēne*/daēna	daēne*	daēnā*

It is not practical in the present survey to list fully the many variants of consonant-stem inflection attested in Avestan. However, two representative paradigms will be presented: that of the Avestan masculine *r*-stem *nar*- “man” (PIE \*H<sub>2</sub>ner-)

(4)	<i>Singular</i>	<i>Dual</i>	<i>Plural</i>
<i>Nominative</i>	nā	nara	narō
<i>Accusative</i>	narəm		nərəuš
<i>Instrumental</i>	nara		
<i>Dative</i>	naire	nərəbiia	nərəbiiō
<i>Ablative</i>	nərət		nərəbiiō
<i>Genitive</i>	narš	narā	narəm
<i>Locative</i>	nairi		
<i>Vocative</i>	narə		

and that of the Avestan neuter *s*-stem *manas*- “thought” (PIE \*mene/os-):

(5)	<i>Singular</i>	<i>Plural</i>
<i>Nominative</i>	manō	manā
<i>Accusative</i>	manō	manā
<i>Instrumental</i>	manañhā	manōbiš
<i>Dative</i>	manañhe	
<i>Ablative</i>	manañhat	
<i>Genitive</i>	manañhō	-manañham
<i>Locative</i>	(manahi)	

Readers are referred to the more comprehensive grammars of Avestan (Hoffmann and Forssman 1996 or Reichelt 1909) for more details concerning the many classes of noun inflection.

### 4.3 Pronominal morphology

#### 4.3.1 Personal pronouns

The personal pronouns have singular, dual, and plural forms, though there are many gaps in attestation. For the accusative and some oblique cases, one must distinguish between tonic and enclitic forms, as elsewhere in Indo-European. The pronouns for the third person are generally supplied by demonstratives (see §4.3.2). Attested Young Avestan forms for the first and second persons are presented in Table 29.3; forms in parenthesis are Gāthic, provided when the Young Avestan form is unattested:

Table 29.3 First- and second-person personal pronouns of Avestan				
	First person		Second person	
	Tonic	Enclitic	Tonic	Enclitic
<i>Singular</i>				
<i>Nominative</i>	azəm		tūm	
<i>Accusative</i>	mam	mā	θβam	θβā
<i>Instrumental</i>			θβā	
<i>Dative</i>	māuuōiia	mē	(taibiiā)	tē
<i>Ablative</i>	(mat)		θβat	
<i>Genitive</i>	mana	mē	tauua	tē
<i>Dual</i>				
<i>Nominative</i>	(vā)			
<i>Accusative</i>	(əəūuuā)			
<i>Genitive</i>		(nā)	yauuākəm	
<i>Plural</i>				
<i>Nominative</i>	vaēm		yūžəm	
<i>Accusative</i>	ahma	nō		vō
<i>Instrumental</i>	(əhmā)		(xšmā)	
<i>Dative</i>	(ahmaibiiā)	nō	yūšmaoiio xšmāuuōiia	vō
<i>Ablative</i>	(ahmat)		yūšmat	
<i>Genitive</i>	ahmākəm	nō	yūšmākəm	vō

A special set of enclitic forms of the third-person pronoun is also attested. It does not distinguish between masculine and feminine, but has distinct neuter forms. It is found only for the accusative, except in the singular, where a dative-genitive form is also found:

(6)	<i>Masc./Fem.</i>	<i>Neuter</i>
<i>Acc. singular</i>	īm, hī, dim	it, diṭ
<i>Dat./Gen. singular</i>	hē	hē
<i>Acc. dual</i>	(ī)	
<i>Acc. plural</i>	hīš, diš	ī, di

#### 4.3.2 Demonstrative pronouns

An example of the inflection of demonstrative pronouns (usually referred to as the “pronominal inflection”) is presented in Table 29.4; the table shows the forms of Avestan *ta-* “this”



**Table 29.4 Demonstrative pronouns of Avestan**

	Masculine	Neuter	Feminine
<i>Singular</i>			
<i>Nominative</i>	hō / hā	taṭ	hā
<i>Accusative</i>	təm	taṭ	taṃ
<i>Instrumental</i>	tā	tā	aētaiia
<i>Dative</i>	aētahmāi	aētahmāi	
<i>Ablative</i>	aētahmāt	aētahmāt	
<i>Genitive</i>	aētahe	aētahe	aētaṇha / aētaiiā
<i>Locative</i>	aētahmi	aētahmi	
<i>Dual</i>			
<i>Nom./Acc.</i>	tā	tē	
<i>Genitive</i>	aētaiiā	aētaiiā	
<i>Plural</i>			
<i>Nominative</i>	tē	tā	tā
<i>Accusative</i>	tō / ta	tā	ta
<i>Instrumental</i>	(tāiš)	(tāiš)	
<i>Dative</i>	aētaēibiiō	aētaēibiiō	aētābiiō
<i>Genitive</i>	aētaēšam	aētaēšam	aētaṇham
<i>Locative</i>	aētaēšu	aētaēšu	

(fem. *tā-*), compare Sanskrit *tā-*. As in Sanskrit, the nominative singular masculine and feminine of this pronoun is formed from PIE *\*se/o-*, rather than *\*te/o-*. For cases in which the relevant form of *ta-* is not attested, but a form of the similarly inflected pronoun *aēta-* (likewise “this,” compare Sanskrit *etā-*) is attested, the *aēta-* form is provided.

#### 4.4 Verbal morphology

The Avestan verbal system, like that of Proto-Indo-European, is built around the verbal root. From such a root may be derived a set of tense-aspect stems (though not all roots are found in all tense categories), including the present stem, the aorist stem, and the perfect stem. To these stems are built the moods of Avestan, which continue more or less directly the like-named Proto-Indo-European mood categories. Not all tense stems form the basis for all moods. The moods include the indicative, the injunctive, the subjunctive, the optative, and the imperative. Finally, the endings are added to the mood-stem. The endings encode person, number, and voice – in addition, they play a role in the encoding of some moods. The Avestan categories indicated by the endings are much like those of Proto-Indo-European itself – person (first, second, third), number (singular, dual, plural), voice (active and middle, perhaps also stative and passive).

The endings themselves fall into four well-defined sets, each used in the expression of one or more tense/aspect categories: (i) primary endings (used in the indicative present, indicative future, and in part in the subjunctive); (ii) secondary endings (used in the indicative imperfect, indicative aorist, indicative pluperfect, injunctive, optative, and in part in the subjunctive); (iii) imperative endings (used in the imperative); and (iv) perfect endings (used in the indicative perfect). In the active, these endings have the forms which are presented in (7) (absence of a form indicates lack of attestation; — indicates that no form is expected):

(7)		<i>Primary</i>	<i>Secondary</i>	<i>Imperative</i>	<i>Perfect</i>
	<i>Singular</i>				
	<i>First</i>	-mi, (-a)	-(a)m	—	-a
	<i>Second</i>	-hi	-h	-Ø, di	-θa
	<i>Third</i>	-ti	-t, -t	-tu	-a
	<i>Dual</i>				
	<i>First</i>	(-uuahī)	(-uuā)	—	
	<i>Second</i>				
	<i>Third</i>	-tō, -θō	-təm		-atarə
	<i>Plural</i>				
	<i>First</i>	-mahi	-ma	—	-ma
	<i>Second</i>	-θa	-ta	-ta	
	<i>Third</i>	-ṇti, -ati -aiṇti	-n, -at, -ārə, -ārəš	-ṇtu, -əṇtu, -aṇtu	-arə, -ərəš

Using the verb *bar* “carry” as an example of a simple thematic present, the expected forms of the present indicative active would be as follows:

(8)	<i>Singular</i>	<i>Dual</i>	<i>Plural</i>
<i>First</i>	barāmi	barāuuahi	barāmahi
<i>Second</i>	barahi	baratō	baraθa
<i>Third</i>	baraiti	baratō	barəṇti

In Gāthic Avestan, the first singular ending *-mi* is found only with athematic stems. Thematic stems such as *bara-* show the archaic ending *-ā* instead.

The injunctive present active of *bar*, which is identical to the imperfect except for the absence of the so-called *augment* (an *a-* prefix), is presented below

(9)	<i>Singular</i>	<i>Dual</i>	<i>Plural</i>
<i>First</i>	barəm	barāuuu	barāma
<i>Second</i>	barō	baratəm	barata
<i>Third</i>	baraṭ	baratəm	barən

The subjunctive present active, which shows some variation as to whether or not it takes primary or secondary endings in some persons, is illustrated in (10):

(10)	<i>Singular</i>	<i>Dual</i>	<i>Plural</i>
<i>First</i>	barāni	—	barāmahi
<i>Second</i>	barāhi	barātō	barāθa
<i>Third</i>	barāiti / barāṭ	barātō	barāṇti / barən

The optative active present of *bar* is as follows (the duals are not attested):

(11)	<i>Singular</i>	<i>Plural</i>
<i>First</i>	—	baraēma
<i>Second</i>	barōiš	baraēta
<i>Third</i>	barōiṭ	baraiiən

Avestan attests a large number of present stem classes and several different types of aorist and perfect. Again, readers are referred to the standard grammars of Avestan for further details.

Avestan, like Sanskrit, presents well-known difficulties in distinguishing between infinitivals and case forms (usually “datives”) of verbal abstracts. The only infinitival form not directly traceable to a nominal case-form origin is the infinitive in *-diiāi/-δiiāi* (compare Sanskrit *-dhyai*), which may be built either to the verbal root (as in Gāthīc *dərəidiiāi* “to hold” < *dar*) or to a present tense stem (as in Young Avestan *vazadiiāi* “to drive” < *vaz*, present tense stem *vaza-*). The bulk of the remaining infinitives of Avestan represent descendants of Proto-Indo-European “directives” in *\*-ay* built to a variety of verbal abstracts, including (i) root nouns (*buiie* “to become” < *\*b<sup>h</sup>uH<sub>2</sub>-ai*; Skt. *bhuvé*), built on the Avestan verbal root *bū*; (ii) *t*-abstracts (*stē* “to be” < *\*H<sub>1</sub>s-t-ai*, built on the root *ah*; and (iii) *s*-abstracts (Gāthīc *srāuuaiiefjhe* “to recite,” as if from *\*kloweyes-ai*, built on the present causative stem *srāuuaiia-* of *sru-* “hear.” Infinitives built on other abstracts (*men*-stems, *wen*-stems, and *ti*-stems, for example) are also attested.

The participle system is quite robustly attested. The present and aorist systems show participles in *-nt/-at-* (added to the tense stem), continuing PIE *\*-ent/-ont/-yt-* participles. In the perfect system, the suffix is *-uuāh-* in the strong cases, and *-uš-* in the weak cases (cf. Skt. *-vāns/-uṣ-*).

The PIE *\*-to*-participle (and its variant in *\*-no-*) is also well attested in Avestan, showing up normally with the zero-grade of the verbal root. It has a “passive” meaning with inherently transitive verbal roots and an active meaning with inherently intransitive ones, thus *kərəta-* “made” < *kar* “make” and *gata-* “gone” from *gam* “go”. Proto-Indo-European *\*-no-* is found, for example, in *pərəna-* “filled” (i.e., “full”) from *\*p<sub>l</sub>h<sub>1</sub>no-* (with a root vocalism analogous to the nasal-infix present), built on the Avestan root *par* “fill” (cf. Skt. *pūrṇa-*).

## 4.5 Numerals

As in other archaic Indo-European languages, the numerals 1 to 4 are inflected for case and number (1 being invariably singular, 2 invariably dual, 3 and 4 invariably plural), while higher numerals up to 19 are not. The Young Avestan numbers 1–10 are as follows:

(12)	1	aēuua-	6	xšuuuāš
	2	duua-	7	hapta
	3	θraii-, tišr- (fem.)	8	ašta
	4	caθβar-/ catur-, cataṅra- (fem.)	9	nauua
	5	paṇca	10	dasa

For the teens, compounds are used, much as in English. The second element of these compounds is *dasa*, thus 12 is *duua.dasa* and 15 is *paṇca.dasa*.

The decades 20 to 90 show a variety of formations and are generally inflected. The Young Avestan decades, with some revealing case forms provided, are presented below (see Hoffmann and Forssman 1996:175):

(13)	20	vīṣas, vīsaiti	60	xšuuuāšti-
	30	θriṣas, θrisatəm; θrisatanəm	70	haptāiti-
	40	caθβarəsātəm	80	aštāiti-
	50	paṇcāsātəm, paṇcāsaṭbīš-ca	90	nauuaiti-

Finally, the numerals 100 and 1,000 are inflected as regular *a*-stems, their stem-forms being: *sata-* 100 and *hazaṅra-* 1,000. A noninflecting numeral for 10,000, *baēuuarə*, is also found.

## 5. SYNTAX

### 5.1 Word order

The study of word order in Avestan reveals a typical archaic Indo-Iranian system, the “basic” order of which can be clearly determined only by a detailed investigation of a number of technical details. Such an investigation has not been fully undertaken for Avestan at this time. It is apparent that the placement of major nominal arguments of the verb – when they are not clitic pronouns or so-called WH-words (i.e., interrogatives and relatives) – is determined by a variety of pragmatic systems (topic, focus), rather than by the role of the argument in the clause (subject, object, etc.). While such systems are sometimes referred to as “free word order,” it would be a mistake to take such a label too literally. Many restrictions on word order do in fact exist, one of the best known of which is *Wackernagel’s Law* (see §5.2). Another, less well-known restriction, concerns the placement of WH-elements, relative pronouns, and complementizers. These elements always occur either sentence-initially, or with a single focused constituent to their left. The latter construction can be seen in the Old Avestan example (*Yasna* 28.1):

- (14) *vaṇhōuš*      *xratūm*      *managhō*      *yā*      *xšnəuuīšā*  
 good-GEN.SG.   insight-ACC.SG.   thought-GEN.SG.   which-INSTR.SG.   you may satisfy  
*gəušcā*      *uruuənəm*  
 cow-GEN.SG.=and   soul-ACC.SG.  
 “With which you may satisfy the insight of good thought and the soul of the cow”

In this example the noun phrase *vaṇhōuš xratūm managhō* “the insight of good thought” has been fronted into sentence-initial position around the relative pronoun (*yā*) as a focusing process. Such constructions are much more rare in Iranian than in Indic, and are virtually limited, within Avestan, to Old Avestan texts. Nevertheless, their widespread occurrence in a wide variety of archaic Indo-European languages allows us to see these Old Avestan examples as a valuable syntactic archaism.

Given the highly restricted placement possibilities for WH-elements, it seems most profitable to posit that such items always occupy the same position in the clause (the so-called *complementizer* slot). They sometimes occur after a single focused constituent as a consequence of the fronting of that constituent for emphasis. Thus, like the Wackernagel’s Law clitics, WH-elements are rigidly fixed in place. Word order is thus obviously not “free” in any meaningful sense.

### 5.2 Clitics

It is necessary to distinguish between three classes of clitic elements in archaic Indo-European languages, including Avestan (Hale 1987a and b). *Sentential clitics* include sentence-level connectives (the conjunction “and,” Avestan *ca*, and the disjunction “or,” Avestan *vā*) and adverbial particles. *Emphatic clitics*, such as Avestan *zī*, indicate focus on the element to which they attach. Finally, *pronominal clitics* are stressless versions of the personal pronouns, usually found in a limited number of case forms. A listing is provided in the discussion of pronominal morphology above. Each of these types of clitics is normally found in so-called second position. The observation that these elements show such a restricted distribution is credited to Bartholomae, who demonstrated the relevant phenomenon using Avestan data in his *Arische Forschungen* (1886). Wackernagel (1892) expanded the data set used by Bartholomae

to include extensive materials from Sanskrit and Greek. From his study the phenomenon of second-position placement of clitics has come to be called Wackernagel's Law.

Each of the clitic types identified above occupies second position for rather different reasons and through distinct mechanisms, such that the definition of "second" turns out to vary somewhat with the class of clitic under discussion (Hale 1987b). Wackernagel's Law is thus the epiphenomenal by-product of a diverse set of processes. Crucial with respect to several of these processes is one overarching principle – that clitic elements, being prosodically deficient through their stresslessness, must not occur at the left edge of a (prosodic) constituent. If the syntax places the clitic in such a position, the prosodic phonology repositions the clitic rightward until it has an appropriately stressed host on its left. We can see the effects of this operation quite clearly in the case of conjunctive clitics like Avestan *ca* "and." Examine the following conjoined sentences, for example (from *Yāšt* 19.51):

- (15) A. *ā.dim haθra haṅəuruuiaiṭ apəm napā auruaṭ.aspō*  
 preverb.him at once grabbed Apam Napat quick-horsed  
 "Quick-horsed Apam Napat grabbed at him at once"  
 B. *taṭ.ca iziieiti apəm napā auruaṭ.aspō*  
 it.and desired Apam Napat quick-horsed  
 "and quick-horsed Apam Napat desired it"

It is clear that the sentence of (15) represents the conjunction of two clauses, the first one being *ā.dim haθra haṅəuruuiaiṭ apəm napā auruaṭ.aspō*, the second *taṭ iziieiti apəm napā auruaṭ.aspō*. The conjunction itself (*ca*) is not part of the content of the second clause, but rather the link between the two (though of course it is related to the second clause, indicating that that clause is conjoined to what precedes). Thus, syntactically, we might identify the basic (i.e., underlying) structure of (15) as being something like the following:

- (16) *ā.dim haθra haṅəuruuiaiṭ apəm napā auruaṭ.aspō* [*ca* [*taṭ iziieiti apəm napā auruaṭ.aspō*]]

The syntactic structure gives rise to the following problem: a clause cannot begin with a clitic, which requires a host on its left, yet the second conjunct in (16) is a clause which begins with the clitic *ca*. The clitic thus is shifted phonologically rightward to the first available position which would give it an appropriate host – in this case to the spot immediately after *taṭ*. The result is the Wackernagel's Law placement of *ca* seen in (15B). This phonological process has been referred to as a "prosodic flip" (Halpern 1992).

In the case of emphatic clitics in Wackernagel's Law position, the facts are somewhat different. Avestan, like other archaic Indo-European languages, provides a number of mechanisms for emphasizing a particular constituent of the sentence. These include adding a particle, such as Avestan *ciṭ* (Skt. *cit*), to the constituent, as in the following example (from *Yāšt* 5.86):

- (17) *θβəm nara-ciṭ yōi taxma jaiðiiānte...*  
 you men-EMPHATIC.PARTICLE REL. bold entreat...  
 "Even bold men entreat you..."

The subject (*nara yōi taxma* "bold men") has been given a degree of emphasis by the addition of the particle *ciṭ* (which takes second position within the subject noun phrase by the same "prosodic flip" processes described above). However, in this same sentence the direct object (*θβəm* "you" [acc. sg.]) has also been focused, in this case syntactically, by being fronted into clause-initial position. The pragmatics of the two processes of focusing can be somewhat

distinct, as can be seen from the English translation of (17). However, since both addition of a particle and syntactic fronting encode emphasis, it is not surprising to find that sometimes both forms of emphasis are placed on the same element, which is then both accompanied by an emphatic particle and fronted into sentence-initial position. Indeed, the pragmatic force of some emphatic particles is such that they are only appropriately used when the element to which they attach is fronted. Such is the case with the Avestan emphatic particle *zī* (cognate with Sanskrit *hi*). An example of the use of this particle can be seen in the recurrent Avestan formula in (18):

- (18) *aēte zī vācō... ahurō mazdā frāmrəuṣ zaraθuštraī*  
 these EMPHATIC.PARTICLE words... Ahura Mazda spoke to Zarathustra  
 “Ahura Mazda spoke *these words* to Zarathustra”

The placement of an emphatic clitic such as *zī* works in the same ways as the placement of *ci* in (17) – such emphatic clitics take second position within the constituent being emphasized. In the case of the sentence in (18), that constituent is *aēte vācō* “these words” and second position within that constituent is the position immediately following *aēte*. When the entire constituent is fronted into clause-initial position, as is appropriate given the type of emphasis indicated by *zī*, it is clear that the emphatic clitic will end up – accidentally, as it were – in second position in the clause. The emphatic clitics thus appear to occupy the same position as the sentential clitics, when in fact somewhat different processes lie behind their placement.

The precise mechanism whereby pronominal clitics come to occupy second position is again somewhat different, though the details are far too complex and theory-dependent to warrant full treatment in the present discussion (see the essays in Halpern and Zwicky 1996 for interesting speculations on this matter). What is of relevance here, however, is that there are only rarely exceptions to second position placement of such pronominals in Avestan. Just as in Sanskrit, where the number of such exceptions steadily decreases between the earliest Rig-vedic hymns and the later Vedic Prose texts, Old Avestan offers a greater – though still small – number of exceptions to Wackernagel’s Law positioning of pronominal clitics than does Young Avestan. The surviving exceptions in Young Avestan clearly represent archaisms and are themselves systematic – they involve cliticization to the verb, rather than the clause. A formulaic and often cited example, involving the first-person singular dative clitic *mē*, is given in (19):

- (19) *auuə āiiaptəm dazdi mē*  
 this boon grant me  
 “Grant me this boon!”

Given these exceptions, Avestan will offer very real contributions to the much needed study of the diachronic development of the processes which underlie Wackernagel’s Law in the archaic Indo-European languages. This domain has already proven to be one of the most productive for the study of Indo-European diachronic syntax.

## 6. LEXICON

One of the more interesting features of the Avestan lexicon is the split in a number of common vocabulary items between *daēuuic* and *ahurian* terms. The term *daēuuā-* has come

to refer to demonic beings in Avestan, in sharp contrast to the use in Sanskrit of the cognate word *deva-* to refer to the gods. It is of some interest to note in this context that the Sanskrit term *asura-*, clearly the cognate of Avestan *ahura-*, which forms part of the name of Ahura Mazdā “the Wise Lord,” who is the god of the Zoroastrians, came, during the Vedic period, to refer to a set of demonic enemies of the gods (*deva-*; for a discussion of these interesting inversions, see Humbach 1991:21ff.). Daēuuic vocabulary items are used when reference is being made to the properties (usually body parts) or actions of manifestations of evil, the ahurian terms being used when referring to manifestations of good (the creations of Ahura Mazdā). Examples include the following:

(20) <i>Daēuuic term</i>	<i>Ahurian term</i>	<i>Gloss</i>
duuar-	i-	“go”
gah-	x <sup>v</sup> ar-	“eat”
aš(i)-	dōiθra-	“eye”
karəna-	gaōša-	“ear”

The fundamental role of dualism as a guiding principle of Zoroastrian thought is clearly evidenced by such lexical splits.

## 7. READING LIST

The most up-to-date and comprehensive grammar of Avestan is that of Hoffmann and Forssman 1996; the earlier grammar of Reichelt 1909, however, contains a detailed discussion of syntax and other matters not handled by later grammars. Beekes 1988 presents an idiosyncratic “interpretation” of the Old Avestan texts and should be used only by those familiar enough with Avestan philology to appreciate fully the implications of such an approach. All work on Avestan before that of Karl Hoffmann tends to misinterpret linguistically relevant phenomena as a superficial matter of orthographic convention. Hoffmann and Narten 1989 represents the most valuable work on the nature of the textual transmission of the Avesta. The only dictionary making any claim to completeness is Bartholomae 1904. Schlerath 1966, in spite of its name, is not a dictionary, but a set of tools for the study of textual repetitions and parallels, including Vedic parallels, as well as a passage-linked bibliography.

Geldner 1886–1896 is the standard edition of the core Avestan corpus, being based on a large number of valuable manuscripts since gone astray. There are a few texts which were excluded from Geldner’s edition, including the *Aogəmadaēcā*, of which JamaspAsa 1982 provides an edition and translation. Translations of the Gāthās include Humbach *et al.* 1991, Insler 1975, and Kellens and Pirart 1988. An excellent overview of the difficulties involved in interpreting the Gāthās can be gained from the detailed treatment by three Iranists of a single hymn, *Yasna* 33, in Schmidt 1985. For Young Avestan texts, Gershevitch 1967 presents one of the Great Yašts (*Yašt* 10). Wolff 1960 presents the translation of the entire corpus which is contained in Bartholomae’s (1904) dictionary – arranged in text order (rather than by keyword). Finally, Reichelt 1911 provides a number of Old and Young Avestan texts, with glossary and notes.

The texts represent the founding documents of Zoroastrianism, and it is therefore of considerable assistance to familiarize oneself with the fundamental doctrines and history of that religion before attempting to tackle them. Boyce 1979 provides a detailed survey of current practices.



## Bibliography

- Bartholomae, C. 1886. *Arische Forschungen*, vol. II. Halle: Niemeyer.
- . 1904. *Altiranisches Wörterbuch*. Strasburg: Trübner. Reprint (as a single volume together with Bartholomae's 1906 *Zum altiranischen Wörterbuch*) 1979. Berlin/New York.
- Beekes, R. 1988. *A Grammar of Gatha-Avestan*. Leiden: Brill.
- Boyce, M. 1979. *Zoroastrians: Their Religious Beliefs and Practices*. London: Routledge and Kegan Paul.
- Geldner, K. 1886–1896. *Avesta: The Sacred Books of the Parsis*. Stuttgart: Kohlhammer.
- Gershevitch, I. 1967. *The Avestan Hymn to Mithra*. Cambridge: Cambridge University Press.
- Hale, M. 1987a. "Wackernagel's Law and the language of the Rigveda." In C. Watkins (ed.), *Studies in Memory of Warren Cowgill*. Berlin: de Gruyter.
- . 1987b. "Studies in the comparative syntax of the oldest Indo-Iranian Languages." Ph.D. dissertation, Harvard University.
- Halpern, A. 1992. "Topics in the placement and morphology of clitics." Ph.D. dissertation, Stanford University.
- Halpern, A. and A. Zwicky (eds.). 1996. *Approaching Second: Second Position Clitics and Related Phenomena*. Stanford: CSLI Publications.
- Hoffmann, K. 1975–1976. *Aufsätze zur Indoiranistik* (2 vols.). Wiesbaden: Dr. Ludwig Reichert.
- Hoffmann, K. and B. Forssman. 1996. *Avestische Laut- und Flexionslehre*. Innsbruck: Innsbrucker Beiträge zur Sprachwissenschaft.
- Hoffmann, K. and J. Narten. 1989. *Der Sasanidische Archetypus*. Wiesbaden: Dr. Ludwig Reichert.
- Humbach, H. (with J. Elfenbein and P. Skjaervø). 1991. *The Gāthās of Zarathustra* (2 vols.). Heidelberg: Carl Winter.
- Insler, S. 1975. *The Gāthās of Zarathustra*. Leiden: Brill.
- JamaspAsa, K. 1982. *Aogomadaēcā: A Zoroastrian Liturgy*. Vienna: Austrian Academy of Sciences.
- Kellens, J. 1984. *Le verbe avestique*. Wiesbaden: Dr. Ludwig Reichert.
- Kellens, J. and E. Pirart. 1988. *Les textes vieil-avestiques*. Wiesbaden: Dr. Ludwig Reichert.
- Kuiper, F. 1978. *On Zarathustra's Language*. Amsterdam: Noord-Hallandsche.
- Mayrhofer, M. 1986ff. *Etymologisches Wörterbuch des Altindoarischen*. Heidelberg: Carl Winter.
- Narten, J. 1986. *Der Yasna Haptanhāiti*. Wiesbaden: Dr. Ludwig Reichert.
- Reichert, H. 1909. *Awestisches Elementarbuch*. Heidelberg: Carl Winter.
- . 1911. *Avestan Reader*. Strasburg: Trübner. Reprint 1968. Berlin: de Gruyter.
- Schlerath, B. 1966. *Awesta-Wörterbuch: Vorarbeiten I–II*. Wiesbaden: Otto Harrassowitz.
- Schmidt, H-P. (with contributions by W. Lentz and S. Insler). 1985. *Form and Meaning of Yasna 33*. New Haven: American Oriental Society.
- Schmitt, R. (ed.). 1989. *Compendium Linguarum Iranicarum*. Wiesbaden: Dr. Ludwig Reichert.
- Tavadia, J. 1956. *Die Mittelpersische Sprach und Literatur der Zarathustrier*. Leipzig: Otto Harrassowitz.
- Wackernagel, J. 1892. "Über ein Gesetz der indogermanischen Wortstellung." *Indogermanische Forschungen* 1:333–436.
- Wolff, F. 1960. *Avesta: Die heiligen Bücher der Parsen, Übersetzt auf der Grundlage von Chr. Bartholomae's Altiranischem Wörterbuch*. Berlin: de Gruyter.

# Pahlavi

MARK HALE

## 1. HISTORICAL AND CULTURAL CONTEXTS

The term *Pahlavi* is used to describe a variety of closely related Middle Iranian languages, including a more archaic variety – the language of early inscriptions – and a more innovative one, so-called *Book Pahlavi*. Given the sparseness of attestation of the earliest varieties of Pahlavi, this sketch will focus on Book Pahlavi, which is in fact quite richly attested. Book Pahlavi is the name generally used to designate that particular variety of Western Middle Iranian used in Zoroastrian writings. Its use in this function covers both a long temporal span (the third century BC to the eighth/ninth century AD) and a broad geographic area. Given this spatio-temporal range, it can be safely assumed that the language showed considerable variation, particularly in the phonological domain. However, the archaizing writing system in which virtually all of our Pahlavi records survive remained quite stable throughout this period and in these diverse regions. Many details of the interpretation of the Pahlavi records thus remain somewhat speculative, our sketches doubtless (inadvertently, of course) combining features of diverse temporal (and perhaps geographical) strata.

The “golden age” of Pahlavi was almost certainly the third to seventh centuries AD, during which time it served as the “standard” language of the Sasanian realms, both for government and for commerce. The Pahlavi corpus includes texts from an impressively broad range of genres, including (but not limited to) royal inscriptions, encyclopedia-like collections (e.g., the *Bundahišn*), texts of religious instruction and worship, legal documents (both laws and instruments), historical texts, and translations.

## 2. WRITING SYSTEM

The Pahlavi writing system is derived from a cursive Aramaic script. Like Aramaic, the script runs right to left and vowels are not written. As noted in Chapter 29 (see §2), the Pahlavi script formed the basis for the Avestan script, which represents an elaboration of the earlier writing system for the purposes of capturing details of hieratic recitation. The Pahlavi writing system includes a number of symbols which carry multiple values (e.g., a single symbol is used to represent *w*, *n*, and *r*; similarly, a single symbol stands for *g*, *d*, and *y*); moreover, many of the combined forms of individual signs are identical in shape to either simplex or combined forms of other signs. When coupled with the lack of vowel symbols, these factors make the orthography highly ambiguous. In practice, remarkably, the many ambiguities rarely impede interpretation.

**Table 30.1 The Pahlavi script**

	Character	Transliteration		
		Segmental Pahlavi	Aramaic-based logogram	"Corrupt"
1	𐭪	'	A, H	—
2	𐭫	b	B	ḡ, ḏ, ȳ
3	𐭬	g, d, y	G, D, Y	ḡ, ḏ, ȳ, <u>ḡ</u> , <u>ḏ</u> , <u>ȳ</u>
4	𐭭	—	E	—
5	𐭮	w, n, r, '	W, N, O, R, '	—
6	𐭯	z	Z	<u>ḡ</u> , <u>ḏ</u> , <u>ȳ</u>
7	𐭰	k	K	—
8	𐭱	(ȳ)	K	—
9	𐭲	l	L	—
10	𐭳	(t)	L	—
11	𐭴	m	M, Q	—
12	𐭵	s	S	—
13	𐭶	p	P	—
14	𐭷	c	C	<u>ḡ</u> , <u>ḏ</u>
15	𐭸	š	Š	—
16	𐭹	t	T	<u>ḡ</u> , <u>ḏ</u>

The symbols used in Pahlavi writing, including both those used in the regular Iranian vocabulary and those used in Aramaic-derived logograms (on which see below), can be found in Table 30.1. On occasion (usually in specific lexical items), a "reduced" form of certain letters is used – MacKenzie (1971) calls these *corrupt*. They are transliterated by the addition of a bar to the letter used to transliterate the nonreduced glyph.

To understand the Pahlavi script, it is necessary to distinguish carefully between a *transliteration* and a *transcription*. Because of the great ambiguities present in the Pahlavi writing system, both of these representations involve an interpretation, rather than simple remapping, of the graphic sequence. In the case of a *transliteration*, the number of symbols in the transliterated string is the same as the number of symbols in the Pahlavi string – but the mapping is not one-to-one. Since, as mentioned above, in some cases an individual Pahlavi character may represent any of a number of segments, the scholar responsible for the transliteration must make a decision about which of these segments is being represented. By contrast, in the case of a *transcription*, the word itself is interpreted phonemically. A standard target "stage" in the historical development of Pahlavi is used to establish this phonemic representation (usually the third century AD) – the vowels are inserted and the archaizing practices of the script are "corrected for" such that the representation reaches the form believed to be an accurate phonemic representation for this historical stage.

An example will make this clearer. The Pahlavi orthographic sequence [𐭬𐭮𐭷] consists of character 3 from Table 30.1, followed by character 5, followed by character 14. Since character 3 can represent either Pahlavi *g*, *d*, or *y* (ignoring its use in so-called "corrupt" characters), and character 5 can represent, in this position in the word, either *w*, *n*, or *r*, a number of distinct *transliterations* are available for this sequence, including, for example, *dwc*, *gnc*, *gwc*, and *ywc*. Each of these transliterations (which will henceforth be represented in square brackets; care should be taken to distinguish between square brackets in this use

and their use for representing allophones in the phonological discussions of §3) can then be interpreted in light of current hypotheses regarding the phonemic system of Pahlavi and what is known about the phonology of the lexeme in question, giving, in these cases, the following *transcriptions* respectively (see MacKenzie 1971): *duz* [dwc] “thief,” *ganj* [gnc] “treasure,” *gōz* [gwc] “walnut,” and *yōz* [ywc] “cheetah.” One should not, therefore, confuse the transliteration of Pahlavi with the usual transliteration of, for example, ancient Greek into Roman characters. The latter involves a fixed mapping between individual characters in the Greek writing system and individual characters in the Roman alphabet (with added diacritics). Such transliterations are purely mechanical and can be handled by individuals (or computers) with no knowledge of Greek. Pahlavi transliteration is a more interpretive procedure, not at all mechanical. Words frequently cannot be transliterated in isolation from their textual context.

The writing system reached a relatively fixed state at what appears to be quite an early date – perhaps by the third century BC, almost certainly by the second. Since many of the surviving texts were composed well after this date, the script may be, and has been, labeled “archaizing,” in other words reflective of a much earlier pronunciation than that in use at the time of writing (as is the case, for example, with the present writing system of English). Moreover, since words written for the first time in the later period are made to fit the “archaic” patterns established for earlier vocabulary, the writing system is frequently “pseudo-archaizing,” giving the words a historical shape which they, in fact, never had.

Pahlavi scribes frequently used Aramaic-based logograms in place of Iranian lexical items, particularly for verbs and a number of “function” words. That these masks were purely a feature of the orthography – standing in for Iranian words much as Hittite scribes used Akkadian and Sumerian graphic sequences to stand in for Hittite words – and not, for example, loanwords is made quite clear by the utter absence of Aramaic loanwords in Middle Iranian texts transmitted in other scripts (including Manichean Middle Persian, Parthian, and Pāzand), as well as by manuscript variants (some of which contain the native Iranian word, others the Aramaic “mask”). These logograms are transliterated in all capital letters, thus the logogrammic representation of Pahlavi *dux̌t* [dwhtʰ] “daughter” is transliterated BRTE (from Aramaic *brt-h*). The logograms representing verb forms are regularly accompanied by the appropriate Pahlavi inflectional ending, giving a mixed orthographic representation: for example, YBLWN-*ytʰ*, with the logogram YBLWN “carry” (from Aramaic *ybl*) and the Pahlavi third singular present indicative ending, the entire written form standing for Pahlavi *barēd* “he carries.”

### 3. PHONOLOGY

#### 3.1 Phonemic inventory

The phonemic inventory for Pahlavi of roughly the third century AD is, in its general outlines, not highly controversial.

##### 3.1.1 Consonants

The consonant phonemes are presented in Table 30.2; of these, phonemic *y* and *ž* are limited to what appear to be learned loanwords, for the most part from Avestan.

The voiced stops and affricates (/b/, /d/, /j/, and /g/) in the table are generally assumed to have had voiced fricative allophones ([β], [ð], [ž], and [ɣ], respectively). The precise

**Table 30.2 Consonantal sounds of Pahlavi**

Manner of articulation	Place of articulation					
	Labial	Dental/ Alveolar	Palato- alveolar	Palatal	Velar	Glottal
<i>Stop</i>						
<i>Voiceless</i>	p	t			k	
<i>Voiced</i>	b	d			g	
<i>Fricative</i>						
<i>Voiceless</i>	f	s	š		x	h
<i>Voiced</i>		z	ž		ɣ	
<i>Affricate</i>						
<i>Voiceless</i>			č			
<i>Voiced</i>			ǰ			
<i>Nasal</i>	m	n				
<i>Liquid</i>		l, r				
<i>Glide</i>	w			y		

conditioning of these allophones clearly varied somewhat over time, it being generally accepted that this spirantization took place in the first instance in postvocalic, preconsonantal position (V\_C): for example, in a word like *abd* [ʔp̄d] “wonderful” (the square brackets here marking a *transliteration*; see § 2), which is taken to have been pronounced with a medial [β]. The voiced fricative allophones were eventually, perhaps by the third or fourth century AD, the usual realization of the voiced stops in all postvocalic positions.

### 3.1.2 Vowels

The vowels include the following:

	FRONT	CENTRAL	BACK
HIGH	i, ī		u, u <sup>-</sup>
MID	(e), ē		(o), ō
LOW		a, ā	

**Figure 30.1** Pahlavi vowels

As indicated by the parentheses in the figure above, the phonemic status of the short mid vowels is not at present entirely clear.

## 3.2 Interesting and significant diachronic developments

The Iranian voiceless unaspirated stops are generally preserved as such in word-initial position (PIIr. is Proto-Indo-Iranian; PIr. is Proto-Iranian):

- (1) pid [AB', p(y)t'] “father” < PIIr. \*pətā (Av. pitā)  
tār [t'l] “darkness” < PIr. \*tantra- (Av. t̥aθra-)  
kōf [kw̥p] “hill” < PIr. \*kaufa- (Av. kaofa-)

as well as in word-internal position when following a voiceless fricative:

- (2) asp [SWSYA, 'sp] "horse" < PIr. \*aspa- < PIIr. \*ācua- (Av. aspa-)  
hušk [hwšk'] "dry" < PIr. \*huška- (Av. huška-)

After *r* and the vowels, the voiceless stops have become voiced:

- (3) kirb [klp] "body, form" < PIIr. \*kṛpa- (Av. kəhrpa-)  
nab [np] "grandson" < PIIr. \*napāt- (Av. napāt-)  
sard [slt'] "cold" < PIIr. \*śarta- (Av. sarəta-)  
pad [PWN] "to, at, in" < PIIr. \*pati (Av. paiti)  
gurg [gwlg] "wolf" < PIIr. \*wrka- (Av. vəhrka-)

Following nasals the voiceless stops have become voiced as well, though it is of some interest to note that the Pahlavi orthography actually writes the postnasal stops as voiced, while, as the examples above show, in its characteristically archaizing style, it writes the postvocalic and post-*r* stops as voiceless: *frazand* [prznd] "child" (cf. Av. *frazaiṇti*-). This has generally been taken as evidence that the voicing after nasals was earlier than that in other positions.

The Indo-Iranian affricate \*č (< PIE \*k, \*kʰ in palatalizing environments) shows similar developments. Preserved in word-initial position (*čarm* [čm] "leather," cf. Av. *carəman*-), it is voiced after nasals (*panj* [pnc] "five," cf. Av. *pañca*). However, it becomes /z/ after vowels and after *r*:

- (4) az [MN, hc] "from" (Av. haca)  
rōz [rwc] "day" < PIr. \*raučah- (Av. raocah-)

The Indo-Iranian affricate \*č̌ (for PIE \*ǩ) is reflected in Pahlavi, as in Avestan, as *s*:

- (5) sad [100] "hundred" < PIIr. \*śata- (Av. sata-)  
suxr [swhl] "red" < \*PIIr. \*śukra- (Av. suxra-)

In general, the Iranian voiced stops are preserved, though as pointed out above they had fricative allophones in a variety of positions:

- (6) baxt [bht'] "fate, fortune" < PIr. \*baxta- (Av. baxta-)  
abr ['bl, 'pl] "cloud" < PIr. \*abra- (Av. aβra-)  
dēn [dyn'] "religion" < PIr. \*dainā- (Av. daēnā-)  
nazd [nzd] "near" < PIr. \*nazda- (cf. Av. nazdišta- 'nearest')  
garm [glm] "warm" < PIr. \*garma- (Av. garəma-)  
mazg [mzg] "marrow, brain" < PIr. \*mazga- (Av. mazga-)

Proto-Iranian \*b is lenited to Pahlavi *w* when originally intervocalic (OP is Old Persian):

- (7) nēw [nyw'] "good, brave" < PIr. \*naiba- (OP naiba-)  
aswār ['swb'l, 'spw'l] "rider" < PIr. \*aspa-bāra- "horse-born"

Additionally, Proto-Iranian \*g is weakened to *w* when intervocalic (even if \*r precedes):

- (8) murw [mwlw'] "bird" < PIr. \*mṛga- (Av. mərəya-)  
mowbēd [mgwpt'] "Mazdean priest" < PIr. \*magu-pati- (cf. OP magu- "priest")

The Indo-Iranian voiced affricates \*j (from palatalized voiced velars) and \*ǰ (from PIE voiced palatal stops) are both realized as *z* in Pahlavi:

- (9) zan [NYŠE, zn'] "woman, wife" < PIIr. \*jani- (Av. jani-)  
 zofr [zwpł] "deep" < PIIr. \*jafra- (Av. jafra-)  
 zand [znd] "district" < PIIr. \*jantu- (Av. zaṇtu-)  
 zrēh [zlyh] "sea" < PIIr. \*jrayah- (Av. zraiiāh-)

The vowel developments from Proto-Iranian to Pahlavi are quite straightforward. Iranian \*a and \*ā are generally preserved as such:

- (10) ast [ʼst'] "bone" < PIIr. \*ast- (Av. ast-)  
 tan [tn'] "body" < PIIr. \*tanū- (Av. tanū-)  
 kām [k'm] "desire" < PIIr. \*kāma- (Av. kāma-)  
 āb [ʼp'] "water" < PIIr. \*āp- (Av. āp-)

As part of certain consonant cluster simplifications Proto-Iranian \*a shows compensatory lengthening to Pahlavi ā:

- (11) hazār [hc'l] "thousand" < PIIr. \*hazahra- (Av. hazaṇra-)  
 sāl [s'l] "year" < PIIr. \*sard- (Av. sarəḏ-)  
 māl- [m'l-] "rub, sweep" < PIIr. \*marz- (Av. marəz-)

A number of sequences involving intervocalic glides lose their glides and show vowel coalescence such that Pahlavi ā results:

- (12) \*āwa: srāy- [sl'd-] "sing" < PIIr. \*srāwaya- (causative of sru "hear")  
 \*awā: bād [b't'] "may be" < \*bawāti (Av. bauuāiti)  
 Pretonic \*āwi: āškārag [ʼšk'lk'] "obvious, evident" < PIIr. \*āwiš-kāra-ka-  
 "made manifest"

The high vowels as well are generally preserved as such, both long and short:

- (13) \*i: tigr [tgl] "arrow" < PIIr. \*tigra- (Av. tiṅra-)  
 \*i: wīr [wyl] "hero" < PIIr. \*wīra- (Av. vīra-)  
 \*u: pus [pws] "son" < PIIr. \*puθra- (Av. Puθra-)  
 \*ū: stūn [stwn'] "column" < PIIr. \*stūna- (Av. stūna-)

The Proto-Iranian diphthong \*ai gives Pahlavi ē:

- (14) ēw [l] "one" < PIIr. \*aiwa- (Av. aēuua-)  
 wēn- [wyn-] "see" < PIIr. \*waina- (Av. vaēna-)

In addition, several sequences involving the palatal glide result in Pahlavi ē as well:

- (15) \*aya: sē [3] "three" < PIIr. \*θraya-  
 -ēḏ 3rd sg. pres. tense verbal ending < PIIr. \*-ayati  
 \*ahya: kē [MNW] "who" < PIIr. \*kahya "whose" (O.Av. kahiiā, Y.Av. kahe)  
 \*arya: ēr [ʼyl] "noble" < PIIr. arya- (Av. airiia-)

In a strongly parallel manner, the Proto-Iranian diphthong \*au and the sequence \*awa yield Pahlavi ō:

- (16) gōš [gwš] "ear" < PIIr. \*gauša- (Av. gaoša-)  
 nōg [nwk'] "new" < PIIr. \*nawa-ka- (cf. Av. nauua- "new")

Finally, Proto-Iranian \*r is reflected in the majority of cases by Pahlavi ur (though occasionally Pahlavi ir is found):



- (17) purs- [pwrs-] “ask” < Plr. \*pṛsa- (Av. pərəsa-)  
 wazurg [wzlǵ] “big” < Plr. \*wazrka- (OP vazarka-)  
 buland [bwlnd] “high” < Plr. \*brzant- (Av. bərəzant-)

It should be apparent from the above examples that Pahlavi has suffered serious “erosion from the right” – the final syllables (or at least their codas) having more or less uniformly disappeared. The set of developments which gave rise to this loss – a function of the development of strong dynamic stress – are clearly implicated in the massive morphological (and syntactic) restructuring which Pahlavi has undergone.

## 4. MORPHOLOGY

### 4.1 Morphological type

Not surprisingly, given the extensive erosion from the right that words have suffered in the phonological history of Middle Iranian, Pahlavi is much more isolating than its Indo-European inflectional forebears, virtually nothing but traces remaining of the elaborate nominal (case, gender, number) and verbal (tense, aspect, mood) morphological systems of Proto-Indo-Iranian.

### 4.2 Nominal morphology

The nominal inflection system of Old Iranian, with numerous distinct stem-types each showing a somewhat different pattern of inflection, was clearly eliminated in the early stages of Middle Iranian. It is generally held that this system was reduced to two cases, usually called the *casus rectus* (deriving from the earlier nominative) and the *casus obliquus* (from the earlier genitive), and two numbers (a singular and plural, the dual having vanished). The inflection of nominals, with generalization of the *a*-stem endings to the other stem-classes of Iranian, thus took the following form:

	Singular	Plural
(18) <i>Casus rectus</i>	asp < Plr. *aspah (nom. sg.)	asp < Plr. *aspāh (nom. pl.)
<i>Casus obliquus</i>	aspē < Plr. *aspahya (gen. sg.)	aspān < Plr. *aspānām (gen. pl.)

The distinctive oblique case-marker in the singular was apparently lost early, with a merger of the rectus/obliquus contrast in the singular. The *-ān* plural subsequently came to be used in both cases in the plural, though both forms (those with *-ø* and those with *-ān*) are found in both cases in most Pahlavi texts. Various factors appear to favor explicit marking of plurality, including animacy (animate nouns are more likely to have explicitly marked plurals) and whether or not the verb form in the clause unambiguously indicates the number of its subject.

In general, the surviving nominal stem-form in Pahlavi looks like the old oblique without its case ending in *-ē*. In a few cases, however, mostly involving Old Iranian *r*-stem kinship terms, Pahlavi shows both a reflex of the earlier rectus and a reflex of the earlier oblique. For example, the word for “father” is attested both in a form which must have originally been the rectus case (*pid* [AB', p(y)ɾ'] < Plr. \*pitā, nom. sg.) and in a form which must have originally been the oblique (*pidar* [ABYɾl, pyɾl], ultimately – after morphological elimination of the oblique *-ē* case ending – from Plr. \*pitarahya, genitive singular, with the ending generalized from the *a*-stems). The stem-doublets do not appear to be distributed systematically in

the bulk of the surviving Book Pahlavi texts, though detailed philological work in this area remains a desideratum. Additional examples include the following:

- (19) mād [AM] next to mādar [AMYtl, m'tl] “mother”  
 brād [AH, bl't'] next to brādar [bl'tl] “brother”  
 xwah [AHTE] next to xwahar [AHTEL, hw'hl] “sister”  
 duxt [BRTE, dwht'] next to duxtar [dwhtl] “daughter”

In other cases, it is clear that the old nominative case form (rather than the old oblique) is the one that survives in Pahlavi. This usually involves what were originally neuter *r/n*-stems. Examples include the following:

- (20) A. kišwar [kyšwl] “region” (Av. nom. sg. karšuuarə, weak cases made to the stem karšuuān-)  
 B. zafar [zpl] “mouth” (daevic only) (Av. nom. sg. zafarə, oblique not directly attested, but doubtless an *n*-stem)  
 C. ʃagar [ykl] “liver” (Av. nom. sg. yākarə, oblique also not directly attested)

Since the final *-r* of the Pahlavi lexical items was originally found only in the nominative – the obliques having *-n* in its stead (in the well-known Indo-European pattern) – it must be the *casus rectus* which has been generalized in these lexemes.

### 4.3 Pronominal morphology

As was originally the case with nominals, the Pahlavi pronominal system showed a reduction of the rich case system of Proto-Iranian to a simple *casus rectus*: *casus obliquus* system. This system was further reduced, except in the first-person singular, to a single form. Pahlavi distinguishes between tonic and enclitic pronouns in first (oblique), second, and third person:

(21)			<i>Tonic</i>	<i>Enclitic</i>
	<i>Singular</i>	1. <i>Rectus</i>	az (?) [ANE]	—
		1. <i>Oblique</i>	man [L]	-m
		2	tō [LK]	-t
		3	ōy [OLE]	-š
	<i>Plural</i>	1.	amā [LNE]	-mān
		2.	ašmā [LKWM]	-tān
		3.	awēšān [OLEš'n]	-šān

The reading of the first singular rectus form is not entirely clear, some scholars favoring *az*, others *an*. Note that the enclitic plural forms appear to be derived from the enclitic singulars (themselves corresponding more or less directly to Av. *mē*, *tē*, and the RUKI-induced [see Ch. 29, §3.4.4] *šē* realization of the third singular *hē*) by the addition of the nominal pluralizer *-ān*.

### 4.4 Verbal morphology

Pahlavi verbs are inflected for person, number, tense, and mood, though most of these categories are significantly reduced relative to the rich systems present in Old Iranian. With respect to number, there is no trace of the Old Iranian dual endings. The tense system has been reduced to a simple present versus preterite (past) stem (the latter built on the PIE

\**to*-particle), with the loss of the imperfect, the aorist, and the perfect of Old Iranian. In Book Pahlavi, the moods (subjunctive and optative) are attested in only limited distribution, being wholly unattested for certain persons.

In general, the inflection of the present indicative is to be traced back to Old Iranian \**aya*-presents, the endings being characterized by the vowel *ē* < PIr. \**-aya-*. There is some debate regarding the etymological status of endings which have instead the vowels *a* or *o* – some scholars (e.g., Nyberg) believe these reflect Old Iranian simple thematic verbs, others (e.g., Tedesco, Sundermann) that they represent innovations within the history of Pahlavi.

The endings of the present indicative are as follows:

(22)	<i>Singular</i>	<i>Plural</i>
<i>First</i>	-ēm [-ym], -am [-m], -om [-wm]	-ēm [-ym], -am [-m], -om [-wm]
<i>Second</i>	-ē(h) [-yḏ, -yh]	-ēd [-ytʼ]
<i>Third</i>	-ēd [-ytʼ], -at [-tʼ]	-ēnd [-ynd], -and [-nd]

The inflection of the verb *būdan* [YHWWNtnʼ, bwtnʼ] “to be” shows a mixture of forms, some coming from the cognate of Avestan *bauuaiti* “is, becomes” (Pahlavi present stem *baw-*), some corresponding to Avestan *asti/hənti* “is”/“are” (Pahlavi present stem *h-*). In many instances, forms from both stems are found:

**Table 30.3 Pahlavi verb inflection**

	STEM	
	<i>baw-</i>	<i>h-</i>
Indicative		
<i>Singular</i>	1. <i>bawam</i> [YHWWNm]	<i>hom</i> [HWEwm], <i>ham</i> [HWE m], <i>hēm</i> [HWEym]
	2. <i>bawē</i> [YHWWNyḏ, -yh]	<i>hē</i> [HWEʼybḏ, -yh]
	3. <i>bawēd</i> [YHWWNy tʼ], <i>bēd</i> [by tʼ]	<i>hast</i> [ʼYTʼ]
<i>Plural</i>	1. <i>bawēm</i> [YHWWNym]	<i>hēm</i> [HWEʼym]
	2. <i>bawēd</i> [YHWWNy tʼ]	<i>hēd</i> [HWEʼytʼ]
	3. <i>bawēnd</i> [YHWWNd]	<i>hēnd</i> [HWEnd]
Subjunctive		
<i>Singular</i>	2. <i>bawāi</i> [bwpʼy]	<i>hāi</i> [HWEʼʼy], <i>hāh</i> [HWEʼʼh], <i>hā</i> [HWEʼʼ]
	3. <i>bawād</i> [YHWWN(ʼ) tʼ], <i>bād</i> [bʼ tʼ]	<i>hād</i> [HWEʼʼ tʼ], <i>hat</i> [HWE t]
<i>Plural</i>	2. <i>bawād</i> [YHWWNʼ tʼ]	
	3. —	<i>hānd</i> [HWEnd]
Optative		
<i>Singular</i>	3. —	<i>hē</i> [HWEʼyḏ, HWE d]
Infinitive	<i>būdan</i> [YHWWNtnʼ, bwtnʼ]	
Past	<i>būd</i> [YHWWN tʼ, bwtʼ]	

Although there is a basic uniformity in the inflectional endings found on Pahlavi verbs, something of the richness of the Proto-Indo-Iranian verbal system can still be seen in the great diversity of formal relationships holding between the present stem and the infinitive. The Pahlavi present stem often reflects aspects of the Proto-Indo-Iranian present class of which the verbal root was a member. For example, Avestan has as the general present

stem of the verbal root  $x^v ap$ - “sleep” (PIE  $*slep$ -) the reflex of a Proto-Indo-European  $*-skē/o$ - “inchoative”: namely,  $x^v afsa$ -. This root makes a normal past participle of the Proto-Indo-European  $*-to$ -type: thus,  $x^v apta$ -. These forms correspond directly to the Pahlavi present/infinitive stem pair  $xufs$ - :  $xuftan$  [ $hwps$ - :  $HLMWNtn'$ ,  $hwptn'$ ], where the  $s$  of the present stem can only be motivated by a historical account. Analogies between the infinitival stem (generally identical to the past tense form) and the present stem are common: corresponding to the  $s$ -inchoative present stem  $purs$ - [ $pwrs$ -] “ask” (Av.  $pərəsa$ - < PIE  $*pr̥k-skē/o$ -) is the infinitive  $pursīdan$  [ $pwrsytn'$ ], clearly analogically remade, since it contains the derivational affix  $-s$ , originally limited to the present system. In other cases, the present stem appears to have been analogically remade based on the infinitive (or past) stem.

Traces of a wide range of the present-tense formation categories of Old Indo-Iranian can be seen in this way in Pahlavi. Some examples include the following:

1. *Zero-grade thematic presents*: for example,  $kuš$ - :  $kuštan$  [ $kwš$ - :  $NKSWNtn'$ ,  $kwštn'$ ] “kill” (cf. the Avestan present stem  $kuša$ -)
2. *Full-grade thematic presents*: for example,  $yaz$ - :  $yaštan$  [ $yc$ - :  $YDBHWNtn'$ ,  $yštn'$ ] “worship” (cf. Av. present stem  $yaza$ - : past participle  $yašta$ -)
3. *Lengthened-grade causatives*: for example,  $tāz$ - :  $tāxtan$  [ $t'c$ - :  $t'htn'$ ] “cause to run” (as if from Old Iranian  $*tācaya$ -)
4. *Reduplicated presents*: for example,  $dah$ - :  $dādan$  [ $dh$ - :  $YHBWNtn'$ ,  $d'tn'$ ] “give, create” (cf. Av. present stem  $daḍā$ - : past participle  $dāta$ -)
5. *Nasal-infix presents*: for example,  $hanj$ - :  $hixtan$  [ $hnc$ - :  $hyhtn'$ ] “draw [water]” (cf. Av. present stem  $hiñca$ - : past participle  $hixta$ -)

In some cases, the present stem and infinitive came to diverge considerably, as in  $ōft$ - :  $ōbastan$  [ $'wpt$ - :  $NPLWNstn'$ ,  $'wptn'$ ] “fall”

The present stem here reflects an Iranian  $*awa-pta$ - (cf. the Avestan present stem-forms in  $pta$ - for the root  $pat$ - “fall”), with zero-grade of the root. By contrast, the infinitive (built like the  $to$ -participle, as usual) shows the preconsonantal full-grade of the root, coming thus from  $*awa-pasta$ - ( $*pasta$ - being from  $*pat-ta$ -, with the usual Iranian treatment of Proto-Indo-European  $*tt$ -clusters).

In terms of synchronic morphology – rather than the traces of Iranian morphology which have become lexicalized by the time of Pahlavi – there is a productive causative suffix  $-ēn$ -, no doubt originally denominative, seemingly from  $*-anya$ -. It is added to the present stem: for example,  $abzāy$ - :  $abzūdan$  [ $'p̄z'd$ - :  $'p̄zwtn'$ ] “increase, grow” →  $abzāyēnēd$  [ $'p̄z'dynytn'$ ] “makes increase, makes grow.”

The preterite (or past) tense, as indicated in passing above, was built from the Proto-Indo-European  $*to$ -participle, being used in periphrastic constructions with auxiliary verbs such as  $budan$  “be/become,”  $h$ - “be,” and  $ēstādan$  “stand.” The  $*to$ -participle had the property of being *active* in interpretation with intransitive verbs, but *passive* with transitive verbs. Using the intransitive verb  $šudan$  “go” and the transitive verb  $kardan$  “do, make” as examples, the interpretation of these various periphrases was as follows (cf. Brunner 1977:213ff.):

(23)	Intransitive verb	Transitive verb
Simple preterite	šud h- “went”	kard h- “was made”
Perfect	šud ēst- “has gone”	kard ēst- “has been made”
Pluperfect	šud būd h- “had gone”	kard būd h- “had been made”

The forms of the verb  $h$ - “to be” were optional in this construction and not normally expressed.

Given this method of constructing preterites, the issue of how to express the past tense of a simple *active* transitive verb arises. That is, what corresponds in the past tense to the present tense clause *awēšān kušēm* “I slay them”? It seems clear that in Pahlavi the relevant construction was *man kušt hēnd* “they were slain by me,” though it remains somewhat unclear whether the meaning of this construction was, for the Pahlavi speaker, passive (as I have expressed in my translation) or active (i.e., “I slew them”). Hand-in-hand with this lack of clarity as to the real interpretation of this construction is a lack of clarity as to its syntactic structure (is it an ergative-like structure or a passive one?). It seems likely that there is no one resolution to these matters for “Pahlavi” in all its temporal, geographical, and stylistic diversity.

## 4.5 Numerals

The Pahlavi numerals 1 to 10 are presented below:

(24) 1 ēk [’ywk’]	6 šaš [ŠTA]
2 dō [TLYN’]	7 haft [hpt’]
3 sē [TLTA]	8 hašt [hšt’]
4 čahār [ALBA, ch’l]	9 nō [TŠA]
5 panj [pnc]	10 dah [ASLYA]

The decads generally reflect the diverse formations found in Proto-Indo-Iranian (and Proto-Indo-European) for this function (many of the higher numbers are represented only by digits, not spelled out, in the texts):

(25) 20 wīst [wyst’]	70 haftād
30 sih	80 haštād
40 čehel	90 nawad
50 panjā [pnc’h]	100 sad
60 šast	1,000 hazār [hc’l]

The attested ordinals include:

(26) 1st fradom [AWLA, pltwm]	6th šašom [6wm]
2nd dudīgar [dtykl]	7th haftom [7wm]
3rd sidīg(ar) [styk’, stykl]	8th haštom [8wm]
4th čahārom [ch’lwm]	9th nohom [9wm, nhwm]
5th panjom [5wm]	10th dahom [dhwm]

## 5. SYNTAX

### 5.1 Word order

The loss of the elaborate case system of Old Iranian has, not surprisingly, led to a considerably more restrictive use of word order in Pahlavi. The “normal” order – in the sense both that it is statistically most common and that it involves no particular pragmatic focus on any element of the sentence – is *Subject–Object–Verb*. The possibility of fronting elements from within the verbal complex to sentence-initial position for emphasis of some type (not at this time explored in any great detail) certainly exists.

A regular exception to this ordering statement arises in the case of arguments expressed by the enclitic personal pronouns presented above (see §4.3). These elements tend strongly to

occur in *Wackernagel's Law position* (i.e., after the first constituent of their clause) regardless of their syntactic function. When used in subject function, they are often redundant, as in (27):

- (27) guft-aš      Ohrmazd    ō    Spitāmān    Zarduxšt  
       gwptš      'whrmzd    'L    spyt'm'n    zltwhšt  
       spoke-he    Ohrmazd    to    Spitāmān    Zarduxšt  
       "Ohrmazd said to Spitāmān Zarduxšt"

This example also shows the fronting of the verb to sentence-initial position. The "unmarked" order of such a sentence, without the redundant enclitic pronominal *-aš*, would be as follows:

- (28) Ohrmazd    ō    Spitāmān    Zarduxšt    guft  
       'whrmzd    'L    spyt'm'n    zltwhšt    gwpt  
       Ohrmazd    to    Spitāmān    Zarduxšt    spoke  
       "Ohrmazd said to Spitāmān Zarduxšt"

No deep understanding of the pragmatic differences indicated by the order in (27) versus that in (28) has as yet been attained, nor is it generally understood why, in the case of this particular clause (in which the creator makes a pronouncement of religious significance to the prophet Zoroaster), the "marked" order in (27) is in fact regular.

## 6. READING LIST

The standard dictionary for Pahlavi, which includes the necessary tools to allow one to actually decipher the script as well, is that of MacKenzie 1971. Those attempting to read Pahlavi texts are also likely to find the glossaries in Nyberg 1964 and Boyce 1977 of considerable assistance. For inscriptions, the necessary lexical material can be found in Gignoux 1972.

There is no comprehensive grammatical description of Pahlavi, though considerable historical and descriptive information can still be usefully gleaned from Salemann 1895–1901. Beyond this, there is a grammatical sketch in Nyberg 1964, an extensive discussion of a variety of syntactic issues in Brunner 1977, and scattered observations on grammar in the various "survey" articles (Henning 1958, Sundermann 1989).

Extensive text editions from a wide range of genres with, in some cases, quite detailed notes can be found in Nyberg 1964. Boyce 1975 presents a wealth of Manichaean Middle Persian (and Parthian) texts, the basic grammatical structure of which differs only in matters of fine detail from that presented here. For specific texts see, for example, Gignoux 1984 and Humbach and Skjaervø 1978–1983. Tavadia 1956 and Boyce 1968 present useful surveys of just what types of texts are found in the corpus.

Finally, Dresden 1966 presents one of a number of Pahlavi manuscripts – this one of the very significant text known as the *Denkart* – which are available in facsimile. Such facsimile editions are more critical than usual in this field, where disputes about the details of the interpretation of what is written on the page are inevitable.

## Bibliography

- Bartholomae, C. 1904. *Altiranisches Wörterbuch*. Strasburg: Trübner. Reprint (as a single volume together with Bartholomae's 1906 *Zum altiranischen Wörterbuch*) 1979. Berlin/New York.  
 Boyce, M. 1958. "Middle Persian literature." In *Iranistik*, 2 Literatur, pp. 67–76. Handbuch der Orientalistik I. Leiden: Brill.

- \_\_\_\_\_. 1975. *A Reader in Manichaean Middle Persian and Parthian: Texts with Notes*. Leiden: Brill.
- \_\_\_\_\_. 1977. *A Word-List of Manichaean Middle Persian and Parthian*. Leiden: Brill.
- Brunner, C. 1977. *A Syntax of Western Middle Iranian*. Delmar, NY: Caravan Books.
- Dresden, M. 1966. *Denkart: A Pahlavi Text*. Wiesbaden: Otto Harrassowitz.
- Gignoux, P. 1972. *Glossaire des inscriptions pehlevies et parthes*. *Corpus Inscriptionum Iranicarum* Suppl. Series, vol. I. London: Lund Humphries.
- \_\_\_\_\_. 1984. *Le livre d'Arda Viraz: translittération, transcription et traduction du texte pehlevi*. Paris: Editions Recherche sur les Civilisations.
- Henning, W. 1933. "Das Verbum des Mittelpersischen der Turfanfragmente." *Zeitschrift für Indologie und Iranistik* 9:158–253.
- \_\_\_\_\_. 1958. "Mitteliranisch." In *Iranistik*, 1 Linguistik, pp. 20–130. *Handbuch der Orientalistik* I. Leiden: Brill.
- Humbach, H. and O. Skjaervø. 1978–. *The Sassanian Inscription of Paikuli*. Wiesbaden: Dr. Ludwig Reichert.
- MacKenzie, D. 1971. *A Concise Pahlavi Dictionary*. London: Oxford University Press.
- Nyberg, H. 1964. *A Manual of Pahlavi* (2 vols.). Wiesbaden: Otto Harrassowitz.
- Salemann, C. 1895–1901. "Mittelpersisch." In W. Geiger and E. Kuhn (eds.), *Grundriss der iranischen Philologie*, vol. I, section 1, pp. 249–332. Strasburg: Trübner.
- Sundermann, W. 1989. "Mittelpersisch." In R. Schmitt (ed.), *Compendium Linguarum Iranicarum*. Wiesbaden: Dr. Ludwig Reichert.
- Tavadia, J. 1956. *Die mittelpersische Sprache und Literatur der Zarathustrier*. Leipzig: Otto Harrassowitz.



# Phrygian

CLAUDE BRIKHE

## 1. HISTORICAL AND CULTURAL CONTEXTS

When the Phrygians emerged as a political entity in the middle of the eighth century BC, they occupied the central part of Anatolia. They were often settled on sites which were previously occupied by the Hittites (Gordion, the city of Midas, Boğazköy, etc.).

Their arrival has long been associated with the collapse of the Hittite Empire (around 1200 BC). However, archeological findings more often than not reveal a hiatus on the relevant sites between the last Hittite level, generally punctuated by a catastrophe involving fire, and the first Phrygian level. In Hattusas/Boğazköy, the capital of the empire, following a fire on the citadel, Büyükkale (c. 1180 BC), about four centuries of abandonment seem to have preceded the Phrygian occupation (see Gunter 1991:106). In Gordion, however, where no break is observed between the end of the Bronze Age and the beginning of the Iron Age, the first traces of a “European” presence are found from the very start of the first millennium (see Sams 1994).

Even so, Tiglathpilesar I of Assur (1112–1072 BC) speaks of his battles against the Mushkis, on the upper Tigris (cf. *Der kleine Pauly*, vol. IV, col. 822, s.v. *Phryger* [G. Neumann]). And the Assyrian chronicles (Sargon II) mention the presence of Mita the Mushki king, between 717 and 709 BC, in southeast Anatolia: the monarch’s name, which corresponds to that of the Phrygian sovereign at the time (Midas II), and the presence in Tyana of Paleo-Phrygian texts encourage us to consider *Mushkis* as the designation for the Phrygians (generalized from one of their constituents perhaps?) by the Easterners. In order to reconcile these conclusions and the archeological evidence, perhaps one should assume a somewhat longer time period between the first appearances of the Phrygians in Asia Minor and their settlement (see, with the bibliography, Brixhe 1995, §3.3, and 1991:44–45).

The Phrygians came from the Balkans – certain mythical accounts from Macedonia or Western Thrace recall their stay in that region (cf. the Gardens of Midas, near Mount Bermion, in Herodotus 8.138). The Greek historian Herodotus (7.73) notes their coexistence with the Macedonians, and it seems that during their migration, they left behind, from Pelagonia to the areas surrounding the Athos, several tribes known by various names (Brugoi, Briges, Brukes, etc.) which immediately evoke the designation *Phrygian*.

In their language, in fact, the Proto-Indo-European aspirates display voiced reflexes, and *\*b<sup>h</sup>* had become *b*. The name which we give them, Φρύγες, has been transmitted to us by the Greeks, who developed *p<sup>h</sup>* from PIE *\*b<sup>h</sup>*. The Phrygians must have called themselves something like *Bruges* (cf. the *Briges* of Herodotus 7.73).

The Phrygians have left two large corpora of written documents, widely spaced in time – the earlier called Paleo-Phrygian and the later Neo-Phrygian. On evidence for a distinct Middle Phrygian form, see §1.2.

## 1.1 Paleo-Phrygian

The following Paleo-Phrygian inscriptions are cited from Brixhe and Lejeune 1984 (completed by Brixhe 1991 for T-03), from which work I have adopted the symbols assigned to epigraphic regions (M, W, B, etc., see below). The peripheral texts of §1.1.1 are identified according to their original site.

The Paleo-Phrygian documents, collected in 1984 by Claude Brixhe and Michel Lejeune, were spread across a vast area: (i) to the west of Great Phrygia (W in Brixhe and Lejeune 1984), with Midas City (M); (ii) in Bithynia (B) (iii) in Galatia, Gordion (G) and environs (C); (iv) in Pteria (P), a region of Boğazköy; (v) in Cappadocia, on the site of ancient Tyana (T). Two objects are of unknown origin (Dd).

Since 1984 have been discovered: (i) a seal of unknown origin (see Masson 1987); (ii) a graffito on a silver vessel (sixth century) found in a tumulus about 20 kilometers to the west of Uşak (see Brixhe 1989–1990); (iii) ten graffiti on silver vessels (late seventh to early sixth century) found in a tumulus in the plain of Elmalı (Lycia; see Varinlioğlu 1992); (iv) a “Spinnwirtel” near Thyateira (see Dinç and Innocente 1999) (v) a seal and some graffiti (sixth century?) recovered in Eskişehir-Dorylaion (see Darga 1993).

If one adds to this collection about 80 or so unpublished items from Gordion and Dorylaion, we have about 340 documents, of which several appearing on the same stone may belong to a single text. The Paleo-Phrygian inscriptions are distributed very unequally, with about 250 (including unpublished ones) being furnished by Gordion alone.

To the extent that one can judge given their contexts and our understanding of the language, the inscriptions are of various sorts: cult texts; a royal affirmation of suzerainty (?T); an apotropaic formula (G-02); seals; marks of ownership (graffiti on pottery); and perhaps notations of gift exchange.

So far as the archeology and historical cross-references (for T, for instance) allow one to judge, Paleo-Phrygian texts date from the beginning of the eighth century (e.g., G-104, G-237, G-249) down to the period immediately preceding the Macedonian conquest. After the sixth century, there probably exist only graffiti on pottery (see Brixhe 1993:21); for the high chronology of the first documents, see now Manning *et al.* 2001 and Voigt *et al.* 2001.

### 1.1.1 Additional evidence

The distribution of Paleo-Phrygian documentation is further extended if one takes into account several peripheral texts that belong to the same time frame (perhaps representing dialectal Phrygian):

1. A sinistroverse document, which could no doubt be attributed to the fourth century BC, found long ago at Üyücek to the south of Tavşanlı (in the far west of Phrygia) but presently lost (see Cox and Cameron 1932).
2. A bilingual sinistroverse inscription consisting of thirteen epichoric lines, preceded and followed by, respectively, five and two Greek lines discovered at Vezirhan in Bithynia (to the north of Bilecik) and belonging to the fifth–fourth centuries (see Neumann 1997).
3. A set of texts provided by the excavations in Daskyleion (the westernmost of the relevant sites, Mysia) and comprising: (i) two steles (sixth/fifth centuries; one heavily damaged) and (ii) eight graffiti on pottery (second half of the sixth century–first half

of fourth century; see Brixhe 1996; Gusmani and G. Polat 1999; Gusmani and Y. Polat 1999).

The oldest documents from Gordion (beginning of the eighth century) come from the close of the most robust period of Phrygian history. In the eighth century BC, the Phrygian Empire was the dominant power in Asia Minor, and excavations from Gordion reveal a lifestyle and architecture resembling that seen among its Greek neighbors to the west, as well as certain correspondences with the society portrayed in the Homeric poems (see De Vries 1980:42). One can see Phrygians intervening in Cappadocia (cf. the relations between Sargon II and Mita, king of Mushkis noted above), and neither Lydia nor Caria seems to have acted as a barrier between the Phrygians and the Greeks – tradition tells us about cultural and matrimonial relations with the Hellenic world in ancient times (see Brixhe 1995, §3.2).

According to tradition, the people called the Cimmerians arrive at the beginning of the seventh century BC. What then becomes of Midas's empire? It is possible that it might have been divided up into small principalities (see Brixhe 1991:45); but the Phrygians, as the spread of their monuments shows, continue to occupy a vast space and, no doubt, to exert a cultural influence over Anatolia. Towards the end of the seventh century, they fall under the yoke of first the Lydians, and later the Persians, until the Macedonian conquest.

During these dark periods, we know very little about the Phrygians. They undoubtedly continued to have relations with the Greeks to the west, through Lydia and Caria (cf. above); but to the northwest, they came into direct contact with the Ionian colonies of the Propontide. These relations, probably more intense than they seem, no doubt explain the earliness of the lexical (ἡ σορός) and anthroponymic borrowings revealed by the Dokimeion document, which will be mentioned.

## 1.2 Middle Phrygian?

The last Paleo-Phrygian graffiti from Gordion in all likelihood predate the time of Alexander. However, Th. Drew-Bear discovered a funerary stele in the area of Dokimeion which likely belongs to the very end of the fourth century BC; it bears a long inscription of eight dextroverse lines, apparently recording that a certain Nikostratos had the monument erected for one Kleumakhos. Not being written in the local alphabet, but already with the Classical Attic Greek alphabet, could it perhaps represent a Middle-Phrygian stage, intermediary to Paleo- and Neo-Phrygian? See Brixhe 1993:326–327, and 1994:167.

## 1.3 Neo-Phrygian

The Phrygians remain silent for the next several centuries – probably until the beginning of the Christian era – when their language then reappears exclusively in a funerary context. Currently, 113 documents are known (for details, see Brixhe 1994, §§1/2/3), about half of which are Greek–Phrygian bilingual and typically consist of an epitaph coupled with an imprecation against possible depredators.

The Neo-Phrygian speech area (delimited by Konya, the northern tip of Tuz Gölü, Eskşehir, Kütahya, Dinar, with the highest density on the western border and to the north/northwest of the northern tip of Tuz Gölü) was considerably smaller than that of Paleo-Phrygian. Besides the influx of Balkanic peoples which seems not to have ceased until the Roman era, but which could have affected only the northwestern part of the domain, two events had a profound repercussion on the linguistic situation of the region: the Macedonian invasion and, soon after 280 BC, the settlement of the Celtic-speaking Galatians in the northeast.

The Phrygian elite (like the Galatian) was quickly Hellenized linguistically; the Phrygian tongue was devalued and found refuge only in the countryside, in the weakly urbanized perimeter defined above. Although in the Paleo-Phrygian era it had acceded to all of the written registers, public and private, sacred and profane, in the Neo-Phrygian period the language was confined to the sacred domain, having become the language of a colonized people.

The ancient sources last speak of the Phrygian language in the fifth century AD (see Friedrich 1941, col. 868–869); but it is quite possible that it was not actually eliminated until the Arab incursions in the seventh century (Brixhe 1987:11).

## 1.4 Greek evidence

Apart from the documents which underpin the discussions of §§1.1–1.3, there are two additional sources for the Phrygian language:

1. A certain number of Greek glosses survive, especially in the work of the lexicographer Hesychius, but also in literary sources; see Brixhe 1982:243–244; 1990:93; Haas 1966:157–172, among others.
2. Greek inscriptions of Phrygia provide (i) various terms of Phrygian origin (e.g., τὸ βέννος “association of the believers in a god”; (ii) anthroponyms and toponyms (see Brixhe 1983:129; 1987:110–116, 157–158; 1993:342).

## 1.5 Phrygian within the Indo-European family

Phrygian shares several isoglosses with the Anatolian Indo-European languages: for example, the ending *-s* for the third-person singular of the preterite verb, and the middle verb ending *-r*, though neither is absolutely clear in Phrygian nor exclusively shared with the Anatolian subfamily.

One notices several similarities with Latin, such as the use of the preposition-preverb *ad(-)*, the extension of the infix *-k-* to the present stem (Neo-Phrygian ἀδδακετ no doubt corresponding to Latin *afficiat* “(s)he affects,” subjunctive), the ending *-tor* of the middle third-person singular (Latin *-tur*).

Unquestionably, however, Phrygian is most closely linked with Greek. Non-exclusive isoglosses include: (i) the relative pronoun *yos/ios/ιός* (see §4.2); (ii) the augment (see §4.3); (iii) the stem *pant-* “all” (Paleo-Phrygian *panta*, §1.1.1, 2, Vezirhan, l. 4; Middle Phrygian παντης, l. 7; Neo-Phrygian παντα, no. 35). Exclusive isoglosses include: (i) the *-s* ending of the nominative singular of *a*-stem masculine nouns (see §4.1.2); (ii) the denominative verbs in *\*-ye/o-* built on *o*-stems (Greek κακῶω “I mistreat,” Paleo-Phrygian *kakoioi/kakuioi*); (iii) the participial suffix *-meno-*; (iv) the pronoun *auto-*; (v) the stem *kako-*; (vi) the conjunction αἰ, having the same conditional usage as Doric and Aeolic αἰ.

These features betray very close prehistoric ties between the two languages, Phrygian and Greek, as well as the fact that they belong, no doubt, to the same dialectal subgroup of early Indo-European. The odds are that Midas’ titlature in M-01a, *Midai lavagtaei vanaktei*, where appear two functions that are also found in the Mycenaean Greek documents (*lawagetas* and *wanax*), does not correspond to Greek borrowings, but rather reflects the existence of a common heritage (on the stem *wanaks-* see Brixhe 1990:73–75).

We will dismiss, at least temporarily, the idea of a Thraco-Phrygian unity. Thraco-Dacian (or Thracian and Daco-Mysian) seems to belong to the eastern (satem) group of Indo-European languages and its (their) phonetic system is far less conservative than that of Phrygian (see Brixhe and Panayotou 1994, §§3ff.).

## 2. WRITING SYSTEMS

### 2.1 Paleo-Phrygian

The Paleo-Phrygian documents are written in an alphabetic script which resembles the archaic Greek alphabets (from which it is derived for the most part), characterized by the total absence of the letters that mark the aspirated stops in Greek, and presenting several regional variations:

1. A common set of seventeen symbols (conventionally transcribed by Latin letters): *a, b, g, d, e, v, i, k, l, m, n, o, p, r, s, t, u*, with the rectilinear *i* – as in Greek – and the splitting of the Semitic *waw* into F and Y.
2. Apart from a few symbols, the values of which are not evident, two letters are limited to certain regions: (i) a sign for the palatal glide /y/ (𐤀, 𐤁, of variable orientation), transliterated as *y*, the usage of which is optional and which, with the rectilinear *i*, evokes the pair formed by the two Greek *iôtas* (rectilinear and serpentine); (ii) a letter which probably corresponds to an affricate /ts/ or a variant thereof (𐤂, 𐤃; see Brixhe 1982:229–235), and which recalls the Ionian *sampi*, formally as well as functionally (cf. Woodard 1997:175–184).

A little less than one-third of the texts are sinistroverse (right to left); a few are written *boustrophedon* (alternating direction every other line). Exceptionally among the graffiti, more often on stones, words are separated by punctuation signs (three or four super-script dots).

The “peripheral” texts of §1.1.1 use an alphabet which diverges from the above on only two points: (i) each has a distinct symbol for /y/ (that of 1 being almost identical to that of 2), though it is likely that the graffiti of 3 use the common sign or a variant thereof; (ii) according to the editors of the texts, 1 and 2 have two symbols corresponding to voiceless sibilants: one of the two could represent /ts/ and replace the common sign (𐤂, 𐤃) which is here absent.

### 2.2 Neo-Phrygian

The script used is the Greek alphabet of the period, taking into account the phonological needs of the language, with a fund of seventeen letters: A, B, Γ, Δ, E, Z, I, K, Λ, M, N, O, Π, P, Σ, T, Υ. The letter Ω is rare except in the formula ΔΕΩΣ/ΖΕΜΕΛΩΣ (“gods/men” or “heavenly gods / infernal gods”). Also rare is H which often appears to be equivalent to E. The characters Ξ and Ψ and the “aspirated” signs, Φ, Χ, and Θ, are extremely rare or absent (on the value of the Greek characters, see Ch. 24, §2.3).

The texts are always written without separation of words.

## 3. PHONOLOGY

Though nearly a thousand years separates the earliest Paleo-Phrygian and the latest Neo-Phrygian documents, the conservatism inherent in the writing and in the formulary character of the texts allows one readily to verify that these are indeed two states of a single language (see Brixhe 1993:330–333). Having been spoken over such a vast area, Phrygian must have presented some local variations, even certain dialectal differentiations (see §1.1.1),

but our ignorance of much of the language, and the formulary nature of the evidence, prevents us from reaching definite conclusions on this point (see Brixhe 1993:337–338). The reader should bear in mind that Phrygian is still far from being well understood. At present only the simple texts are relatively clear – the short Paleo-Phrygian dedications and the Neo-Phrygian imprecations for instance. Phrygian is a language which is still undergoing decipherment.

### 3.1 Paleo-Phrygian consonants

The consonantal inventory of Paleo-Phrygian is presented in (1):

#### (1) Paleo-Phrygian consonantal phonemes

p	t	k
b	d	g
	ts	
	dz	
	s	
m	n	
	l	r
w	y	

#### 3.1.1 Obstruents

It was long claimed that the Phrygian consonantal system was dominated by a mutation (*Lautverschiebung*) of stops – the Proto-Indo-European aspirates becoming voiced, the voiced stops becoming voiceless, and the voiceless becoming aspirates. This is, however, a highly unlikely hypothesis; see Lejeune 1979 and, especially, Brixhe 1994:171–172.

The voiceless stop phonemes /p/, /t/, and /k/ developed from the voiceless stops of Proto-Indo-European – \*p, \*t, and \*k/\*k<sup>w</sup> respectively (on the Proto-Indo-European stops, see Ch. 17, §2.1.1): for example, *podas* (G-02), *matar* (e.g., W-04). The voiced stops, /b/, /d/, /g/, have two Proto-Indo-European sources: (i) the plain voiced stops \*d, \*g (e.g., again, *podas*); no certain examples exist for \*b, other than, possibly, the one that provides a *Lallname* like *Baba(s)* (*passim*); and (ii) the voiced aspirates \*b<sup>h</sup>, \*d<sup>h</sup>, \*g<sup>h</sup>: for example, *bagun* (G-136, if < \*b<sup>h</sup>ago-, cf. Neo-Phrygian (αβ)βερετ(ορ), root \*b<sup>h</sup>er-); *edaes* (root \*d<sup>h</sup>eh<sub>1</sub>, *passim*), and so forth.

The voiceless affricate /ts/ (?), written ↑, developed from \*k occurring before the front vowels /i/ and /e/, and is probably matched by a voiced /dz/ which arose from \*g and \*g<sup>h</sup> in the same context (see Brixhe 1982:229 ff.).

#### 3.1.2 Sonorants

Paleo-Phrygian has a pair of nasals, /m/ and /n/, with a neutralization of the contrast in word-final position in favor of the dental (just as in Greek). The liquids /l/ and /r/, and the glides /w/ (traditionally transcribed as ν) and /y/ complete the sonorant inventory.

### 3.2 Neo-Phrygian consonants

As far as can be determined, several consonantal changes have occurred by the Neo-Phrygian period:

1. The voiced affricate /dz/ may have become a voiced fricative /z/, supplying a voiced counterpart to /s/ (which itself might possibly have merged with the voiceless affricate, though no examples are available).
2. Word-final nasals have been eliminated. However, since they were preserved in the orthography, they at times appear by hypercorrection in unexpected positions: compare the dative of the masculine/neuter demonstrative σεμουν (the norm, < \**se-me/o-*) with the feminine σαν (for σα/σαι); see Brixhe 1978b:13–14, 19–20.
3. While the palatal glide /y/ has been preserved in all positions, the labiovelar /w/ (now written ΟΥ) seems to have disappeared before a rounded vowel.

### 3.3 Paleo-Phrygian vowels

The vocalic inventory of Paleo-Phrygian consists of five short vowels, /i/, /e/, /a/, /o/, and /u/, and at least four long (not distinguished from the short ones in writing): /i:/, /a:/, /o:/, and /u:/. The inherited long mid vowel \**ē* merged with \**ā* (cf. *matar* from \**mātēr*), a merger which is perhaps structurally linked to the appearance of a secondary /e:/ arising by monophthongization of the Proto-Indo-European diphthong \**ei* (see Brixhe 1990:70–71; on vowels in general, see Brixhe 1983:115 ff.; 1990).

When occurring before another vowel, the high vowels /u(:)/ and /i(:)/ were pronounced with an off-glide, [w] and [y] respectively, either written (*kuliya*, G-101, -127; *t/guvatis*, G-133) or not (*agartioi*, G-02a; *tuaveniy*, M-01f). In addition, the mid vowel *o* was raised to *u* before the word-final nasal *-n* (e.g., *avtun*, W-01b, corresponding to Greek αὐτόν).

The language possesses both so-called “short” and “long” diphthongs: /oi/, /ai/, /au/; and /o:i/, /a:i/. It was noted above that \**ei* had perhaps already yielded /e:/; and it is possible that \**ou* had undergone a similar fate: \**ou* > /o:/ > /u:/.

### 3.4 Neo-Phrygian vowels

Neo-Phrygian has only five isochronous vowels (Brixhe 1983:118–121; 1990): (i) /a/ (which arises from the /a(:)/ and /a(:)i/ of the earlier language); (ii) /e/ (from /e/); (iii) /o/ (from /o/); (iv) /u/ (from /u(:)/, /o:/ [primary or secondary], and /o:i/; see Brixhe 1990:97); and (v) /i/ (from /i(:)/ and secondary /e:/ (< \**ei*)). The Neo-Phrygian vocalic system is further characterized by a neutralization of the contrast of /e/ and /i/ in hiatus (cf. δεως/διως [*passim*]). In addition, the language exhibits a neutralization of the contrast /e/ ~ /i/ and /o/ ~ /u/ (in favor of the high vowel) in absolute-final position (bear in mind the elimination of the nasal in this position): compare, for example, *κακουν* (the norm) for *κακον* (*passim*) or the adverb *κακιν* (with a purely graphic nasal, no. 14) for *κακε(ν)* (e.g., no. 88); the inflectional ending of the Paleo-Phrygian athematic singular dative, *-ei* or *-ey*, can appear in Neo-Phrygian as *-ei* (historical), *-i* (phonetic), or *-e* (an inverse spelling related to the interchangeability, in final position, of *e* and *i*), see Brixhe 1990:78–79.

### 3.5 Middle Phrygian vowels

The stage I have proposed to identify as Middle Phrygian (see §1.2) perhaps preserves an intermediary phase of certain vowel changes. For example, in the phrase εϣ [σ]ϣ σοροι (ll. 2–3), the spelling [σ]ϣ (for Paleo-Phrygian \**sai* or \**say*) suggests that the final syllable of σοροι probably continues a historical spelling; the diphthong /-o:i/ has already lost its second element, but has not yet become /u(:)/ (cf. Neo-Phrygian σορου, no. 21).



## 4. MORPHOLOGY

### 4.1 Nominal morphology

Although we have not yet identified a complete Phrygian paradigm, it remains possible to outline the inflectional system of the language. Phrygian, being an early Indo-European language, inflects nouns for case, gender, and number. At least four morphological cases can be identified (nominative, accusative, genitive, dative), three genders (masculine, feminine, neuter), and two numbers (singular and plural). In typical Indo-European fashion, Phrygian attests thematic and athematic stems, as well as a class of nouns formed with the stem-vowel *a* (PIE *\*-eh<sub>2</sub>*). On the Indo-European categories, see Chapter 17, §3.5.

#### 4.1.1 Thematic stems

Among nominals, the inflection of thematic stems is best understood (see Brixhe 1990:94–97; 1999, §5). The following sequences of thematic vowel + ending are identified (Paleo-Phrygian forms in Latin, Middle and Neo-Phrygian in Greek characters, with phonetic transcription in brackets):

(2)	<i>Singular</i>	<i>Plural</i>
<i>Nominative</i>	-os / -ος	-oi / -οι
<i>Accusative</i>	-un / -ου(v), -υ(v) ([ <i>-u</i> ])	
<i>Genitive</i>	-ovo / -ου ([ <i>-u</i> ])	
<i>Dative</i>	-oy/-oi ([ <i>o:y/o:i</i> ]) / -ου(v) ([ <i>-u</i> ])	-ωσ(ι)

#### 4.1.2 *a(:)*-stems

Both masculine and feminine forms are attested in this inflectional category which corresponds to the Greek first declension. Paleo-Phrygian provides masculine singular forms: (i) a nominative in *-as* (alternating with *-a*, see below); (ii) perhaps a genitive in *-vo* (cf. *leravo* or *lelavo*, W-10); and (iii) a dative in *-ai* (*Midai*, M-01a). Feminine forms include the following: (i) a nominative singular in *-a* (*Imeneia*, G-183b); (ii) an accusative *-αν* ([*-an*]); ... *μανκων ιαν εστας*, “... the m. that he erected,” no. 31); (iii) a genitive in *-ας* ([*-as*]; e.g., Middle Phrygian *μεκας*; Neo-Phrygian *σκερεδριας*, no. 56); (iv) a dative in *-ay* (*avtay*, W-01b), *-αι* (*σαι* ... *μανκαι*, no. 35) / *-a* (*σα* ... *μανκα*, no. 82); (v) and, in the plural, a dative in *-ais* (*mekais*, G-239).

#### 4.1.3 Athematic stems

In the case of the athematic stems, we are less well-off. Note that the lowering of *\*ē* to *\*ā* gave certain paradigms a novel character; consider, for example, the word for “mother”: nominative *matar*, accusative *materan*, dative *materey*; or the man’s name *μᾶν*, genitive *μεινος*, in the Greek texts of the Roman era. The quantitative contrast of *\*ē* versus *\*e* was transformed into a contrast in quality and quantity (Paleo-Phrygian), and then simply into a contrast in quality (Neo-Phrygian).

### 4.2 Pronouns

The Phrygian documents shed some light on the phenomena of deixis and anaphora. Those Proto-Indo-European demonstrative/anaphoric stems *\*se/o-* and *\*te/o-*, which are associated in Greek with the article paradigm, may have become specialized in Phrygian:

1. *\*se/o-* as the demonstrative: nominative/accusative neuter singular *si* (equivalent to Latin *hoc*); accusative, genitive, and dative feminine singulars  $\sigma\alpha\nu$ ,  $\sigma\alpha\varsigma$ ,  $\sigma\alpha(1)$ ; dative masculine and neuter singular  $\sigma\epsilon\mu\omicron\upsilon\nu$  (< *\*se-me/o*).
2. *\*te/o-* as the anaphoric: compare the correlation in Neo-Phrygian  $\iota\omicron\varsigma \nu\iota \dots$ ,  $\tau\omicron\varsigma \nu\iota \dots$  (“whoever . . . , that one . . . ,” no. 6, 25); or the genitive in Paleo-Phrygian *tovo* (< *\*to-wo*) which became  $\tau\omicron\upsilon$  in Neo-Phrygian (no. 87) and with which the dative merged after /o:i/ became /u/ (see §3.4).

In addition, the Proto-Indo-European reflexive stem *\*swe-* seems to have also provided an anaphoric – compare dative  $\omicron\iota$  – and perhaps the possessive as well: Neo-Phrygian  $\omicron\upsilon\alpha$  (no. 2, 33, 36).

Just like Greek, Phrygian has a stem *auto-* expressing identity and used emphatically (cf.  $\alpha\upsilon\tau\omicron\varsigma$ , no. 33, 36), and which likely combined with the preceding pronominal to form a reflexive (?), cf. *ven-avtun* (W-01b) and  $\epsilon\text{-}\alpha\upsilon\tau\alpha\iota$  (no. 116, l.12).

Finally, the relative is *ios/yos/ios* (*passim*). The Phrygian indefinite is represented by masculine  $\kappa\omicron\varsigma$  and the neuter  $\kappa\iota\nu$ . The indefinite relative pronouns are  $\iota\omicron\varsigma \nu\iota$  or  $\iota\omicron\varsigma \kappa\epsilon$ .

On Phrygian pronouns as a whole, see Brixhe 1978b:6–22; 1990:95–97; 1997, §§5.1.1, 5.1.2, 6.1, and 6.2.

### 4.3 Verbal morphology

Phrygian verbs are morphologically marked for tense, voice, and mood, and by inflectional endings which encode the typical Indo-European distinction of three persons and two numbers.

In addition to present tense, the Phrygian documents provide evidence of a possible future tense stem in *-s-* (*egeseti* P-04a,  $\epsilon\gamma\epsilon\sigma\iota\tau$  no. 58). Phrygian undoubtedly possesses a preterite tense formed with a prefix (the Indo-European “augment” known from Greek, Armenian, and Indo-Iranian) and having a third-person singular marked by *-s-*: consider, for example, *e-daes/ε-δαες* (cognate with Latin *fecit*, < PIE\**d<sup>h</sup>eh<sub>1</sub>*), and the compounds *en-e-parkes/εν-ε-παρκες* (Brixhe and Lejeune 1984: 14),  $\pi\omicron\sigma\text{-}\epsilon\text{-}\kappa\alpha\nu\epsilon\varsigma$  (no. 116, l. 7). A perfective stem characterized by reduplication also occurs: for example,  $\tau\epsilon\text{-}\tau\iota\kappa\mu\epsilon\nu\omicron\varsigma$  (*passim*),  $\gamma\epsilon\text{-}\gamma\alpha\rho\iota\tau\mu\epsilon\nu\omicron\varsigma$  (nos. 33, 36, 79).

Phrygian distinguishes a voice contrast of active versus middle: for example,  $\alpha\delta\delta\alpha\kappa\epsilon\tau$  –  $\alpha\delta\delta\alpha\kappa\epsilon\tau\omicron\rho$  (< *\*d<sup>h</sup>h<sub>1</sub>-* or *\*d<sup>h</sup>eh<sub>1</sub>-*);  $\alpha\beta\beta\epsilon\rho\epsilon\tau$  –  $\alpha\beta\beta\epsilon\rho\epsilon\tau\omicron\rho$  (< *\*b<sup>h</sup>er-*); see Brixhe 1979:177ff.

In addition to the indicative, we have reason to suspect that Phrygian has alongside a subjunctive mood (?  $\alpha\delta\delta\alpha\kappa\epsilon\tau$ ,  $\alpha\beta\beta\epsilon\rho\epsilon\tau$ ) also an optative: thus, the third-person singular *kakoioi* and *kakuioi* (G-02c, P-04b), which is a denominative of *kako-*, “make *kako-*.” To the optative ending *-oi* may correspond plural *-oyen*, as in *[...]toyen* (W-04). In the imperative mood, only third-person forms are attested: in the active voice, a singular *-to/-του* (< *\*tōd*) and a plural *-vou/-(τ)τινου*; in the middle, a singular *-do/-δου* (< *\*d<sup>h</sup>ō?*); see Brixhe 1979:177–184; 1990:90–91.

The Phrygian documents preserve a single participial morpheme – the middle participle suffix *-meno-*.

## 5. SYNTAX

### 5.1 Word order

In the case of Paleo-Phrygian, as well as Neo-Phrygian, to the extent that major subject constituents can be identified, the language generally remains faithful to the Indo-European order SV (see Brixhe 1983:126).

Phrygian has both proclitics and enclitics, at times occurring in sequence, as, for example, in nos. 33 and 76

- (3) ακ κε οι  
proclitic preposition – enclitic conjunction – enclitic anaphoric object of the preposition ακ

Compare, however, the order found in no. 88

- (4) πουρ Ουανακταν κε  
proclitic preposition – object of the preposition – enclitic conjunction

On the order of the Phrygian clitics, see Brixhe 1997, §7.

## 5.2 Case usage

As in other ancient Indo-European languages, Phrygian prepositions (e.g., εν, αδ, ας, με, πορ/πορ/πουρ) require their nominal objects to be inflected in particular noun cases (see Brixhe 1997, §2).

The phonetic changes which occurred in the evolution of Neo-Phrygian from Paleo-Phrygian (see §§3.2, 3.4) fused final *-on* (accusative), *-owo* (genitive), and *-o:i* (dative) into /u/ and thus led to the merger of the accusative, the genitive, and the dative cases in the singular (then no doubt in the plural) for the thematic and then the other stems. Compare the feminine final *-ας* and *-αν*, where one would expect *-α(ι)* (see Brixhe 1978b:13–14, 19–20; 1997, §2.4).

## 6. LEXICON

The irregularity of the punctuation in Paleo-Phrygian and the general absence of word division in Neo-Phrygian are clearly obstacles to text segmentation. Nevertheless, it has been possible to isolate a number of lexical units, to which we are able to assign meaning when the text is straightforward: consider the famous Paleo-Phrygian dedication of the so-called Tomb of Midas (M-01a):

- (5) *Ates... Midai lavagtaei vanaktei edaes*  
“Ates... has dedicated [this monument] to Midas, lavagtas and vanax”

or a Neo-Phrygian protasis such as (no. 61):

- (6) ιος vi σεμουν κνουμανε κακουν αδδακετορ...  
“whoever will damage this tomb...”

Still, the latter translation hides our ignorance of the precise meaning of κνουμανει, the dative of the neuter κνουμαν. Similarly, we do not know the exact meaning of a series of terms relating to architecture (often funerary): for example, *iman* (appellative in G-136, etc.); *meka* (G-239, P-03, etc.), Middle Phrygian μεκα (l. 1), μανκα (Middle Phrygian l. 1; Neo-Phrygian nos. 2, 18, 26, etc.); Neo-Phrygian σκερεδρια/σκελεδρια (nos. 56, 67).

Regarding place names (toponyms), the persistence of a characteristic Phrygian suffix should be noted, which in the Greek transcriptions has the form *-αειον*, with variants *-αιον* and *-αεον*, for example, Κοτιαειον (the present Kütahya; cf. Zgusta 1984, §594/3).

Greek inscriptions in the Phrygian area are extremely valuable for the understanding of personal names (anthroponyms; Zgusta 1964:552–555). These seem to have remained

quite stable throughout the history of the language; though names of Hittito-Luwian origin increased in frequency. Typical Phrygian names include *Iman* (e.g., G-210),  $\mu\alpha\nu - \mu\epsilon\nu\omicron\varsigma$  (Zgusta 1964, §466/1); *Aladis* (G-109), *Voine(s)* (G-129, G-228),  $\omicron\upsilon\epsilon\nu\alpha\omicron\upsilon\iota\alpha/-\eta$  or  $\omicron\upsilon\epsilon\nu\alpha\upsilon\iota\alpha/-\eta$  (Zgusta 1964, §1153/1–3),  $\omicron\upsilon\alpha\nu\alpha\xi\omicron\varsigma/\omicron\upsilon\alpha\nu\alpha\xi\iota\omega\nu/\omicron\upsilon\alpha\nu\alpha\xi\omega\nu$  (Zgusta 1964, §1138/1–3),  $\Xi\epsilon\upsilon\nu\alpha/\Xi\epsilon\upsilon\nu\eta/\Xi\epsilon\upsilon\nu\epsilon$  (no. 116). Hittito-Luwian names include *Mamutas* (G-229), *tuvatis/guvatis* (G-133),  $\tau\rho\omicron\kappa\omicron\nu\delta\alpha\varsigma$  (Zgusta 1964, §1512/31). Also found are infantile terms such as *Ata(s)* (G-107, G-128, etc.),  $\Gamma\alpha, \Delta\alpha, \Nu\alpha/\Nu\eta$  (Zgusta 1964, index), *Mama* (G-173), and so forth.

Phrygian contacts with speakers of Anatolian languages such as Hittite and Luvian brought about an interesting morphological phenomenon: as opposed to the masculine/feminine morphological contrast of Phrygian (with two forms), the Anatolian languages have only one common animate gender (with a unique form; see Ch.18, §4.2.1; Ch. 19, §4.1). This divergence would generate a Phrygian tendency to have a single ending for the names of both men and women, hence the wavering between these two categories: for example, masculines *Ata* or *Voine* next to *Atas* and *Voines* and conversely, no doubt, feminine forms such as  $\Nu\alpha\nu\alpha\varsigma$  next to  $\Nu\alpha\nu\alpha$  (see Brixhe 1983:128; 1994:176).

At a later date, the Galatian impact (see §1.3) was quite modest. These Celtic speakers perhaps provided the Neo-Phrygian personal name  $\beta\omega\delta\omicron\rho\iota\varsigma$  (no. 34) and the lexeme  $\tau\epsilon\upsilon\tau\omicron\upsilon\varsigma/\tau\epsilon\upsilon\tau\omega\varsigma\iota$  (nos. 33, 36, 116), if the latter represents a reflex of *\*teuta*, which provided the western Indo-European dialects (such as Celtic) with a noun meaning “people” (see Brixhe 1993:338, 340, rectified by 1997, §2.5).

It is with Greeks that Phrygian contacts would be the longest and the most intense – stretching from the second millennium BC to the Roman period. Among the linguistic manifestations of these contacts are relatively ancient lexical borrowings such as Middle Phrygian  $\sigma\omicron\rho\omicron\iota$  (l. 3, dative) and Neo-Phrygian  $\sigma\omicron\rho\omicron\upsilon$  (no. 21, dative), from Greek  $\eta\ \sigma\omicron\rho\omicron\varsigma$  (a vessel hence “cinerary urn,” then “Sarcophagus”); or Middle Phrygian  $\kappa\omicron\rho\omicron$  (ll. 6 and 7), Neo-Phrygian  $\kappa\omicron\rho\omicron\upsilon$  (no. 92), from Greek  $\chi\omicron\rho\omicron\varsigma$  (“land, country”). More recent borrowings include the Neo-Phrygian  $\theta\alpha\lambda\alpha\mu\epsilon\iota$  (no. 4), from Greek  $\eta\ \theta\alpha\lambda\acute{\alpha}\mu\eta$  (“den”).

## Bibliography

- Brixhe, Cl. 1978a. “Études néo-phrygiennes I,” *Verbum* 1 (1): 3–21.  
 ———. 1978b. “Études néo-phrygiennes II,” *Verbum* 1 (2): 1–22.  
 ———. 1979. “Études néo-phrygiennes III,” *Verbum* 2:177–192.  
 ———. 1982. “Palatalisations en grec et en phrygien.” *BSL* 77:209–249.  
 ———. 1983. “Epigraphie et grammaire du phrygien: état présent et perspectives.” In E. Vineis (ed.), *Le lingue di frammentaria attestazione* (= *Die indogermanischen Restsprachen*, Atti del Convegno della Società Italiana di Glottologia e della Indogermanische Gesellschaft, Udine 1981), pp. 109–133. Pisa: Giardini.  
 ———. 1987. *Essai sur le grec anatolien au début de notre ère*. 2nd ed. Nancy: Presses universitaires de Nancy.  
 ———. 1989–1990. “La plus occidentale des inscriptions phrygiennes,” *Incontri linguistici* 13:61–67.  
 ———. 1990. “Comparaison et langues faiblement documentées: l'exemple du phrygien et de ses voyelles longues.” In J. Kellens (ed.), *La reconstruction des laryngales*, pp. 59–99. Liège: Université de Liège.  
 ———. 1991. “Les inscriptions paléo-phrygiennes de Tyane: leur intérêt linguistique et historique.” In Br. Le Guen-Pollet and O. Pelon (eds.), *La Cappadoce méridionale jusqu'à l'époque romaine. Etat des recherches*, pp. 37–46 and pl. I. (Actes du Colloque d'Istanbul, Institut Français d'Etudes Anatoliennes, 13–14 avril 1987). Paris: Editions Recherche sur les Civilisations.  
 ———. 1993. “Du paléo- au néo-phrygien.” *Académie des Inscriptions et Belles-Lettres, Comptes Rendus*, pp. 323–344.  
 ———. 1994. “Le phrygien.” In Fr. Bader (ed.), *Langues indo-européennes*, pp. 165–178. Paris: CNRS Editions.

- . 1995. "Les Grecs, les Phrygiens et l'alphabet." In Al. Fol *et al.* (eds.), *Studia in honorem Georgii Mihailov*, 101–114. Sofia: Institute of Thracology.
- . 1996. "Les documents phrygiens de Daskyleion et leur éventuelle signification historique," *Kadmos* 35: 125–148.
- . 1997. "Les clitiques du néo-phrygien." In Gusmani *et al.* 1997, 41–70.
- . 1999. "Prolégomènes au corpus néo-phrygien," *BSL* 94:285–315.
- Brixhe, Cl. and Th. Drew-Bear. 1997. "Huit inscriptions néo-phrygiennes." In Gusmani *et al.* 1997, 71–113.
- Brixhe, Cl. and M. Lejeune. 1984. *Corpus des inscriptions paléo-phrygiennes*. Paris: Editions Recherche sur les Civilisations.
- Brixhe, Cl. and G. Neumann. 1985. "Découverte du plus long texte néo-phrygien: l'inscription de Gezler Köyü." *Kadmos* 24:161–184.
- Brixhe, Cl. and A. Panayotou. 1994. "Le thrace." In Fr. Bader (ed.), *Langues indo-européennes*, pp. 179–203. Paris:
- Brixhe, Cl. and M. Waelkens. 1981. "Un nouveau document néo-phrygien au Musée d'Afyon." *Kadmos* 20:68–75.
- Cox, C. W. M. and A. Cameron. 1932. "A native inscription from the Myso-Phrygian Borderland," *Klio* 25:34–49.
- Darga, M. 1993. "Fouilles de Şarhöyük-Dorylaion (1989–1992)," *XV. Kazi sonuçları toplantısı (Anakara 1993)* I, 481–501. Ankara: Turk Ministry of Culture.
- DeVries, K. 1980. *From Athens to Gordion. The Papers of a Memorial Symposium for Rodney S. Young*, pp. 33–49. University Museum Papers I. Philadelphia: University of Pennsylvania.
- Diakonoff, I. and V. Neroznak. 1985. *Phrygian*. Delmar, NY: Caravan Books.
- Diñç, R. and L. Innocente. 1999. "Ein Spinnwirtel mit phrygischer Inschrift," *Kadmos* 38:65–72.
- Friedrich, J. 1932. *Kleinasiatische Sprachdenkmäler*. Berlin: Walter de Gruyter.
- . 1941. *RE* XX-1, col. 868–891, s.v. *Phrygia*: *Sprache* et *Phrygia*: *Geschichte*.
- Gunter, A. 1991. *Gordion Excavations. Final Reports III: The Bronze Age*. Philadelphia: University of Pennsylvania.
- Gusmani, R. 1958, 1959. "Studi frigi." *Rendiconti dell' Istituto Lombardo* 92:835–928; 93:17–49.
- Gusmani, R. and G. Polat. 1999. "Manes in Daskyleion," *Kadmos* 38:137–162.
- Gusmani, R. and Y. Polat. 1999. "Ein neues phrygisches Graffito aus Daskyleion," *Kadmos* 38:59–64.
- Gusmani, R., M. Salvini, and P. Vannicelli. 1997. *Frigi e frigio*: Atti del 1° Simposio Internazionale. Roma: Consiglio Nazionale delle Ricerche.
- Haas, O. 1966. *Die phrygischen Sprachdenkmäler*. Sofia.
- Laminger-Pascher, G. 1984. *Beiträge zu den griechischen Inschriften Lykaoniens* (= *Ergänzungsband zu TAM*, Nr. 11 = *Denksch.* Wien 173). Vienna.
- Lejeune, M. 1979. "Regards sur les sonores indo-européennes en vieux phrygien." In *Mélanges E. Laroche*, pp. 219–224, Paris.
- Manning, St. W., B. Kromer, P. I. Kuniholm and M. W. Newton. 2001. "Anatolian tree rings and a new chronology for the east Mediterranean Bronze–Iron Ages," *Science* 294:2532–2535.
- Masson, O. 1987. "Le sceau paléo-phrygien de Mane," *Kadmos* 26:109–112.
- Mitchell, S. 1993. *Anatolia. Land, Men and Gods in Asia Minor*. Oxford: Oxford University Press.
- Neroznak, V. 1978. *Paleobalkanskije jazyki*. Moscow: Nauta.
- Neumann, G. 1988. *Phrygisch und Griechisch* (= *Sitzungsber. Wien* 499). Vienna: Österr. Akad. Wissenschaften Verlag.
- . 1997. "Die zwei Inschriften auf der Stele von Vezirhan." In Gusmani *et al.* 1997:13–32.
- Orel, Vl. 1997. *The Language of Phrygians. Description and Analysis*. New York: Delmar.
- Sams, G. 1994. *The Gordion Excavations. Final Reports IV: The Early Phrygian Pottery*. Philadelphia: University of Pennsylvania.
- Varinlioglu, E. 1992. "The Phrygian Inscriptions from Bayındır," *Kadmos* 31:11–20.
- Voigt, M., G. K. Sams, and K. DeVries. 2001. Paper read to the Van Symposium (6.08.2001; in press).
- Woodard, R. 1997. *Greek Writing from Knossos to Homer*. Oxford: Oxford University Press.
- Zgusta, L. 1964. *Kleinasiatische Personennamen*, Prague: Tschechoslowakische Akademie der Wissenschaften.
- . 1984. *Kleinasiatische Ortsnamen*, Heidelberg: Winter.

# Latin

JAMES P. T. CLACKSON

## 1. HISTORICAL AND CULTURAL CONTEXTS

Latin – the language of Ancient Rome – takes its name from Latium, a region encompassing Rome on the west coast of Italy and bordered by the river Tiber to the northwest, the Apennines to the northeast and the Pontine marshes to the south. The Roman antiquarian Varro dated the founding of Rome to 753 BC, but there is archeological evidence for settlement much earlier than this, and it was only later, in the sixth century BC, that Rome became an organized and sophisticated city-state. Latium itself did not achieve political unity until it came under Rome's dominance in the fourth century BC, but the Latini – as the inhabitants of Latium are termed – appear to have shared cultural and religious practice, as well as their language, from well before the period of the first city-states.

The increasing control over Latium was the first stage of Rome's rise to power throughout the Italian peninsula a dominance achieved through conquest, alliance, and colonization. By the second century BC, Rome's military power was great enough to make possible the conquest and annexation of territory outside Italy, including North Africa, Spain, Southern France and Greece. Civil wars throughout much of the first century BC led to the end of the Roman Republic and the foundation of the Roman Empire under Augustus. Imperial rule continued for over four hundred years, and under Trajan (AD 98–117) and Hadrian (AD 117–138) the empire reached its maximum extent, stretching from Britain to Egypt and encompassing much of Europe and all of the area surrounding the Mediterranean.

The influence and spread of the Latin language mirrored the power and extent of the Roman Empire and, even today, Latin retains a great deal of cultural prestige. It is still widely taught and it is retained in use by many different scientific, legal, and religious institutions around the world. Moreover, over 500 million people currently speak as their first language a language derived from Latin.

Latin is one of a number of Indo-European languages which were spoken in ancient Italy. It shares several features with the Faliscan language which was spoken to the north of Latium, most importantly the existence of two thematic genitive singular morphemes *-ī* and *-osio* (the latter attested in an Early Latin inscription recently discovered in Satricum), and the formation of a future tense with a morpheme containing a reflex of Indo-European *\*-bʰ-*. Although we have only limited knowledge of the Faliscan language (our sources are mostly epitaphs and a few early vase-inscriptions, none of them extensive) these morphological agreements are sufficient to lead to a general consensus that Latin and Faliscan form a subgroup. More contentious is the relationship between Latin-Faliscan and the Sabellian group languages, which include Oscan and Umbrian (see Ch. 33). Many scholars have judged that the similarities between Latin and Sabellian justify the reconstruction of an



Italic subgroup of Indo-European, but it has proved difficult to demonstrate conclusively that these similarities result from genetic affiliation, and have not arisen through convergence of separate branches of Indo-European over time. Our present state of knowledge of Sabellian and the early history of Latin is not sufficient to allow a definite answer to this question.

Latin has a long history. The earliest documents date from the seventh century BC, and there is a continuous literary tradition from the third century BC through to the medieval period. The following chronological stages of the language are generally recognized, although there is no clear agreement on exact dating:

1. *Early or Old Latin*: The language from the earliest times to c. 100 BC. A distinction is often recognized between *Pre-literary Latin*, the scantily attested language of the earliest documents, and *Pre-Classical Latin*, the language of the first extensive literary works from c. 240 BC to c. 100 BC.
2. *Classical Latin*: The language of official inscriptions and literature from c. 100 BC to c. AD 14. Following the models of the works of Cicero and the writers of the age of Augustus, Classical Latin became the standard for later canons of "Latinity," and it has remained the prestige form of written compositions right up to the present day.
3. *Post-Classical and Late Latin*: The language of writers after the Classical period. Writers of the hundred years following the death of Augustus (AD 14) have traditionally been judged harshly in comparison with their predecessors, and their language is sometimes termed *Silver Latin*, although the linguistic differences between Classical and Silver Latin are not great. The leveling effect of the standard language, and the increasing artificiality of much of the literary language makes it difficult to define exactly when the stage of *Late Latin* begins, but it is clear that by AD 400 even the standardized form of the language shows substantial differences from Classical Latin.

A further classification often encountered is *Vulgar Latin*. This is not a chronological stage per se but rather a catch-all category which is used to denote an informal register of Latin spoken by those who had received little or no literate education, as opposed to Classical Latin, the formal standard language of the elite. However, our knowledge of the spoken registers of Latin is severely limited, since the written record is always prone to influence from the standard, and the construction of a uniform "Vulgar Latin" probably oversimplifies a very complex linguistic situation. Different communities of speakers used different nonstandard varieties, and the relationship between the spoken registers and the artificial written language changed considerably over time. We should consequently bear in mind that, rather than a simple opposition between Classical Latin and Vulgar Latin, there was a much more complex relationship between an ever-evolving standard and a number of different spoken registers. In short, there is no Vulgar Latin, only "vulgar" forms present in a greater or lesser degree in individual texts.

Nowadays few scholars would recognize another distinct variety of Latin in the language used in Christian texts, although the case for *Christian Latin* was promoted enthusiastically by some Dutch scholars in the first half of the twentieth century. It is true that Christian texts show a preponderance of lexical items which do not occur as frequently elsewhere, but these words all refer to new institutions or beliefs, and in fact it can be shown that the language of the early Christian writers shows the same properties and variety as contemporary non-Christian registers.

The best evidence for dialectal varieties of Latin comes from the earlier periods of its history. Surviving inscriptions from Praeneste, an early city-state of Latium and rival to Rome, support the statements from ancient sources suggesting that Praenestine was distinct from Roman Latin. However, as Roman power and prestige grew, dialects outside Rome



became stigmatized as rustic and were subsumed in the Classical period under the influence of the standard language. All later Latin is written by those who have had at least some education in the standard and it is difficult to detect any major regional differences across the empire.

## 2. WRITING SYSTEM

The Latin alphabet is derived, like the Faliscan alphabet, from a variety of the Etruscan script, itself an adaptation of the Western Greek Euboean alphabet. The earliest Latin script has the following letters:

### (1) ABCDEFZHIKLMNOPQRSTVX

In the use of the letters *C*, *K*, and *Q* this script shows its clear affinities with the Etruscan alphabet used in South Etruria (see Ch. 39, §2). The three letters are all used to represent both the voiceless and the voiced velar stop, but a convention loosely followed on several early inscriptions (and also found in Faliscan inscriptions) governs their distribution: *C* occurs before front vowels; *K* before /a/; and *Q* before rounded back vowels (note also the names of the letters *cē* for *C*, *kā* for *K*, and *qū* for *Q*). The very earliest Latin inscriptions use a digraph *FH* for /f/, following Etruscan practice, but this is soon replaced by the simple *F*; the Faliscan alphabet uses a new sign ↑ for /f/. The letters *B*, *D*, and *O* are not used in Etruscan texts; the Etruscan language appears to have lacked a phonemic contrast between voiced and voiceless stops, and to have had only one back vowel, written *V*. However, these letter forms are attested on Etruscan abecedaria, and so there is no need to posit a direct borrowing from Greek. The letter *X* is used for the combination /k/ + /s/. The letter *Z* is used in the Faliscan alphabet but, apart from its presence in a Latin abecedarium, it occurs in only one very fragmentary archaic Latin inscription, with uncertain value, and it dropped out of use completely by the third century BC. The position of *Z* in the Latin alphabet was taken by a new letter – *G* – invented in order to differentiate the writing of voiceless and voiced velar stops. In the first century BC the need to represent the sounds of Greek loanwords led to the adoption of the letters *Y* and *Z* directly from the Greek alphabet. The distinct writing of *V* and *J* for the consonants /w/ and /y/ as opposed to *U* and *I* for the vowels /u/ and /i/ is first made systematic only in the fifteenth century AD.

**Table 32.1 The Archaic Latin alphabet**

Character	Transcription	Character	Transcription
Ɑ	a	Ɱ	m
Ɱ	b	Ɐ	n
Ɱ	c/g	Ɒ	o
Ɱ	d	ⱱ	p
Ɱ	e	Ⱳ	q
Ɱ	f	ⱳ	r
Ɱ	z	ⱴ	s
Ɱ	h	Ⱶ	t
Ɱ	i	ⱶ, ⱷ	v
Ɱ	k	ⱸ	x
Ɱ	l		

### 3. PHONOLOGY

#### 3.1 Consonants

The phonemic inventory of Latin consonants is presented in Table 32.2:

Table 32.2 The consonantal phonemes of Latin							
Manner of articulation	Place of articulation						
	Bilabial	Labiodental	Dental	Palatal	Velar	Labiovelar	Glottal
<i>Stop</i>							
<i>Voiceless</i>	p		t		k	k <sup>w</sup> (?)	
<i>Voiced</i>	b		d		g	g <sup>w</sup> (?)	
<i>Fricative</i>		f	s				h
<i>Nasal</i>	m		n		ŋ (?)		
<i>Liquid</i>							
<i>Lateral</i>			l				
<i>Nonlateral</i>			r				
<i>Glide</i>				y		w	

##### 3.1.1 “Labiovelars”

The phonological status of the sounds written with the digraphs *QU* and *GU* is debated. There is no conclusive argument for favoring a realization [k<sup>w</sup>] and [g<sup>w</sup>] over [kw] and [gw], or indeed for supposing that *GU* must represent the voiced counterpart to *QU*, especially since the distribution of *GU* is much more restricted than *QU* (it only occurs after a nasal). Both sounds can derive from earlier unitary phonemes \*k<sup>w</sup> and \*g<sup>w</sup>, but *QU* also continues the cluster \*kw. It is true that metrical texts normally require that a word such as *liquidus* “liquid” be scanned with the first syllable short, thus implying [li.k<sup>w</sup>i-] (with syllable boundary represented by the dot), but there are some texts where this word is scanned with the first syllable long implying a pronunciation [lik.wi-].

##### 3.1.2 The velar nasal

The sound [ŋ], which occurs only in syllable codas, is represented in the orthography by *N* before velars and *G* before nasal consonants. This orthographic convention may imply that [ŋ] is an allophone of both /n/ and /g/, but distinctive triplets such as *amnī* “river” (dat. sg.), *annī* “year” (gen. sg.), and *agnī* [aŋni:] “lamb” (gen. sg.) could be taken to support the existence of a separate phoneme /ŋ/.

##### 3.1.3 Glides

In Proto-Indo-European the phonemes /i/ and /u/ have consonantal and vocalic allophones, and it is likely that this is continued into Early Latin. This is reflected in the Latin script, where a single letter, *I*, is used to represent both the vowel [i] and the glide [y], and *V* is used for [u] and [w]. Indeed, for the sake of metrical convenience, Latin poets of the Classical period occasionally interchange [i] / [y] and [u] / [w]; the word for “knees” – *genua* – is generally a trisyllable, but in verse it is found as a disyllable [genwa], and the name *Iūlius* is found scanned as both a trisyllable and quadrisyllable. However, apart from this poetic

license, by the Classical period [i] / [y] and [u] / [w] are no longer allophones but separate phonemes – note the minimal pairs *iambus* [iam-] “iambus” : *iam* [ya] “now,” and *uoluit* [-lurt] “he wanted” : *uoluit* [-lwrt] “he rolls.”

Metrical evidence from Latin poetry shows that when the glide /y/ is intervocalic it is usually pronounced as a double consonant although not normally so written – thus *maior* “greater” was pronounced [mayyor]. The glide /y/ is not usually found following a consonant, except in compounds such as *con-iungo* “join together”; hence *medius* “middle” was pronounced [mɛdius] not [medyus]. Conversely the glide /w/ is found in clusters such as [sw] and possibly [kw] and [gw] (see §3.1.1), with the consequence that in some forms the orthography is ambiguous: for example, *sua* [sua] “one’s own” (fem. nom. sg.) is written the same as *suāvis* “sweet” [swa:wis].

### 3.2 Vowels

Latin has five short vowels /i/, /e/, /a/, /o/, and /u/ and five long vowels /i:/, /e:/, /a:/, /o:/, and /u:/. Short and long vowels are paired in many areas of Latin morphology, thus one means of forming the perfect tense employs the rule that the root vowel is lengthened:

- (2) /i/ : /i:/ *uincō* “I conquer,” *uicī* “I conquered”  
 /e/ : /e:/ *ueniō* “I come,” *uenī* “I came”  
 /a/ : /a:/ *scabō* “I scratch,” *scābī* “I scratched”  
 /o/ : /o:/ *fodiō* “I dig,” *fōdī* “I dug”  
 /u/ : /u:/ *fugiō* “I flee,” *fūgī* “I fled”

However, for the high and mid vowels, length differences also involved a change in quality; /i/ was probably realized as [ɪ] but /i:/ as [i:], and /e/ as [ɛ], but /e:/ as [e:], making /i/ actually closer in quality to /e:/ than to /i:/. This skewed phonetic realization was to have effects on the vowel system in the spoken registers from which the Romance languages originated. Vowel length was lost as a distinctive feature, becoming an automatic concomitant of the word stress, and under the stress short /i/ and long /e:/ merged as a high-mid front vowel [e]. In many areas a similar merger also took place between /u/ and /o:/, and one can suppose that there was a similar disparity between the phonetic values of long and short vowels on the back axis also.

In Classical Latin there was also a series of nasalized vowels, /ĩ/, /ē/, /ā/, /ō/, and /ū/, which were restricted in occurrence to (i) word-final position, where in the standard orthography they are written *im*, *em*, *am*, *om*, *um*; or (ii) before a sequence of nasal + continuant. All nasal vowels were inherently long; they do not contrast with short nasal vowels.

In Early Latin there are a number of distinctive diphthongs: /ei/, /ai/, /oi/, /au/, /ou/, and /eu/ (the last attested only in a single inscription). These all underwent monophthongization in nonstandard varieties of Latin, but in Classical Latin /au/ was generally maintained, /ai/ was continued as a diphthong /ae/, and, in a few lexical forms, the earlier diphthong /oi/ was continued as /oe/. In a very small number of words a new diphthong /eu/ also arose as a result of contractions. The remaining diphthongs were monophthongized; the exact details are complex, but in essence (i) *ei* > *ē*; (ii) *ou* and *eu* > *ū*; and (iii) *oi* > *ū* word-internally (in most cases, although sometimes *oi* > *oe*), *oi* > *ī* word-finally and following [w].

### 3.3 Accent

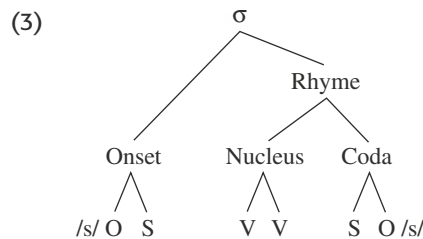
The word accent of Early Latin was a stress accent placed on the first syllable of every word. The effects of the accent can be seen in the syncope which affected many short final

vowels and a series of quality changes in vowels in noninitial syllables known collectively as “vowel weakening.” One such change is the merger of non-initial, non-final short vowels in open syllables to /i/, leading to morphophonemic alternations of the type *faciō* “I make,” compound form *re-ficiō* “I restore.” Sometime before the beginning of the Classical period, however, the place of the word accent had changed, and the accentuation of all Latin words, bar a few special exceptions, can be predicted by the rule of the penultimate: the accent falls on the penultimate syllable, unless that syllable is metrically light (i.e., open with a short vowel nucleus) in which case the accent falls on the antepenultimate.

The nature of the word accent of Classical Latin is disputed. Ancient grammarians, who largely follow Greek models of description, uniformly use the terms applicable to the pitch accent-type of Greek to describe the Latin accent, although metrical practice and the continuing evidence for syncope throughout the Classical and post-Classical period (and in the reconstructed early development of the Romance language) strongly suggest that the word accent continued to be stress.

### 3.4 Syllable structure

Latin syllables can be described under the schema of (3), where O = obstruent (stop or /f/), S = sonorant, V = vocalic element (here long vowels are counted as equivalent to a double short vowel):



In initial or final clusters with /s/ the obstruent must be a voiceless stop; thus, the nominative singular of the word for “town” – written *urbs* in most modern editions of Latin texts – was pronounced, and sometimes written, *urps*. Some clusters are avoided: for example initial /dr-/ is rare, occurring only in loanwords and onomatopoeia; and /-ts/ is always replaced by /-s/.

Syllables with VV nucleus and an element in the coda, sometimes termed *superheavy* or *overlong*, are prone to simplification; compare the stem formations of the following verbs:

- (4) *gerō* “I manage,” perfect *gessī*, supine *gestum*  
*hauriō* “I drain,” perfect *hausī*, supine *haustum*

In both verbs the *-r-* of the present tense comes from earlier *\*s* which changed to *-r-* when in intervocalic position (the unchanged *-s-* is preserved in the stem of the supine, which is formed with a suffix *-tum*). The perfect stem is formed with a suffix *-s-*; the root-final *-s-* is preserved before this in *gessī* with short vowel nucleus, but in *haussī* (still so written in some early texts) the geminate is simplified. A further example of the avoidance of syllables with long nucleus and coda is the regular shortening of long vowels in final syllables before *-t*, *-r*, *-l*, *-nt*, and *-m* which took place in the Early Latin period; the effect of this change can be seen in some of the nominal and verbal paradigms given in the next section.

For the purposes of poetic meter, a syllable with one or more branches in the rhyme is counted as heavy (i.e., a syllable with a long vowel nucleus or final consonant). The relationship between syllables with long nucleus and syllables with coda is also evident from historical developments. When consonant clusters are simplified, preceding short vowels are often lengthened: *īdem* “the same” is a regular development from *\*isdem*. Such *compensatory lengthening* is, of course, paralleled in the histories of many other languages. Comparatively rare is the opposite process, whereby length is transferred from the vowel to the following consonant, but there are several examples of such a change in Latin, a standard example being the divine name *Iuppiter* < *\*Iūpiter* < *\*Dyew-pater* (the change of medial vowel reflects the process of vowel weakening discussed above). There is evidence for similar “length metatheses” throughout the period of Latin, and it is possible that this unusual change has its origin in variant pronunciations in different social registers and subsequent hypercorrection.

Latin words may be built up of one or more admissible syllables, with the following corollaries: (i) no monosyllabic word may consist of a final unchecked short vowel unless it is a clitic; (ii) word-final voiced stops are only found in nonlexicals and /st/ is also permitted as a word-final cluster.

## 4. MORPHOLOGY

### 4.1 Word formation

Latin morphology is widely used in linguistic textbooks as a standard example of the fusional or inflectional type. In Latin nominal morphology, different categories are generally expressed solely through suffixes, which encode both number and case. Thus, the suffixes *-is* and *-bus* are used in different declensional classes as the cumulative exponents of the categories DATIVE PLURAL and ABLATIVE PLURAL, and they cannot be further analyzed into separate morphs for DATIVE or ABLATIVE and PLURAL.

Verbal morphology is also largely fusional, and person, number, and mood marking is always encoded through suffixation, although in many verbal paradigms infixation, reduplication, and ablaut of the root morpheme play a role in the formation of the tense stem.

Some verbal paradigms also approach a degree of analyticity; compare the following examples of present and future passive forms from the root *amā-* “love”:

- (5) *amat* “he loves,” present active indicative third singular  
*amātur* “he is loved,” present passive indicative third singular  
*amābit* “he will love,” future active indicative third singular  
*amābitur* “he will be loved,” future passive indicative third singular

One possible analysis of the future passive *amābitur* would be as follows:

- (6) *amā-*      *-bi-*      *-t-*      *-ur*  
 LOVE    FUTURE    3RD.SINGULAR    PASSIVE

Unfortunately, although this analysis could be made to work for the first and third persons, in the second-person forms the markers of person and voice are fused:

- (7) *amās* “you (singular) love”  
*amāris* “you (singular) are loved”  
*amātis* “you (plural) love”  
*amāmini* “you (plural) are loved”

Latin word structure is of the Indo-European type, the basic scheme of word formation (ignoring, for the moment, compound forms, which will be discussed further below) is as follows:

- (8) word = lexical root + (derivational suffix)<sup>x</sup> + inflectional ending

All well-formed nominals and verbs (barring a handful of indeclinable forms) show a root and ending, and most also incorporate at least one derivational suffix. It is possible to multiply the number of derivational suffixes, and to derive verbs from nominal roots or nominals from verbal roots; thus, for example, *dictātōrius* “belonging to a dictator” can be analyzed as follows:

- |     |              |  |                                 |
|-----|--------------|--|---------------------------------|
| (9) | <i>dic-</i>  | root DIC “say”                         | cf. <i>dīcere</i> “to say”      |
|     | <i>-tā-</i>  | frequentative verb suffix              | cf. <i>dictāre</i> “to dictate” |
|     | <i>-tōr-</i> | agent noun suffix                      | cf. <i>dictātor</i> “dictator”  |
|     | <i>-ius</i>  | fused adjectival suffix and inflection |                                 |

A nonproductive pattern found in a few nouns and verbs attaches inflectional endings directly to the lexical root, as *dux* “leader,” analyzable as lexical root (DUC-) + inflection (-s). Roots and suffixes may show alternations before further derivational suffixes. For example, from the same stem *dictātōr-* mentioned above, a feminine *dictātrīx* “female dictator” is formed, with a regular loss of the medial vowel of the agent suffix *-tōr-* before *-īc-*, the suffix denoting a female.

## 4.2 Nominal morphology

Latin nouns are marked for number and case, and adjectives also for gender.

### 4.2.1 Gender and number

There are three genders, traditionally termed masculine, feminine, and neuter, and two numbers, singular and plural. Gender for nouns denoting humans and gods, and, to a lesser extent, animals overlaps with the semantic criterion of sex, so that *mulier* “woman” is feminine, although it contains no specific feminine morpheme, and *agricola* “farmer,” which belongs to the predominantly feminine declension class with nominative singular in *-a*, is nevertheless masculine. For words designating inanimates, however, all three genders are found: *nix* (feminine) “snow,” *lapis* (masculine) “stone,” *iecur* (neuter) “liver.”

### 4.2.2 Case

The category of case is more complex. Classical Latin has six paradigmatic cases, traditionally labeled nominative, vocative, accusative, genitive, dative, and ablative (note that the traditional term *ablative* is potentially misleading, since this last case also serves as the instrumental and, in part, the locative). In the plural, the dative and ablative are syncretic in all declensions, and all neuter nouns have syncretic nominative, accusative, and vocative. Oscan and Umbrian have a paradigmatic seventh case, the locative; in Latin this is replaced for most nouns by the syntagm of preposition and ablative. However, proper names referring to towns and small islands retain a locative form, as do three nouns denoting place (*rūs*, “countryside,” *domus* “home,” and *humus* “ground”). In Classical Latin the form of the

locative is always syncretic with another case (in the singular, with the genitive in declensions I and II, but with the ablative in declension III; in the plural, with the dative-ablative in all declensions).

### 4.2.3 Nominal declensions

Latin has five morphologically distinct declensions, which largely continue inherited types, with the exception of declension V which represents a Latin innovation. Representative paradigms are given in Table 32.3; note that the declension III has two main subtypes:

Table 32.3 Latin nominal paradigms					
I mēnsa "table"	II lupus "wolf"	IIIa rēx "king"	IIIb turris "tower"	IV manus "hand"	V rēs "thing"
<i>Singular</i>					
Nom. mēnsa	lupus	rēx	turris	manus	rēs
Voc. mēnsa	lupe	rēx	turris	manus	rēs
Acc. mēnsam	lupum	rēgem	turrim	manum	rem
Gen. mēnsae	lupī	rēgis	turris	manūs	reī
Dat. mēnsae	lupō	rēgī	turrī	manuī	reī
Abl. mēnsā	lupō	rēge	turrī	manū	rē
<i>Plural</i>					
Nom. mēnsae	lupī	rēgēs	turrēs	manūs	rēs
Voc. mēnsae	lupī	rēgēs	turrēs	manūs	rēs
Acc. mēnsās	lupōs	rēgēs	turrīs	manūs	rēs
Gen. mēnsārum	lupōrum	rēgum	turrium	manuum	rērum
Dat. mēnsīs	lupīs	rēgibus	turribus	manibus	rēbus
Abl. mēnsīs	lupīs	rēgibus	turribus	manibus	rēbus

Nouns of neuter gender only occur in declensions II, III, and IV. Their case endings are the same except in the syncretic nominative-vocative-accusative:

(10)		II	IIIa	IIIb	IV
<i>Singular</i>		iugum "yoke"	genus "kind"	rēte "net"	genū "knee"
	Nom.-voc.-acc.	iugum	genus	rēte	genū
	Gen.	iugī	generis	rētis	genūs
<i>Plural</i>					
	Nom.-voc.-acc.	iuga	genera	rētia	genua
	Gen.	iugōrum	generum	rētium	genuum

### 4.2.4 Comparatives and superlatives

Adjectives, as well as adverbs, have an additional category of gradation, so that alongside the unmarked "positive" degree of an adjective such as *longus* "long" there are also paradigmatic forms for the comparative *longior* "longer, too long" and a superlative *longissimus* "longest,



very long”; the adverb *longē* “far off” similarly forms comparative *longius* and superlative *longissimē*.

### 4.3 Pronouns

The Latin pronominal system has the same categories as the nominal system. Personal pronouns are marked for number and case, and demonstrative, anaphoric, interrogative, and relative pronouns are marked for number, case, and gender. Apart from the vocative case, which is syncretic with the nominative in all pronominal declensions, the dimensions of each category are the same.

#### 4.3.1 Personal pronouns

The declension system of personal pronouns in Latin is synchronically anomalous as the following paradigms show (note that there is no third-person pronoun other than the reflexive *sē*; oblique forms of the anaphoric pronoun *is*, *ea*, *id* are used to supply the deficiency):

(11)	SINGULAR		PLURAL		
	<i>First</i>	<i>Second</i>	<i>First</i>	<i>Second</i>	<i>Third</i>
	<i>person</i>	<i>person</i>	<i>person</i>	<i>person</i>	<i>person</i>
<i>Nom.</i>	ego	tū	nōs	uōs	
<i>Acc.</i>	mē	tē	nōs	uōs	sē
<i>Gen.</i>	meī	tuī	nostrī	uestrī	suī
			nostrum	uestrum	
<i>Dat.</i>	mihi	tibi	nōbīs	uōbīs	sibi
<i>Abl.</i>	mē	tē	nōbīs	uōbīs	sē

Some of the irregularities of these declensions continue inherited patterns found in other Indo-European languages; others are unique to Latin. A notable development which occurred only in Latin (and, as far as we have evidence, in Faliscan) is the syncretism of accusative and ablative of *ego*, *tū*, and *sē*, which in Early Latin are written *mēd*, *tēd*, and *sēd* (*med* is found as accusative in Faliscan); no other Indo-European language shows accusative forms ending in \*-d for these pronouns. Although Latin lost the distinction between a morphologically separate set of accented and clitic forms, as found in Greek and Sanskrit, it is likely that the personal pronouns could either carry the stress accent or not, depending on context and emphasis.

#### 4.3.2 Demonstrative pronouns

Classical Latin has a fairly rich system of demonstrative and anaphoric pronouns. There is a three-way deictic contrast between the demonstratives *hic*, *haec*, *hoc* “this” (indicating proximity to the speaker); *iste*, *ista*, *istud* “that” (indicating proximity to the hearer); and *ille*, *illa*, *illud* “that” (indicating distance from both speaker and hearer). The declension of these pronouns shows two principal peculiarities: (i) they have distinct endings for all genders in the genitive and dative singular, and for the masculine and neuter nominative; and (ii) some case forms show the amalgamation of pronominal stems with various deictic particles, which have a more independent existence in Early Latin: for example, *hic*, *haec*, *hoc* “this”:

(12)	SINGULAR			PLURAL		
	<i>Masculine</i>	<i>Feminine</i>	<i>Neuter</i>	<i>Masculine</i>	<i>Feminine</i>	<i>Neuter</i>
<i>Nom.</i>	hic	haec	hoc	hī	hae	haec
<i>Acc.</i>	hunc	hanc	hoc	hōs	hās	haec
<i>Gen.</i>	huius	huius	huius	hōrum	hārum	hōrum
<i>Dat.</i>	huīc	huīc	huīc	hīs	hīs	hīs
<i>Abl.</i>	hōc	hāc	hōc	hīs	hīs	hīs

The final *-c*, found in the neuter plural and all cases except the genitive of the singular, derives from an earlier enclitic deictic particle *-ce*, which in Early Latin is also found attached to other forms, as accusative plural masculine *hōsce*. For the two pronouns denoting identity, *ipse*, *ipsa*, *ipsum* “-self” and *īdem*, *eadem*, *idem* “the same,” entirely new paradigms have been generated ultimately from the combinations of the anaphoric *is* with the particles *-pse* and *-em*.

### 4.3.3 Relative, interrogative, and indefinite pronouns

In Latin the same stem is used for relative, interrogative, and indefinite pronouns, which share exactly the same declension outside of the (i) nominative singular, masculine and feminine, and (ii) nominative-accusative singular neuter. The relative pronoun’s declension is as follows:

(13)	SINGULAR			PLURAL		
	<i>Masculine</i>	<i>Feminine</i>	<i>Neuter</i>	<i>Masculine</i>	<i>Feminine</i>	<i>Neuter</i>
<i>Nom.</i>	quī	quae	quod	quī	quae	quae
<i>Acc.</i>	quem	quam	quod	quōs	quās	quae
<i>Gen.</i>	cuius	cuius	cuius	quōrum	quārum	quōrum
<i>Dat.</i>	cuī	cuī	cuī	quibus	quibus	quibus
<i>Abl.</i>	quō	quā	quō	quibus	quibus	quibus

The distinct forms of the relative pronoun ultimately continue the Indo-European stem *\*k<sup>w</sup>o-*, and the distinct forms of the interrogative pronoun, *quis*, *quid*, continue a stem *\*k<sup>w</sup>i-*. The indefinite *quis*, *qua*, *quid*, neuter plural *qua* or *quae*, is originally an *i*-stem also, but it shows a separate feminine form. In Early Latin there is evidence for a more wide-ranging difference between the relative and the interrogative; for instance a distinct dative-ablative plural form *quīs* of the relative pronoun is widely attested. However, it appears that the declensions were confused from an early stage – witness the ubiquity of the accusative singular masculine *quem* as relative pronoun, which must continue *\*k<sup>w</sup>im*.

## 4.4 Verbal morphology

Latin finite verbs are marked for person, number, tense/aspect, mood, and voice. There are three persons and two numbers, singular and plural; three moods occur, indicative, subjunctive, and imperative.

### 4.4.1 Tense and mood

There are six tenses of the indicative and four of the subjunctive, built from two separate stems, here termed the *present stem* and *perfect stem*. The interrelationship between tense and mood is illustrated in (14):

## (14) The interrelationship between tense and mood in Latin

TENSE	MOOD		
	<i>Indicative</i>	<i>Subjunctive</i>	<i>Imperative</i>
<i>Present stem</i>			
<i>Present</i>	Present indic.	Present subjunc.	Present imperative
<i>Imperfect</i>	Imperfect indic.	Imperfect subjunc.	
<i>Future</i>	Future indic.		Future imperative
<i>Perfect stem</i>			
<i>Perfect</i>	Perfect indic.	Perfect subjunc.	
<i>Pluperfect</i>	Pluperfect indic.	Pluperfect subjunc.	
<i>Future perfect</i>	Future perfect indic.		

The Latin verbal system does not grammaticalize aspect to the same degree as some other Indo-European languages, such as Greek and Slavic. The contrast between imperfect, used to indicate an uncompleted or ongoing event, and the perfect, indicating a finished or accomplished event, could be viewed as aspectual, but the notion of aspect is not necessary for the description of the rest of the system. Indeed, it is possible to dispense with aspect as a descriptive category of the Latin verb altogether. The present stem marks states of affairs which take place now (the present), at the same time as some past moment (the imperfect), and at the same time as some future moment (the future). In contrast, the perfect stem tenses mark states of affairs before the present (the perfect), before some specific moment in the past (pluperfect), or in the future (future perfect). The contrast between the imperfect and perfect tense is consequently a contrast between something viewed as contemporaneous with a certain moment in the past, and something viewed as anterior to a moment in the present. The imperfect, therefore, is the appropriate tense to describe ongoing events in the past, and the perfect for completed actions.

## 4.4.2 Voice

The array of tense and mood formations presented above applies equally to the two voices of the verb, active and passive. The active present stem and passive present stem are the same, but perfect formations of passive verbs are always periphrastic, built from the perfect passive participle and auxiliary verb *esse* “to be.” Accordingly, in the perfect passive system, verbs also encode the gender of the subject, as well as the number and person.

## 4.4.3 Conjugation classes

The formation of the different tense and mood paradigms and personal endings is the same for all verbs conjugated in the perfect system, but in the present system there are four main conjugation classes which differ in personal endings and in the formation of the future tense and present subjunctive.

## 4.4.3.1 Present stem system

The conjugation of the present system is illustrated in (15)–(17) using the paradigms of the verbs *amāre* “to love” (Conjugation I); *spondēre* “to pledge” (Conjugation II); *regere* “to rule” (Conjugation III); and *uenire* “to come” (Conjugation IV).

## (15) Latin present indicative

		<i>Conjugation I</i>	<i>Conjugation II</i>	<i>Conjugation III</i>	<i>Conjugation IV</i>
<i>Singular</i>	<i>1st</i>	amō	spondeō	regō	ueniō
	<i>2nd</i>	amās	spondēs	regis	uenīs
	<i>3rd</i>	amat	spondet	regit	uenit
<i>Plural</i>	<i>1st</i>	amāmus	spondēmus	regimus	uenīmus
	<i>2nd</i>	amātis	spondētis	regitis	uenītis
	<i>3rd</i>	amant	spondent	regunt	ueniunt

As can be seen, in the present indicative the four conjugations share the same set of personal endings, which are basically those of Conjugation III, but differences arise from fusion of the endings with stem-vowels in Conjugations I, II, and IV.

## (16) Latin present subjunctive

		<i>Conjugation I</i>	<i>Conjugation II</i>	<i>Conjugation III</i>	<i>Conjugation IV</i>
<i>Singular</i>	<i>1st</i>	amem	spondeam	regam	ueniam
	<i>2nd</i>	amēs	spondeās	regās	ueniās
	<i>3rd</i>	amet	spondeat	regāt	ueniat
<i>Plural</i>	<i>1st</i>	amēmus	spondeāmus	regāmus	ueniāmus
	<i>2nd</i>	amētis	spondeātis	regātis	ueniātis
	<i>3rd</i>	ament	spondeant	regant	ueniant

In the subjunctive, the personal endings are the same as those of the indicative (15) except for the first-person singular, which is marked by *-m* rather than *-ō*. These two endings are a survival of a much more pervasive system of differentiation of primary (= non-past, non-subjunctive) and secondary (= either + past, or + subjunctive, or both) endings, which is more widely attested in some Early Latin texts. The distribution of the *-m* and *-ō* morphemes is still largely governed by the original primary/secondary distinctions, except in the future indicative.

## (17) Latin future indicative

		<i>Conjugation I</i>	<i>Conjugation II</i>	<i>Conjugation III</i>	<i>Conjugation IV</i>
<i>Singular</i>	<i>1st</i>	amābō	spondēbō	regam	ueniam
	<i>2nd</i>	amābis	spondēbis	regēs	ueniēs
	<i>3rd</i>	amābit	spondēbit	regēt	ueniet
<i>Plural</i>	<i>1st</i>	amābimus	spondēbimus	regēmus	ueniēmus
	<i>2nd</i>	amābitis	spondēbitis	regētis	ueniētis
	<i>3rd</i>	amābunt	spondēbunt	regent	uenient

For this future, not only is there a difference in the first-person singular morpheme in Conjugations III and IV as opposed to Conjugations I and II, but there is also a radically different stem formation. The future in *-bō* of Conjugations I and II matches a formation found in Faliscan, which has a future formed in *-fo* (inscriptional *pipafo* “I will drink” and *carefo* “I will lack”), but which does not have a clear Indo-European origin.

Note that there is also a subclass of Conjugation III of the type *facere* “to do,” which forms present indicative *faciō* and *faciunt*; subjunctive *faciam* *faciās* and so forth; future *faciam* *faciēs* and so on.

The other present stem tense and mood forms of *amō* are as follows:

(18) <i>Imperfect indicative</i>	<i>amāb</i> -am, -ās, -at, -āmus, -ātis, -ant
<i>Imperfect subjunctive</i>	<i>amār</i> -em, -ēs, -et, -ēmus, -ētis, ent
<i>Imperative I</i>	<i>amā</i> , <i>amāte</i>
<i>Imperative II</i>	<i>amātō</i> , <i>amātō</i> , <i>amātōte</i> , <i>amantō</i>

In the imperative paradigms there is no form for the first person and the forms given in (18) are respectively (i) second singular and plural for the imperative I; and (ii) second singular, third singular, second plural, and third plural for imperative II. In the second person there is consequently a difference between two different imperative forms. This is not a difference of aspect, but rather one of relative tense. Where the two forms are used in conjunction, the future imperative (imperative II) is used to refer to an event following the present imperative – note, for example, the following commands from Plautus' play *Pseudolus*:

- (19) *cape hās tabellās, tūte hinc narrātō tibi*  
 “Take these tablets and find out for yourself from them”

*Cape*, present imperative “take!” refers to the initial action, and *narrātō*, future imperative “tell!,” refers to an action consequent on this, reading what is written on the tablets.

#### 4.4.3.2 Perfect stem system

The perfect stem is generally distinguished from the present in one of four ways: (i) through the addition of a suffix (-s- or -u-), (ii) through reduplication of the initial consonant or consonant cluster of the root syllable, (iii) through change (usually lengthening) of the nucleus of the root syllable, or (iv) through suppletion. One class of perfects, however, has stems which are identical to those of the present. The perfect is further marked by a special set of personal endings in the perfect indicative. As examples of the different types of perfect formation and the endings, the perfects of the four verbs considered above are presented in (20). The perfect of *amāre* is formed with the suffix -u- (*amāu*-), that of *spondere* by reduplication (*spopond*-), that of *regere* with the suffix -s- (*rēx*-), and that of *uenire* by vowel lengthening (*uēn*-):

#### (20) Latin perfect indicative

		<i>Conjugation I</i>	<i>Conjugation II</i>	<i>Conjugation III</i>	<i>Conjugation IV</i>
<i>Singular</i>	<i>1st</i>	<i>amāu</i> -ī	<i>spopond</i> -ī	<i>rēx</i> -ī	<i>uēn</i> -ī
	<i>2nd</i>	<i>amāu</i> -istī	<i>spopond</i> -istī	<i>rēx</i> -istī	<i>uēn</i> -istī
	<i>3rd</i>	<i>amāu</i> -it	<i>spopond</i> -it	<i>rēx</i> -it	<i>uēn</i> -it
<i>Plural</i>	<i>1st</i>	<i>amāu</i> -imus	<i>spopond</i> -imus	<i>rēx</i> -imus	<i>uēn</i> -imus
	<i>2nd</i>	<i>amāu</i> -istis	<i>spopond</i> -istis	<i>rēx</i> -istis	<i>uēn</i> -istis
	<i>3rd</i>	<i>amāu</i> -ērunt	<i>spopond</i> -ērunt	<i>rēx</i> -ērunt	<i>uēn</i> -ērunt

Representative examples of perfects formed through suppletion, and the perfect with unchanged stem, can also be added: *ferō* “I carry,” perfect *tulī*; *bibō* “I drink,” perfect *bibī*. The third plural ending -*ērunt* of Classical Latin probably represents a conflation of two competing morphs -*ēre* and -*erunt*, both well attested in Early Latin and still used in later archaizing texts.

The remaining tenses and moods of the perfect stem take either the secondary endings found in the present system or, in the case of the future perfect, the primary endings of the present system, as sketched out in the examples from *amāre* given below:

(21) <i>Future perfect</i>	amāuer-o, -is, -it, -imus, -itis, -int
<i>Pluperfect</i>	amāuer-am, -ās, -at, -āmus, -ātis, -ant
<i>Perfect subjunctive</i>	amāuer-im, -is, -it, -imus, -itis, -int
<i>Pluperfect subjunctive</i>	amāuiss-em, -ēs, -et, -ēmus, -ētis, -ent

In Early Latin the future perfect and perfect subjunctive were better distinguished, since the perfect subjunctive showed a long vowel in the ending, *amāueris*, *amāuerit*, and so forth.

#### 4.4.3.3 *Passive voice marking*

The Latin passive is marked against the active morphologically and semantically. For the morphological marking, compare the third singular present indicative active *amat* with the passive *amātur*. Semantically the prototypical use of a verb in the passive is to promote the object of a transitive verb to subject position: active *Caesar amat Cicerōnem* “Caesar loves Cicero”; passive *Cicerō amātur* “Cicero is loved.” However, third singular forms of the passive of intransitive verbs can also be used impersonally (the so-called *impersonal passive*): active *Caesar it* “Caesar goes”; passive *itur* “a journey is made” (lit. “it is gone”).

A large number of verbs (termed *deponents* in traditional grammar) only show passive morphology of finite forms but are not semantically passive. Many of these correspond to middle or reflexive verbs in other languages: thus, *irāscor* “I become angry,” *utor* “I use,” *reor* “I think,” and *morior* “I die.” Deponent verbs do not have separate active paradigms, but they do use some nonfinite active forms (see §4.4.4), such as the present and future participles; note also that the gerundive of deponent verbs is semantically passive.

As discussed above, the present passive system is formed using the same stem as the present active system. Thus, for *amō* “I love” the passive present indicative, present subjunctive and future indicative are as follows:

#### (22) Latin passives of the present system

<i>Present indicative</i>	amor, amāris, amātur, amāmur, amāminī, amantur
<i>Present subjunctive</i>	amer, amēris, amētur, amēmur, amēminī, amentur
<i>Future indicative</i>	amābor, amāberis, amābitur, amābimur, amābiminī, amābuntur

The perfect stem of the passive is different from the perfect active stem, and all moods and tenses are formed through periphrasis with the perfect passive participle and present stem forms of the copula verb *esse*. The perfect passive indicative and subjunctive of *amō* are given for illustration (in all persons the forms agreeing with a masculine subject are given; agreement for a feminine or neuter subject would be different):

#### (23) Latin passives of the perfect system

<i>Perfect indicative</i>	amātus sum, amātus es, amātus est, amātī sumus, amātī estis, amātī sunt
<i>Perfect subjunctive</i>	amātus sim, amātus sīs, amātus sit, amātī sīmus, amātī sītis, amātī sint

#### 4.4.3.4 *Diachronic developments*

In the subliterate registers of Late Latin, the complex tense and mood system of Latin undergoes many changes, and the end result of these is reflected in the modern Romance languages. The most pervasive changes are the increasing spread of periphrastic formations at the expense of synthetic paradigms. A striking example is the future indicative. In all

languages ways of referring to future events are prone to remarking with more direct or vivid constructions and, as we saw earlier, the formation of the Classical Latin future is anomalous, with different exponents found in different conjugations. It is therefore no surprise that the synthetic formation of the future becomes increasingly marked in Late Latin and is eventually completely replaced in Romance languages.

#### 4.4.4 Nonfinite verbals

The nonfinite verb system is less orderly than the finite. There is a present and future participle active, and a perfect participle passive; the present participle is formed from the present stem, but the future participle is generally formed from the same stem as the perfect passive participle (which, following Aronoff, I shall call the *t*-stem), even where the verb is suppletive. A future passive participle, denoting necessity or obligation, and termed the *gerundive* in traditional grammar, is also formed from the present stem. Thus, *ferō* “I carry” has a present participle *ferēns*, and gerundive *ferendus*, *-a*, *-um*, but a future active participle *lātūrus*, *-a*, *-um* and perfect passive participle *lātus*, *-a*, *-um*.

There are six tense- and voice-marked infinitives, of which three – present active, perfect active, and present passive – are synthetic while the others are periphrastic: (i) the future active infinitive = future active participle + *esse* “to be”; (ii) the perfect passive infinitive = perfect passive participle + *esse* “to be”; and (iii) the future passive infinitive, made through the curious periphrasis of the supine (on which see below) + *irī*, the passive infinitive of *eō* “I go.”

There are also two defective verbal nouns: the first, traditionally called the *gerund*, is in form identical to the neuter singular forms of the gerundive and provides the oblique cases to the present infinitive active. The second, the *supine*, also has active meaning and is formed from the *t*-stem and has two distinct forms (originally case forms) *-um* (thus *lātum* from *ferō*) and *-ū* (*lātū*). In Classical Latin the *-um* supine is only used as an optional means of expressing purpose clauses after verbs of motion (for example, *spectātum ueniunt* “they come to watch”) and the *-ū* supine is used after certain adjectives (for example, *mīrābile dictū* “amazing to describe”).

### 4.5 Compounds

Nominal compounding is a productive process of word formation in Latin. However, at the earliest stage of the language, there was only a relatively small number of compounds. Most of these are either (i) exocentric compounds with a numeral or negative element as first member, such as *bi-dēns* “sacrificial animal” (lit. “having two teeth,” formed from the prefix *bi-* “two” and *dēns* “tooth”), *in-ermis* “safe” (lit. “without weapons,” formed from the negative prefix *in-* and *arma* “weapons”); or else they are (ii) verbal-governing compounds of the type of *ponti-fex* “priest” (lit. “one who makes a bridge,” formed from *pōns* “bridge” and a verbal noun from the root of *faciō* “I make”), or *rēm-ex* “oarsman” (lit. “one who drives the oar,” from *rēm* “oar” and a verbal noun from the root of *agō* “I drive”). The huge influence of Greek literary texts led to a revival of compounding in Latin, and many new compounds and new types of compounding are found in works of all periods under Greek influence, many of them calques of actual Greek compounds.

Alongside “true” compounding, a number of quasi-compounds are found in Latin of all periods through juxtaposition and univerbation of adjective and noun, or noun and dependent genitive. Examples include the following: *rēs-publica* “republic” (*rēs* “affairs,” *publica* “public”); *pater-familiās* “head of the household” (*pater* “father,” *familiās*, continuing an archaic genitive form, “of the household”); and *aquae-ductus* “aqueduct” (*aquae* “of water,” *ductus* “conveyance”).



Verbal compounds are nearly all of the type preverb + verb: for example, *re-ficere* “to restore” (*re-* “back, again” and *facere* “to make”), *inter-currere* “to run between” (*inter-* “between” and *currere* “to run”). Preverbs were originally independent adverbial elements, and most preverbs have adverbial or prepositional counterparts; compare *retrō* “backwards” and *inter* “between.” There are a small number of compound verbs of the type adverb + verb, noun + verb, and verb + verb. Where these occur they generally arise out of earlier juxtapositions, such as *animaduertere* “to notice” (from *animam aduertere* “to turn one’s attention”), and *ne-scīre* “not to know” (from *ne* “not” and *scīre* “to know”). The analysis of a small class of verbal compounds of the type *cale-facere* “to make warm” is a long-standing problem of Latin linguistics; the second element is clearly the verb *facere* “to make,” but it is disputed whether the first element derives from the verb *calēre* “to be warm.”

## 4.6 Numerals

The only declined Latin numerals are the following: *ūnus*, -a, -um 1 (masculine and neuter genitive *ūnī* and feminine genitive *ūnae*); *duo*, -ae, -o 2 (genitive *duōrum* -ārum); *trēs*, *tria* 3 (genitive *trium*); and the terms for multiple hundreds, *ducentī*, -ae, -a 200 (genitive *ducentōrum* -ārum), *trecentī*, -ae, -a 300 (genitive *trecentōrum* -ārum), and so forth, which all decline like adjectives in concord with their head noun. The cardinals from 4 through 10 are as follows:

(24)	4	quattuor	8	octō
	5	quīnque	9	novem
	6	sex	10	decem
	7	septem		

These and all other numbers are invariable, with the exception of the word for “1,000,” *mille*, which is indeclinable in the singular but has a declined plural *mīlia* after which the head noun is placed in the genitive plural.

The numeral system is decimal; higher numerals are formed through combination of thousands, hundreds, decads, and units. Noteworthy are the numbers 18 and 19 which are formed through a subtractive system, *duodēuīgintī* “18” (literally “2 from 20”) and *undēuīgintī* “19” (“1 from 20”).

Ordinals are declined as adjectives, the masculine and neuter forms having second declension inflection, the feminine having first declension (see §4.2.3). The ordinal numbers “first” through “tenth” are presented in (25):

(25)	<i>First</i>	<i>prīmus</i>	<i>Sixth</i>	<i>sextus</i>
	<i>Second</i>	<i>secundus</i>	<i>Seventh</i>	<i>septimus</i>
	<i>Third</i>	<i>tertius</i>	<i>Eighth</i>	<i>octāuus</i>
	<i>Fourth</i>	<i>quārtus</i>	<i>Ninth</i>	<i>nōnus</i>
	<i>Fifth</i>	<i>quīntus</i>	<i>Tenth</i>	<i>decimus</i>

## 5. SYNTAX

### 5.1 Word order

Classical Latin does not have an obligatory word order, and in some Latin literary works metrical and stylistic considerations lead to considerable variation in word order with scrambling of words belonging to separate constituent phrases. In the following line of Vergil (*Aeneid* I.109) there is an extreme example of displacement of the relative pronoun, which occurs

after the subject and verb of its clause and interrupts a prepositional phrase, which is itself in a nonstandard order:

- (26) saxa                vocant            Itali        mediis    quae                in   fluctibus   arās  
       rock                call                Italian    middle    which                in   wave        altar  
       ACC.PL.NEUT.   3RD PL.PRES.   NOM.PL.   ABL.PL.   ACC.PL.NEUT.        ABL.PL.   ACC.PL.  
       “Rocks in the middle of the waves which the Italians call ‘The Altars’”

In this sentence the word order is clearly highly marked and artificial, but such sentences could still be understood by Roman audiences.

There is, however, a preferential (“unmarked”) order of constituents observable in Classical Latin prose. In sentences the order is normally Subject–Object–Verb (SOV), but other unmarked patterns are of the “head-first” rather than “head-final” type. Thus, Latin typically has prepositions, not postpositions; and adjectives (except for subjective adjectives of the type *bonus* “good,” etc.) usually follow the head noun, as do relative clauses. It appears that Latin is in a transitional phase from a “head-final” to a “head-first” language; and it is certainly true that later texts increasingly show a preponderance of SVO-type sentences, while there is evidence for an earlier unmarked pattern for adjectives preceding the noun, and postpositions (retained in some phrases such as *mē-cum*, lit. “me.ABL.-with,” that is, “with me”). It is not clear, however, whether the verb-final preference of Classical prose is a wholly artificial, archaizing construct or whether it does reflect certain registers of speech.

Since Latin word order is not obligatory, emphatic positions in the sentence may be taken by any constituent which needs to be highlighted for pragmatic reasons. The position of focused elements also interacts with the word-order rule termed *Wackernagel's Law*, whereby unstressed elements occupy the second position in their clause. In Classical Latin this rule was reinterpreted with the effect that focused elements, whether initial or not, became preferential hosts for unstressed elements of different types: particles, some personal pronouns, and the copula verb *esse*. As illustration, consider the following sentence from Caesar (*Bellum Gallicum* 1.44.8; note that there is ellipsis of the main verb, for which I have supplied “he said”):

- (27) prōvinciam   suam                hanc                esse    Galliam,        sicut    illam  
       province        his                this                to be    Gaul                just as    that  
       ACC.SG.FEM.   ACC.SG.FEM.   ACC.SG.FEM.   INF.   ACC.SG.FEM.                ACC.SG.FEM.  
       nostram  
       our  
       ACC.SG.FEM.  
       “[He said that] this Gaul was his province, just as that [Gaul was] ours”

In this clause, the clitic *esse* splits the constituent *hanc Galliam*. The placement of *esse* reflects the fact that *hanc*, standing in antithesis to *illam*, is emphasized.

## 5.2 Subordination

Classical Latin has a number of different subordinate clause types and subordinating procedures. It is likely that at the earliest period of Latin, subordination was a less important phenomenon; but already by the beginning of the second century BC, official Latin prose inscriptions show a highly developed system of subordination. Latin subordinate clauses can be formed with or without an explicit subordinator, but where a subordinator is present the subordinating verb must be marked as indicative or subjunctive (the imperative is

sometimes found when the main verb is also imperative). Where a subordinator is not present the subordinate clause is marked either through the use of the subjunctive mood or through one of the nonfinite verb forms (infinitive, participle, gerundive, gerund, or supine; see §4.4.4).

The following sentences are given to show some of the range of subordinate types; they are all taken from Classical Latin prose works:

- (28) arma                      capiās                      oportet  
 arms                      you take                      it is necessary  
 ACC.PL.NEUT.    2ND.SG. PRES.SUBJUNC.    3RD.SG.PRES.INDIC.  
 “You ought to take up arms”

In (28) there is no subordinator; the subordinate clause is marked solely through the use of the subjunctive; this construction is largely restricted to sentences where there is a simple modal predicate.

- (29) ingemescunt              nōn    quod              doleant              sed  
 they groan              not    because              they are in pain              but  
 3RD.PL.PRES.INDIC.              SUBORDINATOR              3RD.PL.PRES.SUBJUNC.  
 quia              omne              corpus              intenditur  
 because              all              body              is stretched  
 SUBORDINATOR    NOM.SG.NEUT.    NOM.SG.    3RD.SG.PRES.INDIC.PASS.  
 “They groan, not because they are in pain, but because their whole body is stretched”

In (29) two parallel subordinate clauses show different moods (subjunctive in the first clause, indicative in the second), because the first clause describes a potential or alleged cause, and the second the actual cause.

- (30) eum              hominem    occīdendum              cūrāuit  
 this              man              to be killed              he arranged  
 ACC.SG.MASC.    ACC.SG.    ACC.SG.MASC.GDVE    3RD.SG.PERF.INDIC.  
 “He arranged for this man to be killed”

In (30) the gerundive is used to mark the embedded clause. Note that this construction with the verb *cūrō* “I arrange” is found principally with the gerundive, and never with the future active participle.

In some genres of Classical Latin prose there is a marked preference for so-called “periodic” sentences, which comprise a number of subordinate and coordinate clauses, often featuring several layers of embedding. To a large extent this is an artificial device, but it is facilitated by rules for the tense marking of subjunctives in dependent clauses. Subjunctives encode both the tense of the matrix clause and the tense of the dependent clause according to the following system, known in traditional grammar as the rules for *Sequence of Tense*:

(31) Latin sequence of tense

<i>Tense of matrix clause</i>	<i>Tense of dependent clause</i>		
	<i>Past</i>	<i>Present</i>	<i>Future</i>
<i>Non-past</i>	Perfect subjunctive	Present subjunctive	Periphrastic future (-ūrus sim, sīs, etc.)
<i>Past</i>	Pluperfect subjunctive	Imperfect subjunctive	Periphrastic future (-ūrus essem, essēs, etc.)

The rich array of nonfinite verbal forms discussed above (see §4.4.4) also play an important role in subordination, as we have already seen. The most important of these verbal forms is the infinitive, which not only is used after “control” verbs such as *uolō* “I want,” *incipiō* “I begin,” *cōnor* “I try,” but which also plays a major role after other predicate types, most importantly in the construction of reported speech. The subject of the infinitive when used in this way is usually in the accusative, hence giving the traditional name of *Accusative and Infinitive* (AcI) construction. The syntactic value of the subject of the infinitive is interesting since it behaves as if it were an argument of the matrix clause. Thus if the subject of the AcI clause is also the subject of the matrix clause, the equivalence is marked through the reflexive pronoun *sē*:

- (32) 

dixit	sē	librōs	eōs	in	ignem
he said	himself	books	those	in	fire
3RD.SG. PERF.INDIC.	ACC.SG.	ACC.PL.	ACC.PL.		ACC.SG.
coniectūrum		esse			
will be throwing		to be			
ACC.SG.MASC. FUT.ACT.PART.		PRES.INF.			

  
 “He said that he would throw the books in the fire”

Note the formation of the future infinitive (*coniectūrum esse*) through a combination of the future participle and auxiliary, and that the participle is marked for agreement (see §5.3) with *sē*.

The subject of the infinitive can also be raised to be subject of the verb of the matrix clause, if the verb would otherwise be an impersonal passive:

- (33) 

dicitur	Appius	itā	precātus	esse
is said	Appius	thus	having prayed	to be
3RD.SG.PRES. INDIC.PASS.	NOM.SG.MASC.		NOM.SG.MASC.PERF.PART.	PRES.INF.

  
 “It is said that Appius prayed in this way”

In this example the subject of the infinitive is also subject of the matrix clause and consequently takes nominative case marking, as does the participle of the periphrastic infinitive; *dicitur* consequently behaves as a “pseudo-control” verb.

### 5.3 Agreement

The rich nominal and verbal morphology of Latin is dependent on a strict system of agreement. Adjectives show the same number, gender, and case as their head noun and verbs agree in number and person with their subjects. When different subjects of a verb are conjoined the verb shows agreement according to the person hierarchy *first* > *second* > *third*; hence a verb with first- and second-person subjects will normally take first-person endings.

Agreement patterns in conjoined nominal phrases are more interesting: in phrases where there is a single adjective but two conjoined nouns of different genders, the adjective will either be marked for agreement with the closest noun, or will be marked masculine or neuter. The choice between masculine and neuter is governed by the animacy of the referents: if the two conjoined nouns refer to animates, adjectives take the masculine; if inanimates, adjectives are marked neuter. For example (Livy *Ab urbe conditā* 32.29.1):

- (34) 

mūrus	et	porta	dē	caelō	tācta	erant
wall	and	gate	from	sky	touched	had been
NOM.SG.MASC.		NOM.SG.FEM.		ABL.SG.MASC.	NOM.PL.NEUT.	3RD PL.IMPF.

  
 “The wall and gate had been touched from the sky”

Here the participle *tācta* is assigned neuter gender, although the two conjoined nouns to which it refers are respectively masculine and feminine.

## 6. LEXICON

The Latin lexicon has long been thought to be highly conservative, and it does retain a number of roots which are only found in a few other Indo-European languages. Thus the word for “believe” *crēdō*, which developed from a periphrasis of the words for “heart” and “put,” is found elsewhere only in Indo-Iranian and Celtic languages. A word for “drunk” *ēbrius* may continue a root meaning “drink” attested only in Hittite and Tocharian. The verb *spondeō* “I pledge” only has cognates in Hittite, Greek, and Tocharian. Latin *rēs* “property” is matched by forms in Umbrian and Indo-Iranian alone. The pair *hostis* “stranger, enemy” and *hospēs* “guest, host” continue forms also found in the Slavic languages and (in the case of *hostis*) Germanic.

However, the Latin vocabulary also contains a number of loanwords; some of these – for example, *ficus* “fig,” *citrus* “citron-tree,” *menta* “mint,” and *cupressus* “cypress-tree” – almost certainly represent borrowings from lost Mediterranean languages. For others, we can identify the source language involved with more certainty. A number of Latin words were borrowed from the neighboring Faliscan and Sabellian languages, and these may be identified through their distinctive phonology or non-Latin phonological developments: *bōs* “ox, cow,” *popīna* “cook-shop,” *lacrima* “tear,” *uafer* “wily,” *rōbus* “red,” and *rūfus* “red.” Although these loans are often said to have come from Sabine, a Sabellian language for which we have very little direct evidence, it is possible that they represent borrowings from different languages at different times: note that the loanwords *rōbus* and *rūfus* both continue the same original formation, presumably borrowed through different Italic sources. It is also possible that some of these terms are actually derived from dialectal varieties of Latin.

A second source for the enrichment of the Latin vocabulary was Etruscan, although here too we run into problems of identification of individual words owing to lack of available evidence on the Etruscan lexicon. Some words, such as *histrīō* “actor,” are explicitly stated as Etruscan by ancient authorities, and Etruscan is also the most likely origin for others such as *persōna* “mask.” It is often supposed that many words which derive ultimately from Greek were loaned first from Greek to Etruscan, and from there into Latin. Unfortunately, in most cases this is difficult to prove since the Etruscan evidence is lacking, but the representation of Greek voiced stops as Latin voiceless stops suggests that there was an Etruscan intermediate stage in some words, such as *catamītus* “catamite” (borrowed from the Greek name Γανυμήδης) and *sporta* “basket” (Greek σπύριδα accusative singular).

Greek civilization was the dominant cultural influence on Rome throughout much of its history, and it is no surprise that the greatest influence on the Latin lexicon was from Greek. Greek loanwords entered the language from the very earliest stages, not only denoting the material objects and professions which were associated with Greek trade – for example, *mācina* “crane,” *nauta* “sailor,” *ancora* “anchor” – but also reflecting the influence of Greek culture in all areas of civilized life: *balneum* “bath,” *poena* “punishment,” *camara* “ceiling,” *poēta* “poet.” In the Classical period Latin writers looked to Greek models to expand what they saw as the poverty of expression in their native tongue, and through direct borrowing and widespread calquing the Latin lexicon was enormously expanded, and new technical vocabularies were created in many fields, including grammar, rhetoric, philosophy, and medicine. The Latin Bible translations and early Christian works also incorporated many Greek terms from their exemplars, and Christian terms such as *angelus* “angel,” *diabolus*

“devil,” and *presbyter* “priest” have entered into many Western European languages from Greek via Latin.

## Bibliography

### *Introduction, grammars and collections of articles*

- Baldi, P. 1999. *The Foundations of Latin*. Berlin: Mouton de Gruyter.
- Coleman, R. 1987. “Latin and the Italic languages.” In B. Comrie (ed.), *The World's Major Languages*, pp. 180–202. Oxford: Oxford University Press.
- Dangel, J. 1995. *Histoire de la langue latine*. Paris: Presses Universitaires de France.
- Giacomelli, G. 1963. *La lingua falisca*. Florence: Leo S. Olschki.
- Giacomelli, R. 1993. *Storia della lingua latina*. Roma: Jouvence.
- Leumann, M. 1977. *Lateinische Laut- und Formenlehre. Lateinische Grammatik*, vol. I (7th edition). Munich: Beck.
- Meiser, G. 1998. *Historische Laut- und Formenlehre der lateinischen Sprache*. Darmstadt: Wissenschaftliche Buchgesellschaft.
- Palmer, L. 1954. *The Latin Language*. London: Faber and Faber.
- Panagl, O. and T. Krisch (eds.). 1992. *Latein und Indogermanisch: Akten des Kolloquiums der Indogermanischen Gesellschaft, Salzburg, 23.–26. September 1986*. Innsbruck: Institut für Sprachwissenschaft der Universität Innsbruck.
- Vineis, E. 1998. “Latin.” In A. Ramat and P. Ramat (eds.), *The Indo-European Languages*, pp. 261–321. London/New York: Routledge.

### *Script*

- Wachter, R. 1987. *Altlateinischen Inschriften: sprachliche und epigraphische Untersuchungen zu den Dokumenten bis etwa 150 v. Chr.* Bern: Peter Lang.
- Wallace, R. 1989. “The origins and development of the Latin alphabet.” In W. Senner (ed.), *The Origins of Writing*, pp. 121–135. Lincoln, NE: University of Nebraska Press.

### *Phonology*

- Allen, W. 1978. *Vox Latina* (2nd edition). Cambridge: Cambridge University Press.
- Cser, A. 1996. “Latin syllable structure.” *Working Papers in the Theory of Grammar* 3:35–47.
- Zirin, R. 1970. *The Phonological Basis of Latin Prosody*. The Hague: Mouton.

### *Morphology*

- Aronoff, M. 1994. *Morphology by Itself: Stems and Inflectional Classes*. Cambridge, MA: MIT Press.
- Coleman, R. 1992. “Italic.” In J. Gvozdanović (ed.), *Indo-European Numerals*, pp. 389–445. Berlin/New York: Mouton de Gruyter.
- Ernout, A. 1953. *Morphologie historique du latin*. Paris: Klincksieck.
- Kent, R. 1946. *The Forms of Latin*. Baltimore: Linguistic Society of America.
- Neue, F. and C. Wagener. 1892–1905. *Formenlehre der lateinischen Sprache* (3rd edition, 4 vols.). Leipzig: Reisland.

### *Syntax*

- Adams, J. 1976. “A typological approach to Latin word order.” *Indogermanische Forschungen* 81:70–99.
- Calboli, G. (ed.). 1989. *Subordination and Other Topics in Latin: Proceedings of the Third Colloquium on Latin Linguistics, Bologna, 1–5 April 1985*. Amsterdam/Philadelphia: John Benjamins.
- Hofmann, J. and A. Szantyr. 1965. *Lateinische Syntax und Stilistik. Lateinische Grammatik*, vol. II. Munich: Beck.

- 
- Kühner, R. and C. Stegmann. 1966. *Ausführliche Grammatik der lateinischen Sprache. Zweiter Teil: Satzlehre* (2 vols.). Hannover: Hahn.
- Pinkster, H. 1990. *Latin Syntax and Semantics*. Translation of *Latijnse syntaxis en semantiek* by H. Mulder. London: Routledge.

*Dictionaries and lexical studies*

- Biville, F. 1990–1995. *Les emprunts du latin au grec: approche phonétique*. Louvain: Peeters.
- Glare, P. 1982. *Oxford Latin Dictionary*. Oxford: Clarendon Press.
- 1900–. *Thesaurus linguae Latinae*. Leipzig: Teubner.



# Sabellian languages

REX E. WALLACE

## 1. HISTORICAL AND CULTURAL CONTEXTS

The term “Sabellian” refers to a group of genetically related languages that were spoken throughout a substantial portion of pre-Roman Italy. Oscan and Umbrian are considered the major representatives of this group because they are attested by the largest corpora of inscriptions. The former was spoken in the southern half of the Italian peninsula, in the territories of Samnium, Campania, Lucania, and Bruttium; the latter was spoken east of the Tiber River in Umbria. Other Sabellian languages include Paelignian, Marrucinian, Vestinian, Marsian, Volscian, Hernican, Aeolian, and Sabine—languages which were spoken in central Italy in the hill districts lying east and southeast of Rome. Recently, South Picene, a language spoken in southern Picenum and in northern Samnium, and Pre-Samnitan, the language of Sabellian peoples who inhabited southern Campania before the arrival of the Oscan-speaking Samnites, have been added to the inventory of Sabellian tongues.

Archeological evidence has not yet shed sufficient light on the dates at which or the routes by which, Sabellian speakers moved into the Italian peninsula. By the beginning of the historical period (c. 700 BC), however, Sabellian speakers had spread over a considerable portion of central Italy, from Umbria and Picenum in the northeast to the Sorrentine peninsula in the southwest (see Map 4). Sabellian tribes were still on the move during the fifth and fourth centuries. Roman historical sources document the invasion of Campania and the capture of Capua, Cumae, and Paestum by Oscan-speaking Samnites. By the middle of the fourth century they had pushed south into Lucania and Bruttium, and southeast into Apulian territory. At the beginning of the third century there were Oscan speakers in Sicily. The Mamertini, a band of mercenary soldiers, crossed the straits in 289 BC and wrested control of the Sicilian city of Messana from the Greeks.

The Sabellian languages did not survive Roman expansion. Those languages spoken in central Italy succumbed to Romanization earlier than did those in the north and south. Sabellian speakers in central Italy had probably shifted to Latin before the end of the Roman Republic (c. 30 BC). In some areas Sabellian was more tenacious. Evidence from the city of Pompeii indicates that Oscan was still being spoken there when the city was destroyed by Vesuvius in AD 79. However, it is unlikely that any Sabellian language survived much beyond the first century AD, by which time the territories of the Sabellians were securely incorporated into the Roman Empire both politically and culturally.

The Sabellian languages are documented almost exclusively by inscriptions. The texts belong to standard epigraphical types: dedications, epitaphs, proprietary inscriptions, inscriptions on public works, religious regulations, contracts, curses, trademarks, legends on coins, and so forth. A few Sabellian vocabulary items are preserved by Roman and Greek

writers of the late Republic and early imperial period, but they do not add substantially to our knowledge of any Sabellian language (Vetter 1953:362–378).

Oscan owns the largest corpus of texts, approximately six hundred and fifty inscriptions. They cover a span of six hundred years, from the sixth century BC to the first century AD. Most of the inscriptions belong to the period between 300 and 89 BC, the latter being the date of the final Sabellian uprising against Rome. The nucleus of the corpus, over 30 percent of the texts, comes from the Campanian cities of Capua and Pompeii. One of the most important Oscan inscriptions was also discovered in Campania, the so-called *Cippus Abellanus*, a limestone plaque recording an agreement between the cities of Nola and Abella regarding the common use of a sanctuary of Heracles. The longest Oscan text, the *Tabula Bantina*, is from the Lucanian town of Bantia. This bronze tablet is incised with a list of statutes concerning municipal administration.

Even though the number of Umbrian inscriptions does not exceed forty, the corpus is one of the most important in ancient Italy. Umbrian was the language of the *Tabulae Iguvinae* (Iguvine Tablets), seven bronze tablets that were discovered in Gubbio (Roman Iguvium) in the fifteenth century. The tablets were incised with the ritual regulations and cultic instructions of a religious fraternity, the Atiedian brotherhood. They date from the first half of the third century (for Tablets I–Vb7) to the end of the second century (for Tablets Vb8–VII). Despite the relative lateness of these texts, it is likely that many of the ritual procedures and regulations stem from an earlier tradition (see Rix 1985).

The remaining Sabellian languages are much less well represented. For most, there are only a few short and often fragmentary inscriptions.

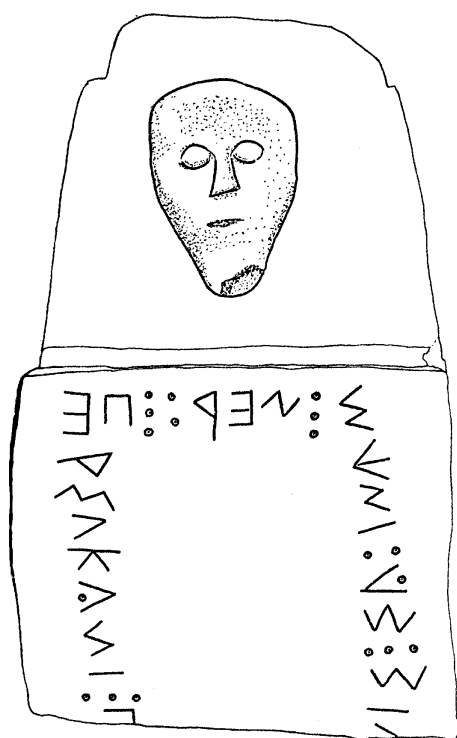
Examples of Sabellian inscriptions are given below (Figs. 33.1–2). According to standard epigraphical conventions, texts written in native Sabellian alphabets are transcribed in bold-face type; texts written in a Republican Latin alphabet appear in italics. The *editio minor* of Sabellian inscriptions is Rix 2002. Vetter 1953 and Poccetti 1979 remain invaluable for epigraphic and linguistic commentary. An *editio maior* of the *Tabulae Iguvinae* was published by Prosdocimi in 1984. Shorter Umbrian texts are collected in Rocca 1996. Marinetti 1985 is the *editio maior* for South Picene inscriptions.

The Sabellian languages, together with Latin and Faliscan, belong to the Italic branch of the Indo-European language family. The evidence for an Italic subgroup consists of three significant morphological innovations that are shared exclusively by Sabellian and Latino-Faliscan:

#### (1) Innovations shared by Sabellian and Latino-Faliscan

- A. Imperfect subjunctive suffix *\*-sē-*, e.g., Oscan **fusíd** “should be” 3RD SG. IMPF. SUBJUNC., Latin *foret* 3RD SG. IMPF. SUBJUNC. < *\*fusēd*
- B. Imperfect indicative suffix *\*-fā-*, e.g., Oscan **fufans** “they were” 3RD PL. IMPF., Latin *portābant* “they were carrying” 3RD PL. IMPF. (*\*-fā-* > *-bā-* in Latin)
- C. Verbal adjective formation in *\*-ndo-*, e.g., Oscan **úpsannam** ACC. SG. FEM. “to be built,” Umbrian **pihaner** GEN. SG. MASC. “to be purified” (*\*-nd-* > *-nn-* in Sabellian), Latin *operandam* ACC. SG. FEM. “to be built”

The Sabellian languages share several significant morphological innovations, among which are the spread of the *i*-stem genitive singular ending *\*-eis* to *o*-stem and consonant-stem inflection; the spread of the *o*-stem accusative singular *\*-om* to consonant-stems; personal and reflexive pronominal forms with accusative singular **-om/-om** (e.g., Umbrian *tiom* “you,” **míom** “me,” Oscan *siom* “himself”); and the development of a mediopassive infinitive suffix in **-fir/-fi** (Oscan **sakrafir** “to be consecrated,” Umbrian *pihafi* “to be purified”).



**Figure 33.1** South Picene inscription. South Picene, Rix Sp TE 6, stele (fragmentary)

**[–?]nis : safinúm : nerf : persukant : p[–?]**

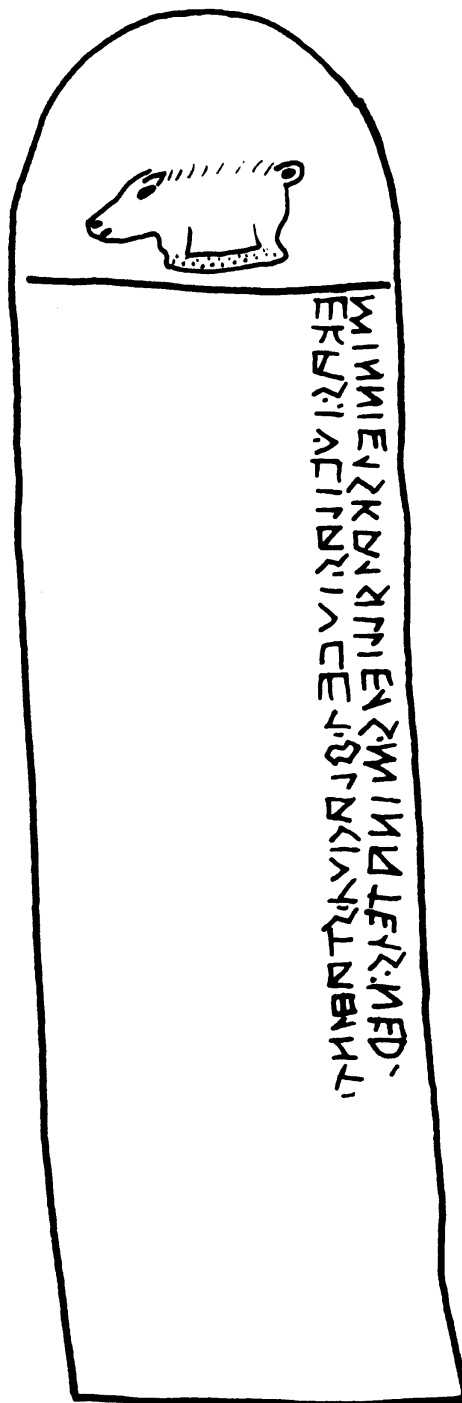
[nis-[name ?] "Sabines"-gen. pl. masc. "leaders"-acc. pl. masc. "?"-3rd pl. pres. act. "[names?] they ? the leaders of the Sabines [?]"

Prominent phonological innovations include the syncope of *\*o* in word-final syllables (*\*g<sup>h</sup>ortos* > Oscan **húr<sup>z</sup>** "enclosure"), the raising of inherited mid vowels (e.g., *\*ē* to *\*ē̄*, Proto-Sabellian *\*fēs<sup>nā</sup>* giving Oscan **físnú** "sanctuary," cf. Latin *fēstus* "festal"), and the change of Proto-Indo-European labiovelars to labials (*\*k<sup>w</sup>*is > Oscan **pis** "who").

Interrelationships among the Sabellian languages are difficult to determine because there is so little evidence for the languages in central Italy. However, the split into two Sabellian subgroups, one closely aligned with Umbrian, the other with Oscan, is not supported by the evidence. Instead, the territories occupied by Sabellian speakers form a linguistic continuum with Umbrian positioned in the north, Oscan in the south, and the Sabellian languages in central Italy constituting a transitional linguistic area where the languages have both Oscoid and Umbroid features (Wallace 1985). Exactly how South Picene fits into this schema is currently under deliberation (Meiser 1987; Adiego Lajara 1990).

## 2. WRITING SYSTEMS

The Sabellian languages were written in a variety of different alphabets. The type of alphabet employed depended on two factors: when a Sabellian tribe became literate and from what



**Figure 33.2** Oscan inscription. Oscan, Rix Cp25a, funerary stele (side a)

**ekas : iúvilas . iuvei . flagiúi . stahint . / minnieis kaisillieis . minateis : ner .**

"these"-nom. pl. fem. "Iovilas"-nom. pl. fem. "Jupiter"-dat. sg. masc. "Flagius"-dat. sg. masc. "be standing"-3rd pl. pres. act. "Minis"-gen. sg. masc.

"Kaisillies"-gen. sg. masc. "Minaz"-gen. sg. masc. "commander"-abbreviation for gen. sg. masc.

"These Iovilas are set up for Juppiter Flagius. [They belong to] Minis Kaisillies, [son of] Minas, commander."

**Table 33.1 National Oscan alphabet, c. 250 BC**

Character	Transcription	Character	Transcription
𐌆	a	𐌛	m
𐌇	b	𐌜	n
𐌈	g	𐌝	p
𐌉	d	𐌞	r
𐌊	e	𐌟	s
𐌋	v	𐌠	t
𐌌	z	𐌡	u
𐌍	h	𐌢	f
𐌎	i	𐌣	í
𐌏	k	𐌤	ú
𐌐	l		

source – Greek, Etruscan, or Latin – the alphabet was borrowed. Some Sabellian tribes borrowed from more than one source.

Oscan inscriptions were written in three different alphabets. Inscriptions from Campania and Samnium were composed in an alphabet that was borrowed from Etruscans who had colonized the Campanian plain in the sixth century BC. In the territories of Lucania and Bruttium, Oscan inscriptions were written in an alphabet of the East Greek type. A few inscriptions from the first century BC, including the important *Tabula Bantina*, were written in a Republican Latin alphabet.

The Oscan alphabet that developed from Campanian Etruscan sources was formed during the last half of the fifth century BC. This alphabet spread rapidly throughout Oscan-speaking Campania and into Samnium and was eventually codified as the so-called *national* Oscan script (see Table 33.1). At the beginning of the third century, two new signs were incorporated into the *abecedarium* in order to represent more accurately the phonology of Oscan mid vowels. Diacritics were added to the letters **i** and **u** to create signs for the vowels /ē/, /ē/ 𐌣 and /o/ 𐌤. These signs are transcribed as **í** and **ú** respectively.

The Sabellian-speaking tribes in central Italy, most of whom became literate via contact with Romans, borrowed the Latin alphabet. In a few instances, there is evidence for changes in the inventory of signs. In Paelignian, for example, the sign *delta* was modified by means of a diacritic and then employed on several inscriptions to represent the outcome of the palatalization of a voiced dental stop (\**dy* > [j]), transcribed as 𐌞, for example, Paelignian *petieḌu* “Petiedia” (nom. sg. fem.).

Umbrian was written in several different local versions of an Etruscan alphabet (Cristofani 1979). One of the earliest Umbrian inscriptions, that inscribed on a statue of Mars, was written in an alphabet similar to the one used in the central Etruscan city of Orvieto. The alphabet of Umbrian inscriptions from Colfiorito may also have come from this area, as is indicated by the fact that *gamma* was used for the voiceless velar /k/ rather than *kappa*. In contrast, the Iguvine Tablets I through Vb7 were inscribed in an Etruscan-based alphabet that did not have the letter *gamma*. This alphabet had a north Etruscan source, perhaps Perugia or Cortona (see Table 33.2).

The chief characteristic of the Umbrian alphabet used for the Iguvine Tablets I–Va is the absence of the signs *gamma* and *omicron*. The voiced stop /g/ was represented by *kappa*, and

**Table 33.2 Umbrian alphabet, Iguvine Tablets I–Vb7, c. 250 BC**

Character	Transcription	Character	Transcription
ⱱ	a	Ɱ	m
Ɱ	b	Ɱ	n
Ɱ	ř	1	p
Ɱ	e	Ɱ	r
Ɱ	v	2	s
Ɱ	z	Ɱ	t
Ɱ	h	Ɱ	u
Ɱ	i	8	f
Ɱ	k	d	ç
Ɱ	l		

*upsilon* was used for the mid vowel /o/. Interestingly, the signs *beta* and *delta* were a part of this script, although it is not clear whether they were inherited from the Etruscan alphabet that served as a model or were reborrowed from another source. *Delta* was used for a voiced fricative /z/ (ř) rather than for the voiced stop /d/, which was represented by *tau*. Both *pi* and *beta* shared the function of representing the voiced stop /b/, e.g., **hapinaf**, **habina** (acc. pl. fem.) ‘lambs.’ The inherited Etruscan inventory of signs was further modified in order to represent the native Umbrian phoneme /š/. The letter d (transcribed ç), of uncertain origin, was assigned this function.

Tablets Vb8, VI, and VII and a small number of Umbrian inscriptions belonging to the second and first centuries BC were written in a Republican Latin alphabet. The inventory of signs was augmented by the addition of S' (/š/, transcribed š), a Latin *sigma* modified by an oblique stroke appended in the upper left quadrant of the sign space.

Sabellian inscriptions composed in an Etruscan-based alphabet were generally written sinistrogade (right to left), but some were written dextrograde and a few others were laid out in boustrophedon (“as the ox plows”) style, every other line alternating in direction. Oscan inscriptions in the Greek alphabet were consistently written from left to right, as were the Sabellian inscriptions in the Latin alphabet, including Tablets Vb8, VI, and VII of the *Tabulae Iguvinae*.

Most Sabellian inscriptions in Etruscan-based alphabets use some form of punctuation to separate words, although a few of the earliest inscriptions are written *scriptio continua*. Punctuation between words is customarily a single point appearing at mid-line level, but word-dividers also take the form of double or triple puncts, the latter being particularly common on South Picene inscriptions in order to avoid confusion with the sign for /f/ : (see Figure 33.1, South Picene). In contrast, Oscan inscriptions written in the Greek alphabet rarely use punctuation for word boundaries; *scriptio continua* is the norm.

### 3. PHONOLOGY

Despite the genetic affiliation of the Sabellian languages, the phonological systems of each language developed distinctive characteristics. The Oscan sound system was more conservative, the Umbrian system more innovative.

**Table 33.3 The consonantal phonemes of Oscan**

Manner of articulation	Place of articulation						
	Bilabial	Labiodental	Dental	Palatal	Velar	Labiovelar	Glottal
Stop							
<i>Voiceless</i>	p		t		k		
<i>Voiced</i>	b		d		g		
Fricative		f	s				h
Nasal	m		n				
Liquid							
<i>Lateral</i>			l				
<i>Nonlateral</i>			r				
Glide				y		w	

Throughout the remaining sections of this chapter, the following abbreviations are used in glossing examples: G (gentilicium); PN (praenomen); DN (name of a god or goddess).

### 3.1 Oscan consonants

The Oscan consonantal inventory consists of fifteen members. There are three sets of stops – labials, dentals, and velars – with each set having a contrast in voicing. The three fricatives are all voiceless, and the nasals, liquids, and semivowels voiced.

These phonemes are illustrated by the examples of (2):

#### (2) Oscan consonant phonemes

**pús** (“who”) /p/, **tanginúd** (“decree”) /t/, **kúmbened** (“it was agreed”) /k/  
**blússii(eis)** (“Blossius” G) /b/, **deded** (“he gave”) /d/, **genetaí** (“Genita” DN) /g/  
**faamat** (“he calls”) /f/, **súm** (“I am”) /s/, **heriad** (“he should wish for”) /h/  
**maatreís** (“mother”) /m/, **niir** (“commander”) /n/  
**leígúss** (“statute”) /l/, **regatureí** (“the director,” epithet of Jupiter) /r/  
**iúveí** (“Jupiter” DN)/y/, **veru** (“gate”) /w/

The fricative /h/ was probably restricted to word-initial position. The fact that non-etymological **h** appears occasionally to mark vocalic hiatus supports this view; consider Oscan **stahínt** /stāɛnt/.

Intervocalic /s/ was phonetically voiced. The evidence is provided by inscriptions written in the Latin alphabet where the sign *z* is employed to write the sound derived from original \*s, for example, *ezum* [ezum] “to be” (pres. inf.), *egmazum* [egmazum] “affairs” (gen. pl. fem.). It is possible that the fricative /f/ was also voiced intervocalically, but the writing system provides no evidence in this instance.

#### 3.1.1 Palatalization

All geographical varieties of Oscan palatalize consonants (except for /f, s, w/) in the environment of a following /y/. Palatalization was marked in the national alphabet by gemination of the palatalized consonant: for example, **mamerttiaís** “of Mamers (name of month)”



<\*-ty-, **meddikkiai** “the office of meddix” (title of political official) <\*-ky-, **kúmbennieís** “assembly” <\*-ry-, **vítelliú** “Italia” <\*-ly-. The dialect of Bantia, which is attested by the *Tabula Bantina* (c. 90–80 BC), shows a more advanced stage of development. Dental and velar stops were assibilated and the glide was lost, thus, *bansae* “Bantia” (town in Apulia) <\*-ty-; *meddixud* <\*-ky-. Moreover, palatalized liquids were spelled without any indication of palatalization, e.g., *famelo* [-eʎo] “servant” <\*-ly-; *herest* [-eʎe-] “he will wish for” <\*-ry-.

### 3.1.2 Anaptyxis

Another feature characteristic of Oscan phonology is the anaptyxis of vowels to break up clusters consisting of sonorant (liquids, nasals) and some other consonant. Anaptyxis occurred in sonorant plus consonant clusters, for example, **aragetud** “silver” (abl. sg.) < \**argentōd*, as well as in consonant plus sonorant clusters, for example, **paterei** “father” (dat. sg. masc.) < \**patrei*, provided the preceding vowel was short. In the case of so-called *anterior* anaptyxis, the quality of the anaptyctic vowel was determined by the quality of the vowel preceding the sonorant, for example, **aragetud** and **herekleís** “Herakles” (gen. sg. masc.) < \**herkleís*. On the other hand, in *posterior* anaptyxis the quality of the anaptyctic vowel was determined by the quality of the vowel following the sonorant, as in **paterei** and **tefúrúm** “burnt offering” (acc. sg.) < \**tefrom*.

## 3.2 Oscan vowels

The Oscan vowel system is made up of eleven phonemes. There are three pairs of phonemes in the front region, each pair being distinguished by the features of height and length: /i/ and /ī/; /e/ and /ē/; /a/ and /ā/. The inventory of back vowels is half that of the front region: two high vowels, /u/ and /ū/, and one mid vowel, /o/. The low vowels /a/ and /ā/ fill out the system. In the national Oscan script, long vowels in word-initial/radical syllables are distinguished from short ones (see §3.3) by double writing of the vowel sign, though this orthographic practice is by no means consistently employed, even within the same inscription.

### (3) Oscan vowel phonemes

**viíbis** (PN) /ī/, **tanginúđ** (“decree”) /i/ (no examples in initial syllables)  
**fiísnam** (“temple, sanctuary”) /ē/, **ídík** (“it”) /e/  
**teer[úm]** (“territory”) /ē/, **pedú** (“foot”) /e/  
**fluusai** (“Flora” DN) /ū/, **purasiai** (“concerned with fire”) /u/  
**púd** (“which”) /o/ (no examples of /ō/ are attested)  
**slaagid** (“boundary”) /ā/, **tanginúđ** (“decree”) /a/

In addition to these simple vowel phonemes Oscan also has five diphthongs, /ai/, /ei/, /oi/, /au/, and /ou/.

### (4) **kvaísturei** (“quaestor”) /ai/, **deívai** (“divine”) /ei/, **múínikú** (“common”) /oi/ **avt** (“but”) /au/, **lúvkei** (“grove”) /ou/

Although the evidence is not conclusive, it is likely that the distinction in vowel length noted above was maintained only in word-initial/radical syllables (Lejeune 1970:279). It is also likely that distinctions in vowel quality were neutralized in word-final syllables. Etymological \*ā in absolute final position and etymological \*o, \*ō, and \*u in final syllables,

both open and closed, are spelled either **ú** or **u** in the national Oscan alphabet, the variation in spelling being tied to the writing habits of local scribes. The use of **ú** or **u** to spell what originally were four different sounds suggests that they all developed phonetically to a mid vowel having a quality between that of [u] and [o], perhaps [ɔ] (Lejeune 1970:300–305).

At the beginning of the third century BC, the Oscan vowel system was augmented by a sound that developed from short /u/ after dental consonants. In the national alphabet this sound is spelled **iu**, for example, **tiurri** “tower” (acc. sg.) < \**turrim*, compare Latin *turrim*; this spelling probably represents a palatalized [u], in other words [tyurre]. However, there is some evidence to suggest that by the end of the third century the pronunciation of this phone had developed to a front rounded vowel [ü]. For representing this sound, third- and second-century Oscan inscriptions written in the Greek alphabet use *upsilon* (υ), which had the value [ü] in Greek at the time, for example, Νιυμσδιης /n(y)ümsdieis/ (gen. sg. masc.), Νυμψιυ/nümpsim/ (acc. sg. masc.). In order to keep the high-back vowels /u, ü/ graphically distinct from [ü], they were spelled with the digraph ου, e.g., ουπσενς /üpsens/ “they built” (3rd pl. perf.).

### 3.3 Umbrian consonants

The Umbrian consonantal inventory displays several substantive differences when compared with that of Oscan. In addition to the dental fricative /s/, Umbrian has a voiceless palato-alveolar spirant that developed from the prehistoric combinations \**ky*, \**ki*, \**ke*, for example, **çerfie** /šerfye/ “Serfia” (epithet of deities) dat. sg. masc. Perhaps the most interesting innovation in the system was the change that introduced yet another fricative. This new sound, which was probably a voiced retroflex spirant /ʒ/, developed from intervocalic \**d* and also from intervocalic \**l* when adjacent to palatal vowels (Meiser 1986:213). In the native alphabet the sound is represented by the sign ʀ (ř); in the Latin alphabet it is spelled with the digraph *rs*, for example, **teřa**, *dirsa* /deřa/ “gives” (3rd sg. pres. subunc.).

**Table 33.4** The consonantal phonemes of Umbrian

Manner of articulation	Place of articulation								
	Bilabial	Labio-dental	Dental	Palato-alveolar	Retroflex	Palatal	Velar	Labio-velar	Glottal
Stop									
Voiceless	p		t				k		
Voiced	b		d				g		
Fricative									
Voiceless		f	s	š					h
Voiced					ʒ				
Nasal	m		n						
Liquid									
Lateral			l						
Nonlateral			r						
Glide						y		w	

These sounds can be illustrated by the following examples:

### (5) Umbrian consonant phonemes

*poplom* (“people. nation”) /p/, **tuta**, *totam* (“community, state”) /t/, **kumaltu**, *comoltu* (“let him grind”) /k/  
**krapuvi**, *grabouie* (“Grabovius,” epithet of Jupiter) /b/, **teṛa**, *dirsa* (“he should give”) /d/, *grabouie* /g/  
**fust** (“he will give”) /f/, *stahu* (“stand”) /s/, **ṣerfie**, *ṣerfia* (“Serfius, Serfia,” epithet of deities) /š/, **habia** (“he should take hold of”) /h/  
*matrer* (“mother”) /m/, *nerf* (“commander”) /n/  
**kumaltu**, *comoltu* (“let him grind”) /l/, **rufru** (“red”) /r/, **teṛa** (“he should give”) /z/  
**iuviu**, *iouiu* (“of Jupiter”) /y/, **verufe**, *uerir* (“gate”) /w/

In Umbrian **h** is weakly articulated. The sound was lost in medial environments before the historical period, and the character **h** was frequently used to mark both vocalic hiatus and vowel length, for example, *stahu* /stāu/ “I stand” (1st sg. pres. act.), *ahatripursatu* /ā tripuṣatu/ (3rd sg. impv. II) “dance the three-step.” In word-initial position **h** may also have been lost. Spellings with and without **h** are found in the earliest sections of the Tablets, for example, **eretu** “wished for” (abl. sg. neut.), as are examples of **h** appearing where unexpected on etymological grounds, for example, **ebetrafe** acc. pl. fem. + postposition versus **hebetafe** (a place name).

### 3.3.1 Word-final consonants

Particularly characteristic of Umbrian are changes affecting word-final consonants. In the oldest Umbrian inscriptions word-final **d** is not spelled, for example, **dede** “gave” 3rd sg. perf. Word-final **s** is spelled sporadically in Iguvine Tablets I–Vb7, indicating that it too was weakened. In those Iguvine Tablets written in the Latin alphabet, original word-final **s** was rhotacized to **r**, for example, *popler* (gen. sg. masc.) “people, nation” < \**popleis* (Meiser 1986:277); furthermore, word-final **m**, **n**, **f** (< \*-ns), and **r**, including **r** from original **s**, were in the process of being lost. The writing of word-final **f** in these Tablets is illustrative; **f** is regularly, but not always, omitted in polysyllabic words and in monosyllables ending in a consonant cluster. In other monosyllables, however, **f** is generally written. The result is a sentence such as the following, in which final **f** is spelled in two words but not in two others (*rofu*, *peiu*): *abrof trif fetu heriei rofu heriei peiu* (VIIa 3) “let him sacrifice three boars, either red or spotted.”

### 3.4 Umbrian vowels

The basic inventory of Umbrian vowels is similar to that found in Oscan, though with two additional phonemes. The first is a long mid /ō/, corresponding to short /o/; the second is a short mid vowel which is phonetically lower than /o/, perhaps /ɔ/. As in Oscan, the distinction between long and short vowels is maintained in word-initial or radical syllables, etymological long vowels being shortened in medial and final syllables (Meiser 1986:150).

Umbrian has no diphthongs corresponding to those found in Oscan cognates; all diphthongs inherited from Proto-Sabellian were monophthongized before the historical period and merged with existing long or short vowel phonemes. New diphthongs subsequently arose in Umbrian as the result of phonological changes, for example, /dēytu/ **deitu** (3rd sg. impv. II) “speak,” **aitu** /aytu/ (3rd sg. impv. II) “set in motion.”

## (6) Umbrian vowel phonemes

**persnihmu** (“pray”) /i/, **atiersir** (“Atiedian”) /i/  
**sehmeniar** (“of?”) /ē/, **aves**, **avis**, **auēis** (“bird”) /e/  
**esuna**, **eesona** (“religious”) /ē/, **ařfertur** (“chief priest”) /e/  
**kumnahkle** (“meeting place”) /ā/, **ařfertur** /a/  
**pihaz**, **pihos** (“purified”) /ɔ/  
**uhtur**, **oht** (“auctor,” title of political office) /ō/, **poplom**, **puplum** (“people, nation”) /o/  
**struhçla** (“offering”) /ū/, **fust**, **fust** (“he will be”) /u/

## 3.5 Sabellian accent

Very little is known about the word accent of any Sabellian language. Nevertheless, it is possible to make informed inferences about accentuation based on orthographic practices and on certain phonological processes that affected the Sabellian languages, in particular Oscan and Umbrian, at various stages in their development. In all Sabellian languages short vowels were lost before word-final \*s. Short vowels in open medial syllables were also syncopated before the historic period. This vocalic instability suggests that Sabellian had a stress accent which was positioned on the initial syllable of words. The fact that vowel length is indicated only in initial/radical syllables in both Oscan and Umbrian (with rare exceptions) suggests that word-initial/radical accent was still in place during the historical period (Meiser 1986:150; for Oscan antepenultimate accent, see Schmid 1954).

## 4. MORPHOLOGY

The Sabellian languages are classified typologically as fusional, inflecting languages. All inflectional categories are signaled through endings attached to nominal and verbal stems. Several word classes, such as conjunctions, pre- and postpositions, sentential adverbs, and the cardinal numerals four and above, are uninflected.

## 4.1 Nominal morphology

The nominal system is composed of nouns, adjectives, and pronouns. With the exception of a handful of forms, all members inflect for the grammatical features of case, number, and gender. Sabellian has seven cases (nominative, vocative, accusative, dative, ablative, genitive, locative), two numbers (singular and plural), and three gender categories (masculine, feminine, and neuter). Nouns are generally assigned to one of the three genders on the basis of their stem-type. For example, *a*-stems are feminine, *o*-stems and *u*-stems either masculine or neuter, *men*-stems neuter, and so forth. There are, however, exceptions, particularly in the case of animate nouns, which are assigned gender based on sex, not on form. Adjectives, most pronouns, and the cardinal numerals from one to three inflect so as to agree in gender and case with the noun which they modify, for example, Umbrian **tutaper ikuvina** “for the Iguvine state” (abl. sg. fem.); Umbrian **tref sif kumiaf** “three pregnant sows” (acc. pl. fem.).

## 4.1.1 Nominal classes

Nouns are formally organized into subsystems – declensions – according to the formation of the stem (see Table 33.5). Sabellian has four major vocalic-stem declensions: *a*- (Oscan **aasaí**

**Table 33.5 Sabellian noun stems***a-stems*

	Oscan	Umbrian
NOM. SG.	<b>víú</b> , <i>touto</i>	<b>muta</b> , <b>mutu</b>
VOC. SG.	—	<i>Tursa</i> , <i>prestota</i>
ACC. SG.	<b>víam</b> , <b>toutam</b>	<b>tuta</b> , <i>totam</i>
DAT. SG.	<b>deívaí</b>	<b>tute</b> , <i>tote</i>
ABL. SG.	<b>eítiu vad</b> , <b>toutad</b>	<b>tuta</b> , <i>tota</i>
GEN. SG.	<b>vereias</b>	<b>tutas</b> , <i>totar</i>
LOC. SG.	<b>víaí</b> , <i>bansae</i>	<b>tafle</b> , <i>tote</i>
NOM. PL.	<b>aasas</b> , <b>scriftas</b>	<b>pumperias</b> , <i>iuengar</i>
ACC. PL.	<b>viass</b> , <b>eituas</b>	<b>vitlaf</b> , <i>uitla</i>
DAT./ABL./LOC. PL.	<b>kerssnaís</b>	<b>tekuries</b> , <i>dequrier</i>
GEN. PL.	<b>eehiianasúm</b>	<b>urnasiaru</b> , <i>pracatarum</i>

*o-stems*

	Oscan	Umbrian
NOM. SG.	<b>húrz</b>	<b>Ikuvins</b>
VOC. SG.	<b>Statie</b> , <i>Silie</i>	<i>Serfe</i> , <i>Tefre</i>
ACC. SG.	<b>húrtúm</b>	<b>puplum</b> , <i>poplom</i>
DAT. SG.	<b>húrtuí</b>	<b>kumnacle</b> , <i>pople</i>
ABL. SG.	<b>sakaraklúd</b>	<b>puplu</b> , <i>poplu</i>
GEN. SG.	<b>sakarakleís</b>	<b>katles</b> , <i>popler</i>
LOC. SG.	<b>tereí</b> , <i>comenei</i>	<b>kumne</b> , <i>pople</i>
NOM. PL.	<b>Núvlanús</b>	<b>Ikuvinus</b> , <i>Iouinur</i>
ACC. PL.	<b>feihúss</b>	<b>vitluf</b> , <i>uitlu</i>
DAT./ABL./LOC. PL.	<b>Núvlanúís</b>	<b>veskles</b> , <i>uesclir</i>
GEN. PL.	<b>Núvlanúm</b>	<b>pihaklu</b> , <i>pihaclo</i>

*i-stems*

	Oscan	Umbrian
NOM. SG.	<b>aídil</b>	<b>ukar</b>
VOC. SG.	—	—
ACC. SG.	<b>slagím</b>	<b>uvem</b> , <i>uerfale</i> (NEUT.)
DAT. SG.	—	<i>ocre</i>
ABL. SG.	<b>slagid</b>	<b>ocri-per</b>
GEN. SG.	<i>aeteis</i>	<i>ocrer</i>
LOC. SG.	—	<b>ukre</b> , <i>ocre</i>
NOM./VOC. PL.	—	<b>puntes</b> , <i>sakreu</i> (NEUT.)
ACC. PL.	—	<b>avif</b> , <b>avef</b> , <b>perakneu</b> (NEUT.)
DAT./ABL./LOC. PL.	<b>luisarifs</b>	<b>avis</b> , <i>aves</i>
GEN. PL.	<b>[a]ittium</b>	<i>peracrio</i>

*consonant-stems*

	Oscan	Umbrian
NOM. SG.	<b>meddíss</b>	<b>ařfertur</b> , <b>pir</b> (NEUT.)
VOC. SG.	—	<b>Iupater</b>
ACC. SG.	—	<i>capirso(m)</i> , <b>pir</b> (NEUT.)
DAT. SG.	<b>medíkeí</b>	<b>nomne</b>

Table 33.5 (cont.)

	Oscan	Umbrian
ABL. SG.	—	<b>kapiře</b>
GEN. SG.	<b>medikeís</b>	<i>nomner/matres</i>
LOC. SG.	—	—
NOM./VOC. SG.	<b>humuns</b>	<b>frater/uasor</b> (NEUT.)
ACC. SG.	—	<b>capif</b> , <i>tuderor</i> (NEUT.)
DAT./ABL./LOC. SG.	—	<b>capirus</b>
GEN. SG.	<b>fratrúm</b>	<b>fratrum</b>

“altar” [loc. sg. fem.]), *o*- (Umbrian *poplom* “people” [acc. sg. masc.]), *i*- (Umbrian **uvi-kum** “with a sheep” [abl. sg. + postposition **-kum**]), and *u*-stems (Umbrian **trifu** “tribe” [acc. sg.]). In addition, four major consonant-stem declensions occur: stop- (Oscan **aitatum** “one’s age” [acc. sg.]), *s*- (Umbrian **meřs** “law” [nom. sg. neut.]), *r*- (Oscan **patir** “father” [nom. sg. masc.]), and *n*-stems (Umbrian **umen** “ointment” [acc. sg. neut.]). Sabellian probably also had another vocalic-stem declension, *ē*-stems (Umbrian **re-per** “according to the ceremony” [abl. sg. fem.] + postposition **-per**). Unfortunately, the evidence is limited to a few words in Umbrian, and it is consequently impossible to determine to what extent these constituted a special inflectional class.

Within these basic inflectional categories there exist several distinct paradigmatic subclasses. For example, *o*-stems, *i*-stems, and consonant-stems split into subgroups based on the gender of the noun – neuters having inflectional endings which are distinct from masculines and feminines in the nominative and accusative singular and plural:

#### (7) Oscan *o*-stem masculines and neuters

	masculine	neuter
NOM. SG.	<b>húrz</b>	<b>tefúrúm</b>
ACC. SG.	<b>húrtúm</b>	<b>dunum</b>
NOM. PL.	<b>Núvlanús</b>	<b>veru</b>
ACC. PL.	<b>feihúss</b>	<b>veru</b>

In addition, *o*-stems and *i*-stems developed subtypes as a result of sound changes that eliminated short *\*o* and short *\*i* in word-final syllables before *\*s* and, in the case of *\*o*, also in the environment *\*yom*. Owing to these changes, *o*-stems that were built historically with a *\*yo*-suffix came to have an inflectional pattern that was distinct from other types of *o*-stems. This latter group, in turn, is distinguished depending on whether the nominative singular retained or lost its original word-final *\*s*. Compare, for example, the nominative and accusative singulars in (8):

#### (8) Subclasses of Umbrian *o*-stem nouns

	<i>*to</i> -stems	<i>*ro</i> -stems	<i>*lo</i> -stems	<i>*yo</i> -stems
NOM. SG.	<i>tašez</i> / <i>tašets</i> /	<i>ager</i>	<b>katel</b>	<b>Vuvćis</b>
ACC. SG.	<i>ehiato</i> ( <i>m</i> )	<b>kaprum</b>	<b>katlu</b> ( <i>m</i> )	<b>graboui</b> ( <i>m</i> )

### 4.1.2 Diachronic developments

The paradigms given in Table 33.5 also serve to illustrate the main features of the diachrony of the nominal system in the Sabellian languages, namely the formal merger of cases both within and across paradigms. The *i*-stem genitive singular ending, Oscan **-eís**/Umbrian **-e(s)**, was taken over by *o*-stems and consonant-stems. The accusative singular ending **-om/-úm**, originally at home in *o*-stem inflection, spread into the consonant-stems. In Oscan the similarities between these two inflectional classes are even greater because the consonant-stems also borrowed the *o*-stem ablative singular **-úd/-ud**, for example, Oscan **tanginúd** (abl. sg.) “decree,” *ligud* (abl. sg.) “law.”

Generally, however, the formal merger of cases in Umbrian is considerably more advanced than in Oscan. Sound changes in Umbrian, in particular the monophthongization of diphthongs and the loss of word-final consonants, eliminated distinctions between case endings: consider, for example, Umbrian *a*-stem **tote** (dat. sg. fem.) “state,” **tote** (loc. sg. fem.) “state,” compare Oscan *a*-stem **anagtiai** (dat. sg. fem.) “Angitia” (name of goddess), **aasaí** (loc. sg. fem.) “altar”; Umbrian *a*-stem **uestisia** (acc. sg. fem.) “offering cake,” **uestisia** (abl. sg. fem.), compare Oscan *a*-stem **viam** (acc. sg. fem.) “road,” **toutad** (abl. sg. fem.) “state.”

### 4.1.3 Adjectives

Adjectives are organized into paradigmatic classes on the same basis as nouns, although the number of stem-types is more restricted. Adjectives are inflected as *o*-stems, *a*-stems, *i*-stems, and consonant-stems (no *u*-stems or *ē*-stems occur). Together *o*-stems and *a*-stems form one adjective declension, the masculine and neuters taking *o*-stem inflection (as in Oscan **túvtíks** “of the community, state” [nom. sg. masc.], *touticom* [acc. sg. neut.]) and the feminines taking *a*-stem inflection (Oscan *toutico* [nom. sg. fem.] with *-o* from *\*ā*). In contrast, *i*-stem and consonant-stem adjectives occur in all three gender classes (e.g., *i*-stem, Umbrian *perakri* “fit for sacrifice” [abl. sg. masc.], **perakre** [acc. sg. fem.]).

The inflectional category of degree, comparative and superlative, is marked by suffixes added to the adjective stem. The regular suffixes are **-tro-** and **-imo-** respectively, for example, Umbrian **mestru** (nom. sg. fem.) “greater,” Oscan **maimas** (nom. pl. fem.) “greatest.”

### 4.1.4 Pronouns

The Sabellian pronominal system includes personal, reflexive, demonstrative, emphatic, anaphoric, interrogative, indefinite, and relative pronouns. The pronouns for first and second persons are not marked for gender, but the rest of the forms in the pronominal system are assigned gender based on that of the noun with which they are in agreement or to which they refer, for example, Umbrian **este persklum** “this ceremony” (acc. sg. neut.).

Sabellian pronouns show significant differences in inflection when compared with nouns and adjectives. These differences are particularly strong in the personal pronouns, but are manifest also in other pronominal categories. For example, the dative singular of the first- and second-person pronouns has unique endings **-he**, **-fe/-fei**, for example, Umbrian **mehe** “to me,” **tefe** “to you,” Oscan **t(e)fei** “to you,” compare Latin *tibi*. Furthermore, the dative singular and the locative singular of demonstratives and relatives are marked by distinctive endings in Umbrian, dative **-smi**, **-smei**, locative **-sme**, for example, Umbrian demonstrative **esmi-k**, *esmei* “this” (dat. sg.), relative **pusme** “who, which” (dat. sg.), demonstrative *esme* “this” (loc. sg.). The pronominal neuter nominative/accusative singular is distinguished from nominals by its case ending **-d**, Oscan **púd** “which,” Umbrian **puře** “which” <*\*pod-id*.



Outside of the personal pronouns, Sabellian pronominal formations exhibit either *a-*, *o-*, or *i-*stem inflection. The relative and indefinite pronouns have the stems **po-** and **pi-**:

#### (9) Oscan and Umbrian relative pronouns

	Oscan	Umbrian
NOM. SG. MASC.	—	<i>poi, porsī</i>
NOM. SG. FEM.	<b>paí</b>	—
NOM. SG. NEUT.	<b>púd</b>	<b>puře</b>
ACC. SG. FEM.	<b>paam</b>	—
DAT. SG. MASC.	<b>pui</b>	<b>pusme</b>
ABL. SG. FEM.	<b>pad</b> , <i>poizad</i>	<i>pōra</i>
NOM. PL. MASC.	<b>pús</b>	<b>pure</b>
NOM./ACC. PL. NEUT.	<b>paí</b>	<i>porse</i>
ACC. PL. FEM.	—	<i>pafe</i>

Demonstrative formations typically have *a-/o-*stem inflection: for example, Paelignian *ecuc* “this” (nom. sg. fem.) < \**ekā-k(e)*, Oscan **ekas** “this” (nom. pl. fem.), both with stem \**eko-/ekā-*; Oscan **eíseis** “his” (gen. sg. masc.), Umbrian **erēr** “this” (gen. sg. masc.), with stem \**eiso-*; Umbrian **estu** “that” (acc. sg. masc.), with stem \**isto-*; Oscan **eksuk** “this” (abl. sg. neut.); Umbrian *eso* “this” (nom. sg. fem.) < \**eksā*, with stem \**ekso-/eksā-*:

#### (10) Oscan and Umbrian demonstrative pronouns (stem \**i-/ei-*)

	Oscan	Umbrian
NOM. SG. MASC.	<i>izic</i>	<b>erek</b>
NOM. SG. FEM.	<b>iiuk</b>	—
NOM. SG. NEUT.	<b>ídik</b>	<b>eřek</b>
ACC. SG. MASC.	<i>ionc</i>	—
ACC. SG. FEM.	<b>íak</b>	<i>eam</i>
NOM. PL. MASC.	<i>iusc</i>	—
ACC. PL. FEM.	<i>iafc</i> (Marrucinian)	<b>eaf</b>
NOM./ACC. PL. NEUT.	<i>ioc</i>	<b>eu</b>

The Sabellian anaphoric pronoun is built with the stem \**i-/ey-*, for example, Oscan *izic* “he” (nom. sg. masc.), Umbrian **erek**, *erēc* “he” (nom. sg. masc.).

In the prehistory of the Sabellian languages many of these pronominal forms were augmented by means of particles. The accretion of these particles to pronominal forms had the effect of producing paradigms with inflectional endings that appear, at first glance, to have little in common with those of the nominal system. In many instances the inflectional ending of a pronominal form cannot easily be recognized until the particle has been removed, for example, Umbrian **erarunt** “the same” = **erar** (gen. sg. fem.) + particle **-unt** < \**esās-ont*; Umbrian **erak** “this” = **era** (abl. sg. fem.) + particle **-k** < \**esād-k(e)*; Umbrian **puře** “which” = **puř** (nom. sg. neut.) + particle **-e** < \**pod-i*.

## 4.2 Verbal morphology

The Sabellian verb is inflected for the categories of tense, voice, mood, person, and number. There are three persons (first, second, third), two numbers (singular, plural), and two voices (active, mediopassive). The mood categories are indicative, imperative, and subjunctive. Five

different tense forms are attested for Sabellian verbs: present, imperfect, future, perfect, and future perfect. The basic symmetry of the Sabellian system and the fact that it is quite similar to that of Latin suggest the occurrence of another tense form, the pluperfect, compare Latin *portauerat* “had carried.”

### 4.2.1 Aspectual stems

The finite verb system is formally organized into subsystems based on two stem-types that mark a distinction in aspect, the *infectum* (present system) and the *perfectum* (perfect system). Present, imperfect, and future tense forms are built on the stem of the infectum, the perfect and the future perfect on that of the perfectum:

(11)	INFECTUM	PERFECTUM
PRES.	<i>didet</i> “he gives” (Vestinian)	PERF <i>fe&lt;id&gt;</i> “he should do”
FUT.	<i>didest</i> “he will give”	FUT. PERF. <i>fefacust</i> “he will have done”
IMPF.	<i>fufans</i> “they were”	PLUPERF. ? —

### 4.2.2 Verb endings

The grammatical categories of person, number, and voice are signaled by affixes traditionally called “personal endings.” These are of two basic sets, one for active and one for mediopassive voice (see Table 33.6). The active set of endings has two forms depending on the tense of the verb to which it is attached: the so-called primary endings are used for present, future, and future perfect tenses; while secondary endings are used for imperfect and perfect indicative, and for all tenses of the subjunctive. In the passive voice, only Umbrian shows a primary versus secondary distinction, for example, 3rd sg. mediopass. – primary **herter** “it is desirable” (3rd sg. pres.); secondary **emantur** “they should be accepted” (3rd sg. pres. subjunc.).

**Table 33.6 Sabellian personal endings**

PRIMARY	
1ST SG. ACT.	Umbrian <b>suboca-u</b> “I invoke”
2ND SG. ACT.	Umbrian <b>herie-s</b> “you will desire”
3RD SG. ACT.	Vestinian <i>dide-t</i> “he gives”
3RD SG. MEDIOPASS.	Oscan <i>uinc-ter</i> “he is convicted”
1ST PL. ACT.	—
2ND PL.	—
3RD PL. ACT.	Umbrian <i>furfa-nt</i> “they shear”
3RD PL. MEDIOPASS.	Umbrian <i>ostens-endi</i> “they will be presented”
SECONDARY	
1ST SG. ACT.	Oscan <b>manaf-úm</b> “I entrusted”
2ND SG. ACT.	—
3RD SG. ACT.	Oscan <b>prúfatte-d</b> “he approved”
1ST PL. ACT.	South Picene <b>adstaeo-ms</b> “we have set up”
2ND PL. ACT.	Umbrian <b>benuso</b> /-us-so/ “you all will have come”
3RD PL. ACT.	Paelignian <i>coisat-ens</i> “they took care of”
3RD PL. MEDIOPASS.	Umbrian <b>ema-ntur</b> “they should be accepted”

The Sabellian languages also possess a third singular mediopassive suffix **-r** for use in impersonal constructions, for example, Umbrian **ferar** (3rd sg. pres. subjunc.). mediopass. “there is a carrying,” *ier* (3rd sg. pres. mediopass.) “there is a going.”

### 4.2.3 Verbal classes

The Sabellian verb is organized into paradigmatic classes, or conjugations, based on the form of the verb-stem found in present tense inflection. If verbs such as “to be” (Oscan **súm** “I am”) and “to go” (Umbrian *est* “he will go”) are excluded as “irregular,” five basic conjugational patterns can be established: *a*-conjugation (Oscan **faamat** “he calls”); *e*-conjugation (Umbrian **tusetu/tusitu** “let him pursue,” Oscan **fatium** “to speak,” **licitud** “let it be permitted”); *i*-conjugation (Umbrian **seritu** “let him watch out for”); *y/i*-conjugation (Umbrian **façiu** /fašyo(m)/ “to sacrifice,” Oscan **fakiiad** “he should make”); and *e/ø*-conjugation (Oscan *agum* “to move,” *actud* “let him move,” Umbrian **aitu** “let him move”). Forms of the *y/i*- and *e/ø*-conjugations such as the Oscan imperatives *factud*, *actud* are derived from earlier forms in which medial vowels were present – short *i* for *factud* < \**fakitōd*, short *e* for *actud* < \**aketōd*.

### 4.2.4 Verb tense

Tense is typically signaled by a combination of stem-type (perfectum versus imperfectum; see §4.2.1) and suffixation. Outside of the present and perfect there are special tense-forming suffixes. The imperfect has **-fa-**, the future **-(e)s-**, and the future perfect **-us-**: for example, Oscan **fu-fa-ns** “they were” 3rd pl. impf., Oscan *deiua-s-t* “he will swear” (3rd sg. fut.), Oscan **tribarakatt-us-et** “they will have built” (3rd pl. fut. perf.).

Sabellian perfect tense stems of active voice are formed by a number of different morphological operations: (i) reduplication (Oscan **deded** “he gave” [3rd sg. perf.], *fefacid* “he should do” [3rd sg. perf. subjunc.]; Umbrian *dede* “he gave” [3rd sg. perf.]); (ii) suffixation (**-tt-**: Oscan **prúfatted** “he approved” [3rd sg. perf.]; **-nçi-/ -nši-**; Umbrian *purdinšiu* “he will have presented” [3rd sg. fut. perf.]; **-f-**: Umbrian **andirsafust** “he will have made a circuit” [3rd sg. fut. perf.]); and (iii) radical vowel lengthening (Oscan **uupsens** “they built” [3rd pl. perf.]). Some perfects are formed from the bare verb-stem, minus the suffix used to generate the present: for example, Umbrian **anpelust** “he will have slain” (3rd sg. fut. perf.) built to a present that is characterized by a suffix **-ne**, **anpentu** “let him slay” < \**-pennetōd* < \**-pelnetōd*. In the mediopassive, the perfect is formed by a periphrastic construction involving the past participle plus a form of the verb “to be”: for example, Oscan **prúftúset** (“they have been approved” [3rd pl. perf. mediopass.]; Oscan *scriptas set* “they have been written” [3rd pl. perf. mediopass.]; Umbrian **pihaz fust** “it will have been purified” [3rd sg. fut. perf. mediopass.]). Interestingly, there is one perfect mediopassive formation that is not a periphrastic, Oscan *comparascuster* “it will have been discussed,” a future perfect found in the *Tabula Bantina*. Presumably this formation is an independent (and late?) Oscan creation.

In some cases, in particular derived verbs, the stem of the perfect is built directly from the present. For example, *a*-stem presents generally form **-t(t)**-stem perfects in Oscan and in the Sabellian languages of central Italy: thus, Oscan **duunated** “he presented” (3rd sg. perf.); Paelignian *coisatens* “they took care of” (3rd pl. perf.); Marrucian *amatens* “they seized” (3rd pl. perf.); Volscian *sistiatens* “they set up” (3rd pl. perf.). Still, even here there are exceptions. The verb-stem **opsa-** “build” forms a perfect by lengthening the radical vowel and truncating the present stem vowel **a**, thus Oscan **uupsens** “they built” (3rd pl. perf.).

In Umbrian, *a*-stems form their perfects by means of the suffix *-f-*, *andirsafust* “he will have made the circuit” (3rd sg. fut. perf.). In many cases the type of perfect formation cannot be predicted by the paradigmatic class of the present. For example, the verb “to give” forms a reduplicated present (Vestinian *didet* “he gives” [3rd sg. pres.]) and a reduplicated perfect (Oscan *deded* “he gave” [3rd sg. perf.]), while the verb “to make” forms a *y/i-* present but a reduplicated perfect, *fakiia* “he should make” (3rd sg. pres. subjunc.), *fefacid* “he should make” (3rd sg. perf. subjunc.).

#### 4.2.5 Nonindicative moods

Subjunctive mood is indicated by suffixes which are attached to the verb-stem preceding the personal endings. Present subjunctive is marked by *-a* in Umbrian for all present classes except *a*-conjugation, which shows *-ia*: for example, *e*-conjugation *habi-a* “he should hold” (3rd sg. pres. subjunc.); compare *a*-conjugation *porta-ia* “he should carry” (3rd sg. pres. subjunc.). In Oscan *-i* is used for *a*-conjugation, *deiua-i-d* “he should swear” (3rd sg. pres. subjunc.), *-a* for all other conjugation classes, for example, *púti-a-ns* “they should be able” (3rd pl. pres. subjunc.).

The imperfect subjunctive is attested only in Oscan and Paelignian. The suffix used is Oscan *-sí*, Paelignian *-se* (< \**sē*): Oscan *fu-sí-d* “he should be” (3rd sg. impf. subjunc.); Paelignian *upsa-se-ter* “it was built” (3rd sg. impf. subj. mediopass.). For the perfect subjunctive active, the suffix is *-i/i*, Oscan *tribarakatt-i-ns* “they should build” (3rd pl. perf. subjunc.).

Imperative mood forms have two special sets of person, number, and voice endings. So-called *imperative I* endings are used for commands that are to be carried out immediately following the time of speaking:

##### (12) Imperative I

2ND SG. ACT.	Umbrian <i>anserio</i> “observe”
3RD SG.	—
2ND PL. ACT.	Umbrian <i>eta-tu</i> “go,” Paelignian <i>ei-te</i> “go”
2ND PL. MEDIOPASS.	Umbrian <i>katera-mu</i> “arrange in order”
3RD PL.	—

*Imperative II* endings are reserved for commands to be carried out at some undefined point in the future. This type is particularly common in the Iguvine Tablets, where sets of ritual instructions are set forth to be carried out whenever the religious observance is required:

##### (13) Imperative II

2ND SG. ACT.	Umbrian <i>ene-tu</i> “begin”
2ND SG. MEDIOPASS.	Umbrian <i>persni-mu</i> “pray”
3RD SG. ACT.	Oscan <i>liki-tud</i> “let it be permitted”
3RD SG. MEDIOPASS.	Oscan <i>censa-mur</i> “let him be assessed”
2ND PL. ACT.	Umbrian <i>ambre-tuto</i> “circumambulate”
2ND PL. MEDIOPASS.	Umbrian <i>pesni-mumo</i> “pray”
3RD PL. ACT.	Umbrian <i>habi-tuto</i> “let them hold”
3RD PL. MEDIOPASS.	Umbrian <i>pesni-mumo</i> “let them pray”

#### 4.2.6 Nonfinite verbals

An important component of the Sabellian verbal system consists of a constellation of nonfinite formations. These include present infinitives, both active and mediopassive (Umbrian *erom* “to be” [pres. act.]; Umbrian *pihafi* “to be expiated” [pres. mediopass.]); present and past participles (Umbrian *zeref* “sitting” [pres. act.]; Umbrian *çersnatur* “having dined” [past. mediopass.]); supines (Umbrian *anzeriatu* “to observe”); and the so-called gerundive (Oscan *úpsannam* “to be built”).

### 4.3 Derivational morphology

Complex Sabellian words are formed by means of the morphological processes of affixation and compounding. Affixation, in particular suffixation, appears to have been more productive than compounding.

#### 4.3.1 Suffixation

Several suffixes are used productively to form nouns in Oscan and Umbrian. The suffix **-iuf** (nom. sg.)/**-in-** (other cases) produces nouns with abstract or concrete meanings, for example, Oscan *tribarakkiuf* “a building,” compare *tribarakattens* “they built.” The extended suffix **-tiuf/-tin-** has the same morphological function, for example, Oscan *medicatinom* “judgment,” Umbrian *natine* “tribe,” compare Praenestine Latin *nationu* “childbirth” (gen. sg.). The suffix **-tur** is used to form agent nouns from verb-stems, for example, Oscan *regaturei* “the director” (epithet of Jupiter) dat. sg. from *\*regā-* “direct,” *ařfertur* “flamen, chief priest” from *\*ad-fer-* “to carry.” The suffix **-etia**, which is added to noun stems to build abstracts, is attested in Umbrian by several formations that serve to indicate terms of elected office, for example, *kvestretie* “in the term of office as quaestor” (loc. sg.).

One productive adjective-forming suffix is **-(a)sio-** “relating to, pertaining to,” used to form adjectives from nominal stems: for example, Oscan *kerssnasias* “concerned with banquets” (nom. pl. fem.), compare Oscan *kersnu* “banquet” (nom. sg. fem.); *purasiai* “concerned with fire” (loc. sg. fem.), compare Umbrian *pir* “fire” (nom./acc. sg. neut.). The suffix **-ano-** is also used to form adjectives from nouns; most of the examples attested in inscriptions are formed from ethnic or topographical names, for example, Oscan *Abellanús* “from the city of Abella” (nom. pl. masc.), Umbrian *Treblanir* “leading to Trebula” (abl. sg. neut.).

Verbs are productively formed in all Sabellian languages by means of the suffix **-a** or by extensions of this suffix, **-ia**, **-ta**, etc. Formations in **-a**, a suffix used primarily to build verbs from nouns and adjectives, are widely attested: thus, Umbrian *kuratu* “accomplished” (acc. sg. neut. mediopass. part.); Paelignian *coisatens* “they supervised” (3rd pl. perf.) < *\*koišā-*, compare Latin *cūra* “concern”; Oscan *deiuaid* “he should swear” (3rd sg. pres. subjunc.) < *\*deiuā-*, compare Oscan *deivai* “divine” (dat. sg. fem.); Umbrian *pihatu* “let him purify” (3rd sg. impv. II), compare Volscian *pihom* “religiously unobjectionable” (nom. sg. neut.); Oscan *teremnattens* “they set a limit on” (3rd pl. perf.) < *\*termnā-*, compare Oscan *teremniss* (acc. pl.), Latin *termen* “limit”; Umbrian *osatu* “let him build” (3rd sg. impv. II) < *\*opesā-*, compare Latin *opus* “work.” This suffix, as well as variants derived from it, are also used in the formation of deverbative verbs: for example, Umbrian *andirsafust* “he will have made the circuit” (3rd sg. fut. perf.) < *\*am-did-ā-*; Umbrian *kumbifiatu* “deliver instructions” (2nd sg. impv. II) < *\*kom-bif-iā-*, compare Latin *fidit* “he puts confidence in”; Umbrian *etaians* “they go” (3rd pl. pres. subjunc.) < *\*ey-tā-*.

### 4.3.2 Compounds

Sabellian compound formations consist in large part of words with an adverbial first constituent. In fact, the only pervasive type of verbal composition attested in Sabellian involves the use of adverbial elements: for example, Umbrian **aha-uendu** “let him turn away” (3rd sg. impv. II), **am-pendu** “let him slay” (3rd sg. impv. II), **re-vestu** “let him examine” (3rd sg. impv. II), etc. There is also a substantial number of nominals formed by means of an adverbial first constituent. The best attested are built with the privative element **a-**, **an-** “not”: for example, Oscan *an-censto* “unburnt” (nom. sg. fem.); Umbrian *a-uirseto* “unseen” (nom. sg. neut.), **an-takres** “unground” (abl. pl.), **a-snata** “not wet” (acc. pl. neut.), **a-seçeta** “uncut” (abl. sg. fem.).

Nominal compounding is not well represented in Sabellian. There are a couple of good examples of possessive compounds with numerals as the first member, for example, Umbrian *petur-purs-us* (dat. pl.) “animals” (i.e., “having four feet”); *du-pursus* (dat. pl.) “having two feet.” But aside from these, there are few formations that qualify as compounds from a synchronic point of view, though several forms derive historically from compounds: thus, Oscan **meddiss** “meddix” (a title of magistracy), which was originally an adjectival compound with first member *\*med-* “law” and second member *\*dik-* “speaking,” compare Latin *iūdex* < *\*iowes-dik-* “speaking the law.” The semantics of **meddiss**, the fact that it refers to a magistracy, suggests that it was no longer interpreted synchronically as a compound.

### 4.3.3 Locative case formation

An especially interesting morphological development is found in the Oscan and Umbrian case system. The postposition Oscan *-en* “in, upon,” Umbrian **-en**, **-e**, **-em** “in, upon” governs the locative case in one of its primary functions. When this postposition was added to the locative of *o*-stem nominal forms in Oscan, or to the locative of vowel stems in Umbrian, the vowel of the case ending and initial vowel of the postposition contracted, as in Oscan **húrtin** /hortēn/ “in the precinct” < *\*hortey-en*. This contracted form of ending + postposition was then reanalyzed as a new form of the locative case. That such was indeed the case is indicated by noun phrases in which this “ending” is attached to both adjectives and nouns, for example, Oscan **húrtin Kerriin** “in the precinct of Ceres” (loc. sg. masc.), Umbrian *ocrem Fisiem* “on the Fisian Mount” (loc. sg.); and by instances in which the postposition has been added to a noun already marked with the original postposition, for example, Umbrian *toteme Iouinem* “in the Iguvine community.” In this instance, *toteme* can be segmented diachronically as *tote* (loc. sg. fem.) + postposition *-em* + postposition *-e*.

## 4.4 Numerals

Lack of evidence prevents a comprehensive treatment of numerals in Sabellian. Cardinal numbers are well represented only by “two” and “three,” which inflected for gender, case, and number: Umbrian **sif trif** “three sows” (acc. pl. fem.), **triia tefra** “three pieces of burnt offering” (acc. pl. neut.). The number “four” **pettiur** is found on one Oscan inscription (Rix Sa17). Unfortunately, the inscription is fragmentary and the context in which the word occurs is no longer recoverable. The number “twelve” is attested in Umbrian in the form of a copulative compound “ten + two,” *desen-duf* (acc. pl.). Other cardinals can only be pieced together from derived formations. For example, the Umbrian nominal forms **pumpeṛias** “representing 5 decuriae” and **puntes** “groups of five” point to *\*pompe* as the form for the cardinal “five.”

In addition to the cardinals, a few ordinals and multiplicative adverbs are attested. Umbrian has forms for the first three ordinals: **prumum**, *promom* “first” (acc. sg. neut.),

**etre** “second” (dat. sg. fem.), and **tertiam-a** “third” (acc. sg. fem.) + postposition **-a**. Multiplicatives are also attested in Umbrian: **sumel** “once,” **duti** “two times,” **triuپر** “three times,” and **nuvis** “nine times.”

## 5. SYNTAX

### 5.1 Case usage

In Sabellian the role of noun phrases in a sentence is denoted by the inflectional feature case. The complements of the verb are marked by nominative case for subject, accusative case for direct object, and dative case for indirect object or beneficiary. Nominative is also used for adjectival and nominal predicates in copular sentences, and accusative case for the objects of certain prepositions and for goal of motion. Vocative is the case of direct address. The remaining oblique case forms, genitive, ablative, and locative, are used for adnominal (genitive = possession, partitive) or adverbial functions (ablative = place from which, source; locative = place where, time when).

### 5.2 Word order

The order of the major constituents in a Sabellian sentence is predominantly Subject–Object–Verb (SOV), but almost all possible permutations of this basic order are attested in inscriptions. Changes from basic SOV order do not affect the grammaticality of a sentence and are usually motivated by considerations of focus (topicalization), prosody (speech rhythm), or aesthetics (style).

The order of elements within a noun phrase depends on the type of modifier. Typically, adjectives occupy postnominal position (Oscan **lígatúis núvlanúis** “legates from Nola”), while genitive noun phrases are placed before the modified noun (Oscan **herekleís fíisnu** “temple of Herakles”), though adjectives can also appear in prenominal position (Oscan **múínikeí tereí** “in common territory”) and genitives can follow their head noun (Sakaraklúm **herekleís** “sanctuary of Herakles”). Numerals and pronominal modifiers are almost invariably placed before the noun (Umbrian **tref hapinaf** “three lambs”; Oscan **eíseí tereí** “in that territory”). Definite relative clauses usually follow the antecedent noun phrase, but there are examples in which the relative clause is preposed; sample relative clauses are given in (14):

#### (14) Relative clauses in Sabellian

- A. **púst.**      **feihúis.**                      **pús.**                      **fisnam.**  
 behind walls-ABL. PL. MASC. which-NOM. PL. MASC. temple-ACC. SG. FEM.  
**amfret**  
 surround-3RD PL. PRES.  
 “Behind the walls which surround the temple” (Oscan Rix CA)
- B. **pafe.**                      **trif.**                      **promom. haburent.**  
 which-ACC. PL. FEM. three-ACC. PL. FEM. first-ADV. will catch-3RD PL. FUT. PERE.  
**eaf.**                                      **acersoniem /**  
 these-ACC. PL. FEM.                      Acedonia-LOC. SG. FEM. + POSTPOSITION  
**fetu**  
 sacrifice-3RD SG. IMPV. II  
 “Which three [victims] they will have caught first, these he shall  
 sacrifice at Acedonia” (Umbrian VIIa 52)



The Sabellian languages possess both prepositions and postpositions, Umbrian exhibiting a good selection of the latter: *-ař* “to, toward”; *-co*, *-ku* “with”; *-en*, *-e*, *-em* “into, to, upon”; *-per* “for”; *-to*, *-ta*, *-tu* “from.” In the other Sabellian languages, however, postpositions are much less common (see Oscan *censtom-en* “for the census” and 4.3.3). In the case of prepositional phrases with adjective modifiers, it is common to find the preposition standing between the adjective and the noun: thus, Umbrian *nertru-co persi* (abl. sg. masc.) “at the left foot,” compare Latin *magnō cum dolōre* (abl. sg. masc.) “with great sorrow.”

### 5.3 Agreement

There are three basic rules of agreement in Sabellian:

1. Pronominal modifiers and adjectives, both attributive as well as predicative, modify their head noun in terms of the inflectional features of gender, number, and case, for example, *sif kumiaf* (fem. acc. pl.) “pregnant sows.”
2. A relative pronoun agrees with the head of its antecedent noun phrase in gender and number, while case is determined by the role of the relative word within its clause (see the examples in [14] above).
3. Verbs are marked for person and number based on the person and number of their subject. So, in the Oscan sentence of (15) below, the verb form *centsazet* “they will assess” (3rd pl. fut.) is marked for third-person plural in order to agree with the nominative plural subject *centstur* “censors.”

- (15) *pon centstur bansae tautam*  
 when-CONJ. censors-NOM. PL. MASC. Bantia-LOC. SG. FEM. people-ACC. SG. FEM.  
*centsazet*  
 “assess”-3RD PL. FUT.  
 “When the censors will assess the people at Bantia” (Oscan Rix TB)

Deviations from these rules of agreement do occur and can usually be attributed to factors such as “agreement through sense.” So, for example, in the following sentence from the Iguvine Tablets the main verb *prusikurent* “they will have proclaimed” (3rd pl. fut. perf.) is marked for plural based on the collective sense of the grammatically singular subject noun phrase *mestru karu* (nom. sg. fem.) “the greater portion” = “majority.”

- (16) *sve mestru karu fratri*  
 if-CONJ. greater-NOM. SG. FEM. portion-NOM. SG. FEM. brothers-GEN. PL. MASC.  
*Atiēriu pure ulu*  
 Atiedian-GEN. PL. MASC. who-NOM. PL. MASC. there-ADV.  
*benurent prusikurent rehte*  
 come-3RD PL. FUT. PERF. proclaim-3RD PL. FUT. PERF. properly-ADV.  
*kuratu eru*  
 has been executed-PERF. PASS. INF.  
 “If a majority of the Atiedian brothers who will have come there will have proclaimed that [the ceremony] has been executed properly” (Umbrian Va 24–26)

### 5.4 Main clauses

The mood of a Sabellian verb in main clauses is semantically determined. Statements of fact take the indicative mood. Subjunctive mood is used for wishes (Oscan *nep pútiad* “([I hope] he is not able”) and for prohibitions (*ni hípid* “let him not hold”). Commands

and prescriptions appear in the imperative (Umbrian *anserio* “observe”; Oscan **factud** “let him make”).

## 5.5 Subordinate clauses

### 5.5.1 Modal distribution

In dependent clauses the distribution of the subjunctive and indicative moods is a function of the type of subordination involved. In indirect commands the subjunctive mood is used as a replacement for the imperative. In Umbrian this type of subordination does not take an introductory conjunction.

- (17) **kupifiatu**                      **rupiname**                      **erus**  
 order-3RD SG. IMPV. II   Rubinia-ACC. SG. FEM. + POSTPOSITION   *erus*-ACC. SG. NEUT.  
**tera**                                      **ene**                      **tra sahta**  
 distribute-3RD SG. PRES. SUBJUNC.   and-CONJ.   Trans Sancta-ACC. SG. FEM.  
**kupifiaia**  
 order-3RD SG. PRES. SUBJUNC.  
 “At Rubinia he shall order him to distribute the *erus* and to give the command  
 at Trans Sancta” (Umbrian Ib 35)

Indirect questions use both indicative and subjunctive depending on whether the event described in the question is considered a fact or a possibility, but there is at least one example, cited in (18), of the use of a subjunctive as a replacement for the indicative mood of the direct question.

- (18) **ehvelklu**                      **feia...**                      **sve**                      **rehte**  
 vote-ACC. SG. NEUT.   take-3RD SG. PRES. SUBJUNC.   if-CONJ.   properly-ADV.  
**kuratu si**  
 execute-PERF. MEDIOPASS. PART. + be-3RD SG. SUBJUNC. (= PERF. PASS.)  
 “Let him take a vote on whether [the ceremony] has been properly executed”  
 (Umbrian Va 23)

The spread of the subjunctive mood at the expense of the indicative appears to have been in progress during the historical period.

### 5.5.2 Subordinating conjunctions

Temporal clauses are introduced by a variety of conjunctions: Umbrian *arnipo* “until”; Umbrian **ape** “when”; Umbrian *ponne*, **pune**, Oscan **pun** “when”; Oscan *pruter pan*, Umbrian *prepa* “before”; Umbrian *post pane* “after.” Adverbial clauses of purpose are signaled by the conjunction **puz** Oscan, *pusi* Umbrian “so that” and a subjunctive mood verb in the subordinate clause. The conjunction meaning “if,” **sve** Umbrian, **svai** Oscan, marks the protasis of a conditional clause.

### 5.5.3 Infinitival complements

Infinitives are used to represent the main verb of a statement that is subordinated in indirect discourse. The subject in the subordinated clause shifts from nominative to accusative case, and the tense of the infinitive is determined by the tense of the verb in direct discourse.

Thus, in the Umbrian example of (19), the perfect periphrastic infinitive is used because the tense of the verb in direct discourse was perfect:

- (19) **prusikurent**                      **rehte**  
 proclaim-3RD PL. FUT. PERF. properly-ADV.  
**kuratu eru**  
 execute-PERF. MEDIOPASS. PART. ACC.SG. NEUT. + be-PRES. ACT. INF. =  
 PERF. PASS. INF.  
 “(A majority of the brotherhood) will have proclaimed that it [the ceremony]  
 has been properly executed” (Umbrian Va 26)

Infinitives also serve as the complements of verbs that have meanings within the semantic range of “wish,” “be necessary,” “be fit,” etc. The examples cited below are from Umbrian.

- (20) A. **pune**                      **puplum**                      **aferum**                      **heries**  
 when-CONJ.    people-ACC. SG. MASC.    purify-PRES. INF.    wish-2ND SG. FUT.  
**avef**                      **anzeriatu**                      **etu**  
 birds-ACC. PL. FEM.    observe-SUPINE    go-2ND SG. IMPV. II  
 “When you will wish to purify the people, go to observe the birds”  
 (Umbrian Ib 10)
- B. *perse*                      *mers*                      *est*                      *esu*  
 if-CONJ.    right-NOM. SG. NEUT.    is-3RD SG. PRES. ACT.    this-ABL. SG. MASC.  
*sorsu*                      *persondru*  
 pig-ABL. SG. MASC.    excellent-ABL. SG. MASC.  
*pihaclu*                      *pihafi*  
 victim of purification-ABL. SG. NEUT.    be purified-PRES. MEDIOPASS. INF.  
 “If it is right that it be purified with this excellent pig as a victim of purification”  
 (Umbrian VIb 31)

Supines are used as complements to verbs of motion; see **anzeriatu** in the first sentence of (20).

### 5.5.4 Sequence of tenses

In indirect commands, indirect questions, adverbial clauses of the purpose type, and subordinate clauses within indirect discourse, the tense of the subjunctive is governed by the tense of the main verb, so-called *consecutio temporum* “sequence of tenses.” Present tense in the main clause requires present tense of the subjunctive in the subordinate clause; past tense in the main requires an imperfect subjunctive in the subordinate clause. In the Oscan example of (21), the verb in the subordinate clause is imperfect subjunctive because the governing verb is in the perfect tense:

- (21) **kúmbened**                      **thesavrúm**                      **pún**  
 agree-3RD SG. PERF.    treasury-ACC. SG. MASC.    when-CONJ.  
**patensíns**                      **múíníkad**                      **tanginúd**  
 open-3RD PL. IMPF. SUBJUNC.    common-ABL. SG. FEM.    consent-ABL. SG. FEM.  
**patensíns**  
 open-3RD PL. IMPF. SUBJUNC.  
 “It was agreed [that] when they opened the treasury they should open it by  
 joint agreement” (Oscan Rix CA)

### 5.5.5 Relative clause formation

There are two important Sabellian syntactic processes that concern relative clause formation – *attraction* and *incorporation*. Attraction refers to the process whereby the antecedent of a relative pronoun is attracted into the case of the relative, or the case of the relative is modified to agree with that of its antecedent (so-called *reverse attraction*). Incorporation refers to movement of the antecedent out of the main clause and into the relative clause.

In the Oscan sentence of (22), both syntactic processes are at work: (i) *ligud*, which serves as the antecedent of the relative pronoun *poizad*, is incorporated into the relative clause; and (ii) the relative pronoun *poizad*, which is the underlying direct object accusative of the verb *anget<.>uzet*, is attracted into the ablative case of the antecedent.

- (22) *censamur.* *esuf... poizad.*  
 assess-3RD SG. PRES. MEDIOPASS. IMPV. II self-NOM. SG. MASC. which-ABL. SG. FEM.  
*ligud / iusc. censtur.*  
 law-ABL. SG. FEM. this-NOM. PL. MASC. censors-NOM. PL. MASC.  
*censaum. anget<.>uzet*  
 assess-PRES. ACT. INF. propose-3RD PL. FUT. PERF.  
 “He himself shall be assessed by the law which these censors shall have proposed to take the census” (Oscan Rix TB)

## 6. THE LEXICON

The basic layer of the Sabellian lexicon is made up of words inherited from Proto-Indo-European. Many of these words are attested in both branches of Italic as well as in other Indo-European languages:

### (23) Sabellian words of Proto-Indo-European origin

- A. “father”: Oscan **pater** NOM. SG. MASC., South Picene **patereih** DAT. SG. MASC., Latin *pater*
- B. “mother”: Oscan **maatreis** GEN. SG. FEM., South Picene **matereih** DAT. SG. FEM., Latin *māter*
- C. “brother”: Umbrian *frater* NOM. PL. MASC., Latin *frāter*
- D. “carries”: Umbrian *ferest* 3RD SG. FUT., Volscian *ferom* PRES. INF., Marrucian *feret* 3RD PL. PRES., Latin *fert* 3RD SG. PRES.
- E. “says”: Oscan **deikum** PRES. ACT. INF., Latin *dicit* 3RD SG. PRES.
- F. “be”: Oscan **súm**, *sim* 1ST SG. PRES., *est* 3RD SG. PRES., Umbrian *est*, Volscian *estu* 3RD SG. IMPV. II, South Picene **esum** 1ST SG. PRES., Pre-Samnite **esum**, Latin *sum*, *est*
- G. “foot”: Umbrian **peri** ABL. SG. MASC., Oscan **pedú** GEN. PL. MASC., Latin *pēs*

Other Sabellian vocabulary items have solid etymological connections with languages in other branches of Indo-European but lack Latino-Faliscan cognates:

### (24) Inherited Sabellian vocabulary not found in Latino-Faliscan

- A. “son”: Oscan **puklui** DAT. SG. MASC., Paelignian *puclois* DAT. PL. MASC., Marsian *pucle[s]* DAT. PL. MASC., Sanskrit *putras*, cf. Latin *filius*
- B. “daughter”: Oscan **futír** NOM. SG. FEM., Greek θυγάτηρ, Sanskrit *duhitā*, cf. Latin *fīlia*
- C. “fire”: Umbrian *pir* NOM./ACC. SG. NEUT., Oscan **purasiái** “having to do with fire” LOC. SG. FEM., Greek πῦρ, English *fire*, cf. Latin *ignis*

- D. “water”: Umbrian **utur** “water” NOM./ACC. SG. NEUT., Greek ὕδωρ, cf. Latin *aqua* but note also Oscan **aapa** “water”
- E. “community”: Oscan *touto* NOM. SG. FEM., Umbrian **totam** ACC. SG. FEM., Marrucian *toutai* DAT. SG. FEM., cf. Venetic **teuta[m]** ACC. SG. FEM., Lithuanian *tauta* “people,” Gothic *piuda* “people,” Old Irish *tuath* “people”

A small set of vocabulary items are restricted to Italic. A substantial number of these shared vocabulary items are associated with religion and ritual practices: for example, Latin *sacer* “sacred,” Oscan **sakrím** “victim” (acc. sg.); Latin *sanctum* “consecrated,” Oscan **saahútum** (acc. sg. neut.); Latin *pius* “obedient,” *piat* “he propitiates,” Volscian *pihom* “religiously unobjectionable” (nom. sg. neut.), Umbrian **pihatu** “let him purify” (3rd sg. impv. II); Latin *feriae* “days of religious observance,” Oscan **fiúsíais** (dat pl. fem.). A few items in this category, however, belong to “secular” levels of the lexicon: thus, Latin *cēna* “dinner,” Oscan **kersnu** (nom. sg. fem.); Latin *habet* “he has, holds,” Oscan **hafiest** (3rd sg. fut.); Latin *ūtī* “to use,” Oscan **úttiuf** “use” (nom. sg.); Latin *familia* “family,” Oscan *famelo* “household” (nom. sg. fem.); Latin *cūrat* “he superintends,” Paelignian *coisatens* (3rd pl. perf.), Umbrian **kuraia** (3rd sg. pres. subjunc.).

Loanwords entered the Sabellian languages from three main sources: Greek, Etruscan, and Latin. The earliest layer of loanwords in Oscan resulted from contact with Greeks and Etruscans in southern Italy. A considerable portion of these loans are the names of deities or their divine epithets: for example, **Herekleís** “Herakles” (gen. sg.), compare Etruscan **hercle**, Greek Ἡρακλῆς; **Herukinaí** (dat. sg.), compare Greek Ἐρυκίνη, epithet of Aphrodite; Ἀππελλουνή *“Apollo”* (dat. sg.), **Appelluneís** (gen. sg.), compare Doric Greek Ἀπέλλων. Outside of *nomina sacra*, there is a handful of cultural borrowings: for example, Oscan **kúiníks** “quarts” (nom. pl.), compare Greek χοῖνιξ “quart (dry measure)”; Oscan **thesavrúm** “storehouse” (acc. sg.), compare Greek θησαυρός. Other words, ultimately of Greek origin, made their way into Sabellian via Etruscan intermediation, for example, Oscan **culchna** (nom. sg.) “kylix,” cf. Etruscan **culizna**, Greek κνίχνα.

Greek loans, particularly the names of divinities, penetrated also into the Sabellian languages of central Italy. A late second-century Paelignian inscription (Ve 213) reveals the names of two Greek divinities: *Uranias* “Urania” (gen. sg.), *Perseponas* “Persephone” (gen. sg.).

Etruscan may be the source for one of the most important sacred terms in Sabellian. The word for “god” that is attested in the central Sabellian languages (Marrucian *aisos* “gods” [nom. pl. masc.], Marsian *esos* [nom. pl. masc.], Paelignian *aisis* [dat. pl. masc.]) and in Oscan (**aisu(s)is** dat. pl. masc.) is based on the root **ais-**, which is the uninflected form of the word in Etruscan, *ais* “god.”

In the third and second centuries BC, as the influence of Roman Latin became progressively more pervasive, Latin loanwords began to appear in all levels of the Sabellian lexicon, but most importantly in the spheres of politics and the law. Oscan and Umbrian public officials appear in inscriptions with the titles of magistracies borrowed from Rome: Latin *quaestor* gives Umbrian **kvestur** (nom. sg.), Oscan **kvaísstur** (nom. sg.); Latin *censor* provides Oscan **keenzstur** (nom. sg.); and Latin *aedilis* is taken over as Oscan **aídil** (nom. sg.). The Oscan word for assembly is replaced by Latin *senātus*, thus Oscan *senateis* (gen. sg.). Oscan *ceus* “citizen” is based on Latin *cīuis*. The Oscan *Tabula Bantina*, inscribed at the beginning of the first century BC, attests a formidable array of borrowings and calques based on Latin legal and political terminology. The borrowings in this text are a barometer of Rome’s growing cultural, political, and linguistic supremacy in first-century Italy and of the Sabellian languages’ declining linguistic fortunes.

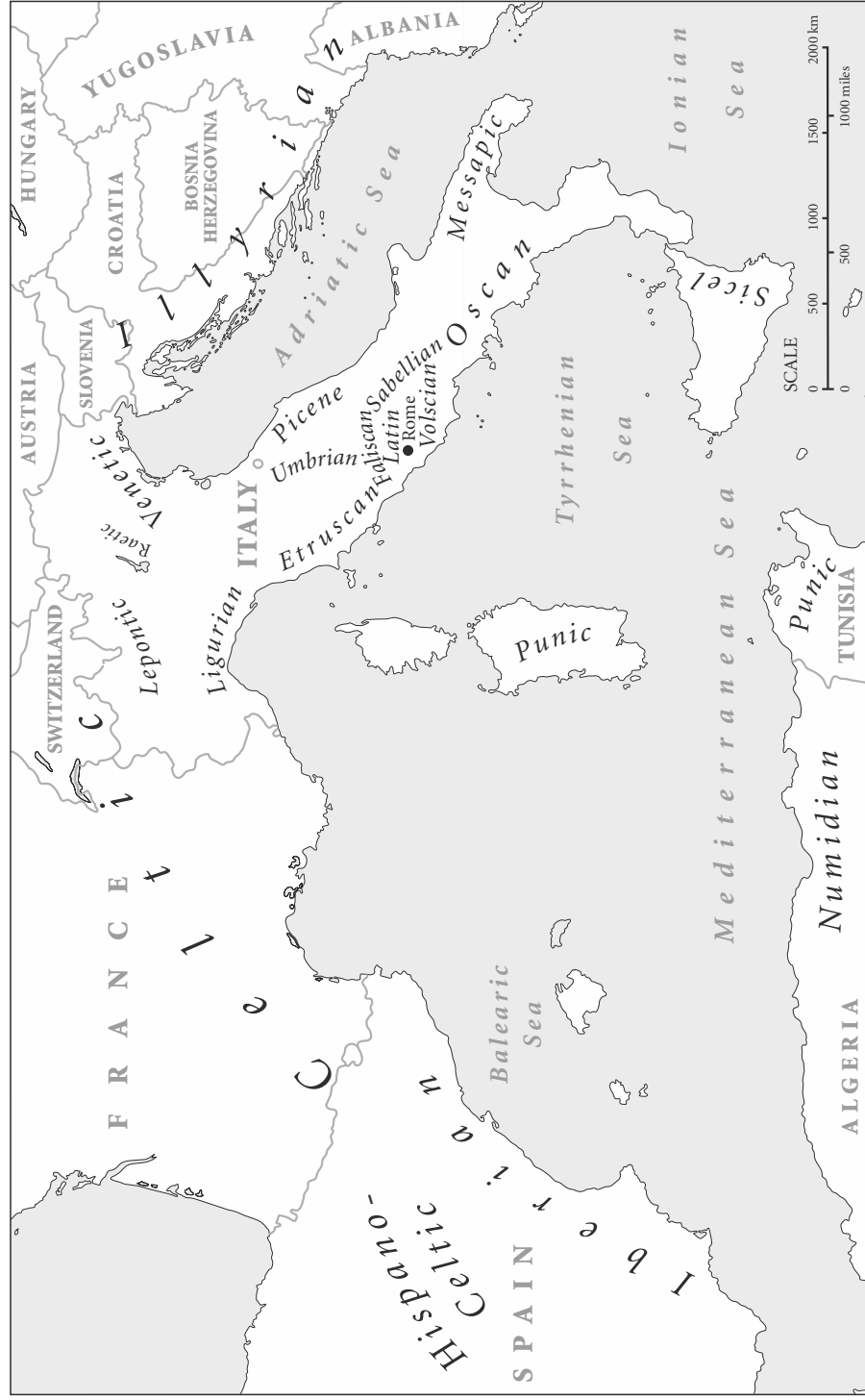
## Bibliography

- Adiego Lajara, I. J. 1990. "Der Archaismus des Sudpikenischen." *Historische Sprachforschung* 103:69–79.
- Benediktsson, H. 1960. "The vowel syncope in Oscan-Umbrian." *Norsk Tidsskrift for Sprogvidenskap* 19:157–295.
- Briquel, Dominique. 1972. "Sur des faits d'écriture en Sabine et dans l'ager Capenas. Mélanges de l'Ecole française de Rome." *Antiquité* 84:789–845.
- Buck, C. Darling. 1928. *A Grammar of Oscan and Umbrian* (2nd edition). Boston: Ginn.
- Colonna, G. 1974. "Nuovi dati epigrafici sulla protostoria della Campania." *Atti della riunione scientifica*, pp. 151–169.
- Cowgill, W. 1970. "Italic and Celtic superlatives and the dialects of Indo-European." In G. Cardona, H. Hoenigswald, and A. Senn (eds.), *Indoeuropean and Indoeuropeans*, pp. 113–153. Philadelphia, PA: University of Pennsylvania Press.
- . 1973. "The source of Latin *stāre*, with notes on comparable forms elsewhere in Indo-European." *Journal of Indo-European Studies* 1:271–303.
- . 1976. "The second plural of the Umbrian verb." In G. Cardona and N. H. Zide (eds.), *Festschrift for Henry Hoenigswald*, pp. 81–90. Tübingen: Narr.
- Cristofani, M. 1979. "Recent advances in Etruscan epigraphy and language." In D. and F. R. Ridgway (eds.), *Italy before the Romans: The Iron Age, Orientalizing and Etruscan Periods*, pp. 373–412. London/New York/San Francisco: Academic Press.
- Durante, M. 1978. "I dialetti medio-italici." In A. L. Prosdocimi (ed.), *Lingue e dialetti dell'Italia antica*, pp. 789–823. Popoli e civiltà dell'Italia antica VI. Roma: Biblioteca di Storia Patria.
- García-Ramon, J. L. 1993. "Zur Morphosyntax der passivischen Infinitive im Oskisch-Umbrisch." In H. Rix (ed.), *Oskisch-Umbrisch. Texte und Grammatik. Arbeitstagung der Indogermanischen Gesellschaft und der Società Italiana di Glottologia vom 25. bis 28. September 1991 in Freiburg*, pp. 106–124. Wiesbaden: Dr. Ludwig Reichert.
- Gusmani, R. 1966. "Umbrisch *pihafi* und Verwandtes." *Indogermanische Forschungen* 71:64–80.
- . 1970. "Osco sipus." *Archivio glottologico italiano* 55:145–149.
- Jeffers, R. 1973. "Problems with the reconstruction of Proto-Italic." *Journal of Indo-European Studies* 1:330–344.
- Jones, D. M. 1950. "The relation of Latin to Osco-Umbrian." *Transactions of the Philological Society*, pp. 60–87.
- . 1962. "Imperative and jussive subjunctive in Umbrian." *Glotta* 40:210–219.
- Joseph, B. D. and R. E. Wallace. 1987. "Latin *sum* and Oscan *súm*, *sim*, *esum*." *American Journal of Philology* 108:675–93.
- Lejeune, M. 1949. "Sur le traitement osque de \* -ā final." *Bulletin de la Société Linguistique de Paris* 45:104–110.
- . 1970. "Phonologie osque et graphie grecque." *Revue des Études Anciennes* 72:271–316.
- . 1975. "Réflexions sur la phonologie du vocalisme osque." *Bulletin de la Société Linguistique de Paris* 70:233–251.
- Marinetti, A. 1981. "Il sudpiceno come italico (e come 'sabino?'): nota preliminare." *Studi etruschi* 49:113–158.
- . 1985. *Le iscrizioni sudpicene. I: Testi*. Lingue e iscrizioni dell'Italia antica 5. Firenze: Leo S. Olschki.
- Meiser, G. 1986. *Lautgeschichte der umbrischen Sprache*. Innsbruck: Institut für Sprachwissenschaft der Universität Innsbruck (IBS 51).
- . 1987. "Pälignisch, Latein und Südpikenisch." *Glotta* 65:104–125.
- . 1993. "Uritalisches Modusyntax: zur Genese des Konjunktiv Imperfekt." In H. Rix (ed.), *Oskisch-Umbrisch. Texte und Grammatik. Arbeitstagung der Indogermanischen Gesellschaft und der Società Italiana di Glottologia vom 25. bis 28. September 1991 in Freiburg*, pp. 167–195. Wiesbaden: Dr. Ludwig Reichert.
- Negri, M. 1976. "I perfetti oscoumbri in -f-." *Rendiconti dell'Istituto Lombardo di Scienze e Lettere, classe di lettere e scienze morali e storiche* 110:3–10.
- Nussbaum, A. 1973. "*Benuso*, *couortuso*, and the archetype of Tab. IG. I and VI–VIIa." *Journal of Indo-European Studies* 1:356–369.

- . 1976. "Umbrian 'pisher." *Glotta* 54:241–253.
- Olzsha, K. 1958. "Das umbrische Perfekt auf -nki-." *Glotta* 36:300–304.
- . 1963. "Das f-Perfektum im Oskisch-Umbrischen." *Glotta* 41:290–299.
- Poccetti, P. 1979. *Nuovi documenti italici a complemento del Manuale di E. Vetter*. Pisa: Giardini.
- Porzio Gernia, M. L. 1970. "Aspetti dell'influsso latino sul lessico e sulla sintassi osca." *Archivio glottologico italiano* 55:94–144.
- Poultney, J. W. 1959. *The Bronze Tables of Iguvium*. Baltimore: American Philological Association.
- Prodocimi, A. L. 1978a. "L'Oско." In A. L. Prodocimi (ed.), *Lingue e dialetti dell'Italia antica*, pp. 825–911. Popoli e civiltà dell'Italia antica VI. Roma: Biblioteca di Storia Patria.
- . 1978b. "L'Umbro." In A. L. Prodocimi (ed.), *Lingue e dialetti dell'Italia antica*, pp. 587–787. Popoli e civiltà dell'Italia antica VI. Roma: Biblioteca di Storia Patria.
- . 1984. *Le tavole iguvine, I*. Firenze: Leo S. Olschki.
- Rix, H. 1976a. "Die umbrischen Infinitive auf -fi und die urindogermanische Infinitivendung \*-dhiōi." In A. Morpurgo Davies and W. Meid (eds.), *Studies in Greek, Italic, and Indo-European Linguistics offered to Leonard Palmer*, pp. 319–331. Innsbruck: Institut für Sprachwissenschaft der Universität Innsbruck.
- . 1976b. "Subjonctif et infinitif dans les complétives de l'ombrien." *Bulletin de la Société Linguistique de Paris* 71:221–220.
- . 1976c. "Umbrisch **ene...kupifiaia**." *Münchener Studien zur Sprachwissenschaft* 34:151–164.
- . 1983. "Umbro e Proto-Oско-Umbro." In E. Vineis (ed.), *Le lingue indoeuropee di frammentaria attestazione. Atti del Convegno della Società Italiana di Glottologia e della Indogermanische Gesellschaft, Udine, 22–24 settembre 1981*, pp. 91–107. Pisa: Giardini.
- . 1985. "Descrizioni di rituali in Etrusco e in Italico. L'Etrusco e le lingue dell'Italia antica." In A. Moreschini (ed.), *Atti del Convegno della Società Italiana di Glottologia, Pisa, 8 e 9 dicembre 1984*, pp. 21–37. Pisa: Giardini.
- . 1986. "Die Endung des Akkusativ Plural commune im Oskischen." In A. Etter (ed.), *o-o-pe-ro-si, Festschrift für Ernst Risch zum 75. Geburtstag*, pp. 583–597. Berlin/New York: de Gruyter.
- . 1992a. "La lingua dei Volsci. Testi e parentela." *I Volsci. Quaderni di archeologia etrusco-italica* 20:37–49.
- . 1992b. "Una firma paleo-umbra." *Archivio glottologico italiano* 77:243–252.
- Rix, H. 2002. *Sabellische Texte. Die Texte des Oskischen, Umbrischen und Südpikenischen*. Heidelberg: C. Winter.
- Rocca, G. 1996. *Iscrizioni umbre minori*. Firenze: Olschki.
- Schmid, W. 1954. "Anaptyze, Doppelschreibung und Akzent im Oskischen." *Zeitschrift für vergleichende Sprachforschung* 72:30–46.
- Untermann, J. 1973. "The Oско-Umbrian preverbs a-, ad-, and an-." *Journal of Indo-European Studies* 1:387–393.
- Untermann, J. 2000. *Wörterbuch des Oskisch-Umbrischen*. Heidelberg: Carl Winter.
- Wallace, R. E. 1985. "Volscian sistiations." *Glotta* 63:93–101.
- Vetter, E. 1953. *Handbuch der italischen Dialekte. I. Texte mit Erklärung, Glossen, Wörterverzeichnis*. Heidelberg: Carl Winter.
- Von Planta, R. 1892–1897. *Grammatik der oskisch-umbrischen Dialekte* (2 vols.). Strasburg: Trübner.







**Map 4.** The Ancient Languages of Italy and Surrounding Regions (for the Greek dialects of Italy and Sicily, see Map 3)

# Venetic

REX E. WALLACE

## 1. HISTORICAL AND CULTURAL CONTEXTS

The Venetic language is attested by approximately 350 inscriptions that have come to light in the territory of pre-Roman Venetia in northeastern Italy. The inscriptions cover a span of nearly five hundred years, dating from the final quarter of the sixth century to the middle of the first century BC. The spoken language did not survive Roman colonial expansion and the spread of Latin into the northeastern portions of the Italian peninsula during the second and first centuries BC. Venetic has no modern descendants.

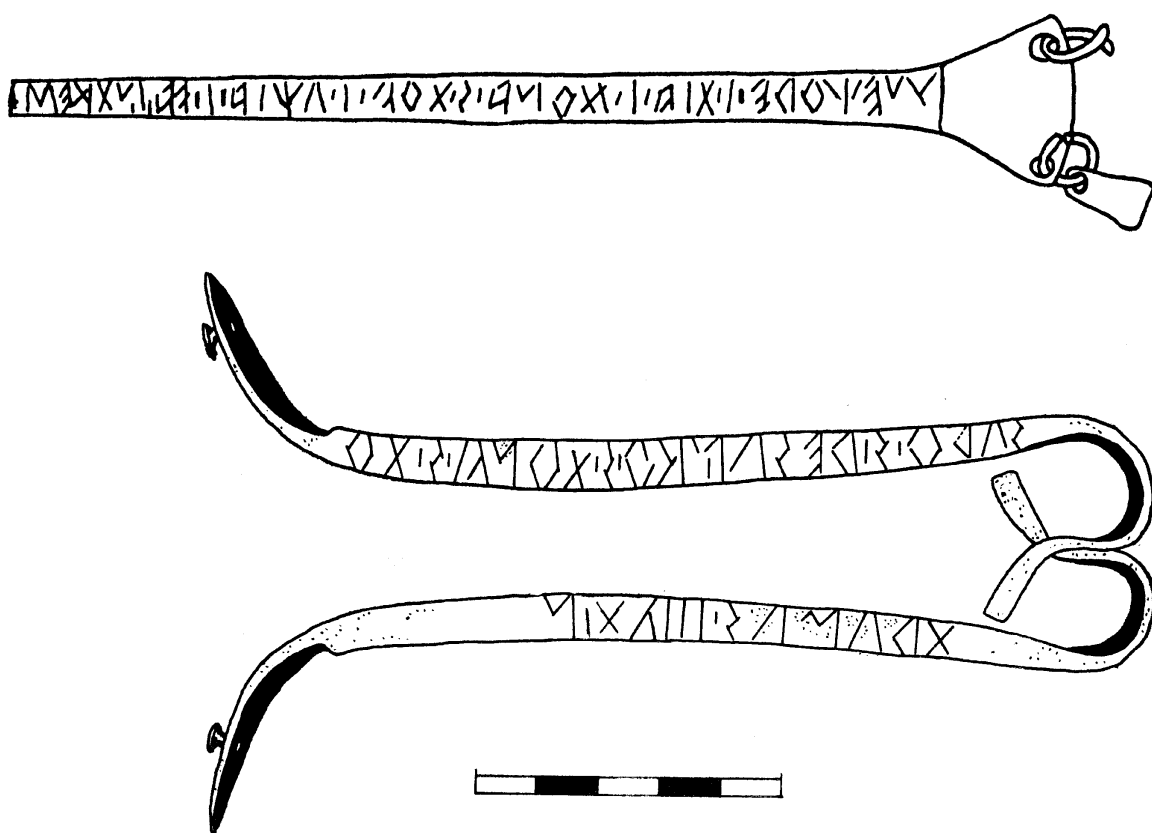
Venetic inscriptions have been found at sites scattered throughout most of pre-Roman Venetia as well as in territories lying to the north and east. The community of Adria, which is situated in the Po River valley a few kilometers inland from the Adriatic Sea, marks the southern limit. The rock inscriptions at Würmlach and the votive texts from Gurina, both sites located in the valley of the Gail River in Austrian Carinthia, mark the northernmost boundary. Venetic inscriptions have been uncovered as far east as Trieste at the head of the Adriatic.

The most abundant source for Venetic inscriptions is the sanctuary of the goddess Reitia at Baratella just east of Este. The religious sanctuary at Lâgole di Calalzo in the valley of the Piave River is another principal source, yielding nearly a quarter of the total number of Venetic texts. Important inscriptions come also from Padova and Vicenza in the south, from Montebelluna and Belluno located along the Piave River, from Oderzo, situated east of the Piave at the head of the Adriatic Sea, and from Gurina in the valley of the Gail River.

According to Livy, the Veneti arrived in northeastern Italy as exiles from the Trojan War. Livy's account of the arrival of the Veneti is fictitious (he was a native of Venetic Padua), but the date of the arrival of Venetic-speaking peoples implied by his tale is likely to be accurate. Archeological evidence points to the development of an independent Iron Age culture in this area shortly after the beginning of the first millennium BC (Fogolari 1988; Ridgway 1979).

The corpus of Venetic inscriptions consists almost exclusively of two epigraphic types, votive inscriptions and funerary inscriptions, with each type accounting for approximately one-half of the total number of inscriptions.

Votive texts were inscribed on objects such as bronze plaques, small replicas of alphabetic tablets, bronze writing implements, and the handles of bronze pails, all of which were commissioned for dedication at religious sanctuaries. The following are typical votive inscriptions (see Fig. 34.1). Inscriptions in the native Venetic alphabet are printed in bold-face type; those in the Latin alphabet are in italics. Inscriptions are cited from Pellegrini and Prosdocimi 1967 = PP; Prosdocimi 1978 = P\*.



**Figure 34.1** Venetic votive inscriptions

A, Este, PP Es 57, bronze stylus

**me go re.i.tiia.i. dona.s.to vhugia.i. va.n.tkeni [a]**

"me"-acc. sg. "Reitia"-dat. sg. fem. "gave"-3rd sg. past "Fugia"-dat. sg. fem.

"Vantkenia"-nom. sg. fem. "Vantkenia gave me [as a gift]  
to Reitia on behalf of Fugia"

B, Lagole, PP Ca 7, bronze handle

**suro.s. resun.k.o.s. tona.s.to trumus.i.iatin**

"Suros"-nom. sg. masc. "Resunkos"-nom. sg. masc. "gave"-3rd sg. past "Trumusiats"-acc. sg. fem "Suros Resunkos gave [me as a gift]  
to Trumusiats"

The oldest Venetic funerary inscriptions from Este are incised on stone cippi in the shape of obelisks (see Fig. 34.2A). Inscribed funerary stelae with figures sculpted in relief are characteristic of Padova (see Figure 34.2B). Less impressive, but more numerous, are the funerary inscriptions scratched on the bodies or on the covers of terracotta vases that served as repositories for the ashes of the deceased (see Fig. 34.2C).

In addition to the aforementioned epigraphic types, a few inscriptions have been found, less than ten in number, that belong to other epigraphical categories. For example, PP Pa 19 is a manufacturer's advertisement stamped on a large storage container (dolium), **keutini/ceutini** "[from the workshop] of Keutinos."

Dating Venetic inscriptions is often problematic because the archeological contexts in which they were discovered were not adequately recorded. In lieu of dating by archeological criteria, most Venetic texts are dated, albeit very roughly, on the basis of a few key

characteristics of the writing system. Venetic texts from Este are divided into four chronological periods based on these orthographic/paleographic features:

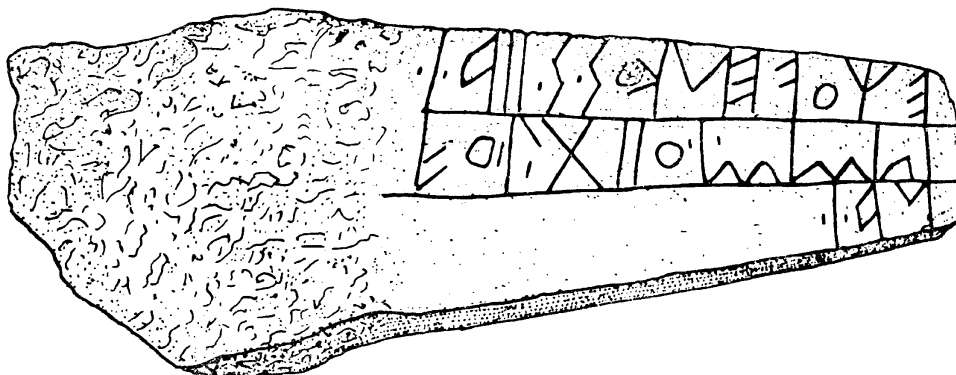
- (1) Archaic, c. 525–475 BC (no syllabic punctuation)
- Ancient, c. 475–300 BC (syllabic punctuation, /h/ = 𐌒, 𐌓)
- Recent, c. 300–150 BC (innovative /h/ = ·𐌒·)
- Latino-Venetic, c. 150–50 BC (use of Latin alphabet)

Venetic is a member of the Indo-European language family, but its often-mentioned affiliation with the languages of the Italic branch, in particular with Latin, is difficult to determine. On the basis of existing evidence the precise position of Venetic within Indo-European remains an open question (see Beeler 1981; Carruba 1976; Euler 1993; Lejeune 1974:173; Polomé 1966:71–76; Untermann 1980:315–316).

Venetic shows the “centum”-style treatment of the Proto-Indo-European (PIE) dorsal stops. The Proto-Indo-European palatals and velars merge as velars (Venetic **ke** “and” from PIE *\*k<sub>e</sub>*; Venetic **lo.u.ki** “grove” from PIE *\*lowkos*); but there is a distinctive reflex for Proto-Indo-European labiovelars (Venetic **-kve** “and” from PIE *\*k<sup>w</sup>e*).

A third-person singular mediopassive ending in **-r** may also be attested, but the verb forms that have this suffix appear to be functionally active (transitive) rather than mediopassive, for example, **tuler donom** “brings/brought (?) a gift [as an offering].”

Several features that are common to the Indo-European languages of the west are found in Venetic. The Proto-Indo-European laryngeal consonants appear as **a** in Venetic in the environment between consonants, as in Italic and Celtic: for example, Venetic **vha.g.s.to** “set up [as an offering],” Latin *facit*, Oscan **fakiiad**, all from the zero-grade of the Proto-Indo-European root *\*d<sup>h</sup>eh<sub>1</sub>-* with *\*k-* extension (< *\*d<sup>h</sup>h<sub>1</sub>-k-*). Venetic probably also shares with



**Figure 34.2** Venetic epigraphs

A, Este, PP Es 2, cippus

**ego vhu.k.s.sia.i. vo.l.tio.m.mnina.i.**

“I”-nom. sg. “Fugsya”-dat. sg. fem. “Voliomnina”-dat. sg. fem. “I [belong] to Fugsia Voltiomnina”

B, Padova, PP Pa 2, stele

**plede.i. ve.i.gno.i. kara.n.mniio.i. e.kupetari.s. ego**

“Pledes”-dat. sg. masc. “Veignos”-dat. sg. masc. “Karanmnis”-dat. sg. masc.

“funerary monument?”-nom. sg. “I”-nom. sg. “I [am] the *ekupetaris* (funerary monument ?) belonging to Pledes Veignos Karanmnis”

C, Este, PP Es 77, terracotta vase

**va.n.t.s.a.froi**

“Vants”-nom. sg. masc. “Afros”-dat. sg. masc. “Vants, for Afros”

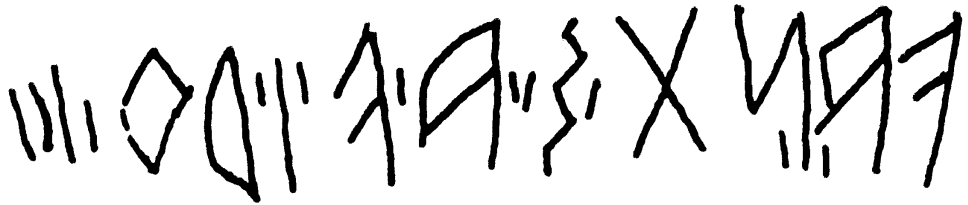


Figure 34.2 (cont.)

Latin and with Celtic an *o*-stem genitive singular ending *-ī*, for example, **keutini** “[from the workshop] of Keutinos,” **lo.u.ki** “of grove,” but the diagnostic importance of this isogloss is not securely established (see §4.1.1).

Linguistically, Venetic inscriptions are relatively homogeneous over the whole of the Venetic-speaking world. Very little evidence points to regional dialect differences. There is, however, one phonological isogloss that can be extracted from extant documents, namely the treatment of the nasals **n** and **m** in word-final position.

In prehistoric Venetic, *\*m* and *\*n* merged as **n** in word-final position throughout most of Venetia, for example, **.e.kvo.[.n.]** “horse” < *\*ēkwom*. However, in the valley of the Piave River near Cadore the spelling of word-final nasals is in a state of flux. The bilabial **m** appears regularly in word-final position in inscriptions incised in the local writing system, for example, **dono.m.** “gift.” However, a few inscriptions written in a system strongly influenced by the one used at Este show **n** in this position, thus, **donon..** Regardless of how the **m** versus **n** variation is to be explained diachronically (preservation of original **m** with **n** introduced via dialect borrowing or reintroduction of **m** through contact with Latin speakers), this is a phonological feature that serves to set off the region of Cadore from the rest of Venetia during the Recent Venetic period.

2. WRITING SYSTEMS

2.1 The Venetic alphabet

Venetic texts prior to the Latino-Venetic period were written in an alphabetic script that was introduced into southern Venetia from Etruria during the first half of the sixth century BC (Cristofani 1979:388–389; Pandolfini and Prosdocimi 1990:244–289). The source of the Venetic alphabet was a northern Etruscan script of the “reformed” type, namely one that had eliminated the letters *beta*, *gamma*, *delta*, and *omicron* from the canonical list of letters forming the teaching alphabet. The fact that the Etruscan alphabet introduced into Venetia lacked these letters forced those responsible for adaptation to use the letters *phi*, *khi*, and *theta* with a new function, namely as signs for the voiced stops /b/, /g/, and /d/ respectively. At some point during the formative stages, the letter *omicron* was “reborrowed,” most likely from a Greek source, and added to the very end of the alphabetic series, thus yielding the earliest form of the native Venetic alphabet, the so-called *alphabet princeps* (see Table 34.1).

Local differences in the spelling of the dental stop phonemes /t/ and /d/ developed during the latter half of the sixth century and the first decades of the fifth century as the *alphabet princeps* spread throughout Venetia. Other communities altered the spelling of the *alphabet princeps* in diverse ways, thus giving rise to the local writing traditions attested by Venetic inscriptions (see Table 34.2).

During the Recent Venetic period (c. 300–150 BC) orthographic changes and stylistic developments that altered the shapes of certain letters introduced greater geographical diversity into Venetic orthography.

One interesting diachronic change concerns the spelling of the labiodental phoneme /f/. In the northern Etruscan writing system of the sixth century, /f/ was spelled by

Table 34.1 Venetic <i>alphabet princeps</i> (c. 550 BC)	
A	a
Э	e
Ɑ	v
Ɱ	z
Ɐ	h
Ɒ	d
ⱱ	i
Ⱳ	k
ⱳ	l
ⱴ	m
Ⱶ	n
ⱶ	p
ⱷ	s
ⱸ	r
ⱹ	s
ⱺ	t
ⱻ	u
ⱼ	b
ⱽ	g
Ȿ	o
Ɀ	spelling for /f/



**Table 34.2 Spelling of Venetic dental stops**

	/t/	/d/
Este, Lågole	×	✕
Padova	⊙, ◇	↗
Vicenza	×	↗
Cadore	×	↘

means of a digraph **vh** 𐌶𐌵. Venetic inherited and maintained this digraphic spelling in most local writing traditions. However, at Cadore, after the sound /h/ was lost, the digraphic spelling of /f/ was simplified to *heta* 𐌺 = /f/, for example, 𐌸𐌶𐌶𐌶𐌺𐌺 = /futtos/ (PP Ca 15).

Inscriptions from Este dating to the period before 300 BC typically show the letter *heta* in its older shapes 𐌵, 𐌶. Near the end of the fourth century **h** is stylistically streamlined to a form without horizontal strokes ·l̥, a form that is all but identical to *iota* with syllabic punctuation (see §2.2). For example, the personal name **vhaba.i.tonia** shows both **i** (iota with syllabic punctuation) and **h** with precisely the same form ·l̥. The motivation for this stylistic change is not clear, but the innovative form of **h** spread rapidly from Este throughout most of Venetia during the first decades of the Recent Venetic period. Interestingly, this innovation failed to gain a foothold at Lågole and at Idria in the Julian Alps. Even more remarkable is the fact that at Idria the letter **h** 𐌵 was used with the same functions that the letter ·l̥ had in other local writing systems: it represented the second part of *i*-diphthongs and the second part of the digraphic spelling of the sound /f/, for example, **la.i.v.na.i** = 𐌶𐌵𐌶·𐌶𐌵𐌶𐌶 (PP Is 1).

Venetic inscriptions were written *scriptio continua*, without spaces separating words, though in modern copies of the texts word breaks are generally indicated. The most common direction of writing was sinistroke, but dextrograde writing was not unusual. A few Venetic inscriptions were written in boustrophedon style (“as the ox plows”), with every other line alternating in direction. The precise layout and arrangement of inscriptions on obelisks, stelae, and bronze plaques depended to some extent on the aesthetic considerations of the sponsor or of the craftsman responsible for the work (see Fig. 34.2a and b).

## 2.2 Syllabic punctuation

The most striking feature of the Venetic orthography was “syllabic punctuation.” This was a form of punctuation (indicated in transliteration by a period) in which all syllable-initial vowels (word-initial vowels and vowels in hiatus), with the usual exception of **i**, and all syllable-final consonants, including the final element of diphthongs, received a mark in the form of a short vertical stroke or, less often, a point, for example, 𐌶̣, 𐌶̣̣. Punctuation was generally placed both in front of and behind the letter, as noted above, but at Lågole inscriptions are found in which punctuation is marked with a single point, usually placed after the letter.

Syllabic punctuation is not found on the earliest Venetic inscriptions and thus must be a secondary development postdating the borrowing of the alphabet from Etruscans. It appears first on Venetic inscriptions from the fifth century and is an obligatory feature of the writing system from this period onward. The probable source of syllabic punctuation is the scribal school affiliated with the religious sanctuary of Apollo in the southern Etruscan city of Veii (Wachter 1986). Syllabic punctuation is used on votive inscriptions at this sanctuary from c. 600 to c. 500 BC and it is likely that this orthographic feature was introduced into Venetia by



**Figure 34.3** Latino-Venetian inscription

*Frema..Ivantina..Ktulistoi vesces*

"Frema"-nom. sg. fem. "Ivantina"-nom. sg. fem. Ktulistoi-dat. sg. masc. "vesces?"-nom. sg. fem.(?) "Frema Ivantina, for Ktulistos, vesces?"

means of contacts between the scribes of the Etruscan sanctuary at Veii and scribes affiliated with Venetic religious communities. One scenario suggests that syllabic punctuation was first adopted by scribes at the sanctuary of Reitia at Este and then, because of this sanctuary's prominence, spread throughout Venetia via other important religious sanctuaries.

## 2.3 The Latin alphabet

During the final period of the Venetic language, inscriptions ceased to be written in the native Venetic alphabet and were written instead in a Latin alphabet characteristic of the late Roman Republic. These late Venetic inscriptions composed in the Latin script employed the usual features of Republican orthography: dextrograde ductus, punctuation between words, and letter-forms typical of "cursive" Latin orthography (see Fig. 34.3).

## 3. PHONOLOGY

### 3.1 Consonants

The inventory of Venetic consonants consisted of sixteen, possibly seventeen, phonemes. There were two sets of stop consonants with four distinctive points of articulation – labial, dental, velar, and labiovelar. The labials, dentals, and velars had a contrast in voicing. The fricatives in the system were voiceless. The nasals, liquids, and semivowels were voiced. Table 34.3 summarizes the Venetic consonant system:

Table 34.3 The consonantal phonemes of Venetic							
Place of articulation							
Manner of articulation	Bilabial	Labiodental	Dental	Palatal	Velar	Labiovelar	Glottal
<i>Stop</i>							
<i>Voiceless</i>	p		t		k	k <sup>w</sup>	
<i>Voiced</i>	b		d		g		
<i>Fricative</i>		f	s				h
<i>Nasal</i>	m		n				
<i>Liquid</i>							
<i>Lateral</i>			l				
<i>Nonlateral</i>			r				
<i>Glide</i>				y		w	

The consonantal phonemes are illustrated by the examples of (2). Here and throughout the remaining sections of this chapter, the following abbreviations are used in glossing examples: ID (personal name); PT (patronymic); DN (name of a god or goddess).

## (2) Venetic consonantal phonemes

**per.** (“by, through?”) /p/, **te.r.monio.s.** (“of the boundaries”) /t/, **ke** (“and”) /k/, **.e.kvo[.]n[.]** (“horse”) /kʷ/  
**bu.k.ka** (“Bukka” ID) /b/, **de.i.vo.s.** (“god”) /d/, **.e.go** (‘I’) /g/  
**vha.g.s.to** (“he made”) /f/, **donasan** (“they gave”) /s/, **ho.s.tihavo.s.** (“Hostihavos” ID) /h/  
**murtuvoi** (“dead”) /m/, **dono.m.** (“gift”) /n/  
**lo.u.derobo.s.** (“children”) /l/, **re.i.tiia.n.** (“Reitia” DN) /r/  
**iorobo.s.** (“?”) /y/, **vo.l.tiomno.i.** (“Voltiomnos” ID) /w/

In addition to these sounds, the letter *san* M (transcribed by *ś*) probably represented a phoneme distinct from /s/, most likely a palatal fricative /š/ or a dental affricate /tʰ/ (Lejeune 1974:152–157). Unfortunately, neither the status nor the quality of the sound represented by *ś* can be securely determined.

## 3.2 Vowels

There were at least five vowels in the Venetic phonemic inventory, all differing in quality. The writing system did not distinguish vowel quantity but it is possible that Venetic maintained the distinction in length that it inherited from Proto-Indo-European. If so, the Venetic vowel system had a five-way distinction in quality accompanied by a distinction in quantity at each position.

## (3) Venetic vowel phonemes

**vivoi** (“living”) /ī/, **tribus.iiate.i.** (“Tribusiatis” epithet of Reitia) /i/  
*pater* (“father”) /ē/, **te.r.monio.s.** (“of the boundaries”) /e/  
**vhratere.i.** (“brother”) /ā/, **vha.g.s.to** (“he made”) /a/  
**dono.n.** (“gift”) /ō/, **hostihavo.s** (“Hostihavos” ID) /o/  
**.u.** (“on behalf of”) /ū/, **klutiiari.s.** (“Klutiaris” PT) /u/

The simple vowel phonemes listed above were complemented by six diphthongs:

- (4) **de.i.vo.s.** (“gods”) /ei/, **te[.]u[.]ta** (“community, nation”) /eu/  
**bro.i.joko.s.** (“Broijokos” ID) /oi/, **vhouge** (“Fougonta” ID) /ou/  
**.a..i.su.n.** (“god”) /ai/, **augar** (“?”) /au/

Of these the diphthong **eu** was subject to both geographical and chronological restrictions, found in a handful of words from Lâgole and also attested once at Padova. All of the examples date to the Recent Venetic period or later, which makes interference via contact with non-Venetic (Celtic?) speakers a likely culprit (see Lejeune 1974:110–111), though a sound change **ou** > **eu** (geographically restricted?) cannot be ruled out of the picture.

## 3.3 Diachronic developments

The inventory of vowels remained relatively stable throughout the history of the language. However, in the Latino-Venetic period, particularly in Venetic inscriptions written in the Latin alphabet, there is evidence for sporadic monophthongization: **ou** > *o* /ō/, *Toticinai* (dat. sg. fem.), and **ei** > *e* /ē/, *Trumusiata* (dat. sg. fem.). Since **ou** and **ei** develop to /o/ and /e/ in nonurban Latin inscriptions, it is possible that these changes were contact-induced.

The major features of the diachronic phonology of Venetic vowels are the changes affecting the suffix *-yo-*. In the prehistoric period *\*o* was lost before word-final *\*s* in the environment *\*C-yos*; thus, *\*Cyos* > *\*Cis*. Onomastic formations in *\*-yo-*, for example, **ve.n.nonni.s.** (nom. sg. masc.) < *\*-nyos* and **klutiiari.s.** (nom. sg. masc.) < *\*-ryos* illustrate this development. In the historic period, the *i* resulting from loss of *\*o* in this suffix was also lost before word-final *-s*, for example, **.e.ge.s.t.s.** (nom. sg. masc.) < *\*egestis* < *\*egestys*, compare **.e.ge.s.tio.i.** (dat. sg. masc.). This change is characteristic of the Recent Venetic and Latino-Venetic periods, though it seems to have affected different areas of the Venetic-speaking world at different times and with varying degrees of intensity (Lejeune 1974:111–125).

The inventory of consonantal phonemes was subject to reorganization as a result of several phonological changes. The earliest documented change involved the loss of the glottal fricative **h**. The sound disappeared in all Venetic-speaking areas between c. 350 and 300 BC.

By the beginning of the Latino-Venetic period the distinction between **s** and **ś** also seems to have been eliminated. In Venetic inscriptions written in the Latin alphabet, both sounds are represented by means of Latin *sigma*, though it should be kept in mind that the lack of an orthographic distinction here could be attributed to underdifferentiation on the part of the Latin spelling system, Latin having a single sibilant sound in its phonemic inventory.

### 3.4 Accent

No direct evidence is available to determine the accentual system of Venetic. It is possible to infer, however, from the syncope of short vowels in noninitial syllables, that Venetic had a stress accent system with stress positioned on or near the initial syllable (Lejeune 1974:125; Prosdocimi 1978:318).

## 4. MORPHOLOGY

Venetic was, like all ancient Indo-European languages, an inflecting language. Inflectional categories were specified by means of suffixes attached to nominal and verbal stems.

### 4.1 Nominal morphology

The Venetic nominal system, comprising nouns, adjectives, and pronominal forms possesses the inflectional features of case, number, and grammatical gender. There are three genders (masculine, feminine, and neuter) and two numbers (singular and plural). The total sum of cases in the nominal system cannot be securely determined because the extant inscriptions are so few, and because the inscriptions that are attested belong to such restricted epigraphic types. As a result, there are serious gaps in all nominal paradigms. From the evidence at hand, however, it is possible to recognize five cases: nominative, dative, accusative, genitive, and ablative.

#### 4.1.1 Nominal classes

Venetic adjectives and nouns are organized into inflectional classes based on the sound characterizing the stem. There are five vocalic-stem classes: *o*-stems (**ke.l.lo.s.** nom. sg. masc.); *a*-stems (**vhugiia** fem. sg. masc.); *u*-stems (**.a.i.su.n.** “god” acc. sg.); *i*-stems (**trumusijatin** acc. sg. fem.); and *e*-stems (**.e.nogene.s.** nom. sg. masc. vs. **.e.nogene.i.** dat. sg.). The *o*-stems split into subtypes: stems in *-yo-* had the vowel(s) of the nominative singular syncopated,

for example, *yo*-stem **.a.kut.s.** (nom. sg. masc.) < \**akutis* < \**akutyos*, compare **.a.kutiio.i.** (dat. sg. masc.). Consonant-stems had three inflectional types: stop-stems (**va.n.t.s.** nom. sg. masc.); *r*-stems (**lemetore** < **.i.** > dat. sg. masc.); and *n*-stems (**mo.l.do** nom. sg. masc. with loss of final **-n**, compare **pupone.i.** dat. sg. masc.).

(5) Venetic *o*-, *yo*-, and *a*-stems

	<i>o</i> -stems	<i>yo</i> -stems	<i>a</i> -stems
NOM. SG.	<b>vo.l.tiiomno.s.</b>	<b>.a.kut.s.</b>	<b>vhrema</b>
ACC. SG.	<b>.e.kvo[.]n[.]</b>	—	<b>re.i.tia.n.</b>
DAT. SG.	<b>vo.l.tiiomno.i.</b>	<b>.a.kutiio.i.</b>	<b>vh.u.k.s.siia.i.</b>
ABL. SG.	<b>leno</b>	<b>vo.l.tio</b>	—
GEN. SG.	<b>keutini</b>	—	—
NOM. PL.	—	—	—
ACC. PL.	<b>de.i.vo.s.</b>	<b>te.r.monio.s.</b>	—
DAT./ABL. PL.	<i>andeticobos</i>	—	—

(6) Venetic *r*-, *n*-, and stop-stems

	<i>r</i> -stems	<i>n</i> -stems	stop-stems
NOM. SG.	<b>lemetor</b>	<b>molo</b>	<b>va.n.t.s.</b>
ACC. SG.	—	—	—
DAT. SG.	<b>lemetore.i.</b>	<b>pupone.i.</b>	<b>va.n.te.i.</b>
ABL. SG.	—	—	—
GEN. SG.	—	—	—
NOM. PL.	<b>.a.nsores</b>	—	—
ACC. PL.	—	—	—
DAT./ABL. PL.	—	—	—

The evidence for the *o*-stem genitive singular **-i** rests on a small number of forms, almost all of which are problematic in one way or another (see Untermann 1960, 1980). The least controversial example of this case ending is stamped, along with a version in Latin, on the body of a large storage container (PP Pa 19), namely **keutini**, Latin *ceutini*, “[from the workshop] of Keutinos.” But since this inscription belongs to the latest Venetic period, it may not be possible to rule out Latin influence here, even though the name appears to be of local origin (Prosdocimi 1978:303). The only other reasonably good example of this **i**-ending is **lo.u.ki**, which is found on an inscription from Padova (PP Pa 14; Prosdocimi 1979) as the object in a prepositional phrase **.e.n.to.l.lo.u.ki** “within the grove” (/entol/ for \**entos* via assimilation?). Unfortunately, this text and its interpretation are not at all clear and so the analysis of **lo.u.ki** as a genitive must be viewed with some caution.

The publication of an inscription discovered near Oderzo (Prosdocimi 1984 \*Od 7) offers a more interesting entry in the discussion of *o*-stem genitives in Venetic. The text, which is cited below, is incised on an oval-shaped funerary stone. Side (b) has a bipartite onomastic phrase in the nominative case; side (a) is inscribed with a single word.

(7) Oderzo, P \*Od 7, oval-shaped funerary stone

(b) **padros . pompeteguaiois.**

(a) **kaialoiso**

Side (a) has been interpreted as the genitive singular of an *o*-stem idionym **kaialo-** (Gambiari and Colonna 1988:138; Lejeune 1989). This interpretation may prove to be correct

but it is not without difficulties because the Proto-Indo-European form of the *o*-stem genitive singular is *\*-osyo*, not **-oiso** (cf. the Latin *o*-stem genitive singulars *ualesiosio popliosio*). A satisfactory explanation for the change in this putative Venetic ending *\*-osyo* > **-oiso** has not yet been offered (for suggestions, see Gambiari and Colonna 1988:138; Lejeune 1989:64; Eska 1995:42–43). Interestingly, forms with what appear to be the same ending **-oiso** are attested on Lepontic inscriptions (for which, see Gambiari and Colonna 1988; Eska 1995), so that a final determination concerning Venetic **kaialoiso** must be made with due consideration of the Celtic evidence (see now Eska and Wallace 1999).

#### 4.1.2 Pronouns

Venetic inscriptions have thus far yielded only three pronominal forms. Two forms belong to the first-person pronoun: **ego** (nom. sg.) and **megeo** (acc. sg.). The third form is a pronominal adjective *sselboisselboi* “himself” (dat. sg.), which is interesting not only because of its double spelling of the sibilant and its reduplicative structure, but also because of its etymological connection to forms found in Gothic *silba* and Old High German *selbselbo*.

### 4.2 Verbal morphology

Venetic verbs are inflected for tense (present, past), mood (indicative, imperative, and possibly subjunctive), voice (active, mediopassive), person (first, second, third), and number (singular, plural).

#### 4.2.1 Verbal classes

The number of inflectional classes for present tense verbs cannot be determined. The past tense forms **dona.s.to** “gave,” **donasan**, presuppose *a*-stem inflection in the present (**dona-**). **atisteit** “sets up” is customarily analyzed as a present tense form built from the zero-grade of the PIE root *\*steh<sub>2</sub>*- “stand” + prefix **ati-**, but exactly how and with what morphemes the stem **-stei-** has been formed is not at all clear (Lejeune 1974; Prosdocimi 1978; Untermann 1980).

**Dona.s.to**, **donasan**, and **vha.g.s.to** “offered” form their past tense stems by means of a suffix **-s-**, and so may be parsed as **dona-s-to**, **dona-s-an**, **vhag-s-to**. For etymological reasons **doto** “gave” probably also qualifies as a past tense form. In most Proto-Indo-European languages the past tense (aorist) of the verb “give” is a root formation and Venetic **doto** appears to have a similar structure (**do-to**), compare Greek *édōke* (3rd sg. act.), *édoto* (3rd sg. mediopass.) “he gave” and Vedic *adāt* (3rd sg. act.), *adita* (3rd sg. mediopass.) “he gave.”

The tense of the verb forms **tole.r.**, **tule.r.**, **tola.r.** “brought” (?) is more difficult to gauge because the suffixes **-e/a-r** and their functions are not clear. The fact that the verbs **tole.r.**, **tule.r.**, **tola.r.** are used in votive texts, contexts in which past tense forms are preferred to presents by a significantly large margin, points to a past tense formation. However, neither the suffixes **-e/a-**, nor the bare mediopassive ending (?) **-r**, are characteristic of past tense formations.

#### 4.2.2 Verb endings

The inflectional features of person, number, and voice are marked by “personal endings.” The ending for active voice is attested by the third singular **-t** (**atisteit**). It is also likely that the endings were split into sets based on tense stems, a set of primary endings for present and a set of secondary endings for past (sg. pres. **-t**, sg. past **-to**, pl. past **-an**).

The third singular past ending **-to** looks like the secondary mediopassive ending found in Greek *-to* and Sanskrit *-ta*. The Venetic ending may share with these a common etymological source, but it is not clear that it has middle force in Venetic, and it seems to correspond functionally to the active third plural ending **-an**.

#### (8) Venetic verb forms: summary

present	<b>atisteit</b> (“sets up”)
past	<b>dona.s.to</b> (“gave”), <b>donasan</b> <b>vha.g.s.to</b> (“made”), <b>doto</b> (“gave”) <b>tole.r.</b> (?), <b>tola.r.</b> <b>tule.r.</b>

#### 4.2.3 Nonfinite verbals

The nonfinite forms of the verb system are even less well represented than the finite forms. There is one possible example of a present participle in **-nt-**, **horvionte**, but its root form, meaning, and case are not readily apparent. Other participle forms in **-nt-** appear in onomastic formations, for example, **vho.u.go.n.ta.i.** (dat. sg. fem.), **vho.u.go.n.te[i.]** (dat. sg. masc.), both from the root **vhoug-** “flee,” compare Greek *p<sup>h</sup>eúgont-* “flees,” Latin *fugient-* (3rd-*iō*). A Latino-Venetic inscription from the first century (PP Es 113) contains the only possible example in Venetic of a deverbal adjective in *-to-*, *poltos* “distressed.”

### 4.3 Naming constructions

The basic form for personal names, of both women and men, is the individual name or idionym (**va.n.t.s.** masc.; **vhugia** fem.). Additional names were commonly added to the idionym to create two- or three-part onomastic phrases (**suro.s.resu.n.ko.s.** masc.; **va.n.t.s.mo.l.do.n.ke.o.kara.n.mn.s.** masc.).

Some idionyms were originally substantives, and their derivational history is clear. For example, the idionym **vho.u.go.n.t-** is in origin a participial formation in **-ont-** built to the verb root **vhoug-** “flee” (see §4.2.3). *\*domator-*, an idionym presupposed by the derived name **tomatoriio.i.** dat. sg. masc. (initial **t** by distant assimilation?), is built from the stem *\*doma-* by means of an agent noun suffix **-tor**, compare Latin *domitor* “tamer” (< PIE *\*domh<sub>2</sub>-* “tame”).

Feminine idionyms are generally secondary formations. Most are derived from masculine *o*-stem idionyms by replacing the stem-vowel **-o** with **-a**, for example, masculine **vhugiio-** gives feminine **vhugiia**. Feminines built to consonant-stems generally add **-a** to the final consonant of the masculine stem, thus, masculine **vhougont-** provides feminine **vho.u.go.n.ta.**

The forms making up the second and third members of Venetic personal names are derived from idionyms by means of a limited set of suffixes belonging to either *o*-stem (for masculine) or *a*-stem (for feminine) inflection: for example, **-io**: **vho.u.go.n.tio.i.** (dat. sg. masc.); **-ia**: **vhuk.s.sia.i.** (dat. sg. fem.); **-ko**: **ossoko.s.** (nom. sg. masc.); **-ka**: **vho.u.go.n.ttiaka** (nom. sg. fem.); **-(V)nko**: **.a.r.bo.n.ko.s.** (nom. sg. masc.); **-na**: **vho.u.go.n.tna** (nom. sg. fem.); and **-kno**: **bo.i.kno.s.** (nom. sg. masc.).

The familial relationships specified by the second and third members of personal name constructions are the subject of serious disagreement. One of the interpretations currently under debate regards the formations built with **-io/-ia**, **-ko/-ka**, **-kno**, etc. as patronymics



(Lejeune 1974:53–57). Thus, in bipartite constructions the second member of the phrase specified the patronymic of the idionym, for example, **va.n.t.s. mo.l.donke.o.** “Vants, (son) of Moldo,” while in tripartite constructions the third member referred to the grandfather of the idionym, for example, **ka.n.te.s. vo.t.te.i.iio.s. a.kut.s.** “Kantes, (son) of Vottos, (the son) of Akutos” (for a dissenting view, see Untermann 1980).

Feminine constructions having derived forms in **-na** as the second or third member indicate a different type of relationship. The **na**-suffix is specialized to designate the gamonymic (Lejeune 1974:60–63). Thus, in the phrase **ne.r.ka lemeto.r.na**, the second member specifies the “wife of Lemetor.” Three-member constructions, such as **vhugii.a.i. a.n.detina.i. vhuginia.i.**, indicate both gamonymic and patronymic, thus “Fugia, (wife) of Andetos, (daughter) of Fugs.”

## 4.4 Compounds

Several nominal compounds are attested in the Venetic onomastic system. There are native formations such as **ho.s.ti-havo.s.**, **volti-genei**, **vo.l.to-pariko.s.**, and **eno-kleves**, as well as formations of Celtic origin, for example **ve.r.ko.n.darna** < \**Wer-kon-daros*. Outside of the anthroponymic formations, however, the inscriptions give us only a single example of a nominal compound, **.ekvopetari[.].s.** plus variants **.e.kupetari.s.**, **.e.p.petari.s.**, *ecupetaris*, and *equpetars*.

This compound undoubtedly refers to a funerary monument of some type, perhaps for members of an equestrian social class, suggested, of course, by the fact that the first element is the stem **.ekvo-** “horse.” Nevertheless, this compound continues to generate considerable discussion, not only because the second constituent **pet-** has yet to be given a convincing etymological explanation, but also because it is not clear how the variants **.ekvo-**, **e.p.-**, etc. are to be connected to one another, if at all (see Brewer 1985; Lejeune 1971a; Prosdocimi 1978:297–301; Pulgram 1976).

# 5. SYNTAX

## 5.1 Case usage

In typical Indo-European fashion, the role of Venetic noun phrases (NPs) is denoted by the inflectional feature case. The complements of the verb are marked by nominative case for subject, accusative case for direct object, and dative case for indirect object and for beneficiary. The genitive case is used to indicate possession. Accusative and ablative serve as the cases to mark NPs as the objects of prepositions, the case of the object being determined by the preposition: **per** “by, through (?)” and **.u.** “on behalf of” governed the accusative case; **.o.p** “because of (?)” took the ablative.

## 5.2 Word order

Nothing definitive can be said about the underlying order of the major constituents (subject, direct object, verb) in a Venetic sentence. Only votive inscriptions have finite verb forms, and the order of the constituents attested for this sentential type may be the result of syntactic processes such as topicalization.

At Este, *iscrizioni parlanti* (“speaking inscriptions”) are found with SVO (Subject–Verb–Object), OVS, and OSV orders:

- (9) Este, PP Es 48, stylus

SVO: **vhug.siaa vo.l.tio.n.mnin.(a) dona.s.to r(e).i.tia.i. mego**

“Fugsia”-NOM. SG. FEM. “Volutionmnina”-NOM. SG. FEM. “give”-3RD SG. PAST

“Reitia”-DAT. SG. FEM. “me”-1ST PRO. ACC. SG.

“Fugsia, wife of Voltiomnos, gave me to Reitia”

Este, PP Es 54, stylus

OSV: **mego (v)hugia dona.s.to re.i.tia.i.**

“me”-1ST PRO. ACC. SG. “Fugia”-NOM. SG. FEM. “give”-3RD SG. PAST “Reitia”-DAT. SG. FEM.

“Fugia gave me to Reitia”

Este, PP Es 53, stylus

OVS: **mego dona.s.to re.i.tia.i. ner(.)ka lemeto.r.na**

“me”-1ST PRO. ACC. SG. “give”-3RD SG. PAST “Reitia”-DAT. SG. FEM. “Ner(i)ka”-NOM. SG. FEM. “Lemetorna”-NOM. SG. FEM.

“Nerka, wife of Lemetor, gave me to Reitia”

Numerically, OVS is the most prominent, followed by OSV. These orders could be the result of the movement of the direct object pronoun **mego** “me” into sentence-initial position, which is a common position for the first-person pronoun in votive inscriptions of this type in all of the languages of ancient Italy. As a result, it is quite possible that the basic order at Este was SVO, which has the smallest actual number of attestations, and that the various permutations of this basic order are the result of syntactic movement rules: SVO becomes OSV by fronting the direct object, OVS by subject–verb inversion. This would bring the basic order of the major constituents at Este in line with what is attested for votive inscriptions at Lagolè (Berman 1973).

The order of elements within a noun phrase depends upon the type of modifier present. As far as can be determined, adjectives are generally positioned before the head noun (**te.r.mon.io.s. de.i.vo.s.** “gods of the boundary”?). In onomastic noun phrases, however, the patronymic and gamonymic modifiers followed the idionym (**vhugia.i. a.n.detina.i. vhuginia.i.** “Fugia, (wife) of Andetos, (daughter) of Fugs”).

### 5.3 Agreement

The Venetic verb is marked with an inflectional ending which agrees with its subject in number and person (third person unless a pronominal non-third-person subject is used); thus, below, the verb **doto** takes the third singular ending **-to**, having the singular subject **vhrema.i.s.tina**.

- (10) Este, PP Es 41, stylus

**vhrema.i.s.tina doto re.i.tia.i.**

“Fremaistina”-NOM. SG. FEM. “gave”-3RD SG. PAST ACT. “Reitia”-DAT. SG. FEM.  
(a divinity)

“Fremaistina gave [me] to Reitia”

Agreement is also found in Venetic noun phrases. An attributive adjective agrees with its head noun in case, number, and gender, for example, **te.r.mon.io.s. de.i.vo.s.** (masc. acc. pl.) “gods of the boundary” (?). In onomastic phrases the modifiers of the idionym similarly show agreement (see §5.2).

## 5.4 Coordination

Unfortunately, Venetic inscriptions do not attest any examples of sentential subordination. There is, however, some evidence for coordination. Coordinate noun phrases and coordinate sentences were linked by one of two conjunctions, **kve** or **ke**. The two forms appear to be functionally similar but differ in terms of their syntax. **kve** is judged to be enclitic on etymological grounds (**vivoi oliialekve murtuvoi** “for [him] living and **oliiale** (?) dead”); **ke** may have been proclitic (**.<a>.i.mo.i. ke lo.u.derobo.s.** “for Aimos and [her] children”).

## 6. LEXICON

Apart from personal names and theonyms the number of vocabulary items in the known Venetic lexicon amounts to approximately fifty words. So few lexemes cannot provide an adequate picture of the lexicon; this condition is only exacerbated by the fact that the vocabulary is drawn basically from two text-types.

The “core” element of the Venetic lexicon consists of those words which have etymological connections to lexemes in other Indo-European languages. The words listed in (11) have solid Indo-European comparanda.

### (11) Venetic words with cognates in Indo-European

**dono.m./dono.n.** ACC. SG. NEUT. “gift,” cf. Latin *dōnum*, Oscan **dūnum** “gift”

**doto** “gave,” cf. Greek *dídōsi* “gives,” *édoto* “gave”

**dona.s.to** “presented (as a gift),” Latin *dōnat* “presents (as a gift),” Oscan

**duunated** “presented (as a gift)”

**vha.g.s.to** “offered,” cf. Latin *facit* “makes,” Oscan **fakiad** “makes”

**<v>hratere.i.** DAT. SG. MASC. “brother,” cf. Latin *frāter* “brother,” Umbrian **frater**

NOM. PL. MASC. “brothers,” Greek *phrētēr* “brotherhood”

**hostei** DAT. SG. MASC. “host,” cf. Latin *hostis* “guest”

**vivoi** DAT. SG. MASC. “living,” cf. Latin *uīuus* “alive,” Oscan **bivus** NOM. PL. MASC. “alive”

**murtuvoi** DAT. SG. MASC. “dead,” cf. Latin *mortuus* “dead”

**kve** “and,” cf. Latin *que* “and,” Greek *te* “and”

In addition to vocabulary with sound Indo-European pedigrees, there is a handful of words with probable etymological connections within Indo-European. For example, the root **vol-**, found in the ablative form **vo.l.tiio**, is most likely connected with the Proto-Indo-European root *\*wel-* “wish, desire.” **vo.l.tiio** is probably an adjective built from a *nomen actionis* *\*wl-ti-* (Lejeune 1974:88). Similarly, the root **mag-**, which forms the base of the Venetic noun **magetlon**, **mag-** plus instrumental suffix **-(e)tlo-**, referring in all likelihood to an offering of some type, may be etymologically connected with the root attested in Latin *mactus* “honored, adored.”

Venetic also has a small cache of vocabulary items that are without Indo-European etymologies. An interesting example is the nominal form *vesces* (nom. sg.), **ve.s.ke.ś.** (nom. sg.), **ve.s.ketei** (dat. sg.), which is used as either an attribute of, or an appositional noun phrase referring to, masculine and feminine names. The meaning of this form remains unclear, at least in part because it lacks an etymological connection within Indo-European (for an attempt, see Lejeune 1973; contra, see Untermann 1980). The Venetic noun **.a.i.su.n.** (acc. sg.), **.a.i.su.s.** (acc. pl.), which is assigned the meaning “god(s)” on the basis of comparison

with forms found in the Sabellian languages, e.g., Paelignian *aisis* “gods,” Marrucinian *aisos*, etc., could be a western Indo-European formation. However, it is worth noting that the stem **ais-** is also found in Etruscan (*ais, eis* “god”) and may well have been borrowed into Venetic and Sabellian through contact with Etruscan speakers.

During the second and first centuries BC, Roman presence in territories beyond the Po Valley intensified. One result of contact between Romans and the Veneti was the introduction of Latin loanwords into Venetic. The best examples are *miles* “soldier” and **liber.tos**, “freedman.” It is also worth mentioning that the kinship term *filia* “daughter,” which is often assumed to be a native Venetic word (Lejeune 1967), may actually be a loan from Latin. The inscription on which this word appears is incised in a Latin alphabet and can thus be dated to c. 150–50 BC. Admittedly, the status of this word in the Venetic lexicon cannot be securely determined on the basis of this inscription alone, but the fact that a loan from Latin cannot be ruled out serves as a reminder that the shift from Venetic to Latin as the language of choice in this area was well underway at this time.

## Bibliography

- Beeler, M. S. 1981. “Venetic revisited.” In Y. Arbeitman and A. Bomhard (eds.), *Bono Homini Donum. Essays in Historical Linguistics in Memory of J. Alexander Kerns*, part I, pp. 65–71. Amsterdam: John Benjamins.
- Berman, H. 1973. “Word order in Venetic.” *Journal of Indo-European Studies* 1:252–256.
- Brewer, W. A. 1985. “Notes on Venetic .e.kvopetari.s.” *Zeitschrift für vergleichende Sprachforschung* 98:54–58.
- Carruba, O. 1976. “La posizione linguistica del venetico.” *Athenaeum, fascicolo speciale*, pp. 110–121.
- Cristofani, M. 1979. “Recent advances in Etruscan epigraphy and language.” In D. and F. R. Ridgway (eds.), *Italy before the Romans: The Iron Age, Orientalizing and Etruscan Periods*, pp. 373–412. London/New York/San Francisco: Academic Press.
- Eska, J. F. 1995. “Observations on the thematic genitive singular in Lepontic and Hispano-Celtic.” In J. R. Eska, R. Geraint Gruffydd, and N. Jacobs (eds.), *Hispano-Gallo-Brittonica. Essays in Honour of Professor D. Ellis Evans on the Occasion of his Sixty-fifth Birthday*, pp. 33–46. Cardiff: University of Wales Press.
- Eska, J. F. and R. Wallace. 1999. “The linguistic milieu of \*Oderzo 7.” *Historische Sprachforschung* 109:122–136.
- Euler, W. 1993. “Oskisch-Umbrisch, Venetisch und Lateinisch.” In H. Rix (ed.), *Oskisch-Umbrisch. Texte und Grammatik. Arbeitstagung der Indogermanischen Gesellschaft und der Società Italiana di Glottologia vom 25. bis 28. September 1991 in Freiburg*, pp. 96–105. Wiesbaden: Dr. Ludwig Reichert.
- Fogolari, G. 1988. “La cultura.” In G. Fogolari and A. L. Prosdocimi (eds.), *I Veneti antichi. Lingua e cultura*, pp. 15–195. Padova: Studio Editoriale Programma.
- Gambiari, F. G. and G. Colonna. 1988. “Il bicchiere con iscrizione arcaica da Castelletto ticino e l’adozione della scrittura nell’Italia nord-occidentale.” *Studi etruschi* 54:119–164.
- Lejeune, M. 1966. “Le verbe vénète.” *Bulletin de la Société Linguistique de Paris* 61:191–208.
- . 1967. “Fils et Filles dans les langues de l’Italie ancienne.” *Bulletin de la Société de Linguistique de Paris* 62:67–86.
- . 1971a. “Problèmes de philologie vénète. XIV: Les épitaphes *ekupetaris*.” *Revue Philologie* 45:7–26.
- . 1971b. “Sur l’enseignement de l’écriture et de l’orthographe vénète à Este.” *Bulletin de la Société Linguistique de Paris* 66:267–298.
- . 1973. “The Venetic vocabulary of relations between persons.” *Journal of Indo-European Studies* 1:345–351.
- . 1974. *Manuel de la langue vénète*. Heidelberg: Carl Winter.
- . 1989. “Notes de linguistique italique. XXXVIII. Notes sur la dédicace de Satricum; XXXIX. Génitifs en -osio et génitifs en -i.” *Revue des Études Latines* 67:60–77.

- Pandolfini, M. and A. L. Prosdocimi. 1990. *Alfabetari e insegnamento della scrittura in Etruria e nell'Italia antica*. Biblioteca di "Studi etruschi" 20. Firenze: Leo S. Olschki.
- Pellegrini, G. B. and A. L. Prosdocimi. 1967. *La lingua venetica I. Le iscrizioni*. Firenze: Istituto di glottologia dell'Università di Padova – Circolo Linguistico Fiorentino.
- Polomé, E. 1966. "The position of Illyrian and Venetic." In H. Birnbaum and J. Puhvel (eds.), *Ancient Indo-European Dialects*, pp. 59–76. Berkeley/Los Angeles: University of California Press.
- Prosdocimi, A. L. 1972. "Venetico." *Studi etruschi* 40:193–245.
- . 1978. "Il venetico." In A. L. Prosdocimi (ed.), *Lingue e dialetti dell'Italia antica*, pp. 257–389. Popoli e civiltà dell'Italia antica VI. Roma: Biblioteca di Storia Patria.
- . 1979. "Venetico. L'altra faccia di Pa 14, il senso dell'iscrizione e un nuovo verbo." In *Studi in onore di Carlo Battisti*, pp. 279–307. Firenze: Leo S. Olschki.
- . 1984. "Una nuova iscrizione venetica da Oderzo (\*Od 7) con elementi celtici." In *Studi di antichità in onore di Guglielmo Maetzke, volume secondo*, pp. 423–442. Roma: Giorgio Bretschneider.
- . 1988. "La lingua". In G. Fogolari and A. L. Prosdocimi (eds.), *I Veneti antichi. Lingua e cultura*, pp. 225–420. Padova: Studio Editoriale Programma.
- Pulgram, E. 1976. "Venetic. *e.kupethari. s.*" In A. Morpurgo Davies and W. Meid (eds.), *Studies in Greek, Italic, and Indo-European Linguistics: Offered to Leonard R. Palmer on the Occasion of his Seventieth Birthday, June 5, 1976*, pp. 299–304. Innsbruck: Institut für Sprachwissenschaft der Universität Innsbruck.
- Ridgway, F. R. 1979. "The Este and Golasecca cultures: a chronological guide." In D. and F. R. Ridgway (eds.), *Italy before the Romans: The Iron Age, Orientalizing and Etruscan Periods*, pp. 419–487. London/New York/San Francisco: Academic Press.
- Schmidt, K. H. 1963. "Venetische Medialformen." *Indogermanische Forschungen* 68:160–169.
- Untermann, J. 1960 "Zur venetischen Nominalflexion." *Indogermanische Forschungen* 65:140–160.
- . 1961a. *Die venetischen Personennamen*. Wiesbaden: Otto Harrassowitz.
- . 1961b. "Zur venetischen Nominalflexion." *Indogermanische Forschungen* 66:105–124.
- . 1980. "Die venetische Sprache." *Glotta* 58:281–317.
- Wachter, R. 1986. "Die etruskische und venetische Silbenpunktierung." *Museum Helveticum* 43:111–126.

# Continental Celtic

JOSEPH F. ESKA

## 1. HISTORICAL AND CULTURAL CONTEXTS

The term *Continental Celtic* does not refer to a single linguistic entity – it is not a synonym for Gaulish – but to the entirety of the Celtic linguistic documentation from the ancient European continent. At the present time we can distinguish a discrete language called *Hispano-Celtic* (also known as *Celtiberian*), spoken in the north central meseta of the Iberian peninsula, from *Gaulish*, varieties of which were spoken from Asia Minor in the east through central Europe southward into the northern Italian peninsula and extending to the English Channel and eventually, with the Belgic migrations, over it into Britain. The variety of Gaulish spoken around the northern Italian lake district, usually called *Lepontic*, and that spoken in Asia Minor, usually called *Galatian*, are viewed by some as separate languages, though this view has weakened in recent years (Eska 1998b; cf. Uhlich, forthcoming). It is now commonly believed that Hispano-Celtic first separated from the Proto-Celtic speech area in central Europe sometime in the early first millennium BC, and developed henceforth on its own. The remainder of the Proto-Celtic speech area then developed as a dialect continuum as speakers spread across Europe and into Asia Minor. The traditional view is that this continuum subsequently divided into a *Goidelic* branch and a *Gallo-Brittonic* branch, but an increasing number of scholars have begun to stress the prehistoric unity of Insular Celtic as opposed to Gaulish. Galatian and British, the ancestor of the Brittonic languages, are very poorly attested and will not be discussed further herein.

It is often very difficult to date Continental Celtic inscriptions precisely. While there is some evidence for morphological innovations within the periods of attestation of the various languages, individual inscriptions usually can be identified only as earlier or later on the basis of the script employed (earlier inscriptions being engraved in non-Roman scripts), or on other epigraphic or extra-linguistic grounds. The earliest records are found in Lepontic, which is attested from c. 600 BC to the end of the millennium. Cisalpine Gaulish, probably differentiated from Lepontic only chronologically, is attested in eight inscriptions from the last two centuries of the first millennium BC. Transalpine Gaulish, first attested in the third century BC, was engraved in Greek characters until it gave way after the Roman conquest to Roman characters. The language probably ceased to be spoken in the second half of the first millennium AD. Though the last to be attested, from c. 200 BC to the second century AD, Hispano-Celtic is, by and large, the most conservative variety of Continental Celtic. As with Gaulish, earlier and later periods are distinguished through the employment of non-Roman or Roman scripts and other extra-linguistic means.

The various corpora of Continental Celtic are fragmentary and primarily epigraphic. Inscriptions and graffiti are engraved on stone buildings and monuments, metal plaques (usually lead in Gaul and bronze in Iberia), domestic implements, ceramic ware, and coins. The longest inscriptions are legal or magical-religious in substance. Shorter inscriptions include dedications, funeral monuments, proprietary statements, and expressions of various human sentiments and activities concerning, for example, affection, sex, and drinking. Secondary sources for Continental Celtic are individual lexical items or formulae recorded by classical or medieval writers and lexical items borrowed into ancient languages or surviving as substrate forms, especially in the Romance languages, but also in Basque. It is clear from the subject matter of the surviving records that the languages/dialects were in use at all levels of society. Occasionally, marked surface clausal configurations provide some evidence of a higher, poetic or more formal, register.

As mentioned above, it is probable that Lepontic and Galatian are not discrete languages, but regional dialects of Gaulish. Otherwise there is only sporadic evidence that is indicative of dialectal differentiation. Some scholars, in view of the existence of a few forms that have resisted the Gaulish labialization of Proto-Indo-European (PIE)  $*k^w$  to  $p$ , believe that an archaic dialect of the language may have been preserved. However, since these forms are all month or (ultimately) divine names, it is more likely that they resisted the sound change because of their sacred character, as is not uncommon cross-linguistically.

## 2. WRITING SYSTEMS



























As alluded to above, the Continental Celtic languages were recorded in the earlier periods of their attestation, and sometimes entirely, in various local indigenous scripts before the employment of Roman characters was adopted (see Campanile 1983).

### 2.1 Hispano-Celtic

The large majority of the Hispano-Celtic corpus was engraved in an adaptation of the semi-moraic, semi-segmental Iberian script. Stops are noted with moraic characters that do not indicate voicing and include an inherent vocalism neutral to quantity; thus, there are five characters to write, for example, the dental stops  $t/d$  plus each of the five vowels: (transcribed) *Ta*, *Te*, *Ti*, *To*, *Tu*. The remaining consonants – nasals, liquids, semivowels, and sibilants – are noted with segmental characters, as are the vowels, which do not indicate quantity. The forms of the characters attested at Botorrita are given in Table 35.1.





In (1) are listed some alternative characters. The transcription is the traditional one. See further Tovar (1975), Untermann (1975:71–74), and Lejeune (1993).

**Table 35.1 The Celtic adaptation of the Iberian script**

a		Ca		Pa		Ta		m		n	
e		Ce		Pe		Te		l		í	
i		Ci		Pi		Ti		s		ś	
o		Co		Po		To					
u		Cu		Pu		Tu					



## (1) The alternative nasal and rhotic sets

There are a number of points to be noted about the mechanics of the script. There were two geographic zones which employed differing sets of characters to write the nasals. Broadly speaking, <ṁ> and <ṇ> were employed in the west, <m> and <n> in the east. Nasals are sometimes not written before stops; it is probable that this represents the transference of nasality to a preceding vowel (see Eska 2002a). The character <r> is attested only in some late coin legends; it does not contrast with <ṛ>. It now seems clear that <ś> represents PIE \*s unchanged whereas <s> represents \*s in voiced environments and \*d in certain medial environments and in final position. Some scholars, therefore, elect to transcribe ś as <z> or <ḍ> (and hence then elect to transcribe Ṃ as <s> instead of traditional <ś>). It is not yet clear whether this character represents more than one sound (or phoneme). Geminate consonants are written as single. The sequence <ei> is employed to write the inherited diphthong *ei*, and sometimes *e* from unstressed \**i* (perhaps phonetically a raised [e]), as well as the phoneme which continues PIE \**ē* in final syllables, which eventually became *ī* (perhaps phonetically a lowered [i:]).

Owing to the moraic quality of the stop characters, stop + liquid groups are difficult to represent. A variety of solutions are found, as listed in (2):

- (2) A. An empty vowel (having no phonetic reality) may be written which copies the quality of the following phonemic vowel: e.g., *enTaía* /entra:/
- B. The liquid and following phonemic vowel may be metathesized orthographically: e.g., *ConTeṛPia* /kontrebia:/
- C. The liquid may be elided orthographically: e.g., *ConPouTo* /konblowto/

The moraic quality of the stop characters also makes it difficult to determine the manner in which final stops were written; for example, it is unclear whether the third singular primary ending \*-*ti* is continued intact or with the vowel apocopated in the verbal form *aśeCaTī*. Owing to the influence of the segmental character of the Roman script, but prior to its adoption, syllabic characters came to be followed by a separate character denoting the inherent vowel: for example, in *moñiTuuCooś*. On the use of empty vowels in the Celtic adaptation of the script, see De Bernardo Stempel (1996).

The origin of the Iberian script, which was deciphered by Gómez-Moreno (1922), remains a subject of debate (see de Hoz 1983). While it is agreed that there are Phoenician and Greek elements underlying the script, it is uncertain whether they were integrated simultaneously or whether an original script based upon one was renewed with elements of the other.

## 2.2 Lepontic

The entirety of the Lepontic and Cisalpine Gaulish corpora are engraved in variants of the north Etruscan script. The script is segmental, but shares various features of the Iberian script. Neither the voicing of stops nor the quantity of vowels is noted. Nasals are rarely noted before stops; as with this feature in Hispano-Celtic, in which it is sporadic, it is probable that this represents the transference of nasality to a preceding vowel (see Uhlich 1999:280 and 293 and Eska 2002b: 263–269). Table 35.2, adapted after De Marinis (1991:94), records the Lugano script, in which the corpus is engraved. See further Lejeune (1971:8–27; 1988:3–8).

The infrequently attested characters <χ> and <θ> were inserted into the script in order to introduce a voicing distinction for the dental and velar stops. Whether the new character



**Table 35.2 The Lugano script**

6–5 Centuries	Transcription	3–2 Centuries
ⱱ Ⱳ	a	ⱱ Ⱳ
ⱳ	e	ⱳ ⱴ
Ⱶ ⱶ	v	—
ⱷ	z	ⱷ
ⱸ	h	—
ⱹ	θ	—
ⱺ	i	ⱺ
ⱻ	K	ⱻ ⱼ
ⱽ Ȿ	l	ⱽ
Ɀ	m	Ɀ Ȿ
Ⲁ	n	Ⲁ ⲁ
Ⲃ	P	Ⲃ
Ⲅ ⲅ	ś	Ⲅ ⲅ Ⲇ
ⲇ Ⲉ	r	ⲇ Ⲉ ⲉ
Ⲋ ⲋ Ⲍ ⲍ	s	Ⲋ ⲋ
Ⲏ ⲏ	T	Ⲏ
Ⲑ ⲑ	u	Ⲑ
Ⲓ	X	Ⲓ
Ⲕ ⲕ	o	Ⲕ ⲕ Ⲍ

represents the voiceless or voiced stop varies among inscriptions. The phonetic value of the character <v> has been much disputed, but may well represent /ϕ/ from PIE \*p; see Eska (1998a). The character <ś> and the twice attested <z> represent a sound (or two acoustically similar sounds) known as the *tau Gallicum* (see §3.3.1.1).

## 2.3 Gaulish in Greek characters

Prior to the Roman conquest of Transalpine Gaul, the Massiliote Greek script was employed to write in Gaulish. Noteworthy orthographic features of Greek-character Gaulish are the use of the digraph <ου> for Roman <u> and <v>, and the occasional use of <ει> for <i>, <η> for <ε>, and <ω> for <o> (i.e., long-vowel graphemes for short vowels). The *tau Gallicum* sound (see §3.3.1.1) is variously written: <θ, θθ, σ, σσ, τ, ττ, σθ>. See further Lejeune (1985:427–434, 441–446).

## 2.4 Gaulish in Roman characters

Gaulish was engraved in Roman characters in both capitals and cursive script with the expected values. The *i-longa* is frequently attested, but it does not seem to be differentiated in value from <i>; it is now conventionally transcribed as <j>. The *tau Gallicum* sound (see §3.3.1.1) is written with a wide variety of mono-, di-, and trigraphs: <t, tt, th, tth, d, dd, đ, đđ, ts, ds, s, ss, ss, sc, sd, st>. In some later Gaulish inscriptions, the appearance of final <-m> has been attributed to Roman influence (i.e., perhaps the engraver was principally a Latin speaker).

### 3. PHONOLOGY

Since the Continental Celtic languages are not only fragmentarily attested, but also often engraved in scripts which are phonologically ill-suited to them, it is difficult to establish complete phonemic inventories. It is often necessary to rely upon Indo-European and Insular Celtic etymologies to determine the expected phonology of a form. Readers should keep in mind that the descriptions presented in this and subsequent sections may be incomplete.

#### 3.1 Hispano-Celtic

##### 3.1.1 Consonants

The consonantal phonemic inventory of Hispano-Celtic is as follows:

##### (3) Hispano-Celtic consonantal phonemes

	t	k	k <sup>w</sup>
b	d	g	g <sup>w</sup>
m	n		
	s		
	l		
	r		
	y	w	

The sound represented by the character <s> (NB that /s/ is represented by <ś>), whose status as a phoneme remains to be determined, is not included in (3). Phonetic values for it that have been suggested include the fricatives [z] or [ð] (Villar 1995a:65–82) and affricates [tʃ] or [dʒ] (Ballester 1993–1995).

##### 3.1.2 Vowels

The monophthongs and diphthongs of Hispano-Celtic are listed in (4):

##### (4) Hispano-Celtic vocalic phonemes

<i>Monophthongs</i>		<i>Diphthongs</i>	
i ī	u ū	ai	au
e ē	o	ei	eu
a ā		oi	ou

It is possible, but uncertain, that PIE \*ē is preserved in unstressed syllables; the element -*řeś*, which is normally assumed to continue \**h<sub>3</sub>rēǵs* “king,” occurs several times as the second member of compound forms. Elsewhere, PIE \*ē has been raised to merge, at least phonemically, with ī. In some later inscriptions, PIE \*ei has been monophthongized to ē. A gap in the vowel system was caused by the raising of PIE \*ō to ū in mono- and final syllables and its lowering to ā elsewhere. Unstressed \*i has a tendency to be lowered to e: for example, *aře-* “fore-” from \**pr<sub>h</sub>i-*.

### 3.1.3 Consonant clusters

Groups of stop + *s* are routinely written as <ś>, which suggests that such groups assimilated to -*ss*-. The group *ks* appears to have sometimes been preserved, however, at least to judge from Roman character spellings which employ the character <x>. The inherited group \**ln* also assimilates to *ll*. Other groups are generally preserved. Noteworthy is the fact that nasals do not always assimilate to the place of following stops, for example *TinPiTus* from \**dē-en-b-*. The form *ConPouTo* is peculiar since the basic form of the prefix is *kom-* (but see now Eska 2002a: passim).

## 3.2 Lepontic

### 3.2.1 Consonants

The consonantal inventory of Lepontic is set out in (5):

#### (5) Lepontic consonantal phonemes

(ϕ/)	p	t	k	(k <sup>w</sup> )
b		d	g	(g <sup>w</sup> )
m		n		
		s		
		l		
		r		
		y	w	

The sound(s) spelled by the characters <ś> and <z>, usually called the *tau Gallicum*, is not listed in (5), but is discussed at some length below (see §3.3.1.1). Though it is ordinarily considered to continue the sequence \**ts* immediately, <ś> is apparently also used to spell the outcome of the group \**-ksy-* in the accusative singular *našom*, the Lepontic adaptation of the Greek neuter nominative-accusative adjective Νάξιον (Náksion). It is possible that early Lepontic continued PIE \**p* as the bilabial fricative [ϕ] and preserved PIE \**k<sup>w</sup>* in forms such as *Kuašoni*; the latter might, however, contain *g<sup>w</sup>* from PIE *g<sup>wh</sup>*.

### 3.2.2 Vowels

The inventory of Lepontic monophthongs and diphthongs is identical to that of Hispano-Celtic; see (4). The gap in the vowel system is as with Hispano-Celtic (see §3.1.2). PIE \**ei* is preserved in final position, but elsewhere has been monophthongized to *ē*.

### 3.2.3 Consonant clusters

Consonant groups do not assimilate, save for \**-nd-* > *-nn-* and the predecessors of the *tau Gallicum*.

## 3.3 Gaulish

Since the Gaulish corpus is the largest of the Continental Celtic languages and is attested over the longest chronological period, it is difficult to ascertain a synchronic phonemic inventory. Readers should be aware that the phonemic inventory presented in (6) and (7) is a composite.

### 3.3.1 Consonants

The consonantal phonemes of Gaulish appear to be as follows:

#### (6) Gaulish consonantal phonemes

p	t	k	(k <sup>w</sup> )
b	d	g	
m	n		
	s		
	l		
	r		
	y	w	

The labiovelar *k<sup>w</sup>* is preserved only in a few archaic forms.

#### 3.3.1.1 *Tau Gallicum*

The *tau Gallicum* is not included in (6). Based upon the diversity of graphemes with which it is written, it is usually assumed to have been a dental affricate, fricative, or sibilant. This is supported by etymological considerations, as the *tau Gallicum* often immediately continues *\*ts* and *\*ds*, and ultimately *\*st* (including *\*st* < *\*tst* < *\*-t-t-* and *\*-d-t-*). It is commonly believed that the most likely phonetic value for it is [tʰ], but other suggestions include [tʰ], [θ] (or retracted [θ]), [θs], and [tʰ]. It is usually assumed that the *tau Gallicum*, even when written as a di- or trigraph, was a single segment, but in view of the fact that it is cognate with Insular Celtic *-ss-*, it is probable that it often was a geminate. The most complete discussion of the *tau Gallicum* is that of Evans (1967:410–420), but see also Eska (1998c).

### 3.3.2 Vowels

The monophthongs and diphthongs of Gaulish are listed in (7):

#### (7) Gaulish vocalic phonemes

<i>Monophthongs</i>		<i>Diphthongs</i>	
ī ī	u ū	ai	au
e ē	o ō	ei	eu
a ā		oi	ou

The diphthongs *ai* and *eu* appear only in older forms; in later forms, *ai* is contracted to *ī* and *eu* merges with *ou*, which subsequently contracts to *ō*. The diphthong *oi* is attested early, then is contracted to *ī*; it reemerges later, as does *ei* (PIE *\*ei* having become *ē* in Gaulish), as the result of the loss of intervocalic *\*p* and *\*w*. There is a tendency for long diphthongs to shorten: for example, *ā*-stem dative singular *\*-āi* > *-ai* > *ī*; and *u*-stem dative singular (from the locative) *\*-oū* > *-ou*. Unstressed *i* frequently is lowered to *e*: for example, *\*p<sub>h</sub>h<sub>x</sub>í-* “fore-” > *are-*; and dative or instrumental plural *-bi* > *-be*.

## 3.4 Allophonic variation

Though the Continental Celtic languages – as far as the scripts employed will allow – are usually written phonemically, occasional quasi-phonetic orthographies occur which provide some evidence for allophonic variation in Hispano-Celtic and Gaulish.

In Hispano-Celtic, there is a strong tendency towards labialization of *o* to *u* when adjacent to a nonfinal labial: for example, the *o*-stem dative plural is often written <-uPós> and the first plural present ending is written <-mu(s)>. That *-o-* occurs at all may be the product of phonemic or conservative orthography; but the *o*-stem accusative singular *-om*, for example, is always written with <-o->.

In Gaulish, the velar stop /k/ becomes the fricative [x] before *s* and *t*. Mid vowels in hiatus with non-high vowels tend to be raised: for example, *to = me = declaī* < \*lā- + \*-e; compare *coetic* and *cuēt[ic]*, both with prevocalic /ko/-; and Λουερνιος /luernios/ < \*lo-erno- < \*h<sub>2</sub>lop-erno-.

Hispano-Celtic and Gaulish share a tendency for *e* to raise to *i* before nasal + stop clusters (Gaulish more so). It is presumed that in all of Continental Celtic nasals were realized as [ŋ] before velars. This view is supported by Gaulish inscriptions engraved in Greek characters which employ <γγ> (the Greek grapheme for [ŋg]), for example, ΕΣΚΙΥΓΟΡΕΙΣ for [eskiŋgori:ks].

There is substantial evidence for phonetic lenition in both Hispano-Celtic and Gaulish. In Hispano-Celtic, /s/ = <ś> is normally spelled as <s> in voiced environments (perhaps here being [z]). The clearest evidence for phonetic lenition is provided by genitive singular *TuaTeŋós* and nominative plural *Tua[Te]rés* < \*dugater- < \*d<sup>h</sup>uġh<sub>2</sub>ter- “daughter,” which exhibit the change of [g] > [ɣ] > ∅. The absence of indication for voicing or manner of articulation in the Iberian script and the rarity of quasi-phonetic orthography in Roman character inscriptions conceal any further evidence.

In Gaulish, there are two forms which provide evidence for [s] > ∅ / V \_\_ V: dative or instrumental plural *suioŋebe* < \*swesor- “sister”; and *sioxt* < 3rd sg. preterite \*sesog- + -t (base \*seg- “add”; see Eska 1994c). In later Gaulish, [g] also is often deleted intervocally. Gaulish is also well known for orthographic variation between <c> and <g> (similar variation between other homorganic stops is much less common); it remains uncertain whether this represents phonetic or orthographic variation, though, since the large majority of tokens involve the substitution of a voiceless for a voiced stop, Gray (1944:227) may be correct in suspecting that the voiced stop phonemes of Gaulish were phonetically voiceless. This orthographic variation would then be another type of quasi-phonetic orthography. There are also several examples in which /t/ in lenited position is engraved with one of the graphemes employed to write the *tau Gallicum* (see §3.3.1.1): for example, *eđđic* (cf. *etic*) “and”; *gnatha* (cf. *nata*) “daughter”; and *bueđ* (cf. *buēt*) “be” – suggesting that the lenited allophone of /t/ was either identical, or acoustically similar, to the *tau Gallicum* consonant.

### 3.5 Accent

There is little, if any, direct evidence for the placement of stress in any of the Continental Celtic languages. In Hispano-Celtic, the failure of final *-m* to labialize a preceding *-o-* indicates that it was very weakly articulated, which suggests that the stress may have been fixed towards the beginning of the word. Likewise, in later Gaulish there was a tendency for final *-s* and *-n* to be dropped. However, French toponyms suggest that stress could be variably placed; there are numerous examples in which two different French toponyms are descended from a single, but variably stressed, Gaulish ancestor, for example, *Nemours* from *Nemáusus*, but *Nîmes* from *Némaus*. Falc’hun (1981:294–313) has suggested that penultimate stress was more archaic and that antepenultimate stress was an innovation which spread from the Mediterranean. The placement of stress in Gaulish has also been discussed recently by De Bernardo Stempel (1994; 1995) and Schrijver (1995:20–21).

## 4. MORPHOLOGY

### 4.1 Word formation

Like other ancient languages of the Indo-European family, the Continental Celtic languages are fusional. Words are composed of a basic morpheme to which derivational prefixes and suffixes may be affixed. There is some evidence that multiple prefixation, as is common in the Insular Celtic languages, was productive. A stem-vowel could be added to the end of this complex, after which the inflectional ending, if any, was attached.

### 4.2 Nominal morphology

Nominals, which include nouns, adjectives, and pronouns, are inflected for case, gender, and number. There is evidence for all eight classical Indo-European cases – nominative, accusative, genitive, dative, locative, instrumental, ablative, and vocative – but not in all numbers and declensions, and not in all languages. The familiar three genders – masculine, feminine, and neuter – of the Indo-European family are well documented, as are the singular and plural numbers. There is some slight evidence that the dual also existed.

#### 4.2.1 Nominal stem-classes

##### 4.2.1.1 Hispano-Celtic

The nominal inflection of Hispano-Celtic as presently attested is given in Table 35.3. Uncertain identifications are followed by a question mark.

The *o*-stem genitive singular in *-o* is an innovation via a proportional analogy with the pronominal paradigm. Compare the Proto-Celtic *ā*-stem genitive singular syntagm \**sosyās bnās* ‘this woman’ with *o*-stem \**sosyo wirī* ‘this man’. In order to extrapolate the nominal

**Table 35.3 Hispano-Celtic nominal inflection**

	<i>ā</i> -stem	<i>o</i> -stem	<i>i</i> -stem	<i>u</i> -stem	<i>n</i> -stem <sup>1</sup>	<i>n</i> -stem <sup>2</sup>	<i>r</i> -stem	<i>nt</i> -stem	<i>C</i> -stem
Singular									
<i>Nom.</i>	-a	-oś	-iś		-u	-i			-ś
<i>Acc.</i>	-am	-om	-im					-nTam	
<i>Nom.-acc. neut.</i>		-om							
<i>Gen.</i>	-aś	-o	-eś?		-unoś	-inoś	-efoś	-nToś	-oś
<i>Dat.</i>	-ai	-ui	-e/-ei?	-uei	-unei	-inei			-ei
<i>Loc.</i>		-ei							
<i>Instr.</i>		-u?			-unu?				
<i>Abl.</i>	-as	-us	-is	-ues?	-unes				-es
Plural									
<i>Nom.</i>	-aś?	-oi?	-iś				-efeś		-eś
<i>Acc.</i>	-aś	-uś?		-uś?					
<i>Nom.-acc. neut.</i>		-a							
<i>Gen.</i>	-aum	-um							
<i>Dat.</i>		-o/uPoś							

**Table 35.4** Lepontic nominal inflection

	<i>ā</i> -stem	<i>o</i> -stem	<i>i</i> -stem	<i>n</i> -stem	C-stem
Singular					
<i>Nom.</i>	-a	-os	-is	-u	
<i>Acc.</i>	-am	-om			
<i>Nom.-acc. neut.</i>		-om			
<i>Gen.</i>		-oiso, -i			
<i>Dat.</i>	-ai	-ui	-ei?	-onei/-oni	
Plural					
<i>Nom.</i>		-oi		-ones	
<i>Acc.</i>					-eś
<i>Dat.</i>		-oPos		-onePos	

genitive singular in *-o* one need only notice that in the *ā*-stem inflection the pronominal and nominal endings are identical after the *-y-* in the demonstrative (see Prosdocimi 1991: 158–159; Eska 1995:41–42). The identification of *o*- and *n*-stem forms in *-u* as instrumental singulars has been proposed by Villar (1993–1995). The *o*-stem nominative plural in *-oi* is perhaps attested once (or twice) in a single inscription. A single accusative plural form in *-uś* could be either an *o*- or *u*-stem. In the animate *n*<sup>1</sup>-stems, the lengthened-grade suffix *\*-ō(n)-*, proper only to the nominative singular, has been extended throughout the paradigm.

#### 4.2.1.2 Lepontic

The nominal inflection of Lepontic as attested is given in Table 35.4. Uncertain identifications are followed by a question mark.

The *o*-stem genitive singular in *-oiso* is attested only in very early forms. It appears to continue Indo-European pronominal *\*-osyo*; Colonna (in Gambari and Colonna 1986:138) and Lejeune (1989:64) treat the Lepontic ending, which is also attested once in Venetic (see Ch. 34, §4.1.1 [7]) (but see now Eska and Wallace 1999), as a metathesized variant. Eska (1995:42) suggests that it is the result of a crossing with the Lepontic descendant of the Proto-Indo-European pronominal genitive plural *\*-oisōm* (cf. *Hisp.-Celt. soisum*). De Hoz (1990) suggests that, in addition to earlier *-oiso* and later *-ī*, Lepontic also had an *o*-stem genitive singular in *-ū* from ablative singular *\*-ōd*. These forms have traditionally been interpreted as animate *n*-stem nominative singulars (see Eska 1995, especially pp. 34–37 for a critique of de Hoz's proposal). Attested once, the *n*-stem dative singular *-oni* seems to represent an early instance of the locative in dative function (see now Eska and Wallace 2001). The consonant-stem accusative plural ending *-eś* (attested once) presumably has been remade by analogy with the vocalism of the nominative plural ending, since inherited *\*-ns* would have yielded Proto-Celtic *\*-ans* > *\*-ās*. The spelling of the sibilant with <ś> perhaps indicates that an epenthetic *\*-t-* was inserted into the inherited *\*-ns* group (perhaps *\*-ens* > *\*-ents* > <-és> = /-ēts/), as is attested elsewhere in the accusative plural ending of Luwian (so also in *Cis. Gaul. acc. pl. arTuasś*).

#### 4.2.1.3 Gaulish

The nominal inflection of Gaulish as attested is given in Table 35.5. Multiple exponents of a single ending are given in chronological order of attestation. The inflectional morphemes

**Table 35.5** Gaulish nominal inflection

	<i>ā</i> -stem	<i>o</i> -stem	<i>i</i> -stem	<i>u</i> -stem	<i>n</i> -stem	<i>r</i> -stem	C-stem
Singular							
<i>Nom.</i>	-a	-os, -o	-is	-us	-u	-ir	-s
<i>Acc.</i>	-an, -em -en, -im	-om, -on	-in			-erem	
<i>Nom.-acc. neut.</i>		-on	-e?	-u	-an		
<i>Gen.</i>	-as, -ias	-i	-ios?				-os
<i>Dat.</i>	-ai, -i	-ui, -u	-e	-ou			-i
<i>Loc.</i>		-e					
<i>Instr.</i>	-ia	-u					
<i>Voc.</i>	-a	-e					
Dual							
<i>Nom.-acc.</i>		-o					
Plural							
<i>Nom.</i>	-as	-oi, -i	-is	-oues			-es
<i>Acc.</i>	-as	-os, -us					
<i>Nom.-acc. neut.</i>		-a					
<i>Gen.</i>	-anom	-on	-iom			-ron	
<i>Dat.</i>	-abo, -abi?	-obo, -obe?				-rebo, -rebe?	-bi
<i>Instr.</i>	-abi?	-obe?				-rebe?	

attested only in north Etruscan or Greek characters are here transcribed into Roman characters. Uncertain identifications are followed by a question mark.

The *ā*-stem inflection in later Gaulish has been deeply affected by the inherited *ī*-stem inflection. Accusative singular forms with *e*-vocalism in the *ā*- and *r*-stems appear to be the result of the raising of /a/ before the final nasal, as is also indicated for Old Irish. The final *-m* of *ā*-stem accusative singular *-em* is usually taken to be archaic. The *ā*-stem dative singular in *-ī* is the result of contraction of *-ai* < \**-āi*. The *ā*-stem genitive plural in *-anom* is attested in only one inscription and could, therefore, represent a local innovation. Owing to the difficulty of interpreting the documents, it is unclear whether *ā*-, *o*-, and *r*-stem forms in *-bi*, *-be* are dative or instrumental plural. The *o*-stem dative singular in *-ū* could represent either the apocope of *-i* from earlier *-ūi* or syncretism of the dative, instrumental (and ablative?) singular. The neuter nominative-accusative singular *n*-stem in *-an* regularly continues \**-m*. The consonant-stem dative singular in *-i* continues the inherited locative singular ending.

#### 4.2.2 Pronouns

Partial paradigms of a variety of pronominals are attested in Continental Celtic. The demonstrative stem \**so/ā*- is attested in Hispano-Celtic and Gaulish, with the initial \**s*-, originally only in the masculine and feminine nominative singular, extended throughout the paradigm. It seems to have been fully stressed in Hispano-Celtic; it is unclear whether it ever was stressed in Gaulish. Gaulish also had a reduplicated formation attested in nominative-accusative neuter singular *sosin* and *sosio*. This *-sin* element also seems to be found in several forms



which appear to be ancestors of the Insular Celtic article, namely, *in=sinde*, *indas* (with early loss of initial *s*–; the sign = represents a clitic boundary), and *o=’nda* (contracted in composition with a preposition).

The relative pronominal stem *\*yo-* appears as a stressed and inflected form in Hispano-Celtic. In Gaulish, it has been reduced to an uninflected subordinating clitic particle =*yo*.

The anaphoric pronominal stem *\*ei-* appears to continue its inherited function in two Gaulish forms, namely, *eianom* and *eiabi*. It also can function as a clitic object pronoun. Many scholars believe that the nominative can also be attached as a clitic to a verb for emphasis, for example, neuter singular *buet=id*, though some would segment the sequence otherwise.

Seemingly related to the anaphoric stem is a series of forms which may ultimately be related to the Latin pronoun *iste*. These are *Hisp.-Celt. iſTe* and *ſTam* and *ſTena* (with aphaeresis?), *Lep. iſos*, and *Gaul. ison* and *isoc* (with attached deitic *\*=kē?*).

Hispano-Celtic also has a pronominal stem *o-* attested in the the forms *osiaſ* (fem. gen. sg.?) and *osaſ* (fem. acc. pl.?) which perhaps displays a different ablaut grade of the anaphoric stem.

There are very few personal pronouns attested. The only ones which have been securely identified are the clitic accusatives, Gaulish first singular =*me*, first plural =*snj* and first singular dative =*mi* < *\*moj*. The attested possessive pronouns are first singular *imon* and *mon* and second singular *to*. It also seems probable that the first singular nominative form =*mi* (< acc. *\*mē*) and second singular = *tu* are attested as emphasizing pronouns, though they have been otherwise interpreted (see §4.3.6).

Finally, the deictic stem *\*kēi-* is attested in the Gaulish syntagm *du=ci*, literally “to here,” employed as a connective “and.”

### 4.3 Verbal morphology

In typical Indo-European fashion, the Continental Celtic verb is marked for tense, voice, and mood.

#### 4.3.1 Tense

In the verbal system, there is good evidence for the present, preterite, and future tenses; Meid (1994:392–393) suggests that Hispano-Celtic also continued the Indo-European imperfect, but this is uncertain. The present tense is attested in a number of common Indo-European formations.

The preterite is composed of forms which continue Indo-European perfects, *s*-aorists, and renewed imperfects. There is also at least one example of suppletion (see Schmidt 1986 and Eska 1990). Owing to phonological reductions, the Continental Celtic *s*-preterite has in some cases been augmented with a thematic (i.e., *o*-stem) ending; compare unaugmented *Gaul. 3rd sg. prinas* “he sold” < *\*k<sup>w</sup>ri-n-h<sub>2</sub>-s-t* (which would have been homophonous with the second singular), and augmented *Gaul. 3rd sg. legasit* “he placed” < *\*leg<sup>h</sup>-eh<sub>2</sub>-s-t + \*-et*. The Continental Celtic *t*-preterite is of multiple origin. Like the Insular Celtic *t*-preterite, it continues the Indo-European *s*-aorist in certain forms: for example, *Gaul. 3rd sg. ſoberte* < *\*to-b<sup>h</sup>er-s-t + 3rd sg. perf. -e*, which was affixed to characterize the form as third singular overtly once the *-t* was regrammaticalized as the exponent of tense. A perfect ending was also affixed to inherited imperfect forms in order to recharacterize them as preterites: for example, *Lep. 3rd sg. KariTe* “he placed” < *\*k-r-ye-t + -e* after the apocope of primary *\*-i* (at least after voiceless consonants) caused the present and imperfect to fall together.

The attested Continental Celtic future forms all continue the Indo-European desiderative in  $^{*}(h_1)sye/o-$ : for example, Gaul. 1st sg. *marcosior* (a derivative of *marc* “horse,” of uncertain meaning) and 3rd sg. *bissjet*. A reduplicated formation appears to be attested in Gaul. 1st sg. *siaxsiou* <  $^{*}si-sag-sj\ddot{u}$ . (Pierre-Yves Lambert has proposed that Gaul. *lilous* is a third singular reduplicated perfect, but this is very uncertain.)

#### 4.3.2 Aspect

There is a small amount of evidence for perfective aspect in Hispano-Celtic and Gaulish. In Hispano-Celtic, the perfectivizer *Con-* is prefixed in the verbal adjective *ConśCiliTom* “cut up.” It is attested in the Gaul. 3rd sg. perf. *To = śo = KoTe* “he offered it.” The prefix *ek-* likewise is a perfectivizer in the Gaul. 3rd sg. perf. *to = me = declai* “she set me up” <  $^{*}de-ek-$ . The prefix *ro-*, the most common perfectivizer in Insular Celtic, may occur in reduced form in Gaul. 3rd sg. perf. *readdas* “he dedicated (it).” These prefixes are all attested in this function in Insular Celtic.

#### 4.3.3 Voice

The large majority of verbs presently attested in the Continental Celtic corpus, if not all, are active in voice. There are not a few forms which terminate in *-r*, but the majority of these are deponents, and hence active in voice. Two forms which have been claimed to be passive in voice are Gaul. 3rd pl. *diligentir* and Hisp.-Celt. 3rd pl. *PinToř*.

#### 4.3.4 Mood

There is good evidence for the indicative, subjunctive, and imperative moods in Continental Celtic. Subjunctives are characterized by the suffixes *-se/o-* (e.g., Hisp.-Celt. 3rd sg. pres. *CaPiseTi*) or *-ā-* (e.g., Gaul. 2nd sg. pres. *lubijas* “you enjoy/love”). The subjunctive mood can also be characterized by the thematic vowel (e.g., Gaul. 3rd sg. *buet/bwet* “he may be”).

Imperatives are attested in both the simple and so-called future type. The former, which are certainly attested in only the second singular, take the form of the bare present stem, for example, Gaul. *gabi* “take!” <  $^{*}g^h ab^h-ye-\emptyset$  or continue the imperative in  $^{*}-si$ , for example, *jexs*. The latter, which was characterized by the affixation of  $^{*}-tōd$  to the simple imperative in Proto-Indo-European, appears in Hispano-Celtic in the third singular with the ending *-Tus*. This ending has also been claimed to underlie Gaulish third singular and plural forms in *-tutu* and *-ntutu*, respectively, with an iterated ending as attested in Umbrian.

Lambert (1994:63) suggests that the optative mood is also attested in Gaulish, indicated by the exponent *-si-* in the form *ni = tixsintor*.

#### 4.3.5 Verbal stem-classes

Owing to the fragmentary nature of the corpus, there is insufficient material available to try to attempt to reconstruct the verbal conjugations of Continental Celtic.

#### 4.3.6 Verb endings

The endings of the verb are also far from complete. Those attested for the present tense are given in Table 35.6, those of the preterite in Table 35.7.

In the Gaulish first singular, both thematic *-u* and athematic *-mi* are attested. Some first-person verbs terminate with the sequence *-umi*, which some have taken to represent a fusion of the two endings (cf. Sanskrit *-āmī*). It is also possible, however, that the segment *-mi* in

**Table 35.6 Present endings of Continental Celtic**

		Hispano-Celtic	Gaulish
<i>Singular</i>	1.		-u, -mi
	2.		-s
	3.	-Ti, -t	-t
<i>Singular deponent</i>	1.		-or
	3.		-toi?
<i>Plural</i>	1.	-mus, -mu	
	2.		-tes, -tis
	3.	-nTi	-nt
<i>Plural deponent</i>	3.		-ntor
<i>Plural passive</i>	3.	-nToī?	-ntir?

**Table 35.7 Preterite endings of Continental Celtic**

		Hispano-Celtic	Lepontic	Gaulish
<i>Singular</i>	3.	-es?	-e, -u	-s, -t, -e, -u, -ai?, -i?
				-us
<i>Plural</i>	3.			
<i>Plural deponent</i>	3.	-nTo?		

such endings is a clitic emphasizing pronoun. This seems likely since first singular verbs can terminate with both *-u* and *-umi* in the same text.

Villar (1995a:31–33 = 1995b:17–19) has proposed that some Hispano-Celtic forms in *-es* may continue third singular perfects to which secondary *\*-t* was affixed, as in Old Latin, which then was voiced to *\*-d* and subsequently developed into the phone(me) represented by *-<s>*. The Gaulish ending *-ai*, later contracted to *-ī*, apparently is third singular to judge by context. Forms in *-us* have traditionally been interpreted as third plural, made by the affixation of a pluralizing *-s* to third singular *-u*, but this has recently been challenged by de Hoz (1995).

#### 4.3.7 Nonfinite verbals

Like the Insular Celtic languages, Continental Celtic did not have true infinitives, but employed nominalized verbs. There are three attested in Hispano-Celtic, all formed with the exponent *-un-* and inflected for the dative case (it is not clear whether this suffix continues *\*-w(e)r/n-* or *\*-mn-*).

There is a variety of participial forms attested. A single Gaulish inscription has four examples of the present active participle in *\*-nt-*, all of which terminate in *-ontias* (*ā*-stem gen. sg. or nom. or acc. pl.). The same inscription contains a single example of a form in *-mno-* which has been interpreted as a mediopassive participle (though this would require syncope in *\*-mano-* < *\*-mh<sub>1</sub>no-*). More widely attested is the passive participle in *\*-to/ā-*. It is attested in Hispano-Celtic as a verbal adjective and often in Gaulish in anthroponyms.

## 4.4 Derivational morphology

The principal method of derivation in all of the Continental Celtic languages is affixation. Prefixation is as common as suffixation. Compounding is also very frequent, especially in the formation of anthroponyms (Schmidt 1957; Evans 1967). There is one particularly interesting example of a *dvandva* compound (see Ch. 26, §4.4.2.1) in Gaulish, genitive plural *TeuoχTonion* “of gods and men” < \**deiwo-* + \**d<sup>h</sup>g<sup>h</sup>onyo-*.

## 4.5 Numerals

There is a little evidence for numerals in Continental Celtic. In Hispano-Celtic, the attested cardinals are *Tifis* “three” (masc. acc.), *šueš* “six,” and *CanTom* “one hundred.” A single ordinal, “tenth,” is attested as acc. sg. *TeCameTam*. For some forms which arise in onomastics, see Tovar (1954).

There are only two cardinals attested in Gaulish: *tidres* “three” (fem. acc.) and possibly *trjcontis* “thirty.” Compositional forms include *cintu-* “first,” *tri-* “third,” *petru-* “four” and *pompe-* “five.” We are fortunate that a nearly complete set of ordinals for 1–10 have been preserved; these are listed in (8):

(8)	1 <sup>st</sup>	cintuxo[s]	6 <sup>th</sup>	suexos
	2 <sup>nd</sup>	allos	7 <sup>th</sup>	sextametos
	3 <sup>rd</sup>	tr[itios]	8 <sup>th</sup>	oxtumetos
	4 <sup>th</sup>	petuar[ios]	9 <sup>th</sup>	namet[os]
	5 <sup>th</sup>	pinpetos	10 <sup>th</sup>	decametos

A further ordinal, Latinized dative-ablative singular *petrudecameto* “fourteenth,” indicates that the tens were formed by compounding. One final form, probably the fraction “one-third,” which appears to be calqued on Latin acc. pl. *trientēs*, is *trianis*.

# 5. SYNTAX

Owing to the fragmentary nature of the Continental Celtic corpus, we have a much less complete picture of syntax than of phonology or morphology. This section, then, does no more than present a selection of the principal constructions that are attested. We are, however, in the fortunate position, owing to the varying degrees of conservatism of the individual languages, of being able to observe the evolution of Celtic clausal configuration in fieri (see Eska 1994b). The languages are addressed in order of increasing innovation. In the examples, translations are provided only when fairly secure.

## 5.1 Hispano-Celtic

### 5.1.1 Word order

Hispano-Celtic is an SOV (Subject–Object–Verb) language, exhibiting “*pro-drop*” (i.e., the subject can be expressed merely by verb inflection), as is usually reconstructed for Proto-Indo-European:

(9)	uTa	ošCues	šTena	ueřsoniTī
	CONN.	PRO.NOM.SG.	NP.ACC.PL.	V.3RD SG.
	“whoever carries out these things”			

It is noteworthy that imperative verbs are also clause-final, for example:

- (10) TeCameTam TaTus  
 NP.ACC.SG. V.3RD SG. IMPV.  
 “let him offer a tithe”

Verbs are not bound to clause-final position, however; they may be raised to clause-initial position for various pragmatic purposes, for example:

- (11) iom ašeCaTi amPiTinCounei šTena  
 CONN. V.3RD SG. NP.DAT.SG. NP.ACC.PL.

It is possible for a non-core argument to appear to the right of the verb. In the following example, the noun phrase (NP) to the right of the verb is in a disjunctive relationship (see §5.1.2) with a core argument:

- (12) ioś uřanTiom = ue auseTi ařaTim = ue  
 PRO.NOM.SG. NP.ACC.SG. = DISJ. V.3RD SG. NP.ACC.SG. = DISJ.

Though it is an SOV language, Hispano-Celtic is not rigorously head-final. While attributive genitives do precede their head nouns, for example,

- (13) A. ologas togias  
 NP.GEN.SG. NP.ACC.PL.  
 B. tiaso togias  
 NP.GEN.SG. NP.ACC.PL.

subordinate clauses usually follow matrix clauses (see [17]), and adjectives follow their head nouns, for example,

- (14) A. TiriCanTam PeřCuneTaCam  
 N.ACC.SG. ADJ.ACC.SG.  
 B. šleiTom ConśCiliTom  
 N.ACC.SG. ADJ.ACC.SG.

In prepositional phrases, both prepositions and postpositions are attested. Individual pre- and postpositions are consistent in their placement:

- (15) A. eś ueřTai  
 from NP.DAT.SG.  
 B. TiriCanTam eni  
 NP.ACC.SG. in/at

### 5.1.2 Clitics

The corpus does not provide any examples of pronominal clitics. The only clitics attested to date are the connective =Cue, =*que* < \*=*k<sup>w</sup>e* and the disjunctive =ue < \*=*we*. In the earlier language, they are attached to each member of a serial correlation, as in (16A), but in the later language are attached only to the final member, as in (16B):

- (16) A. PouśTom = ue Córuiinom = ue maCási[.]m = ue ailam = ue  
 NP.ACC.SG. = DISJ. NP.ACC.SG. = DISJ. NP.ACC.SG. = DISJ. NP.ACC.SG. = DISJ.  
 B. eniorosei equeisui = que  
 NP.DAT.SG. NP.DAT.SG. = CONN.

### 5.1.3 Coordination

The attested corpus exhibits a variety of connectives with which clauses can be coordinated, *uTa* (cf. Sanskrit *utá*), *to* (cf. Old Hittite *ta*), and *iom*. Asyndeton is also common.

### 5.1.4 Subordination

Subordinate clauses generally, but not always, follow main clauses. In the following example, the subject of the subordinate clause (17B) is a stressed relative pronoun which agrees with the NP in the main clause (17A) to which it is bound:

- |      |    |                     |                  |                      |                   |                 |
|------|----|---------------------|------------------|----------------------|-------------------|-----------------|
| (17) | A. | <i>iom</i>          | <i>CuśTaiCoś</i> | <i>aśsnaś</i>        | <i>CuaTi</i>      |                 |
|      |    | CONN.               | NP.NOM.SG.       | NP.FEM.ACC.PL.       | V,3RD SG.         |                 |
|      | B. | <i>iaś</i>          | <i>osiaś</i>     | <i>ueŕTaToś = ue</i> | <i>Temei = ue</i> | <i>ŕoPiśeTi</i> |
|      |    | REL.PRO.FEM.ACC.PL. | PRO.GEN.SG?      | ADV. = DISJ.         | ADV. = DISJ.      | V,3RD SG        |

The attested corpus also contains an interesting example of the Proto-Indo-European correlative construction (cf. Sanskrit *yá-... sá-...*):

- |      |    |                 |               |               |                |
|------|----|-----------------|---------------|---------------|----------------|
| (18) | A. | <i>iomui</i>    | <i>liśTaś</i> | <i>TiTaś</i>  | <i>sisonTi</i> |
|      |    | REL.PRO.DAT.SG. | NP            | NP            | V,3RD PL       |
|      | B. | <i>śomui</i>    | <i>iom</i>    | <i>aśsnaś</i> | <i>PionTi</i>  |
|      |    | DEM.PRO.DAT.SG. | CONN.         | NP            | V,3RD PL       |

### 5.1.5 Agreement

Presently, all evidence points to subject–verb agreement for person and number, and noun–adjective agreement for case, number, and gender.

## 5.2 Lepontic

### 5.2.1 Word order

The Lepontic corpus presently contains only three verbal sequences. One of them is archetypally SOV in structure:

- |      |                                 |                                |              |             |
|------|---------------------------------|--------------------------------|--------------|-------------|
| (19) | <i>uvamoKozis Plialeθu</i>      | <i>uvl TiauiPos ariuonePos</i> | <i>siTeś</i> | <i>TeTu</i> |
|      | NP.NOM.SG.                      | NP.DAT.PL.                     | NP.ACC.PL.   | V,3RD SG    |
|      | “U. B. offered s. to the U. A.” |                                |              |             |

The underlying configuration of the remaining two verbal sequences, which both occur in the same inscription, is unclear owing to movement:

- |      |    |  |               |              |               |
|------|----|--|---------------|--------------|---------------|
| (20) | A. | <i>PelKui</i>                              | <i>Pruiam</i> | <i>Teu</i>   | <i>KariTe</i> |
|      |    | NP.DAT.SG.                                 | NP.ACC.SG.    | NP.NOM.SG.   | V,3RD SG.     |
|      |    | “D. set up the b. for B.”                  |               |              |               |
|      | B. | <i>iśos</i>                                | <i>KaliTe</i> | <i>Palam</i> |               |
|      |    | PRO.NOM.SG.                                | V,3RD SG.     | NP.ACC.SG.   |               |
|      |    | “he (likewise) erected the memorial stone” |               |              |               |

It is, of course, necessary to analyze both clauses together. It is unclear whether they are SOV underlyingly, with postposition of the accusative argument in (20B), or SVO, with raising of both the dative and accusative arguments in (20A). What can be said with certainty,

however, is that, unlike Hispano-Celtic, a core argument can appear to the right of the verb at the surface, as in (20B).

Lepontic adjectives follow the nouns they determine, for example:

- (21) uinom                      naśom  
       N.NOM-ACC.SG.      ADJ.NOM-ACC.SG.  
       “Naxian wine”

### 5.2.2 Clitics

There are no clitic pronominals attested in the Lepontic corpus. The connective =Pe < \* = k<sup>w</sup>e is attested; it attaches to the final member of a serial correlation, for example:

- (22) laTumarui              saPsuTai = Pe  
       NP.DAT.SG.          NP.DAT.SG. = CONN.  
       “for L. and S.”

### 5.2.3 Agreement

Lepontic shows subject–verb agreement for person and number and noun–adjective agreement for case, number, and gender.

## 5.3 Gaulish

### 5.3.1 Word order

It is difficult to be sure about the underlying configuration of the Gaulish clause owing to the wide diversity of surface configurations attested; verb-initial, verb-medial, and verb-final are all found. Some of this variation could be due to dialectal or chronological differences, and much, no doubt, is the result of movement for pragmatic purposes and syntactic rules (see now Eska, forthcoming). There are only a handful of verb-final clauses attested, and the majority of verb-initial clauses contain imperative verbs. Those which are not imperative, for example,

- (23) regu = c                      cambion  
       V.1ST SG. = CONN?      NP.ACC.SG.  
       “I straighten the bent thing”

cannot be diagnosed as underlyingly verb-initial clauses, however, since they can also be analyzed as SVO clauses with *pro*-drop. It is clear, however, that Gaulish was not a verb-second language, as the following inscription, with two NPs preceding the verb, demonstrates (the bracketed character is superfluous):

- (24) ratin briuationom      frontu tarbetis[o]nios      ie{i}uru  
       NP.ACC.SG.              NP.NOM.SG.                      V.3RD SG.  
       “F. T. dedicated the r. of the b.”

The large majority of Gaulish clauses are verb-internal at the surface, for example:

- (25) martialis dannotali      ieuru              ucuete              sosin celicnon  
       NP.NOM.SG.              V.3RD SG.      NP.DAT.SG.      NP.ACC.SG.  
       “M. D. dedicated this edifice to U.”

A very important feature to take notice of is that, whenever a clitic pronominal object (see §5.3.3) is present in the clause, it must be syntactically hosted (i.e., adjacent) to the verb; this constraint on second-position clitics is known as *Vendryes' Restriction*. Since *Wackernagel's Law* was strongly grammaticalized in Celtic (at least by this time), this had the effect of ensuring that the verb occupied clause-initial position. In such cases, the verb either occupies absolute initial position in the clause, for example,

- (26) *sioxt = i                      albanos                      panna(s)                      extra tud(on)                      ccc*  
       V. = PRO.NEUT.        NP.NOM.SG.        NP.FEM.ACC.PL.        PP                      NUM.  
       NOM.-ACC.PL.  
       “A. added them, vessels beyond the allotment (in the amount of) 300”

or is preceded only by a null-position, semantically empty, sentential connective, the original purpose of which was to host the clitic phonologically (as familiar from Anatolian; see Ch. 18, §5.1), for example,

- (27) *to = me = declai                      obalda natina*  
       CONN. = PRO.1ST SG.ACC.=V.3RD SG.        NP.NOM.SG.  
       “O., (their) dear daughter, set me up”

It is commonly agreed that Vendryes' Restriction had a large role to play in the development of the VSO configuration of the Insular Celtic languages.

As one would expect in a language which is – predominately, at least – not verb-final, other syntactic configurations strongly tend to be head-initial. Genitives follow their head nouns, for example:

- (28) A. *ratin                      briuationom*  
       N.ACC.SG.        N.GEN.PL.  
       “the fort of the b.”  
       B. *aTom̃                      TeuoxTonion*  
       N.ACC.SG.        N.GEN.PL.  
       “the border of gods and men”

Likewise, the unmarked position for adjectives appears to be after their head nouns,

- (29) A. *τοουτιο{υ}ς                      ναμαυσατις*  
       N.NOM.SG.        ADJ.NOM.SG.  
       “citizen of Nîmes”  
       B. *ματρεβο                      ναμαυσικαβο*  
       N.DAT.PL.        ADJ.DAT.PL.  
       “to the Matres of Nîmes”

and PPs are always prepositional:

- (30) A. *in                      alixie*  
       in                      NP.LOC.SG.  
       “in Alisia”  
       B. *extra                      tud(on)*  
       beyond        NP.ACC.SG.  
       “beyond the allotment”

A good example of a passive clause, though verbless, has been identified by Prosdocimi (1989):





### 5.3.4 Agreement

Noun–adjective agreement is marked for case, number, and gender. Subject–verb agreement is normally marked for person and number, but there is a single example in which agreement for number may be lacking:

- (37) eluontiu      ieuru      aneuno oclicno      luguriḡ aneunicno  
 NP.DAT.SG.    V.3RD SG?    NP.NOM.SG.      NP.NOM.SG.  
 “To E., A. O. and L. A. dedicated [this stele]”

In this inscription, a compound subject appears not to agree with an apparently third singular verb. However, it has been noted that final postvocalic *-s* apparently has been lost in the language of this text, the addition of which to the verb would make it third plural. The lack of subject–verb agreement might, therefore, be illusory. It should also be borne in mind that, cross-linguistically, it is not uncommon for a singular verb to be used with conjoined subjects.

## 6. LEXICON

With the exception of onomastic material, there have been remarkably few etyma of foreign origin identified in the Continental Celtic lexicon. These Celtic languages appear to have much more frequently been loaning than borrowing languages. Within the onomastic material of foreign origin, Latin, Iberian, and Greek elements (in descending order of frequency) are found in the Hispano-Celtic speech area. As one would expect, Latin elements are common among the Gauls, especially in the later period, and some Greek influence is also felt (see Meid 1980). Greek elements are not uncommon in the Galatian speech area. A so-called Mediterranean substratum has been alleged to be the source of some borrowings into Gaulish and Lepontic.

The most noteworthy borrowing into Continental Celtic is the Lepontic patronymic suffix *-alo/ā-*, which is otherwise unknown in Celtic. It has been connected to the Raetic or Etruscan genitive singular in *-al* (otherwise Prosdocimi 1991:163–176). One further surprising borrowing is Hispano-Celtic *śilaPuř*, apparently “silver,” which is attested twice beside native *afCaTo-*. The etymon is found elsewhere in Indo-European, in Germanic and Balto-Slavonic, and also in Basque. It has been maintained to be of ultimate Semitic origin.

## 7. READING LIST

The individual corpora of the Continental Celtic languages are in the process of publication. The Hispano-Celtic corpus is to be part of Jürgen Untermann’s *Monumenta Linguarum Hispanicarum*; vol. I (1975) contains the Celtic coin legends, and vol. II (1980) contains one Celtic inscription (B.3.1). The remainder of the Celtic corpus appears in vol. IV (1997). For subsequently published inscriptions, see Jordán Cólera (2001). Wodtko (2000) provides a Hispano-Celtic lexicon. The Lepontic corpus as known in 1970 is treated by Lejeune (1971); Tibiletti Bruno (1981) may also be consulted, but is inferior to Lejeune’s work. The most recent collection, which focuses upon all of Cisalpine Celtic, is Solinas (1995); it concentrates almost exclusively on epigraphic matters. The most recent discussion of the Lepontic corpus is Motta (2000). The Gaulish corpus is published as the *Receuil des Inscriptions Gauloises*; the volumes treat the inscriptions in Greek characters (Lejeune 1985; supplemented by Lejeune 1988–1995), north Etruscan characters and Roman characters on

stone (both in Lejeune 1988), the calendrical inscriptions (Duval and Pinault 1986), the coin legends (Colbert de Beaulieu and Fischer 1998), and the inscriptions on movable objects, which are largely engraved in Roman cursive (Lambert 2002b). In addition, Marichal (1988) has collected the graffiti from La Graufesenque in similar format. Delamarre (2003) provides a useful dictionary. Billy (1993) is useful for locating Gaulish lexical items embedded in non-Celtic texts. The sparse Galatian materials have been treated by Weisgerber (1931) and more recently by Schmidt 1994. A new collection has been prepared by Phillip Freeman (2001). The language of the British coin legends has been discussed by De Bernardo Stempel (1991). Tomlin (1987) prints two possible British defixio texts.

Eska and Evans (1993) discusses the various categories of inscriptions in the Continental Celtic corpus and interesting features of the individual languages, but is somewhat dated due to recent discoveries. Schmidt (1983) also surveys some of the important features of Continental Celtic. Particularly important now for Hispano-Celtic grammar are Villar (1995a; 1995b). Jordán Cólera (1998) provides a general introduction. Lambert (2002a) treats Gaulish grammar and provides an excellent selection of the various categories of inscriptions in the corpus, though usually only his own interpretations.

For an alternative treatment of Continental Celtic phonology to that presented herein, see McCone (1996). Certain pronominal forms are discussed in Schrijver (1997). The features of Continental Celtic clausal configuration are treated by Eska (1994b). Eska (1994a) is an exploratory treatment of Vendryes' Restriction.

I should like to thank Joshua Katz and Peter Schrijver for their substantial comments on a preliminary version of this chapter.

## Bibliography

- Ballester, X. 1993–1995. “Sobre el valor fonético de  $\zeta$  en celtibérico.” *Kalathos* 13–14:319–323.
- Billy, P.-H. 1993. *Thesaurus linguae Gallicae*. Hildesheim: Olms; Zürich: Weidmann.
- Campanile, E. 1983. “Considerazioni sugli alfabeti dei Celti continentali.” *Annali del seminario di studi del mondo classico. Sezione linguistica* 5:63–74.
- Colbert de Beaulieu, J.-P. and B. Fischer. 1998. *Recueil des inscriptions gauloises iv, légendes monétaires*. Paris: CNRS Editions.
- De Bernardo Stempel, P. 1991. “Die Sprache altbritannischer Münzlegenden.” *Zeitschrift für celtische Philologie* 44:36–55.
- . 1994. “Zum gallischen Akzent. Eine sprachinterne Betrachtung.” *Zeitschrift für celtische Philologie* 46:14–35.
- . 1995. “Gaulish accentuation. Results and outlook.” In Eska *et al.* 1995, pp. 16–32.
- . 1996. “Die Stummvokale. Eine Bilanz für das Keltiberische.” In W. Meid and P. Anreiter (eds.), *Die grösseren altkeltischen Sprachdenkmäler*, pp. 212–226. Innsbruck: Institut für Sprachwissenschaft der Universität Innsbruck.
- De Hoz, J. 1983. “Origine ed evoluzione delle scritture ispaniche.” *Annali del seminario di studi del mondo classico. Sezione linguistica* 5:27–62.
- . 1990. “El genitivo céltico de los temas en -o-. El testimonio lepóntico.” In F. Villar (ed.), *Studia Indogermanica et palaeohispanica in honorem A. Tovar et L. Michelena*, pp. 315–329. Vitoria-Gasteiz: Universidad del País Vasco/Salamanca: Universidad de Salamanca.
- . 1995. “Is -s the mark of the plural of the puterite in the Gaulish verb?” In Eska *et al.* 1995, 58–67.
- Delamarre, X. 2003. *Dictionnaire de la langue gauloise*, second edition. Paris: Errance.
- De Marinis, R. C. 1991. “Golasecca culture and its links with the Celts beyond the Alps.” In V. Kruta, O. Hermann, B. Raferty, *et al.* (eds.), *The Celts*, pp. 93–102. London: Thames and Hudson.
- Duval, P.-M. and G. Pinault. 1986. *Recueil des inscriptions gauloises iii, Les calendriers (Coligny, Villards d’Heria)*. Paris: CNRS Editions.
- Eska, J. F. 1990. “The so-called weak or dental preterite in Continental Celtic, Watkins’ Law, and related matters.” *Historische Sprachforschung* 103:81–91.

- . 1994a. "On the crossroads of phonology and syntax. Remarks on the origin of Vendryes' Restriction and related matters." *Studia Celtica* 28:39–62.
- . 1994b. "Rethinking the evolution of Celtic constituent configuration." *Münchener Studien zur Sprachwissenschaft* 55:7–39.
- . 1994c. "More on Gaul. siöxt = i." *Études Celtiques* 30:205–210.
- . 1995. "Observations on the thematic genitive singular in Lepontic and Hispano-Celtic." In Eska *et al.* 1995, pp. 33–46.
- . 1998a. "PIE \*p > ø in Proto-Celtic." *Münchener Studien zur Sprachwissenschaft* 58: 63–80.
- . 1998b. "The linguistic position of Lepontic." *Proceedings of the Berkeley Linguistics Society* 24S:2–11.
- . 1998c. "Tau Gallicum." *Studia Celtica* 32:115–127.
- . 2002a. "Symptoms of nasal effacement in Hispano-Celtic." *Palaeohispanica* 2:141–158.
- . 2002b. "Aspects of nasal phonology in Cisalpine Celtic." In *Studia linguarum* 3. *Memoriae A. A. Korolev dicata*, A. S. Kassian and A. V. Sidel'tsev (eds.), 253–275. Moscow: Languages of Slavonic Culture.
- . Forthcoming. "On basic configuration and movement in the Gaulish clause." In P.-Y. Lambert and G.-J. Pinault (eds.), *Actes du colloque international "Gaulois et celtique continental"*. Paris: EPHE.
- Eska, J. F. and D. Ellis Evans. 1993. "Continental Celtic." In M. J. Ball with J. Fife (eds.), *The Celtic Languages*, pp. 26–63. London: Routledge.
- Eska, J. F., R. Geraint Gruffydd, and N. Jacobs (eds.). 1995. *Hispano-Gallo-Brittonica. Essays in Honour of Professor D. Ellis Evans on the Occasion of his Sixty-Fifth Birthday*. Cardiff: University of Wales Press.
- Eska, J. F. and R. E. Wallace. 1999. "The linguistic milieu of \*Oderzo 7." *Historische Sprachforschung* 112: 122–136.
- Evans, D. Ellis. 1967. *Gaulish Personal Names. A Study of Some Continental Celtic Formations*. Oxford: Clarendon Press.
- Falc'hun, F. 1981. *Perspectives nouvelles sur l'histoire de la langue bretonne*. Paris: Union Générale d'Éditions.
- Freeman, P. 2001. *The Galatian language*. Lewiston: Edwin Mellen.
- Gambari, F. G., and G. Colonna. 1986. "Il bicchiere con iscrizione arcaica da Castelletto Ticino e l'adozione della scrittura nell'Italia nord-occidentale." *Studi etruschi* 54:119–164.
- Gómez-Moreno, M. 1922. "De epigrafía ibérica. El plomo de Alcoy." *Revista de filología Española* 9:341–366.
- Gray, L. H. 1944. "Mutation in Gaulish." *Language* 20:223–230.
- Jordán Cólera, C. 1998. *Introducción al celtibérico*. Zaragoza: Universidad de Zaragoza.
- . 2001. "Chronica Epigraphica Celtiberica I." *Novedades en epigraphía celtibérica. Palaeohispanica* 1:369–391.
- Lambert, P.-Y. 2002a. *La langue gauloise. Description linguistique, commentaire d'inscriptions choisies*, second edition. Paris: Errance.
- . 2002b. *Recueil des inscriptions gauloises ii/2, Textes gallo-latins sur instrumentum*. Paris: CNRS Editions.
- Lejeune, M. 1971. *Lepontica*. Paris: Société d'Éditions "Les Belles Lettres."
- . 1985. *Recueil des inscriptions gauloises i, Textes gallo-grecs*. Paris: CNRS Editions.
- . 1988. *Recueil des inscriptions gauloises ii/1, Textes gallo-étrusques, textes gallo-latins sur pierre*. Paris: CNRS Editions.
- . 1988–1995. "Compléments gallo-grecs." *Études celtiques* 25:79–106, 27:175–177, 30:181–189, 31:99–113.
- . 1989. "Notes de linguistique italique. xxxix. Génitifs en -osio et génitifs en -i." *Revue des études latines* 67:63–77.
- . 1993. "D'Alcoy à Espanca. Réflexions sur les écritures paléohispaniques." In Michel Lejeune. *Notice biographique et bibliographique*, pp. 53–86. Leuven: Centre International de Dialectologie Général.
- Marichal, R. 1988. *Les graffites de la Graufesenque*. Paris: CNRS Editions.
- McCone, K. 1996. *Towards a Relative Chronology of Ancient and Medieval Celtic Sound Change*. Maynooth: Dept. of Old Irish, National University of Ireland.

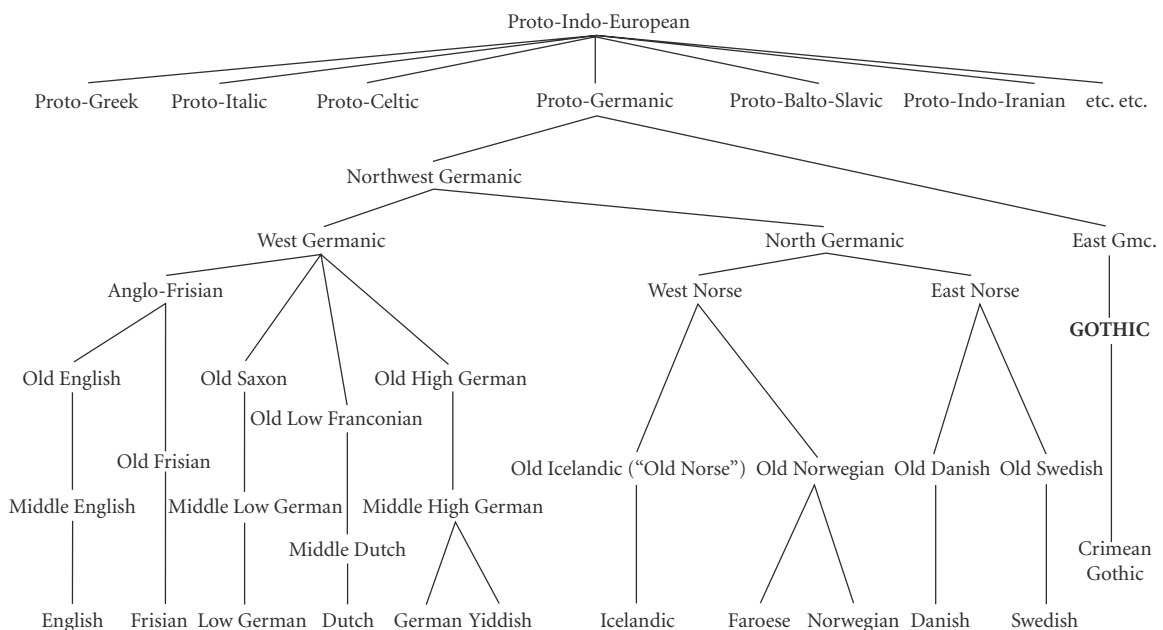
- Meid, W. 1980. *Gallisch oder Lateinisch? Soziolinguistische und andere Bemerkungen zu populären gallo-lateinischen Inschriften*. Innsbruck: Institut für Sprachwissenschaft der Universität Innsbruck.
- . 1994. "Die 'grosse' Felsinschrift von Peñalba de Villastar." In R. Bielmeier and R. Stempel with R. Lanzweert (eds.), *Indogermanica et Caucasica. Festschrift für Karl Horst Schmidt zum 65. Geburtstag*, pp. 385–394. Berlin: Walter de Gruyter.
- Motta, F. 2000. "La documentazione epigrafica e linguistica." In R. C. de Marinis and S. Biaggio Simona (eds.), *I leonti tra mito e realtà*, 2, pp. 181–222. Locarno: Armando Dadò.
- Prosdocimi, A. L. 1989. "Gaulish  $\sigma\upsilon\nu\epsilon\mu\epsilon\tau\omicron\varsigma$  and  $\sigma\omicron\sigma\iota\nu\ \nu\epsilon\mu\eta\tau\omicron\nu$ . À propos of RIG i 154." *Zeitschrift für celtische Philologie* 43:199–206.
- . 1991. "Note sul celtico in Italia." *Studi etruschi* 57:139–177.
- Schmidt, K. H. 1957. "Die Komposition in gallischen Personennamen." *Zeitschrift für celtische Philologie* 26:33–301.
- . 1983. "Grundlagen einer festlandkeltischen Grammatik." In E. Vineis (ed.), *Le lingue indoeuropee di frammentaria attestazione. Die Indogermanischen Restsprachen*, pp. 65–90. Pisa: Giardini.
- . 1986. "Zur Rekonstruktion des Keltischen. Festlandkeltisches und inselkeltisches Verbum." *Zeitschrift für celtische Philologie* 41:159–179.
- . 1994. "Galatische Sprachreste." In E. Schwertheim (ed.), *Forschungen in Galatien*, pp. 15–28. Bonn: Dr. Rudolf Habelt.
- Schrijver, P. 1995. *Studies in British Celtic Historical Phonology*. Amsterdam: Rodopi.
- . 1997. *Studies in the History of Celtic Pronouns and Particles*. Maynooth: Dept. of Old Irish, National University of Ireland.
- Solinas, P. 1995. "Il Celtico in Italia." *Studi etruschi* 60:311–408.
- Tibiletti Bruno, M. G. 1981. "Le iscrizioni celtiche d'Italia." In E. Campanile (ed.), *I Celti d'Italia*, pp. 157–207. Pisa: Giardini.
- Tomlin, R. S. O. 1987. "Was ancient British Celtic ever a written language? Two texts from Roman Bath." *Bulletin of the Board of Celtic Studies* 34:18–25.
- Tovar, A. 1954. "Numerales indoeuropeos en Hispania." *Zephyrus* 5:17–22.
- . 1975. "Les écritures de l'ancienne Hispania." In *Le déchiffrement des écritures et des langues*, pp. 15–23. Paris: L'Asiathèque.
- Uhlich, J. 1999. "Zur sprachlichen Einordnung des Lepontischen." In S. Zimmer, R. Ködderitzsch, and A. Wigge (eds.), *Akten des zweiten deutschen Keltologen-Symposiums*, pp. 277–304. Tübingen: Max Niemeyer.
- . Forthcoming. "On the linguistic classification of Lepontic." In G.-J. Pinault and P.-Y. Lambert (eds.), *Actes du colloque international "gaulois et celtique Continental"*. Paris: Ecole Pratique des Hautes Etudes.
- Untermann, J. 1975. *Monumenta Linguarum Hispanicarum i, Die Münzlegenden*. Wiesbaden: Dr. Ludwig Reichert.
- . 1980. *Monumenta Linguarum Hispanicarum ii, Die Inschriften iberischer Schrift aus Südfrankreich*. Wiesbaden: Dr. Ludwig Reichert.
- . 1997. *Monumenta Linguarum Hispanicarum iv, Die tartessischen, keltiberischen und lusitanischen Inschriften*. Wiesbaden: Dr. Ludwig Reichert.
- Villar, F. 1993–1995. "El instrumental en celtibérico." *Kalathos* 13–14:325–338.
- . 1995a. *Estudios de Celtibérico y de Toponimia Prerromana*. Salamanca: Universidad de Salamanca.
- . 1995b. *A New Interpretation of Celtiberian Grammar*. Innsbruck: Institut für Sprachwissenschaft der Universität.
- Weisgerber, L. 1931. "Galatische Sprachreste." In R. Helm (ed.), *Natalicium Johannes Geffcken zum 70. Geburtstag 2. Mai 1931 gewidmet von Freunden, Kollegen und Schülern*, pp. 151–175. Heidelberg: Carl Winter.
- Wodtko, D. S. 2000. *Monumenta Linguarum Hispanicarum v, Wörterbuch der keltiberischen Inschriften*. Wiesbaden: Dr. Ludwig Reichert.

# Gothic

JAY H. JASANOFF

## 1. HISTORICAL AND CULTURAL CONTEXTS

Gothic, mainly known from a Bible translation of the fourth century AD, is the only Germanic language that has come down to us from antiquity in a reasonably complete state of preservation. Lacking direct descendants itself, it is closely related to the early medieval dialects ancestral to Modern English, German, Dutch, and the Scandinavian languages (Danish, Swedish, Norwegian, Icelandic, Faroese). The family tree of the Germanic languages can be drawn as follows:



**Figure 36.1** The Germanic languages

As can be seen from this figure, Gothic is the sole representative of the East Germanic branch of the family. The more numerous North and West Germanic languages are much later: Old English and Old High German are first substantially attested in the eighth century, while Old Saxon and Old Low Franconian date from the ninth and tenth centuries,

respectively. The remaining “Old” Germanic languages – Old Frisian and the early Scandinavian dialects – are essentially languages of the High Middle Ages, contemporary with Middle English and Middle High German. It is thus not surprising that Gothic presents a significantly more conservative appearance than its Germanic sister dialects. The only comparably archaic remains of an early Germanic language are the Early Northwest Germanic inscriptions of the third, fourth, and fifth centuries, mostly from Denmark and written in the indigenous runic alphabet (see Ch. 37). These, however, are only tantalizing fragments, often deliberately obscure and topheavy with personal names.

Like other East Germanic tribes such as the Vandals, Burgundians, Gepids, and Heruls, the Goths originally lived in the area of present-day Poland and eastern Germany; their own traditions placed their earliest home in southern Sweden. Moving toward the mouth of the Danube and the Black Sea shortly before 200 AD, they first began to make serious raids into Roman territory in the middle of the third century. A hundred years later they had expanded significantly eastwards and split into two sub-peoples: the *Ostrogoths* (“East Goths”), located beyond the Dniester, who controlled most of the modern eastern Ukraine; and the *Visigoths* (meaning unclear; *not* “West Goths”), who remained centered in the southwest of the Ukraine and adjacent parts of Moldova and Rumania. It was in the latter area, toward the middle of the fourth century, that the Arian Christian Wulfila (Ulphilas, Ulphilas) began his ultimately successful effort to convert the Goths to Christianity. Wulfila (Gothic for “Little Wolf”) was himself a native speaker of Gothic, and like many missionaries then and now, recognized the value of translating the Christian scriptures into the language of his intended converts. For this purpose he devised a Greek-based alphabet which remained in use for as long as Gothic continued to be written (see §2). The surviving remains of Wulfila’s translation, amounting to somewhat less than half of the New Testament, constitute the great bulk of the Gothic corpus that has come down to us. Although the Christian Gothic community over which Wulfila presided as bishop was still small at the time of his death (*c.* 382), he laid the groundwork for future missionary work so effectively that Arian Christianity soon became something like a national religion among the Germanic tribes of eastern and central Europe. Yet, interestingly, the Bible seems never to have been translated into Vandal, or Burgundian, or Herulian; evidently these East Germanic languages were close enough to Gothic to make such endeavors unnecessary.

The career of the Goths in the upheavals that accompanied the end of the Western Roman Empire was short but spectacular. The Visigoths, after sacking Rome in 410, established themselves in southern Gaul and subsequently in Spain; here their kingdom lasted until the Moorish conquest of 711, although all our documents from Visigothic Spain are in Latin. The Ostrogoths, in the meantime, established a short-lived kingdom in Italy under their great ruler Theodoric (492–526). Unlike their Spanish cousins, the “Italian” Goths appear to have cultivated their fledgling literary tradition during their half-century of independence. It is to sixth-century Italy, and not to Spain, that we owe our surviving manuscripts of the Gothic Bible, including the famous 188-page Codex Argenteus now housed in Uppsala, Sweden. Also of Italian origin are the few surviving non-Biblical Gothic monuments, which include a fragmentary commentary on the Gospel of John (the so-called *Skeireins* or “explanation”), a calendar, and two very short legal documents. Following the Byzantine reconquest of Italy in 552, the Ostrogoths – and with them the Gothic language – disappear from history.

Or nearly disappear. By chance, a ninth- or tenth-century parchment (the Salzburg–Vienna Alcuin Ms.) has come down to us containing two incomplete versions of the Gothic alphabet and a few verses from the Gothic Bible, the latter accompanied by a mixed transcription/ translation into Old High German. A curious feature of this document is that the Gothic letters bear names, which closely resemble the names of the corresponding runes in Old English and Old Norse. We can only guess at the specific circumstances under which



this information came to be recorded, but one thing seems certain: the descendants of the Ostrogoths who withdrew over the Alps in the middle of the sixth century somehow managed to retain a shadow of their linguistic and religious identity, albeit tenuously, for a period of three or four hundred years.

Another Gothic “survival” turns up much later in a very different corner of Europe. In the middle of the sixteenth century AD, Ogier van Busbecq, the ambassador of the emperor Charles V to the court of the Turkish sultan Suleiman the Magnificent, recorded eighty-six words of a language spoken in the sultan’s Crimean dominions that reminded him of his native Flemish. Most of the lexical items written down by Busbecq are, in fact, obviously Germanic, and one, *ada* “egg,” appears to show the distinctively East Germanic sound change of *\*-jj-* to *-ddj-* (see §3.6.4). It is usually held, therefore, that the Crimean Goths were the last remnants of the Gothic population that once occupied the northern shore of the Black Sea, and that their language was a direct descendant of the Gothic of the fourth century. Unfortunately, by the time anyone thought to extend Busbecq’s vocabulary, Crimean Gothic had disappeared.

## 2. WRITING SYSTEMS

Apart from Busbecq’s word list and two or three problematic runic inscriptions, the entire surviving Gothic corpus is written in Wulfila’s alphabet. Table 36.1 shows the letters as they appear in our most important Gothic manuscript, the Codex Argenteus:

**Table 36.1 Wulfila’s alphabet**

	Transcription	Numerical value	Name
Ɑ	a	1	aza
Ɱ	b	2	bercna
Ɐ	g	3	geuua
Ɒ	d	4	daaz
ⱱ	e	5	eyz
Ⱳ	q	6	quertra
ⱳ	z	7	ezec
ⱴ	h	8	haal
Ⱶ	p	9	thyth
ⱶ	i, ī	10	iiz
ⱷ	k	20	chozma
ⱸ	l	30	laaz
ⱹ	m	40	manna
ⱺ	n	50	noicz
ⱻ	j	60	gaar
ⱼ	u	70	uraz
ⱽ	p	80	pertra
Ȿ	—	90	—
Ɀ	r	100	reda
ⳉ	s	200	sugil
Ⳋ	t	300	tyz
ⳋ	w	400	uuinne
Ⳍ	f	500	fe
ⳍ	x	600	enguz
Ⳏ	h	700	uuaer
ⳏ	o	800	utal
Ⳑ	—	900	—



The essentially Greek inspiration of this alphabet is shown by a number of features, including:

1. The form of the letters, about two-thirds of which closely resemble their uncial Greek counterparts;
2. The order of the letters and their associated numerical values;
3. Greek orthographic practices, such as the (late) use of *ai* to stand for the monophthong [ɛ], and the use of *g* to stand for the the velar nasal [ŋ] before velar consonants.

Wulfila did not, however, adhere slavishly to his Greek model. In several instances he assigned altogether new values to Greek letters which would otherwise have been useless in Gothic. This was the case with Greek **F** ([w]), which became Gothic *q* ([k<sup>w</sup>]), and with **Ψ** (psi), which was probably the source of the Gothic character *hv* ([h<sup>w</sup>]). Curiously, Wulfila chose not to use the letters **Φ** (phi) and **Θ** (theta) to write the Gothic voiceless fricatives [f] and [θ], respectively, despite the fact that **Φ** and **Θ** had precisely these values in fourth-century Greek. Instead, he employed **Φ** to write Gothic [θ] and borrowed the Latin letter **F** to write Gothic [f]. The new phonetic value of **Φ** led to its being moved to the alphabetic position formerly occupied by **Θ**, while the new Latin-derived *f* took over the place vacated by **Φ**. Other Latin letters that found their way into the Gothic alphabet were *r* and *h*, as well as the variant of the *s*-character used in the Codex Argenteus (other Gothic manuscripts show an *s* that is decidedly more Greek-looking). In addition, several Gothic letters have been claimed to come from the runic alphabet – *u*, for example, which Wulfila used in place of the Greek digraph **OY**. But the extent to which runic writing played a role in the creation of the Gothic alphabet is highly controversial, not least because many of the characters in the runic alphabet are very similar to their Latin counterparts.

### 3. PHONOLOGY

#### 3.1 Consonants

The most highly structured part of the Gothic consonant system consists of a symmetrically organized subsystem of twelve stops and fricatives (the term *coronal* is used here to denote the dental, alveolar, and palatal regions):

(1)	<i>Labial</i>	<i>Coronal</i>	<i>Velar</i>	<i>Labiovelar</i>
<i>Voiceless stops</i>	/p/	/t/	/k/	/k <sup>w</sup> / <q>
<i>Voiceless fricatives</i>	/f/	/p/	/h/	/h <sup>w</sup> / <hv>
<i>Voiced stops/Fricatives</i>	/b/	/d/	/g/	/g <sup>w</sup> / <gw>

Of the voiceless stops, the labial /p/ is infrequent outside obvious Greek and Latin loanwords (e.g., *prauſetus* “prophet,” *pund* “pound”). The labiovelar /k<sup>w</sup>/, which Wulfila’s native-speaker intuition led him to write with a single character (*q*), patterns phonotactically as a single consonant (cf. *qrammiþa* “moistness,” with initial *qr-*) and is best analyzed as a unitary phoneme. The voiceless fricatives include /h/ and /h<sup>w</sup>/ (likewise a unitary phoneme), which, phonetically, were probably indistinguishable from the English sounds spelled *h* and *wh* – in other words, simple glottal fricatives with no significant velar occlusion. (This was doubtless also the case in syllable-final position, as, e.g., in *sahv* “saw” [1st, 3rd sg.], *nahts* “night” and *sahvt* “saw” [2nd sg.]; the development of [h] to velar [x] in this position in German [cf. *Nacht*, etc.] had no parallel in Gothic). Historically, however, they arose from older \*x and

\*x<sup>w</sup>, and structurally their place is still clearly with the oral fricatives /f/ and /p/, with which they share important distributional properties.

The sounds denoted by the letters *b*, *d*, *g(w)* were voiced stops in some environments and voiced fricatives in others. The stop reading is certain after consonants (e.g., *windan* [windan] “wind,” *siggwan* [sing<sup>w</sup>an] “sing,” *þaurban* “need” [þorban]), and probable, at least for *b* and *d*, in word-initial position (*barn* [b-] “child,” *dags* [d-] “day”). After vowels, single *b*, *d*, and *g* are fricatives (e.g., *sibun* [siβun] “seven,” *bidjan* [biðjan] “ask,” *ligan* [liɡan] “lie.” The stop /g<sup>w</sup>/ is found only after nasals (in words like *siggwan*) and in the geminate combination -ggw- (e.g., *bliggwan* [-gg<sup>w</sup>-] “strike”); there is thus no fricative allophone [g<sup>w</sup>].

The remaining Gothic consonants include two sibilants and a standard complement of nasals, liquids, and glides:

(2)	Labial	Coronal	Velar
<i>Nasals</i>	/m/	/n/	([ŋ] <g>)
<i>Voiceless sibilant</i>		/s/	
<i>Voiced sibilant</i>		/z/	
<i>Liquids</i>		/r/, /l/	
<i>Glides</i>	/w/	/y/	

The voiced sibilant /z/ is not found in word-initial position. The velar nasal [ŋ], spelled <g> in imitation of Greek practice, is the automatic realization of /n/ before velar and labiovelar stops. The graphic sequence -ggw- is thus ambiguous, representing both [-gg<sup>w</sup>-] and [-ŋg<sup>w</sup>-].

## 3.2 Vowels

Gothic has five short and seven long vowels, along with a single diphthong:

(3)	Short		Long		
	Front	Back	Front	Back	
	High	/i/ <i>	/u/ <u>	/i:/ <ei>	/u:/ <u>
	High-mid			/e:/ <e>	/o:/ <o>
	Low-mid	/ɛ/ <ai>	/ɔ/ <au>	/ɛ:/ <ai>	/ɔ:/ <au>
	Low	/a/ <a>		/a:/ <a>	
	Diphthong	/iu/			

### 3.2.1 Short vowels

Among the short vowels, /ɛ/ and /ɔ/ are only marginally phonemic, being in most cases mere positional variants of underlying /i/ and /u/ before -r, -h, and -h (breaking; see §3.4.2). But both have a general distribution in foreign (i.e., Greek and Biblical Semitic) words (e.g., *aikklesjo* [ɛkkle:sjo:] “church,” Greek ἐκκλησία; *apaustaulus* [apɔstɔlus] “apostle,” Greek ἀπόστολος), and /ɛ/ serves as the normal reduplication vowel in native Gothic preterites of the type *letan* – *lailot* [lelo:t] “let,” *aukan* – *aiauk* [ɛɔ:k] “increase.” The use of the graphic diphthong <ai> to stand for a front monophthong is based directly on late Greek practice; the parallel use of <au> for [ɔ] is an innovation of Wulfila’s system.

### 3.2.2 Long vowels

The long vowels include the high-mid vowels /e:/ and /o:/, which lack short counterparts and are unambiguously indicated by the letters *e* and *o*. The Gothic alphabet, however, does not mark length as such. The long versions of [a], [ɛ], [ɔ], and [u] are not written differently from their short equivalents; orthography alone gives no indication that *pahta* “(s)he thought,” *air* “early,” *hauhs* “high,” and *bruþs* “young woman” represent [pa:hta], [ɛ:r], [hɔ:hs], and [bru:þs], respectively, with distinctive length (note that the modern editorial practice of writing *þâhta*, *áir*, *haúhs*, and *brúþs* to indicate length, and writing *ái* and *áu* for short /ɛ/ and /ɔ/, has no basis in ancient usage). The case of /i/ and /i:/, which are orthographically distinguished as <i> and <ei> (cf. *bitan* “bitten” [nom. sg. neut.] vs. *beitan* “to bite” [inf.]), is exceptional. Wulfila’s practice probably reflects a qualitative difference between the two *i*-vowels, perhaps comparable to that between the relatively low [-ɪ-] and the relatively high [-i:-] of German *bitten* “ask” versus *bieten* “offer.”

The seven long vowels show considerable differences of patterning and distribution. Low central /a:/ is rare, being confined in the native Gothic lexicon to etymological sequences of *\*-anh-*, which yielded [-ãh-] in Proto-Germanic and subsequently lost its nasalization in Gothic (cf. 3.4.4). The lower-mid vowels /e:/ and /ɔ:/, on the other hand, are relatively common; they represent the Proto-Germanic diphthongs *\*ai* and *\*au* and pattern as the *o*-grade counterparts of /i/ and /u/. There is little basis for the view, rooted in a coincidence of Germanic etymology and Greek orthography, that “long” *ai* and *au* actually represent synchronic diphthongs in Wulfila’s Gothic. The only true Gothic diphthong is /iu/.

### 3.3 Accent

The position of the word accent is not overtly indicated. To judge from the other Germanic languages, ordinary words were stressed on their first syllable. But in verbal compounds consisting of a prefix and a lexical verb, the prefix was proclitic, so that the accent probably remained on the initial syllable of the verbal root (cf. *af-niman* [af-níman] “take away” and *and-niman* [and-níman] “receive,” with the accentuation of the simplex *niman* [níman] “take”). The accent pattern of the corresponding nominal compounds (e.g., *anda-numts* “reception,” *anda-numja* “receiver”) is uncertain.

### 3.4 Synchronic phonological processes

A number of automatic phonological rules, reflecting historical sound changes, affect the surface form of Gothic words.

#### 3.4.1 Word-final devoicing

This rule applies exclusively to fricatives, converting [b], [d], [g], and [z] to [f], [p], [x], and [s] in absolute-final position: for example, *gaf* < *\*gab*, third singular preterite of *giban* “give”; *baþ* < *\*bad*, third singular preterite of *bidjan* “ask”; *maujos* < *\*maujoz*, genitive singular of *mawi* “girl.” The devoicing of [g] to [x] is not noted orthographically (cf. *mag* [max] “is able”), presumably because the [g] : [x] contrast was not phonemic and there was no letter in ordinary use to denote the voiceless velar fricative (Wulfila’s use of the letter *x* is virtually confined to the divine name *Xristus* “Christ”). No devoicing is found in forms of the type *band* “bound” and *ward* “word,” showing that the final consonant was a stop in these environments.

### 3.4.2 Breaking

This is the traditional name (German *Brechung*) for the regular lowering of synchronically underlying *\*i* and *\*u* to *ai* [ɛ] and *au* [ɔ] before *-r*, *-h*, and *-hv*; for example, *wairþan* “become,” first singular preterite *warþ*, first plural preterite *waurþum*, participle *waurþans*, paralleling the regular pattern seen in *hilpan* “help” *halp*, *hulpum*, *hulþans*.

### 3.4.3 Hiatus lowering

This is the regular but comparatively rare process by which long high and high-mid vowels were replaced by their low-mid counterparts when immediately followed by another vowel: as in *saian* [se:an] < *\*sean* [se:an] “sow”; *stauida* [stɔ:iða] < *\*stoida* [stɔ:iða], third singular preterite of *stojan* “judge.”

### 3.4.4 Loss of *-n-* before *-h-* with compensatory lengthening

This process is found not only after *-a-* (cf. *þahta* < *\*þanhta*; see §3.2.2), but also after *-u-* (cf. *þuhta* < *\*þunhta*, third singular preterite of *þugkjan* “seem”) and *-i-* (cf. *þeihana* < *\*þinhan* “prosper”). The nasalized vowels that originally resulted from *\*-Vnh-* sequences fell together with non-nasal /a:/, /u:/, and /i:/ in Wulfila’s language.

## 3.5 Morphophonemic processes

Phonological processes that have been *morphologized*, i.e., restricted to specific morphemes and/or morphological categories, include the following:

### 3.5.1 Grammatical change

Grammatical change (German *grammatischer Wechsel*) is the traditional name for the alternation of word-internal voiceless and voiced fricatives (or stops derived from fricatives) under conditions originally governed by Verner’s Law (see §3.6.2): for example, *hafjan* “lift” versus *uf-haban* “lift up”; *fra-wairþan* “perish” versus *fra-wardjan* “destroy”; third singular *aiþ* [ɛ:h] “has” versus third plural *aigun* [ɛ:gun]. Voiced : voiceless pairs of this type are much rarer in Gothic than in the other early Germanic languages. But Gothic has a number of derivational suffixes which vary according to *Thurneysen’s Law*: a voiced fricative appears when the preceding syllable begins with a voiceless consonant, and vice versa: for example *aupida* “desert” versus *diupiþa* “depth”; *wulþags* “glorious” versus *stainahs* “stony”; *fraistubni* “temptation” versus *waldufni* “power”.

### 3.5.2 Ablaut

Ablaut, or apophony, is the system of morphologically governed vowel alternations inherited by Gothic and the other Germanic languages from Proto-Indo-European (PIE). The clearest examples are seen in the formation of the principal parts of strong verbs, as in *wairþan* (< PIE *\*wert-*; “e-grade”), *warþ* (< PIE *\*wort-*; “o-grade”), *waurþum* (< PIE *\*wrt-*; “zero-grade”), *waurþans* (likewise < PIE *\*wrt-*). But ablaut changes are also associated with other derivational and inflectional processes, ranging from the inflection of *n*-stem nouns (e.g., acc. sg. *auhsan* “ox” < pre-Germanic *\*ukson-*; dat. sg. *auhsin* < *\*uksen-*; gen. pl. *auhsne* < *\*uksn-*) to the formation of causatives from underlying strong verbs (e.g., *frawairþan* → *frawardjan*, *sitan* “sit” → *satjan* “set”).

### 3.5.3 Sievers' Law

Sievers' Law describes the regulated distribution – observable in both *ja*-stem nouns and adjectives, and in verbs with infinitives in *-jan* – of *-ji-* after “light” sequences (i.e., sequences of the form *\*-VC-*) and *-ei-* [i:] after “heavy” sequences (i.e., sequences of the form *\*-V̄C-* and *\*-VCC-*): e.g., *harjis* “army” versus *hairdeis* “shepherd”; third singular *satjīþ* “sets” versus *frawardeiþ* “destroys.” In its Proto-Indo-European form, Sievers' Law mandated the realization of underlying *\*-y-* as *\*-iy-* after heavy sequences; the *-ei-* of *hairdeis* and *frawardeiþ* is the contraction product of pre-Germanic *\*-iji-*.

### 3.5.4 Dental substitution

Suffix-initial *-d-* is replaced by *-s-* after an immediately preceding root-final *-t-* or *-d-*, or by *-t-* after any other root-final obstruent. In the former case the root-final *-t-* or *-d-* itself becomes *-s-*; in the latter case the root-final obstruent is represented by the corresponding voiceless fricative: for example, *witan* “know,” preterite *wissa*; *þaurban* “need,” preterite *þaurfta*; *magan* “be able,” preterite *mahta*. Contrast the “normal” pattern seen in *munan* “think,” preterite *munda*; *satjan*, preterite *satida*; etc. These alternations reflect the special treatment of dental + dental clusters in Proto-Indo-European, and the failure of voiceless stops to undergo the Germanic Consonant Shift (see §3.6.1) when preceded by an obstruent.

### 3.5.5 Clitic-related effects

Word-final *-s* usually becomes *-z-* before vowel-initial enclitics, especially *-(u)h* “and” and the relativizing particle *-ei-*: e.g., *hwazuh* “each” < nominative singular masculine *hwās* “who” + *-uh* (cf. Lat. *quisque*), where the final *-s* is a devoiced etymological *\*-z*; and *þizei* “whose” < genitive singular masculine *þis* “his” + *-ei-*, where the *-z* is analogical. Similar effects are seen in the behavior of prefixes; compare the variant forms in *us-hafjan* “lift up,” *uz-anan* “breathe out,” and *ur-reisan* “arise.” The final *-h* of *-(u)h* sometimes assimilates to a following *-þ-*, as in *wesunupþan* (= *wesun-uh-þan*) “but there were,” *sumaiþþan* (= *sumai-h-þan*) “but some,” etc.

## 3.6 Diachronic developments

### 3.6.1 Grimm's Law

As a Germanic language, Gothic shared in the characteristic phonological developments that set Germanic apart from the rest of the Indo-European family. The most conspicuous sound change in the prehistory of Germanic was *Grimm's Law* or the *Germanic Consonant Shift*, which took place in three steps:

- (4) A. PIE voiceless stops *\*p*, *\*t*, *\*k̑* (+ *\*k*)<sup>1</sup> *\*k<sup>w</sup>* became the voiceless fricatives *\*f*, *\*þ*, *\*x* (> *h*), *\*x<sup>w</sup>* (> *\*h<sup>w</sup>*) when not preceded by an obstruent
- B. PIE voiced stops *\*b* (rare), *\*d*, *\*ǵ* (+ *\*g*), *\*g<sup>w</sup>* became the voiceless stops *\*p*, *\*t*, *\*k*, *\*k<sup>w</sup>*
- C. PIE voiced aspirated stops *\*b<sup>h</sup>*, *\*d<sup>h</sup>*, *\*ǵ<sup>h</sup>* (+ *\*g<sup>h</sup>*), *\*g<sup>wh</sup>* became the voiced fricatives *\*ḃ*, *\*ḋ*, *\*ḡ*, *\*ḡ<sup>w</sup>*, which further developed to voiced stops in some environments

Examples are legion: compare (A) Go. *fotus* (Eng. *foot*), *þrija* (Eng. *three*), *haur̥n* (Eng. *horn*), *hwata* (Eng. *what*) beside Lat. *pēs*, *trēs*, *cornu*, *quod*; (B) Go. *tunþus* (Eng. *tooth*), *kaurn* (Eng. *corn*), *qius* (Eng. *quick*) beside Lat. *dēns*, *grānum*, *uīuus* (< \**g<sup>w</sup>*īwos); (C) Go. *beitan* (Eng. *bite*), (*ga*)-*daursan* (Eng. *dare*), *gaitis* (Eng. *goat*), *warmjan* (Eng. *warm*, with *w*- < \**g<sup>w</sup>*-) beside Skt. *bhid-* “split,” *dhṛs-* “be bold,” Lat. *haedus* (< \**x*- < \**k<sup>h</sup>*- < \**g<sup>h</sup>*-), Skt. *gharmá-* (< \**g<sup>w</sup>**h*-) “hot drink.”

The voiceless stops, however, remained unchanged after \**s* (cf. Go. *steigan* “climb” beside Gk. στεῖχω [*steík<sup>h</sup>*ō] “id.”) or when preceded by another stop (cf. Go. *-haftis* “having, having taken” beside Lat. *captus* “taken”).

### 3.6.2 Verner’s Law

The Germanic Consonant Shift applied both word-initially and word-internally (Proto-Indo-European word-final stops were lost). In word-internal position, however, the voiceless fricatives produced by the shift, together with the inherited sibilant fricative \**s*, were potentially subject to *Verner’s Law*. The effect of this rule was to convert \**f*, \**þ*, \**x*, \**x<sup>w</sup>*, and \**s* to the corresponding voiced fricatives \**b*, \**d*, \**g*, \**g<sup>w</sup>*, and \**z* when the preceding vowel *did not* bear the pre-Germanic (equivalent to the Proto-Indo-European) movable accent. Thus, the Proto-Indo-European word for “father,” which was accented on the second syllable (cf. Skt. *pitár-*, Gk. πατήρ [*patér*]), gave \**faþér* by Grimm’s Law and \**faðér* (> Go. *fadar*) by Verner’s Law, while the word for “brother,” which had initial accent (cf. Skt. *bhrátar-*, Gk. φράτηρ [*phrátēr*]), became \**brôþēr* by Grimm’s Law and retained its voiceless \**-þ-* in Gothic (*broþar*). Following the operation of Verner’s Law, the pre-Germanic system of “free” accent was replaced by the attested Germanic system of fixed initial stress (see §3.3), so that the original condition for the voicing of word-internal fricatives can no longer be detected synchronically in Gothic or in any other Germanic language.

### 3.6.3 Further obstruent developments

The obstruent system that emerged from the operation of Grimm’s and Verner’s Laws was subject to further changes within the Germanic period, notably the following:

1. The weakening of \**x* and \**x<sup>w</sup>* to \**h* and \**h<sup>w</sup>*.
2. The “strengthening” of \**b*, \**d*, \**g*, and \**g<sup>w</sup>* to stops after nasals and, at least in the case of \**b* and \**d*, word-initially.
3. The development of the fricative \**g<sup>w</sup>* to \**w* in most remaining environments (though \**g<sup>w</sup>* was dissimilated to \**g* before a following \**u*; note the Gothic pair *magus* “boy” < \**mag<sup>w</sup>*uz vs. *mawi* “girl” < \**mag<sup>w</sup>*i).
4. The change of \**s* to \**z*, regardless of the original position of the accent, in absolute final position.

The resulting Proto-Germanic system was hardly modified in Gothic at all, save by the introduction of final devoicing and by the substitution of [b], [d], [g] for [b̥], [d̥], [g̥] after non-nasal consonants (*waurd*, etc.; see §3.4.1).

### 3.6.4 Sonorant developments

The Proto-Indo-European consonant system also included the liquids \**r* and \**l*, the nasals \**m* and \**n* (the latter with a velar allophone [ŋ]), the glides \**y* and \**w*, and the three so-called laryngeals \**h*<sub>1</sub>, \**h*<sub>2</sub>, and \**h*<sub>3</sub>, of uncertain phonetic value. The liquids were preserved unchanged in Germanic and Gothic. This was also true of the nasals except before \**h*

and in absolute final position, where *\*-m* and *\*-n* fell together and eventually disappeared. But the fate of the glides *\*y* and *\*w* was more complicated. Word-initially and postconsonantly, *\*y* and *\*w* were preserved as Germanic *\*j* and *\*w*, respectively (cf. Go. *juk* [Eng. *yoke*], *winds* [Eng. *wind*] beside Lat. *iugum*, *uentus*). After vowels, however, there were two basic treatments:

1. Germanic *\*-Ø-* and *\*-w-*, respectively (cf. Go. *bau-an* “dwell” < *\*bhū-ye/o-*; *aiws* “age, time” beside Lat. *aeuom*). A specifically Gothic change subsequently deleted *\*-w-* after the rounded vowel *o* (cf. *stojan* “judge” < *\*stōwjan*, pret. *stauida* < *\*stōida* < *\*stōwida*).
2. Germanic *\*-jj-* and *\*-ww-*, respectively, whence Gothic *-ddj-* and *-ggw-*, respectively: e.g., Gmc. *\*twajjōn* “of two” (gen.), Go. *twaddje* (cf. Skt. *dvayoh* “id.”); Gmc. *\*trewwaz* “true,” Go. *triggws* (cf. Old Prussian *druwīt* “believe”). The seemingly irregular doubling or *Verschärfung* of *\*-y-* and *\*-w-* to *\*-jj-* and *\*-ww-* is now thought to reflect the original presence of a Proto-Indo-European laryngeal after the glide.

Apart from their role in *Verschärfung*, laryngeals had much the same treatment in Germanic as in the other Indo-European languages; their typical fate was to disappear with compensatory lengthening of an immediately preceding vocalic element in the same syllable. The vocalic element in question might be a vowel proper (*\*e*, *\*a*, etc.) or a syllabic liquid (*\*ṛ*, *\*ḷ*) or nasal (*\*ṁ*, *\*ṅ*) – the syllabic liquids or nasals being non-contrastive sounds which served in Proto-Indo-European as allophones of consonantal *\*r*, *\*l*, *\*m*, *\*n*.

### 3.6.5 Vocalic developments

#### 3.6.5.1 Proto-Indo-European

Following the loss of laryngeals, the Proto-Indo-European dialect ancestral to Germanic had five short and five long vowels:

	Short		Long	
	Front	Back	Front	Back
High	i	u	ī	ū
Mid	e	o	ē	ō
Low	a		ā	

(It is no longer customary to include a central mid vowel *\*ə* in the inventory of Proto-Indo-European short vowels. The sound denoted by this symbol in older handbooks was a subphonemic support vowel; cf., e.g., *\*ph<sub>2</sub>tér* [p<sub>3</sub>h<sub>2</sub>té:r], which was eventually phonologized as /a/ in most Indo-European languages.) In addition, there were four short and four long syllabic liquids and nasals:

(6) *ṛ*, *ḷ*, *ṁ*, *ṅ*      *ī̄*, *ū̄*, *ṁ̄*, *ṅ̄*

and six short and six long *i-* and *u-*diphthongs:

(7) *ei*    *ai*    *oi*      *ēi*    *āi*    *ōi*  
*eu*    *au*    *ou*      *ēu*    *āu*    *ōu*

This is the inventory of syllabic nuclei that must be taken as the point of departure for the history of the Proto-Indo-European vowel system in Germanic.

#### 3.6.5.2 Proto-Germanic

The number of vowels and vowel-like elements was greatly reduced over the course of the three millennia or so that passed between dialectal Proto-Indo-European and



Proto-Germanic. An early development was the shortening of the long diphthongs and the long syllabic liquids and nasals, which merged with their short counterparts; syllabic liquids and nasals were subsequently eliminated altogether by the change of *\*ŕ, \*ŕh, \*ŕ* to the vowel + consonant sequences *\*ur, \*ul, \*um, \*un*: e.g., Gothic *fulls* “full” < *\*fulnaz* < *\*pŕnós* < *\*pŕh<sub>1</sub>-nó-s*; *hund* “100” < *\*hundan* < *\*kŕntóm*; *haurn* “horn” < *\*hurnan* < *\*kŕnóm*. Among the vowels proper, the *\*a : \*o* distinction was lost in both the long and short subsystems, the longs merging as *\*ō* (cf. Go. *broþar, bloma* “flower” beside Lat. *frāter, flōs*) and the shorts as *\*a* (cf. Go. *akrs* “field,” *ahtau* “eight” beside Lat. *ager, octō*). (It is interesting to note that a similar confusion of *a-* and *o-*vowels occurred in the neighboring Indo-European languages, Celtic and Balto-Slavic.) There was also a change of short *\*e* to *\*i* in certain environments: for example, before nasal clusters (*\*-nt-, \*-mb-, etc.*), and before an *\*i* in the next syllable (cf. Old High German *bintan*, Gothic *bindan* “bind” < *\*b<sup>h</sup>énd<sup>h</sup>onom*; OHG *ist*, Go. *ist* “is” < *\*éstī*; but OHG *geban*, Go. *giban* “give” < *\*g<sup>h</sup>éb<sup>h</sup>onom*; forms are cited from Old High German to show the still recoverable difference between Germanic *\*e* and *\*i*, which was effaced entirely in Gothic). These developments were paralleled in the treatment of the diphthongs: *\*ai* and *\*oi* merged as *\*ai*; *\*au* and *\*ou* merged as *\*au*; *\*ei* gave *\*ī* (i.e., /ii/, spelled <ei>; cf. Go. *steigan* [OHG *stigan*] beside Gk. *στεῖχω* (*steík<sup>h</sup>ō*)); and *\*eu* gave the new diphthong *\*iu* before an *\*i* in the following syllable (cf. OHG 3rd sg. *biutit* “offers” < Gmc. *\*biudīþ*, but inf. *beotan, biotan* < Gmc. *\*beudan*). Within the long vowel subsystem, *\*ē* was phonetically lowered to approximately the sound heard in English *sad* (i.e., [æ]), while the phonetic place of the old *\*ē* was taken over by a new vowel *\*ē<sub>2</sub>*, of obscure origin.

The result of the foregoing, in the end, was the vowel system reconstructible for Proto-Germanic:

#### (8) Proto-Germanic monophthongs

	Short		Long	
	Front	Back	Front	Back
High	i	u	ī	ū
Mid	e	[o]	ē <sub>2</sub>	ō
Low		a		æ

Some authorities set up a secondary short *\*o* for Proto-Germanic, but there is no evidence for such a vowel in the prehistory of Gothic, and it can equally well be explained as a common innovation of the North and West Germanic dialects. The low vowel *\*æ* is commonly also written *\*ē* or *\*ē<sub>1</sub>*.

#### (9) Proto-Germanic diphthongs

ai   au  
       eu  
       iu

In addition, there were also nasalized *\*ā<sup>N</sup>, \*ī<sup>N</sup>, \*ū<sup>N</sup>*, and probably – at least in final syllables – other nasalized vowels as well. All were purely allophonic.

##### 3.6.5.3 Gothic

The main Gothic innovations in the treatment of the Germanic short vowels were the complete merger of *\*e* and *\*i* as *i* (cf. Go. *giban* beside OHG *geban*, etc.) and the subsequent



creation of new low-mid vowels by “breaking” before *-r*, *-h*, and *-hv* (see §3.4.2). The long vowels were somewhat more extensively restructured, with *\*ā* and *\*ē* falling together as the high-mid vowel written *e* (cf. Go. *her* “here” [OHG *her*, *hiar*] < *\*he<sub>2</sub>r*, identical in vocalism with first plural preterite *gebum* “we gave” [OHG *gābum*] < *\*gābum*), and a new *ā* joining the system through the denasalization of *\*ā<sup>N</sup>*. Here as in the shorts, the system was expanded by the addition of new low-mid vowels – this time through the monophthongization of *\*ai* and *\*au* (cf. §3.4.2). As a byproduct of the general shift of short *\*e* to *\*i*, the two remaining diphthongs, *\*eu* and *\*iu*, fell together as *\*iu* in Gothic (cf. *-biudan*, *-biudip* beside OHG *biotan*, *biutit*).

Gothic shows major changes vis-à-vis Proto-Germanic in its treatment of final syllables. Proto-Germanic generally preserved the vowels of late Proto-Indo-European final syllables intact; thus, for example, the *o*-stem nominative singular in *\*-os* was still *\*-az* in Proto-Germanic (cf. Runic Norse *-aR*; and see Ch. 37, §2.1), and the first singular present in *\*-ō* (< *\*-oh<sub>2</sub>*) remained as *\*-ō*. In addition to normal long and short endings, however, Proto-Germanic also had final syllables with hyperlong or “trimoric” long vowels; these mainly arose from prehistoric sequences of two vowels in hiatus (e.g., PGmc. *\*galikō* “similarly,” with trimoric or “circumflex” *\*-ō* from PIE *\*-o-h<sub>2</sub>ad*). Gothic is often said to have undergone a “law of three moras” or *Dreimorengesetz*, under which short vowels were lost (cf. nom. sg. *dags* “day” < *\*dagaz*) in final syllables, normal (bimoric) long vowels were shortened (cf. 1st sg. *nima* “I take”), and trimoric long vowels became bimoric longs (cf. *galeiko*). But this generalization is not completely valid: *\*-u(-)* was never lost at all (cf. *sunus* (< *\*-uz*) “son,” *faihu* (< *\*-u*) “cattle”), and even bimoric long vowels retained their length before *\*-z* (acc. pl. *gibos* “gifts” < *\*-ōz* < late PIE *\*-ās* < *\*-ah<sub>2</sub>(m)s*). As in every other Germanic language, the *Auslautgesetze* of Gothic still present many problems.

## 4. MORPHOLOGY

### 4.1 Nominal morphology

From a morphological point of view, Gothic is an averagely conservative older Indo-European language, similar in overall complexity to, e.g., Old Church Slavonic. Nouns come in three genders (masculine, feminine, neuter) and distinguish five cases (nominative, vocative, genitive, dative, accusative). There are singular and plural forms, but no dual (though the dual survives in personal pronouns; see §4.1.4). A number of features familiar from other Indo-European languages, such as the identity of the nominative and accusative cases in the neuter, and the identity of the nominative and vocative in the plural, appear in Gothic as well.

#### 4.1.1 Nominal case development

Proto-Indo-European had eight cases: nominative, accusative, instrumental, dative, ablative, genitive, locative, and vocative. Of these, the ablative was lost in Germanic (it survives in adverbs like Gothic *galeiko* “similarly”; see §4.3), and the dative and the locative merged to form the synchronic dative. The instrumental, which was still a separate case in Proto-Germanic, was absorbed by the dative in the post-Germanic history of Gothic; thus, a form which patterns as a dative in Gothic may in principle go back to a Proto-Indo-European dative, locative, or instrumental.

### 4.1.2 Nominal stem-classes

Gothic declensions are conveniently classified according to the original stem-final element, which is usually best preserved in the dative plural and/or accusative plural. The most important types, as in the other Germanic languages, are (i) *a*- and *ja*-stems; (ii) *ō*- and *jō*-stems; (iii) *i*-stems; (iv) *u*-stems (collectively termed *strong*); and (v) *n*-stems (traditionally termed *weak*). The basic paradigms are given in Table 36.2.

In the *ja*-stems, the difference between *hairdeis* and *harjis* is due to Sievers' Law (see §3.5.3). The endings of *i*-, *u*-, and *n*-stems show traces of stem-final ablaut: *anstim* : *anstaïs* : *ansteis* (< \**-ey-es*); *sunum* : *sunaus* : *sunīwe* (< \**-ew-ōm*); *guma* (< \**-ō(n)*) : *gumins* : *gumans*; and *namō* (< \**-ō(n)*) : *namins* : *namna*. Minor declensional types include relics of other consonant-stem classes, especially *r*- and *nt*-stems (e.g., *broþar*, gen. *broþrs*, nom. pl. *broþrjus*; *nasjands* "savior," gen. *nasjandis*, nom. pl. *nasjands*).

#### 4.1.2.1 Ablaut and accent patterns

Proto-Indo-European nouns, with the exception of *o*-stems (> Gmc. (*j*)*a*-stems) and *ā*-stems (> Gmc. *ō*-stems), were characterized by complex alternations of ablaut and accent which affected the root, the derivational suffix that optionally followed the root, and the grammatical ending proper or *desinence*. Four or five such ablaut/accent patterns can be reconstructed for stems containing a suffix (e.g., \**-t(e/o)r-*, \**-(e/o)n-*, \**-w(e/o)nt-*, \**-t(e/o)i-*, etc.). Thus, for example, the oldest recoverable declension of the Proto-Indo-European word for "father" (Go. *fadar*) was of the *hysterokinetic* type, with nominative singular \**ph<sub>2</sub>-tér* (zero-grade root, accented *ē*-grade suffix, zero desinence), accusative singular \**ph<sub>2</sub>-tér-ŋ* (accented *e*-grade suffix, invariant desinence), and genitive singular \**ph<sub>2</sub>-tr-és* (zero-grade suffix, accented *e*-grade desinence). Quite different from this was the declension of the word for "sowing, seed" (Go. *seþs*; *i*-stem), which was *proterokinetic*, with nominative singular \**séh<sub>1</sub>-ti-s*, accusative singular \**séh<sub>1</sub>-ti-m* (accented *e*-grade root, zero-grade suffix, invariant desinence), and genitive singular \**sh<sub>1</sub>-tēi-s* (zero-grade root, accented *e*-grade suffix, zero-grade desinence). Root nouns – nouns lacking a derivational suffix – displayed comparable inner-paradigmatic allomorphy, as in the Proto-Indo-European word for "foot" (Go. *fotus*): nominative singular \**pód-s* (*ō*-grade root, invariant desinence), accusative singular \**pód-ŋ* (*o*-grade root, invariant desinence), genitive singular \**péd-s* (*e*-grade root, zero-grade desinence).

Little remains of this complexity in Germanic and Gothic. Root ablaut was almost completely abandoned within paradigms (*seþs* and *fotus* generalized the vocalism of the nominative singular), and suffixes and desinences fused to form what can be described synchronically as "*i*-stem endings," "*u*-stem endings," "*n*-stem endings," etc. Only the *n*-stems, which underwent a period of great expansion in Germanic, retain something of the variety of Indo-European ablaut patterns, as can be seen by comparing the morphological differences between *guma*, *hairto*, and *namo* (see Table 36.2; the feminine *n*-stem types – *qino* and *managei* – are entirely a Germanic innovation).

#### 4.1.2.2 Gothic *ō*- and *jō*-stems

The Proto-Indo-European *o*- and *ā*-stems (i.e., thematic and *eh<sub>2</sub>*-stems respectively) lacked the ablaut alternations of the other stem-types – a fact no doubt partly responsible for their frequency and productivity around the family. In Gothic the *ō*-stems (< *ā*-stems) in particular retain a fairly transparent declension, with the historical desinences added to the still-preserved stem-vowel (e.g., dat. sg. *gibai* < \**-āi* < \**-eh<sub>2</sub>-ei*; nom. pl. *gibos* < \**-ās*

**Table 36.2 Gothic nominal stems**

*a-* and *ja-*stems (*hlaifs* [masc.] “bread,” *waurd* [neut.] “word,” *hairdeis* [masc.] “shepherd,” *harjis* [masc.] “army,” *kuni* [neut.] “race”):

<i>Sg. nom.</i>	<i>hlaifs</i>	<i>waurd</i>	<i>hairdeis</i>	<i>harjis</i>	<i>kuni</i>
<i>voc.</i>	<i>hlaif</i>	<i>waurd</i>	<i>hairdi</i>	<i>hari</i>	<i>kuni</i>
<i>gen.</i>	<i>hlaibis</i>	<i>waurdis</i>	<i>hairdeis</i>	<i>harjis</i>	<i>kunjis</i>
<i>dat.</i>	<i>hlaiba</i>	<i>waurda</i>	<i>hairdja</i>	<i>harja</i>	<i>kunja</i>
<i>acc.</i>	<i>hlaif</i>	<i>waurd</i>	<i>hairdi</i>	<i>hari</i>	<i>kuni</i>
<i>Pl. nom.</i>	<i>hlaibos</i>	<i>waurda</i>	<i>hairdjos</i>	<i>harjos</i>	<i>kunja</i>
<i>gen.</i>	<i>hlaibe</i>	<i>waurde</i>	<i>hairdje</i>	<i>harje</i>	<i>kunje</i>
<i>dat.</i>	<i>hlaibam</i>	<i>waurdam</i>	<i>hairdjam</i>	<i>harjam</i>	<i>kunjam</i>
<i>acc.</i>	<i>hlaibans</i>	<i>waurda</i>	<i>hairdjans</i>	<i>harjans</i>	<i>kunja</i>

*ō-* and *jō-*stems (*giba* [fem.] “gift,” *bandi* [fem.] “bond,” *mawi* [fem.] “girl”):

<i>Sg. nom.</i>	<i>giba</i>	<i>bandi</i>	<i>mawi</i>
<i>voc.</i>	<i>giba</i>	<i>bandi</i>	<i>mawi</i>
<i>gen.</i>	<i>gibos</i>	<i>bandjos</i>	<i>maujos</i>
<i>dat.</i>	<i>gibai</i>	<i>bandjai</i>	<i>maujai</i>
<i>acc.</i>	<i>giba</i>	<i>bandja</i>	<i>mauja</i>
<i>Pl. nom.</i>	<i>gibos</i>	<i>bandjos</i>	<i>maujos</i>
<i>gen.</i>	<i>gibo</i>	<i>bandjo</i>	<i>maujo</i>
<i>dat.</i>	<i>gibom</i>	<i>bandjom</i>	<i>maujom</i>
<i>acc.</i>	<i>gibos</i>	<i>bandjos</i>	<i>maujos</i>

*i-* and *u-*stems (*gasts* [masc.] “guest,” *anstis* [fem.] “favor,” *sunus* [masc.] “son”):

<i>Sg. nom.</i>	<i>gasts</i>	<i>anstis</i>	<i>sunus</i>
<i>voc.</i>	<i>gast</i>	<i>anstis</i>	<i>sunau, -u</i>
<i>gen.</i>	<i>gastis</i>	<i>anstais</i>	<i>sunaus</i>
<i>dat.</i>	<i>gasta</i>	<i>anstai</i>	<i>sunau</i>
<i>acc.</i>	<i>gast</i>	<i>anst</i>	<i>sunu</i>
<i>Pl. nom.</i>	<i>gasteis</i>	<i>ansteis</i>	<i>sunjus</i>
<i>gen.</i>	<i>gaste</i>	<i>anste</i>	<i>sunuwe</i>
<i>dat.</i>	<i>gastim</i>	<i>anstim</i>	<i>sunum</i>
<i>acc.</i>	<i>gastins</i>	<i>anstins</i>	<i>sununs</i>

*n-*stems (*guma* [masc.] “man,” *hairto* [neut.] “heart,” *namo* [neut.] “name,” *qino* [fem.] “woman,” *managei* [fem.] “multitude”):

<i>Sg. nom.</i>	<i>guma</i>	<i>hairto</i>	<i>namo</i>	<i>qino</i>	<i>managei</i>
<i>voc.</i>	<i>guma</i>	<i>hairto</i>	<i>namo</i>	<i>qino</i>	<i>managei</i>
<i>gen.</i>	<i>gumins</i>	<i>hairtins</i>	<i>namins</i>	<i>qinons</i>	<i>manageins</i>
<i>dat.</i>	<i>gumin</i>	<i>hairtin</i>	<i>namin</i>	<i>qinon</i>	<i>managein</i>
<i>acc.</i>	<i>guman</i>	<i>hairto</i>	<i>namo</i>	<i>qinon</i>	<i>managein</i>
<i>Pl. nom.</i>	<i>gumans</i>	<i>hairtona</i>	<i>namna</i>	<i>qinons</i>	<i>manageins</i>
<i>gen.</i>	<i>gumane</i>	<i>hairtane</i>	<i>namne</i>	<i>qinono</i>	<i>manageino</i>
<i>dat.</i>	<i>gumam</i>	<i>hairtam</i>	<i>namnam</i>	<i>qinom</i>	<i>manageim</i>
<i>acc.</i>	<i>gumans</i>	<i>hairtona</i>	<i>namna</i>	<i>qinons</i>	<i>manageins</i>

< \*-eh<sub>2</sub>-es; etc.). The *jō*-stems mostly follow the same pattern, but include the significant subtype represented by *mawi*, which historically contains an ablauting proterokinetic suffix \*-ī-/yā- < \*-ih<sub>2</sub>-/-yeh<sub>2</sub>- (nom. sg. -i < \*-ih<sub>2</sub>, gen. sg. -jos < \*-yeh<sub>2</sub>-s; cf. Sanskrit nom. *devī* “goddess,” gen. *devyās*; Greek nom. *τράπεζα*, gen. *τραπεζης*, see Ch. 24, §4.1.1.1).

#### 4.1.2.3 Gothic a- and ja-stems

The *a*- and *ja*-stems (continuing the Proto-Indo-European thematic stems) show greater phonetic erosion than the *ō*- and *jō*-stems, especially in the singular; thus, for example, the accusative singular in Germanic, \*-an (< PIE \*-om), was reduced to zero (Go. *dag*), while the corresponding sequence \*-(i)jan (< \*-(i)yom) was reduced to -i (*hari*, *hairdi*). In the genitive singular, Gothic -is (-jis, -eis) is a late borrowing from the pronominal declension (cf. gen. sg. *þis*, *hvis* < PIE \*tes(y)o, \*k<sup>w</sup>es(y)o); the other Germanic languages have forms pointing to \*-os(y)o.

#### 4.1.3 Nominal endings

The historical endings proper show considerable phonetic reduction in Gothic: PIE \*-es gave -s in the nom. pl. *sunjus* (< \*-ew-es); PIE \*-i (locative) gave zero in the dative singular *gumin* (< \*-en-i); PIE \*-m gave zero in the masculine and feminine accusative singular of all stem-classes.

The endings of the dative plural and genitive plural call for special comment. The Gothic dative plural in -m continues the Proto-Germanic instrumental plural in \*-mi(z), which has close counterparts in Baltic (Lithuanian -mi) and Slavic (Old Church Slavic -mi), but contrasts with forms in \*-b<sup>h</sup>i(s) in the other Indo-European languages. The origin of the masculine and neuter genitive plural in -e is a mystery. Most feminines form their genitive plural in -o < \*-ōn < \*-ōm, and \*-ōm is the ending for all three genders in the other Germanic languages (cf. Old High German -o, Old Saxon -o, Old English -a, Old Icelandic -a) and elsewhere in Indo-European (cf. Latin -um, Greek -ων, etc.). The *e*-colored Gothic ending, presumably from \*-ēn, is an unexplained innovation.

#### 4.1.4 Pronouns

Demonstrative and interrogative pronouns show points of contact with *a*- and *ō*-stem nouns, but with a great many idiosyncrasies (see §4.1.4.1). Below are given the paradigms of *sa* (masc.), *so* (fem.), *þata* (neut.) “this; the” (definite article) and *hvas*, *hvo*, *hva* “who, what.” Note the existence of a special instrumental form in the interrogative.

(10)		Masc.	Fem.	Neut.	Masc.	Fem.	Neut.
	Sg. nom.	sa	so	þata	hvas	hvo	hva
	gen.	þis	þizos	þis	hvis	*hvizos	hvis
	dat.	þamma	þizai	þamma	hamma	hvizai	hamma
	acc.	þana	þo	þata	hana	hvo	hva
	instr.				(= dat.)	(= dat.)	hve
	Pl. nom.	þai	þos	þo			
	gen.	þize	þizo	þize			
	dat.	þaim	þaim	þaim			
	acc.	þans	þos	þo			

Based on these are the more emphatic demonstrative *sah*, *soh*, *þatuh* “this . . . here” and the indefinite *hazuh*, *hoh*, *hah* “each,” which consist of the forms of *sa* and *hvas* followed by

-(u)h “and” (see §3.5.5). In lieu of a separate relative pronoun, Gothic uses *sa* with the conjunction *ei* “that” (nom. *saei*, *soei*, *batei*, gen. *þizei*, *þizozei*, etc.). Other demonstratives, interrogatives, and indefinites, including *jains* “that . . . there,” *hvarjis* “which,” and *hvarjizuh* “each,” are declined as strong adjectives (see §4.1.5).

The personal pronoun of the third person is a weakened demonstrative with separate masculine, feminine, and neuter forms; the declension is similar to that of *sa* and *hvas*. The first- and second-person pronouns, on the other hand, are morphologically unique. Here and here alone in Gothic declension, there are separate dual forms.

(11)		“he”	“she”	“it”	“I”	“you”
	Sg. nom.	is	si	ita	ik	þu
	gen.	is	izos	is	meina	þeina
	dat.	imma	izai	imma	mis	þus
	acc.	ina	ija	ita	mik	þuk
	Du. nom.				wit	jut (?)
	gen.				ugkara	igqara
	dat.				ugkis	igqis
	acc.				ugkis	igqis
	Pl. nom.	eis	ijos	ija	weis	jus
	gen.	ize	izo	ize	unsara	izwara
	dat.	im	im	im	uns, unsis	izwis
	acc.	ins	ijos	ija	uns, unsis	izwis

There is also a third-person reflexive pronoun, indifferent to gender and number, with gen. *seina*, dat. *sis*, and acc. *sik*.

#### 4.1.4.1 Pronominal idiosyncrasies

Although many of the Proto-Indo-European demonstrative and interrogative pronouns also had stems in \*-o- (masculine and neuter) and \*-ā- (feminine), their declension was marked by a number of idiosyncratic features. Thus, the Gothic pronominal dative plural in *-aim* (*þaim*, etc.) shows the normal dative plural marker *-m* (see §4.1.3) added to an augmented stem form *þai-*, which otherwise surfaces without a case ending as the nominative plural masculine form. Other stem-extending elements in the Gothic pronominal system are *-mm-* < \*-zm- (dat. sg. masc./neut. *þamma*; cf. Sanskrit *tasmai*) and *-z-* (gen. sg. fem. *þizos*, dat. sg. fem. *þizai*, gen. pl. masc./neut. *þize*; cf. Sanskrit *tasyās*, *tasyai*, *teṣām*). The accusative singular masculine in *-ana* (*þana*, etc.) shows the addition of a particle *-a* < \*-ō to the old accusative in \*-n. The peculiar nominative singular forms *sa* (masc.) and *so* (fem.) go back to a defective stem *\*so-*, fused into a single paradigm with *\*to-* since Indo-European times. The use of a suppletive stem in the nominative singular of the unmarked Proto-Indo-European demonstrative recalls the contrast between *ik* versus *mik*, *mis*, *meina*, or *weis* versus *uns(is)*, *unsara* in the personal pronouns.

#### 4.1.5 Adjectives

Gothic shares with the other Germanic languages the peculiarity of declining adjectives in two ways. The *weak* declension is used with the demonstrative/article *sa*; the forms are the same as those of the masculine, feminine, and neuter *n*-stem nouns *guma*, *qino*, and *hairto* (see Table 36.2): for example, *sa blinda magus* “the blind boy,” genitive *þis blindins magaus*, etc.; *so blindo mawi* “the blind girl,” genitive *þizos blindons maujos*, etc. The *strong*

declension appears in all other environments. The endings are basically those of ordinary (j)a- and (j)ō-stems, but with a heavy admixture of pronominal forms:

(12)		<i>Masc.</i>	<i>Fem.</i>	<i>Neut.</i>
	Sg. <i>nom.</i>	blinds	blinda	blind, blindata
	<i>gen.</i>	blindis	blindaizos	blindis
	<i>dat.</i>	blindamma	blindai	blindamma
	<i>acc.</i>	blindana	blinda	blind, blindata
	Pl. <i>nom.</i>	blindai	blindos	blinda
	<i>gen.</i>	blindaize	blindaizo	blindaize
	<i>dat.</i>	blindaim	blindaim	blindaim
	<i>acc.</i>	blindans	blindos	blinda

The strong:weak distinction between adjectives is one of the most characteristic features of Germanic. The strong adjectives continue the basic type, inherited from Proto-Indo-European. Their declension, originally no different from that of (j)a-, (j)ō-, i- or u-stem nouns, was heavily influenced by the demonstrative pronouns before the breakup of Proto-Germanic. The weak adjectives, on the other hand, are a completely new category. The suffix *\*(e/o)n-* originally served to form “individualized” derived nouns of the type Latin *Cato*, gen. -*ōnis*, literally “Smarty,” or Greek Στράβων (*Strābōn*), gen. -ωνος (-*ōnos*), literally “Squint-eyes,” from o-stem adjectives (cf. *catus* “smart,” στράβός (*strabós*) “squint-eyed”). The pre-Germanic ancestor of a phrase like Gothic *sa blinda magus* thus probably once meant something like “the blind person, a boy.” But by late Proto-Germanic and Gothic, the distribution of the two types had become completely grammaticalized, the weak form being *de rigueur* after the definite article and the strong form being almost mandatory elsewhere.

In principle, most adjectives also form a comparative and a superlative. The comparative is always declined according to the weak paradigm; it is marked by a suffix *-iza* (nom. sg. masc.; fem. *-izei*, neut. *-izo*) or, less frequently, *-oza* (*-ozei*, *-ozo*). The superlative ends in *-ists* or *-osts* and is declined both strong and weak: for example, *manags* “much”: comparative *managiza*: superlative *managists*; *arms* “miserable”: *\*armoza*: *armosts*. A few common adjectives have suppletive comparative and superlative forms, e.g., *gōps* “good”: *batiza* “better”: *batists* “best”; *mikils* “large”: *maiza* “larger”: *maists* “largest.”

## 4.2 Verbal morphology

The Gothic verbal system is similar to that of the other Germanic languages, but with a number of conspicuously archaic features. In addition to the singular and plural, there are special dual forms in the first and second persons. The only tenses are the present and preterite; to express future time Gothic uses the simple present rather than a periphrastic construction like English *I will go* or German *ich werde gehen*. No purely morphological distinction is made between forms meaning “I went” and “I was going/used to go,” or between “I went” and “I have gone.” The active:passive distinction, marked periphrastically in the other early Germanic languages, is expressed in Gothic, at least in the present tense, with the aid of a special inflected passive. There are three moods – indicative, optative, and imperative; the imperative is remarkable for having third- as well as second-person forms. The nonfinite forms of the verb, consisting of an infinitive, a present active participle, and a past passive participle, conform to the Germanic standard.

### 4.2.1 Strong versus weak

As in the declensional system (see §§4.1.2, 4.1.5), most verbs can be classified as *strong* or *weak*. The terms are traditional, going back to Jakob Grimm in the early nineteenth century. (As used by Grimm, “strong” referred to vowel-stem nouns and vowel-changing verbs, while “weak” referred to consonant-stem [typically *n*-stem] nouns and consonant-suffixing verbs). Formally, verbs are distinguished as strong or weak depending on how they form their preterite and past participle. Strong verbs, which are almost always primary, are characterized by a participle in *-an(a)-* (nom. sg. masc. *-ans*) and by ablaut or reduplication (occasionally both) in the preterite. Weak verbs, typically denominative or derived from another verb, are marked everywhere outside the present by a dental suffix, normally *-d-*.

To generate the complete paradigm of a normal strong or weak verb, it is necessary to know four potentially different stem-forms, corresponding to the four *principal parts* of traditional grammars:

1. The *infinitive* (e.g., *niman* “take,” *satjan* “set”), reflecting the stem of the present indicative and optative (active and passive), and of the imperative and present participle;
2. The *first singular preterite* (e.g., *nam*, *satida*), underlying the rest of the preterite singular;
3. The *first plural preterite* (e.g., *nenum*, *satidedum*), underlying the rest of the preterite plural and dual, along with the preterite optative;
4. The *past participle* (e.g., *numans*, *satips* [stem *satida-*]).

### 4.2.2 Strong verbs

The principal parts of strong verbs fall into seven well-defined patterns or classes. The first six are characterized by ablaut:

(13)	Class	Infinitive		1st sg. pret.	1st pl. pret.	Past part.
	<i>I</i>	<i>beitan</i>	“bite”	<i>bait</i>	<i>bitum</i>	<i>bitans</i>
	<i>II</i>	<i>-biudan</i>	“offer”	<i>-baup</i>	<i>-budum</i>	<i>-budans</i>
	<i>III</i>	<i>bindan</i>	“bind”	<i>band</i>	<i>bundum</i>	<i>bundans</i>
		<i>wairpan</i>	“become”	<i>warþ</i>	<i>waurþum</i>	<i>waurþans</i>
	<i>IV</i>	<i>niman</i>	“take”	<i>nam</i>	<i>nenum</i>	<i>numans</i>
		<i>bairan</i>	“bear”	<i>bar</i>	<i>berum</i>	<i>baurans</i>
	<i>V</i>	<i>giban</i>	“give”	<i>gaf</i>	<i>gebum</i>	<i>gibans</i>
	<i>VI</i>	<i>faran</i>	“go”	<i>for</i>	<i>forum</i>	<i>farans</i>

(*wairpan*, *waurþans*, etc.; *bairan*, *baurans*, etc. show the breaking of *i* to *ai* and *u* to *au*; see §3.4.2).

Class VII is reduplicated, usually without ablaut; the reduplication vowel is *-ai-* (= short /*ɛ*/; see §3.2.1):

(14)	<i>VII</i>	<i>skaidan</i>	“separate”	<i>skaiskaiþ</i>	<i>skaiskaidum</i>	<i>skaidans</i>
		<i>aukan</i>	“increase”	<i>aiauk</i>	<i>aiaukum</i>	<i>aukans</i>
		<i>letan</i>	“let”	<i>lailot</i>	<i>lailotum</i>	<i>letans</i>
		<i>hvoþan</i>	“boast”	<i>hvaihvoþ</i>	<i>hvaihvoþum</i>	<i>hvoþans</i>

A very few strong verbs have infinitives in *-jan* or *-nan*, which affects their conjugation in the present but not in the preterite or past participle: for example, *bidjan* – *baþ* – *bedum* – *bidans*



“request”; *hafjan* – *hof* – *hofum* – *hafans* “lift”; *fraihnan* – *frah* – *frehum* – *fraihans* “ask” (note also *standan* – *stop* – *stopum*, with infixed *-n-* in the present stem).

The class membership of a given strong verb is generally predictable from the vocalism and root structure of the infinitive. Note that classes III–V are in complementary distribution: in class III the root ends in a nasal + obstruent or liquid + obstruent cluster; in class IV it ends in a single liquid or nasal; in class V it ends in a stop or fricative. Class VII includes all strong verbs with *ai*, *au*, *e* (cf. also *saian* “sow” < \**sean* [see §3.4.3], pret. *saiso*) or *o* in the infinitive.

### 4.2.3 Weak verbs

The weak verbs are likewise traditionally grouped into classes:

(15) Class	Infinitive		1st sg. pret.	1st pl. pret.	Past part.
<i>I</i>	<i>satjan</i>	“set”	<i>satida</i>	<i>satidedum</i>	<i>satip̃s</i>
<i>II</i>	<i>salbon</i>	“anoint”	<i>salboda</i>	<i>salbodedum</i>	<i>salboþs</i>
<i>III</i>	<i>haban</i>	“have”	<i>habaida</i>	<i>habaidedum</i>	<i>habaiþs</i>
<i>IV</i>	<i>fullnan</i>	“become full”	<i>fullnoda</i>	<i>fullnodedum</i>	—

A small number of weak verbs with infinitives in *-jan*, such as *waurkjan*, pret. *waurhta* “make” and *þagkjan*, pret. *þahta* (< \**-anh-*) “think,” lack the union vowel *-i-* in the preterite and past participle. Class I weak verbs with a heavy first syllable (e.g., *hausjan* “hear”) or more than one syllable before the infinitive ending (e.g., *mikiljan* “magnify”) substitute *-ei-* for *-ji-* in the present, exactly as in *ja*-stem nouns (3rd sg. *hauseiþ*, *mikileiþ*). Class IV weak verbs in *-nan*, which are intransitive, lack past participles; their inflection is like that of *niman* in the present but like that of *salbon* in the preterite (see table 36.3). The mood sign of the optative is /i:/, which appears as *-ei-* in the preterite and contracts with the preceding stem vowel to give *-ai-* (*nimai-*, *satjai-*, etc.) or *-o-* (*salbo-*) in the present.

### 4.2.4 Preterito-presents

By far the largest class of irregular verbs are the so-called *preterito-presents* – verbs whose presents resemble strong preterites and whose synchronic preterites are weak. Given below are representative forms of *witan* “know,” *munan* “think,” *magan* “be able,” and *þaurban* “need”:

(16)	<i>Pres. indic.</i>	<i>sg.</i>	1	wait	man	mag	þarf
			2	waist	mant	magt	þarft
			3	wait	man	mag	þarf
		<i>pl.</i>	1	witum	munum	magum	þaurbum
			2	wituþ	munuþ	maguþ	þaurbuþ
			3	witun	munun	magun	þaurbun
	<i>opt. sg.</i>		2	witeis	muneis	mageis	þaurbeis
			3	witi	muni	magi	þaurbi
	<i>part.</i>			witands,	munands,	magands,	þaurbands,
				fem. -ei	fem. -ei	fem. -ei	fem. -ei
	<i>Pret. indic. sg.</i>		1	wissa	munda	mahta	þaurfta
		<i>pl.</i>	1	wissedum	mundedum	mahtedum	þaurftedum

Also irregular are *wisan* – *was* – *wesum* “be,” with a suppletive and anomalous present (sg. *im*, *is*, *ist*, pl. *sijum*, *sijuþ*, *sind*; opt. *sijai-*), and *wiljan* – *wilda* – *wildedum* “want,” which



**Table 36.3 Gothic strong and weak verb paradigms**

			Active			
Pres. indic.	sg.	1	nima	satja	salbo	haba
		2	nimis	satjis	salbos	habais
		3	nimiþ	satjiþ	salboþ	habaiþ
	du.	1	nimos	satjos	salbos	habos
		2	nimats	satjats	salbots	habats
	pl.	1	nimam	satjam	salbom	habam
		2	nimiþ	satjiþ	salboþ	habaiþ
		3	nimand	satjand	salbond	haband
Pres. opt.	sg.	1	nimau	satjau	salbo	habau
		2	nimais	satjais	salbos	habais
		3	nimai	satjai	salbo	habai
	du.	1	nimaiwa	satjaiwa	salbowa (?)	habaiwa
		2	nimats	satjats	salbots	habats
	pl.	1	nimaima	satjaima	salboma	habaima
		2	nimaiþ	satjaiþ	salboþ	habaiþ
		3	nimaina	satjaina	salbona	habaina
Pres. impv.	sg.	2	nim	satei	salbo	habai
		3	nimadau	satjadau	salbodau	habadau
	du.	2	nimats	satjats	salbots	habats
		pl.	1	nimam	satjam	salbom
	2		nimiþ	satjiþ	salboþ	habaiþ
3	nimandau	satjandau	salbondau	habandau		
Pres. part.			nimands, f. -ei	satjands, f. -ei	salbonds, f. -ei	habands, f. -ei
Pres. inf.			niman	satjan	salbon	haban
Pret. indic.	sg.	1	nam	satida	salboda	habaida
		2	namt	satides	salbodes	habaides
		3	nam	satida	salboda	habaida
	du.	1	nemu	satidedu	salbodedu	habaiedu
		2	nemuts	satideduts	salbodeduts	habaieduts
	pl.	1	nemum	satidedum	salbodedum	habaiedum
		2	nemuþ	satideduþ	salbodeduþ	habaieduþ
		3	nemun	satidedun	salbodedun	habaiedun
Pret. opt.	sg.	1	nemjau	satidedjau	salbodedjau	habaiedjau
		2	nemeis	satidedeis	salbodedeis	habaiedeis
		3	nemi	satidedi	salbodedi	habaiededi
	du.	1	nemeiwa	satidedeiwa	salbodedeiwa	habaiedeiwa
		2	nemeits	satidedeits	salbodedeits	habaiedeits
	pl.	1	nemeima	satidedeima	salbodedeima	habaiedeima
		2	nemeiþ	satidedeiþ	salbodedeiþ	habaiedeiþ
		3	nemeina	satidedeina	salbodedeina	habaiedeina
Passive						
Pres. indic.	sg.	1	nimada	satjada	salboda	habada
		2	nimaza	satjaza	salboza	habaza
		3	nimada	satjada	salboda	habada
	pl.	1–3	nimanda	satjanda	salbonda	habanda
Pres. opt.	sg.	1	nimaidau	satjaidau	salbodau	habaidau
		2	nimaizau	satjaizau	salbozau	habaizau
		3	nimaidau	satjaidau	salbodau	habaidau
	pl.	1–3	nimaindau	satjaindau	salbondau	habandau
Pres. part.			numans, fem. -a	satips, fem. -da	salboþs, fem. -da	habaiþs, fem. -da

inflects in the present like a preterite optative (*wiljau*, *wileis*, etc.). Note, too, the irregular preterite *iddja*, pl. *iddjedum*, suppletive *gaggan* “go.”

#### 4.2.5 Verb endings

The inflection of the individual moods and tenses in Gothic conforms closely to what would be expected in an archaic Germanic language. In the present system, both strong and (class I) weak verbs preserve the inherited distribution of the thematic vowel (*-i-* in *nimis*, *nimiþ*; *-a-* in *nimam*, *nimand*, part. *nimands*; *-a* < *\*-ō* (< *\*-o-h<sub>2</sub>*) in 1st sg. *nima*). The only athematic present to survive in Gothic was the verb meaning “to be,” which preserves a trace of the athematic ending *\*-mi* in the first singular form *im* (on Indo-European thematic and athematic morphology see ch. 17, §3.4). The optative of an athematic present underlies the paradigm of *wiljan* (see §4.2.4).

The verb endings themselves are well anchored in Indo-European comparative grammar, including those of the present optative, which differ in part from the terminations of the indicative (e.g., 1st sg. *nimau* < *\*-oih<sub>1</sub>-m*, 3rd sg. *nimai* < *\*-oih<sub>1</sub>-t*, with the Proto-Indo-European *secondary* endings). In the other Gothic modal category, the imperative (no trace of the Indo-European subjunctive survives in Gothic), the second singular and second plural go back to well-established preforms in *\*-e* and *\*-ete*, while the third-person forms in *-adau* and *-andau* have close, though not exact, counterparts in Sanskrit and Hittite. The special passive forms *nimada* (3rd sg., extended to the 1st sg.), *nimaza* (2nd sg.), and *nimanda* (3rd pl., extended to the 1st, 2nd pl.) continue earlier middles in *\*-toi*, *\*-soi*, and *\*-ntoi*, with exact equivalents in Greek and Sanskrit. A significant innovation of the passive in Gothic and Germanic was the generalization of the *a*-colored variant of the thematic vowel throughout the paradigm.

All preterites are inflected alike outside the indicative singular. The plural (and dual) endings contain the vowel *-u-*, which arose by regular sound change in the third plural (*-un* < *\*-nt*) and was morphologically extended as a union vowel. In the singular, strong preterites and preterito-presents have the reduced endings of the Proto-Indo-European perfect (1st sg. *\*-a* (< *\*-h<sub>2</sub>a*), 2nd sg. *\*-t(h)a* (< *\*-th<sub>2</sub>a*), 3rd sg. *\*-e*). The singular of the weak preterite has special endings, of which only the first-person form in *\*(d)ōn* is wholly uncontroversial.

#### 4.2.6 Diachrony of the Gothic verb

The Gothic verbal system retains a number of significant archaisms vis-à-vis the other Germanic languages, such as the inflected passive, the third-person imperative, and the special dual forms of the first and second person. Yet in comparison with the Indo-European parent language, Gothic shares the characteristic Germanic features of *reduction* and *regularization*: reduction in the number of grammatical categories, and regularization in the number of ways that these categories can be expressed.

##### 4.2.6.1 Tense-aspect

The Proto-Indo-European tense-aspect system included three preterite-like formations: (i) the *imperfect*, built to the present stem and sharing its imperfective (iterative, durative, etc.) nuance; (ii) the *aorist*, formed from a distinct stem and denoting a punctual action or process; and (iii) the *perfect*, likewise formed from its own stem and properly denoting the state resulting from a process. Proto-Germanic reduced this system more drastically than most of the other early Indo-European languages, completely eliminating the imperfect and aorist and converting the perfect into a simple preterite.

## 4.2.6.2 Strong verbs

The past tense which arose from the Indo-European perfect was the Germanic and Gothic *strong preterite*, which betrays many traces of its origin. The perfect in Proto-Indo-European was characterized by reduplication with *\*-e-*, special endings, and *o* : zero ablaut; the accent was on the *o*-grade root in the indicative singular and on the endings elsewhere. In general, Germanic gave up reduplication in verbs where ablaut was preserved, but retained reduplication in the minority of cases where ablaut distinctions were impossible. The strong preterites of classes I–III illustrate the typical treatment:

(17) Class	PIE (sg./pl.)	Germanic	Gothic
I	*b <sup>h</sup> eb <sup>h</sup> oid-/*b <sup>h</sup> eb <sup>h</sup> id-'	*bait-/*bit-	bait/bitum
II	*b <sup>h</sup> eb <sup>h</sup> oud <sup>h</sup> -/*b <sup>h</sup> eb <sup>h</sup> ud <sup>h</sup> -'	*baud-/*bud-	baup/budum
III	*b <sup>h</sup> eb <sup>h</sup> ond <sup>h</sup> -/*b <sup>h</sup> eb <sup>h</sup> nd <sup>h</sup> -'	*band-/*bund-	band/bundum
	*wewórt-/*wewrt-'	*warþ-/*wurd-	warþ/waurþum

There is a complication in classes IV (*niman*, *bairan*) and V (*giban*), where the singular has the regular *o*-grade (*nam*, *bar*, *gaf* < \*(*ne*)nóm-, \*(*b<sup>h</sup>e*)b<sup>h</sup>ór-, \*(*g<sup>h</sup>e*)g<sup>h</sup>ób<sup>h</sup>-), but the plural, which would have been inconvenient or unpronounceable with the expected zero-grade (\**nnum*, \**brum*, \**gbum*), inserts an *\*-æ-* of uncertain origin (*nemum*, *berum*, *gebum*). Class VI is deviant; the nucleus consists of verbs which had Proto-Indo-European *\*-a-* in the present and made their perfects by lengthening *\*-a-* to *\*-ā-* (cf. Go. *skaban* “scrape,” pret. *skof*, *skobum*, matching Lat. *scabō* “scratch,” perf. *scābī*). Class VII, with retained reduplication, is largely composed of verbs which were incapable of ablaut, or whose vocalism in the perfect fell together with their vocalism in the present (*skaidan* – *skaiskaip*, *aukan* – *aiauk*, etc.). Ablaut and reduplication aside, a peculiarity of the strong preterite in Gothic is the elimination of inherited *grammatischer Wechsel* (see §3.5.1) between singular and plural. Note the contrast between, on the one hand, Gothic *warþ* – *waurþum*, with *-þ-* in both singular and plural, and, on the other, Old English *wearþ* – *wurdon*, with etymological *\*-ð-* in the plural.

The regularization and regimentation characteristic of the preterite are equally typical of the present (and of the derived present infinitive, which continues a Proto-Indo-European verbal noun in *\*-ono-*; Go. *bairan* = Skt. *bhāraṇam* “(act of) carrying”). Of the numerous ways that roots could form presents in Proto-Indo-European, one was greatly extended at the expense of the others in Germanic – the *primary thematic* type, marked by accented *e*-grade of the root and the suffix-like thematic vowel *\*-e/o-* (*\*-e-* before obstruents, *\*-o-* elsewhere). Thus, the standardly cited examples *beitan* (< \**b<sup>h</sup>éide/o-*), *biudan* (< \**b<sup>h</sup>éud<sup>h</sup>e/o-*), *bindan* (< \**b<sup>h</sup>énd<sup>h</sup>e/o-*), *niman* (< \**néme/o-*), and *giban* (< \**g<sup>h</sup>éb<sup>h</sup>e/o-*) all go back to *e*-grade thematic preforms; the comparative evidence, however, indicates that at least \**b<sup>h</sup>eid-* “split” and \**b<sup>h</sup>eud<sup>h</sup>-* “awake” formed their presents differently in Proto-Indo-European (cf. Lat. *fi-n-dō*, Skt. *budh-ya-te*). In classes I–V the monotony of the usual pattern is broken only by a handful of old *ye/o-* and *ne/o-* presents like *bidjan* and *fraiḥnan* (see §4.2.2). Even the more seriously aberrant classes VI and VII, consisting of inherited *o*-grade presents (e.g., *faran*) and verbs with inherent *a*-vocalism (*skaban*, etc.), have been considerably normalized.

The past participle of strong verbs goes back to a zero-grade verbal adjective in *\*-ana-* < *\*-onó-*, which was generalized at the expense of the competing participial suffix *\*-tó-*. Classes I–III thus show the same vocalism in the participle as in the preterite plural (*bitans*, *-budans*, *bundans*, *waurþans*). In classes IV and V, where the vocalism of the preterite plural is an innovation (Go. *nemum*, *gebum*, etc.), the vowel of the participle is secondary as well (*numans*, *gibans*). The pattern of the non-ablauting verbs of class VII, which have the same

vowel in the participle as in the present (*skaiþans*, *haitans*, etc.), was copied in class VI (*farans*).

#### 4.2.6.3 Weak verbs

The two most important classes of weak verbs, represented by *satjan* (class I) and *salbon* (class II), go back to Proto-Indo-European presents in *\*-eye/o-* and *\*-āye/o-* (earlier *\*-eh<sub>2</sub>ye/o-*), respectively. The suffix *\*-eye/o-* made causatives and denominatives in the parent language; typical Gothic reflexes are *satjan* itself (< *\*sod-éye/o-*) and *fulljan* “fill (tr.)” Proto-Indo-European *\*-āye/o-* made both denominatives like *salbon* itself (< *salba* “unguent”) and iteratives of the type *hvarbon* “walk back and forth” (< *hvaiban* “walk”).

Since derived verbs had no perfects in Proto-Indo-European, they lacked ablauting or reduplicated preterites in Germanic. New preterites were therefore needed, and these were of a characteristic innovated type, marked by an added dental element. The origin of this formation, the *weak preterite*, is the most widely discussed morphological problem in Germanic. Although there is no solution that is generally agreed upon, many arguments favor the old view that the weak preterite goes back to a periphrastic formation involving the verb “to do” (Gmc. *\*dōn*, pret. *\*ded-/dād-*). Particularly striking is the resemblance of the Gothic plural forms in *-dedum*, *-deduþ*, *-dedun* to the Old High German free-standing preterite plural *tātum*, *tātut*, *tātun* “we, you, they did.” The “long” endings *-dedum*, *-deduþ*, and so forth are a Gothic specialty; the other Germanic languages simplified *\*-dād-* to *\*-d-* under the influence of the singular.

The *\*-da-* of the weak past participle goes back to PIE *\*-tó-*, which was favored over *\*-ana-* < *\*-ono-* because of its resemblance – probably originally accidental – to the preterite marker *\*-d(ād)-*. The vowel that preceded the participial suffix was extracted from the stem of the (pre-Germanic) present: class I presents in *\*-eye/o-* were given participles in *\*-e-tó-* (Go. *satipþ* < *\*satidaz* < *\*sod-e-tó-*) and class II presents in *\*-āye/o-* were given participles in *\*-ā-tó-* (Go. *salboþs* < *\*salbōdaz* < *\*solp-ā-tó-*). The pattern of employing *\*-e-* (> Gmc. *\*-i-*) and *\*-ā-* (> Gmc. *\*-ō-*) as “linking vowels” before the dental of the participle eventually became characteristic of the preterite proper as well (cf. Go. *satida*, *satidedum* and *salboda*, *salbodedum*).

The stage was thus set for two further developments:

1. The weak verbs of class III, which were marked by an etymologically obscure diphthong *\*-ai-* in some of their present forms (cf. Go. *habaiþ* “has”), extended this element to the preterite and past participle (cf. Go. *habaiþ* – *habaida* – *habaiþs*).
2. The preterito-presents (see §4.2.4) – old stative perfects that escaped the normal Germanic development of the perfect to a preterite – were provided with weak preterites based on their inherited participles in *\*-tó-* (cf. Go. *witan*, part. *\*wissa-* (< *\*wid-tó-*), pret. *wissa*; *þaurban*, part. *þaurfts*, pret. *þaurfta*).

### 4.3 Adverbs

Gothic adverbs are productively made from adjectives by means of the suffixes *-ba*, of obscure origin (e.g., *bairhtaba* “brightly” from *bairhts* “bright”) and *-o*, historically the ending of the *a*-stem ablative singular (e.g., *galeiko* “similarly” from *galeiks* “similar”). Adverbs of location are commonly associated in semantically related groups, as, for example, *þar* – *þadei* – *þapro* “there” – “thither” – “thence”; *inna* – *inn* – *innaþro*, *innana* “within” – “to within” – “from within.” Like adjectives, adverbs can have comparatives and superlatives;

the comparative form ends in *-is* (e.g., *airis* “earlier,” *hauhis* “higher”), showing a more archaic variant of the suffix (from PIE *\*-yes-/-yos-/-is-*) than the *n*-extended form found in adjectives (see §4.1.5).

#### 4.4 Numerals

The *numerals* in Gothic present a characteristic mixture of inflected and invariant forms. The numbers from 1 (*ains*) to 3 (*\*preis*) are adjectives with masculine, feminine, and neuter forms; 2 (*twai*) has the notable genitive form *twaddje* (< *\*twajj-*), apparently the replacement of an old genitive dual. From 4 (*fidwor*) onwards there are no gender distinctions and only optional inflection for case. Noteworthy among the higher numerals are the decades from 20 to 60, which incorporate the *u*-stem noun *tigus* (cf. *taihun* “10”) “a tenfold” (e.g., *twai tigus* “20,” etc.). Both 100 (*hund*) and 1,000 (*þusundi*) are nouns.

### 5. SYNTAX

#### 5.1 Syntax and the Greek text

Because almost the whole Gothic corpus is a literal translation from the Greek, it is extremely difficult to tell how much of Wulfila’s syntax is authentically Gothic and how much is Greek in Gothic disguise. Thus, for example, the supposed dative absolute construction seen in the recurrent phrase (*at*) *andanahtja waurþanamma* “when evening had come on” has often been dismissed as artificial because the dative absolute in Gothic invariably translates a similar construction – the genitive absolute – in Greek (οφίως γενομένης).

Relatively safe conclusions can be drawn, on the other hand, about the placement of enclitic particles and pronouns, which frequently pattern quite differently in the two languages. In Mark 8.23, for example, where the Greek reads

- (18) ἐπηρώτα      αὐτόν    εἰ    τι      βλέπει  
          he was asking   him    if    anything   he sees  
          “He asked him whether he saw anything”

the Gothic has

- (19) *frah ina ga-u-hva-sehvi*

with both the question particle *-u* (here = “whether”) and the indefinite/interrogative pronoun *hva* (here = “anything”) infixed into the compound verb *ga-saihvān* “see” (perfective). Such *tmesis*, or “cutting,” of a compound is an Indo-European feature that was lost from New Testament Greek, but remains fairly common in Gothic, especially when the inserted element is *-uh* “and” (cf. *uz-uh-hof* “and he raised” < *us-hafjan* “raise”).

#### 5.2 Word order

Larger-scale questions about word order are harder to answer. The best evidence comes from cases where a word-for-word translation was simply impossible. Thus, in II Timothy 3.12, the Greek mediopassive verb διωχθήσονται “they will suffer persecution” could only be rendered by a two-word sequence in Gothic, with separate words for “will suffer” (*winnand*) and “persecution” (*wrakos*). Here and in similar cases, Wulfila put the object before the verb

(*wrakos winnand*); when the object was a pronoun, on the other hand, he put the verb first (cf. Matthew 27.5 ἀπ'ἑξῆς “he hanged himself,” rendered *ushaiḥah sik* in Gothic). Occasional details like these, gleaned from a minute comparison of the Greek and Gothic texts, provide our safest points of reference for the study of Gothic syntax.

### 5.3 Prepositions

Gothic has a full complement of *prepositions*, some of which govern the dative (e.g., *mih* “with,” *us* “out of,” *fram* “from”), some the accusative (e.g., *faur* “for,” *and* “along,” *pairh* “through”), and some more than one case, including the genitive (e.g., *ana* “at” [+ dat.], “to” [+ acc.]; *in* “in” [+ dat.], “into” [+ acc.], “on account of” [+ gen.]).

As in most early Indo-European languages, the inventory of prepositions overlaps considerably with the set of *preverbs* – preposition-like elements optionally prefixed to verbs to form compounds (e.g., *ana-biudan* “command,” *faur-biudan* “forbid”; *af-niman* “take away,” *and-niman* “receive”). Although prepositions and preverbs can be traced historically to a single category, the two are synchronically quite distinct in Gothic; thus, for example, the common preverbs *fra-* (sometimes meaning “away, forth”) and *ga-* (sometimes meaning “together” and sometimes merely perfectivizing) lack prepositional counterparts. As in the oldest Greek and Sanskrit, verbal compounds in Gothic sometimes display *tmesis* – the interposition of a restricted range of words and particles between the verb and prefix: for example, *ga-u-hva-selvi* “whether he might have seen anything” (*ga-saitvan* “see” [perfective], *-u* = question particle, *hva* = indefinite/interrogative pronoun); *uz-uh-hof* “and he raised” (*us-haffjan* “raise,” *-uh* “and”). Phrase-internal facts like these are among our safest points of reference for the study of Gothic syntax.

### 5.4 Conjunctions

Gothic retains the inherited enclitic *-(u)h* (PIE *\*-k<sup>w</sup>e*) “and”; the normal free-standing word for “and” is *jah* (< *\*yo-k<sup>w</sup>e*), with cognates elsewhere in Germanic. The ubiquitous subordinating conjunction is *ei*, which in isolation introduces purpose clauses and which combines with other words to form complex conjunctions of the type *patei* “that,” *akei* “but,” *faurpizei* “before,” *mihþanei* “while,” and so forth. Other common conjunctions include *aiphau* “or,” *auk* “for,” *ip* “but,” and *unte* “until,” *swe* “as,” and *pau* “than,” all inherited or composed of inherited materials.

### Note

1. Germanic belongs to the *centum* division of IE languages, in which the PIE “palatals” *\*k̑, \*g̑, \*g̑h* and the less common “velars” *\*k, \*g, \*gh* fell together into a single velar series.

### Bibliography

- Bammesberger, A. 1986–1990. *Untersuchungen zur vergleichenden Grammatik der germanischen Sprachen*. Heidelberg: Carl Winter.
- Bennett, W. 1980. *An Introduction to the Gothic Language*. New York: Modern Language Association of America.
- Braune, W. and E. Ebbinghaus. 1981. *Gotische Grammatik* (19th edition). Tübingen: Niemeyer.
- Jasanoff, J. 1994. “Germanic.” In F. Bader (ed.), *Langues indo-européennes*. Paris: CNRS Editions.
- Krahe, H. and E. Seebold. 1967. *Historische Laut- und Formenlehre des Gotischen* (2nd edition). Heidelberg: Carl Winter.

- Krahe, H. and W. Meid. 1967–1969. *Germanische Sprachwissenschaft* (7th edition). Berlin: De Gruyter.
- Krause, W. 1968. *Handbuch des Gotischen* (3rd edition). Munich: Beck.
- Lehmann, W. 1986. *A Gothic Etymological Dictionary* (based on third edition of S. Feist, *Vergleichendes Wörterbuch der gotischen Sprache*). Leiden: Brill.
- Mossé, F. 1956. *Manuel de la langue gotique: grammaire, textes, notes, glossaire* (2nd edition). Paris: Aubier.
- Robinson, O. 1992. *Old English and its Closest Relatives*. Stanford: Stanford University Press.
- Seebold, E. 1970. *Vergleichendes und etymologisches Wörterbuch der germanischen starken Verben*. The Hague/Paris: Mouton.
- Streitberg, W. 1965. *Die gotische Bibel. I. Teil: Text* (5th edition). Heidelberg: Carl Winter.

# Ancient Nordic

JAN TERJE FAARLUND

## 1. HISTORICAL AND CULTURAL CONTEXTS

Germanic languages prior to AD 500 are attested in two major types of documents, the Gothic Bible translation and runic inscriptions. The bulk of the runic inscriptions are in a language different from Gothic. Most of them are found in Scandinavia, but there is some controversy as to whether the language represents a common Northwest Germanic stage or a separate North Germanic variety (see §1.2). Without further implications and without prejudice in favor of one or the other view, I will henceforth refer to this language as *Ancient Nordic*.

### 1.1 Prehistory

There is considerable controversy over the absolute chronology of the Indo-European settlement of Northern Europe and of the development of a separate branch of Germanic languages. But most archeologists and historical linguists seem to have reached the consensus that southern Scandinavia and northern Germany were inhabited by speakers of an Indo-European language by the beginning of the third millennium BC (Østmo 1996), and that a distinct branch of Indo-European had evolved by c. 500 BC. From this region the Germanic-speaking people spread north into Sweden and Norway and south into the European continent.

The Germanic area was never politically unified; there has never been a Germanic nation (Haugen 1976:100). The Germanic-speaking people were farmers and cattle-herders organized in loosely knit bands of extended families and clans. During the late Roman period, pre-Christian Scandinavia was a stable society with a strict social hierarchy. Marriage, funerals, and inheritance were conducted according to fixed laws and regulations (Grønvik 1981).

The earliest known group to have left the Germanic homeland was that of the Goths, who moved south and east, their dialect(s) becoming the East Germanic group of languages, of which the Gothic language of Wulfila's Bible translation is the best-known and most completely attested variety (see Ch. 36). After the departure of the Goths, the other Germanic tribes stayed in contact for some hundred years still, until the dialects spoken on the continent (West Germanic) began to develop features that would separate them from the more conservative dialects spoken in Scandinavia (North Germanic).

### 1.2 North or Northwest Germanic?

As for the actual identity of the language of the runic inscriptions, four main views can be identified in the literature:



1. Ottar Grønvik argues, mainly on the basis of the development of the vowel systems, that North and West Germanic must have split off from each other during the first couple of centuries of our era. Since the inscriptions are Scandinavian, the language is distinctly North Germanic.
2. Hans Kuhn (followed by Haugen and others) finds that the runic language also has so many “Western” features that it is most probably the common ancestor of North and West Germanic. According to this view, the Northwest Germanic unity was maintained until the Anglo-Saxon settlement of England in the fifth century.
3. Elmer Antonsen agrees with Grønvik that the split between North and West Germanic took place before AD 500 (200–300 according to Antonsen), but asserts that the split consisted only in innovations in West Germanic. The Scandinavian dialects maintained the archaic form of the common parent language, which is what we find in the runic inscriptions.
4. E. A. Makaev (followed by Krause and Kufner) considers the runic inscriptions to have been written in some kind of *koine*, a common ritual pan-Germanic language.

### 1.3 Language variation

Many scholars have remarked on the homogeneity of the language of the inscriptions, and it is this homogeneity which has led to the theory of a *koine* (see §1.2). The chief problem with the *koine* scenario is the absence of a unifying social and political organization that would support scribal education and language codification. Moreover, it seems that the linguistic homogeneity of the inscriptions may simply be due to a common geographic origin (Southern Scandinavia).

On closer inspection, however, the language may not be as uniform as previously assumed. The number of securely interpreted forms is very limited, and there may well have been dialect differences between, for instance, East and West Scandinavian, as well as historical differences, that are not reflected in the attested material. In addition, it must be kept in mind that part of our assumed knowledge of Ancient Nordic comes from reconstruction based on other Indo-European languages and younger stages of Germanic. In many cases the results of this reconstruction have favored certain readings over others. This has no doubt made the language appear more uniform than it actually is. There are no securely interpreted forms in the total body of inscriptions that would preclude a certain amount of dialect variation. There is, in other words, no reason to assume that the rune carvers did not write on the basis of their own spoken language.

### 1.4 The documents

All of the extant material in Ancient Nordic consists of inscriptions in the older runic alphabet (the *older futhark*). None of the inscriptions refer directly to historical persons or events, therefore an absolute chronology based on the linguistic documents alone is impossible. The dating of the inscriptions is partly based on archeological findings, and partly on relative chronology of linguistic forms. The oldest inscriptions can be dated to the end of the second century AD; towards the eighth century the older futhark was replaced by the *younger futhark* and eventually by the Latin alphabet. Standard corpora of inscriptions in the older futhark (Krause and Jankuhn 1966, Krause 1971, Antonsen 1975) consist of some 120–130 items. Of these, between 100 and 105 (depending on the dating and interpretation) can be said to be written in Ancient Northwest or North Germanic. The rest either belong to a later stage of the language (sixth and seventh centuries), or have a distinctly East Germanic (Gothic) form.

Through the entire period, inscriptions were made on movable artifacts such as spear-heads, arrow shafts, swords, shields, combs, buckles, clasps, and rings. From the last part of the period we have bracteates, a kind of gold medallion, with inscriptions. From the fourth century on, there are inscriptions on stone, usually gravestones and memorial monuments. This custom seems to have originated in Norway and spread to Sweden and Denmark. No inscription on stone in the older runic alphabet has been discovered outside of Scandinavia.

All of the inscriptions are short, varying from a single rune to the five-line inscription of fifteen words on the Tune stone. The content may be a short description (one word) of the object carrying the inscription, or of the owner. The stone carvings usually contain the name of the person commemorated, or the name of the person who erected the stone, or both, often in the form of a complete sentence or phrase. Some inscriptions seem to have a metrical form.

Many of the inscriptions are uninterpretable. Some contain just a few runes, which, although identifiable, do not make sense. Others may be longer, but contain so many unclear runes that an interpretation hardly amounts to more than guesswork.

## 1.5 Corpus and transliteration

The present survey of Ancient Nordic is based on a corpus consisting of the runic inscriptions from *c.* AD 500 and earlier. Those inscriptions which runologists have not been able to interpret are omitted from my corpus, as are those which have engendered widely differing interpretations by experts. For the remaining inscriptions, I have followed accepted readings as presented by Krause (1971) and Antonsen (1975).

By convention, runes are transliterated by boldface lower case letters. This has been done in the present work mainly in the phonology section, where the original spelling is relevant. In the morphology and syntax sections, Ancient Nordic forms are printed in italics. Vowel length is not indicated in the runic alphabet (see §3.1). In forms given in italics below, vowel length will be indicated (by a macron) only in grammatical morphemes and only in the morphology section. Although proper names often have a transparent meaning, they are generally not glossed, but their gender is indicated as PNm (masculine) or PNf (feminine).

An Ancient Nordic inscription is traditionally identified by the name of the place where it is found. This name is given in parentheses after each cited form.

## 2. WRITING SYSTEM

### 2.1 The runes and the futhark

The symbols used to write Ancient Nordic are called *runes*. There are twenty-four runes, at least twenty-two of them representing phonemes of the language. The runes were organized in a specific order, like an alphabet; such a runic alphabet is called a *futhark*, from the values of the first six runes. Although there was some individual variation, the futhark was remarkably uniform throughout the area and through the four centuries of use.

**Table 37.1 The Northwest Germanic futhark**

ᚠ	ᚢ	ᚦ	ᚨ	ᚫ	ᚭ	ᚱ	ᚴ	ᚷ	ᚹ	ᚻ	ᚾ	ᚿ	ᛀ	ᛃ	ᛆ	ᛈ	ᛊ	ᛋ	ᛏ	ᛒ	ᛖ	ᛗ	ᛙ	ᛚ	ᛞ	ᛟ	ᛠ	ᛡ
f	u	þ	a	r	k	g	w	h	n	i	j	p	è	z	s	t	b	e	m	l	ŋ	d	o					

The order of the runes is known from several inscriptions containing the full list. Their value can be deduced from their use in identifiable words, and from their correspondence with letters in the Mediterranean alphabets. In addition, each rune has its own name, beginning with the sound that it represents. The twenty-four runes are organized into three groups of eight runes each. The groups are called *ættir* (sg. *ætt* “family,” or the word may also be related to *átta* “eight”).

There is a close correspondence between what may be assumed to be the phonetic value of the runes and the reconstructed phonological system of the language. The only real uncertainty resides in 𐌆, which probably represents a long, low unrounded vowel, contrasting in Proto-Germanic with a long, low rounded vowel (Antonsen 1975:2f.). This is a contrast that does not exist in the short vowel system of Proto-Germanic, where /a/ is the only low vowel. The rune eventually became superfluous through phonological development, which explains why it is found almost only in the futharks, and hardly in any complete word (with one possible exception). One other rune which may not have represented a separate phoneme is 𐌛.

The reflex of Germanic /z/ (from /s/ by Verner's Law) is written 𐌿. This letter was earlier considered to represent a palatalized /t̪/, since it later merged with /r/. It could not be /z/, it was assumed, since it did not undergo final devoicing (as its Gothic equivalent did: Gothic *dags* “day” vs. Old Norse *dagr*). But since there is no other reason to posit a transitional stage between /z/ and /r/, we will follow Antonsen (1975), among others, in transcribing it <z> and considering it a voiced sibilant.

The writing is usually from left to right, but the opposite direction and bidirectional writing (*boustrophedon*) are also used. Words are usually not spaced.

## 2.2 Origin

The futhark is a phonologically based writing system of the same type as the Greek and Latin alphabets. Many of the symbols have a clear Latin or Greek base, such as 𐌺, 𐌾, 𐌰, 𐌱, 𐌴, 𐌶. In addition, 𐌸 and 𐌹 can have a Latin, but not a Greek, origin. Conspicuously, runes that represent phonemes not found in Latin show no similarity to Latin or Greek letters: 𐌲, 𐌳, 𐌴, 𐌵. The most likely root of the runic script may therefore be the Latin alphabet, combined with the creativity and ingenuity of its inventor (notice that the runic script, unlike the Latin alphabet, distinguishes between /i/ and the semivowel /j/, and between /u/ and the semivowel /w/), who also found inspiration in the Greek alphabet and perhaps in North Italian writing systems.

Who the inventor was and when and where s/he lived, we of course do not know. The date of invention must be prior to AD 150, but perhaps not much earlier, since this is the earliest date of a securely identified inscription (the Meldorf Fibula from before the middle of the first century AD may contain runes; in which case the date of the first appearance of runic inscriptions has to be pushed back more than a century). On the other hand, it is not unlikely that the runes were first exclusively written on wooden objects that are now lost, as the angular shape of the runes may indicate that they were originally designed for carving in wood. Their inventor must have been a Germanic-speaking person, since the futhark is particularly well suited for representing an early Germanic phonological system. If the invention took place not too long before the earliest inscriptions, it is plausible that the locale was somewhere near the center of their greatest diffusion, namely Denmark (as claimed by Moltke [1985:64]). It is clear, however, that the runes could not have been invented by someone who did not have contact with the classical cultures of the Mediterranean. On the other hand, it is not likely that the futhark would have been invented in the immediate vicinity of the Latin or the Greek world, since in that case one could simply

have adopted the Latin or the Greek alphabet, which in fact the High Germans and Wulfila the Goth did.

### 3. PHONOLOGY

#### 3.1 Vowels

The runic alphabet contains five vowel symbols (plus the ambiguous *ē*). These correspond exactly to the Ancient Nordic vowel system with the five canonical vowels /i, u, e, o, a/. In addition there is a length contrast, which is not indicated by the runic letters, but which can be reconstructed on a comparative basis. Each short vowel except /e/ has a long counterpart. In accented syllables, reflexes of Proto-Germanic *\*e:/* have become /a:/. The vowel system of Ancient Nordic can therefore be represented thus:

(1)		<i>i</i>	<i>i:</i>	<i>u</i>	<i>u:</i>	<i>e</i>	<i>o</i>	<i>o:</i>	<i>a</i>	<i>a:</i>
	HIGH	+	+	+	+	—	—	—	—	—
	LOW	—	—	—	—	—	—	—	+	+
	ROUND	—	—	+	+	—	+	+	—	—
	LONG	—	+	—	+	—	—	+	—	+

Redundancy rule: [+ ROUND] > [+ BACK] (i.e., all rounded vowels are back vowels).

There are three diphthongs, /ai/, /au/, /iu/; in addition, a fourth attested diphthong, *eu*, is probably an allophonic variant of /iu/.

##### 3.1.1 Vowels in unaccented syllables

Ancient Nordic has already acquired the common Germanic accentual pattern, whereby the accent falls on the root syllable of words, while affixes remain unaccented. As a result of this fixed accent, Ancient Nordic has a different vowel inventory in accented and unaccented syllables: /i/ and /e/ have merged and are written *i*, and there is no short /o/ in unaccented syllables (the short /o/ in accented syllables is the result of *a*-umlaut).

Among unaccented long vowels, there is a contrast *u/o*, but the /a:/ has been fronted and is written *e*. The diphthong /ai/ is monophthongized in unaccented syllables and is also represented by *e*. There is no attestation of /au/ in unaccented syllables, but there is probably a reflex of /eu/ in *Kunimundiu* (PNm; Tjurkö).

In unaccented open final syllables of original Indo-European bisyllabic words, short vowels (except /u/) were lost prior to attested Ancient Nordic. This is shown by the first- and third-person singular preterite of strong verbs, *unnam* “undertook” (Reistad), *was* “was” (Kalleby); and by the third-person singular present form of “be”: *ist* (Vetteland).

An epenthetic vowel /a/ is sometimes inserted in consonant clusters containing a liquid: **worahto** (= *worhto* “wrought”; Tune), **harazaz** (= *Hrazaz* PNm; Eidsvåg), **harabanaz** (= *Hrabnaz* “raven,” PNm; Järsberg), **witadahalaiban** (= *witandahlaiban* “bread-ward”; Tune). This was probably a synchronic process which became nonproductive, as these forms have not been passed down to later stages of Nordic; compare Old Norse *orta*, *hrafn*. Contemporary forms without the epenthetic vowel are also found: **hrazaz** (Rö). In later inscriptions an epenthetic vowel is also used in certain other consonant clusters.

### 3.2 Semivowels

The semivowels, or glides, are /j/ and /w/. The former is sometimes written **ij**. This is always the spelling in the case of a three-moraic rhyme: **raunijaz** “tester, prober” (Øvre Stabu), **holtijaz** “son of Holt” (Gallehus), **þirbijaz** (PNm; Barmen). After one or two morae, both forms occur: **harja** (PNm; Vimose comb), **auja** “luck” (Sjælland), **bidawarijaz** (PNm; Nøvling), **gudija** “priest” (Nordhuglo).

### 3.3 Consonants

Ancient Nordic’s consonant inventory is comprised of stops, fricatives, nasals, and liquids.

#### 3.3.1 Obstruents

The runic alphabet has nine letters representing obstruents. As with vowels, this matches the phonological contrasts exactly. The obstruents (stops and fricatives, voiced and voiceless) have three contrasting points of articulation: labial, dental, and velar. Among the voiced obstruents, stops and fricatives occur as allophonic variants (each allophonic pair being spelled with the same runic symbol).

(2)		LABIAL			DENTAL			VELAR		
		<b>b</b>	<b>p</b>	<b>f</b>	<b>d</b>	<b>t</b>	<b>þ</b>	<b>g</b>	<b>k</b>	<b>h</b>
	VOICE	+	–	–	+	–	–	+	–	–
	STOP		+	–		+	–		+	–

Thus, **d** is seen to alternate with **þ** in the same morpheme in different environments: **laþodu** (Trollhättan) versus **laþoþ** (Halskov) “invitation (acc.),” where the alternating consonant is a fricative in both cases, but with voicing alternation (voiced and voiceless respectively). In summary, **b**, **d**, and **g** represent a voiced stop word-initially, after nasals, and after /l/; but a voiced fricative intervocally, after /r/, and perhaps word-finally. The **p** is very rare, and does not occur in any full word in the inscriptions from our period.

There also exists a pair of dental sibilants: unvoiced /s/ and voiced /z/. The voiced sibilant never occurs word-initially; it eventually merged with /r/.

#### 3.3.2 Sonorants

As with the obstruents, there is a series of nasals with three points of articulation: /m/, /n/, /ŋ/. The phonemic status of /ŋ/ is not quite clear; it may be an allophonic variant of /n/ before velars. In addition there occur liquids, /l/ and /r/. See also the above discussion of glides (§3.2).

## 4. MORPHOLOGY

Ancient Nordic is a typical archaic Indo-European language in that it has a rich inflexional morphology. Grammatical categories are to a large extent expressed by means of suffixation. Apart from the inherited ablaut system, there is little morphophonological variation. The complex morphophonology of younger Nordic languages is due to sound changes such as umlaut and syncope, which took place after AD 500. Ancient Nordic therefore appears to have a more agglutinative character than its descendants.

## 4.1 Nominal morphology

Nouns, adjectives, pronouns, and determiners are inflected for gender, number, and case.

### 4.1.1 Nominal stem-classes

Ancient Nordic nouns and adjectives belong to several declensional classes; the class is determined by the stem suffix (a stem consisting of a root plus [optionally] one or more suffixes, to which an ending is then attached [see below], in typical Indo-European fashion). Three stem-types can be identified: (i) vowel; (ii) vowel + *n*; (iii) zero (consonant stems). Four different vowel stems occur, *a*-, *ō*-, *i*-, and *u*-stems; and two different *n*-stems, *an*- and *ōn*-stems.

There are three genders, marked, to a degree, by the stem-vowel: *a*-stems and *an*-stems are masculine or neuter; *o*-stems and *on*-stems are feminine; *i*-stems are masculine or feminine; *u*-stems are masculine, feminine, or neuter; consonant-stems are masculine or feminine.

The stem suffix is followed by an ending indicating number and case. As in other Indo-European languages, the two categories can be expressed by a single morpheme. The number/case morpheme varies according to gender and partly according to stem-class. There is a singular/plural distinction, and at least four cases are marked: nominative, accusative, dative, and genitive. Already at the stage of Ancient Nordic, the stem-vowel and the number/case ending may have coalesced, so that the stem-vowel is not always identifiable synchronically.

No single noun or adjective is attested in all its number/case forms in the runic corpus. By comparing different words in different forms, however, it is possible to establish complete paradigms for some declensional classes. Most of the remaining lacunae can be filled in on the basis of comparison with Gothic and with later stages of Nordic and West Germanic; see Table 37.2, in which vowel length is indicated for the endings only:

Table 37.2 Ancient Nordic nominal stems				
	Nominative	Accusative	Dative	Genitive
<i>a</i> -stems: masculine				
Sg.	eril-az <sup>†</sup>	stain-a “stone”	Wodurid-ē PNm hanh-ai “horse”	Godag-as PNm
Pl.	*-ōz	*-an	*-amz/-umz	*-ō
<i>a</i> -stems: neuter				
Sg.	lin-a “linen”	horn-a “horn”	-kurn-ē “grain, corn”	*-as
Pl.	hagl-u	*-u	*-amz/-umz	*-ō
<i>o</i> -stems: feminine				
Sg.	laþ-u “summons”	run-ō “rune”	Birging-ū PNf	*-ōz
Pl.	*-ōz	runōz	*-amz/-umz	*-ō
<i>i</i> -stems: masculine and feminine				
Sg.	-gast-iz “guest”	hall-i “stone”	win-ē “friend”	ungand-īz “unbeatable”
Pl.			*-amz/-umz	*-o

(cont.)

**Table 37.2 (cont.)**

	Nominative	Accusative	Dative	Genitive
<i>u-stems: masculine and feminine</i>				
Sg.	Haukoþ-uz	mag-u	Kunimundiu	mag-ōz
	PNm	“son”	PNm	
Pl.	*-iuz	*-un	*-umz	*-ō
<i>u-stems: neuters (as above but without the nominative singular -z, thus:)</i>				
Sg.	alu			
<i>an-stems: masculine (distinct neuter forms are not attested)</i>				
Sg.	gudij-a	*-an	-hlaib-an	Keþ-an
	“priest”		“bread”	PNm
Pl.	*-niz	*-an	*-umz	arbij-ano
				“heirs”
<i>on-stems: feminine</i>				
Sg.	Bor-ō	*-ōn	*-ōn	Ingij-ōn
	PNf			PNf
Pl.	*-ōn	*-ōn	*-ōmz/-umz <sub>i</sub>	*-ōno
<i>Consonant stems: feminine</i>				
Sg.	swestar			
	“sister”			
Pl.	dohtriz			
	“daughters”			

<sup>†</sup>The word *erilaz*, which occurs in several inscriptions, has an obscure meaning. It has been suggested that it is the name of a tribe or an ethnic group, that it means “rune-master,” or that it is a proper name.

In the superlative adjective *asijostez* “dear, lovable” (Tune; see Grønkik 1981), the masculine plural nominative appears as *-ēz*, which is a specifically adjectival ending.

In a couple of inscriptions, a proper name occurs in its root form. This may be taken either as a vocative case (Krause 1971:48) or as a separate West Germanic form (Antonsen 1975:26) – nominative singular lost its ending early on in West Germanic.

Younger West Germanic dialects (Old High German, Old English) have a separate instrumental case, therefore such a case would be expected also in early Northwest Germanic, but there is no syntactic position attested in which the instrumental would be required. Consequently, we have no evidence of the possible existence of such a case form.

#### 4.1.2 Pronouns and determiners

Only personal pronouns in the first-person singular are securely attested in the corpus. The nominative occurs several times, usually in the form *ek*, but also *ik*, which may be a West Germanic form or may reflect an unaccented pronunciation. In enclitic position the forms *-eka* or *-ika* are used. The dative form *mez* is also attested.

Determiners may have adjectival endings, as the first-person possessives *minas* (masc. sg. gen.) and *minu* (fem. sg. nom.), or they may have pronominal endings, as the first-person possessive *mininō* and the demonstrative *hinō* “this,” which are both masculine singular accusative. No other determiners are securely attested.

## 4.2 Verbal morphology

### 4.2.1 Verbal stems

Though there are very few verb forms attested in the corpus, both strong and weak verbs are represented (see Ch. 36, §4.2.1). Among strong verbs, the following ablaut series and stages are attested (cf. Ch. 36, §4.2.2):

(3)	Present	Preterite singular	Participle
I.	writu “write”		
IV.		-nam “took”	
V.	gibu “give” ligi “lie”		
		was “was”	
VI.			slaginaz “slain”

The weak verbs form their preterite by adding *-d-* to the stem (plus the person/number ending). Most of the verbs that are attested in the corpus have a stem-forming suffix *-(i)j-* added to the root. This suffix appears as a vowel *-i-* when it occurs in front of the preterite marker *-d-*: *faihidō*, *tawidō*, *satidō* (cf., with no stem-vowel, *worhtō*).

### 4.2.2 Finite verbs

The finite verbs are attested in the indicative present and preterite, and in the optative present. Verbs are conjugated for three persons and two numbers. No secure second-person forms seem to be attested, and no dual forms. The person/number endings that are found are illustrated in (4):

(4)		Strong verbs		Weak verbs	
		Present	Preterite	Present	Preterite
<i>Indicative</i>					
Sg.	1.	writ-u “write”	-nam “took”	taw-ō “make”	tawid-ō
	3.				tawid-ē
Pl.	3.				dalid-un “prepared”
<i>Optative</i> <sup>†</sup>					
Sg.	2.				watē “wet”
	3.	ligi “lie”			skapi “scathe”

<sup>†</sup>These forms are all from the Strøm whetstone, the interpretation of which is rather controversial (cf. Grønvik 1996).

One verb belonging to the reduplicating class of strong verbs is attested in the first-person singular present: *haitē* “I am called” (which derives from the old middle conjugation). The verb “to be” occurs in the third-person singular indicative present, *ist*, and preterite, *was* (according to Antonsen [1975] the word *em* [1st. sg. pres. indic. of “to be”] occurs in *ek erilaz Asugisalas em* “I am Asugisala’s *erila*” [Kragehul]; but this reading is very insecure and has been challenged by Knirk [1977], among others).



### 4.2.3 Participles

The past participle of strong verbs has a root vowel from the relevant ablaut series, and the suffix *-in-* (plus nominal inflexion): *slaginaz*. The past participle of weak verbs is formed by means of the suffix *-d-* (plus nominal inflection): *hlaiwidaz* (cf. 4.3.2). The present participle is formed in *-and-* (plus nominal inflexion): *witanda-*.

## 4.3 Derivational morphology

### 4.3.1 Prefixation

The prefix *un-* is used to denote negation or absence of a quality: *Unwodiz* “calm, peaceful” (PNm; Gårdlösa), compare *wodiz* “furious, raging”; *Ungandiz* “unbeatable” (PNm; Nordhuglo).

### 4.3.2 Suffixation

Proto-Germanic had several derivational suffixes inherited from Indo-European. Some of these became unproductive before the Ancient Nordic stage and thus have been lexicalized, for example, *-s-* in *laus-* “loose” (cf. Greek λύω “I loose” and Latin *luo* “I pay, atone”). Other derivational suffixes were grammaticalized to become inflexional endings, for example, *-d-*, which formed the basis of the past participle of the weak verbs.

The following derivational suffixes seem to be more or less productive, with an identifiable meaning in Ancient Nordic:

- (5) A. *-j-*: agent nominal or patronymic, *raunijaz* “tester, prober” (Øvre Stabu), *holtijaz* “son of Holt” (Gallehus)
- B. *-ing-*: (place of) origin, *iubingaz* “from \*Yd” (Reistad)
- C. *-op-/od-*: action nominal, *laþodu* “invitation” (Trollhättan)
- D. *-san/-son-*: diminutive, *Hariso* (PNf; Himlingøje I)

## 4.4 Compounding

Despite the small size of the corpus, the Ancient Nordic material offers a large number of compounds, constructed of nouns and adjectives. The first member of the compound ends in the stem-vowel: *-a-*, *-i-*, or *-u-*.

1. *Noun + noun*. The second member is the head of the word, while the first member functions as a modifier: *walha-kurne* “Celtic corn” (i.e., “foreign gold”; Tjurkö); *widu-hundaz* “forest dog” (Himlingøje II).

2. *Adjective + noun*:

- 2A. The noun is the head: *Wodu-ride* “furious rider” (Tune); *Hagi-radaz* “giver of suitable advice” (Garbølle), from *hag-* “suitable” + *rad-* “advice” – this example could also belong to type 2C.
- 2B. The adjective is the head: *witanda-hlaiban* “bread-ward” (Tune). The first member is an adjective (present participle) derived from a verb meaning “to see to, pay attention to,” and the second member is the noun “bread.”
- 2C. Headless, or exocentric, compounds, typically *i-*-stems: *alja-markiz* “foreigner” (Kårstad), from *alj-* “other” + *mark-* “land”; *glæ-augiz* “bright-eyed” (Nebenstedt).

3. *Noun + adjective*. The second member, the adjective, is the head: *saira-widaz* “with gaping wounds” (Rö), from *sair-* “wound” + *widaz* “wide, open”; *flagda-faiknaz* “threatened by deceit” (Vetteland).

4. *Proper names*. The great majority of the nominal compounds in the corpus are proper names. Most of these were semantically transparent (which, however, does not necessarily mean that they are still interpretable), and for some (the oldest ones?), the composition is also motivated: *Woduride* “furious rider” (Tune); *Hadu-laikaz* “battle-player” (Kjølevik). Other names look more like arbitrary juxtapositions, thus several names in *-gastiz* “guest,” for example, *Hlewa-gastiz* (Gallehus), from *hlew-* “lee, protection” (Ottar Grønvik [personal communication] suggests that the apparent arbitrariness of these names is due to our lack of knowledge of the ancient society; if *hlewa-*, for instance, refers to some kind of sanctuary or temple, *Hlewagastiz* might mean “priest”).

## 5. SYNTAX

Among the inscriptions from before c. AD 500 which have been deciphered and interpreted in a sufficiently secure and noncontroversial way, it is possible to identify forty-three combinations of words that can be considered syntactic constructions (divided among thirty-one inscriptions). It goes without saying that it is impossible to present anything even remotely reminiscent of a full syntactic description of the language on the basis of this small corpus. The material should rather be seen as illustrative of certain syntactic features. None of the constructions in the corpus represents crucial counterevidence to what may be expected from an Indo-European language of this period (if it did, it should probably be taken as evidence that the inscription has been misinterpreted; for a discussion of a younger inscription from such a perspective, see Faarlund 1990:166). On the other hand, even this limited database gives us an indication as to which choices the grammar of Ancient Nordic has made among alternatives exploited differently by various Indo-European languages.

There is no example of a subordinate sentence or of sentence conjunction in the corpus.

### 5.1 Noun phrase structure

#### 5.1.1 Noun phrase word order

In the Ancient Nordic material there are twenty-seven complex noun phrases. The dominant ordering pattern is head-dependent. This is the case in all of the examples with an adjective: *Hlewagastiz holtijaz* “H. (son) of Holt” (Gallehus); *Swabaharjaz sairawidaz* “S. with gaping wounds” (Rö). In *Owlþupewaz ni wajemariz* “O. of no bad fame” (Thorsberg) the adjective is itself modified. Possessive and demonstrative determiners also follow the head noun: *magoz minas* “son mine” (Vetteland); *swestar minu* “sister mine” (Opedal); *halli hino* “stone this” (Strøm). A dependent genitive also usually follows its head: *erilaz Asugisalas* (Kragehul); *þewaz Godagas* “servant of G.” (Valsfjord); *gudija Ungandiz* “priest of U.” (Nordhuglo). In two instances, where the head noun denotes the monument bearing the inscription and the genitive the person commemorated, the genitive precedes the noun: *Ingijon hallaz* “Ingio’s stone” (Stenstad); ... *an waruz* “...’s enclosure” (Tomstad; all of the attested examples with genitive nouns or possessive determiners are consistent with an observation by Smith [1971] that animate heads require a following genitive and inanimate ones a preceding genitive; see also Antonsen 1975:24). The only quantifier attested precedes its head: *prijoz dohtriz* “three daughters” (Tune).

### 5.1.2 Apposition

By far the most commonly occurring complex noun phrases in the corpus are appositional constructions. Most of these consist of a first-person singular pronoun + a noun phrase (NP). The second member is usually a proper name or a nominalized adjective functioning as a proper name: *ek Unwodiz* (Gårdlösa) “I U.”; *mez Wage* “me W.(dat.)” (Opedal); *ek Hrazaz* (Rö). The second member can also be a complex NP: *ek Hlewagastiz holtijaz* “I H. of Holt” (Gallehus); *ek gudija Ungandiz* “I the priest of U.” (Nordhuglo). In *Woduride witandahlaiban* “W. the bread-ward” (Tune) and *Boro swestar minu* “B. my sister” (Opedal), the first member of the apposition is a proper name. There are even three-member appositions, consisting of a first-person singular pronoun + a proper name + a further identification or characterization: *ek Hagustaldaz bewaz Godagas* “I H. the servant of G.” (Valsfjord); *ek Wagigaz erilaz Agilamundon* (Rosseland).

### 5.1.3 Agreement

As can be seen from these examples, aside from dependent genitives, all dependents agree with their heads in gender, number, and case.

## 5.2 Prepositional phrase structure

The Ancient Nordic corpus preserves four instances of a preposition followed by an NP complement; no postpositions occur. Only two different prepositions are attested, *an(a)* “on” and *after* “after.” They both govern the dative case: *ana hanhai* “on horse” (Möjbro); *an walhakurne* “on Celtic corn” (Tjurkö); *after woduride witandahlaiban* “after (i.e., in commemoration of) W. the bread-ward” (Tune).

## 5.3 Verb phrase structure

### 5.3.1 Complements

The verb *haitan* “to be called” takes a predicate complement in the nominative: *Uha haite* “(I) am called U.” (Kragehul); *ek erilaz Sawilagaz hateka* “I, the *erila*, am called S.” (Lindholm).

Transitive verbs take a noun phrase in the accusative as their object: *ek Hlewagastiz holtijaz horna tawido* “I H. of Holt made the horn” (literally, “horn (acc.) made”; Gallehus); *ek erilaz runoz waritu* “I the *erila* wrote the runes” (literally, “runes (acc.) wrote”; Järsberg). In addition, prepositional phrases occur as verb complements: *ana hanhai slaginaz* “slain on the horse” (literally, “on horse slain”; Möjbro); *ek Wiwaz after Woduride witandahlaiban worhto* “I Wiwa wrought in commemoration of Wodurida” (literally, “I Wiwa after Wodurida bread-ward wrought”; Tune).

In *ek Hrazaz satido staina ana . . . r . . .* “I H. set stone (acc.) on . . .” (Rö), there is a prepositional phrase (with an illegible complement) in addition to an accusative object. And [*falh*] *Woduride staina* “dedicated the stone to W.” (literally, “dedicated Wodurida [dat.] stone [acc.]”; Tune) is a double object construction with a dative object preceding the accusative (the runes preceding **woduride** here are partly missing; Grønvik [1981] argues very convincingly for the emendation of a verb form *falh*, preterite indicative third person of *\*felhan* “to dedicate”).

The direct object is sometimes omitted when it refers to the object bearing the inscription or to the runes themselves: *Bidawarijaz talgide* “B. carved” (Nøvling); *Hagiradaz tawide* “H. made” (Garbølle).

### 5.3.2 Auxiliary verbs and passive voice

In the two occurrences of a complex verb form, the auxiliary follows the main verb (supporting an OV analysis of the language; see §5.4): *flagdafaikinaz ist* “is threatened by deceit” (Vetteland); *haitinaz was* “was called” (Kalleby).

These two sentences must be interpreted as passives. The passive auxiliary may be omitted, however, as in *ana hanhai slaginaz* “slain on the horse” (Möjbro), and *... iz hlaiwidaz þar* “... i buried here” (Amla).

There are no attested occurrences of the inflectional passive which is found in Gothic and in non-Germanic Indo-European languages (the only trace of the Indo-European middle voice is perhaps the verb *haitē* “I am called”).

## 5.4 Word order

### 5.4.1 Verb position

The examples above having a single complement – be it a predicate complement, an accusative object, or a prepositional phrase – may be taken as evidence that Ancient Nordic is a verb-final (OV) language (there are, however, no postpositions in the corpus, only prepositions; note also the predominant head-dependent order in NPs [see 5.1.1]). This is by no means surprising, since this is the order which can be reconstructed for Proto-Indo-European, and since there are traces of an underlying verb-final pattern in Old Norse.

In contrast, the two sentences above with double complements (*ek Hrazaz satido staina ana ... r ...* “I H. set stone (acc.) on ...”; and [*falh*] *Woduride staina* “dedicated the stone to W.”) appear to suggest a VO order (as do several other sentences in the corpus). It is worth noting, however, that in all the examples with a nonfinal verb, the verb is finite, and it is in first or second position. This is consistent with a rule of verb movement, shifting the finite verb into second position, as in later stages of Germanic and in all of the modern Germanic languages (except English): *ek Hagustadaz hlaiwido magu minino* “I H. buried my son” (Kjølevik).

The sentences with the verb in first position are subjectless sentences (cf. §5.4.2), except *wate halli hino horna* “wet this stone, horn!” (Strøm), where the verb is in the optative mood and perhaps fronted for emphasis.

Since we find no verb in any other position than first, second, or last, and since we find no nonfinite verb preceding its complement, it can be concluded that Ancient Nordic is V2 (verb-second) and OV (verb-final) at the same time, just like Modern German.

### 5.4.2 Subject position

There are eighteen sentences in the corpus having a finite verb and a nominative subject. In fifteen of these the subject is in first position, as in *Bidawarijaz talgide* “B. carved” (Nøvling) and *ek Hlewagastiz holtijaz horna tawido* “I H. of Holt made the horn” (Gallehus). More examples are provided by sentences cited above. In *wate halli hino horna* “wet this stone, horn!” (Strøm), the verb is in the optative and in first position. In *wurte runoz an walhakurne Heldaz Kunimundiu* “wrought runes on the Celtic corn, H. for K.” (Tjurkö), the subject has been focused and moved to the right. In *Hariuha hait-eka farawisa* “H. I am called, the travel-wise” (Sjælland), the subject is expressed as an enclitic on the verb. And in *ek erilaz Sawilagaz hait-eka* “I, the erila, am called S.” (Lindholm), the clitic repeats the subject in first position.

There is no doubt that the apparent regularity with regard to the position of the subject must be due to the homogeneous nature of the material, consisting solely of epigraphic texts. Ancient Nordic must have a rather free word order, like its relatives in Germanic and other Indo-European language groups.

### 5.5 Pro-drop?

On the basis of epigraphic material alone it is impossible to determine securely whether the language has *pro-drop* or not – that is, whether the subject can be omitted and represented by verbal inflection alone, even when it is not recoverable from the context. It is true that there is not one single occurrence of a pronoun as a subject in the corpus; all of the subject pronouns attested occur as constituents of appositional constructions (cf. §5.1.2). Moreover, we do find five occurrences of a missing subject. Four of these, however, follow immediately after other lexical material in which the subject referent is mentioned: *Hariuha haiteka farawisa gibu auja* “H. I am called, the travel-wise, give (1st per.) luck” (Sjælland). In this sentence, *gibu* is first-person present, and the subject is the same as that of *hait-*, namely *-eka* “I.” In *haitinaz was* “was called” (Kalleby), the subject can be inferred from a preceding genitive noun, *þrawijan* (PNm). In the case of *Uha haite* “I am called Uha” (Kragehul), preceding is *ek erilaz Asugisalas*. The sentence *wurte runoz an walhakurne* “wrought runes on the Celtic corn” (Tjurkö) occurs together with the two names *Heldaz Kanimundiu*, in the nominative and dative, respectively, on the same stone (Grønvik 1987:151); the subject is therefore recoverable (*Heldaz*). This leaves us with one short inscription with two words: *tawo lapodu* “make (1st per.) the invitation” (Trollhättan). Bearing in mind that this is epigraphic material, we certainly have no evidence to conclude that Ancient Nordic is a language in which subject pronouns can be freely omitted.

### 5.6 Nonverbal sentences?

Examples have already been given of deleted auxiliaries. The question is whether this is due to the epigraphic style (comparable to modern newspaper headlines – cf. “Ten killed in car crash”), or part of the regular grammar of the language (as in, e.g., Modern Russian). The question is further complicated by apparent appositional constructions consisting of two nominative NPs (cf. §5.1.2). When these stand by themselves in an inscription, they may also be read as a copular sentence with an omitted copula: *ek Unwodiz [em]* “I am Unwodi”; *ek gudija ungandiz [em]* “I am Ungandi’s priest”; and so forth.

## 6. LEXICON

The vocabulary in the Ancient Nordic inscriptions consists almost exclusively of inherited Germanic items. In the extant material there is no certain example of a word with a distinctly non-Germanic form, or a loanword from a non-Germanic language, although we know from later attestations that, for example, Celtic words had been adopted during the early Iron Age.

### Acknowledgments

I am grateful to Ottar Grønvik, Jan Ragnar Hagland, Kathy Holman, Brit Mæhlum, and Arne Torp, who have read a previous version of this chapter and given me many valuable comments and suggestions.

## Bibliography

- Antonsen, E. H. 1975. *A Concise Grammar of the Older Runic Inscriptions*. Sprachstrukturen. Reihe A. Historische Sprachstrukturen 3. Tübingen: Niemeyer.
- Bammesberger, A. 1990. *Die Morphologie des urgermanischen Nomens*. Heidelberg: Carl Winter.
- Faarlund, J. T. 1990. "Syntactic and pragmatic principles as arguments in the interpretation of runic inscriptions." In J. Fisiak (ed.), *Historical Linguistics and Philology*, pp. 165–186. Berlin/New York: Mouton de Gruyter.
- Grønvik, Ottar. 1981. *Runene på Tunesteinen*. Oslo: Universitetsforlaget.
- . 1987. *Fra Ågedal til Setre*. Oslo: Universitetsforlaget.
- . 1996. *Fra Vimose til Ødemotland. Nye studier over runeinnskrifter fra førkristen tid i Norden*. Oslo: Universitetsforlaget.
- Haugen, E. 1976. *The Scandinavian Languages. An Introduction to their History*. London: Faber and Faber.
- Knirk, J. E. 1977. Review of *A Concise Grammar of the Older Runic Inscriptions*, by E. H. Antonsen. *Maal og Minne*, pp. 172–184.
- Krause, W. 1971. *Die Sprache der urnordischen Runeninschriften*. Heidelberg: Carl Winter.
- Krause, W. and H. Jankuhn. 1966. *Die runeninschriften im älteren Futhark*. Göttingen: Vandenhoeck and Ruprecht.
- Moltke, E. 1985. *Runes and their Origin. Denmark and Elsewhere*. Copenhagen: Nationalmuseets Forlag.
- Morris, R. L. 1988. *Runic and Mediterranean Epigraphy*. NOWELE Supplement 4. Odense: Odense University Press.
- Østmo, E. 1996. *The Indo-European Question in a Norwegian Perspective. A View from the Wrong End of the Stick*. Journal of Indo-European Studies Monograph 17.
- Smith, J. R. 1971. *Word Order in the Older Germanic Dialects*. Ann Arbor: University Microfilms.
- Syrett, M. 1994. *The Unaccented Vowels of Proto-Norse*. NOWELE Supplement 11. Odense: Odense University Press.

# Classical Armenian

JAMES P. T. CLACKSON

## 1. HISTORICAL AND CULTURAL CONTEXTS

Armenian forms an independent branch of the Indo-European language family. Although Armenian was spoken in areas adjacent to those inhabited by speakers of Anatolian languages, it shares few significant linguistic features with the Anatolian subgroup of Indo-European. Its closest linguistic relatives are Greek and the Indo-Iranian subgroup. These three branches of Indo-European show shared developments in their morphology and vocabulary which are not found in other Indo-European languages: for example, the use of the augment *\*e-* to mark past tense verb forms; the use of a marker *\*-b<sup>h</sup>i (s)* for the instrumental case; and the prohibitive particle *\*mē*.

Some scholars have thought that the agreements between Armenian and Greek are sufficient to allow the reconstruction of a Helleno-Armenian subgroup of Indo-European, but their arguments are not conclusive, since it cannot be clearly proved that the agreements represent shared common innovations. Others, relying on an ancient tradition that the Armenians were a “colony of the Phrygians” (Herodotus 7.73) have tried to identify developments shared by Armenian and Phrygian, but have met with little success. Some of the phonetic developments which have been claimed for Phrygian also took place in Armenian, but all too often these sound changes rest upon very uncertain etymologies, and the close link between the languages is called into question by several well-established Phrygian forms. For example, the Phrygian form *matar* is generally taken to be a nominative singular meaning “mother,” from Proto-Indo-European *\*mātēr*; the cognate Armenian form is *mayr*. Note that *matar* shows a development of *\*ē* to *a* (found also in other Phrygian words) which is at odds with the development of *\*ē* to *i* found in Armenian.

The position of Armenian as a separate branch of Indo-European was not recognized until 1875 by Heinrich Hübschmann (Hübschmann 1875). Before that date most comparativists believed Armenian to be an Iranian language, mistakenly taking the large number of Iranian loanwords in Armenian to represent the inherited vocabulary. Their inability to isolate the “native” stratum of vocabulary in Armenian is understandable: only a small number of words are directly inherited from Proto-Indo-European, and these have undergone a series of complicated and intricate sound changes which make many forms unrecognizable.

Only a small amount of information about the prehistory of Armenian can be deduced from linguistic material. The large influx of Iranian vocabulary will be discussed more fully below. Some Iranian words may have been borrowed into Armenian as early as the sixth century BC, but the greatest period of influence was the Parthian period in the first four centuries of the Christian era. While the Armenian lexicon shows the influence of Iranian, the phonemic inventory of the language is strikingly similar to Georgian, the Kartvelian language



historically spoken in areas to the north of Armenia. It is unlikely that this situation is the result of chance, but it must result from a long period of contact between the speakers of the two languages. The morphological categories and syntax of Kartvelian languages may also have influenced Armenian. However, there is little lexical interchange between Kartvelian and Armenian, although some Iranian loanwords in Old Georgian appear to have entered the language via Armenian.

The extralinguistic facts relevant to the prehistory of the Armenian people are also obscure. Speakers of Armenian appear to have replaced an earlier population of Urartian speakers (see Ch. 5) in the mountainous region of Eastern Anatolia. The name *Armenia* first occurs in the Old Persian inscriptions at Bīsotūn dated to c. 520 BC (but note that the Armenians use the ethnonym *hay* [plural *haykʻ*] to refer to themselves). We have no record of the Armenian language before the fifth century AD. The Old Persian, Greek, and Roman sources do mention a number of prominent Armenians by name, but unfortunately the majority of these names are Iranian in origin, for example, Dādrši- (in Darius' Bīsotūn inscription), Tigranes, and Tiridates. Other names are either Urartian (Haldita- in the Bīsotūn inscription) or obscure and unknown in literate times in Armenia (Araxa- in the Bīsotūn inscription).

Armenia officially adopted Christianity in the early years of the fourth century AD (the traditional date is 301–304). Conversion to Christianity provided the impetus for the creation of an alphabet (see below) and the translation of the Bible into the Armenian language in the fifth century. The Bible translation and the historical and theological works of the fifth century provided the model for the classical language, which was the medium of educated written discourse for Armenians until the nineteenth century, and is still used as a liturgical language in the Armenian Church.

Modern Armenian consists of a large number of different local dialects, usually grouped into two principal branches, Eastern and Western. Sub- and nonliterary written material from the thirteenth-century Armenian Kingdom of Cilicia (often termed *Middle Armenian*) shows that the separation of the East and West dialect groups had already taken place at that date, and reveals the wide range of variation in the spoken language. However, the language of the Armenian Bible translation and early authors is strikingly uniform and may result from a deliberate attempt to create a standard. It seems a priori unlikely that the inhabitants of the different valleys and plains in the mountainous region of the Armenian Highlands should have spoken a uniform language in the fifth century, and some passages in classical authors can be interpreted as references to dialectal differences in the Armenian lexicon (cited by G. B. Djahukian at Greppin and Khachaturian 1986:9f.).

## 2. WRITING SYSTEM

Classical Armenian is written in an alphabet of thirty-six letters (increased to thirty-eight letters in the tenth century). The alphabet was specially created for the language and was used for no other language until recent times. The exact circumstances and date of the creation of the Armenian alphabet are not exactly known. The traditional account, given in the earliest sources, attributes the creation of writing to the saint Mesrop (also called Maštocʻ) in the early years of the fifth century (the dates AD 404 and 406–407 are frequently cited). Koriwn, contemporary and biographer of Mesrop, relates that the saint adapted a previous writing system invented by a Syrian bishop, Daniel, and this has led to speculation that an earlier script for Armenian existed, despite the complete absence of any attested remains. It is possible that Koriwn was referring to a different alphabet, such as Aramaic; pre-Christian inscriptions found in Armenia are written in Greek or Aramaic.



The earliest surviving specimens of the Armenian script are inscriptions in stone in the now ruined church of Tekor and on mosaic pavements excavated in Jerusalem. These are not dated, but art historians have been able to ascribe their contexts to between the end of the fifth and the beginning of the sixth century. The earliest Armenian manuscript of the Gospels was copied in AD 887, but there are palimpsests, manuscript fragments and a papyrus which are of an earlier date. The early examples of the script show only capital letters (termed *erkat'agir* “iron-writing” in Armenian).

The relationship of the letter-forms of the Armenian script to other scripts of the Near East has been a subject of much dispute. Many scholars now concur with the view that Mesrop used the Greek alphabet as a model. This is supported by the following observations: (i) the script is written from right to left; (ii) the order of the letters for which there are Greek correspondences follows that of the Greek alphabet; (iii) some of the letter-forms correspond to those of a cursive form of Greek, for example: *բ* <B> for *b* (compare Greek *β*); (iv) the digraph *ու* <OW> is used to represent the vowel [u] in imitation of Greek *ου* for [u]. However, it is difficult to find appropriate models for most of the letters for which there are no Greek equivalents, and Mesrop's original contribution to the formation of the alphabet should not be underestimated. The alphabet has an almost perfect one-to-one correspondence with the phonemes of Classical Armenian. Linguists working on Armenian normally use a particular transliteration system (for which see Schmitt 1972) which I will follow here.

**Table 38.1 The Armenian alphabet**

Character	Transcription	Character	Transcription
Ա	a	Ճ	č
Բ	b	Թ	m
Գ	g	Յ	y
Դ	d	Ն	n
Ե	e	Շ	š
Զ	z	Ո	o
Է	ê	Չ	č'
Ը	ə	Պ	p
Թ	t'	Ք	j
Ժ	ž	Բ	r
Ի	i	Ս	s
Լ	l	Վ	v
Խ	x	Տ	t
Վ	c	Ր	r
Կ	k	Յ	c'
Հ	h	Խ	w
Ջ	j	Փ	p'
Ղ	ł	Բ	k'

### 3. PHONOLOGY

#### 3.1 Consonants

The phonemic inventory of Classical Armenian consonants is presented in Table 38.2. Where the traditional transliteration scheme is at odds with the International Phonetic Alphabet, I have indicated the IPA equivalent in square brackets.

**Table 38.2 The consonantal phonemes of Classical Armenian**

Manner of articulation	Place of articulation						
	Labial	Dental	Alveolar	Palato-alveolar	Palatal	Velar	Glottal
Stops and affricates							
<i>Voiceless</i>	p	t	c [ts]	č [tʃ]		k	
<i>Voiced</i>	b	d	j [dz]	ǰ [dʒ]		g	
<i>Aspirate</i>	p' [pʰ]	t' [tʰ]	c' [tsʰ]	č' [tʃʰ]		k' [kʰ]	
Fricatives							
<i>Voiceless</i>						x	h
Sibilants							
<i>Voiceless</i>		s		š [ʃ]			
<i>Voiced</i>		z		ž [ʒ]			
Nasals	m	n					
Liquids			r, r̄		l	ł	
Glides	w/v				y [j]		

The phonetic interpretation of several Armenian phonemes is not clear-cut, and different explanations are possible. A notorious problem has long been the identification of the manner of articulation of the different stop/affricate series. The aspirates /p'/, /t'/, etc. are unproblematic; these sounds are usually transcribed as Greek aspirates and are used to transcribe Greek aspirates. The presence of a voiceless velar fricative phoneme /x/, distinct from /k'/, makes it clear that these Armenian consonants cannot be ascribed a fricative pronunciation. The other two series have been variously interpreted. The series /b/, /d/, etc. has voice as a distinctive feature: they are used to transcribe voiced stops in other languages and are themselves transcribed as voiced stops. Similar evidence enables us to know that the series /p/, /t/, etc. are unvoiced. The straightforward interpretation would therefore be that the three stop series were respectively aspirated, voiced, and voiceless. However, this leads to serious problems for the explanation of diachronic phonological developments, in particular for the emergence of the Modern Armenian dialects. In Modern Western Armenian, members of the Classical Armenian series /p/, /t/, etc. have become voiced obstruents, while members of the series /b/, /d/, etc. have become voiceless. A *simultaneous* diachronic development

- (1) voiced stops > voiceless stops  
voiceless stops > voiced stops

has been rightly rejected as impossible. The change would have to have been instantaneous in the dialects concerned in order for the two series not to be confused. It is therefore assumed that either one of the Classical “voiced” or “voiceless” series, or both, also had some extra feature which would allow one or both series to stand in opposition to a “plain” voiced or voiceless series. The diachronic development could therefore be as follows (taking, for the sake of illustration, the voiceless series to have an extra distinctive feature):

- (2) 

<i>Stage I</i>	<i>Stage II</i>	<i>Stage III</i>
voiced	voiceless	voiceless
voiceless + X	voiceless + X	voiced

It is not difficult to find possible features which would fit the bill. Many Modern Eastern Armenian dialects show a three-way distinction between aspirates, voiceless ejectives, and voiced obstruents. Phonetic investigation has also indicated that in some Eastern Armenian

dialects the voiced series is also aspirated. It is not clear, however, that any of the three possible systems is correct for fifth-century Armenian: (i) ejective / voiceless aspirate / voiced aspirate; (ii) voiceless / voiceless aspirate / voiced aspirate; or (iii) ejective / voiceless aspirate / voiced (see Vaux 1998:238f.)

Classical Armenian, like many Modern Armenian dialects, had two phonemically distinct varieties of *r*: /*r̄*/ is a rolled alveolar trill, and /*r*/ is an unrolled approximant. The difference between /*r̄*/ and /*r*/ is neutralized before immediately following /*n*/, where only /*r̄*/ can appear. The Armenian version of the grammatical work attributed to the Greek grammarian Dionysios Thrax lists /*r̄*/ as a double consonant, and this, together with the Armenian use of /*r̄*/ for [rr] in Iranian loanwords, has led some scholars to interpret /*r̄*/ as a geminate. However, genuine geminate consonants are extremely rare in Armenian, and it is therefore preferable to consider /*r̄*/ as an independent unit phoneme.

The phonemic opposition between /*l*/ and /*h*/, the palatal and velar lateral approximants, may have been neutralized before a following consonant (where the velar lateral is usually written), and possibly also in word-final position after /*y*/ (where there is some alternation in spelling in early biblical manuscripts). In Modern Armenian /*h*/ has developed to a voiced uvular fricative.

There is some uncertainty over the phonemic status and the phonetic value of the Armenian letters transcribed as *v* and *w*. In Classical Armenian they are nearly in complementary distribution; *v* occurs in word-initial (and sometimes morpheme-initial) position and after *o*, whereas *w* is found: (i) as part of the digraph *ow* for the vowel [u]; (ii) after *a*, *e*, *i*; (iii) in the position C.V in oblique cases of polysyllables ending in *-i*, for example, *ordwoy* “of the son” (genitive singular of *ordi*). Both sounds also appear to contrast with the digraph *ow* in the position C.V; note the following pairs:

- (3) *anowan*, genitive singular of *anown* “name” : *anvan* “invincible”  
*anowoy*, genitive singular of *aniw* “wheel” : *hanwoy*, genitive singular of *hani* “grandmother”

In the traditional pronunciation of Classical Armenian all three sounds are pronounced as [v].

## 3.2 Vowels

Figure 38.1 presents the phonemic inventory of Classical Armenian vowels:

	FRONT	CENTRAL	BACK
HIGH	i		ow [u]
HIGH-MID	ê	ə	
MID	e		o
LOW		a	

**Figure 38.1** The vowel phonemes of Classical Armenian

The vowel system of Classical Armenian is relatively straightforward. Vowel length is not distinctive. There are six full vowels: /*a*/, /*e*/, /*ê*/, /*i*/, /*o*/, /*ow*/, as well as /*ə*/ (schwa), which can never occur in a stressed syllable. The vowel /*ê*/ derives diachronically from \**ey*, and in some paradigms the rule *e* + *y* > /*ê*/ is still operative. There are six diphthongs /*ea*/, /*aw*/, /*iw*/, /*ew*/, /*ay*/, and /*oy*/, and two triphthongs /*eyw*/ and /*iyw*/.

of these diphthongs is disputed. The diphthong /ea/ is traditionally pronounced [ya], and /oy/ is traditionally pronounced [ui] in all positions except word-final. Word-final /-oy/ and /-ay/ are traditionally rendered as [-o] and [-a] respectively; this is almost certainly a later development, but it should be noted that in some Classical Armenian paradigms [-o] and [-a] derive synchronically from /-oy/ and /-ay/: compare the pronominal forms *k'o* “your” from /k'oy/ (genitive *k'oyoy*) and *na* “he, she, it” from /nay/ (written *nayn* with the enclitic definite article *-n*).

The vocalism of Armenian is partly dependent on the prosodic feature of stress. In Armenian the stress was always placed on the final syllable of an accented word (the few exceptions to this rule either result from recent univerbation, or are pronominal forms or interjections). High vowels and some diphthongs undergo a regular and predictable raising or reduction when not lying under the stress accent. The synchronic rules for vowel alternation are broadly as below:

(4) In stressed syllables	In unstressed syllables
i	ə
ow	ə
ê	i
oy	ow
ea	e

Consider the following examples: (i) *hin* “old,” genitive *hnoy* (read as *hənoy*); (ii) *sowt* “false,” derived verb *stem* “I lie” (read as *sətem*); (iii) *gitem* “I know,” but *angêt* “ignorant”; (iv) *yoy*s “hope,” genitive *yowsoy*; (v) *sirec'i* “I loved,” aorist of *sirem* “I love,” 3rd singular *sireac'* “(s)he loved.”

### 3.3 Phonotactics

In Classical Armenian texts the vowel ə (*schwa*) is not written except in word-initial position before a cluster of nasal or *t* followed by a consonant. This may give the impression that the language admitted complex and lengthy consonant clusters, for example, *čšmarit* “true,” *sksanim* “I begin,” *mkrtem* “I baptize,” *mštnjean* “eternal.” However, the traditional pronunciation of Classical Armenian, and the writing of *schwa* at line-endings in some manuscripts reveal that Armenian avoided complex consonant clusters in syllable-initial position. In fact, no syllable could begin with more than a single consonant. Initial combinations of the type *sibilant* + *obstruent* were pronounced with *schwa* preceding the cluster: orthographic *stin* “breast” = [əstin], *sksanim* “I begin” = [əskəsanim]. Note that such initial clusters could also be read with *schwa* separating the sibilant and obstruent in some derived terms, such as *stem* “I lie” = [sətem] from *sowt* “false.” In combinations of the type *obstruent* + *liquid*/*nasal*, the *schwa* was inserted after the obstruent: orthographic *glowx* “head” = [gəlux]; *grem* “I write” = [gərem]; *gnam* “I go” = [gənam], etc. Certain clusters of two consonants are admitted in syllable-final position, but the exact rules governing the occurrence of such clusters are not exactly known (see further Godel 1975:9–23). As stated above, geminate consonants are almost entirely excluded in Armenian; where geminates appear to occur they generally straddle a morpheme boundary.

### 3.4 Historical phonology

The development of the Classical Armenian sounds from the Indo-European parent language involved a number of intricate and sometimes unusual sound changes. However, the paucity of inherited vocabulary, and uncertainty over the correct etymologies of much

of the Armenian vocabulary often makes it difficult, if not impossible, to reconstruct the conditioning factors for a sound change. An illustrative example of the difficulties may be provided by the fate of Proto-Indo-European initial \**y* in Armenian: scholars have argued for a development to *l-*, *j-*, *ǰ-*, and *ϕ*.

The obstruent system of Armenian has no exact parallel in any other Indo-European language. What are traditionally reconstructed as the voiced stops of Proto-Indo-European are represented in Armenian by the voiceless series *p*, *t*, *c*, *č*, *k*, parallel to their outcome as voiceless stops in Germanic. The traditionally reconstructed voiceless series is continued, at least in some positions, by aspirates *pʰ*, *tʰ*, *cʰ*, *čʰ*, *kʰ*, and the voiced aspirate series by the Armenian “voiced” series *b*, *d*, *j*, *ǰ*, *g* (as seen above, it may be better to describe this series as voiced aspirate).

Among the unusual sound changes of Armenian is the regular metathesis of clusters of the type *obstruent + liquid* to *liquid + obstruent*; this occurs even in initial position, for example, *artasowk* ‘tears’ < \**draḱu-*. But, the most famous sound change, familiar from many textbooks on historical linguistics, is the development of the cluster \**dw-* to *erk-* in initial position as in the word for “two” *erkow* < \**dwō*. It is still not fully clear by what steps the sound change took place; most explanations envisage a loss of occlusion of \**d* > *r* and concomitant “hardening” of \**w* to a velar (the development of \**w* to Armenian *g* is found in other words).

## 4. MORPHOLOGY

### 4.1 Word formation

Armenian is an inflectional language of a recognizable Indo-European type. Morphological marking is mostly encoded through suffixation, although some morphological categories are expressed through prefixes and in several noun classes case-marking is shown through internal vowel changes, sometimes combined with suffixation and sometimes not.

Most of the morphological processes in Armenian are fusional, as in other ancient Indo-European languages. In Modern Armenian, morphological marking is far closer to an agglutinative type, and several morphological processes of Classical Armenian could also be described as agglutinative: for example, the instrumental case in most of the noun declensions (using the *o*-declension as a representative example):

- (5) 

<i>nom. sg.</i>	<i>ji</i>	<i>nom. pl.</i>	<i>ji-kʰ</i>
<i>instr. sg.</i>	<i>ji-ov</i>	<i>instr. pl.</i>	<i>ji-ov-kʰ</i>

The case-marking here is, however, crucially different from the system of case-marking in Modern Armenian, and, apparently, also from the majority of agglutinative languages, in that the plural marker *-kʰ* follows, rather than precedes, the instrumental marker *-ov*. Agglutinative-type patterning is also found in verbal paradigms.

The basic unit of word formation is the lexical root, which may be mono- or polysyllabic. In some verbs the aorist stem is identical with the lexical root, for example, root *tes-* “see,” third singular aorist *e-tes*; but this pattern is of limited productivity in Classical Armenian, and most present and aorist verbal stems are formed through suffixation. Most noun and adjective stems are also derived through suffixation of the root, even where the root stands as the second member of a compound. In many instances, the form of the nominative singular is coincidental with that of the root, and the suffix is only apparent in oblique cases. For example, the root *gorc-* means “work,” from which are derived (synchronically) a noun *gorc*

“work,” a verb *gorcem* “I work” and a compound adjective *angorc* “lazy.” The noun *gorc* and adjective *angorc* cannot, however, be described as root formations, as they have different declensions: *gorc* is an *o*-stem (genitive plural *gorcoc*'), whereas *angorc* is an *i*-stem (genitive plural *angorcic*'). A more exact citation form would consequently be *gorc(o)* and *angorc(i)*.

Unlike the earlier Indo-European languages, Classical Armenian has given up vowel alternation (ablaut) within a lexical root as a productive derivational marker. Ablaut alternations are, however, still found in some of the inherited vocabulary items, for example, *barjr* “high” and *erknaberj* “sky-high”; *snanim* “I nourish” and *san* “nursling”; but the only productive use of vowel alternation is in the formation of reduplicated compounds, for which see below.

Preverbs and compounding are also employed to form lexical stems, as in other Indo-European languages. However, Armenian is unusual among languages of that family in that nouns and verbs may also be derived directly from an inflected nominal form, or from a complete syntagm. For example, *kanambi* “having a wife” is derived from the instrumental singular, *kanamb*, of *kin* “woman, wife”; a common word for “night,” *c'ayg* (*o*-stem), derives from the prepositional phrase *c'-ayg* “until (*c'*-) dawn” (*ayg*, normally *ow*-stem); the adjective *č'k'met* “innocent, free from sin” is formed from a complete sentence:

- (6)    *č'=ik'*                      *met*  
          not=any.NOM.SG.    sin.NOM.SG.  
          “there is no sin”

## 4.2 Nominal declensions

The Armenian nominal declension has seven cases: nominative, accusative, genitive-dative, locative, ablative, instrumental. A few personal names borrowed from Greek show a distinct vocative form, modeled on the Greek vocative, but the nominative normally serves as the case of address. There are two numbers, singular and plural. There are no gender distinctions, even in pronouns.

The nominal declensions of Armenian are noteworthy for the large degree of partial and complete syncretism that is found. The nominative and accusative are syncretic in all declensions (except for personal pronouns) in the singular, but distinct in the plural, where the accusative and locative are always syncretic. The genitive-dative and ablative are always syncretized in the plural, but in most declensions they are distinct in the singular, while the locative is syncretized with the genitive-dative in most declensions in the singular, having a distinct case-marker only in one declension class.

In Classical Armenian there is a fairly large number of different nominal declensional paradigms. Even so, for the instrumental singular and all plural cases the case-markers themselves are either the same or morphophonemic alternatives:

- (7)    *instr. sg.*                      -w / -v / -b  
          *nom. pl.*                      -k'  
          *acc. pl., loc. pl.*              -s  
          *gen.-dat.-abl. pl.*            -c'  
          *instr. pl.*                      -wk' / -vk' / -bk'

The plural declensions of all nouns and nearly all pronouns are consequently nearly isomorphic. The declensions differ in the markers of the oblique singular cases and the vocalism of the element preceding the instrumental singular and plural and the genitive-dative-ablative plural.

It is possible to divide the regular declensional paradigms into three broad patterns.

1. The first type shows an invariable stem; the nominative singular is zero-marked and ends in a consonant, or the vowel *-ow*, or is a monosyllable ending in *-i*. The case-markers of the instrumental and genitive-dative-ablative plural are preceded by the stem-vowel: *a*, *i*, *o* or *ow*. There is also a separate subclass of the *a*-declension, restricted to personal names, which I have not listed here.

(8)	<i>a</i> -declension “year”	<i>i</i> -declension “heart”	<i>o</i> -declension “horse”	<i>ow</i> -declension “advice”
<i>Singular</i>				
<i>Nominative</i>	am	sirt	ji	xrat
<i>Accusative</i>	am	sirt	ji	xrat
<i>Genitive-dative</i>	ami	srti	jioy	xratow
<i>Locative</i>	ami	srti	ji	xratow
<i>Ablative</i>	amê	srtê	jioy	xratowê/xratê
<i>Instrumental</i>	amaw	srtiw	jiov	xratow
<i>Plural</i>				
<i>Nominative</i>	amkʼ	sirtkʼ	jikʼ	xratkʼ
<i>Accusative</i>	ams	sirts	jis	xrats
<i>Genitive-dative</i>	amacʼ	srticʼ	jiocʼ	xratowcʼ
<i>Locative</i>	ams	sirts	jis	xrats
<i>Ablative</i>	amacʼ	srticʼ	jiocʼ	xratowcʼ
<i>Instrumental</i>	amawkʼ	srtiwkʼ	jiovkʼ	xratowkʼ

The noun *xrat* is unusual in showing both ablative singular forms *xratê* and *xratowê*; most of the nouns in the *ow*-declension show only one or the other form.

2. The second type can be termed the “mixed” type; it comprises only polysyllabic nouns with a nominative singular in *-i*. Unlike the previous inflection type, the stem undergoes modifications in different cases: the final *-i* of the nominative changes to *-e* before following *-a*-, and *-w*- before following *-o*-. The inflectional endings are mostly the same as the *o*- and *a*-declensions, but several nouns of this class show a locative and ablative formed with the marker *-j* which is also found in some irregular noun declensions and in a few pronominal forms.

(9)	<i>wo</i> -declension “son”	<i>ea</i> -declension “place”
<i>Singular</i>		
<i>Nominative</i>	ordi	teḷi
<i>Accusative</i>	ordi	teḷi
<i>Genitive-dative</i>	ordwoy	teḷwoy
<i>Locative</i>	ordi/ordwoj	teḷwoj
<i>Ablative</i>	ordwoy	teḷwoy/teḷwojê
<i>Instrumental</i>	ordwov	teḷeaw
<i>Plural</i>		
<i>Nominative</i>	ordikʼ	teḷikʼ
<i>Accusative</i>	ordis	teḷis
<i>Genitive-dative</i>	ordwocʼ	teḷeacʼ
<i>Locative</i>	ordis	teḷis
<i>Ablative</i>	ordwocʼ	teḷeacʼ
<i>Instrumental</i>	ordwovkʼ	teḷeawkʼ

3. The third type covers nouns with a variable stem. The nominative singular ends *-Cn*, *-Cr* or *-Cł* (*C* = any consonant or *w*). In cases outside the nominative-accusative singular the stem changes either through vowel-insertion or vowel-alternation in the predestinential syllable. In the *r*- and *ł*-declensions the insertion vowel is normally *e*, but the *n*-declension has several subclasses with different internal vocalism.

(10)	<i>r</i> -declension “bone”	<i>ł</i> -declension “star”	<i>n</i> -declension “finger”      “blood”	
<i>Singular</i>				
<i>Nominative</i>	oskr	astł	matn	ariwn
<i>Accusative</i>	oskr	astł	matn	ariwn
<i>Genitive-dative</i>	osker	asteł	matin	arean
<i>Locative</i>	osker	asteł	matin	arean
<i>Ablative</i>	oskerê	astełê	matnê	arenê
<i>Instrumental</i>	oskerb	astełb	matamb	areamb
<i>Plural</i>				
<i>Nominative</i>	oskerkʻ	astełkʻ	matownkʻ	ariwnkʻ
<i>Accusative</i>	oskers	astełs	matowns	ariwns
<i>Genitive-dative</i>	oskeracʻ	astełacʻ	matancʻ	areancʻ
<i>Locative</i>	oskers	astełs	matowns	ariwns
<i>Ablative</i>	oskeracʻ	astełacʻ	matancʻ	areancʻ
<i>Instrumental</i>	oskerbkʻ	astełbkʻ	matambkʻ	areambkʻ

The forms given for the genitive-dative-ablative plural and the instrumental plural for the *r*- and *ł*-declension are illustrative. In early texts, forms ending *-rcʻ* and *-rawkʻ* and *-ławkʻ* are also found. The *n*-stem declension has a number of subclasses which show different patterns of vowel alternation before the *-n*- in the genitive-dative and locative of the singular and the nominative and accusative-locative in the plural. There are a number of other minor and irregular declension patterns which show variations on the above types. Several nouns also show different plural and singular declension patterns, for example: the noun *now* “daughter-in-law” declines as an *o*-stem in the singular, but an *n*-stem (nominative plural *nowankʻ*) in the plural.

### 4.3 Pronominal declensions

As in many other Indo-European languages, in Armenian the declension of deictic pronouns shows a considerable degree of integration with the nominal paradigms, whereas the personal pronouns are synchronically anomalous. The most noteworthy structural feature of the pronominal inflection is the absence of syncretism between genitive and dative singular, which is found in all nominal declensions; the personal pronouns show a difference between genitive and dative plural as well. Many of the pronominal case-markers are the same as those used in the nominal declensions, but other markers are also found, most importantly *-r* as a marker of the genitive singular, and *-owm* for the dative, locative, and ablative.

The pronominal declensions show several examples of compound case-marking, where one case is built up from another inflectional form. This process is also found in some nominal declensions (for example, the ablative form *tetwojê* formed from the locative *tetwoj* in the *wo*-declension cited above), but is more widespread among the pronouns. The case affected is always the ablative, which adds its characteristic marker *-(an)ê* to the form of the locative case in the singular, but to the genitive-dative in the plural. For example:



- (11) *nmanê* ablative singular of *na* “he, she it,” formed from *nma*, locative singular  
*owmek’ê* ablative singular of *ok’* “anyone,” formed from *owmek’*, locative singular  
*noc’anê* ablative plural of *na*, formed from *noc’a*, genitive-dative plural

(There are also sporadic examples of the unextended form *noc’a* used as an ablative plural in the earliest texts.)

Demonstrative pronouns in Armenian show a three-way deixis corresponding to proximity to speaker, addressee, and other. Pronouns marked for proximity to the speaker share a stem *s-*: anaphoric *sa*, demonstrative pronoun and adjective *ays*, *soyn* “the same.” For proximity to the addressee the stem is *d-*: anaphoric *da*, demonstrative *ayd*, *doyn* “the same.” And for nonproximity to speaker and addressee the stem is *n-*: anaphoric *na*, demonstrative *ayn*, and *noyn* “the same.” The same deixis system operates for the indeclinable clitic definite articles *-s*, *-d*, and *-n*.

Interrogative and indefinite pronouns are marked as human/nonhuman: *ov* “who?”, *omn* “someone,” *ok’* “anyone,” opposed to *zi* “what?”, and *imn/inč’* “something, anything.”

As representative of the variety of the pronominal declension, there follow the classical forms of the first singular and plural personal pronoun and the anaphoric pronoun *na*:

#### (12) Personal and anaphoric pronouns

	“I”	“we”
<i>Nominative</i>	es	mek’
<i>Accusative</i>	is	mez
<i>Genitive</i>	im	mer
<i>Dative</i>	inj	mez
<i>Locative</i>	is	mez
<i>Ablative</i>	inên	mênj
<i>Instrumental</i>	inew	mewk’

	“he, she, it”	“they”
<i>Nominative</i>	na	nok’a
<i>Accusative</i>	na	nosa
<i>Genitive</i>	nora	noc’a
<i>Dative</i>	nma	noc’a
<i>Locative</i>	nma	nosa
<i>Ablative</i>	nmanê	noc’anê
<i>Instrumental</i>	novaw	nok’awk’

### 4.4 Verbal conjugations

The Armenian verbal system shows separate categories for person, number, tense/aspect, voice, and mood. The verbal paradigm is built around the opposition of a present and aorist stem. From the present stem are formed the present indicative, imperfect indicative, present subjunctive, imperative, and infinitive, while from the aorist stem are formed the aorist indicative, aorist subjunctive, and imperative. Various nominal and adjectival formations, including the past participle, are formed from both the aorist and present stem.

Indicative forms encode both tense and aspect. The imperfect and aorist indicative both predominantly refer to situations in the time preceding the utterance, and the present indicative refers to events contemporaneous with the utterance (the present and imperfect can also have modal uses in, for example, conditional sentences). Reference to events in future

time is usually made with the subjunctive. The basis of the opposition between present and aorist stems is aspectual. This is most clearly seen in the imperative: the present imperative is used (in all cases except for the existential verb) with the negative *mi* as a prohibitive, but the aorist imperative is only used as a positive imperative. However, the basis of the aspectual nuance in the subjunctive and indicative is not always clear; in the Bible translation, Greek aorist subjunctives are translated by both aorist and present subjunctives, and Greek present subjunctives similarly. It has been suggested that the present/aorist opposition was differently organized in Armenian and Greek and that in Armenian the aorist is the marked aspect (see Meillet 1909:104–113 [= 1962:93–102] and Vogt 1930 [= 1988:8–24]).

Classical Armenian has a curiously skewed system of voice marking. In the present indicative and imperative, one class of verbs, with first-person singular *-em*, regularly forms mediopassives in *-im*. However, verbs with first-person singular *-im* (other than those which serve as mediopassives to verbs in *-em*), *-am*, and *-owm* are not marked for voice. In the imperfect, indicative voice is not marked in any verb paradigm. But in the aorist all indicative, subjunctive and imperative forms are marked as active or mediopassive with a separate set of endings. There is no marking of voice in the infinitive or past participle. In addition, there is a large number of deponent verbs which only show mediopassive forms.

#### 4.4.1 Present tense

In the present stem system, nearly all verbs fall into four different classes: the *-em*, *-im*, *-am*, and *-owm* conjugations. Examples of conjugation are as follows:

##### (13) Present and imperfect indicative

<i>Present indicative</i>				
	“I bring”	“I am brought”	“I hope”	“I take”
1st sg.	berem	berim	yowsam	aṛnowm
2nd sg.	beres	beris	yowsas	aṛnows
3rd sg.	berê	beri	yowsay	aṛnow
1st pl.	beremkʻ	berimkʻ	yowsamkʻ	aṛnowmkʻ
2nd pl.	berêkʻ	berikʻ	yowsaykʻ	aṛnowkʻ
3rd pl.	beren	berin	yowsan	aṛnown
<i>Imperfect</i>				
1st sg.	berei		yowsayi	aṛnowi
2nd sg.	bereir		yowsayir	aṛnowir
3rd sg.	berêr		yowsayr	aṛnoyr
1st pl.	bereakʻ		yowsayakʻ	aṛnowakʻ
2nd pl.	bereikʻ		yowsayikʻ	aṛnowikʻ
3rd pl.	berēin		yowsayin	aṛnowin

The endings of the present subjunctive conjugation are identical with those of the present indicative, but the stem is formed by the addition of a suffix *-icʻ-* to the present stem, illustrated here with the first and second persons of the singular:

##### (14) Present subjunctive

1st sg.	bericʻem	bericʻim	yowsaycʻem	aṛnowcʻowm
2nd sg.	bericʻes	bericʻis	yowsaycʻes	aṛnowcʻows

Note that the subjunctive of the *-am* conjugation (*yowsayc'em*) is formed with the quasi-active personal endings *-em*, *-es*, etc.; this is the case even for deponent verbs which exclusively take passive endings in the aorist.

The present imperative has special endings in the second-person singular:

- (15) berer berir yowsar aṛnowr

#### 4.4.2 Aorist tense

There is only a single set of inflectional endings for the active and mediopassive in the aorist system. The prefix *e-* (termed the “augment”) is attached to finite aorist verb forms which would otherwise be monosyllabic. As an example, consider *tesī*, the aorist of *tesanem* “I see”:

##### (16) Aorist indicative

	Active	Passive
1st sg.	tes-i	tes-ay
2nd sg.	tes-er	tes-ar
3rd sg.	e-tes	tes-aw
1st pl.	tes-ak'	tes-ak'
2nd pl.	tes-êk'	tes-ayk'
3rd pl.	tes-in	tes-an

The aorist subjunctive is formed with the suffix *-c'* and a special set of endings, some of which recall the present *-em* and *-im* conjugations:

##### (17) Aorist subjunctive

	Active	Passive
1st sg.	tes-ic'	tes-ayc'
2nd sg.	tes-c'es	tes-c'is
3rd sg.	tes-c'ê	tes-c'i
1st pl.	tes-c'owk'	tes-c'owk'
2nd pl.	tes-ĵik'	tes-ĵik'
3rd pl.	tes-c'en	tes-c'in

The aorist imperative has different endings for mono- and polysyllabic aorist stems in the singular. For polysyllabic stems, the aorist active imperative is formed through loss of the final consonant of the stem. For monosyllabic stems the endings are as follows:

##### (18) Aorist imperative

	Active	Passive
2nd sg.	tes	tes-ir
2nd plural	tes-êk'	tes-arowk'

### 4.5 Nonfinite verbal formations

All Armenian verbs can form a verbal noun, traditionally termed the infinitive, through the addition of *-l* to the present stem, with the slight complication that verbs in the *-im* conjugation have an infinitive in *-el*. The infinitive is not marked for aspect or voice, and it behaves like a noun in that it declines (as an *o*-stem), can be marked by an article and by

dependent genitives, and is governed by prepositions. As the complement after verbs the infinitive never receives a definite article or defining genitive.

All Armenian verbs are also capable of forming a past participle, usually by the addition of the suffix *-eal* to the aorist stem; in one class of verbs, those with present *-em/-im*, aorist *-ecʻi/-ecʻay*, the past participle can be optionally formed from the present stem. The past participle is also unmarked for voice. The subject of the participle frequently stands in the genitive case when it has a transitive meaning.

Other suffixes also attach to the present or aorist stem to form verbal adjectives. The most closely integrated into the verbal system is the so-called future-participle formed with a suffix *-ocʻ* added directly to the present infinitive. These forms are always found as predicates with the copula verb to denote immediacy, necessity, or obligation; they are not marked for voice. The suffix *-i* is also added directly on the infinitive to form passive adjectives, such as *sireli* “lovable” from *sirel* “to love.”

Two other “quasi-participial” forms should be mentioned:

1. The first is an *a*-stem construction in *-otʻ/-awtʻ* (both spellings are found in the earliest manuscripts), formed from either present or aorist stem: both *tesan-otʻ* and *tes-otʻ* “see-ing” are found in early texts.
2. The second is an *o*-stem formation in *-own*, built from the present stem. These forms are predictable in meaning and used with the same syntactic constructions as the verbs from which they are derived, but neither is freely productive.

## 4.6 Derivational morphology

Armenian mainly forms derived verbs and nouns through suffixation. There are a large number of different suffixes, many highly productive in Classical Armenian with largely predictable meanings. Nouns can be derived from nouns, verbs, or indeed whole syntagms (see §4.1 and the example of [6]). For most nouns the nominative singular also serves as the stem to which suffixes are attached; for deverbative formations the aorist stem of the verb is usually used. Some suffixed forms may themselves serve as a base for further suffixation, with the consequence that a single lexical root may have a large number of derivatives.

Some examples of Armenian patterns of nominal suffixation can be shown from the following derivatives of *gorc* “work, action, manufacture,” found in fifth-century Armenian (compounds have been excluded):

1. *gorcawor* “workman, anyone who works,” formed with the suffix *-awor* which forms nouns denoting occupation or profession. From this is further built *gorcaworowtʻiwn* “work (in the abstract), labor” with the extremely common abstract noun suffix *-owtʻiwn*.
2. *gorci* “tool,” formed with the suffix *-i* which is sometimes used, as here, to denote an instrument. A derivative of *gorci* can be made by the adjectival suffix *-akan*, yielding *gorciakan* “instrumental.”

There are also a number of highly productive derivational suffixes used to form verbs. Extremely common is the suffix used to form causatives: present *-owcʻane-*, aorist *-owcʻ-* (third singular *-oycʻ*, aorist imperative *-o*), added to the aorist stem of the verb: thus, *darñnam* “I turn (intr.),” aorist *darjay* gives *darjowcʻanem* “I turn (tr.).”

## 4.7 Compounds

Compounding is a productive process of word formation in Classical Armenian (see Meillet 1913b = 1962:159–184 for the best survey); indeed, many of the derivational suffixes of

Armenian (for example, *-awor* mentioned above) evolve from generalized compound forms. For all compounds the head of the compound occurs as the second member. The first member of a compound, if a noun or adjective, normally stands in the stem-form which, for most items, is identical with the nominative singular. When the second element of a compound does not begin with a vowel, the productive pattern is to insert a liaison vowel *-a-* between the two members of the compound. However, a number of compounds are formed without the liaison vowel *-a-*, and in derivatives of compounds the liaison vowel is often dropped. The principal productive types of compounding found in Armenian are the following:

1. Exocentric compounds of the type *modifier + head noun*: *mec* “big” + *town* “house” > *mecatown* “rich”; *an-* “without-” + *mit(k’)* (*a*-stem) “mind” > *anmit* “mad, senseless.” Exocentric compounds frequently follow the same declension class as their head noun, but many are declined as *i*-stems: in the Bible translation *anmit* is found declined both as an *i*-stem and as an *a*-stem.
2. Endocentric compounds of the type *modifier + head noun*: *aysawr* “today,” from *ays* “this” and *awr* “day.” This type of compound is of limited productivity in fifth-century Armenian.
3. Governing compounds, with a verbal element as the second member. This type is highly productive in Armenian. As examples, consider: *jowkn* “fish” + *orsam* “I hunt” > *jknors* “fisher” (*a*-stem); *andam* “limb” + *lowcanem* “I loose” (aorist *lowci*, third singular *eloyc*) > *andamaloyc* “paralytic, having been loosed as to the limbs.” This second example shows a compound which appears to be exocentric with the first element as its head: “having loosened limbs.”
4. There are a small number of copulative compounds in Armenian; usually these show the conjunction *ew* between the two elements: *ert’ewek* (*a*-stem) “coming and going,” derived from the stems of *ert’am* “I come” and *eki* suppletive aorist of *gam* “I go.”
5. A productive means of forming words with intensive or distributive meanings is through reduplication of the same lexical element, sometimes with associated vowel or consonant changes (see Leroy 1986 for full survey): *mecamec* “very big” (*mec* “big”); *dasadas* “in divisions” (*das* “division”); *kerakowr* “food” (suppletive aorist *ker-ay* “I ate”); *atxamatax* “diverse goods for sale” (*atx* “box, baggage”).

## 4.8 Numerals

The numeral system of Armenian is decimal. The numbers 1–16 and the decads, hundreds and 1,000 and 10,000 are expressed by single lexical items; other numbers are formed through juxtaposition and combination using *ew* “and.” Suffixed forms of the cardinal numbers are used to express ordinal, collective, multiplicative, and iterative numerals. The numbers 1–4 are inflected in all cases, but higher numbers rarely show inflection in early texts. Some of the lower numerals follow as examples: *mi* “1,” *erkow* “2,” *erek* “3,” *č’ork* “4,” *hing* “5,” *vec* “6,” *ewt’n* “7,” *owt* “8,” *inn* “9,” *tasn* “10,” *k’san* “20,” *k’san ew inn* “29.”

## 5. SYNTAX

There is only space here to sketch out a few of the more remarkable features of Armenian syntax; some topics of relevance, such as the role of aspect in the verbal system, have already been discussed. Other topics, such as the syntax of coordinate and subordinate clauses, will

be omitted from what follows since the syntactic elements are largely familiar from other Indo-European languages; thus Armenian uses particles to introduce subordinate clauses which have an internal structure similar to that of main clauses.

## 5.1 Word order

In Classical Armenian, word order has mostly a pragmatic, rather than syntactic, function. Modern Armenian is a fairly rigorous head-final language, but the earlier language had different preferred orders depending on the nature of the syntactic constituent. Armenian has prepositions, rather than postpositions; in noun phrases the unmarked order is *adjective – head noun*, but *head noun – dependent genitive*. Armenian prose exhibits great variety in the position of the verb in the sentence, with verb-initial placement particularly frequent in historical narrative.

## 5.2 Concord

The rules for concord in Classical Armenian are not straightforward, particularly for noun phrases. Modern Armenian has moved away from the Indo-European pattern, in which all constituents in a noun phrase are marked for concord or dependence, towards a system in which there is only one marker for the whole phrase. The earlier language appears to stand halfway between the two types. Adjectives sometimes agree with their head nouns, but sometimes they do not. Meillet (1900 = 1962:39–55) worked out the following general tendencies:

- (19) A. adjectives following their head noun show concord  
 B. monosyllabic adjectives preceding their head noun show concord unless the noun is nominative or accusative-locative plural

As Meillet noted, these rules do not always apply, and sometimes adjective and noun show partial concord: they are both in the same case, but the adjective is singular and the noun is plural, for example, Mark 5:42:

- (20) *ew zarmac'an mecaw zarmanaleawk'*  
 and amaze.AOR.PASS.3RDPL. great.INSTR.SG. amazement.INSTR.PL.  
 "and they were amazed with great amazement"

Clearly this could also explain the lack of concord between adjectives and nouns in the nominative and accusative plural, since the nominative-accusative singular is zero-marked.

Noun phrases involving numerals also follow a peculiar pattern of concord. When joined with one of the inflected numerals, 2 to 4, head nouns are marked as plural. With a higher numeral, head nouns are mostly only marked as plural nominative or accusative if they precede the numeral. When they follow the head noun they are marked as singular, and sometimes a following verb is inflected as singular not plural, for example, Luke 8:2:

- (21) *ewt'n dew=n eal êr*  
 seven devil.NOM.SG.=ART. leave.PAST.PART. be.IMPF.3RDSG.  
 "the seven devils had left"

The marking of nominal dependents in noun phrases also shows divergence from the Indo-European type. A sporadic feature observable in some Armenian texts (most frequently the early historical writers) but avoided in others (e.g., the Gospel versions) is case attraction, whereby all constituents of a noun phrase, including adnominal dependents, are attracted

into the same case as the head noun (see Hübschmann 1906:478–480 = 1976:434–436, and Vogt 1932 = 1988:25–49). For example, at Genesis 6:7 many Armenian manuscripts read

- (22)  $y=eresac^{\circ}$   $erkrê$   
 from=face.ABL.PL. earth.ABL.SG. (*eresk* “face” is *plurale tantum*)

to render Greek ἀπὸ προσώπου τῆς γῆς “from the face of the earth,” rather than the “expected” (in terms of Indo-European syntax) construction:

- (23)  $*y=eresac^{\circ}$   $erkri$   
 $*from=face.ABL.PL.$  earth.GEN.SG.

This “case attraction” is most frequent when the head noun stands in the ablative or instrumental, although there are also examples where the head noun is in the locative. The case into which the adnominal dependent is attracted is always unambiguous.

### 5.3 Case usage

Case usage in Classical Armenian is broadly similar to that found in other older Indo-European languages, but there are a few important areas of divergence. Except for a few fossilized phrases, cases only have local functions in conjunction with prepositions. Some grammatical functions are also marked by prepositions: the use of the ablative with the preposition *i* (prevocalic *y-*) “from” as the case of the agent after passive verbs is not surprising, but a more interesting phenomenon is the use of the preposition *z-* to mark the accusative. When a noun or pronoun in the accusative is definite, the preposition *z-* precedes the noun, but indefinite items are not so marked. For example:

- (24)  $etes$   $kin$   
 see.AOR.3RD SG. woman.ACC.SG.  
 “he saw a woman”

but

- (25)  $etes$   $z=kin=n$   
 see.AOR.3RD SG. PREP.=woman.ACC.SG.=ART.  
 “he saw the woman”

### 5.4 Cliticized articles

Armenian marks definiteness with three clitic particles *-s*, *-d*, *-n*, termed articles, which are unmarked for case, number, or gender, but which are marked for proximity, correlating with the deictic pronouns *ays* “this (near speaker),” *ayd* “that (near addressee),” and *ayn* “that” (see Jungmann 1964 and 1965). In early Armenian texts these articles frequently have a weak deictic force, for example, Mark 13:1,

- (26)  $tes$   $orpisi$   $en$   $k'arink'=s$   
 see.AOR.IMPV. what-sort be.PRES.3RDP. stone.NOM.PL.=ART.  
 “Look, how wonderful the(se) stones are!”

but they are used also without any perceivable deictic force, for example, Mark 14:38:

- (27)  $ogi=s$   $yawžar ê$   $bayc' marmin=s$   $tkar ê$   
 spirit=ART. willing be.PRES.3RD SG. but body=ART. weak be.PRES.3RD SG.  
 “the spirit is willing but the flesh is weak”

Relative phrases have a syntax similar to that of noun phrases, even to the extent that an entire relative phrase can be marked with an article, which is attached to the first accented word in the relative phrase, whatever part of speech it is, for example, Mark 14:65:

- (28) *ov ê ayn or ehar=n z=k'ez?*  
 who be.PRES.3RDSG. that.NOM.SG. who hit.AOR.3RDSG.=ART. PREP=YOU.ACC.  
 “Who is the one that hit you?”

## 5.5 Syntax of the past participle

A peculiar and much discussed aspect of Armenian syntax is the construction used with the past participle and the periphrastic perfect, formed from the combination of the past participle and copula. For intransitive verbs, or the passive of transitive verbs, the subject of the participle is usually in the nominative, and the copula agrees in number with the subject, for example, Matthew 4:24:

- (29) *or netēal ein*  
 who afflict.PAST.PART. be.IMPF.3RDPL.  
 “(those) who had been afflicted”

However, the construction with transitive verbs is highly unusual. The logical object is in the accusative case, but the logical subject is placed in the genitive case; in the periphrastic perfect the copula always takes the third-person singular form, for example, Matthew 6:8:

- (30) *minč'ew jer xndreal inč' ic'ê*  
 before YOU.GEN.PL. [seek.PAST.PART] [anything.ACC.] [be.SUBJUNC.3RDSG.]  
*i nmanê*  
 from he.ABL.SG.  
 “before you seek anything from him”

When a participle phrase precedes a different main verb the subject of the participle remains in the genitive, even if it is the subject of the main verb of the sentence, for example, Matthew 9:2:

- (31) *ew tēseal yisowsi z=hawats noc'a asê...*  
 and see.PAST.PART. JESUS.GEN.SG. PREP.=faith.ACC.PL. he.GEN.PL. say.PRES.3RDSG.  
 “And Jesus seeing their faith said . . .”

## 6. LEXICON

The lexicon of Classical Armenian has a number of different components. Only a small proportion of the lexicon is inherited directly from Proto-Indo-European. This includes a number of basic vocabulary items: the lower numerals (the word for “100” *hariwr* is not Indo-European, but its source is not clear); many of the terms for kinship relations; body parts; livestock; adjectives denoting physical properties; verbs denoting common human activities or experiences. A few noteworthy shifts of meaning have taken place: PIE *\*ekwo-* “horse” > Armenian *êš* “donkey”; PIE *\*g<sup>h</sup>eh<sub>2</sub>-* “step, go” > Armenian *kam* “stand,” PIE *\*h<sub>2</sub>nēb<sup>h</sup>* “boss, hub” > Armenian *aniw* “wheel.”

There are a very large number of Iranian loanwords in Armenian, over a thousand separate lexical items not counting derivatives or compounded forms. The Iranian influence on the Armenian language is comparable to the influence of Norman French and Latin on English.



In Armenian, not only is the larger part of vocabulary of administration, military life, and religion borrowed from Iranian, but also adjectives and prepositions and a number of adjectival, adverbial, and nominal suffixes. Even phrasal combinations of noun and verb are calqued from Iranian. The loans can be divided into two different strata: (i) during the Parthian period (c. 200 BC to AD 400), cultural and political contacts between the Armenians and Iranians were closest, and there was a large influx of words from Parthian including common terms such as *mah* “death,” *ašxarh* “land,” *šat* “very,” *seaw* “black” and *spitak* “white”; (ii) in the later Sasanian period, contact was much less close and loanwords from this period are not well integrated into the Armenian lexicon.

With the advent of Christianity, more loanwords entered the language, principally from Greek and Syriac, resulting from increased contact with fellow Christians and the use of the Greek and Syriac Bibles in the early Armenian Church.

Although a large portion of the Armenian vocabulary can be traced to its Indo-European, Iranian, Greek or Syriac origin, much remains obscure, for example, the words *sag* “goose,” *zok'anč* “mother-in-law” and *glowx* “head.”

## Bibliography

### *Introduction, grammars and collections of articles*

- Ajello, R. 1997. “Armenian.” In A. Giacalone Ramat and P. Ramat (eds.), *The Indo-European Languages*, pp. 197–227. New York: Routledge.
- Clackson, J. 1994. *The Linguistic Relationship between Armenian and Greek*. Oxford: Blackwell.
- de Lamberterie, C. 1992. “Introduction à l’arménien classique.” *LALIES* 10:234–289.
- Djahukian, G. B. 1986. “Introduction.” In J. A. C. Greppin and A. A. Khachaturian (eds.), *A Handbook of Armenian Dialectology*. Delmar, NY: Caravan Books.
- Godel, R. 1975. *An Introduction to the Study of Classical Armenian*. Wiesbaden: Dr. Ludwig Reichert.
- Hübschmann, H. 1875. “Ueber die Stellung des Armenischen im Kreise der indogermanischen Sprachen.” *Zeitschrift für vergleichende Sprachforschung* 23:5–49.
- . 1976. *Kleine Schriften zum Armenischen*. Edited by R. Schmitt. Hildersheim: Olms.
- Jensen, H. 1959. *Altarmenische Grammatik*. Heidelberg: Carl Winter.
- Meillet, A. 1913a. *Altarmenisches Elementarbuch*. Heidelberg: Carl Winter.
- . 1962. *Études de linguistique et de philologie arméniennes I: Recherches sur la syntaxe comparée de l’arménien, suivies de la composition en arménien*. Lisboa: Librairie Bertrand.
- . 1977. *Études de linguistique et de philologie arméniennes II*. Louvain-la-Neuve: Orientaliste.
- Renoux, C. 1993. “Langue et littérature arméniennes.” In M. Albert, R. Beylot, René-G. Coquin, et al. (eds.), *Christianismes orientaux. Introduction à l’étude des langues et des littératures*, pp. 107–166. Paris: Editions du Cerf.
- Schmitt, R. 1981. *Grammatik des Klassisch-Armenischen mit sprachvergleichenden Erläuterungen*. Innsbruck: Institut für Sprachwissenschaft der Universität.
- Solta, G. R. 1963. “Die Armenische Sprache.” In G. Deeters (ed.), [*Armenisch und kaukasische Sprachen*.] *Handbuch der Orientalistik I/VII*. Leiden: Brill.
- Tumanjan, E. G. 1971. *Drevnearmjanskij jazyk*. Moscow: Nauka.
- Vogt, H. 1988. *Linguistique caucasienne et arménienne*. *Studia Caucasologica II*. Edited by E. Hovdhaugen and F. Thordarson. Oslo: Norwegian University Press.

### *Script*

- Russell, J. R. 1994. “On the origins and invention of the Armenian script.” *Le Muséon* 107:317–333.
- Sanjian, A. K. 1996. “The Armenian alphabet.” In P. T. Daniels and W. Bright (eds.), *The World's Writing Systems*, pp. 356–363. New York/Oxford: Oxford University Press.
- Schmitt, R. 1972. “Empfehlungen zur Transliteration der armenischen Schrift.” *Zeitschrift für vergleichende Sprachforschung* 86:296–306.

### Phonology

- Feydit, F. 1982. *Considérations sur l'alphabet de Saint Mesrop et recherches sur la phonétique de l'arménien* (2nd edition). Vienna: Mechitharisten.
- Jensen, H. 1957. "Die altarmenische Aussprache der Buchstaben *kh ph th*." In J. Irmscher (ed.), *Aus der byzantinischen Arbeit der Deutschen Demokratischen Republik*, vol. I, pp. 27–38. Berlin: Akademie-Verlag.
- Vaux, B. 1998. *The Phonology of Armenian*. Oxford: Oxford University Press.
- Vogt, H. 1958. "Les occlusives de l'arménien." *Norsk Tidsskrift for Sprogvidenskap* 18:143–161.

### Morphology

- Connolly, M. J. 1972. "A descriptive normalisation of the Classical Armenian nominal system." *Revue des Études Arméniennes* NS 9:1–45.
- Klingenschmitt, G. 1982. *Das altarmenische Verbum*. Wiesbaden: Dr. Ludwig Reichert.
- Leroy, M. 1986. "Le redoublement comme procédé de formation nominal en arménien classique." In M. Leroy and F. Mawet (eds.), *La place de l'arménien dans les langues indo-européennes*, pp. 62–75. Louvain-la-Neuve: Peeters.
- Meillet, A. 1909. "Recherches sur la syntaxe comparée de l'arménien. IV. Emploi des formes personnelles des verbes." *Mémoires de la Société de la Linguistique de Paris* 16:92–131.
- . 1913b. "De la composition en arménien." *Mémoires de la Société de la Linguistique de Paris* 18:245–270.
- Olsen, Birgit Anette. 1999. *The Noun in Biblical Armenian*. Berlin/New York: Mouton de Gruyter.
- Stempel, R. 1983. *Die infiniten Verbalformen des Armenischen*. Frankfurt am Main/Bern/New York: Peter Lang.
- Vogt, H. 1930. "Les deux thèmes verbaux de l'arménien classique." *Norsk Tidsskrift for Sprogvidenskap* 4:129–145.

### Syntax

- De Lamberterie, C. 1997. "L'article dans la relative en arménien classique." In E. Crespo and J. L. Garcie Ramon (eds.), *Berthold Delbrück y la sintaxis indoeuropea hoy: Actas del Coloquio de la Indogermanische Gesellschaft, Madrid, 21–24 de septiembre de 1994*, pp. 311–326. Madrid: Ediciones de la Universidad Autónoma de Madrid; Wiesbaden: Dr. Ludwig Reichert.
- Hübschmann, H. 1906. "Armeniaca." *Indogermanische Forschungen* 19:457–480.
- Jungmann, P. 1964. "L'emploi de l'article défini avec le substantif en arménien classique." *Revue des Études Arméniennes* NS 1:47–99.
- . 1965. "L'emploi de l'article défini avec le substantif en arménien classique (suit et fin)." *Revue des Études Arméniennes* NS 2:43–116.
- Klein, J. S. 1996. *On Personal Deixis in Classical Armenian: A Study of the Syntax and Semantics of the n-, s-, and d- Demonstratives in Manuscripts E and M of the Old Armenian Gospels*. Dettelbach: Röhl.
- Lyonnet, S. 1933. *Le parfait en arménien classique*. Paris: Champion.
- Meillet, A. 1898. "Recherches sur la syntaxe comparée de l'arménien. I. les démonstratifs." *Mémoires de la Société de la Linguistique de Paris* 10:241–273.
- . 1900. "Recherches sur la syntaxe comparée de l'arménien. II. Les règles d'accord de l'adjectif." *Mémoires de la Société de la Linguistique de Paris* 11:369–389.
- Ouzounian, A. 1992. *Le discours rapporté en arménien classique*. Louvain-la-Neuve: Peeters.
- Vogt, H. 1932. "Les groupes nominaux en arménien et en géorgien ancien." *Norsk Tidsskrift for Sprogvidenskap* 5:57–81.
- . 1937. "Les formes nominales du verbe arménien. Étude sur la syntaxe de l'arménien classique." *Norsk Tidsskrift for Sprogvidenskap* 8:5–70.

### Dictionaries and lexical studies

- Ačařyan, H. 1971–1979. *Hayeren armatakan bařaran* (2nd edition, 4 vols.). Erevan: University Press.

- Awetik'ean, G., X. Siwrmélean, and M. Awgerean. 1836–1837. *Nor baṙgirk' haykazean lezowi* (2 vols.). Venice: St. Lazar. Reprint 1990, Osnabrück: Biblio.
- Hübschmann, H. 1895–1897. *Armenische Grammatik. I Teil: Armenische Etymologie*. Leipzig: Breitkopf and Härtel.
- Jungmann, P. and Jos. J. S. Weitenberg. 1993. *A Reverse Analytical Dictionary of Classical Armenian*. Berlin/New York: Mouton de Gruyter.
- Künzle, B. O. 1984. *Das altarmenische Evangelium/L'Évangile arménien ancien* (2 vols.). Bern: Peter Lang.
- Solta, G. R. 1960. *Die Stellung des Armenischen im Kreise der indogermanischen Sprachen*. Vienna: Mechitharisten.

# Etruscan

HELMUT RIX

## 1. HISTORICAL AND CULTURAL CONTEXTS

Etruscan, the language of the Etruscans, is attested between 700 BC and AD 50 in the area of northwest central Italy between the Arno, the Tiber, and the Tyrrhenian Sea. A few Etruscan texts come from other areas of Italy (especially from Campania and Emilia) and from Corsica, and isolated examples are known from Provence, Tunisia, Greece, and Egypt.

The most important source of Etruscan is the *c.* nine thousand inscriptions. The majority are funerary inscriptions, which often consist of no more than the name of the deceased. The second largest group is formed by the likewise mostly short texts on objects of daily life which indicate the owner or the manufacturer, or the object as a present or a dedication. Readily comprehensible are the labels inscribed next to figures in pictorial representations. The longer inscriptions are legal or ritual in character. The quasi-bilingual from Pyrgi (with a parallel text in Phoenician) reports the dedication of a cult building; the Perugine cippus records a contract about a piece of land; the clay tablet of Capua (which, with 300 preserved words, is the longest Etruscan inscription) preserves a ritual calendar; and the recently published (Agostiniani and Nicosia 2000) bronze tablet of Cortona seems to contain, as I think, a record of the treatment of tenant farmers after the sale of an estate rented by them. A calendar of rituals is also described in the one noninscriptional, and at the same time longest (1,500 words), Etruscan text – a linen book, which was torn up and used as wrappings on a mummy in Egypt and of which a good half is preserved (often called the Zagreb mummy after its present location). Interesting secondary sources for the lexicon and for textual interpretation are glosses (meanings of Etruscan words given by Latin and Greek authors; e.g., *aesar . . . etrusca lingua deus*, [“*aesar* . . . the Etruscan word for god”] Suetonius, *The Life of Augustus* 97) and loanwords in Latin (*satelles* “body guard” < Etr. *zat[i]laθ* “striker”).

The prehistory of the Etruscans has been disputed for two thousand years. Historians of the fifth century BC (Herodotus 1.94, Hellenicus in Dionysius of Halicarnassus 1.28.3–4) had claimed immigration from the Aegean; the orator Dionysius of Halicarnassus (first century BC) argued from the lack of related languages (but see below) for the autochthony of the Etruscans in Italy. Until now archeological arguments (Pallottino 1988:77–101) have been as poorly conclusive as linguistic.

In the course of their history (seventh to first centuries BC) the Etruscans never formed a centrally governed state. Rather they lived in separate city-states, which were first ruled by monarchs and which later, from around 500 BC, became oligarchies, and were tied to each

other through common cult festivals. The Etruscans who possessed citizenship, the Πάσεννα (Dionysius of Halicarnassus 1.30.3; < Etr. *rasna* “army, people”; see Rix 1984b), clearly made up only a part of the population. Beside these there was a rural population (Πενέστοι, Dionysius of Halicarnassus 9.5.4), with personal freedom and economic independence, but without political rights and at least in part of Italic origin. Only in the third to second centuries was this section of the population incorporated into the Etruscan citizenry (Rix 1963:372–376).

Until the beginning of the fifth century BC the Etruscans were the dominating power in upper and central Italy. The defeat at Cumae by the Greeks in 474 BC marks the beginning of the Etruscan decline, which was accelerated by the invasion of the Celts in the fourth century BC. Politically the Etruscans became dependent allies of Rome at the beginning of the third century and two hundred years later Roman citizens. Shortly after the turn of the millennium, Etruscan ceased to be written; around which time the language would also have ceased to be spoken.

The syncope of unaccented internal vowels (see §3.5.2.4) – which around 480 BC changed the structure of many words and may well be connected with the social and political changes of the time – marks the break between *Archaic Etruscan* and *Late Etruscan*. Since the third century, and intensely in the first century, Latin influence is perceptible (orthography, morphology); incorrect texts appear. In spite of changes in the development of the sound system (both some regional changes [see §3.5.1] and fewer affecting the whole of the Etruscan area [see §3.5.2]), there is no evidence that distinct Etruscan dialects developed. This correlates with the political structure of Etruria and speaks for a relatively late spread of the language from a limited area.

To the same language family as Etruscan there belong only two poorly attested languages: Lemnian in the Northeast of the Aegean (sixth century BC; Agostiniani 1986) and Rhaetic in the Alps (fifth to first centuries BC; Schumacher 1992:246–248; Rix 1998). Lemnian and Rhaetic are so close to Etruscan that Etruscan can be used to understand them. The date of the common source language, *Proto-Tyrsenic*, can probably be fixed to the last quarter of the second millennium BC. The location of its homeland is disputed, however; possibilities include: (i) the northern Aegean, whence Proto-Etruscan and Proto-Rhaetic speakers would have come in the course of the Aegean migration westwards at the end of the second millennium (similarly Herodotus [1.94] identifies Lydia as the Etruscan homeland); (ii) central Italy, from which Proto-Lemnian speakers would have migrated eastwards and Proto-Rhaetic speakers northwards. A decisive judgment is not currently possible.

The lack of well-known related languages limits the comparative method's access to Etruscan to the area of loanwords (see §6). Moreover, in reading an Etruscan text, one must first attempt to determine a text's message from its context, and then to correlate the elements of content in the message with the structural elements in the text. Hereby glosses, loanwords, and above all texts in the better-known languages of the same cultural area (Latin, Greek, and so on) can help. From the results, a grammar and a lexicon can be constructed tentatively; these serve to test hypotheses and require continual amendment.

In this way a significant number of elements and rules have been identified more or less securely for the grammar and lexicon of Etruscan, and the meaning of a considerable number of texts and text fragments has been made clear. We are, of course, still far from a complete understanding of the Etruscan language, so that much of what is presented below still needs to be stated more precisely, amended, and corrected.

**Table 39.1 The Etruscan alphabet of archaic inscriptions**

Character	Transcription	Character	Transcription
A	a	𐌀	p
𐌁	c	𐌁	σ
𐌂	e	𐌂	q
𐌃	v	𐌃	r
𐌄	z	𐌄	s
𐌅	h	𐌅	t
𐌆	θ	𐌆	u
𐌇	i	𐌇	ś
𐌈	k	𐌈	Φ
𐌉	l	𐌉	χ
𐌊	m	𐌊	f
𐌋	n		

## 2. WRITING SYSTEMS

The Etruscan writing system is an alphabet, which was created at the end of the eighth century BC, in several local variants, from an alphabet of West Greek origin; it was taught in scribal schools and is attested in inscriptions (see Table 39.1). The West Greek alphabet contained twenty-two letters derived from Phoenician characters plus four additional signs of Greek origin. This form of the Greek alphabet used X for the sequence /k/ + /s/ and Ψ for /k<sup>h</sup>/. A few letters, for which Etruscan had no use, were not used in texts (“lettres mortes”: β, Δ, Ο and Phoenician *samekh* (𐤌) = East Greek Ξ).

The southern variant of the “working” alphabet used three different letters for the three phonetic variants of /k/: (i) *q* (𐌂; Greek *koppa*) before following /u/; (ii) *k* (𐌈; Greek *kappa*) before /a/; and (iii) *c* (𐌁; Greek *gamma*) before /i/ or /e/. This distribution, which continued and generalized an early Greek practice (*koppa* before or after /u/), was possible because Etruscan did not have voiced obstruents and so had no other use for Greek *gamma* (spelling /g/ in Greek). Of the two sibilant phonemes (see §3.1.1), the southern Etruscan script chiefly represents alveolar /s/ with a three-stroke *sigma* (𐌄) (in the far south X [= East Greek Ξ] is also used) and the less common palato-alveolar /ś/ with Greek *san* (𐌈 = Phoenician *šade*; details in Cristofani 1972:469–473; see also Woodard 1997:161–188).

In the northern writing area of Etruria /k/ is at first written simply as *k*. *Sigma* and *san* were used in a way quite the reverse of that in the south – *sigma* represents palato-alveolar /ś/, *san* represents alveolar /s/. Since in the north, alveolar /s/ before consonants had developed prehistorically to palato-alveolar /ś/, this reversal may have arisen by the creator of the northern alphabetic variants beginning with words which he himself pronounced with /ś/ but which in his southern model he found written with *sigma* (for instance *spura* “community”; Rix 1998). In the north in the later period, alveolar /-s/ was occasionally written with *sigma* as a result of Latin influence. Otherwise the north–south opposition with regard to the writing of the sibilants was maintained up to the end of the Etruscan writing tradition.

In contrast to traditional transliteration based on graphemes, sibilant signs are herein transcribed phonemically (as in Rix and Meiser 1991): /s/ as *s*, if *sigma* is written, and as *ś*,

if *san* is written; and /š/ with  $\sigma$ , if *san*, and as  $\acute{\sigma}$ , if *sigma* is written; likewise northern /s/ (which in certain contexts became [š]) is *phonemically* transcribed with *s*.

By 300 BC the inventory of the Etruscan alphabets had decreased significantly. In the sixth century the south gave up *X*, *q*, and *k*. In the north, in the fourth century, *c* won out for representing /k/, as it also did in the northeast by the middle of the third century. By around 250 BC only nineteen of the twenty-six letters of the “school” or teaching alphabet survived uniformly throughout Etruria.

Beside this loss of signs, there was only one addition to the alphabet. The labiodental fricative /f/ was initially represented by the grapheme cluster *vh* (Greek FH) or *hv* (HF) (out of which Latin *F* was simplified). Towards 600 BC in the north, where there are no local attestations of *vh/hv*, there occurs a sign 8 for *f* (Vn 1.1), which from around 500 BC was in general use in the south too. The origin of this sign, which is also used in sixth- to fourth-century Lydian, has not found a satisfactory explanation; the oldest attestation comes from a Sabellic inscription that dates from the end of the seventh century (Poggio Sommavilla; Rix 1996).

The oldest and latest sequences of the alphabet are contrasted in (1):

- (1) *Archaic school alphabet:* a b c d e v z h  $\theta$  i k l m n š o p  $\sigma$  q r s t u  $\phi$  š  $\chi$   
*Late “working” alphabet:* a c e v z h  $\theta$  i l m n p  $\sigma$  r s t u  $\phi$   $\chi$  f

Note that in the northeast in the fourth to third centuries BC, instead of  $\mathbb{W}$  for /m/ a simpler sign was used that looks like the numeral character for “5”  $\Lambda$ ; it was certainly chosen, because *ma $\chi$* , the word for “5,” begins with /m/.

Of the early archaic texts some are written from left to right and some from right to left. Around 600 BC the direction from right to left became standard and was only reversed occasionally in the first century under Latin influence.

Most archaic texts employ *scriptio continua*; only towards 500 BC does word division become more common. This was normally achieved by the use of dots (one dot or two to three dots in a vertical arrangement); spaces alone occur infrequently. The syllabic punctuation used from 600 to 470 BC in the south, in which letters for vowels at the beginning of a syllable and for consonants at the end of a syllable are furnished with dots, is clearly a school rule borrowed from a syllabic writing system (see Rix 1968) and has limited functional value (see Wachter 1986).

The Etruscan numeral characters have the same shape as the Roman ones derived from them: | “1”; X “10”;  $\wedge$  (Roman V) “5”;  $\uparrow$  (Roman L) “50”; and  $\mathbb{X}/\oplus$  (Roman C) “100”. The principle of “subtraction numerals” is also known from Latin: for example, XIX “19,” to which Etruscan *θun-em zaθrum-s* “- 1 20” corresponds. The numeral X is at one and the same time a symbol for the outstretched fingers of two hands and a letter (š) for the initial sound of Etruscan *sar* “10.” In the latter it is possible to see an echo of the acrophonic numeral system of Greek (Π for πέντε “5” and so on). The system as a whole, however, is autonomous.

### 3. PHONOLOGY

Texts and forms cited in the following discussions can be found via the index in Rix and Meiser 1991. A meaning given in brackets (*zusle* [sacrificial animal]) indicates the semantic class of a lexeme, but this cannot be defined further.

Statements about Etruscan phonetics and phonology are based on the sound values of Etruscan letters in other languages: Greek, Phoenician (the source of Etruscan letters); and Latin, Sabellic, Venetic (for which, conversely, Etruscan letters are the source). Amendments and corrections are supplied by the spelling and spelling variations of Etruscan words; in



addition, typology is sometimes helpful. In a poorly accessible, small-corpus language such as Etruscan, however, many questions, especially concerning phonetics, cannot be answered or at least not explicitly so.

In the following discussion, the Archaic Period of the seventh to sixth centuries BC is described first; where appropriate, phenomena first attested in the Late Period, and occasionally prehistoric phenomena, will be included. Subsequently, general changes in the transition to and within the Late Period are described. Context-sensitive developments of little consequence are only treated (and then on an ad hoc basis) where they have relevance for morphology.

### 3.1 Consonants

#### 3.1.1 Obstruents

The obstruents of Etruscan are phonemically voiceless. In word-initial position they were realised as fortes ([+ tense]) and internally as lenes ([− tense, −/+ son]). Latin transcriptions with *p, t, c, f* at the beginning of a word and *b, d, g* internally lead to this reconstruction (*Pabassa, Tidi, Pergomsna, Fraunal, Noborsinia* for *Papaśa, Titi, Percumsna, Fraucnal, Nufrznei* [personal names]).

The *communis opinio* classifies the Etruscan obstruent phonemes essentially on the basis of the sound values of the corresponding Greek characters:

(2)	Graphemes			Phonemes		
Voiceless stops	<p>	<t>	<c/k/q>	/p/	/t/	/k/
Voiceless aspirated stops	<φ>	<θ>	<χ>	/p <sup>h</sup> /	/t <sup>h</sup> /	/k <sup>h</sup> /
Fricatives	<f>	<s> <(ś)>	<σ> <(ś)>	/f/	/s/	/š/

This model (2) leaves unconsidered <h> for the aspirate /h/ and <z> for the affricate /t<sup>s</sup>/ (which is clarified by spellings such as *rutzs*). Nor does it account for the spelling variants <Ki>/<K> and the complementary distribution of <h> (word-initially) and <χ> (word-internally and word-finally).

The alternative model (3) overcomes these shortcomings, but suffers from meager typological support (see Rix 1984a; Boisson 1991):

(3)	Graphemes			Phonemes		
Unmarked stops	<p>	<t>	<c/k/q>	/p/	/t/	/k/
Fricatives	<f>	<θ> <s>	<σ> <χ/h>	/f/	/θ/s/	/š/ /x/
Palatalized stops	<φ>	<θ>	<z>	/p <sup>y</sup> /	/t <sup>y</sup> /	/t <sup>s</sup> / (<*/k <sup>y</sup> /)

The assumption of palatalized rather than aspirated stops allows the morphologically inexplicable alternation *Larθia* : *Larθa* (Late Etruscan *Larθial* : *Larθal*) in the genitive of the praenomen *Larθ* to be understood as orthographic variation. And under the simple assumption of a (prehistoric) development \*/k<sup>y</sup>/ > /t<sup>s</sup>/, this affricate then fits into the system pattern. The phonetic similarity of aspirated and palatalized sounds makes the use of Greek aspirated stop symbols for palatalized stops understandable.

Beside the undisputed fricatives /f/ (labiodental), /s/ (alveolar) and /š/ (palato-alveolar; spelling variants *huśiur* : *huśur*, orthographic *Larθalīsa* for [-alša]), two further fricatives are herein identified: a velar /x/ and an interdental /θ/, written <χ> and <θ>. The fricative nature of /x/ is suggested by the word-initial variant [h]; and by the palatalization /xwa/ > [jwa] (<va>) of the plural suffix -χva (see §4.2.3.2) following a palatal. Evidence may also be provided by the spelling <χσ> in loanwords which contained [k<sup>(h)</sup>s] originally (Greek



ῥΑλέξανδρος > *Alixsantre*, Proto-Italic *\*louksnā* > Umbrian *\*lōxsnā* > Etruscan *lusχnei* “moon”; Meiser 1986:170f.). There are two arguments for the letter *θ* also representing a fricative: (i) the letter occurs too frequently to be only the spelling of a palatalized stop (<φ> for [pʲ] and <ζ> for [tʰ] < *\*kʲ*/ are much less common); and (ii) the fricative dissimilation /xʷa/ > /kʷa/ following /s/ in the plural ending (§4.2.3.2). That two phonemes can be represented by a single letter is not unparalleled.

### 3.1.2 Sonorants

Etruscan has two nasal and two liquid phonemes; glides occur as allophonic variants of high vowels (see §3.2):

(4)	<i>Nasals</i>	m	n
	<i>Liquids</i>	r	l

Within a syllable, the nasals /m/ and /n/ sometimes join with a preceding vowel to create a nasalized vowel and are consequently no longer written (e.g., *Araθ* = *Aranθ*). In loanwords /-n/ is replaced by Etruscan /-m/; thus, *pruxum* from the Greek accusative προχούν “a vessel for pouring.”

Following the vowel /a/, the liquid /l/ shows a velar variant [ɭ], which is not written in archaic texts: *Larθia* Late Etruscan *Larθial*.

The palatalized sonorants /lʲ/, /rʲ/, /nʲ/, written <l(i)> <r(i)> <n(i)>, which occur infrequently and developed in part from geminates, should perhaps be reconstructed. Such an analysis would account for several disparate phenomena: (i) the umlaut in genitive *clens* and the spelling *cliniaras* (gen. pl.), from *clan* “son”; (ii) the variants *tinā*/*tinia* ([tinʲa]), “Jupiter” (as if from *\*tin-na*, derived from *\*tin* “day”; Cristofani 1997, 212); (iii) Late Etruscan *rasnea* “public” from *\*rasn(a)-na*, derived from *rasna* “people”; and (iv) *Melakre* and *Araθa* as the Etruscan renderings of the Greek names Μελέαγρος and Ἀριάδων.

## 3.2 Vowels

The Etruscan vowel system contains four phonemes:

(5)	/i/	/u/
	/e/	
	/a/	

In Archaic Etruscan, a rounded phonetic realization of /a/ as [ã] is suggested by the orthographic omission of [ɭ] after /a/ (see §3.1.2) in word-final position: for example, *Larθia* (/larʲaɭ/, see §4.2.2.2; Agostiniani 1997).

Etruscan shows the diphthongs /ei/, /ai/, /ui/, and /au/, as seen, for example, in *zuslei* “with (a sacrificial animal),” *Hamaithi* “at Hamae,” *papui* “in [name of a month],” *lavtun* “family.” The diphthong /eu/ appears in Late Etruscan.

Before another vowel, the high vowels /i/ and /u/ are phonetically realized as consonantal allophones – the glides of, for example, *vacil* “then,” *avil* “year,” *ilucve* “on the (festival day),” *iane* “?,” *Hirminaia* [a family name].

No phonemic distinction of vowel length occurs in Etruscan (but see §3.5.2.5); vowels are lengthened phonetically when accented and in word-final position. The realization of nonaccented vowels shows some variation: for example, *mulvanice*/*mulvenece*/*mulvunuke* “gave as a present” (for detailed discussion, see de Simone 1970a:66–70).

### 3.3 Syllable structure

In the Archaic Period the syllabic nucleus was always a vowel. After unaccented vowels underwent syncope (see §3.5.2.4), however, both liquids and nasals could also serve as syllabic nuclei (e.g., *Vestrcna* < *Vestiricina*), as could sibilants in pronouns (e.g., *cs*, *pσl*). An Etruscan syllable can begin with a vowel or with one, two, or three consonants; a syllable can end in a vowel or in one or two consonants. Prehistoric apocope (see §3.5.2.1) and late archaic syncope (see §3.5.2.4) caused many previously open syllables to become closed.

### 3.4 Accent

The Etruscan word accent, not represented orthographically, was in the Archaic Period characterized by strong expiration, which led in the end to the loss of unaccented internal vowels (see §3.5.2.4). In native Etruscan words the accent falls on the initial syllable; however, from their use as enclitics, demonstrative pronouns acquire a generalized final accent (see §4.3.2). Foreign words which were borrowed from languages having phonemic vowel length appear to have carried the accent on the last word-internal long vowel: for example, *Zimite* < *Ziumite* (by syncope) < Greek Διομήδης; Greek γρῦμειᾱ > Etruscan \**crumī-na* > Latin *crūmīna* “money bag.” In other words, Etruscan speakers interpreted word-internal length as an indicator of accent.

### 3.5 Diachronic developments

#### 3.5.1 Consonants

Changes in consonant quality are without exception limited by context or by region. Two such changes may be mentioned here: (i) the change of /f/ to /p/ before liquids or nasals (e.g., *θafna* > *θapna* “cup”; *Θuflθa* > *Θuplθa* [a theonym]); and (ii) the depalatalization of word-final /tʰ/ (deaspiration of /tʰ/? ) in an area of the northeast (e.g., *Larθ*, *zilaθ* > *Lart*, *zilat*; see Rix 1989b:1300–1302). There is also an occasional alternation of the letters used to spell fricatives (aspirates ?) and stops (e.g., *zamθic* ~ *zamtic*, *Preχu* ~ *Precu*), though there is no justification for proposing a free alternation or a suspension of a phonemic opposition next to continuants (*pace* de Simone 1970a:175).

#### 3.5.2 Vowels

Several distinct vowel changes can be identified.

##### 3.5.2.1 Apocope

Inflectional phenomena, also attested for Lemnian and Rhaetic, allow the supposition that in the Proto-Tyrsenic period (see §1) word-final vowels were apocopated due to a penultimate accent (see Prosdocimi 1986:608–616): for example, nominative \**seχi* > *seχ*, beside genitive *seχi-s* (see §3.5.2). Compare the later apocope of the final vowel of the enclitic: Archaic Etruscan *-ca* > Late Etruscan *-c* “and” (see §4.3.2).

##### 3.5.2.2 Vowel lowering

From the beginning of the Late Period, the phonetic realization of vowels is lowered: (i) /u/ as [o]; cf. the Latin name of the Etruscan King *Porsenna* (500 BC) and Etruscan *Purze*; and (ii) /i/ as [e] before /a/ or /e/ in the following syllable, except when occurring after /tʰ/ < z >: *ica* > *eca* “this,” *Θihvarie* > *Θefarie* “Tiberius,” *ci* “3,” *frin* “?”, *zilaθ* “praetor.” Note also a

change which occurs in the quality of /a/: thus, *Luvcie* instead of *Laucie* for Italic *Loukios*. See Agostiniani 1986:27–28.

Beyond the aforementioned lowering of /i/ to [e], intervocalic /i/ is lost (cf. §3.5.2.4), except in the northwest, as in the genitive of female names: Archaic Etruscan *Apucuiā*, Volterrān *Felmuia*, but otherwise *Velua*.

### 3.5.2.3 Vowel raising

Around 400 BC /ai/ becomes /ei/, and in the fourth century /ei/, whatever its origin, becomes /e/ before /u/ and word-finally: for example, *Kaikna* (fifth/fourth century) > *Ceicna* (third century; a family name); *Aivas* (fifth/fourth century) > *Eivas* (fourth/third century) > *Evas* (third century), from Greek Αἴας; Archaic Etruscan *Nuzarnai*, Late Etruscan *Peθnei*, *Peθne* (female family names); in final position /ei/ is for the most part restored by analogy.

### 3.5.2.4 Syncope

Unaccented word-internal vowels disappear between 500 and 470 BC, even in closed and word-final syllables: for example, *turuce* “sacrificed” > *turce*; *Larecena* > *Larcna* (family name); *Scanesna* *Scanasna* > *Scansna* (family name); *Aranθ* > *Arnθ* (praenomen). As a result of this syncope, consonantal sonorants become syllabic between consonants: for example, *Spuriena* ([spuryena]) > *Spurina* (family name), *muluvene* > *mulune* “gives as a gift”; *Leθamsul* > *Leθnsl* (theonym); *vacil vacal* > *vacl* “then.” Syncope is not simply a graphic phenomenon (*pace* Pfiffig 1969:53–63), but a phonetic one. The proof is provided by cases in which an anaptyctic vowel later appears as a secondary consequence of syncope; for example, *Hercele* for *Herclē* < \*Hraklē̃s.

Morphologically relevant vowels are preserved analogically or restored: for example, genitive *Aules* instead of \**Auls* by analogy to the nominative *Aule*; preterite *lupuce* after perfect *lupu* “has died.” A vowel before final -/n/ is not syncopated (e.g., *Turan* “Venus”), because it was nasalized and thereby phonetically lengthened (see §3.1.2). In some cases in which the expected syncope has not occurred, no compelling reason can be given for its absence – as in the /a/ preserved in *zilaθ* “praetor.”

### 3.5.2.5 ê of Cortona

The new text of Cortona (see §1; about 200 BC) has shown that the inverted ∃ <ê>, used only at Cortona, represented a phoneme different from the one written with normal <e>. This /ê/ seems to be recent: some examples go back to diphthongs (clitic -σνê < \*-σναι), others to compensatory lengthening (prenoun *Vêl* < \**Vell* < \**Venl*, syncopated from *Venel*); for some there is no motivation. The rest of Etruria ignored this phenomenon at least in the script (Agostiniani[-Nicosia] 2000: 47–52).

## 4. MORPHOLOGY

### 4.1 Word formation

The usual process of word formation in Etruscan is suffixation. Less commonly, word formation may also be accomplished by, in essence, a phonological modification of morphemes. Less productive still is prefixation. Suffixes can be added both to the *root*, a formant that cannot be analyzed further, and to the *base*, a formant that is already suffixed.

Word-building via apparent phonological modification is commonly the result of phonological processes occurring at morpheme junctures, obscuring the boundaries. For example,

the joining of morphemes may create diphthongs which then undergo monophthongization, as in the locative *zusleve* < *zusleva-i*; compare the nominative *zusleva* (see §4.2.2.3). Less common is distant vowel assimilation, or *umlaut*, as in, for example, genitive *clens* < *\*klanias*; by analogy the ablative is *clen* rather than the expected *\*clan* < *\*krania*, beside nominative *clan* < *\*krania* (cf. gen. pl. *cliniaras*).

A possible Etruscan prefix is *e-* in *eprθnevc* (title of an official) beside *purθne*, *purθ* “first” (?); also in *\*etrs-* (Latin *Etrus-ci*) beside *\*turs-* (Greek *Τυρσ-ηνοί*, Umbrian *Turs-com*, Latin *Tusci*). As the precise meanings of these words are not clear, it is impossible to determine the function of the prefix. The prothetic vowel *e-* in *esl-z* “twice” and *eslem* (“-2” = “8” in numerals), from *zal* “two,” is phonetically motivated.

Typologically, Etruscan is not uniform. Many of its morphological processes are agglutinative. In the noun, for instance, number and case are each marked by their own suffixes: *clan* “son,” genitive *clen-s*, plural *clen-ar*, genitive plural *clinii-ar-as*. Certain cases are not formed from the base, but from the genitive, as with the “pertinentive” *clen-ar-as-i* or the ablative *Arnθ-al-s* (see §4.2.2.4); here the genitive is treated like an adjective.

Other morphological processes, however, are more fusional in nature. These generally result from sound changes which have obscured an agglutinative structure. Thus, locative plural *zusleve* beside nominative plural *zusleva* (from *zusle* [a sacrificial animal]) can be traced to a form *zusle-va-i*, in which the locative suffix *-i* has been added to the plural suffix *-(χ)va-*. The allomorphy *-s/-as/-es/-is/-us/-ls* in the genitive I arose as a consequence of the apocope of final vowels (see §3.5.2.1); earlier this genitive was uniformly characterized by *\*-s* (< *\*-si?*).

The *-s/-l* genitive allomorphy (see §4.2.2.2), in contrast, cannot be a consequence of sound change, but is a morphophonemic phenomenon. *Praenomina* (first names in the Etruscan naming system), in which *-s* and *-l* are distributed according to the final phoneme of the base form, reveal the nature of this allomorphy: following dental obstruents (*/s/*, */θ/*) *-l* occurs, otherwise *-s*; thus, *Laris*–*Larisa*, *Larθ*–*Larθal*; *Aule*–*Aules*, *Vel*–*Velus*. As for appellative pairs such as *cilθ-ś*: *cilθ-l* [locality], *suθi-ś*: *suθi-l* “grave,” no functional difference has yet been distinguished. The distribution seen in family names – such as genitive *Velimna-ś* for men: *Velimna-l* for women – is a relatively late development that came into being around 700 BC with the appearance of the Etruscan system of family nomenclature. The *-s/-l* allomorphy can only have arisen as a result of syncretism, perhaps through the merging of a genitive in *-l(a)* with an ablative in *-s* (see §4.2.2.4), and does not argue against an agglutinative morphology.

## 4.2 Nominal morphology

Both nouns and adjectives are here treated under the rubric of nominal morphology.

### 4.2.1 Gender

Unlike the Indo-European languages with which it was in contact, Etruscan has no grammatical gender (see Fiesel 1922). The female sex is indicated by a suffix, either *-θa*, *-θu*, or *-i*: for example, *lautni* “freedman”: *lautni-θa* “freedwoman”; *Racvu* [man’s name]: *Racu-θu*, *Rakv-i* [women’s names]. The suffix *-i* (< Italic *-ī* < *\*ih<sub>2</sub>-*) was borrowed from Italic and was used under Italic influence with family names that were in origin adjectives: for example, *Tarna-i*.

### 4.2.2 Case

Etruscan nouns and adjectives are marked for case and number (singular and plural; see §4.2.3). The following cases have been identified: nominative-accusative, genitive, locative, ablative, and “pertinentive.”

#### 4.2.2.1 Nominative-accusative

The nominative-accusative is the base form of the nominal paradigm and indicates the subject (*mini zinace Aranθ* “Aranth produced me”); the predicate (*ca suθi* “this [is] the grave”); the direct object (*cn suθi ceriχunce* “he erected this grave”); and the *nominativus pendens*. It is governed by the infrequent postposition *-pi* “?”: for example, *Aritimi-pi* “? Artemis.”

#### 4.2.2.2 Genitive

The *genitive I* is formed with one of the allomorphic suffixes *-s*, *-as*, *-es*, *-is*, *-us*, *-ls* (see Rix 1989a). After vowels *-s* occurs; after consonants no morphophonemic rule is apparent. Following prehistoric apocope (see §3.5.2.1) the original word-final vowel of the base was interpreted as part of the ending and was generalized in a number of semantic groups: *-as* in the *-r*-plurals (see §4.2.3.1); *-us* in individual names (*Velθur-us*, *Θanacvil-us*); *-ls* in the south for multiples of ten and *-uś* in the north (*cealχ-ls* : *cealχ-uś* “30,” syncopated from *\*-χvis*; Lemnian *σialχv-is*). Not belonging to any such semantic groups are, for example, *clen-s* “son,” *meθlum-es* “city,” *seχ-is* “daughter.”

The suffix of the *genitive II* (see Nucciarelli 1975) is *-l* < *\*-la* (see §4.2.2.4), as seen in, for example, *spura-l* “community,” *pui-l* < *\*puia-l* “wife,” *muro-l* “urn,” *culs-l* “gate.” In proper names velar [ɬ] is mostly written *al* (Archaic Etruscan *a*): for example, *Larθi-al*, *Larθi-a*, *Velu-al* < *\*Velui-al*.

The genitive is used to indicate (i) nominal dependency (chiefly possession); (ii) the addressee in dedications (*itun turuce Venel Atelinas Tina-s cliniar-as* “Venel Atelinas dedicated this to the sons of Zeus”) and ordinals (*huθ-iś zaθrum-iś* “the 26th”).

#### 4.2.2.3 Locative

The suffix of the locative is *-i*: Archaic Etruscan *zusle-i* > Late Etruscan *zusle*, plural *zusleve* (< *-e-χva-i*) “with [sacrificial animal]”; *zilc-i* “in the praetorship.” When occurring after a vowel, this *-i* suffix escaped the prehistoric apocope (see §3.5.2.1) and was later extended to base forms ending in a consonant.

The locative indicates (i) sojourn in place and time (e.g., *spure* < *-a-i* “in the community”; *uól-i* “during the day”: *uól*); (ii) motion to a place (e.g., *celi* < *-le-i* “to the earth”); and (iii) instrument (e.g., *turza-i* “with [tool of sacrifice]”).

For the purpose of clarifying syntactic-semantic functions, enclitic postpositions are utilized: *-ri*, indicating a benefactive notion (*meθlumeri* < *-e-i-ri* “for the city”); and *-θi*, *-θ*, *-te*, *-ti*, indicating location (e.g., Archaic Etruscan *Hama-i-θi* “at Hamae”; Late Etruscan *spure-θi* < *-a-i-θi* “in the community”; *velθite* < *-a-i-te* “to the earth”; *lauχumneti* < *-na-i-ti* “in the royal house”). These postpositions can also substitute the locative suffix *-i*: for example, *cela-ti* “in the burial chamber.”

#### 4.2.2.4 Ablative

The ablative occurs in three forms (see Rix 1984a:226–227). The *ablative I* is formed with the suffix *-s* and palatalization of the preceding vowel: for example, Archaic Etruscan *lavtunu-is* “family,” *turza-is* (a sacrificial offering); Late Etruscan *fāse-iś* “porridge,” *Apatru-is*, *Tarnes* < *-na-is*, *Tetnis* < *-nie-is* (family names). The *ablative II* is formed with the suffix *-las* > *-ls*: for example, Archaic Etruscan *Veleθna-las*; Late Etruscan *Visna{fia}-ls* (family name), *Arnθ-als* (praenomen).

It is possible that originally the ablative was formed by the addition of a suffix *-s* to the genitive suffix. In the case of the ablative II, it would have been attached to the

ending *-l* of the genitive II, which, prior to the prehistoric apocope, must have been *\*-la* (cf. §4.2.2.5). In the case of the ablative I, the suffix would have been added to the *-s* of the genitive I, whereupon */ss/#* was shortened with palatalization of the preceding vowel.

The rare *ablative III* has no ending and its morphology is therefore identical with that of the nominative-accusative: for example, *faše* “porridge,” *Ravnθu* (praenomen) (an exception is *clen*, nom.-acc. *clan*; see §4.1). This homomorphy arose through a sound change that we are not able to reconstruct. The combination of the endless ablative III forms with the ablative II suffixed forms (in *-als*; *Tute Arnθals*) has led to the suffix of the latter being incorrectly interpreted as a group inflection.

The ablative expresses (i) the agent in passive constructions (e.g., *anc farθnaχe Tute Arnθals Haθli-als Ravnθu* “which was *-ʔ*-ed by Arnth Tute and Ravnthu Hathli”), (ii) origin (e.g., *paci-als* “[stemming] from Paci”); and (iii) the shared whole (partitive: *šin aiser faše-iš* “take, O gods, from the porridge”). The ablative is governed by the postposition *ceχa* “because of”: for example, *clen ceχa* “because of a son.”

#### 4.2.2.5 Pertinentive

The two constructions of the so-called pertinentive case are likewise based on genitive forms. The *pertinentive I* ends in *-(V)si*, the *pertinentive II* in *-(a)le*. An originally uniform morphology can be hypothesized by proposing that the locative suffix in *-i* (see §4.2.2.3) was added to forms of the two genitives. An original structure *\*(a)la* (see §§4.2.2.2; 4.2.2.4) is proposed for the suffix of the genitive II; the diphthong in *\*(a)la-i* developed prehistorically to *-(al)e*. At times the local postposition *-ti/-θi* (see §4.2.2.3) substituted for the locative suffix *-i*: thus, Archaic Etruscan *Misala-la-ti* “in the [area] of Misala” (with genitive II in *-la!*); *Uni-al-θi*, Late Etruscan *Uni-al-ti* “in the [temple] of Juno.”

The pertinentive often functions simply as a genitival locative: for example, *spureθi apa-s-i* “in the community, in [that] of the father”; *zilci Ceisinie-s-i V(elu-s-i)* “in the praetorship of V. Ceisinie”; *Uni-al-θi* “in the [temple] of Juno.” In several syntactic constructions, however, this use is not obvious. For instance, in *mini Spuriaza [Teiθu]rnas mulvanice Alsaiana-s-i* “Spuriaza Teithurna gave me as a present (into the sphere of =) to Alsaiana,” the pertinentive signifies the addressee (that is, functions as a dative); on the stamp marked *Serturie-s-i* “in [the workshop] of Serturie,” it denotes manufacturer (the agent, that is, it functions as an ablative). Expressions of the type *mi mulu Kavie-s-i* “I [am] a present for/from Gavie” are ambiguous.

### 4.2.3 Number

Etruscan nominals are marked for two numbers, singular and plural; *Tinas cliniaras* “Zeus’ sons” (gen.) does not demonstrate a dual (*pace* Agostiniani 1985; *-ia-* belongs to the stem).

Etruscan has two suffixes for forming the plural: (i) *-r* with the variants *-ar*, *-er*, *-ir*, *-ur*; and (ii) *-χva* with the variants *-cva* and *-va/-ua*. The variants *-ar*, *-er*, *-ir*, *-ur*, like the corresponding variants of the genitive (see §4.1), arose as a consequence of the stem-final vowel, apocopated in the suffixless base form, being preserved (or transferred by analogy) before the suffix. The word endings *-ras* and *-rasi* in the genitive and pertinentive demonstrate that *-ra-* was the original form of the plural suffix. The variants of *-χva* ([x<sup>w</sup>a]) are phonetically conditioned.

The *-r*-plural is predominantly, though not exclusively, used with nominals denoting human referents ([+ hum]). The *-χva*-plural occurs solely with nonhuman referents ([− hum]; see Agostiniani 1993:34–38).



By the side of numerals (Agostiniani 1993:38) the  $-\chi va$ -plural is first used in the Late Period, and its use is not consistent: for example, *zusle-va-c mac* “and five *zusle*-sacrificial animals,” but *avils σas* “of six ( $\sigma a$ ) years.” Otherwise the nominative-accusative or the genitive singular is used: Archaic Etruscan *ci zusle* “three *zusle*-sacrificial animals”; Late Etruscan *muσ-l XX* “20 urns.” The use of the  $-r$ -plural does not show this sort of optionality: thus, Archaic Etruscan *ki aiser* “three gods,” Late Etruscan *ci clenar* “three sons.”

#### 4.2.3.1 The $-r$ -plural

This plural suffix, having the semantic characteristic [+ hum], is used with nouns such as the following (i): *ais-er*, genitive *ais-er-as*; from *ais* “god”; (ii) *clen-ar*, genitive *clinii-ar-as*, pertinensive *clen-ar-as*; from *clan* “son”; *papals-er*; from *papals* “grandchild,” *θanó-ur*; from *θanó* “merciful” (referring to gods). Worthy of note is *tušurθi-r* “married couple,” literally “those on the two cushions,” formed from the locative plural *tuš-ur-θi* “on the cushions.” Among  $-r$ -plural substantives having the semantic characteristics [– hum, – anim] are the following: (i) genitive *tiv-r-s/tiu-r-as*; from *tiu* “month” (gen. *tiv-s* “moon”), (ii) locative *tuš-ur-θi*; locative singular *tuš-θi*; from *tuš-θi* “cushion”; (iii) locative *ramu-r-θi*; locative singular *ramu-e(θ)* [a vessel].

Distributive numerals are formed like  $-r$ -plurals, although they do not necessarily accompany substantives which are [+ hum]: for example, *θu* (stem *θun-*) “one,” in *tun-ur clutiva* “a *cluti*-vessel each” (Pe 5.2); further consider *zel-ur*, from *zal* “two”; *ci-ar*, from *ci* “three.”

In family names and in the formation of collectives  $-(V)r$  is replaced by  $-\theta ur$  (having the original meaning “descendant?”): for example, *heva Marcniθur Pupeinal* “all Marcni [children] of Pupeinei”; *maru paχαθur-as* “priest of the Bacchantes.”

#### 4.2.3.2 $-\chi va$ -plurals

Plurals made with this formant having the semantic characteristics [– hum, – anim] include the following: *caper-χva*, from *caper*, a vessel; *θesn-χva*, from *θesan* “morning, day”; locative *sren-χve*, from *sren* “picture”; *culś-cva*, genitive singular *culś-l* “gate”; *luθ-cva*, from *luθ* “altar”; *hupniva*, from *hupni* “burial couch”; *zuθeva*, from *zuθe*, a cult vessel; *murzua*, from *muσ* “urn.”

Two such plurals show the semantic qualification [– hum, + anim]: (i) *fler-χva* (locative *flerχve*); from *fler* “victim,” which is introduced in a sacrificial prayer as *zivas* “living” and is then *θezine* “to be slaughtered”; and (ii) *zusleva* (locative *zusleve*, ablative *zuśleveś*), from *zusle*, a kind of sacrificial animal. The use of the  $-r$ -plural suffix was consequently not (or no longer) determined by the feature [+ anim], but by [+ hum]. There is no valid example of a  $-\chi va$ -plural with the qualification [+ hum]: *marunuχva* is derived from *marunuχ* “office of a *maru* (a cult official),” not from *marunu* “being *maru*” (Agostiniani 1997:4–9, Maggiani 1998:109–113).

### 4.3 Pronouns

The pronominal paradigm is identical to that of the noun except that the accusative is a separate category, distinct in form from the nominative. The accusative suffix  $-ni$  is only (after /i/?) preserved in Archaic Etruscan *mi-ni* “me,” and before enclitic  $-m$  in the archaic adverb *ita-ni-m* “just as” (< “\*but this”). Otherwise, as a consequence of the prehistoric apocope (see §3.5.2.1), the suffix became  $-n$ . Plural forms are rare; only “articulated” forms are certain: nom. *sani-σva* ([saŋniśwa]), built from *sa(c)ni-σa* (see §5.2) with the plural suffix  $-\chi va$ , gen. *Larisali-óvla* (Cortona, see 1), “pert.” *Larθiali-óvle*.

### 4.3.1 Personal pronouns

The following pronominal lexemes are known:

(6)		<i>First person</i>	<i>Second person</i>
	<i>Nominative</i>	mi	*una
	<i>Accusative</i>	mi-ni	un < *una-n
	<i>Locative</i>		une < *una-i

### 4.3.2 Demonstrative pronouns

There are three demonstrative pronouns in Etruscan, among which *σa* only occurs in enclitic position (see §5.2)

The demonstrative pronouns *ica*, *ita* > *eca*, *eta* (see §3.5.2.2) > *ca*, *ta* are at times used as independent words, usually positioned before those words they determine, and at times as enclitics, fusing with the words they determine (serving as “articles”; see §5.2). The following forms are known (those marked with superscript <sup>i</sup> are only attested as an “article”):

(7)	<i>Archaic</i>	<i>Late</i>		<i>Archaic</i>	<i>Late</i>	
	<i>Nominative</i>	ica, ika-	eca	ca	ita	eta ta
	<i>Accusative</i>	ican, ikan	ecn	cn	itan	etan tn
	<i>Genitive I</i>	- <sup>i</sup> cas	ecs	cś	- <sup>i</sup> tas	etas tś
	<i>Genitive II</i>			cla	- <sup>i</sup> tala, - <sup>i</sup> tula	- <sup>i</sup> tla
	<i>Locative</i>			cei	(tei?)	tei
	<i>Ablative</i>			ceś (cś?)		teiś (?)
	<i>Pertinentive</i>				- <sup>i</sup> tale	- <sup>i</sup> tle
	<i>with</i>		[ecl]θ, eclθi	clθ, clθi	- <sup>i</sup> talte, - <sup>i</sup> tultei	
	<i>postposition</i>					

The final-syllable accent (see §3.4) reveals itself in the preservation of final *-a* in the genitive II, in the syncope of unaccented /a/ in the penultimate syllable (e.g., *-<sup>i</sup>tala* > *-<sup>i</sup>tla*), and in the potential disappearance of the word-initial vowels.

The pertinentive demonstrative is used to designate place and time: for example, *clθ σuθiθ* “in this grave”; Archaic Etruscan *iōve-itule*, Late Etruscan *eōv-itle*, place or time of a ritual. The locative forms are, it seems, only instrumental in sense: e.g., *tesne rasne cei* “according to this state regulation” (?).

Archaic Etruscan *itunia* (< \*ita-n(i)-na), *itu-na*, *eta-na-l*, Late Etruscan *ca-n-l*, *c-n-l* are accusative and genitive II adjectives which are derived from an accusative pronoun by means of a formant *-na*; the meaning seems to be the same as that of the pronoun itself.

### 4.3.3 Relative/interrogative pronoun

A pronoun attested by the forms of (8) functions as an interrogative (*ipas ika-m* “but whose is this?”) and as a relative (see §5.5).

(8)	<i>Nominative</i>	ipa	
	<i>Accusative</i>	inpa	
	<i>Genitive I</i>	ipa-s	Archaic Etruscan
	<i>Genitive II</i>	ipal	Archaic Etruscan
	<i>Genitive II</i>	epl	Late Etruscan
	<i>Locative</i>	ipei, ipe	
	<i>Locative</i>	inpein	Archaic Etruscan
	<i>with postposition</i>	ipe-ri	



This could be a derivative of the relative pronoun *in* (see §4.3.4). On the basis of *in-pa*, interpreted as accusative, a stem *i-pa* could have been abstracted and inflected nominally.

#### 4.3.4 Relative pronouns

The relative pronouns *an* and *in* (also *anc*, *inc* with *-c* “and”) are only attested in nominative and accusative function. Their use is conditioned by the quality of the antecedent: [+hum] requires *an*, [-hum] *in* (Agostiniani and Nicosia 2000:100). The contexts in which the reduplicated *ananc*, *ininc* occur, which (like Latin *quisquis*) could be generalizing, are unclear.

#### 4.3.5 Indefinite pronoun

A pronoun expressing an indefinite quantity (cf. Latin *aliquantus*) is seen in nominative *heva*, accusative *hevn*, genitive *hevl*, *heul* (Steinbauer 1999: 95. 427).

The recently published archaic text *ein θui ara enan* “\*not here do/put anything” contains the accusative *ena-n* of an indefinite pronoun. Its genitive *ena-ś* ‘of anything’ in formulas like *spureri meθlumeric enaś* of the Zagreb mummy text (see §1) declares the authorities *spura* ‘community’ and *meθlum* ‘town’ as not specified for a certain community (Benelli 2001:221).

### 4.4 Verbal morphology

There are fewer attestations of verbal than nominal forms. Thus far, study in this area has been almost exclusively focused on interpreting texts and not on clarification of points of morphology and syntax (but see now Wylin 2000). The following section must therefore be considered highly provisional in nature.

The verb paradigm is of simple structure, characterized by only a single dimension. Verbal categories are not combined with one another, but are each formed directly to the root or the base. Speakers are not designated (i.e., there is no category of person), nor is there a number distinction. The absence of person and number distinction is revealed, for example, by the following pairs:

- (9) A. Turis mi une *ame*  
 “Doris *I am* (= I belong) to you,” beside  
 [t]eurat tanna la rezu *ame*  
 “(The) judge thereby *is* Larth Rezu”;  
 B. mi Araθiale *ziχuxē*  
 “I am *written* from/for Aranth,” beside  
 iχ ca ceχa *ziχuxē*  
 “As this is *written* above”;  
 C. Araθ Spuriana σ[uθ]il *hecece* (see 4.4.1.2)  
 “Aranth Spuriana *set up* the burial construction,” beside  
 Arnθ Larθ Velimnaś Arzneal huθiur θuθi acil *hece* (see 4.4.1.2)  
 “Arnth [and] Larth Velimna, children of Arznei, *set up* grave [and] furnishings”

Thus far, the following verbal categories have been identified: (i) present and preterite tenses, with the latter showing a distinction of active and passive voice; (ii) imperative, subjunctive, and necessitative moods, aside from the indicative. Various verbal nouns are also identified.

Formation of denominative verbs is quite productive. Moreover, many nouns serving as *base forms* (see §4.1) can be analyzed as verbal nouns, derived from simpler verbal forms by the attachment of various suffixes: for example, (i) *-u* (see §4.4.3.1), giving *lup-u* “died,”

*mul-u* “gift,” *ziχ-u* “writer, writing”; and (ii) *-θ* (see §4.4.3.2), providing *trin-θ* “speaking, speaker,” *sval-θ\** “who has lived” (not yet analyzable as verbal nouns are *zilaχ\** > *zilχ* “praetorship,” *acas*, “a sacrifice”). There thus arise whole chains of alternately nominal and verbal derivatives.

The most important denominative suffix is *-ane* (the quality of the vowels is uncertain): thus, *mulu-ane\** “to make a present of,” *ziχu-ane\** “to write,” *acilu-ane\** “to manage, get done,” *acna-ane\** “to make into a possession, get.” The suffix *-ie* (Late Etruscan *-i*), which is frequent in verbal bases, may also be denominative: for example, *vat-ie\** “wish,” *θez-ie\** “slaughtering.”

As there are no personal endings, it is not always easy to distinguish nominal from verbal forms. Roots, that is, monosyllabic segments that (unlike bases) cannot undergo further analysis (e.g., *ziχ* “scratch, write”; *mul\** “give as a gift”; *am* “be”; *men* “make”; *trau* “keep” (?); for additional examples see §4.4.2.1), can be inflected both verbally and nominally. Roots used verbally and their derivatives can only be identified as such (when they can be identified) via the syntax. Nouns can be recognized by the occurrence of case suffixes; yet it appears – unless in the few apparent examples there is chance homonymy – as though case suffixes can also be attached to some typically verbal suffixes, such as the preterite suffixes *-ce* and *-χe*: for example, genitive *tlēna-ce-s*, ablative *tlēna-χe-is*.

#### 4.4.1 Tense and voice

##### 4.4.1.1 Present

Forms of the present, which are rare and not easy to identify, are marked with the suffix *-e*. They express the actual or contextualized present: for example, *ame* “I am,” “he is” (see the examples of [9]); *ale* “gives as a present, places.” With bisyllabic bases, no *-e*-suffix occurs, so that the present is then identical in form with the imperative (see §4.4.2.1): *nunθen* “I call” (as in *un mlaχ nunθen* “you, you good one, I call”). The denominative suffix *-ane*, on the other hand, retains final *-e*: for example, Archaic Etruscan *muluvene* > Late Etruscan *mulune* “makes a present of”; Late Etruscan *acilune* “gets done.”

##### 4.4.1.2 Preterite active

The preterite active, reporting past events, is formed with the suffix *-ce*, which in the Archaic Period was preceded by a vowel, of unpredictable quality, which was later syncopated. At present there is insufficient evidence to determine whether this vowel (*a*, *e*, *i*, or *u*) was originally the root-final vowel which was prehistorically apocopated (see §3.5.2.1) or belonged to the suffix. The following are examples of the preterite active: *amucel/amake* > *amce* “was”; *turucl/turice* > *turce* “sacrificed”; *zinacel/zineke* > *zince* “produced”; *hecece* > *hecce/hece* “erected”; *farice* > *farce* “prepared”; denominative *acasce* > *akśke* “sacrificed”; and with a nasal suffix *amavunice* > *amavence* “produced” (lit. “brought into being”); *ziχ(v)anace* > *ziχunce* “had written” (lit. “brought to writing”); Archaic Etruscan *muluvarice* “gave as a present”; Late Etruscan *ceriχunce* (< *\*cer-ie-χ(e)-u-ana-ce*) “built”; *θezince* (< *\*θez-ie-ana-ce*) “slaughtered”; *zilaχnuce* (< *\*zilaχ-an(a)-u-ce*) “was praetor.”

##### 4.4.1.3 Preterite passive

The suffix of the only recently identified preterite passive is *-χe*. Here too, between roots ending in a consonant and the suffix there occurs one of the four Etruscan vowels, but these vowels are nowhere syncopated (to maximize the distinction between the two phonetically similar suffixes *-ce* and *-χe*?). As with the preterite active, it is impossible to determine whether this vowel originally belonged to the root or to the suffix. Examples of the preterite

passive are the following: Archaic Etruscan *zinaχe* “was produced”; *vatieχe* “was wished for”; Late Etruscan *ziχuxē* “was written”; *menaχe* “was prepared”; denominative *farθnaχe* “was prayed for” (?); and with nasal suffix, *muluanix(e)* “was given as a present.”

The passive character of these forms follows from: (i) the number of participants (in each instance only one in a direct case); (ii) passages in which a pronominal subject in the nominative denotes the patient (the agent is in the ablative; see §4.2.2.4):

- (10) A. *mi titasi cver menaχe*  
       “I was created for/by Tita as a present”  
       B. *inpein . . . mlaχuta ziχuxē*  
       “Which . . . as good (the articulated nominative) was carved”

#### 4.4.2 Mood

In addition to the indicative, Etruscan has an imperative, subjunctive, and necessitative mood.

##### 4.4.2.1 Imperative

The imperative, the mood of strict command, occurring frequently in ritual texts, is identical with the verbal base. Monosyllabic roots provide most of the attested imperatives: for example, *ar* “make,” *al* “give,” *tur* “sacrifice,” *trin* “speak,” *śuθ* “lay,” *heχz* “pour.” The remaining imperatives belong to denominative bases formed with *-en* or *-ie* (Late Etruscan *-i*) or, with “reverse” nasalization (see §3.1.2), *-in*: for example, *nunθen* “invoke”; *θezi*, *θezin* “slaughter”; *uśi*, *mutin*, *firin* “?”

##### 4.4.2.2 Subjunctive

The subjunctive mood, expressing wish, obligation, and futurity, is marked by the suffix *-a*. Consider the following examples:

- (11) A. *mula* “he/you should give as a present”  
       B. *scuna* “he should/will put at (somebody’s) disposal”  
       C. *acasa* “you/he should sacrifice” (denominative)

The subjunctive is also used in subordinate clauses with the conjunction *ipa* “that” (see §5.5). In ritual prescriptions of the Zagreb mummy (see §1), subjunctives alternate with imperative forms: *raχθ tura/tur* “you should sacrifice/sacrifice in fire.”

The subjunctive is also used to express prohibition (see Colonna 1989:345):

- (12) A. *ei . . . ara* “he should not make”  
       B. *ei truta* “he should not injure [by means of an evil look]”

##### 4.4.2.3 Necessitative

In the necessitative, which indicates that an action must be carried out, a suffix *-ri* is added to the base; base-final *-ie* appears in Archaic as *i* (*fani-ri*) and Late *e* (*fane-ri*, *θeze-ri*). The nasal *n* is assimilated to the *r* of this suffix as in, for example, *nunθeri* < \**nunθen-ri* (cf. the assimilation in the *preaenomen Venel* > *Venl-is* > Late Etruscan *Vel*). Examples of necessitatives appear in (13):

- (13) A. *acasri* “X is to be sacrificed” (denominative)  
       B. *perpri* “?”  
       C. *ziχri* “is to be written, carved,” Late Etruscan

- D. *nunθeri* “is to be sacrificed (by invocation)”  
 E. *θezeri* “is to be sacrificed (by slaughter)”

As these examples illustrate, the necessitative has a passive sense. Identification of its voice as passive follows from the same phenomena noted for the preterite passive (see §4.4.1.3): *esvita* . . . *spetri* “the *esvita* (articulated nominative; see §5.2) is to be expiated.”

#### 4.4.3 Verbal nouns

Without an accompanying auxiliary, verbal nouns were used as predicates; these are formed with the suffixes *-u*, *-θ*, and *-as*. Locative verbal nouns in *-e* were used as infinitives.

##### 4.4.3.1 Verbal nouns in *-u*

These function as nouns for results of actions and agent nouns (see §4.4), and they are indifferent to voice. With transitive verbs they can be used both passively (*mul-u* “given as a gift, gift”) and actively (*zic-u* “writer”). They serve as predicates of matrix sentences and designate a state which began in the past and continued over a long period of time, often right up to the present (in this respect, they are reminiscent of the Ancient Greek perfect):

- (14) A. *mi mul-u kaviiesi*  
       “I (am) presented / a present for/from [see §4.2.2.5] Gavius”  
 B. *eθ fan-u lavtn precuś ipa*  
       “Thus (?) has decided the Precu family that . . .”

The difference between this formation, with its stative sense, and the preterite, which records past events, is revealed by sentences such as the following:

- (15) *lupu-ce* (PRETERITE) *munisuleθ* . . . *avils LXX lup-u* (VERBAL NOUN)  
       “He died while holding the . . . -office; dead at the age of 70”

Enlarged verbal stems can also provide the base of verbal nouns in *-u*, the final vowel of these enlarged stems disappearing before the *-u*-suffix: *\*zina-ce + -u > zinaku* “produced, product”; *\*cerie-χe* (cf. *vatiēχe*) + *-u > ceriχu* “having erected,” *\*zilaχ-ane + -u > zilaχnu* “been praetor.”

There is no explanation for the locatives *ten-v-e* and *zilaχn-v-e* which are attested once in the context in which the nominatives *tenu* and *zilaχnu* otherwise occur.

##### 4.4.3.2 Verbal nouns in *-θ*

As predicates, the verbal nouns in *-θ* designate an action that is both current and contemporaneous with another action. They are thus comparable with the present active participles of the Indo-European languages:

- (16) A. *celi ουθ heχś-θ vin(u)m*  
       “Lay on the ground, pouring wine”  
 B. *racθ ουθ nunθen-θ*  
       “Lay on the fire, invoking”

Other examples include: *ar-θ* “making,” *trin-θ* “speaking,” and *zarfne-θ* “?” These verbal nouns constitute a special case of the agent nouns in *-θ* such as *zil-aθ* “praetor”; *tevara-θ > [t]eurat* “judge”; *tesin-θ* [a servant]. The alternative suggestion that the above predicates are imperatives II (so Pfiffig 1969:137) explains neither the distribution (why imperative II in particular?) nor the relationship to the agent nouns.

4.4.3.3 Verbal nouns in *-as*

Verbal nouns formed with the suffix *-as*, occasionally also appearing as *-asa* (without there being any distinguishable difference in function), usually occur as the predicates of embedded sentences, denoting a state completed in the past, and hence correspond to a preterite participle. These are formed directly on the root in rare instances. On occasion, the predicate of the matrix sentence is connected with this verbal noun via a coordinating conjunction (*-c*, *-um*; see §5.4):

- (17) A.  $\text{ra}\chi \dots \text{mena}\acute{\text{s}} \dots \text{mula-}\chi \text{ hu}\acute{\text{o}}\text{l}na \text{ vinum}$   
 “Having prepared fire, you/he ought also to give young wine”  
 B.  $\text{ara}\acute{\text{s}} \theta ui \text{ u}\acute{\text{o}}\text{eti cepen fa}\theta in\text{-um}$   
 “having made a ? here in the ?, but then ? (imperative)”

More frequently, this verbal noun is formed from a base having the denominative suffix *-ane* (see §4.4) or the suffix *-θ* of the present participle (see §4.4.3.2); examples of the type *-ane* + *-as* > *-anas* follow:

- (18) A.  $\text{zelarven-as}$  (< *\*zal-ur-u-ane-as*)  
 “Having doubled” (cf.  $\text{zelur}$  “every two,” see §4.2.3.1)  
 B.  $\text{ra}\chi\theta \quad \acute{\text{o}}ut\text{-anas} \quad celi \quad \acute{\text{o}}u\theta$   
 in the fire having placed on the earth place  
 “Having placed in the fire, place on the earth”  
 C.  $\text{husur} \quad \text{ma}\chi \quad \text{acn-anas} \quad \text{arce} \quad \text{manim}$   
 children five having had he made manim  
 “Having had five children, he made manim” (a taboo expression for “he died”)  
 D.  $\text{papalser} \quad \text{acn-anasa} \quad \text{VI} \quad \text{manim} \quad \text{arce}$   
 grandchildren having had 6 manim he made  
 “Having had six grandchildren, he died”

As examples of verbal nouns formed from bases ending in *-θ*, consider *sval-θas* “having lived”; *trin-θasa* “having spoken” and the following:

- (19) A.  $\text{eslz} \quad \text{zila}\chi n\text{-}\theta as \quad \text{avils} \quad \theta unem \text{ muval}\chi ls \quad \text{lupu}$   
 twice having held the praetorship of year minus one fifty dead  
 “Having twice held the praetorship, he died at the age of forty-nine”  
 B.  $\text{arce} \dots \quad \text{zilc} \quad \text{marunuxva} \quad \text{ten-}\theta as$   
 he made ... presidency marunuxva having held  
 “He [died], having held the presidency of the maru”

The verbal noun in *-as* also expresses contemporaneous action in the instance of *sval-as* “living” (*sval-ce* “lived”), the only such verbal noun formed from a stative verb:

- (20)  $\text{zila}\chi nce \quad \text{spure}\theta i \quad \text{apasi} \quad \text{sval-as}$   
 he held the praetorship in the community in that of his father living  
 “He held the praetorship, [while] living in the community of his father”

The locative in *-as-i* serves as a predicate in an embedded locative absolute clause:

- (21)  $\text{clensi} \quad \text{mule}\theta \quad \text{svalasi} \quad \text{zila}\chi nce$   
 in the sons in the mula living he held the praetorship  
 “While the son lived in the mula, he held the praetorship”

#### 4.4.3.4 Verbal nouns in -e

Verbal noun forms ending in -e, all of them late and therefore open to interpretation as locatives of stems in -e or in -a, function as predicates of embedded sentences with two characteristics: (i) the subjects of matrix and embedded sentences are not identical; and (ii) the verbal nouns lack congruence with another constituent of the embedded sentence (as is the case with the locative absolute). The verbal nouns thus function as infinitives. On the wrappings of the Zagreb mummy, ritual acts are sometimes expressed by a combination of these forms with *acil* (ame) “one ought” (Olzscha 1961:155–173): for example, *ture acil* “one ought to sacrifice”; *neχse acil ame* (VII 14) “one ought to?” Other examples of matrix predicates include *nunθene* “to call,” *zixne* “to write, scratch.” Consider also the following:

- (22)    une...    puθs...    zivaś    fler    θezine...    zati       zatlxne  
          for you    placed    the living    victim    to kill    with the axe    to strike dead  
          “For you... [is] placed... the still living sacrificial animal to kill, to strike dead with  
          the ax”

### 4.5 Numerals

The following cardinal numerals are attested: *θu* (1); *zal* (2); *ci* (3); *σα* (4); *maχ* (5); *huθ* (6); *semφ* (7?); *cezp* (8?); *nurφ* (9?); *śar* (10); *zaθrum* (20); *cialχ* / *cealχ* (30); *sealχ* (40); *muvalχ* (50); *semφalχ* (70); and *cezpālχ* (80). Ordinals identified are as follows: *θunśna* (1st); *cis* (3rd); *huθiś* (6th); *śariś* (10th); and *zaθrumiś* / *zaθrumsna* (20th).

## 5. SYNTAX

### 5.1 Word order

The word order phenomena of Etruscan have not yet been extensively studied (see Pfiffig 1969:207–211; Agostiniani 1982:278–280; Schulze-Thulin 1993). Departure from the unmarked word order occurs often, without any discernible reason. That unmarked word order for phrases with a verbal nucleus is Subject–Object–Verb (SOV):

- (23)    A. Laris    Avle...    cn    suθi    cerixunce  
             Laris    Aule    this    grave    they set up  
             “Laris [and] Aule... set up this grave”  
         B. Velyinei    Śelvanśl    turce  
             Velchinei    to Silvanus    she dedicated  
             “Velchinei dedicated [the statue] to Silvanus”  
         C. ita    tmia...    vatixe       Unialastres  
             this    cult space    was wished for    by Juno herself  
             “This cult space... was wished for by Juno herself”  
         D. ipa    murzua...    ein    heczri  
             that    the urns    not    are to be sprinkled  
             “That the urns... are not to be sprinkled [with libation]”

Not infrequently, however, Object and Verb reverse positions (SVO):

- (24)    Vipia...    turce       Veroenas    cana  
             Vibia    dedicated    to Versena    the statue  
             “Vibia... dedicated the statue to Versena”

Objects which consist of or contain a deictic pronoun regularly appear at the beginning of the sentence (topicalization) and draw the verb after them creating the order Object–Verb–Subject:

- (25) mini mulvanice Mamarce Quθaniies  
 me gave Mamarce Kutanie  
 “Mamarce Kutanie gave me [as a present]”

Typical of a language having basic SOV-structure, Etruscan has postpositions: *-pi* “?” (see §4.2.2.1); *-ri* “for”; *θi*, *-θ*, *-te*, *ti* “in” (see §4.2.2.3); *ceχa* “because of” (see §4.2.2.4).

In nominal phrases, evolutionary developments occur between the Archaic and Late Periods which are consistent with a typological shift from SOV to SVO; this is seen most clearly with modifying numerals. In the Archaic Period the numeral is always placed before the substantive it modifies (e.g., *zal rapa* “two *rapa*-offerings,” *ci avil* “three years,” *huθ zusle* “six victims”); in the Late Period, however, the order is almost always reversed (e.g., *halχza θu* “one little *halχ*-vessel,” *clenar zal* “two sons,” *naper ci* “three *naper* (square measure),” although isolated examples of the earlier order still occur (e.g., *hut naper* and *ci avil*). The attributive genitive (as far as it can be identified) behaves similarly: Archaic Etruscan shows the order Genitive–Noun, as in *Marhies acel* “Marhie’s production”; but Late Etruscan has the order Noun–Genitive, *flerχvetr[-] Neθunsl*, “in the rite of Neptunus,” *luθcva Caθaś* “the altars of Catha.” The same is true of the attributive adjective, for which, however, there are no clear Archaic examples; thus Late Etruscan, with the order Noun–Adjective, provides examples such as: *ziχ neθorac* “text concerning the inspection of the liver,” *aiseraś šeus* “of the ? gods.” Compare, however, Late Etruscan *huślna vinum* “young wine” (Adjective–Noun).

In deictic function, the demonstrative pronoun is always placed before the noun it modifies: Archaic Etruscan *ica tmia* “this cult space,” *etula natinusnal* “of this ?”; Late Etruscan *cn θuθi* “this grave,” *clθi mutnaiθi* “in this sarcophagus.”

## 5.2 Clitics

Demonstrative pronouns can also be used enclitically; they are attached to adjectives and genitival forms, merging with these phonetically, and function essentially as “articles.” The enclitic use of the demonstrative is frequently observed in theonyms such as *Selvans Sanχuneta* “Silvanus, the one belonging to Sancus.” If the modified word ends in a vowel, the resulting diphthong is monophthongized in Late Etruscan (e.g., */e-i/ > /i/*). Consider the following examples: Archaic Etruscan *riθna-ita* “the ?” (nom.), *riθna-itula* (gen.), *riθna-itul-te* (pert. with postposition); Late Etruscan *eśvita* (< *\*iśve-ita*) “the ?” (indicating locality), *eśvitle* (< Archaic *iśve-itule*, pert.). Following final *-s* the initial *i-* of the pronoun disappears with palatalization of the vowel before *-s*: for example, Archaic Etruscan *tameresca* < (*-a<sup>i</sup>s-ka* < *-as-ika*) “the master of the house”; *aθeme-i-s-cas* “?” > Late Etruscan *aθumi(s)cś* (gen.); *θapneśś* (< *-nas-ites*, abl.) “from [the contents of] the goblet.”

In addition to *-ita* and *-ica*, *-sa* is also used as an article, being added to the genitives of personal names and to a few adjectives that refer to persons (e.g., *sacni-sa* “the one dedicated,” that is, a member of a *śacni-ca* “cult brotherhood”). After the word-final velar *-l* of the genitive II, a phonetically motivated *i* appears: for example, *Larθial-i-sa* (gen. *Larθal-i-sa*; pert. pl. *Larθial-i-śvle*) “the [son] of Larth”; *Alfnal-i-sa* “the [son] of Alfnei.” The word-final *-s* of the genitive I and the initial fricative of *-sa* form a geminate cluster, only revealed in Latin transcriptions: for example, *Veluśa* < *-s-sa* (gen. *Veluśa*) “the [son] of



Vel”; *Hanuśa*, Latin *Hanossa* (gen. *Hanuśla*) “the [son] of Hanu,” “articulated” again as *Hanuśliśa* “the [son] of Hanossa.” The double genitives of the type *Larθaliśla*, *Veluśla* are not an absurdness of Etruscan, but quite regular forms.

Apart from these demonstrative pronouns, only the copulative conjunctions *-c* and *-m* (see §5.4) are enclitic.

### 5.3 Agreement

Since neither grammatical gender nor personal endings are found in Etruscan, agreement occurs only in case and number in nominal phrases. Adjectives and pronouns carry no plural marking when they occur immediately next to the substantive which they modify and there is no chance of misconstruing their relationship: for example, *ais-er-aś še-u-ś* “of the ? gods,” *clen-ar sval* “sons, living (= in their lifetime),” *icac heramaś-va* “and these statues.” But if the phonetic distance is greater or there is some possibility of ambiguity, the plural is marked on the adjective: thus, *ais-er śic šeuc* . . . [9 words intervening] . . . *θanś-ur* “gods, ? and ? . . . graceful”; *apac atic sanio-va* “father and mother, members of the cult association” (i.e., both, not just the mother).

Case agreement is marked on both adjectives and pronouns: for example, genitive *aiser-aś še-u-ś* “of the ? gods”; locative *tesne raśne* “with regulation, of the state”; locative + pertinentive (functioning as a locative; see §4.2.2.5) *θaure lautneścle* “on the area, that of the family”; *cl-θi mutna-i-θi* “in this sarcophagus”; ablative III *mex θuta* “with one’s own means.”

### 5.4 Coordination

The coordination of words and sentences can be accomplished using the semantically unmarked conjunction *-c* < *-ca/ka* (see §3.5.2.1) and the weakly adversative conjunction *-m*. The conjunction *-c* can be attached to each member of a coordinated phrase (e.g., *apa-ati-c* “both father and mother”) or only to the final member (e.g., Archaic Etruscan *hecece farice-ca* “set up and prepared”; Late Etruscan *śacnicleri . . . śpureri meθlumeri-c* “for the cult association, the community and the city”).

Asyndetic construction is also not uncommon: *Laris Avle Larisal clenar* “Laris [and] Aule, the sons (pl.) of Laris”; *acilune turune ścune* “gets done, makes over (to someone), puts at (someone’s) disposal.”

The coordinating comparative particle is *iχ* “as”: *etnam iχ matam* “just as earlier”; *eisna iχ flereś crapšti* “a sacrifice as for Flere Crapsti.”

### 5.5 Subordination

Clause embedding is accomplished utilizing (i) verbal noun constructions (verbal nouns, participles, and infinitives; for examples see §§4.4.3.1–4.4.3.4); and (ii) subordinate clauses introduced by pronouns and conjunctions. Embedded clauses can function as subjects, objects, adverbials, or attributives.

The only subordinate clauses introduced by a pronoun which are thus far attested are relative clauses; these function attributively, in part with a pronominal antecedent. Such clauses are introduced with *ipa*, *an*, or *in* (also *anc* and *inc*), all of which appear to function in the same way. In shortened relative clauses without a predicate, only *in* occurs:



- (26) A. Vete... ipa amake apa...  
 Vete who was father...  
 "Vete..., who was the father..."
- B. šulušī θuni šerφue acil ipei... χaśri  
 LOCATIVE INFINITIVE is necessary where NECESSITATIVE
- C. Vel... an cn ουθi ceriχunce  
 Vel who this grave set up  
 "Vel..., who set up this grave"
- D. flere in crapsti  
 divinity which in *crap*  
 "divinity, which [is] in *crap*"
- E. Tins in marle  
 of Jupiter who in *marle*  
 "of Juppiter, who [is] in *marle*"

The relative pronoun can be omitted, as in *flereś crapsti* "of the divinity in the *crap*."

The following subordinating conjunctions have been identified: (i) *ipa* "that" (used with a verb in the subjunctive or necessitative mood) and *iχnac* "as" in object sentences; (ii) *iχ*, *iχnac* in comparative sentences; and (iii) *iχ*, *iχnac*, *nac* ("then" >) "as" in adverbial temporal sentences. Consider the following examples:

- (27) A. tezan fusleri... ipa ama... naper XII  
 ruling to be made that there are *naper* 12  
 "A ruling is to be made, that there are 12 *naper* (unit of square measure)"  
 (contract about a plot of land)
- B. eθ fanu lautn precuś ipa murzua... ein heczri  
 thus established the family of Precu that the urns not to be sprinkled  
 "Thus the family Precu established, that the urns... are not to be sprinkled  
 [with a libation]"
- C. eca sren tva iχnac Hercle Unial clan θrasce  
 this picture shows how Heracles of Juno the son became  
 "This picture (shows?), how Hercules (became?) the son of Juno"
- D. iχ ca ceχα ziχuxe  
 as this above was written  
 "As this was written above"

## 6. LEXICON

The major part of the Etruscan lexicon is native. Some words are also attested in Lemnian or Rhaetic, revealing their origin in Proto-Tyrsenic: for example *zal*, Rhaetic *zal* "2," *maχ*, *sealχls* (gen.), Lemnian *mav*, *σialχvis* "5," "40"; *zinace*, Rhaetic *t'inaxe* "he made"; *avils* (gen.) = Lemnian *avis* "of years."

Within the sphere of trade and crafts, Etruscan borrowed some words from Greek (de Simone 1968), such as the names of vessels (often in the accusative) like *aska* from ἀσκάς; *pruχum* from προχούν (acc.). Also from Greek come *spurta* from σπυρίδα (acc.) "basket"; *elaiva*- from ἐλαίᾱ *F* "oil"; and probably also *φersu* "[demon with] mask" (\**φersu-na* > Latin *persōna*) from πρόσωπα "mask." From Greek there also come several slave names, such as *Tinusi* from Διονύσιος; a few theonyms, for example, *Aplu* from Ἀπόλλων; and many mythic names, like *Aχle* from Ἀχιλλεύς and *Castur* from Κάστωρ.

The existence of only a few Latin loanwords has been demonstrated, such as *cela* from *cella* “small room” or *macstr-* from *magister* “master.” Etruscan *cletram* is from Umbrian *kletram* (acc.) “litter.” Numerous Etruscan personal names, however, come from the Italic languages: for example, *Marce* from *Marcus*, *Crespe* from *Crispus*, *Vuvzies* from Umbrian *Vuvçis* “Lucius.” A good number of theonyms are also of Italic origin: *Menerva* from Latin *Minerva*, *Neθuns* from Umbrian \**Nehtuns* “Neptune.”

The transmission of loanwords from Etruscan into Italic conforms to a similar picture: there are many onomastic borrowings (such as Latin *Aulus* from *Avile*, *Aule*), but few borrowings can be demonstrated in the realm of common nouns (Latin *satelles* “body guard” from *zat[i]laθ*). The sociological and cultic contacts between Etruscans and the Italic peoples seem clearly to have been more intimate than their linguistic contacts.

## Bibliography

- Agostiniani, L. 1982. *Le iscrizioni parlanti dell'Italia antica*. Florence: Leo S. Olschki.
- . 1985. “La sequenza *tinās cliniaras* e la categoria del numero in etrusco.” In *Studi linguistici e filologici per Carlo Alberto Mastrelli*, pp. 13–19. Pisa: Pacini.
- . 1986. “Sull’etrusco della stele di Lemno e su alcuni aspetti del consonantismo etrusco.” *Archivio glottologico italiano* 71:15–46.
- . 1993. “La considerazione tipologica nello studio dell’etrusco.” *Incontri linguistici* 16:23–44.
- . 1997. “Considerazioni linguistiche su alcuni aspetti della terminologia magistrale etrusca.” In R. Ambrosini (ed.), *Scribthair a ainm n-ogaim: Scritti in memoria di Enrico Campanile*, pp. 1–16. Pisa: Pacini.
- Agostiniani, L. and F. Nicosia. 2000. *Tabula Cortonensis*. Rome: Bretschneider.
- Beekes, R. and L. van der Meer. 1991. *De etrusken spreken*. Muiderberg: Coutinho.
- Benelli, E. 2001. “Quattro nuove iscrizioni etrusche arcaiche dall’agro chiusino.” *Studi etruschi* 64:213–234.
- Boisson, C. 1991. “Note typologique sur le système des occlusives en etrusque.” *Studi etruschi* 56:175–187.
- Bonfante, G. and L. Bonfante. 1983. *The Etruscan Language. An Introduction*. New York: New York University Press.
- Caffarello, N. 1975. *Avviamento allo studio della lingua etrusca*. Florence: Leo S. Olschki.
- Colonna, G. 1975. “Al proposito del morfema etrusco *-si*.” In *Archaeologia: Scritti in onore di Aldo Neppi Modona*, pp. 156–171. Florence: Leo S. Olschki.
- . 1989. Contribution to “Rivista di epigrafia etrusca.” *Studi etruschi* 55:273–353.
- Cristofani, M. 1971. “Sul morfema etrusco *-als*.” *Archivio glottologico italiano* 56:38–42.
- . 1972. “Sull’origine e la diffusione dell’alfabeto etrusco.” In H. Temporini (ed.), *Aufstieg und Niedergang der römischen Welt*, vol. I, part 2, pp. 466–489.
- . 1991. *Introduzione allo studio dell’etrusco*. Nuova edizione interamente aggiornata. Florence: Leo S. Olschki.
- De Simone, C. 1968. *Die griechischen Entlehnungen im Etruskischen*, vol. I. Wiesbaden: Otto Harrassowitz.
- . 1970a. *Die griechischen Entlehnungen im Etruskischen*, vol. II. Wiesbaden: Otto Harrassowitz.
- . 1970b. “I morfemi etruschi *-ce (-ke)* e *-χe*.” *Studi etruschi* 38:115–139.
- . 1997. “Masculin/féminin dans la théonymie étrusque,” In F. Gautier and D. Briquel (eds.), *Les Étrusques les plus religieux des hommes*. Paris: La documentation française.
- Fiesel, E. 1922. *Das grammatische Geschlecht im Etruskischen*. Göttingen: Vandenhoeck and Ruprecht.
- Maggiani, A. 1998. “Appunti sulle magistrature etrusche.” *Studi etruschi* 62:95–138.
- Meiser, G. 1986. *Lautgeschichte der umbrischen Sprache*. Innsbruck: Institut für Sprachwissenschaft der Universität.
- Nucciarelli, F. 1975. “I genitivi etruschi in *-l*, *-us*.” *Annali della Facoltà di lettere e di filosofia della Università degli Studi di Perugia* 12:1–26.
- Olzscha, K. 1961. “Etruskisch *acil*.” *Studi etruschi* 29:155–173.
- Pallottino, M. 1978. *La langue étrusque. Problèmes et perspectives*. Translated by J. Heurgon. Paris: Société d’Édition Les Belles Lettres.

- . 1988. *Etruskologie: Geschichte und Kultur der Etrusker*. Translated by S. Steingraber. Basle/Boston/Berlin: Birkhäuser.
- Pfiffig, A. 1969. *Die etruskische Sprache. Versuch einer Gesamtdarstellung*. Graz: Akademische Druck- und Verlagsanstalt.
- Prosdocimi, A. 1986. "Sull'accento latino e italico." In A. Etter (ed.), *o-o-pe-ro-si. Festschrift für Ernst Risch zum 75. Geburtstag*, pp. 601–618. Berlin/New York: Walter de Gruyter.
- Rix, H. 1963. *Das etruskische Cognomen*. Wiesbaden: Otto Harrassowitz.
- . 1968. "Zur etruskischen Silbenpunktierung." *Münchener Studien zur Sprachwissenschaft* 23:85–104.
- . 1975. Review of Pfiffig 1969. *Göttingische Gelehrte Anzeigen* 227:117–143.
- . 1983. "Norme e variazioni nell'ortografia etrusca." *AIΩN, sezione linguistica* 5:127–140.
- . 1984a. "La scrittura e la lingua." In M. Cristofani (ed.), *Gli etruschi. Una nuova immagine*, pp. 210–238. Florence: Giunti Martello.
- . 1984b. "Etr. mex rasnal = lat. res publica." In *Studi di antichità in onore di Guglielmo Maetzke*, pp. 455–468. Roma: Bretschneider.
- . 1989a. "Zur Morphostruktural des etruskischen s-Genetivs." *Studi etruschi* 55:169–193.
- . 1989b. "Per una grammatica storica dell'etrusco." In *Secondo Congresso Internazionale Etrusco, Firenze 26 Maggio–2 Giugno 1985*, vol. 3, pp. 1293–1306. Roma: Bretschneider.
- . 1996. "Il testo paleoumbro di Poggio Sommavilla." *Studi etruschi* 61:233–246.
- . 1998. *Rätisch und Etruskisch*. Innsbruck: Institut für Sprachwissenschaft der Universität.
- Rix, H. and G. Meiser. 1991. *Etruskische Texte. Editio minor*. Vol. I, *Einleitung, Konkordanz, Indices*; vol. II, *Texte*. Tübingen: Narr.
- Schulze-Thulin, B. 1993. "Zur Wortstellung im Etruskischen." *Studi etruschi* 58:177–195.
- Schumacher, S. 1992. *Die rätischen Inschriften. Geschichte und heutiger Stand der Forschung*. Innsbruck: Institut für Sprachwissenschaft der Universität.
- Steinbauer, D. 1999. *Neues Handbuch des Etruskischen*. St. Katharinen: Scripta Mercaturae Verlag [appeared after completion of the manuscript].
- Wachter, R. 1986. "Die etruskische und venetische Silbenpunktierung." *Museum Helveticum* 43:111–126.
- Woodard, R. 1997. *Greek Writing from Knossos to Homer*. Oxford: Oxford University Press.
- Wylin, K. 2000. *Il verbo etrusco. Ricerca morfosintattica delle forme usate in funzione verbale*. Roma: Bretschneider. [appeared after completion of the manuscript]

# Early Georgian

KEVIN TUI TE

## 1. HISTORICAL AND CULTURAL CONTEXTS

Georgian is a member of the Kartvelian family, one of the three indigenous Caucasian language families. Its sister languages are Mingrelian and Laz, two closely related languages spoken in western Georgia and northeast Turkey, and Svan, spoken in the highlands of northwest Georgia. There has been much speculation about the relation of Kartvelian to other language families. Typological similarities with Indo-European and Afro-Asiatic and an impressive number of vocabulary items which appear to be shared with these families have led some linguists to include Kartvelian as a peripheral member of the so-called Nostratic macrofamily, a phyletic grouping encompassing many of the principal Eurasiatic language groups. Even if the Nostratic hypothesis is not correct, the Kartvelian languages show the imprint of long-standing contact with Indo-European and Semitic speech communities, going back four thousand years or more. Most specialists locate the Proto-Kartvelian speech community either in or somewhat to the south of modern-day Georgia.

Ancient Near Eastern documents as early as the twelfth–eleventh centuries BC mention tribal groups which are likely to have included Proto-Georgian speakers. The first clear indications of Georgian political entities date from the seventh–sixth centuries BC, by which time Greek colonies are installed in Colchis, on the east coast of the Black Sea, and much of Transcaucasia and Asia Minor is under Persian domination. The two major early Georgian kingdoms – Colchis in the west and Iberia in the east – began to consolidate at this time.

During this period the Aramaic language, the *lingua franca* of the far-flung Persian Empire, was adopted as a medium for written communication in Georgia, as attested in inscriptions in the period preceding the introduction of Christianity. The adoption of Georgian as a written language is usually seen as a consequence of the conversion of the elite to Christianity in the middle of the fourth century.

The oldest Georgian monuments are written in well-formed letters, and the *karagmebi*, abbreviations of common words and sacred names, show considerable uniformity from the earliest texts onward: for example, ო~ო (*oüpalo* “Lord”), š~e (*šeiç’q’ale* “have mercy [on somebody]!”). It is evident that the new literary language built upon an already well-established cultural infrastructure, appropriating the functions, conventions, and status of both the written language of pagan Georgia (Aramaic) and the new state religion (Greek, Armenian, and Syriac).

For the purposes of this chapter, we will employ the following periodization of the Georgian literary language:

- |     |                                     |                     |
|-----|-------------------------------------|---------------------|
| (1) | <i>Early Old Georgian (EG):</i>     | 5th–8th centuries   |
|     | <i>Classical Old Georgian (COG)</i> | 9th–11th centuries  |
|     | <i>Middle Georgian:</i>             | 12th–18th centuries |
|     | <i>Modern Georgian (ModG):</i>      | 18th–20th centuries |

The Early Old Georgian corpus contains eight manuscript texts (all but one of them palimpsests) and about a dozen inscriptions; altogether, it would fill a book of little over two hundred pages. Two dialects are represented in these materials, known to scholars as *Xanmet'i* “superfluous *x*'s” and *Haemet'i* “superfluous *h*'s.” The first term was coined by the tenth-century translator Giorgi Mtac'mideli, and reflects the most salient feature of these texts from the perspective of a Classical Old Georgian speaker: a second-person subject (S2) and third-person object (O3) prefix *x-*, where the Classical language has *h-*, *s-*, or zero. The two *Haemet'i* texts make consistent use of *h-* in these contexts. Consider the first words of Matthew 17:4 in three early translations (on the agglutinative morphology of the verb, see §§4.3; 4.3.3, for a list of abbreviations specific to this chapter, see §6):

- |     |                           |   |
|-----|---------------------------|---|
| (2) | <i>Xanmet'i</i> (c. 500)  | mi= <b>x</b> -u-g-o p'et're da <b>x</b> -rkw-a iesu-s                           |
|     | <i>Haemet'i</i> (c. 750)  | mi= <b>h</b> -u-g-o p'et're da <b>h</b> -rkw-a iesu-s                           |
|     | <i>Hadish</i> (897) (COG) | mi= <b>Ø</b> -u-g-o p'et're da <b>h</b> -rkw-a iesu-s                           |
|     |                           | (Pv- <b>O3</b> -ObVN-answer-S3sg. Peter and <b>O3</b> -say-S3sg.<br>Jesus-DAT.) |
|     |                           | “Peter answered and said to Jesus”  |

The retention of two verb forms with S2 *x-* in all known varieties of Georgian implies that the *Xanmet'i* dialect is especially conservative in this respect. Most *Xanmet'i* texts come from eastern Georgia, and the single *Haemet'i* inscription is in the west. While the two dialects doubtless derive from distinct varieties of spoken Georgian, this by no means implies that they corresponded closely to the Georgian spoken by the individual scribes who produced the documents in our corpus. There is evidence of diglossia as early as the Cambridge fragments of Jeremiah, c. AD 600. In what is otherwise a solidly *Xanmet'i* text, three verbs have O3 prefixes in *h-* rather than *x-*, a *lapsus calami* indicative of a *Haemet'i*-speaking monk copying from a *Xanmet'i* original. The only extended *Haemet'i* text, the eighth-century lectionary fragments, appears to have been composed by a grammatically unsophisticated scribe who already spoke a dialect similar to Classical Old Georgian, to judge by the extremely high error ratio: the *h*-prefix is missing in fully 36 percent of the verbs where it ought to appear (see Sarjeladze 1971:18).

## 2. WRITING SYSTEM

The Early Old Georgian documents are written in the alphabetic script known as *mrglovani* (“rounded”) or *asomtavruli* (“capital letters”), the oldest of the three Georgian scripts. Asomtavruli writing was used throughout the Old Georgian period, but with formal changes which enable paleographers to arrive at an approximate dating of manuscripts almost at first glance. In the earliest manuscripts and inscriptions the letters are well-rounded and wider than those in later documents, and the top part of the letters *b*, *q*, and *u* forms a closed loop.

**Table 40.1 The Early Georgian Asomtavruli script with numerical values**

Character	Transcription	Numerical value	Character	Transcription	Numerical value
Ⴀ	a	1	Ⴁ	r	100
Ⴂ	b	2	Ⴃ	s	200
Ⴄ	g	3	Ⴅ	t'	300
Ⴆ	d	4	Ⴇ	ü	400
Ⴈ	e	5	Ⴉ	p	500
Ⴊ	v	6	Ⴋ	k	600
Ⴌ	z	7	Ⴍ	y	700
Ⴎ	ê(ey)	8	Ⴏ	q'	800
Ⴐ	t	9	Ⴑ	š	900
Ⴒ	i	10	Ⴓ	č	1,000
Ⴔ	k'	20	Ⴕ	c	2,000
Ⴖ	l	30	Ⴗ	j	3,000
Ⴘ	m	40	Ⴙ	c'	4,000
Ⴚ	n	50	Ⴛ	č'	5,000
Ⴜ	y	60	Ⴝ	x	6,000
Ⴟ	o	70	Ⴐ	q	7,000
Ⴒ	p'	80	Ⴓ	ǰ	8,000
Ⴔ	ž	90	Ⴕ	h	9,000

The later Georgian scripts, known as *nusxa-xucuri* (“ecclesiastic minuscule”) and *mxedruli* (“knightly,” i.e., “secular”), evolved from the asomtavruli alphabet in the course of the Classical period.

In terms of its time of creation, relationship to the Greek alphabet, and general morphology, the Georgian asomtavruli script forms a group with the other two early Christian Transcaucasian alphabets: the Armenian and the Caucasian Albanian. All three incorporate the Greek letter order, but without the straightforward appropriation of Greek characters that marked the creation of most Greek-based alphabets. Except for a handful of cases, the letters of the Armenian and Georgian alphabets are either entirely new creations, or radical transformations of Greek characters. The creator (or creators) of the Georgian alphabet placed the additional characters needed for the phonemes lacking a Greek equivalent at the end, after *k*, the equivalent of Greek *chi* (*k<sup>h</sup>*). Other Early Georgian grammatological features calqued on the Greek model include the creation of an equivalent to *eta* (it represents the diphthong /ey/), and the use of an *ou* digraph to represent the phonemes /u/ or /w/ (e.g., *čwen* “we, us,” spelled *čouēn* in Old Georgian).

**Table 40.2 The Early Georgian consonants**

	Stops and affricates			Fricatives		Nasals	Glides and Liquids
	Voiced	Aspirated	Ejective	Voiced	Voiceless		
(Bi-) Labial	b	p	p'	v	—	m	w
Dental	d	t	t'			n	
Alveolar	j [dz]	c [ts]	c' [ts']	z	s		r l
Palato-alveolar	ʃ [dʒ]	č [tʃ]	č' [tʃ']	ʒ [ʒ]	š [ʃ]		y [j]
Velar	g	k	k'				
Uvular	—	q	q'	ɣ [ʁ]	x [χ]		
Glottal					h		

### 3. PHONOLOGY

#### 3.1 Phonemic inventory

As Caucasian languages go, Georgian has a fairly restrained phonemic inventory. The stops and affricates come in triplets (voiced, voiceless aspirated, and voiceless ejective, i.e., glottalized), and the fricatives in pairs (voiced and voiceless). There are five vowels, without any distinction of length: /a/, /e/, /i/, /o/, /u/. The consonants of Xanmet'i Georgian are listed in Table 40.2. The transcription used here is similar to those employed by most Caucasologists and Armenologists. (International Phonetic Alphabet equivalents are included in square brackets.)

#### 3.2 Allophonic variation

The only allophonic alternations of note in Early Georgian are between [w] and [v], and between [i] and [y]. In general, [w] is employed immediately after obstruents and [v] in other contexts (e.g., in representations of the first-person subject prefix: *v-i-c-i* “I know something,” but *x-w-e-ji-eb* “I seek something”). The glide [y] only appears as the non-syllabic alternant of [i] after vowels, when the latter is the initial phoneme of a case suffix.

#### 3.3 Phonotaxis

Although Early Georgian words can contain daunting sequences of consonants, for example, *msxwerp'l*- “victim,” *xtn̄da* “(s)he liked it,” the structure of lexemes is constrained by phonotactic rules. Many groups of consonants represent so-called harmonic clusters, found in all Kartvelian languages. These consist of an anterior stop, affricate, or fricative followed by a posterior (velar or uvular) consonant, other than /q/, sharing the same voice-onset features; some examples are: *ʃɣola*- “leading,” *c'q'al*- “water,” *sxwa*- “other.” A harmonic cluster functions phonotactically as a single consonant. There is also a class of “nonharmonic clusters,” which are the mirror image of harmonic groups: back consonants precede front, and the voice-onset features are different, for example, *k'bil*- “tooth,” *č'de*- “notch.” Sonorants, especially /m/ and /r/, can precede or follow consonants or clusters within the same lexemes.

**Table 40.3 Declension of *mojγw[a]r*–“leader”**

	Short	Long		
		Singular	n-/t-plural	eb-plural
<i>ABS.</i>	<i>mojγwar-Ø</i>	<i>mojγwar-i</i>	<i>mojγwar-n-i</i>	<i>mojγwr-eb-i</i>
<i>ERG.</i>	<i>mojγwar-man</i>	<i>mojγwar-man</i>	<i>mojγwar-t-a</i>	<i>mojγwr-eb-man</i>
<i>DAT.</i>	<i>mojγwar-s</i>	<i>mojγwar-s-a</i>	<i>mojγwar-t-a</i>	<i>mojγwr-eb-s-a</i>
<i>GEN.</i>	<i>mojγwr-is</i>	<i>mojγwr-is-a</i>	<i>mojγwar-t-a</i>	<i>mojγwr-eb-is-a</i>
<i>INSTR.</i>	<i>mojγwr-it</i>	<i>mojγwr-it-a</i>	( <i>mojγwr-it-a</i> )	<i>mojγwr-eb-it-a</i>
<i>ADV.</i>	<i>mojγwr-ad</i>	<i>mojγwr-ad</i>	( <i>mojγwr-ad</i> )	<i>mojγwr-eb-ad</i>
<i>VOC.</i>	—	<i>mojγwar-o</i>	<i>mojγwar-n-o</i>	<i>mojγwr-eb-o</i>

## 4. MORPHOLOGY

### 4.1 Word structure

The typical Kartvelian nominal root is monosyllabic, with the most common shape being XVX or XV (where X = a single consonant or cluster, optionally preceded and/or followed by sonorants; see §3.3): for example, *mc'q'ems*- “shepherd,” *qorc*- “flesh,” *t'ba*- “lake.” Verbal roots can be either monosyllabic or nonsyllabic, some of the latter comprising no more than a single consonant: *-k'ɾjal*- “forbid,” *-c'q'*- “begin,” *-γ*- “receive.” Vowel-initial roots are less common, and tend to be limited to deictics and pronouns, numerals, and words of foreign origin.

### 4.2 Nominal morphology

The Early Georgian common noun is declined for seven cases (absolutive, ergative, dative, genitive, instrumental, adverbial, and vocative) and two numbers (singular and plural). Many noun stems, in particular those with a final syllable containing the vowels /a/ or /e/ followed by an approximant (e.g., *mojγwar*-), undergo syncope of the vowel when the stem is followed by a declensional morpheme of the configuration -VC- (e.g., *mojγwr-is*). Undoubtedly, at one time syncope was automatically conditioned by stress placement or perhaps vowel length; by the earliest texts, however, it was no longer predictable. The full declensional paradigm of a syncope-taking common noun is given here (on the short and long case forms, see §4.2.1.2).

The declension of vowel-final stems is slightly more complicated. As a general rule, the relative strength of vowels when two of them come into contact across a morpheme boundary follows the hierarchy: *o*, *u* > *i* > *e* > *a*. For example, a suffix beginning in /-i/ added to a stem ending in /a-/ or /e-/ will cause the latter to drop, whereas the same /-i/ will change to /-y/ when preceded by a stem-final /o-/ or /u-/: for example, *kwa* + *is* > *kwis* “stone-GEN.”; *xuro* + *is* > *xuroys* “carpenter-GEN.” The long absolutive suffix /-i/, however, always becomes /-y/ when added to a vowel-final noun (e.g., *kwa* + *i* > *kway* “stone-ABS.”). When two vowels of the set /e/, /o/, /u/ meet, both are expressed without reduction or loss: *sarc'muno* + *o* > *sarc'munoo* “faithful-VOC.”



### 4.2.1 Nominal cases

In the present section each of the seven cases is discussed, as well as the long and short case form distinction

#### 4.2.1.1 Absolutive and ergative cases

Early Georgian was a language of the split-ergative type, with ergative-absolutive alignment in certain circumstances, and nominative-accusative alignment in others. The Series II verb forms, marking perfective aspect, assign absolutive case to the subjects of intransitive verbs and the direct objects of transitive verbs. The ergative case is assigned to the subjects of transitive verbs. The imperfective Series I forms, by contrast, assign absolutive case to both transitive and intransitive subjects, and mark the direct object in the dative case.

#### 4.2.1.2 Long and short case forms

The formal and functional distinction between the long and short forms of the cases has received extensive study. To summarize it briefly, the absolutive and ergative endings, and the vowel /-a/ added to the dative, genitive, and instrumental, derive from postposed demonstratives used as direct articles (as we shall see, this process occurred a second time in the history of Georgian). The attested distribution of short and long absolutive noun phrases reflects a no-longer productive indefinite/definite opposition in the nominal system. The principal uses of the bare-stem absolutive are in (i) predicate nominals (*tkwen xq'avt igi kwab-Ø avazak'ta* [Lk 19:46] “you made it **a den** of thieves”); (ii) naming constructions (*romelsa hrkwian betlem-Ø* [Lk 2:4] “which they call **Bethlehem**”); (iii) time and distance expressions, especially when quantified by numerals (*xiq'o mun ormeoc-Ø d'ye-Ø* [Mk 1:13] “he was there **forty days**”); (iv) compound verbs incorporating a noun stem with generic reference (*nu k'ac=k'av* [Mk 10:19] “thou shalt not kill,” literally: “thou shalt not **person**=kill”).

#### 4.2.1.3 Dative case

This case has the widest range of functions. It is assigned to indirect objects, and to the direct objects of Series I verbs. A large number of verbs, mostly statives and passives, assign dative case and indirect-object status to their subjects. As would be expected, these are primarily verbs of sensation (*ma-s x-c'q'ur-i-s* “(s)he-DAT. is thirsty”), of emotion (*ma-s x-u-q'war-s* “(s)he-DAT. loves somebody”), and of possession (*ma-s x-u-c* “(s)he-DAT. has something”). The dative also appears in time and place expressions: *ma-s žam-sa xrkwa iesu* (Mk 3:3) “**At that time**-DAT. Jesus said”; *xiq'o igi ierusalēm-s* (Jn 2:23) “He was **in Jerusalem**-DAT.”

#### 4.2.1.4 Genitive case

The Early Georgian genitive signals a fairly broad range of relationships between nouns: possession, membership, kinship, substance, and so forth. The genitive optionally marks certain argument–verb dependencies when these are nominalized (*xicit nič'-isa k'etil-isa micemay* [Mt 7:11] “you know how to give good things,” lit. “you know the giving **of good things**”), though nonfinite verbs can alternatively assign the same cases as their finite counterparts. The long-form genitive can also indicate motion toward a person, rather like Greek πρὸς + accusative (e.g., *movida iesu-ysa* [Mt 14:29] “he came **toward Jesus**”). The short genitive occurs in compounds (*mγdel-t mojywarni* [Mt 27:62] “chief priests,” lit. “leaders **of the priests**”), and in certain adverbial expressions with a quantifier (*sam gz-is* [Mk 14:30] “three **times**”).

#### 4.2.1.5 Instrumental case

This case marks a wide range of instruments, means, or accessories (*šemosili samosl-ita sp'et'ak'-ita* [Mk 16:5] “dressed **in white garments**”). The short instrumental marks the place from which motion occurs, a usage which opposes it to the allative sense of the adverbial case: *iesu mosrul ars huriast'an-it galilea-d* (Jn 4:47) “Jesus has come **from Judea** (instr.) to Galilee (adv.).”

#### 4.2.1.6 Adverbial case

In addition to the allative function mentioned immediately above, this case is employed to derive adverbial expressions from adjectives and nouns (*brc'q'invale-d* “splendid-ly”). The adverbial case of the verbal noun functions like an Indo-European infinitive (*ic'q'o gamosxma-d romelni xq'iddes t'redebsa* [Lk 19:45] “he began **to expel** those who were selling doves”).

#### 4.2.1.7 Vocative case

This case is believed to be of more recent origin than the other six. Titles and common nouns take the vocative in -o. Proper names are rarely used in direct address in the Early Georgian corpus, but when they are, they are in the bare-stem form (*c'inac'armet'q'wel-o davit*, *gwitxar* [Mrv. 4.3] “Prophet (voc.) **David**, tell us . . .”).

### 4.2.2 Plural marking

Early Georgian has two structurally distinct means of marking nominal plurality. By far the most frequently used is the synthetic *n-/t-* plural declension. The *n*-rectus-plural suffix is limited to the absolutive and vocative, and may be historically related to the plural absolutive suffix of the verb (see §4.3.3 [9]). The single oblique plural morpheme *-t-(a)* can represent the dative, genitive, or ergative cases; the instrumental and adverbial do not appear to have had distinct plural forms in this declension (cf. the instrumental with plural reference in Mt 15:8: *eri ese bag-ita mat-ita p'at'iv mcems* “these people honor me **with their lip(s)**”).

The agglutinative *eb-* plural suffix, followed by the case endings of the singular declension, appears only a couple of dozen times in the Early Georgian corpus, sometimes in conjunction with *n-/t-* plural nouns: *brm-eb-i da q'ruv-n-i* (Mt 15:30) “the blind (*eb*-plural) and the deaf (*n*-plural).” While there is no evidence of a semantic distinction between the two plural morphemes in Early Georgian, only *n*-plural nouns can control plural agreement in the verb and within the noun phrase, whereas *eb*-plurals are syntactically singular: *rabami kw-eb-i ars* “what large stones there are (lit. **is**)” (Mk 13:1).

### 4.2.3 Definite articles

In what appears to be a renewal of the prehistoric means of signaling this category, demonstrative pronouns placed after the first word of the noun phrase serve to indicate definiteness. Broadly speaking, the Early Georgian definite article functions similarly to its French and English counterparts. In the episode of the healing of the man with the withered hand (Mk 3:1–5), for example, the protagonist and his hand are first introduced through indefinite nouns: *da xiq'o mun k'aci romelsa qeli ganqmel xedga* (Mk 3:1) “And there was **a man** who had **a withered hand**.” Further on in the story, when they are mentioned again, the definite articles are employed: *da xrkwa k'acsa mas: ganiratz qeli šeni . . . da k'walad moxego qeli igi* (Mk 3:5) “And he said **to the man**: Stick out your hand . . . and thereupon **the hand** was restored to him.”

#### 4.2.4 Pronouns

##### 4.2.4.1 Personal pronouns and proper names

First- and second-person pronouns, the personal relative/interrogative pronoun *vi-n* “who,” and proper names do not have a distinct ergative case form. In addition, the first- and second-person pronouns lack distinct dative and vocative forms as well, using the bare stem in these contexts:

(3)	1st sg.	1st pl.	2nd sg.	2nd pl.	vin	Proper names
ABS./VOC./ ERG.	me	čwen	šen	tkwen	vi-n “who”	iesu “Jesus”
DAT.	me	čwen	šen	tkwen	vi-s	iesu-s
GEN.	čem-i	čwen-i	šen-i	tkwen-i	vi-s-(a)	iesu-ys-(a)
ADV.	čem-da	čwen-da	šen-da	tkwen-da	—	iesu-d
INSTR.	čem-it-(a)	čwen-it-(a)	šen-it-(a)	tkwen-it-(a)	—	iesu-yt

The genitive-case stem of the personal pronouns serves as a base for possessive adjectives: for example, *mama-man tkwen-man* (father-ERG. your<sub>pl.</sub>-ERG.), *mam-isa tkwen-isa* (father-GEN. your<sub>pl.</sub>-GEN.), etc. “your father.”

##### 4.2.4.2 Interrogative/indefinite pronouns

The principal interrogative pronouns are: *vi-* “who”; *romel-* “which”; *ra-* “what,” and its derivatives *ra-ysa-twis* “why” and *ra-oden-* “how much / how many.” These can be converted into indefinite pronouns by the addition of the suffix *-me*: *vi-n-me* “someone,” *ra-y-me* “something,” etc.

##### 4.2.4.3 Relative pronoun

The relative pronoun *passe-partout* is *romel-*, which can have animate or inanimate antecedents. When the relative clause is necessary for the identification of the referent, *romel-* can be accompanied by a demonstrative, almost always *igi*, which does not decline in this context: *ara ese ars=a romel-sa igi xejiebdes mok'lvad?* (not that-ABS. is=QUES. **which-DAT. DEM.** they-were-seeking to.kill-ADV.; Jn 7:25) “Is this not **the one whom** they sought to kill?”

##### 4.2.4.4 Demonstrative pronouns

The demonstrative pronouns come in three sets, with suppletive absolutive and non-absolutive (oblique) stems. They take the same case and number suffixes as common nouns, save for the archaic ergative singular ending *-n*.

(4)	absolutive	oblique	meaning
I.	ese	ama-	“this”
II.	ege	maga-	“that” (associated with interlocutor)
III.	igi	ma-	“that” (remote); basic 3rd-person pronouns “she,” “he,” “it,” “they”

All of these demonstratives double as definite articles. The set II demonstratives, although commonly encountered in conversation, are relatively rare in writing, and hence sparsely represented in the Early Georgian corpus. As would be expected for pronouns associated with the real or metaphoric locus of the interlocutor, they occur almost exclusively in reported speech. At the conclusion of a discussion, for example, Jesus is quoted as saying:

*ara q'(ove)lta dait'ion sit'q'way ege* (Mt 19:11) “Not everyone will accept **that** teaching” (i.e., the teaching which the interlocutors have just mentioned). The set III demonstratives are also the unmarked third-person pronouns, and as such have a far higher frequency of occurrence than the other two sets combined: *ma-n xrkwa ma-s* (Lk 15:27) “he-ERG. said to him-DAT.”

### 4.3 Verbal morphology

The Early Georgian verb is morphologically more complex than the noun, but its generally agglutinative structure permits an analysis by *morpheme slots* and regularities of co-occurrence. In this section, the longest in the chapter, we will begin with an overview of (i) the three verb classes and (ii) the three paradigm series; then embark on a detailed examination of the morphology, slot by slot, followed by a presentation of the semantics of the tense-aspect-mood paradigms (the=sign is used in the glosses to segment cliticized or incorporated lexical elements, such as preverbs, clitic pronouns, and incorporated noun stems, from the internal morphology of the verb).

#### 4.3.1 Verb classes

Georgian philologists divide the verbs of the classical language into three classes, also known as *voices* or *conjugations*, according to their morphology, semantics, and valence. The same tripartite division is employed here, with one minor change.

##### 4.3.1.1 Transitive class

This class includes all verbs having Series II forms that assign ergative case to their subjects. Almost all of these verbs are in fact transitive, but a goodly number are either monovalent (*man imruša* [Lk 16:18] “he-ERG. committed adultery”) or bivalent with an indirect object but no direct object (*man mas mixugo* “he-ERG. him-DAT. answered”).

##### 4.3.1.2 Intransitive class

The *intransitive class* includes both true passives, derived from transitive roots, and basic intransitives. There are four subgroups in this class:

1. *i-prefixal*: Such verbs are marked by the version vowel *-i-* (see §4.3.3 [6]) before the verb root (slot 6), preceded by a dummy third-person object prefix (see §4.3.3 [4]). Always monovalent, their only argument is a subject assigned absolutive case: *igi x-i-kmn-eb-i-s* (that:ABS. “O3”-PASS.-make-SM-TM-S3sg.) “something is being made, done.”
2. *e-prefixal*: This subgroup is marked by the version vowel *-e-*, and comprises verbs that are almost always bivalent, with a subject assigned absolutive case and an indirect object assigned dative case: *igi mas x-e-kmn-eb-i-s* (that:ABS. that:DAT. O3-OBVN-make-SM-TM-S3sg.) “something is being made, done to/for somebody.”
3. *suffixal*: These verbs are marked by the suffix *-n* or *-d*. Many of these verbs are inchoative, often derived from nouns or adjectives: *igi gan=jlier-d-eb-i-s* (that:ABS. Pv=strong-PASS.-SM-TM-S3sg.) “somebody becomes strong.”
4. *root intransitive*: These verbs have no special marker and constitute a small, nonproductive, and archaic group: *igi še=k'rb-eb-i-s* (that:ABS. Pv=gather-SM-TM-S3sg.) “(group) gathers together.”

**Table 40.4 Early Georgian verb classes ("conjugations")**

	Transitive class	Intransitive class	Atelic class
<i>Semantic characteristics</i>	Agentive, accomplishment verbs	Root intransitive, inchoative, passive	Atelic stative and activity verbs
<i>Syntactic characteristics</i>	Assign ERG. in Series II; inversion in Series III	Never assign ERG.	Simplest (archaic?) Series II forms do not assign ERG.
<i>Range of verb forms</i>	All 3 series	All 3 series	Typically Series I only; rare examples with periphrastic or "borrowed" Series II and III

#### 4.3.1.3 Atelic verb class

The third class, which I have designated "atelic verbs," comprises verbs used to describe an ongoing state or activity, without a foregrounded beginning or end point. The atelic class includes statives (-*kw*-/-*kwn*- "have," -*ši*- "be hungry") and activity verbs (-*kadag*- "preach," -*γayad*- "cry out"). One important morphological difference between these and verbs of the other two classes, consistent with their semantics, is the absence of an opposition between perfective and imperfective forms. Each verb selects a single past indicative and future/conjunctive paradigm, usually from Series I, less often from Series II (the term "conjunctive" [Georgian *k'avshirebiti*] is used by Georgian grammarians to denote a set of verb forms with subjunctive, optative, or future meaning):

(5) <i>present</i> :	x-a-kw-s "somebody has something"	x-gon-i-es "somebody thinks something"
<i>past</i> :	x-a-kwn-d-a [=imperfect]	x-e-gon-a [=aorist]
<i>future/conjunctive</i> :	x-a-kwn-d-e-s [=impf. conjunctive]	x-e-gon-o-s [=optative]

### 4.3.2 Paradigm series

Georgian verb forms are traditionally grouped into paradigms marking a specific tense, mood, and aspect. The Early Georgian transitive or intransitive verb formed thirteen paradigms, as far as can be told from the corpus, of which one is sufficiently rare that its status as a productive form is questionable. The Georgian paradigms are grouped into three sets or *series*, based on their stem morphology and syntactic properties:

#### 4.3.2.1 Series II ("aorist series")

These are the morphologically simplest verb forms, associated with perfective, more precisely, *punctiliar* aspect: in the structuring of the narrative, the event or state is represented as a closed-off point (opposed to the linear sense of the Series I paradigms). In some contexts the punctiliar aspect emphasizes the completion of the narrated event; in others its primary function is to mark the events forming the principal narrative line. The ergative case is only assigned by the Series II forms of transitive verbs.

#### 4.3.2.2 Series I ("present series")

The Series I paradigms include a stem formant (*series marker*) which does not appear in the corresponding Series II forms. The two morphologically basic Series I paradigms mark the present indicative. The other four members of the series contain the stem augment

-d-/-od-/-id- and pair off with the Series II paradigms employing the same tense/mood vowels and person suffixes. The contrast is one of durative (or linear) aspect versus punctiliar; the Series I paradigms emphasize the duration of an event, either to imply noncompletion, or to set the temporal background for a foregrounded event marked by a Series II form. In prehistoric Kartvelian, the Series I paradigms were all intransitive, as reflected in their case-assigning properties (they cannot assign ergative case) and in their morphology (the series markers seem to be the relics of ancient antipassive suffixes).

#### 4.3.2.3 Series III (“perfect series”)

This is the most recent and formally most heterogeneous of the three series. In the Early Georgian period, only transitive verbs had synthetic Series III forms; intransitives formed their perfects analytically, as in Latin (*micemul ars = datum est*). The Early Georgian transitive and intransitive Series III forms are identical to the absolute (monovalent) and relative (bivalent) passives of state, and indeed the semantic distance between the passive and perfect functions of these forms is often not very large: the Series III paradigms are principally resultative in meaning, referring to a state of affairs proceeding from the completion of an earlier action.

(6)		<i>monovalent passive of state</i>	<i>intransitive Series III</i>
	<i>c'eril ars</i>	it is written (present)	it has been written (present perfect)
	<i>c'eril xiq'o</i>	it was written (aorist)	it had been written (pluperfect)
		<i>bivalent passive of state</i>	<i>transitive Series III</i>
	<i>x-u-c'er-i-e-s</i>	it is written to/for somebody (present)	somebody has written it (present perfect)
	<i>x-e-c'er-a</i>	it was written to/for somebody (aorist)	somebody had written it (pluperfect)

One interesting syntactic feature of transitive Series III verbs is known as *inversion*: they assign dative case and indirect-object marking to their semantic subjects, and subject status to their direct objects. The case-shift phenomena associated with transitive verbs in Series I, II and III is illustrated in Table 40.5:

**Table 40.5 Case shift**

	Transitive construction			Intransitive construction		
	Subject	Direct object		Subject		
Series I: ( <i>nom.-acc.</i> )	<b>mama-y</b> father-ABS.	je-sa son-DAT.	x-p'ov-eb-s O3-find-SM-S3sg. “The father finds (his) son”	je-y son-ABS.	x-i-p'ov-eb-i-s O3-PASS.-find-SM-TM-S3sg.	“The son is being found”
Series II: ( <i>erg.-abs.</i> )	<b>mama-man</b> father-ERG.	je-y son-ABS.	p'ov-a find-S3sg. “The father found (his) son”	je-y son-ABS.	x-i-p'ov-a O3-PASS.-find-S3sg.	“The son was found”
Series III: ( <i>dat.-abs.</i> )	<b>mama-sa</b> father-DAT.	je-y son-ABS.	x-u-p'ovn-i-e-s O3-OBVN-find-TM-TM-S3sg. “The father has found (his) son”	je-y son-ABS.	p'ovebul found	ars is “The son has been found”

### 4.3.3 Composition of the verb

The Early Georgian verb can be analyzed as consisting of fourteen slots, which may or may not be filled with a morpheme in a given verb form: (i) six prefixal positions; (ii) the root; and (iii) seven suffixal positions:

#### (7) The fourteen slots of the Early Georgian verb

preverb<sub>1</sub>-preverb<sub>2</sub>=clitic<sub>3</sub>=O<sub>4</sub>-S<sub>5</sub>-version<sub>6</sub>-ROOT<sub>7</sub>-causative/passive/inchoative<sub>8</sub>-ABS.plural<sub>9</sub>-series<sub>10</sub>-imperfect<sub>11</sub>-tense/mood<sub>12</sub>-S<sub>13</sub>=clitic<sub>14</sub>

1. *Slot 1 – preverb with more or less predictable directional meaning*: The most common Early Georgian preverbs are: *mi-* “to, away”; *da-* “down”; *šta-* “down”; *ay-* “up”; *gan-* “out”; *še-* “in”; *c’ar-* “away”; *garda-* “across, downward”; *uk’un-* “backwards.”

2. *Slot 2 – preverb mo- (“hither”)*: Indicates movement toward the source, or point of reference (usually, but not always, the locus of the speaker). The addition of *mo-* to a slot 1 preverb gives combinations such as *še-mo=slva-y* “come in, enter [toward source].” The preverb *da-* can also follow certain preverbs, adding what appears to be a nuance of intensity or iteration, as in *mi-mo-da=x-xed-v-id-a* (thither-hither-**da**-O3-look-SM-IMP.-S3sg.) “circumspectavit” (PJ57). In Modern Georgian, preverbs have the additional function of signaling perfective aspect, as in the Slavic languages. Although this is not the case in Early Georgian, there is nonetheless a perceptible tendency for Series I verb forms to lack preverbs, while Series II forms generally have them. The preverbal slot of certain verbs can also be occupied by incorporated direct objects with generic reference: *yaγad=q’-o* (**cry**=do-S3sg.) “he cried out.”

3. *Slot 3 – preverbal clitic*: In Early Georgian, unlike the modern standard language, the bond between preverbs and verbs was sufficiently loose to permit the optional interposition of certain clitic particles, a phenomenon known as *tmesis*. The ten or so Early Georgian preverbal clitics form two semantic groups: (i) adverbials (*-re-* “a little”; *-oden-* “when”; *-ray-* “while, after”) and (ii) indefinite pronominals (*-vietme-*, *-vinme-* “some [people]”; *-rayme-* “something”). Consider these examples: *še=oden=rižwneboda* (Jn 6:17) “**when** it was getting dark”; *mi=vietme=xuges mc’ignobarta ganta* (Mt 12:38) “**some** of the scribes addressed him.”

4. *Slot 4 – morphological object prefix (Set O)*: The Set O person prefixes cross-reference, in the majority of contexts, an argument assigned the dative case. Given the complexity of Georgian case-assignment rules, this latter could be an indirect or direct object, or even the subject of an indirect or Series III transitive verb. First- and second-person absolutive direct objects also control Set O agreement. There are four Set O prefixes, forming a two-by-two array:

#### (8) Morphological object (Set O) markers

	– hearer	+ hearer
+ speaker	m- (1st singular or exclusive)	gw- (1st inclusive)
– speaker	x- (h-) (3rd person)	g- (2nd person)

What appears to be a dummy third-person object prefix (O3) is attested in all Early Georgian *i*-prefixal passives, even though these are monovalent in surface structure: *mi=x-i-q’wan-a igi angeloz-ta-gan c’iaγ-ta abraham-is-ta* (to=O3?-PASS.-bear-S3sg. he:ABS. angel-GEN.PL.-by bosom-DAT.PL. Abraham-GEN.-DAT.PL.; Lk 16:22) “he was carried by angels to the bosom of Abraham.” One possible explanation is that the *x*-prefix once marked agreement with the demoted deep-structure subject (e.g., “angels” in the above example).



5. *Slot 5 – morphological subject prefix (Set S)*: The Set S markers cross-reference the subjects of verbs with direct syntax, and the direct objects of verbs with indirect syntax. The prefixes indicate person only; number being marked by a suffix in slot 13:

(9) Morphological subject (Set S) markers

	singular (slot 5 or 13)	plural (slot 5 and 13)
1st person	v/w-	v/w- -t
2nd person	x- (h-)	x- (h-) -t
3rd person	-s, -a, -n	-n, -es, -ed

With one exception, the presence of a Set O prefix blocks the expression of the Set S prefix controlled by the morphological subject: for example, *šen me mo-m-c-e* (you me Pv=O1excl.-give-OPT.) “You<sub>sg.</sub> will give it to me.” The exception is the combination of third-person object (O3) and first-person subject markers (S1), in which case both are expressed in surface structure: *me mas mi=x-w-c-e* (I this:DAT. Pv=O3-S1-give-OPT.) “I will give it to him/her.” In later Old Georgian, the order of the person prefixes reverses, with the S1 marker preceding the O3 prefix.

6. *Slot 6 – version vowel*: The grammatical category of *version* (Georgian *kceva*) reflects, roughly speaking, the relation between the action or the absolutive argument (direct object of a transitive verb or subject of an intransitive verb), and either the agent or indirect object. There are four formally distinct version relations, though only a few verbs distinguish all four, and many lack the distinction entirely.

- 6A. *Subjective version*: This formant indicates an activity either done for the benefit of the agent him- or herself, or directed toward a direct object linked to (or even identical to) the subject. It is marked by the version vowel *-i-* in all persons (*sibrjne-man i-šen-a tavisaxli* [999 Proverbs 9:1] “Wisdom built a home **for itself**”). Possibly of the same origin is the marker *-i-* in monovalent prefixal passives, which occupies the version vowel slot.
- 6B. *Objective version*: This marker indicates the presence of an indirect object: for example, *ay=x-u-dgin-o-s mk’widri jma-sa twis-sa* (up=O3-OBVN-stand-OPT.-S3sg. offspring-ABS. brother-DAT. own-DAT.; Mt 22:24) “that he raise up offspring **for his brother**.” It is generally marked by the version vowels *-u-* (3rd-person object) and *-i-* (1st- or 2nd-person object); prefixal passive verbs and four archaic transitives employ *-e-* (all persons).
- 6C. *Superessive version*: This is a less common version indicating the presence of an indirect object denoting some kind of surface *upon which* the action is accomplished: for example, *moxgwares k’icwi igi iesus da da=x-a-sx-es mas samoseli* (Mk 11:7) “They brought the colt to Jesus and set [their] clothing **upon it**.” Superessive version is marked by the vowel *-a-* in all persons.
- 6D. *Neutral version*: Many version-marking verbs have a neutral form, with either the vowel *-a-* or no version marker at all.

In a handful of transitive verbs, the version vowel alternates with zero in the third-person subject forms, an alternation evidently once conditioned by stress placement in verbs with or without a syllabic person suffix: S2sg. *x-a-rkw-Ø* “you said something to somebody” versus S3sg. *x-Ø-rkw-a* “(s)he said something to somebody.”

7. *Slot 7 – verb root*: Many verb roots undergo ablaut, of which the two principal patterns are as follows:



- 7A. *e* -*i*-  $\emptyset$ : These root vowels display the distribution: *e* (Tr. Series I; Intr. aorist S1/2), *i* (Tr. Series II),  $\emptyset$  (other Intr.). Consider, for example,  $\text{\textit{\text{še}}}=\text{\textit{x-k'ərb-s}}$  “he gathers<sub>tr</sub>” (Mt 12:30);  $\text{\textit{\text{še}}}=\text{\textit{x-i-k'ri b-i}}$  “you gather<sub>tr</sub> (habitually)” (Mt 25:24);  $\text{\textit{\text{še}}}=\text{\textit{k'rb-es}}$  “they gathered<sub>intr</sub>” (Mt 13:2).
- 7B.  $\emptyset$ -*a*: The distribution is:  $\emptyset$  (most forms), *a* (Aorist S1/2); thus,  $\text{\textit{mo}}=\text{\textit{k'l-a}}$  “(s)he killed somebody,”  $\text{\textit{mo}}=\text{\textit{v-k'al}}$  “I killed somebody.”

Early Georgian ablaut is believed to be the outgrowth of prehistoric alternations related to syllable quantity, stress placement, and perhaps transitivity.

8. *Slot 8 – passive/inchoative or causative suffix*: Directly following the root is a slot reserved for the valence-altering suffixes -*d/-n* (passive/inchoative) and -*ev/-i(v)/-in* (causative). The former pair of allomorphs is used to form suffixal passives, with -*d* added mostly to stems ending in the sonorants /l/, /r/, or /n/, and -*n* in other contexts. The causative suffixes are often accompanied by the version vowel -*a*-.  $\text{\textit{ay}}=\text{\textit{x-w-a-dg-in-eb}}$  “I raise somebody,” compare  $\text{\textit{ay}}=\text{\textit{w-dg-eb-i}}$  “I rise, get up.”

9. *Slot 9 – plural absolutive suffix*: Series II and Series III verb forms (except for the suffixal passives and root intransitives) add a marker -(*e*)*n*- if the absolutive-case argument, denoting the direct object or intransitive subject, is formally plural (i.e., marked by the pluralizer -*n*, which may be related to -(*e*)*n*-):  $\text{\textit{rayta}}=\text{\textit{mca x-i-did-n-es igi-n-i}}$  (that-OPT. O3-PASS.-big-**PL. Abs.**-S3pl. this-**PL.**-Abs.; Mt 6:2) “that **they** be magnified”;  $\text{\textit{m-i-qsn-en čwen borot'isa-gan}}$  (**O1excl.**-OBVN-release-**PL. Abs. us** evil-from; Mt 6:13) “deliver **us** from evil.”

10. *Slot 10 – series marker (or “present/future stem formant”)*: This is a lexically specified morpheme used to form the Series I stem of most verbs, for example:

- (10) *Series I (imperfect)*:  $\text{\textit{x-c-em-d-es}}$  (O3-strike-**SM-IMP.**-S3pl.) “they were striking him”  
*Series II (aorist)*:  $\text{\textit{x-c-Ø-es}}$  (O3-strike-S3pl.) “they struck him”

The principal series markers are -*eb*-, -*av*-, and -*i*-; the less common allomorphs include -*am*-, -*ev*-, -*em*-, -*ob*-, and -*op*-. According to most experts, the series markers were once antipassive formants, deriving aspectually durative intransitives from transitive forms associated with punctiliar aspect. The vowels of some series markers undergo syncope when followed by certain suffixes, and the markers -*av* and -*am* undergo a vowel mutation that may reflect prehistoric umlaut: compare the forms  $\text{\textit{x-loc-av-s}}$  “somebody implores somebody” (present);  $\text{\textit{x-loc-v-id-a}}$  “somebody was imploring somebody” (S3sg. imperfect), and  $\text{\textit{x-loc-ev-d-Ø}}$  “you were imploring somebody” (S2sg imperfect, < \* $\text{\textit{x-loc-av-id-Ø}}$ ).

11. *Slot 11 – imperfect stem suffix*: The stem augment -*d/-od/-id* is used to form the imperfect and indeed all of the Series I paradigms except for the present and present iterative. The allomorph -*od* is employed by intransitives and some atelics; -*id* follows the series markers -*av* and -*am*; and -*d* appears elsewhere.

12. *Slot 12 – tense/mood vowel*: A vowel (-*e*-, -*o*-, -*i*-) inserted before the person/number (Set S) suffix of certain forms serves to distinguish indicative from conjunctive paradigms. Also occupying this slot is the suffix -*i* of the passive present, and a homophonous (perhaps cognate) suffix employed by statives and the present perfect of transitives in conjunction with an -*e*- element of unclear origin, for example,  $\text{\textit{g-gon-i-e-s}}$  “you think something.” The passive and stative -*i*- are to be further distinguished from the vowel /i/ inserted before the Set S suffix -*n* and optionally before the S1/2 pluralizer -*t* (see 13) in certain paradigms: for example, in the imperfect imperative  $\text{\textit{x-a-did-eb-d-i-n}}$  “may they praise somebody.”

13. *Slot 13 – person/number suffix (Set S)*: While the first- and second-person subject (S1/2) suffixes are the same in all paradigms, the third singular and plural subject (S3) morphemes come in three pairs, correlated to a degree with semantic features of the verb forms. A few

**Table 40.6 Set S (morphological subject) suffix groups**

Set S suffix set (slot 13)	1st and 2nd person	3rd singular	3rd plural
A. PRESENT/CONJUNCTIVE present, present-perfect, imperfect iterative, permansive, optative, imperfect and pluperfect conjunctive	sg. -Ø / pl. -t	-s	-n / -en / -an
B. PAST INDICATIVE imperfect, aorist, pluperfect	sg. -Ø / pl. -t	-a	-es
C. IMPERATIVE/ITERATIVE present iterative, imperfect and aorist imperative	sg. -Ø / pl. -t	-n	-ed

paradigms are distinguished by the Set S suffixes alone (e.g., present indicative and present iterative, optative and aorist imperative); since the S1/2 endings do not vary, only the S3 forms are distinct in these instances.

14. *Slot 14 – postposed clitics*: These include the optative particle *-mca* (used with indicative-mood verbs to give them optative/subjunctive force); the adverbials *-ya* “even, just” and *-ve* “indeed, the very” (e.g., *kvani yayadebden=ve* [Lk 22:60] “**the very** rocks will cry out”); the yes-no question particle *-a*; and the indefinite quantifier *-me* (e.g., *xiq’os=me vin tkwengani k’aci* [Mt 7:9] “would there be **any** man among you?”).

#### 4.3.4 Verb paradigms and their functions

In the present section, Early Georgian verb paradigms and their functions are discussed according to paradigm series (see §4.3.2). In Table 40.7, verb paradigms are illustrated using transitive (TR) and intransitive (INTR.) S3sg. (having a third-person singular subject marker) forms of *mi=c-em-a* “give”; verb slots (see §4.3.3) are indicated by subscript numerals.

##### 4.3.4.1 Paradigm Series I

1. *Present*: This is the unmarked present indicative paradigm, and the most frequently attested in the Early Georgian corpus.

2. *Present iterative*: The present iterative can be formally distinguished from the present in the third person only. It often appears in statements of verities and generalizations. Note the contrast between the present iterative and simple present in the following passage. The present iterative and the permansive, its Series II counterpart, are used to convey a fact known from repeated observation, while the optative (future) and present are used in the description of an event – the Second Coming – which will occur only once:

- (11) *xolo leywisagan isc’avet igavi igi : ras žams rt’oni misni dač’wnian da purceli gamo = val-n xuc’q’odit rametu axlos ar-n zapxuli. egreca tkwen : ras žams hixilot ese q’oveli xuc’q’odit rametu axlos ar-s k’arta zeda*

“From the fig tree learn a lesson: When its branches grow tender (PERMANISIVE) and the leaves **come out** (PRESENT ITERATIVE), you will know that summer **is** (PRESENT ITERATIVE) near. Likewise when you will see (OPTATIVE) all these things, you will know that he **is** (PRESENT) near, at your door” (GL Mt 24:32–33).

**Table 40.7 Early Georgian verb paradigms**

	Punctiliar (Series II)	Linear/durative (Series I)	Resultative (Series III)
<i>present</i> <i>indicative</i>	— — —	<i>present</i> TR. $mi_1 = x_4-c_7-em_{10}-s_{13}$ INTR. $mi_1 = x_4-e_6-c_7-em_{10}-i_{12}-s_{13}$	<i>present perfect</i> TR. $mi_1 = x_4-u_6-c_7-i_{12}-e_{12}-s_{13}$ INTR. $mi = cemul\ ars$
<i>past</i> <i>indicative</i>	<i>aorist</i> TR. $mi_1 = x_4-c_7-a_{13}$ INTR. $mi_1 = x_4-e_6-c_7-a_{13}$	<i>imperfect</i> TR. $mi_1 = x_4-c_7-em_{10}-d_{11}-a_{13}$ INTR. $mi_1 = x_4-e_6-c_7-em_{10}-od_{11}-a_{13}$	<i>pluperfect</i> TR. $mi_1 = x_4-e_6-c_7-a_{13}$ INTR. $mi = cemul\ xiq'o$
<i>future/</i> <i>conjunctive</i>	<i>optative</i> TR. $mi_1 = x_4-c_7-e_{12}-s_{13}$ INTR. $mi_1 = x_4-e_6-c_7-e_{12}-s_{13}$	<i>imperfect conjunctive</i> TR. $mi_1 = x_4-c_7-em_{10}-d_{11}-e_{12}-s_{13}$ INTR. $mi_1 = x_4-e_6-c_7-em_{10}-od_{11}-i_{12}-s_{13}$	<i>pluperfect conjunctive</i> TR. $mi_1 = x_4-e_6-c_7-e_{12}-s_{13}$ INTR. $mi = cemul\ xiq'os$
<i>permansive/</i> <i>habitual</i>	<i>permansive</i> TR. $mi_1 = x_4-c_7-i_{12}-s_{13}$ INTR. $mi_1 = x_4-e_6-c_7-i_{12}-s_{13}$	[ <i>imperfect iterative</i> ] TR. $mi_1 = x_4-c_7-em_{10}-d_{11}-i_{12}-s_{13}$ INTR. $mi_1 = x_4-e_6-c_7-em_{10}-od_{11}-i_{12}-s_{13}$ <i>present iterative</i> TR. $mi_1 = x_4-c_7-em_{10}-n_{13}$ INTR. $mi_1 = x_4-e_6-c_7-em_{10}-i_{12}-n_{13}$	— — —
<i>imperative</i>	<i>aorist imperative</i> TR. $mi_1 = x_4-c_7-e_{12}-n_{13}$ INTR. $mi_1 = x_4-e_6-c_7-e_{12}-n_{13}$	<i>imperfect imperative</i> TR. $mi_1 = x_4-c_7-em_{10}-d_{11}-i_{12}-n_{13}$ INTR. $mi_1 = x_4-e_6-c_7-em_{10}-od_{11}-e_{12}-n_{13}$	— — —

The other principal function of this paradigm is in negative imperatives introduced by the particle *nu* “do not” (2nd person: *nu še = x-jrc'un-d-eb-i-t* [Lk 21:9] “do not be anxious”; 3rd person: *nu še = jrc'un-d-eb-i-n guli tkweni* [Jn 14:1] “Let not your heart be anxious”).

3. *Imperfect indicative*: This is the basic Series I past indicative paradigm, aspectually contrasted with the aorist. It is the only past indicative form for many stative and atelic verbs: *brc'q'in-v-id-a* “glistened”; *jc'-od-a* “trembled”; *x-tn-d-a* “liked.”

4. *Imperfect imperative*: Early Georgian has two positive imperative paradigms, distinguished by aspect. The imperfect imperative is used to direct the listener to engage in some sort of ongoing, repeated activity: *sneulta gan = x-k'urn-eb-d-i-t*, *ganbok'lebulta gan = x-c'med-d-i-t* (Mt 10:8) “cure the sick, cleanse the leprous.” As with the aorist imperative, the imperfect imperative has no S2 prefix: *še = (Ø)-vid-od-e-t ic'rosa mas bč'esa* (Mt 7:13) “enter by the narrow gate.”

5. *Imperfect iterative*: This paradigm is unusually difficult to detect, in that it is formally identical to the Series I conjunctive of intransitive verbs, and – in the first and second person – to the imperfect indicative of transitive verbs. This leaves the S3sg. and S3pl. of the transitive conjugation as the only morphologically unambiguous forms of the imperfect iterative. Only three examples are attested in the Early Georgian corpus, all from the same passage:

- (12) *v-e-vedr-eb-od-i-t da odes igi ševidis vitar igi šišit da jc'olit vdgēt* [L. K'ik'nadze reads *vdgit*] *da guls v-e-t'q'-od-i-t da ymrtisa mimart v-i-loc-v-id-i-t misisa mis gulis mokcevisatwis*

“[If we desire something from an earthly monarch] . . . **we would plead** (IMP. ITER.) to his servant for admission and when he comes (PERMANISIVE), we would stand (PERMANISIVE) as though in fear and trembling, and **we would feel desire** (IMP. ITER.) and **we would pray** (IMP. ITER.) to God that his heart be turned (toward us)” (Mrv 65).

The presence of the permansive indicates that a gnomic/iterative sense is intended. It should be noted that the manuscript in question is relatively late (eighth century), and contains numerous divergences from standard Early Georgian usage. In particular, the O3 prefix *x-* is frequently omitted before the S1 marker, as in the above passage. It may be that the imperfect iterative was an innovation in late Early Georgian, or introduced into this text from the native dialect of the translator.

**6. Imperfect conjunctive:** The imperfect conjunctive can be translated by either a subjunctive or a future indicative, depending on context: (fut. indic.) *da mravalni cruv c'inac'armet'q'welni ay=dg-e-n da x-a-ctun-eb-d-e-n mravaltā* (Mt 24:11) “and many false prophets will arise (OPTATIVE) and **will deceive** (IMPERFECT CONJUNCTIVE) many”; (subjunc.) *tu marjwenê qeli šeni g-a-ctun-eb-d-e-s* (Mt 5:30) šen “if your right hand **deceive** you.” The imperfect conjunctive (and optative) are likewise commonly found in restrictive relative clauses: *xlocevdit romelni mi = g = xweč-d-e-n tkwen* (Mt 5:44) “pray for those who **persecute** you.”

#### 4.3.4.2 Paradigm Series II

**1. Aorist:** The aorist is the unmarked Series II paradigm, the second most common verb form in the Early Georgian corpus, after the present indicative. In narratives the aorist is employed by verbs representing the main story line, presented as a succession of events; in this function it contrasts primarily with the imperfect, as well as the conjunctive paradigms, the pluperfect, etc.

**2. Aorist imperative:** The second-person aorist imperative is formally the simplest of the Early Georgian paradigms, lacking the Set S prefix found in the otherwise identical aorist indicative: for example, *mo=ved* “come!”; compare aorist *mo=x-wed* “you came.”

**3. Permansive (aorist iterative):** This paradigm is employed in parables, statements of regularities, and accepted truths, and as such can be translated by the simple present in English: *mas x-u-rkw-i c'arved da c'ar=vid-i-s* (Mt 8:9) “I **tell** him ‘go,’ and he **goes**.”

**4. Optative (aorist conjunctive):** The optative, like its Series I counterpart, the imperfect conjunctive, can have either future indicative or subjunctive meaning. In the latter sense it commonly appears after subordinating conjunctions.

#### 4.3.4.3 Paradigm Series III

**1. Present perfect:** The Early Georgian present perfect is primarily resultative in meaning, representing a state of affairs extending to the (narrative) present as resulting from some event in the past: for example, *ay=dgomil ars mk'wdretit* (Mt 14:2) “he **has risen** from the dead” (implication: he is still alive); *ege q'oveli da=m-i-marx-av-s siq'rmit čemitgan* (Mt 19:20) “all of these [commandments] I have kept since childhood” (implication: I still do).

**2. Pluperfect:** The basic function of the Early Georgian pluperfect is to mark past anteriority: *šeič'q'nares igi galilevelta rametu q'oveli x-e-xilv-a raodeni x-e-kmn-a ierusalêms* (Jn 4:45) “The Galileans welcomed him, for they **had seen** all that he **had done** in Jerusalem.” The semantic difference between Series III and passive of state is especially slight in the case of

intransitive present perfects and pluperfects: *šek'rebul xiq'o bevreuli eri* (Lk 12:1) “a crowd of thousands **had (was?) gathered.**”

3. *Pluperfect conjunctive*: This rare paradigm is attested only twice in the Early Georgian corpus. In both cases it appears to mark future anteriority: <arya> *x-e-q'iv-n-o-s katamsa vidremde uvar* = *mq'o me sam gzis* (Jn 13:38) “The cock **will not have crowed** before you deny me three times”; *net'ar xiq'wnen romelta ara x-w-e-xilv-o da x(w)urc'mene* (GL Jn 20:29) “Blessed will be those who **will not have seen me** but who will believe in me” (note that the S1 prefix *w-* in *xwexilvo* marks the direct object, in accordance with the inverse syntax governed by transitive verbs in Series III; see §4.3.2.3).

### 4.3.5 Nonfinite verbals

The principal nonfinite forms of the Early Georgian verb are the verbal noun and three participles: active, past passive, and future passive.

#### 4.3.5.1 Verbal noun

This is usually formed by adding the suffix *-a* to the verb root and its series marker (a smaller number of verbs, mostly members of the atelic class, employ the suffix *-il/-ol/-ul*, sometimes with the prefix *si-*). Among other things it can function like an infinitive in nominalized clauses subcategorized by certain verbs: for example, *p'ilat'e xubrjana mi=c-em-a-d gwami misi* (Mt 27:58) “Pilate ordered them **to give** him his (Jesus') body” (lit. “Pilate ordered them **the giving** of his body”).

#### 4.3.5.2 Participles

The active or agentive participle contains a prefix *m-/ma-/me-/mo-* inserted before the stem, and a suffix *-el/-ar/-ul*: *vin ars mi=m=c-em-el-i misi* (Jn 6:64) “who is the **one who will hand him over**” (lit. “who is his **giver**”). The past (or perfect) passive participle is usually formed with the suffix *-il/-ul*; among other uses it is employed in the Series III forms of intransitive verbs: *romelta mi=c-em=ul ars* (Mt 19:11) “[those] to whom it is **given.**” The future passive is formed with the addition of a prefix *sa-* before the stem, and the same suffix as in the corresponding active participle: *xicit sa=c-em=el-i k'etili micemad švilta tkwenta* (Lk 11:13) “you know to give your children good **gifts**” (lit. “that-which-is-to-be-given”).

## 4.4 Diachronic morphological developments

Although the Xanmet'i dialect is the most archaic attested variety of Georgian, hints of changes to come can be detected here and there in Early Georgian texts. Among them are the following.

1. *Uncertainty in the use of O1excl. m-*: While the inclusive/exclusive opposition in the Set O prefixes is maintained in the Xanmet'i gospels, evidence that the first-person inclusive object marker *gw-* is being reinterpreted as a general first plural prefix begins to appear in the Graz Lectionary composed a century later: *uitar igi m-e-t'q'-od-a čwen gzasa zeda; da vitar igi gamo=gw-i-targman-eb-d-a čwen c'ignta* (GL Lk 24:32) “how he spoke to **us** (*m-*) on the road, and how he interpreted the books for **us** (*gw-*).”

2. *Paradigm recruitment for atelic verbs*: In later stages of Georgian, atelic activity verbs have the same range of paradigms as the transitive and intransitive conjugations. In the Early Georgian period, however, the rare Series II and III atelic verbs seem almost to be nonce

formations cobbled together from elements borrowed from the transitive and intransitive conjugations. The early Series II paradigms of atelic verbs display three types of formation:

- (i) periphrastic, formed with *q'opa* “make”: *ḡayad* = *q'o* (Mt 14:30; Jn 7:28) “he cried out” (lit. “he made a cry”)
- (ii) root intransitive morphosyntax (more archaic?) with subject in absolutive: *katami q'iv-a* (Lk 22:60) “the cock-ABS. crowed”
- (iii) transitive morphosyntax (more recent?) with subject in ergative and verb in subject version: *man i-mruš-a* (Mt 5:28) “he-ERG. committed adultery.”

The root *-q'iv-* “crow” is a curious case, having a formally intransitive aorist, but a formally transitive pluperfect conjunctive with inversion: *x-e-q'iv-n-o-s katam-sa* (O3-OBVN-crow-PL.?-TM-S3sg. cock-DAT.; Jn 13:38) “the cock will have crowed.”

## 4.5 Numerals

Georgian has a mixed decimal and vigesimal counting system. Monomorphemic number names are used for counting to ten, followed then by compounds of the form “ten-N-more” (e.g., *at=rva=met'* lit. “10-8-more,” i.e., “18”) up to *oc* “20.” Counting continues by scores (e.g., *otx = me-oc da a(t) = cxra = met'* lit. “4 = score and 10=9= more,” i.e., “99” [Mt 18:12]) up to *as* “100.” Higher units include *at = as* (“10 = 100”) “1,000” and *bevr* “10,000.”

# 5. SYNTAX

## 5.1 Word order

Early Georgian word order gives the impression of being freer than it actually is. While it is indeed the case that very few constituents occupy an *obligatory* position, most do have a preferred position. According to Sarjeladze’s quantitative study (1984:528, 535–536), Old Georgian in general, and Early Georgian in particular, favors head–modifier order both within the clause and within the noun phrase (NP): direct and indirect object after the verb; adjective, article, and possessor after the head noun, for example, *twali<sub>1</sub> šeni<sub>2</sub> marjwenê<sub>3</sub>* (Mt 5:29) “your<sub>2</sub> right<sub>3</sub> eye<sub>1</sub>.” The principal exceptions are interrogative, negative, and numeral modifiers, which generally precede their head. The subject, interestingly, is as likely to follow the verb as precede it, postverbal position being favored by subject NPs referring to new topics: *xolo xiq'wnes mun dedanica mravalni* (Mt 27:55) “But **many women** were there.”

Among the items which have a relatively fixed position are definite articles and sentential clitics such as *tu* “if,” *ra(y)* “when,” which follow the first element in the NP or clause: *atertmet'i igi moc'apeni* (Mt 28:16) “**the** eleven disciples”; *aγ= ra =xesrulnes dyeni igi* (Lk 4:2) “**When** those days were over.”

## 5.2 Coordination and subordination

In addition to the relative pronoun *romel-*, described earlier (see §4.2.4.3), other interrogative pronouns double as subordinators, for example, *raoden-* “how much?”; “as much as”: *xuq'wes mas raodeni xunda* (Mt 17:12) “they did to him **as much as** they wanted.” Subordinate clauses can likewise be introduced by conjunctions of various sorts: *tu* “if,” *rayta* “that,” *vidremde* “until,” etc. Many of these require a verb in the conjunctive or optative. The principal coordinating conjunction is *da*, which operates at the word, phrase, and clause level.

### 5.3 Agreement

Agreement, as distinguished from cross-referencing, occurs within the NP, and also between certain verb forms and absolutive-case NPs within a phrase. In the instance of NP-internal agreement, adjectives, articles, and even genitive-case modifiers reflect the case and number of the head noun: *jujeul-n-i mat-n-i* (alumnus-PL.-ABS. their-PL.-ABS.; BQ III) “their foster children.” In NPs where a modifier is itself modified by a noun in the genitive, the latter may bear three case endings: its own (genitive), a copy of its head’s case (genitive), and the case assigned the head of its head: for example, *saidumlo-y<sub>1</sub> sasupevel-isa<sub>2</sub> ca-ta<sub>3</sub>-ysa<sub>2</sub>-y<sub>1</sub>* (secret-ABS. kingdom-GEN. sky-GEN. PL.-GEN.-ABS.; Mt 13:11) “the secret of the kingdom of the heavens.” The second agreement phenomenon of note is between Series II and Series III verbs and their absolutive arguments. Formally plural absolutive NPs (those marked with the pluralizer *-n-*, as well as first- and second-person pronouns and plural null anaphors) control the probably cognate agreement marker *-(e)n-* in slot 9 of the verb (see §4.3.3 [9]).

## 6. LEXICON

The great majority of lexemes employed in the Early Georgian texts are of indigenous origin, as far as can be told. At the same time, a number of cultures have left their imprint on the Georgian lexicon. The Greek of eastern Christianity has contributed terms such as *ek’lesia* “church” and *angeloz-* “angel”; *nav-* “ship” and *mankana* “machine, device” may go back to Hellenic times, when Greek merchants first established trading posts in Colchis. Persian civilization, with which the Georgians have been in regular contact since well before the Christian period, is the source of a considerable number of words, including many in common use: *p’at’iv-* “honor,” *žam-* “time,” *parto* “wide.” The contribution of Armenian is easy to underestimate, since many words of Persian and Syriac origin (*sp’et’ak’-* “white,” *targm(a)n* “translate”) presumably entered Georgian via their neighbors to the south. The verb root *šên-* “build” and possibly the noun *mgel-* “wolf” (borrowed to replace a tabooed inherited root?) represent prehistoric loans from Armenian.

### Abbreviations

#### *Linguistic terms*

IMP.	imperfect-stem formant
O1excl	1st-person exclusive object marker
O1incl	1st-person inclusive object marker
O3	3rd-person object marker
OBVN	objective version vowel
Pv	preverb
Ques.	question particle
S1	1st-person subject marker
S3pl.	3rd-person plural subject marker
S3sg.	3rd-person singular subject marker
SBVN	subjective version vowel
SM	series marker
TM	tense/mood vowel



Most of this chapter was written in 1996. Since that time, further Early Georgian texts have been made available for study, including the palimpsest Codex Georg. 2 of Vienna, and a new edition of the Graz Lectionary, through the efforts of Jost Gippert (Frankfurt) and Zurab Sarjveladze (Tbilisi). Recently, the archeologist Levan Ch'ilashvili has published the startling claim that several fragmentary inscriptions uncovered during excavations of what he believes was a pagan temple at Nek'risi, in eastern Georgia, are to be dated to the 1st–3rd centuries AD (*Burji Erovnisa* #3, pp. 6–7, 2001). If true, this would be the first evidence that the Georgian alphabet predated the adoption of Christianity as state religion. In my view, there is nothing in either the form of the letters, nor in the grammatical features of the one inscription that has been published, which would compel the attribution of such an early date. It remains to be seen whether further investigation of the inscriptions, and the archeological context in which they were found, will confirm Ch'ilashvili's hypothesis.

## Bibliography

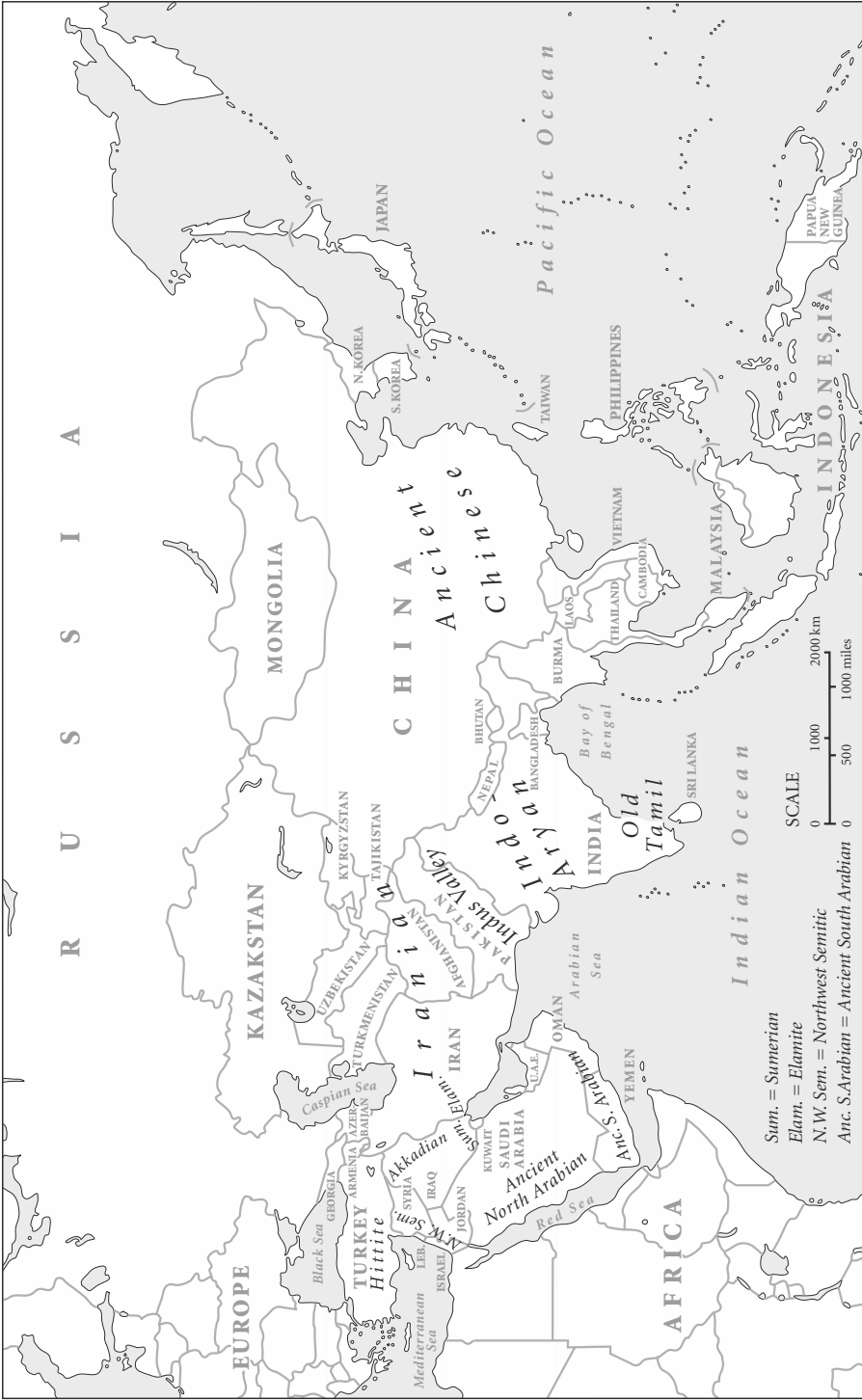
- Abuladze, I. 1973. *kartuli c'eris nimuṣebi. p'aleografiuli albomi* (Examples of Georgian writing. A paleographic album; 2nd edition). Tbilisi: Mecniereba.
- Birdsall, J. N. 1970. "A second Georgian recension of the Protevangelium Jacobi." *Le Muséon* 83:49–72.
- . 1991. "Georgian paleography." In A. C. Harris (ed.), *The Indigenous Languages of the Caucasus, Vol. 1: The Kartvelian Languages*, pp. 85–128. Delmar, NY: Caravan Books.
- Blake, R. 1932. "Khanmeti palimpsest fragments of the Old Georgian version of Jeremiah." *Harvard Theological Review* 25:207–272.
- C'ereteli, G. 1960. *ujvesi kartuli c'arc'erebi p'alest'inidan* (The oldest Georgian inscriptions from Palestine). Tbilisi: Mecniereba.
- Danelia, K. and Z. Sarjveladze. 1997. *Kartuli p'aleografia*. [Georgian paleography.] Tbilisi: Nek'eri.
- Gamq'relidze, T. 1990. *C'eris anbanuri sist'ema da jveli kartuli damc'erloba. anbanuri c'eris t'ip'ologia da c'armomavloba* (Alphabetic writing and Old Georgian script. The typology and origins of alphabet writing). Tbilisi: Tbilisi State University Press.
- Kajaia, L. (ed.). 1984. *xanmet'i t'ekst'ebi, 1: xanmet'i otxtavi* (Xanmeti texts, 1: the Xanmeti Gospels). Tbilisi: Mecniereba.
- Molitor, J. (ed.). 1956. *Monumenta Iberica antiquiora: Textus chanmeti et haemeti ex inscriptionibus, Bibliis et patribus*. Louvain la-Neuve: Corpus Scriptorum Christianorum Orientalium.
- Sarjveladze, Z. 1971. *xanmet' da haemet' t'ekst'ebši dadast'urebul zmnis p'irian pormata sajieblebi* (A tabulation of the finite verb forms attested in the Xanmeti and Haemeti texts). Tbilisi: Mecniereba.
- . 1984. *kartuli salit'erat'uro enis ist'oriis ṣesavali* (An introduction to the history of the Georgian literary language). Tbilisi: Ganatleba.
- Tuite, K. 1990. "Das Präfix x- im Frühgeorgischen." *Georgica* 13/14:34–61.

## Sources

Mt, Mk, Lk, Jn	Xanmet'i gospels (6th century) [Kajaia 1984]
Bol	Bolnisi inscriptions (493–494) [Abuladze 1973]
BQ	Bir el-Qutt inscriptions (429–444) [C'ereteli 1960]
GL	Graz Xanmet'i Lectionary (late 7th century) [Molitor 1956]
H	Haemet'i lectionary (8th century) [Molitor 1956]
Jer	Cambridge & Oxford Jeremiah fragments (c. 600) [Blake 1932; Molitor 1956]
Mrv	Xanmet'i Mravaltavi (early 8th century) [Molitor 1956]
PJ	Protevangelium Jacobi (c. 700) [Birdsall 1970]
999	Palimpsest H-999 (Bible fragments, c. 500) [Molitor 1956]







**Map 5.** The distribution of attested Ancient languages within Asia (for the languages of Asia Minor, see Map 2)

# Ancient Chinese

ALAIN PEYRAUBE

## 1. HISTORICAL AND CULTURAL CONTEXTS

### 1.1 Introduction

“Chinese is only one of a very few languages whose history is documented in an unbroken tradition extending back to the second millennium BC” (Norman 1988: ix). There are two main causes for this situation: (i) the unity of Chinese culture in spite of periods of political disunity; (ii) the use of a script which has been independent of any particular phonetic manifestation of the languages it represented.

Chinese is usually divided into *Ancient Chinese* (*gudai hanyu*) and *Contemporary Chinese* (*xiandai hanyu*; the pronunciation of all Chinese characters is herein given in the modern standard language *putonghua*, in the standard romanization *pinyin*). Ancient Chinese is simply defined as “the language of the writings of the past” (Wang 1979:1). It covers a very long period, from the first Chinese inscriptions known to us, dated to the fourteenth century BC, until the nineteenth century. One then distinguishes generally for Ancient Chinese three basic stages: (i) the Archaic period (*shanggu*), until the second century BC; (ii) the Middle or Medieval period (*zhonggu*), from the first century BC to the middle of the thirteenth century AD; (iii) the Modern period (*jindai*), from the middle of the thirteenth century to the middle of the nineteenth century.

It is during the Archaic period that what is known today as *Classical Chinese* (*wenyan*) is fixed. This language remained as the main written language used in literary texts until the beginning of the twentieth century, but has progressively become a dead language since the beginning of the Medieval period (playing a role like that of Latin in Europe), and the current spoken Sinitic languages have diverged from it considerably.

The Classical period proper begins with Confucius (551–479 BC), and ends around the founding of the Qin Empire in 221 BC. The attested language of the period was probably not very different from cultured speech. The gap between the written and the spoken language began to develop in the Han dynasty (206 BC–AD 220) and increased naturally with time. Before Confucius, the literary language is called *Preclassic*.

It is essentially the classical language par excellence – that of the Warring States period (475–221 BC) – that will be discussed in this chapter. Attention will, however, also be paid to the preclassic language and, above all, to the language as known after the Han dynasty, up until around AD 600.

## 1.2 The linguistic family of Chinese

Chinese is considered today by most specialists of the language, and also by Tibeto-Burmanists, to belong to the *Sino-Tibetan* family of languages (subdivided principally into Chinese [Sinitic] and Tibeto-Burman). Other hypotheses regarding the genetic affiliations of Chinese have been offered in very recent years. Sagart (1994) posits a Sino-Tibetan-Austronesian family (with the three subdivisions of Sinitic, Tibeto-Burman, and Austronesian). Starostin (1989) argues for a Sino-Caucasian macrofamily which includes Sino-Tibetan, Yeniseian, and North Caucasian. Pulleyblank (1995) relates Sino-Tibetan and Indo-European to one another, stressing that “the traces of shared phonological and morphological correspondences at a very deep level are hard to explain except as evidence of common origin.”

## 1.3 Chronological stages

At present there are several possible periodizations of the Chinese language, falling into two major categories according to phonological or syntactic criteria.

The long history of Chinese phonology is divided today into the following periods: (i) *Old Chinese* (this term has replaced what Karlgren [1915–1926] called “Archaic Chinese”), representing the language of around 1000–800 BC; (ii) *Early Middle Chinese* (replacing Karlgren’s “Ancient Chinese”), which represents the literary pronunciation of the sixth century AD; (iii) *Late Middle Chinese*, the language of the Late Tang (618–907) and Early Song (960–1279) periods; (iv) *Early* (or *Old*) *Mandarin*, the language of the Yuan (1279–1368) period (see Pulleyblank 1970–1971; Baxter 1992:14–15).

A different periodization is based on syntactic criteria only. It distinguishes four major stages, *Archaic* (*shanggu*), *Medieval* (*zhonggu*), *Modern* (*jindai*), and *Contemporary* (*xiandai*), which are subdivided as follows: (i) *Pre-Archaic*, the language of the oracle inscriptions on bone and shell (fourteenth–eleventh centuries BC); (ii) *Early Archaic*, bronze inscriptions and early Chinese classics (tenth–sixth BC); (iii) *Late Archaic*, the Classical Chinese par excellence, comprising such well-known texts as the *Analects of Confucius*, and *Mencius* (fifth–second BC); (iv) *Pre-Medieval*, a transitional period, which witnesses the birth and development of an attested vernacular language different from the literary one (first century BC–first century AD); (v) *Early Medieval* (second–sixth AD); (vi) *Late Medieval* (seventh–mid-thirteenth AD); (vii) *Pre-Modern*, another transitional period (mid-thirteenth–fourteenth AD); (viii) *Modern Chinese* (fifteenth–mid-nineteenth AD); (ix) *Contemporary Chinese* (mid-nineteenth century to the present; see Wang 1958:35; Chou 1963:432–438; and especially Peyraube 1988). This syntactic periodization is followed below, except for the section on phonology, which adopts the periodization based on phonological criteria.

For all these stages, inscriptions and texts have survived in great quantities.

## 1.4 Dialectal variation

Regional dialects surely existed in China from the most ancient period. We have an early record of dialectal words, the *Fangyan* by Yang Xiong (53 BC–AD 18), but it tells us very little about how the different words were pronounced.

The classical literary language, and also the vernacular literary language which appeared around the second century AD, apparently established norms preventing the development of any real dialectal or regional literary competitors. However, if we consider the classical period proper, corresponding to the Late Archaic Chinese of the Warring States period (fifth–third centuries BC), we do find some linguistic diversity among the texts. This is probably not

merely the result of different historical phases; various regional dialects most likely were in use as the vehicles of literature in their respective areas.

Pulleyblank (1995:3) distinguishes the following four dialects: (i) a rather archaic form of literary language, based probably on a central dialect, used in historical texts; (ii) an Eastern Lu dialect used in *Lun yu* (*Analects*) and *Mengzi* (*Mencius*); (iii) a Southwestern Chu dialect; (iv) a third-century BC dialect found in philosophical texts.

By the Tang period, and probably earlier, China has acquired a standard common language (*gongtongyu*). This is clearly the case for the written language, but there are numerous indications that it is also the case for the spoken (see Mei 1994).

## 2. WRITING SYSTEMS

The following section is largely inspired by Norman (1988:58–82) and Qiu (1978).

### 2.1 The origin of Chinese writing

It can be reasonably assumed that Chinese writing begins sometime around the seventeenth century BC (see Qiu 1978). The Chinese script already appears as a fully developed writing system in the late Shang dynasty (fourteenth–eleventh centuries BC). We have for this period a great number of texts inscribed or written on tortoise-shells or bones, the script being known as *jiaguwen* (“oracle bone script”).

The next phase of the Chinese script, *jinwen* (“bronze script”), used during the Western Zhou (eleventh century–771 BC), and the Spring and Autumn (770–476 BC) periods, in its basic structure and style is similar to that of the late Shang, and is clearly derived from it. Later scripts will similarly be derived in succession from the preceding form, until the appearance of the “standard script” (*kaishu*), which begins to take shape around the third century AD, and is still in use today (see §2.3).

One could therefore say that Chinese writing is characterized by continuous use of the same system since remote antiquity.

### 2.2 The nature of the writing system

From its very beginning, the Chinese writing system has been fundamentally morphemic, in other words, almost every graph represents a single morpheme. Since the overwhelming majority of Archaic Chinese morphemes are monosyllabic, every graph then represents a single morpheme at the phonological level.

If one excludes a very small number of early graphs, which are apparently arbitrary signs bearing no iconic or phonetic relationship to the word represented, one can claim that the Shang script contains characters of two basic types: (i) those which are semantically representational without any indication of the pronunciation of the words represented; and (ii) those which are in some fashion tied to the pronunciation of the words.

The earliest Chinese writing clearly reveals a basically pictographic origin of the characters. It is now quite difficult, however, to identify what object was originally depicted by many of the symbols; that is, after being simplified and stylized in later stages, they lost their original pictorial quality. Graphic representations were thus originally linked to the words they represented without any reference to the pronunciation of the morphemes in question (the system was iconic). However, as in all fully developed writing systems, phonetic elements were eventually introduced. These were particularly needed for representing abstract notions

**Table 41.1 Development of the Chinese script**

	Shang bone script	Zhou bronze script	Warring States script	Seal script	Clerical script (Han)
1. “child”	𠂔	𠂔	𠂔	𠂔	子
2. “cloud”	𠂔	𠂔	𠂔	雲	雲
3. “water”	𠂔	𠂔	𠂔	水	水
4. “year”	𠂔	𠂔	𠂔	𠂔	年
5. “silk”	𠂔	𠂔	𠂔	絲	絲
6. “be born”	𠂔	𠂔	𠂔	生	生
7. “eye”	𠂔	𠂔	目	目	目
8. “fruit”	𠂔	𠂔	果	果	果
9. “tripod”	𠂔	𠂔	鼎	鼎	鼎
10. “deer”	𠂔	𠂔	𠂔	鹿	鹿
11. “wise”	𠂔	𠂔	聖	聖	聖
12. “buy”	𠂔	𠂔	𠂔	買	買

and grammatical elements which were difficult to represent in pictorial form. Through the application of the so-called rebus principle, a pictograph or some other nonphonetic representational graph could be used for its sound value only. Many of the early graphs are derived by this “phonetic borrowing principle” – indeed, almost all functional words and grammatical elements are so represented.

One thing is certain. The individual graphs of the Shang writing system represent specific *words*, and most probably spoken words, in the Shang language. They do not directly represent *ideas*: “The notion which is sometimes encountered that Chinese characters in some Platonic fashion represent ideas rather than specific Chinese words is patently absurd” (Norman 1988: 60–61).

Later on, beginning in the Early Archaic period, another strategy for character formation is utilized which in subsequent centuries is to become increasingly important – that of phonetic compounding. A character of this type consists of a semantic element combined with a second element used to indicate the pronunciation of the new graph. About 85 percent of Chinese characters are presently of this type.

## 2.3 Evolution of the writing system

After the *jiaguwen* (“oracle bone script”), which already had 4,000 to 5,000 characters, and the *jinwen* (“bronze script”), another type of writing called *zhouwen* (or *dazhuan*) “large seal” appears, in conjunction with the bronze script, in the Spring and Autumn period (770–476 BC). By the end of the period, the use of this script has already spread to virtually all levels of society, leading to the development of many simplified forms – which may be called *demotic* – accelerating conventionalization and the movement away from pictographic symbols.

During the Warring States period (475–221 BC), a growing diversity among the scripts of the various states can be observed, due mainly to political fragmentation. These are called *liuguo wenzi* “scripts of the Six States.” The Qin dynasty, newly established in 221 BC, undertook a script standardization, sanctioning only two scripts – a complex form and a simplified demotic. The former, known as *zhuanshu* “seal script,” also frequently referred to as *xiaozhuan* “small seal,” is derived from the *jinwen* and *zhouwen* mentioned above. The latter of the reformed scripts, more important for the history of Chinese writing, is called *lishu* “clerical script.” It is highly evolved in its graphic form and represents a much simplified version of the standard seal script. All attempts to preserve the pictorial nature of graphs are abandoned, and convenience becomes the overriding principle.

Around the end of the first century BC, the demotic clerical script becomes the official form of writing employed for all purposes, and the use of the ancient script comes almost to an end. By the end of the Han period (c. third century AD), when approximately 10,000 characters exist (the *Shuo wen jie zi* [*Explanation of Graphs and Analysis of Characters*], compiled by Xu Shen around the second century AD, lists 9,353 different characters), the standard form of the script called *kaishu* begins to take shape. It represents a further evolution toward a more regular and convenient form of writing in which the wave-like strokes of the clerical script are replaced by more linear strokes. By the fifth century AD, *kaishu* becomes the standard form of Chinese script for all ordinary purposes, and it is still widely used at the present time.

### 3. PHONOLOGY

The earliest important analyses of Chinese historical phonology owe a great deal to the Swedish sinologist Karlgren, who was the first to apply the methods of European historical linguistics to Chinese. Karlgren provided two complete reconstructions: (i) one which he called Ancient Chinese (now Middle Chinese), based on the *Qieyun* (AD 601) by Lu Fayan, a dictionary of Chinese characters arranged by tone and rhyme, and (ii) the other which he called Archaic Chinese (now Old Chinese), based on the rhymes of the *Shi jing* (*Classic of Poetry*), a collection of 305 poems completed about 600 BC (see Karlgren 1957 for their almost definitive versions).

The hypotheses of Karlgren, after being modified by Jaxontov (1960 [1983]), Pulleyblank (1962, 1984, 1991), Li (1971), Bodman (1980), and Baxter (1980, 1992), have become obsolete today. At present, the most developed systems of reconstruction are those of Pulleyblank (1991) for Middle Chinese, and Baxter (1992) for Old Chinese.

#### 3.1 Reconstruction of Middle Chinese phonology

The source of Middle Chinese is the *Qieyun* dictionary. It represents most probably a single, coherent form of the Chinese language, namely the elite standard which was common to educated speakers from both north and south around the sixth century AD.

##### 3.1.1 Consonants and vowels

The Chinese syllable can be divided into two parts: (i) an initial (*shengmu*), the consonantal onset; and (ii) a final (*yunmu*), further divided into (a) a medial glide (*yuntou*), (b) a main

vowel (*yunfu*), and (c) a coda (*yunwei*). Upon the basis of this structure, the inventory of phonetic segments for Early Middle Chinese in Baxter's (1992:45–61) transcriptions of the traditional categories is presented below. Aspiration of stops and affricates is indicated by superscript <sup>h</sup>; *r* does not represent a separate segment, but a retroflex articulation of the preceding consonant; <sup>y</sup> indicates a palatal articulation of the preceding consonant. The velar nasal [ŋ] is represented by *ng*. Initial *h*- represents a voiced guttural fricative (probably pharyngeal [ħ] or velar [ɣ]), in contrast to *x*-, which is voiceless:

### (1) Middle Chinese initials

<i>Labials</i>	p	p <sup>h</sup>	b	m		
<i>Dentals</i>	t	t <sup>h</sup>	d	n		
<i>Lateral</i>						l
<i>Dental stridents</i>	ts	ts <sup>h</sup>	dz		s	z
<i>Retroflex stops</i>	tr	tr <sup>h</sup>	dr	nr		
<i>Retroflex stridents</i>	tsr	tsr <sup>h</sup>	dzr		sr	zr
<i>Palatals</i>	ts <sup>y</sup>	ts <sup>y</sup> h	dz <sup>y</sup>	n <sup>y</sup>	s <sup>y</sup>	z <sup>y</sup> y
<i>Velars</i>	k	k <sup>h</sup>	g	ng		
<i>Laryngeals</i>	ʔ				x	h

A *final* includes at least a main vowel, which may be followed by a coda, or may be preceded by one or more medials. The basic medials are the glides *-y-* and *-w-*.

The Middle Chinese main vowels are as follows:

### (2) Middle Chinese main vowels

i	ɨ	u
e		o
ɛ		
æ	a	

These main vowels may be followed by the codas of (3):

### (3) Middle Chinese codas

	w		y	i
ng	wng	m	n	
k	wk	p	t	

The combinations *-wng* and *-wk* may be taken literally, or interpreted as labiovelars /ŋ<sup>w</sup>/ and /k<sup>w</sup>/.

### 3.1.2 Tones

Middle Chinese has a system of four tones which, according to Chinese tradition, was first identified and named by Shen Yue in the fifth century: the tones are called *ping* “level,” *shang* “rising,” *qu* “departing,” and *ru* “entering.” Every Chinese syllable is marked by one of these four tonal categories. The entering tone occurs on all syllables which end in one of the three stops *p*, *t*, and *k*. Scholars have also argued that particular voice qualities are associated with the rising and departing tones (Mei 1970, Pulleyblank 1978, Sagart 1986).



### 3.2 Reconstruction of Old Chinese phonology

The reconstruction of an Old Chinese phase is much more problematic than Middle Chinese, since available evidence is more fragmentary. Such a reconstruction is based on two types of evidence: Old Chinese rhyming as reflected in *Shi jing*; and the phonetic series. Only the phonetic series gives us information on the initial consonants or groups of consonants.

#### 3.2.1 Consonants and vowels

The Old Chinese syllable is also analyzed as being composed of an *initial* and a *final* (cf. §3.1.1). Following Baxter (1992:7), whose inventory of phonetic segments is given below, the (i) initial contains a *preinitial* and an *initial*, and the (ii) final contains a *medial*, a *main vowel*, a *coda*, and a *postcoda*. The terms “preinitial” and “postcoda” are introduced since Old Chinese allows consonant clusters in both initial and final position.

Baxter (1992) reconstructs four preinitials (which are now treated as prefixes in Baxter and Sagart 1998; see §4.3.3), thirty-seven initials, three medials, six main vowels, ten codas, and two postcodas. The initials reconstructed for Old Chinese are presented below (this model is largely inspired by Pulleyblank 1962). The spellings *hm*, *hn*, *hng*, and so forth denote the voiceless counterparts of the sonorants *m*, *n*, *ng*:

#### (4) Old Chinese initials

p	p <sup>h</sup>	b	m	hm	w	hw
t	t <sup>h</sup>	d	n	hn	l	hl
					r	hr
					y	hy
					z	s
ts	ts <sup>h</sup>	dz				
k	k <sup>h</sup>	g	ng	hng		
k <sup>w</sup>	k <sup>wh</sup>	g <sup>w</sup>	ng <sup>w</sup>	hng <sup>w</sup>		
ʔ	x	fi				
ʔ <sup>w</sup>						

Three medial elements have been reconstructed: \*-r- (on the hypothesis of Jaxontov 1960), \*-y- (though the reconstruction of the medial \*-y- has now been replaced by a contrast of vowel length), and, marginally, \*-l-.

The six main vowels (after Bodman 1980) are as follows:

#### (5) Old Chinese main vowels

i	i	u
e		o
	a	

The following elements are reconstructed in the coda position:

#### (6) Old Chinese codas

	k	ng
y	t	n
w	wk	
	p	m

The two postcodas are \*-ʔ and \*-s, which are the respective sources of the rising tone and of the departing tone in Middle Chinese (the rising tone hypothesis is offered by Pulleyblank 1962 and Mei 1970; that of the departing tone by Haudricourt 1954).

This reconstructed Old Chinese phonemic system, the most recently proposed and probably the most complete, is far from being universally accepted. Several of Baxter's propositions are considered controversial and are being actively debated (for a detailed account of the controversial questions, see Pulleyblank 1993 and Sagart 1993a). In fact, several specialists still consider that even today one can do no more than approach a reconstruction of Old Chinese.

### 3.3 Significant diachronic processes linking Old and Middle Chinese

The main phonological developments from Old Chinese to Middle Chinese can be summarized as follows (for a more detailed account of these changes, see Appendix A of Baxter 1992:565–582):

1. The preinitial position was lost entirely as the preinitial elements (now prefixes) merged with the following initials to form single initial consonants.
2. The Old Chinese initials were also influenced by the following medials. Dentals developed into palatals when followed by \*-y- or retroflex stops when followed by \*-r-.
3. The vowel system of Old Chinese underwent radical changes under the influence of the medial and the coda.

There remains an important point of debate: did Old Chinese already have tones? In the past it was proposed that tone is an inherent feature of languages that cannot be derived from nontonal elements; accordingly, as Middle Chinese most likely had tones, Old Chinese must also have had tones. This hypothesis is today highly contested. Studies in recent years have shown that some present-day tonal languages (Vietnamese, for instance) are, indeed, derived from nontonal ancestral languages.

If the rising tone (*shang*) of Middle Chinese can be derived from the glottal postcoda of Old Chinese, as proposed by Pulleyblank (1962) and Mei (1970), and if the departing tone (*qu*) can be derived from the postcoda \*-s of Old Chinese, as proposed by Haudricourt (1954), then it turns out that there were no tones in Old Chinese. The two other tones of Middle Chinese can be interpreted as follows: (i) the level tone (*ping*) was the unmarked category consisting of those syllables ending in plain vowels or in other voiced segments; (ii) the entering tone (*ru*), as seen above, consisted of all the syllables ending in one of the three stops *p*, *t*, or *k* (for a different point of view, see Ting 1996, who argues that Old Chinese was already tonal).

## 4. MORPHOLOGY

Chinese is a language of that morphological type called analytic or isolating. Old Chinese morphemes are almost entirely monosyllabic, and most words are monomorphemic.

It is often said that Chinese is a language with an impoverished morphology, a language in which the grammatical processes are almost totally syntactic. Moreover, one usually considers that this lack of morphological marking of grammatical relationships is even more critical in Ancient Chinese than in Contemporary Chinese, since, as noted above, Old Chinese morphemes are almost entirely monosyllabic, and most words are monomorphemic.

Ancient Chinese did indeed possess morphological processes, although none of them was fully productive. These word-formation processes are of the same type as those of

Contemporary Chinese: compounding, reduplication, and affixation. But Ancient Chinese was characterized by yet other derivational processes, ones unknown in Contemporary Chinese.

## 4.1 Compounding

Not all words in Classical Chinese are monosyllabic; compounds occur which consist of two syllables. Most of these compounds are not yet fully lexicalized. They commonly consist of two independent free morphemes which can occur separately. Nevertheless, there are some exceptions – bound compounds occur which have meanings that cannot be deduced from the meaning of the morphemes from which they are composed. The most striking example is the word *junzi* “gentleman,” composed of *jun* “lord” and *zi* “child.”

Beginning in the Han period, Chinese develops a greater number of compounds. As new terms are required by the language, compounding is the chief means by which neologisms are introduced (owing to the death of Chinese derivational processes).

There are also in Classical Chinese some bimorphemic monosyllabic words, which result from the fusion of two morphemes. The negative *fu* is thus considered to be formed from the negative *bu* “not” and the third-person pronoun *zhi* “him, her, it.”

## 4.2 Reduplication

Reduplication is a productive morphological process. Archaic Chinese is quite rich in both total reduplicates and partial reduplicates. For the most part, reduplicated forms are expressive or descriptive adjectives or adverbs. Total reduplicates simply repeat the same syllable twice (e.g., *weiwei* “tall and grand”), whereas partial reduplicates only repeat the final part of the first syllable (as in *tanglang* “praying mantis”).

## 4.3 Affixation

Contrary to what is generally thought, affixation is not unproductive in Ancient Chinese, and may represent a vestige of older stages in which such a process was considerably more productive. Several prefixes, suffixes, and infixes have now been reconstructed for Archaic Chinese. These are derivational morphemes changing the meaning or part of the speech of the words to which they are attached (the ensuing discussion closely follows the treatment of Baxter and Sagart [1998]).

### 4.3.1 Prefixes

The following prefixes are reconstructed:

1. A prefix *\*N-* (causing a following voiceless obstruent to become voiced in Medieval Chinese) when attached to a verb (or even a noun in some cases) seems to produce an intransitive verb or adjective: thus, *kens* “to see” : *\*N-kens* “to appear.”
2. A prefix *\*k-* added to a verb or a noun produces, in several examples, a concrete, countable noun of related meaning: for example, *\*ʔjuj-s* “to fear, be afraid” : *\*k-ʔjuj-ʔ* “ghost, demon.” In some cases in which *\*k-* is added to verbs, forms with *\*k-* appear to refer to concrete actions taking place in a specific time frame: for example, *\*ljuk* “to nourish” : *\*k-ljuk* “to breast feed.”

3. A prefix *\*t-*, in contrast to *\*k-*, often appears to produce a derived mass noun, as in *\*ljuk* “nourish” : *\*t-ljuk* “rice gruel.” The same prefix also appears on some intransitive verbs.
4. A prefix *\*s-* derives causative verbs from noncausative verbs or even nouns. See Mei 1989.

#### 4.3.2 Suffix *\*-s*

The suffix *\*-s*, the source of the departing tone of Medieval Chinese (see §3.3), when added to adjectives or verbs produces derived nouns: for example, *\*drjon* “transmit” : *\*drjon-s* “a record.” Some gradable adjectives also have corresponding noun forms in which the suffix *\*-s* functions like English *-th*, occurring in pairs such as “deep/depth,” “wide/width” (see Downer 1959, Mei 1980).

The same *\*-s* suffix, in some instances, also makes transitive verbs from adjectives or intransitive verbs, or [+ give] dative verbs from [+ receive] dative verbs: thus, *\*dju?* “receive”: *\*dju?* -s “give”; *\*tsjAK* “borrow” : *\*tsjAK-s* “lend.”

#### 4.3.3 Infixes

Two infixes can be reconstructed in Archaic Chinese. The exact function of the first one, *\*-j-*, is difficult to establish, but forms with and without *\*-j-* do appear to be semantically related. The second infix, *\*-r-*, is said to produce forms that are plural or collective in the case of nouns, and iterative, durative, or indicating effort, in the case of verbs (see Sagart 1993b).

## 5. SYNTAX

Syntax is all the more critical in Classical Chinese, as words are not usually formally marked for grammatical category or function; words nevertheless do fall into distinct classes such as noun, verb, preposition, and so forth. Word order and the syntactic behavior of words are thus prominent linguistic issues.

### 5.1 Word order

The three basic word orders in Classical Chinese, as well as in Medieval, Modern, and Contemporary Chinese, are as follows: (i) the subject precedes the predicate; (ii) the verb precedes its object; (iii) modifiers precede the words they modify.

There is little controversy surrounding (i) and (iii), which tolerate only a few exceptions (for instance, the *subject–predicate* order is inverted in exclamatory sentences, as seen below in §5.4.2; see Zhu 1980:191 and Mei 1997 for an inverted order *head–modifier* in Early Archaic). However, things are quite different for (ii), which has been much debated since Li and Thompson (1974) put forward the hypothesis according to which Archaic Chinese was originally of SOV (*Subject–Object–Verb*) order, later being changed to SVO. It has been supposed that Proto-Chinese must have been SOV, and therefore Proto-Sino-Tibetan also since almost all Tibeto-Burman languages have a verb-final order (the only known exceptions being Karen and Bai).

Peyraube (1997a) argues that Pre-Archaic Chinese shows a regular order of SVO and is indeed more thoroughly SVO than later stages (Early or Late Archaic). To suppose, then,

that in a more ancient stage, before the oracle-bone inscriptions, the basic order could have been SOV is purely conjectural, not empirically grounded. Moreover, in the later stages of Early and in Late Archaic Chinese, there is also a strong indication that the SVO order is more basic than SOV (see Peyraube 1997b).

### 5.1.1 VO versus OV

Unlike many European languages which require an overt subject, Chinese does not seem to have such a syntactic requirement. It seems preferable then to frame our discussion in terms of VO versus OV order rather than SVO versus SOV. When the object is a full lexical noun phrase, the basic order in Archaic Chinese is undoubtedly VO, as in the following example:

- (7) jun        bi                shi        guo  
       prince    certainly    lose    state  
       "The prince [will] certainly lose the State" (Zuo Tradition)

There are a few cases in which the noun phrase object is found in preverbal position, but these cases are marginal and the OV order is then a marked [+ contrastive] order. It is the same when the noun object is followed by a preverbal marker, usually *shi* or *zhi*, as in (8):

- (8) jin        Wu        shi                        ju  
       now    Wu    OBJECT-MARKER    afraid  
       "Now [they] are afraid of [the state of] Wu" (Zuo Tradition)

However, there are also cases of OV order in Archaic Chinese, not found in Contemporary Chinese, in which the object is a pronoun: either (i) an interrogative pronoun (9A); (ii) the demonstrative pronoun *shi* "this" (9B); or (iii) a pronoun in a negative sentence (9C):

- (9) A. wu        shei        qi?                qi                tian                hu?  
       I        who    deceive    deceive    Heaven    INTERR.-PCL.  
       "Whom should I deceive? Should I deceive Heaven?" (Confucian Analects)  
       B. zi        zi        sun                sun                shi        shang  
       son    son    grandson    grandson    this    supersede  
       "[His] posterity will supersede this" (Chen gong zi yan, a bronze inscription)  
       C. bu                wu        zhi                ye  
       NEGATION    I        understand    FINAL-PCL.  
       "[You] don't understand me" (Confucian Analects)

Certain observations can be made concerning these various OV orders involving pronouns. First, there are statistical considerations. In a corpus of 2,767 VO or OV sentences drawn from the bronze inscriptions, 88.56 percent of objects (O) are nouns, only 3.3 percent are pronouns (see Guan 1981:88). The ratio of pronoun objects is certainly higher in other documents of the Early Archaic period, and above all in Late Archaic Chinese, but it never exceeds 15 percent of the entire body of VO and OV constructions. Since the OV order is well attested only for pronoun objects, one can conclude with some confidence that the OV order has always been very marginal.

It is also known that in many languages, the position of pronouns is different from that of noun phrases, and that "unstressed constituents, such as clitic pronouns, are often, cross-linguistically, subject to special positioning rules only loosely, if at all, relating to their grammatical relation, so sentences with pronouns can be discounted in favor of those with full noun phrases" (Comrie 1989: 89).

### 5.1.2 Prepositions or postpositions?

In Classical Chinese, prepositional phrases are usually composed of a preposition (see §5.2.2.3) followed by a noun phrase object, as in the following example:

- (10) Zizhi bu neng shou Yan yu Zikuai  
 Zizhi NEGATION can receive Yan from Zikuai  
 “Zizhi cannot receive [the state of] Yan from Zikuai” (Mencius)

These prepositional phrases can be postverbal (as in [10]), or preverbal, as in the following:

- (11) gu yi yang yi zhi  
 therefore with sheep change it  
 “Therefore [I] changed it for a sheep” (Mencius)

Of the two common prepositions of Archaic Chinese, *yu* and *yi*, *yu* has a relatively rigid postverbal position, while *yi* is more preverbal.

More interesting for the problem of word order is that there are cases in which the preposition is found after the noun phrase object. In such instances the “preposition” is thus a postposition. Some scholars (Sun 1991, Mei 1997) have hypothesized that these postpositions are relics of an ancient general order, and, accordingly, that Chinese may have been a postpositional language. Consider the following example:

- (12) shi yi zheng ping  
 this with politics pacify  
 “With this, the politics [will] pacify [the State]” (Zuo Tradition)

However, we should bear in mind that these occurrences are very rare, especially if we exclude the cases in which the object is an interrogative pronoun or the demonstrative pronoun *shi* “this.” The order OP (Object–Preposition) when the object is such a pronoun naturally follows from the rule of positioning these pronouns before the verb (see §5.1.1), as prepositions in Chinese develop diachronically from verbs and still share many properties with them.

To sum up, no OV order needs to be posited for Classical Chinese syntax to capture any sort of linguistic generalization. Classical Chinese has SVO order, just as in the ensuing stages of Medieval, Modern, and Contemporary Chinese.

## 5.2 Parts of speech

Classical Chinese words are traditionally divided into two categories: *shizi* “full words” and *xuzi* “empty words.” The former are content words (carry semantic content) and form an open class; included in this category are nouns, verbs, and adjectives. The latter are function words or grammatical words, used to express grammatical relationships. They include pronouns, adverbs, prepositions, conjunctions, and particles (for the analysis of Classical Chinese word classes presented here, see Liu 1958:18; other scholars [Wang 1979:36–41; Ma 1983:13] distinguish eleven word classes).

As we will see, words can be used in functions customarily reserved for other words. This does not imply, however, as some scholars have assumed, that there are no parts of speech in Classical Chinese, and that words can be used indifferently in any grammatical category.

### 5.2.1 Shizi (full words)

Treated under this heading are nouns, adjectives, and verbs, including auxiliary verbs.

#### 5.2.1.1 Nouns

Chinese nouns typically function as subjects or objects. However, under certain conditions, they may function like verbs, as predicates (13), or like adverbs, as adverbials (14):

- (13) jun    jun    chen    chen    fu    fu    zi    zi  
       ruler ruler minister minister father father son son  
       “The ruler acts as a ruler, the minister as a minister, the father acts as a father, and  
       the son as a son” (Confucian Analects)

- (14) shi    ren    li  
       pig    man    stand-up  
       “The pig, like a man, stood up” (Zuo Tradition)

Localizers (words showing spatial orientation and direction, like *shang* “above,” *nan* “south”), time words (like *ri* “day,” *yue* “month,” etc.) and measure words (indicating standards for length, weight, volume, area, aggregates, containers – like *dou* “bushel,” *bei* “glass”) are better considered as subcategories of nouns, though some scholars treat them as independent word classes; see Norman 1988:91; Wang 1979:38. For a history of measure words in Classical Chinese, see Peyraube 1991.

#### 5.2.1.2 Verbs

Fundamentally, verbs are predicative in nature. Unlike nouns, which are negated by the adverb of negation *fei* “is not,” verbs are negated by the simple adverb *bu* “not.”

Both intransitive (e.g., *yi lai* [lit. doctor come] “the doctor came”) and transitive verbs occur; the latter may take a single object or, sometimes, two – an indirect and a direct:

- (15) gong    ci    zhi    shi  
       prince offer him food  
       “The prince offered him food” (Zuo Tradition)

One particular use of intransitive verbs in Classical Chinese is in a causative function: thus, *huo* “live” : “make (people) live”; *xing* “go” : “put into motion”; *yin* “drink” : “give to drink.”

One verbal subclass is composed of auxiliary verbs. Auxiliary verbs are verbs that take other verbs as their objects and express the modality of the following verb phrase. This modality (ability, possibility, probability, certainty, obligation, volition, etc.) can be characterized as epistemic, deontic, or dynamic. Auxiliary verbs form a closed list and can be classified in the four following semantic groups: (i) verbs expressing mainly possibility and permission, including *ke*, *neng*, *zu*, *de*, *huo*, *keyi*, and *zuyi* (see [16A]); (ii) the four verbs of volition, *gan*, *ken*, *yu*, and *yuan* (see [16B]); (iii) the two auxiliaries of necessity (certainty and obligation), *yi* and *dang*; and (iv) the passive auxiliaries *jian*, *wei*, and *bei*. For a detailed analysis of the modal auxiliary verbs in Chinese, see Peyraube 1999.

- (16) A. tian    zi    bu    neng    yi    tianxia    yu    ren  
       Heaven son negation can OBJECT-MARKER Empire give other  
       “The Emperor cannot give the Empire (to) others” (Mencius)  
       B. Zi    yu    ju    Jiu    Yi  
       Master intend-to live Jiu Yi  
       “The Master intends to live in Jiu Yi” (Confucian Analects)

### 5.2.1.3 Adjectives

Adjectives can be considered as a subcategory of verbs. Indeed, they are intransitive verbs of quality, being negated by the adverb *bu*:

- (17) ming bu zheng ze yan bu shun  
 name not correct then word not justified  
 “If names are not correct, then words cannot be justified” (Confucian Analects)

Like intransitive verbs, adjectives can also have a causative use:

- (18) Wang qing da zhi  
 king beg great it  
 “Your Majesty, [I] beg [you] to make it great” (Mencius)

In addition, adjectives are also typically found as noun phrase modifiers, as in *bai ma* “white horse,” or as verb phrase modifiers, for instance *ji zou* (lit. rapid - run) “run rapidly.”

### Numerals

Finally, one can consider that numerals constitute a subclass of the category of adjectives. They indeed behave syntactically like adjectives; thus, they can form predicates and are negated by the adverb *bu*. Most commonly, however, they function as modifiers of nouns:

- (19) A. nian yi qi shi yi  
 age already seven ten final-part  
 “[He] is already seventy years old” (Mencius)  
 B. wu he ai yi niu?  
 I why begrudge one ox  
 “Why [should] I begrudge one ox?” (Mencius)

## 5.2.2 Xuizi (empty words)

Within this category fall pronouns, adverbs, prepositions, conjunctions, and particles.

### 5.2.2.1 Pronouns

Several types of pronouns can be identified: personal, demonstrative, interrogative, and indefinite.

#### 5.2.2.1.1 Personal pronouns

Personal pronouns characteristically occur in different forms. The most common ones are as follows, with no distinction being made between singular and plural:

- (20) First person wu wo yu  
 Second person ru er ruo nai  
 Third person zhi qi

If it is relatively easy to distinguish third-person *zhi* and *qi* as accusative and genitive respectively, it is not so for pronouns of the other two persons. Several scholars have tried to characterize their different usages according to case (nominative, accusative, or genitive), but dialectal variation also is a factor: in most instances, the different usages of the pronouns depend on the different texts in which they occur.



#### 5.2.2.1.2 Demonstrative pronouns

The most common demonstratives are (i) *shi*, *ci*, *si*, *zhi*, and *zi* “this, these, here”; and (ii) *bi*, *fu*, and *qi* “that, those, there” (Pulleyblank [1995:85] states that *shi* “this, that” is anaphoric with no implication of closeness or remoteness; *ci* and *bi*, on the other hand, form a contrast between “this (here)” and “that (there)”). Here too, it is difficult to explain formal differences without considering dialectal variation. All of these demonstratives can be used as adjectivals (modifying the following nouns or noun phrases), or as subjects or objects.

#### 5.2.2.1.3 Interrogative pronouns

These are divided into two categories: (i) those that replace subjects or objects (which are usually nouns): *shui* “who,” *shu* “which, who,” *he* “what”; and (ii) those that replace predicative verbs or adverbs: *hu* “why, how,” *xi* “why,” *he* “how, why,” *an* “where, how,” *yan* “how, where,” *wu* “how, where.” An interrogative pronoun precedes the verb of which it is the object.

#### 5.2.2.1.4 Indefinite pronouns

This class includes *huo* “some, someone, something,” *mo* “none, no one, nothing” and *mou* “some, a certain one.”

#### 5.2.2.2 Adverbs

Usually positioned in preverbal position, adverbs typically modify the predicate of the sentence. One can distinguish several types: (i) adverbs of degree (*ji* “extremely,” *zui* “most,” *you* “especially,” *shao* “little,” *shen* “very,” etc.); (ii) adverbs of quantification and restriction (*jie* “all,” *ju* “all,” *ge* “each,” *mei* “every,” *wei* “only,” *du* “only,” etc.); (iii) adverbs of time or aspect (*yi* “already,” *ji* “after having,” *chang* “once,” *jiang* “be going to,” *nai* “then,” *fang* “just then”; (iv) adverbs of negation (*bu*, *fu*, *fei*, *wu*, *wei*). *Bu* is the ordinary adverb of negation for verbs and adjectives. *Fu* is said to be the result of the fusion of the negative *bu* plus the object pronoun *zhi* “him, her, it” (only found during the Late Archaic period). *Fei* “is not” is the negation used with nouns. *Wu* “do not” could be a blend of *wu* plus *zhi*. *Wei* is an aspectual negative meaning “not yet” or “never.”

#### 5.2.2.3 Prepositions

Chinese prepositions are all verbal in origin (i.e., arise from verbs through a process of grammaticalization). There are two commonly occurring prepositions in Classical Chinese: (i) *yu* “at, to, in, from, toward, than, by, etc.”; and (ii) *yi* “with, by means of, in order to, because, etc.” The first of these, *yu*, can be locative, ablative, dative, comparative, or passive; the second, *yi*, primarily instrumental, also expresses purpose and several other grammatical relationships. One important characteristic of *yi* is that it can also introduce the direct object of a double-object construction (see [10] above). Additional prepositions are *yong* “with,” *wei* “for, on behalf of, for the sake of, because,” *yu* “with,” *zi* “from,” among still others.

#### 5.2.2.4 Conjunctions

Generally, simple juxtaposition is sufficient to coordinate nouns or noun phrases, as in *fu mu* (lit. father mother) “father and mother”; or verbs or verb phrases in serial verb constructions. However, some coordinative conjunctions also occur, such as *ji* and *yu* “and,” for coordinating noun phrases, or *er* “and” and *qie* “and, moreover” for coordinating verb phrases or clauses.

Liu and Peyraube (1994) have argued that the conjunctions *ji* and *yu* do not directly develop from verbs (as has been claimed), but from prepositions, which are themselves derived from verbs. In other words, two processes of grammaticalization have occurred sequentially: verb > preposition > conjunction (Chinese conjunctions are thus more grammaticalized than verbs).

Subordinating conjunctions also occur: for example, *ru*, *ruo*, or *gou*, all meaning “if” in conditional clauses; *sui* “although, even if” in concessive clauses (see §5.4.5).

#### 5.2.2.5 Particles

This category is usually divided into structural particles (*zhi*, *suo*, *zhe*) and modal particles. For structural particles, see §5.2.3. Modal particles constitute one of the most complex problems in Classical Chinese linguistics; most of the modal notions they express are quite uncertain. Modal particles can occupy the initial, the medial, and, in most cases, the final position of a sentence.

Among the initial particles, we find the following: *qi*, which qualifies a statement as possible or probable; *qi* “how could,” which introduces rhetorical questions requiring a negative answer; and *fu* “as for,” which announces a topic. Medial particles usually express a pause: for example, *zhe* and *ye*. The final particles can be divided according to the sentence-types in which they occur – declarative, interrogative, exclamatory, and so on. In declarative sentences, one often finds *yi* (a particle of the perfect aspect; see Pulleyblank 1995:112–116), *ye* (transforming a statement into an assertion, a judgment), *er*, and *yan*. *Hu*, *yu*, *ye*, and sometimes *zhe*, are more typically used in interrogative sentences. *Zai* occurs in exclamatory sentences.

### 5.3 Elements of sentence structure

#### 5.3.1 Subject and predicate

Classical Chinese sentences can, in general, be divided into two main parts, a subject (most commonly a noun phrase) and a predicate (usually a verb phrase), though the subject may be – and indeed often is – unexpressed:

When the predicate is composed of more than one verb, it is said to be *complex*. Such cases involve serial verb constructions of the type  $V_1 \dots V_2 \dots (V_3) \dots$ . The semantic relationship between verbs in series is varied. It can be a simple narrative sequence (in which case a coordinating conjunction *er* “and” can link the two verb phrases), as in (21A); or the relationship may involve an implication of purpose, as in (21B), where *yi* links the two verb phrases:

- (21) A. Shao wang nan zheng er bu fu  
 Shao prince south invade and not return  
 “Prince Shao invaded the south and did not return” (Zuo Tradition)
- B. Chu ren fa Song yi qiu Zheng  
 Chu people raid Song for save Zheng  
 “The people of Chu raided Song in order to save Zheng” (Zuo Tradition)

Complex predicates may also involve a “pivotal construction,” in which the noun phrase object of the first verb is the subject of the second verb:

- (22) qing jun tao zhi  
 ask Prince attack him  
 “[I] ask [you] the Prince to attack him” (Zuo Tradition)

Existential sentences form a special category of subject–predicate sentences. The predicate is composed of either the verb *you* “there is” or *wu* “there is not.” In such sentences, the subject is often lacking or expressed by a place name.

Nominal predicates will be discussed below, under copular sentences (§5.4.3).

Finally, an interesting characteristic of the subject–predicate constructions is that they can be nominalized by inserting the subordinating particle *zhi* between the two constituents of the construction:

- (23) ren      zhi      ai      ren      qiu      li      zhi      ye  
 person PCL. love person pursue profit him PCL.  
 “One person loving another person [would] pursue profit [for] him” (Zuo Tradition)

### 5.3.2 Object and complement

A transitive verb can take one object (usually a noun phrase) or two, an indirect object (IO), and a direct object (DO). The double-object construction is restricted to those verbs with the semantic feature [+ give], [+ say] or [+ teach]: Apart from the pattern *V + IO + DO*, as in (9) above, two other orders, involving the prepositions *yi* (DO marker) or *yu* (“to,” introducing the IO), are possible for the dative construction: (i) *yi + DO + V + IO* (or *V + IO + yi + DO*), also restricted to verbs which are [+ give], [+ say], [+ teach]; (ii) *V + DO + yu + IO*, used with all kinds of verbs (for a detailed analysis of these constructions, see Peyraube 1987):

- (24) A. Yao    yi                      tianxia    yu    Shun  
       Yao   OBJECT-MARKER   Empire   give   Shun  
       “Yao gave the Empire to Shun” (Mencius)  
       B. Yao   rang   tianxia    yu   Xu   You  
       Yao   leave   Empire   to   Xu   You  
       “Yao left the Empire to Xu You” (Zhuangzi)

Transitive verbs, as well as intransitive ones, may be followed by a *complement* (*buyu*), a term used for adjuncts when they follow the verb. When adjuncts precede the verb, they are denoted as *adverbials* (see §5.3.3). As many “complements” may also be placed in front of the verbs, and are thus “adverbials,” the function of the complement per se is not very important in Classical Chinese (see Ma 1983:135 for discussion).

Complements are divided into two types, depending upon whether or not they are introduced by a prepositional marker. Those that are introduced by such a marker, of course, constitute prepositional phrases. Most notable of this type are the locative complements, usually introduced by the preposition *yu* “at, to”:

- (25) bei      xue    yu    zhong    guo  
       north learn at central state  
       “He went to the north to learn [it] in the Central States” (Mencius)

Compare (26), having a prepositional phrase complement introduced by *yi*, with (19), where *yi* introduces an adverbial:

- (26) yi      zhi    yi      yang  
       change it with sheep  
       “Change it for a sheep” (Mencius)

Among complements not introduced by prepositions, conspicuous are time complements (though these may at times also be introduced by the preposition *yu*) and durative complements, as in the following example:

- (27) Zi Yi zai wei shi si nian yi  
 Zi Yi be-at throne ten four year PCL.  
 “Zi Yi has been on the throne during fourteen years” (Zuo Tradition)

### 5.3.3 Adjectivals and adverbials

Adjectivals (*dingyu*) are modifiers of nouns or noun phrases; adverbials (*zhuangyu*) are modifiers of verbs or verb phrases. As a general rule in Classical Chinese, modifiers precede their heads.

Subordinate relations involving nouns are expressed as follows:  $N_2 + zhi + N_1$ , where  $N_1$  is the head of the phrase,  $N_2$  the modifier (adjectival), and *zhi* the marker of subordination. This marker may be omitted, especially between monosyllables. The same pattern,  $X + zhi + N$ , is found when the modifier  $X$  is a verb or an adjective, as in the following:

- (28) wu duo ren zhi jun  
 insult rob people SUBORD.-PCL. ruler  
 “A ruler who insults and robs [his] people” (Mencius)

Like other modifiers, relative clauses take *zhi* as a marker of subordination.

Two other markers of nominalization are *zhe* and *suo*. The first one may be called an agentive marker. Placed after a verb or a verb phrase, it produces an agent noun phrase: *sha zhe* (kill the-one-who) “The one who kills.” The marker *suo*, placed before the verb, gives a noun phrase referring to the object of a transitive verb, as in: *suo sha* (*suo* kill) “that which was killed”; *qi suo shan* (his *suo* good) “that which he considers to be good.”

Adverbials (verb phrase modifiers), are most commonly (and expectedly) adverbs, though they may also be nouns (as in [8] above), adjectives (29A), or prepositional phrases (29B):

- (29) A. wang zu da bai  
 prince finally great defeat  
 “The prince was finally defeated terribly” (Zuo Tradition)  
 B. wo yu Zhou wei ke  
 I at Zhou become host  
 “I became a host at Zhou” (Zuo Tradition)

No marker is needed between the modifier and the head.

## 5.4 Sentence-types

Sentences are customarily divided into simple and complex types. One can also differentiate declarative, interrogative, imperative, and exclamatory sentences.

As the preceding analyses and examples have been concerned principally with the simple declarative sentence, we will treat here the remaining three types of simple sentences (interrogative, imperative, exclamatory). To these we would also add two particular types of declarative sentences: copular (i.e., nominal predicate sentences) and passive. For complex sentences see §5.4.5.

### 5.4.1 Interrogative sentences

These are of three basic types: (i) yes/no questions; (ii) WH-questions; and (iii) rhetorical questions.

The first type is formed with final question particles which effectively transform statements into questions. As noted above (§5.2.2.5), the most common particles are *hu*, *yu*, *ye*:

- (30) A. *zi yi you yi wen hu?*  
 master also have different hear PCL.  
 “Master, have [you] also heard of different things?” (Confucian Analects)
- B. *wang zhi suo da yu ke de wen yu?*  
 prince PCL. PCL. great desire can obtain hear PCL.  
 “[What] you [the Prince] greatly desire, could obtain a hearing [of it]?”  
 (Mencius)

The second type (WH) contains a question word (one of the interrogative pronouns; see §5.2.2.1.3), generally without a final particle, as in the following:

- (31) *Zi Xia yun he?*  
*Zi Xia say what*  
 “What did Zi Xia say?” (Confucian Analects)

Note that the interrogative pronoun here follows the verb and is not in a preverbal position, as is usually the case.

The third interrogative-type is more complex. Some rhetorical questions are formed with a final particle (*hu*, *yu*, or *ye*), but with an adverb of negation placed before the verb, implying an affirmative answer. Others are formed with the modal particles *qi* or *qi* “how could”. The final particles *hu* and *zai* are also generally used, though they may be omitted:

- (32) *yu qi hao bian zai?*  
 I how-could like debate PCL.  
 “How could I [be one who] loves debating?” (Mencius)

### 5.4.2 Imperative and exclamatory sentences

Imperatives are not syntactically marked as such in Classical Chinese. The subject is usually deleted, but this in itself is not a sufficient diagnostic of the imperative sentence. However, when the imperative is intended to be understood as a request, and not as an order or a prohibition, the verbs *yuan* “wish” or *qing* “beg” are used:

- (33) *wang qing du zhi*  
 prince beg measure it  
 “[My] Prince, please measure it” (Mencius)

The final particle *zai* is the usual marker of the exclamatory sentence. It can be added either to a declarative or to an interrogative. Other particles, like *yi*, may also be used. The subject–predicate order is usually inverted in exclamatory sentences. Consider the following examples:

- (34) A. *xian zai Hui ye!*  
 sage PCL. Hui PCL.  
 “[He] is a sage, Hui!” (Confucian Analects)

- B. si yi Pencheng Kuo!  
 dead PCL. Pencheng Kuo  
 “He is dead, Pencheng Kuo!” (Mencius)

### 5.4.3 Copular sentences

If one defines the copula as an overt word which, when used in equational sentences, links the subject to a nominal predicate, and expresses (i) an equivalence meaning or (ii) a property or classificatory meaning, then one can identify the presence of copulas in Classical Chinese (even if they are not strictly necessary).

The most common way of creating copular sentences is to add the final particle *ye* at the end of a sentence, transforming it from a statement into an assertion or a judgment (see Peyraube and Wiebusch 1995 for a detailed account of the history of copulas in Ancient Chinese, and especially for a discussion of the status of *ye* as a copula):

- (35) bi zhangfu ye wo zhangfu ye  
 that reliable-man PCL. I reliable-man PCL.  
 “They were reliable men, I am a reliable man [too]” (Mencius)

In addition to *ye*, other copulas are attested in Classical Chinese. Thus, the negative copula *fei* “be not” is required in all negative nominal predicate sentences:

- (36) wo fei sheng er zhi zhi zhe  
 I be-not born and know it the-one-who  
 “I am not one who was born with [the possession of] knowledge” (Confucian Analects)

In affirmative copular sentences, the verb *wei*, which also means “to do, to regulate, to act, to consider as,” and so forth, also acts regularly as a copula:

- (37) er wei er wo wei wo  
 you be you I be I  
 “You are you [and] I am I” (Mencius)

Finally, the copular verb *shi* “to be,” still used today, and which comes from the demonstrative pronoun *shi* “this” through a grammaticalization process, is already attested no later than the Qin dynasty (from an astrological document discovered in a tomb at Mawangdui; second century BC):

- (38) shi shi zhu hui ren zhu you si zhe  
 this be bamboo comet man chief have die the-one-who  
 “[When] this will be the bamboo comet [coming], the sovereign will die”

### 5.4.4 Passive sentences

In Classical Chinese there are semantic passives, expressing passivity without any overt morphological marker: a transitive verb can be made passive by placing its object (the patient) in subject position, as in *liang shi* (lit. “supplies eat”) “supplies are eaten.” However, there are also passive structures marked with some marker, such as a preposition, or an auxiliary verb.

In Early Archaic Chinese, there is only one passive construction, formed with the preposition *yu* “by” used to introduce the agent (*V + yu + Agent*). The construction can still be found in Late Archaic, where it is by far the most common means of producing a passive verb:

- (39) zhi yu ren zhe shi ren zhi ren zhe shi yu ren  
 rule by other the-one-who feed other rule other the-one-who feed by other  
 “Those who are ruled by others feed others, those who rule are fed by others”  
 (Mencius)

Two other structures appear in Late Archaic: *wei + V* and *jian + V*. The agent is not expressed, and *wei* and *jian* are best considered to be auxiliary verbs:

- (40) A. chen yi wei ru yi  
 I already AUX.-VERB humiliate PCL.  
 “I was already humiliated” (Lü shi chun qiu)  
 B. Pencheng Kuo jian sha  
 Pencheng Kuo AUX.-VERB kill  
 “Pencheng Kuo was killed” (Mencius)

Still within the Late Archaic period, these constructions are modified so that an agent can be expressed: *wei + Agent (+ suo) + V*; *jian + V + yu + Agent*. The first of these two will become common, beginning in the second century BC. It is probable that the auxiliary *wei* so used to introduce an overt noun phrase agent has in fact been grammaticalized as a preposition meaning “by”:

- (41) hou ze wei ren suo zhi  
 late then by other PCL. control  
 “[If I react] late, [I] will then be controlled by others” (Records of the Historian)

Yet another passive form appears at the end of the Classical period: *bei + V*, where *bei* is a verb meaning “to suffer,” “to be affected.” It will later become an auxiliary verb expressing passivity:

- (42) Cuo zu yi bei lu  
 Cuo finally because-of suffer slaughter  
 “Because of [this], Cuo was finally slaughtered” (Records of the Historian)

One must then wait for several centuries (until the Early Medieval period) before the auxiliary verb *bei* itself comes to be used to introduce a noun phrase agent, and then is grammaticalized into the passive preposition which is still in use today. For a detailed analysis of the passive forms in Ancient Chinese, see Peyraube (1989b).

### 5.4.5 Complex sentences

Complex sentences are composed of two or more clauses joined through coordination or subordination. The joining of clauses can be accomplished without any overt marking, as in the following examples:

- (43) A. lao zhe an zhi pengyou xin zhi shao zhe  
 old the-one-who soothe them friend trust them young the-one-who

huai      zhi  
care for    them

“As for the old, soothe them, as for friends, trust them, as for the young, care for them” (Confucian Analects)

B. bu      duo      bu      yan  
not    snatch    not    satisfy

“[If they] are not snatching, [they] are not satisfied” (Mencius)

A connective may also link the clauses, for instance the conjunction *er* “and, but” or the adverb *yi* “also,” in the case of coordination:

- (44) renmin   shao   er   qin   shou   zhong  
people   few   but   bird   beast   numerous  
“People are few but [wild] animals are numerous” (Han Feizi)

Subordination may be indicated by subordinating conjunctions or particles, which can occur in the first clause, in the second, or in both.

In the case of conditional sentences, the conjunctions *ru*, *ruo* or *gou* “if” may appear in the first clause (*if*-clause), and the markers *ze* or *si* “then” in the main clause (for an exhaustive analysis of the conditionals in Classical Chinese, see Harbsmeier 1981:229–287):

- (45) wang   ruo   yin   qi   wuzui   er   jiu   si      di   ze   niu   yang  
Prince   if   pain   it   no guilt   and   go-to   execution   place   then   ox   sheep  
he   ze      yan  
what   choose   PCL.  
“If [you] the prince were pained by its going without guilt to the place of execution, then what was there to choose between an ox and a sheep?” (Mencius)

In concessive sentences, the most commonly used conjunction of concession is *sui* “although, even if.” In the main clause one often finds *er*, which then has its adversative meaning:

- (46) sui      zhi,      er   bu   bing  
though   outspoken   yet   not   blame  
“Though [he may] be outspoken, [he won’t] be blamed” (Zhuangzi)

In sentences expressing cause, the “because” clause may be introduced by the preposition *yi*, and the main clause may contain the connective *gu* “so, therefore”:

- (47) yi      qi   bu      zheng,   gu      tian   xia   mo   neng   yu  
because   he   NEGATION   compete   therefore   Heaven   under   nobody   can   with  
zhi   zheng  
him   compete  
“Because he does not compete, nobody can compete with him under Heaven”  
(Laozi)

Time clauses are introduced by the prepositions *ji* or *dang*:

- (48) dang   zai   Song   ye,   yu   jiang   you   yuan   xing  
when   be-at   Song   PCL.   I   intend   there-is   far   go  
“When I was in Song, I intended to go far away” (Mencius)



## 5.5 Significant diachronic developments between Late Archaic and Early Medieval Chinese

The above study is concerned chiefly with the Classical Language as it is fixed during the Warring States period, i.e., the Late Archaic period (fifth–second centuries BC). From the time of the Early Medieval period (second–sixth centuries AD), one can consider that the vernacular language is actually distinct from the literary, deserving a separate description of its own. Here we will only discuss certain important grammatical structures which did not exist in Classical Chinese but which developed later in the vernacular, prior to the sixth century AD. For a detailed review of the developments in the language between these two stages, see Peyraube 1996.

The so-called “disposal form” appears around the sixth century, having the following structure: *Noun Phrase*<sub>1</sub>-Agent + *BA* + *Noun Phrase*<sub>2</sub>-Patient + *Verb Phrase* where *BA* is a preposition (*ba*, *jiang*, *chi*, or *zhuo*) which introduces the patient noun phrase. These prepositions were verbs in Classical Chinese meaning “to lead, to take, to hold.” Used for *V*<sub>1</sub> in a serial verb construction *V*<sub>1</sub> + *O* + *Verb Phrase*<sub>2</sub> in the Pre-Medieval period, they were grammaticalized and became prepositions, probably by analogy with the dative construction *yi* + *DO* + *V* + *IO* discussed above (see §5.3.2), where *yi* was already a marker introducing a direct object (see Mei 1990; Peyraube 1989a).

Locative prepositional phrases introduced by the preposition *yu*, which are postverbal in Classical Chinese, begin shifting to preverbal position in Pre-Medieval. The preposition *yu* is replaced by *zai* around the sixth century, when *zai*, a verb meaning “to be at” (and also used as *V*<sub>1</sub> in a serial verb construction *V*<sub>1</sub> + *O*<sub>1</sub> + *V*<sub>2</sub> + *O*<sub>2</sub>) has already been grammaticalized as a locative preposition (see Peyraube 1994).

The resultative construction, of the type *dasi* (lit. beat-die) “beat to death,” also appears in the Early Medieval period, and is not found in Classical Chinese, contrary to what some scholars have argued. What we find prior to the fifth century is a *V*<sub>1</sub> + *V*<sub>2</sub> serial verb construction in which *V*<sub>2</sub> is a transitive verb. The resultative compound arises from this serial verb construction, when the transitive verb has become intransitive (see Mei 1991).

*Classifiers* (CLs) do not exist in Classical Chinese (a classifier being fundamentally a word which, in theory, must occur before a noun and after a demonstrative and/or a number or another quantifier, marking the class to which the associated word belongs). In that period we find only *measure words* (MWs), which are first used in postnominal position, and then in the prenominal position in Late Archaic: *Noun* + *Number* + *MW* > *Number* + *MW* + *Noun*. True classifiers probably begin to appear during the Han period (first century BC), though at that time they still retain many characteristics of the nouns from which they issue, and they are always postnominal.

The grammaticalization process by which classifiers arose, depriving them of their original meanings, is a long one. For a great majority of them, it is completed only around the sixth century. By the time the process has been completed, classifiers have moved into prenominal position: *N* + *Num* + *CL* > *Num* + *CL* + *N* (see Peyraube and Wiebusch 1993).

Several other important developments have taken place between the Classical period and the Tang dynasty. I will mention here briefly the following:

1. *Personal pronouns*: A distinction between the two first-person pronouns *wo* and *wu* (see §5.2.2.1.1) gradually disappears. The third-person pronouns *qi* (genitive) and *zhi* (accusative) are replaced by new forms *yi*, *qu*, and (later) *ta*, and are no longer differentiated according to case. True plural forms, with the markers of plurality *deng*, *cao*, or *bei* following the pronouns, develop.

2. *Negatives*: The great number of adverbs of negation found in Classical Chinese is greatly reduced in the later vernacular. *Fu*, which becomes the most common negative, is no longer construed as a blend of *bu* + *zhi*. *Wu*, likewise, is no longer seen as a blend of *wu* + *zhi*.
3. *Localizers*: Monosyllabic in the Classical period, they become disyllabic, beginning in the Pre-Medieval period, by the addition of a suffix *-tou*.
4. *Disjunctive questions*: These appear in the fifth century, with *wei* serving as a disjunctive question marker, used singly or in pairs: *Noun*<sub>1</sub> + *Verb Phrase*<sub>1</sub> + *wei* + (*Noun*<sub>2</sub>) + *Verb Phrase*<sub>2</sub> or *Noun*<sub>1</sub> + *wei* + *Verb Phrase*<sub>1</sub> + *wei* + (*Noun*<sub>2</sub>) + *Verb Phrase*<sub>2</sub> (see Mei 1978).

## 6. LEXICON

The overall lexicon of Ancient Chinese is quite different from that of Contemporary Chinese. The former is composed of: (i) words that are still attested in the contemporary language, like *shan* “mountain” or *shui* “water”; (ii) words that only exist in the ancient language, and have disappeared from the modern language, such as *yue* “say”; (iii) words that are still used today, but with different meanings, like *zou* “run” (Ancient Chinese) > “walk” (Contemporary Chinese). Of the three types, the first are rare and the last are numerous (see He and Jiang 1980:3).

### 6.1 Historical development of the lexicon

The Ancient Chinese lexicon has changed considerably since the Pre-Archaic period. From the vocabulary of everyday life (lexemes for food, clothing, housing), we find only fifteen words in the oracle bone inscriptions (fourteenth–eleventh centuries BC), seventy-one in the bronze inscriptions (tenth–sixth centuries BC), and 297 in *Shuo wen jie zi* (second century AD). This naturally does not mean that there were only fifteen words denoting these activities in the Pre-Archaic language, and so forth; many other words must have been used which have disappeared leaving no trace (on the varying richness of Chinese vocabulary in different periods, see He and Jiang 1980:9).

According to He and Jiang (1980:136–137), Classical Chinese has an identifiable basic vocabulary of about 2,000 full words, of which 1,100 occur quite commonly. From four major works of the Late Archaic period (*Confucian Analects*, *Mencius*, *Da Xue*, *Zhong Yong*), He and Jiang have isolated 4,466 distinct words, estimating that about half of these are semantically empty (i.e. are proper personal names or place names). There is no implication that the vocabulary of Classical Chinese is impoverished compared to that of Contemporary Chinese – simply different. For example, there is only a single verb meaning “to wash” in Contemporary Chinese (*xi*), whereas there are five in Classical Chinese: *mu* “to wash (the hair)”; *yu* “to wash (the body)”; *hui* “to wash (the face)”; *zao* “to wash (the hands)”; *xi* “to wash (the feet).”

During the long history of Ancient Chinese, several different processes have led to changes in the lexicon. The major processes of internal development include (i) compounding, a highly productive process beginning in the Han period; (ii) semantic extension (e.g., *zu* “foot soldier” > *zu* “all sorts of soldiers”); and (iii) semantic narrowing (e.g., *zi* “child” (boy or girl) > *zi* “son”). In addition, the Ancient Chinese lexicon was enlarged by borrowing words from other languages.

## 6.2 Inherited elements and loanwords

There has been a strong tendency in the past to view the Ancient Chinese lexicon as a monolithic linguistic entity, resistant to influences from all surrounding foreign languages. This is certainly a fallacy. Without going as far as Norman (1988:17) who states, “the fact that only a relatively few Chinese words have been shown to be Sino-Tibetan may indicate that a considerable proportion of the Chinese lexicon is of foreign origin,” we can doubtless rightly assert that the Ancient Chinese lexicon contains numerous loanwords. Nevertheless, the identification of such words and their sources is often uncertain. Below we mention a few noncontroversial examples of loanwords.

There are two common words for “dog” in Ancient Chinese: *quan*, which is probably the native Chinese word, and *gou*, which appears at the end of the Warring States period. *Gou* is a loanword from a language ancestral to the Modern Miao-Yao languages. The word *hu* for “tiger” might have been borrowed from an Austronesian language in prehistoric times (see Norman 1988:17–20). Other words of non-Chinese origin are *xiang* “elephant” (borrowed from a Tai language?); *putao* “grape” (from Old-Iranian?); *moli* “jasmin” (from Sanskrit); *shamen* “Buddhist monk” (from Sanskrit); *luotuo* “camel” (possibly from an Altaic language).

### Bibliography

#### *Principal sources*

- Baxter, W. 1992. *A Handbook of Old Chinese Phonology*. Berlin: Mouton de Gruyter.  
 Baxter, W. and L. Sagart. 1998. “Word formation in Old Chinese.” In J. Packard (ed.), *New Approaches to Chinese Word Formation: Morphology, Phonology and the Lexicon in Modern and Ancient Chinese*, pp. 35–76. Berlin: Mouton de Gruyter.  
 Harbsmeier, C. 1981. *Aspects of Classical Chinese Syntax*. London: Curzon Press.  
 He, J. and S. Jiang. 1980. *Gu hanyu cihui jianghua* (*Talks on the lexicon of Ancient Chinese*). Beijing: Beijing chubanshe.  
 Norman, J. 1988. *Chinese*. Cambridge: Cambridge University Press.  
 Peyraube, A. 1996. “Recent issues in Chinese historical syntax.” In J. Huang and A. Li (eds.), *New Horizons in Chinese Linguistics*, pp. 161–213. Dordrecht: Kluwer Academic Press.  
 Pulleyblank, E. 1991. *Lexicon of Reconstructed Pronunciation in Early Middle Chinese, Late Middle Chinese and Early Mandarin*. Vancouver: University of British Columbia Press.  
 ———. 1995. *Outline of Classical Chinese Grammar*. Vancouver: University of British Columbia Press.

#### *Additional references*

- Baxter, W. 1980. “Some proposals on Old Chinese phonology.” In F. van Coetsem and L. Waugh (eds.), *Contributions to Historical Linguistics: Issues and Materials*, pp. 1–33. Leiden: Brill.  
 Bodman, N. 1980. “Proto-Chinese and Sino-Tibetan: data towards establishing the nature of the relationship.” In F. van Coetsem and L. Waugh (eds.), *Contributions to Historical Linguistics: Issues and Materials*, pp. 34–199. Leiden: Brill.  
 Chou, F. 1963. “Stages in the development of the Chinese language.” In F. Chou (ed.), *Zhongguo yuwen luncong*, pp. 432–438. Taipei: Zhengzhong shuju.  
 Comrie, B. 1989. *Language Universals and Linguistic Typology* (2nd edition). Oxford: Basil Blackwell.  
 Downer, G. 1959. “Derivation by tone-change in Classical Chinese.” *Bulletin of the School of Oriental and African Studies* 22:258–290.  
 Guan, X. 1981. *Xi-Zhou jinwen yufa yanjiu* (*Research on the Grammar of the Bronzes of the Western Zhou*). Beijing: Shangwu yinshuguan.  
 Haudricourt, G. 1954. “Comment reconstruire le chinois archaïque?” *Word* 10 (2/3):351–364.

- Jaxontov, S. 1983–1984 [1960]. “Shanggu hanyu zhong de fufuyin (Consonant clusters in Old Chinese).” *Guowai yuyanxue*, pp. 21–25.
- Karlgren, B. 1915–1926. *Études sur la phonologie chinoise*. Leiden: Brill.
- . 1957. “Grammatica Serica Recensa.” *Bulletin of the Museum of Far Eastern Antiquities* 29:1–332.
- Li, C. and S. Thompson. 1974. “An explanation of word order change: SVO > SOV.” *Foundations of Language* 12:201–214.
- Li, F. 1971. “Shanggu yin yanjiu (Research on Old Chinese pronunciation).” *Tsing Hua Journal of Chinese Studies* 9:1–61.
- Liu, J. 1958. *Hanyu wenyan yufa (A grammar of Classical Chinese)*. Beijing: Zhonghua shuju.
- Liu, J. and A. Peyraube. 1994. “History of some coordinative conjunctions in Chinese.” *Journal of Chinese Linguistics* 22 (2):179–201.
- Ma, C. 1983. *Gudai hanyu yufa (A grammar of Ancient Chinese)*. Qinan: Shandong jiaoyu chubanshe.
- Mei, T. 1970. “Tones and prosody in Middle Chinese and the origin of the rising tone.” *Harvard Journal of Asiatic Studies* 30:86–110.
- . 1978. “Xiandai hanyu xuanzewen jufa delaiyuan (The origin of the disjunction question in Modern Chinese).” *The Bulletin of History and Philology* 49:15–36.
- . 1980. “Sisheng bieyi zhong de shijiancengci (Chronological strata in derivation by tone-change).” *Zhongguo yuwen* 6:427–443.
- . 1989. “The causative and denominative functions of the \*s- prefix in Old Chinese.” In *Proceedings of the 2nd International Conference on Sinology*, pp. 33–51. Taipei: Academia Sinica.
- . 1990. “Tang-Song chuzhishi de laiyan (Origin of the disposal form in the Tang-Song period).” *Zhongguo yuwen* 3:191–216.
- . 1991. “Cong Han dai de ‘dong.sha’ ‘dong.si’ lai kan dongbu jieyou de fazhan (The development of the resultative construction since the Han structures ‘dong.sha’ and ‘dong.si’).” *Yuyanxue luncong* 16:112–136.
- . 1994. “Tang dai Song dai gongtongyu de yufa he xiandai fanyan de yufa (The grammar of Tang-Song koine and the grammar of Modern Chinese dialects).” *Chinese Languages and Linguistics* 2:61–97.
- . 1997. “Hanyu qi ge leixing tezheng de laiyan (Typological changes in the history of the Chinese language).” *Chinese Languages and Linguistics* 4:81–104.
- Peyraube, A. 1987. “The double-object construction in Lunyu and Mengzi.” In *Wang Li Memorial Volumes*, pp. 331–358. Hong Kong: Joint Publishing.
- . 1988. *Syntaxe diachronique du chinois – évolution des constructions datives du 14<sup>ème</sup> siècle av. J.-C. au 18<sup>ème</sup> siècle*. Paris: Collège de France.
- . 1989a. “Zaoqi ba ziju de jige wenti (Some problems on the earliest forms of ‘ba’).” *Yuwen yanjiu* 1:1–9.
- . 1989b. “History of the passive constructions in Chinese until the 10th century.” *Journal of Chinese Linguistics* 17:335–372.
- . 1991. “Some remarks on the history of Chinese classifiers.” *Santa Barbara Papers in Linguistics* 3:106–126.
- . 1994. “On the history of Chinese locative prepositions.” *Chinese Languages and Linguistics* 2:361–387.
- . 1997a. “On word order and word order change in Pre-Archaic Chinese.” *Chinese Languages and Linguistics* 4:105–124.
- . 1997b. “On word order in Archaic Chinese.” *Cahiers de Linguistique Asie Orientale* 26 (1):3–20.
- . 1999. “On the modal auxiliaries of possibility in Classical Chinese.” In H. S. Wang, F. Tsao, and C. Lien (eds.), *Selected Papers from the Fifth International Conference on Chinese Linguistics*, pp. 27–52. Taipei: Crane Publishing.
- Peyraube, A. and T. Wiebusch. 1993. “Le rôle des classificateurs nominaux en chinois et leur évolution historique: un cas de changement cyclique.” *Faits de Langue* 2:51–61.
- . 1995. “Problems relating to the history of different copulas in Ancient Chinese.” In M. Chen and O. Tzeng (eds.), *Linguistic Essays in Honor of William S-Y Wang*, pp. 383–404. Taipei: Pyramid Press.

- Pulleyblank, E. 1962. "The consonantal system of Old Chinese." *Asia Major* 9:58–144; 206–265.
- . 1970–1971. "Late Middle Chinese." *Asia Major* 15:197–239; 16:121–168.
- . 1978. "The nature of the Middle Chinese tones and their development." *Journal of Chinese Linguistics* 6:173–203.
- . 1984. *Middle Chinese: A Study in Historical Phonology*. Vancouver: University of British Columbia Press.
- . 1993. "Old Chinese phonology: a review article." *Journal of Chinese Linguistics* 21 (2):338–380.
- . 1995. "The historical and prehistorical relationships of Chinese." In W. Wang (ed.), *The Ancestry of the Chinese Language*, pp. 145–194. The Journal of Chinese Linguistics Monograph Series, 8. Berkeley: University of California Press.
- Qiu, X. 1978. "Hanzi xingcheng wenti de chubu tansuo (Explorations on the formation of Chinese script)." *Zhongguo yuwen* 3:162–171.
- Sagart, L. 1986. "On the departing tone." *Journal of Chinese Linguistics* 14 (1):90–112.
- . 1993a. "New views on Old Chinese phonology." *Diachronica* 10 (2):237–60.
- . 1993b. "L'infixe -r- en chinois archaïque." *Bulletin de la Société Linguistique de Paris* 88:261–293.
- . 1994. "Proto-Austronesian and Old Chinese evidence for Sino-Austronesian." *Oceanic Linguistics* 33(2):271–307.
- Starostin, S. 1989. "Nostratic and Sino-Caucasian." In V. Shevoroshkin (ed.), *Explorations in Macrofamilies*, pp. 42–66. Bochum: Brockmeyer.
- Sun, C. 1991. "The adposition *yi* and word order in Classical Chinese." *Journal of Chinese Linguistics* 19 (2):202–219.
- Ting, P. 1996. "Tonal evolution and tonal reconstruction in Chinese." In J. Huang and A. Li (eds.), *New Horizons in Chinese Linguistics*, pp. 141–159. Dordrecht: Kluwer Academic Press.
- Wang, L. 1958. *Hanyu shigao (History of Chinese)*. Beijing: Kexue chubanshe.
- . 1979. *Gudai hanyu changshi (Elements of Ancient Chinese)*. Beijing: Renmin jiaoyu chubanshe.
- Zhu, X. 1980. *Gudai hanyu (Ancient Chinese)*. Tianjin: Renmin chubanshe.

# Old Tamil

SANFORD B. STEEVER

## 1. HISTORICAL AND CULTURAL CONTEXTS

Old Tamil stands alongside Sanskrit as one of India's two classical languages. First attested about 254 BC, Old Tamil is the oldest recorded member of the Dravidian languages, a family which today encompasses twenty-four distinct languages. Old Tamil belongs to the southern branch of this family, which includes Malayalam, Irula, Kota, Toda, Kannada, Badaga, Kodagu, and Tulu, as well as Modern Tamil (see Steever 1987).

The era of Old Tamil extends until roughly the seventh century AD, a period of transition to Medieval Tamil. Medieval Tamil differs from Old Tamil in several respects: Old Tamil has two simple tenses – past and non-past – while Medieval Tamil has three: past, present, and future. Old Tamil has relatively few Indo-Aryan lexical borrowings, while Medieval Tamil admits many. During the fourteenth century AD, Medieval Tamil develops into Modern Tamil, a language spoken by nearly 50 million people today. All three periods of Tamil possess a rich literature.

Old Tamil was spoken throughout southern India, in what are now the states of Kerala and Tamil Nadu, as well as in northern Śri Lanka. It is the immediate predecessor of not only Medieval Tamil, but also Malayalam. The western dialects of Late Old Tamil or Early Medieval Tamil, geographically separated from the others by the Western Ghats, developed into Malayalam. Malayalam lost rules of subject–verb agreement so that finite verbs in the modern language lack personal endings. Malayalam also acquired so many Sanskritic loans that aspirated stops now contrast with nonaspirated counterparts. The Old Tamil dialects of Tamil Nadu and northern Śri Lanka developed into Medieval, then Modern Tamil. Late Old Tamil or Medieval Tamil is also likely the predecessor of Irula, a nonliterary language spoken on the slopes of the Nilgiri Mountains in western Tamil Nadu.

Śri Lankan Tamil is more conservative than Continental Tamil, preserving the three-way deictic distinction between *proximal* (*ivan* “this man”), *medial* (*uvan* “the man in between”) and *distal* (*avan* “that man”) of Old and Medieval Tamil. Modern Continental Tamil has reduced the contrast to two by eliminating the medial degree. Śri Lankan Tamil has a synthetic present perfect tense which appears to preserve an Old Tamil present perfect (Steever 1993). Nevertheless, the modern dialects of Śri Lankan and Continental Tamil retain a degree of mutual intelligibility.

Old Tamil exists in three varieties, distinguished by source. *Epigraphic Tamil* is known from rock edicts, cave carvings, and similar inscriptions written in several varieties of Asokan Brahmi script; the earliest of these date to 254 BC. *Mixed Tamil*, also recorded in lithic inscriptions, consists of a mixture of Old Tamil and Sanskrit, which prefigures the medieval



code-mixing style called *maṇippiravāḷam* (lit. “gems and pearls”). Although neither of these varieties is extensively attested, enough survives to date them as contemporaneous with the third type, known today as *caṅka.t tamiz* “*Tamil of the Academy*.” It is richly attested in a large literary corpus and is the variety treated herein. These texts have come down to us through rote memory and on palm-leaf manuscripts (*ōlai*), copied and recopied over the centuries.

Lehmann (1994) divides Old Tamil into three stages – Early, Middle, and Late. Early Old Tamil (250 BC to AD 100) is represented by the grammar *Tolkāppiyam*, and probably by some poems from the anthology *Puraṇānūru* (Four hundred poems on heroism). Middle Old Tamil is the language of bardic poems on themes of love and war (AD 100 to 400) represented in the two collections, *Ettuttokai* (The eight anthologies) and *Pattupāṭṭu* (Ten long songs). Late Old Tamil (AD 400 to 700) is preserved in the twin epics, *Cilappatikāram* and *Maṇimēkalai*, didactic and religious texts, and certain other poems ascribed to that stage. Middle Old Tamil will be the focus of this chapter.

The majority of Old Tamil texts, along with their medieval commentaries, lay forgotten for centuries, resurfacing only during the latter half of the nineteenth century. The medieval period witnessed religious struggles among Hindus, Buddhists, and Jains. When the Hindus ultimately prevailed, they anathematized as irreligious all secular and heterodox texts, destroying them outright or withholding them from copyists. As a result, only a single Buddhist text, the epic *Maṇimēkalai*, survives from the late classical period. Jain texts fared much better because of the ritual of *sastradaṇam* which enjoins rich patrons to commission new copies of old texts to present to scholars at such auspicious occasions as weddings. The *Caṅkam* texts are largely secular, containing poems of love and war; many are informed by a Jain sensibility. Hindu devotional texts such as *Paripāṭal* appear during the Late Old Tamil period. In any event, the classical texts were never known to a wide audience before the modern era. Many were composed for a specific patron and transmitted from teacher to student. When finally committed to writing, the copies were jealously guarded. Frail palm-leaf manuscripts suffered from the extremes of the Indian climate: some crumbled when the leaves were untied, others were thrown into rivers following the death of their owner, and others still probably ended as kindling for cooking fires.

It is only with such Indian scholars as U. V. Caminata Aiyar (1855–1937) that the slow, laborious collecting and editing of these texts began. This paralleled the rediscovery, cataloging, and decipherment of Old Tamil inscriptions by British surveyors and scholars. There is a real sense in which Old Tamil is still being discovered, even among Tamilians. We lack, for example, critical editions of texts (in the Western sense), although we do have editions that may be considered authoritative. Tamil has its own linguistic tradition, anchored in the ancient grammar *Tolkāppiyam* (On ancient composition); but even this text is not fully understood. The linguistic analysis of Old Tamil is still in its infancy, so much so that scholars can still debate the number of cases or tense forms in the language.

## 2. WRITING SYSTEM

Old Tamil was earliest recorded utilizing three different writing systems (see Lehmann 1994). All of these are syllabic scripts that developed from the southern branch of the Ashokan Brāhmī writing system (see Ch. 26, §2). By convention, Old Tamil texts are now transcribed in the modern form of *Tamiz Ezuttu*.

Table 42.1 The Tamil syllabary

VOWEL	ḱ	k	ṅ	c	ñ	t	ṇ	i	n
a	அ	க	ங	ச	ஞ	ட	ண	த	ந
ā	ஆ	கா	ஙா	சா	ஞா	டா	ணா	தா	நா
i	இ	கி	ஙி	சி	ஞி	டி	ணி	தி	நி
ī	ஈ	கீ	ஙீ	சீ	ஞீ	டீ	ணீ	தீ	நீ
u	உ	கு	ஙு	சு	ஞு	டு	ணு	து	நு
ū	ஊ	கூ	ஙூ	சூ	ஞூ	டூ	ணூ	தூ	நூ
e	எ	கெ	ஙெ	செ	ஞெ	டெ	ணெ	தெ	நெ
ē	ஏ	கே	ஙே	சே	ஞே	டே	ணே	தே	நே
ai	ஐ	கை	ஙை	சை	ஞை	டை	ணை	தை	நை
o	ஒ	கொ	ஙொ	சொ	ஞொ	டொ	ணொ	தொ	நொ
ō	ஓ	கோ	ஙோ	சோ	ஞோ	டோ	ணோ	தோ	நோ
au	ஔ	கென	ஙெள	சௌ	ஞௌ	ளட	ணௌ	தௌ	நௌ
None	ஃ	க்	ங்	ச்	ஞ்	ட்	ண்	த்	ந்
VOWEL	p	m	y	r	l	v	z	ḷ	ḥ
a	ப	ம	ய	ர	ல	வ	ழ	ள	ற
ā	பா	மா	யா	ரா	லா	வா	ழா	ளா	றா
i	பி	மி	யி	ரி	லி	வி	ழி	ளி	றி
ī	பீ	மீ	யீ	ரீ	லீ	வீ	ழீ	ளீ	றீ
u	பு	மு	யு	ரு	லு	வு	ழு	ளு	று
ū	பூ	மூ	யூ	ரூ	லூ	வூ	ழூ	ளூ	றூ
e	பெ	மெ	யெ	ரெ	லெ	வெ	ழெ	ளெ	றெ
ē	பே	மே	யே	ரே	லே	வே	ழே	ளே	றே
ai	பை	மை	யை	ரை	லை	வை	ழை	ளை	றை
o	பொ	மொ	யொ	ரொ	லொ	வொ	ழொ	ளொ	றொ
ō	போ	மோ	யோ	ரோ	லோ	வோ	ழோ	ளோ	றோ
au	பௌ	மௌ	யௌ	ரௌ	லௌ	வௌ	ழௌ	NA	றௌ
None	ப்	ம்	ய்	ர்	ல்	வ்	ழ்	ள்	ற்

While preserving the fundamental principles of the syllabic systems descended from Brāhmī, the Tamil orthographic system, *Tamiḻ Ezuttu* “Tamil letter,” has continued to evolve. In *Tamiḻ Ezuttu*, as in related syllabaries, each graph represents a *vowel* or a sequence of *consonant + vowel*. Vowels (*uyir* “breath, soul”) are represented by two main allographs: (i) one for initial position; and (ii) one, or more, used in combination with a consonant



graph. Consonant (*mey* “body”) graphs have a so-called basic form with the syllabic value *consonant* + *a*. This basic form is graphically modified to express other vowel values by adding diacritics above it or below it, or to its left or right. To represent consonant clusters, the basic graph is modified by adding a *pulli*, a small circle, above the basic sign of all but the last graph of the cluster. The *aytam*, symbol for *k*, lacks an inherent vowel component and never occurs with a vowel diacritic.

For further discussion and illustration of the Brahmi and Tamil writing systems, see Daniels and Bright 1996.

### 3. PHONOLOGY

#### 3.1 Consonants

The seventeen consonants of Tamil are as follows:

##### (1) Tamil consonant phones

	Labial	Dental	Alveolar	Retroflex	Palatal	Velar	Glottal
Stop	p	t	<u>r</u>	ṭ	c	g	
Nasal	m	n	<u>n</u>	ṇ	ñ	(ŋ)	
Lateral		l		ḷ			
Tap		r					
Approximant					ɻ		
Glide	v				y		(h)

All sounds are phonemically distinct except those placed within parentheses, which are allophones having a distinct graphemic representation. With the notable exception of /*yaṇṇanam*/ “what manner,” the velar nasal [ŋ], transcribed as *ṇ*, appears to occur as an allophone of other nasals, occurring only before the velar stop /*k*/. The fricative [h], called *aytam* and transcribed as *ḳ*, occurs only between a short vowel and a stop (e.g., /*ahtu*/ “it, that”). As such, it may be regarded as an allophone of /*v*/ since /*v*/ is the only consonant that does not occur in this context.

#### 3.2 Vowels

Old Tamil has ten vowels, five short and five long:

- (2) i, ī                      u, ū  
       e, ē                      o, ō  
                                  a, ā

Diphthongs /*ai*/ and /*au*/ also occur.

For metrical purposes, the long vowels and the diphthong /*ai*/ may be lengthened (through, in effect, the addition of a short vowel). This lengthening may then be repeated; in other words, an already elongated vowel may itself be lengthened. Consider, for example, /*ciṛār*/ “small ones” (*Parī* 3.6) becoming /*ciṛāār*/ (*Aka* 107.17) and further /*ciṛāāār*/ (*Pura* 291.2).

#### 3.3 Morphophonemic variation

When morphemes combine and compound words are formed, several kinds of morphophonemic changes (*sandhi*) may occur. Among the more common are the following:

1. Loss of a final segment: *pāṭṭu* “song” + *-āl* instrumental case > *pāṭṭ-āl* “by song”; *maram* “wood” + *vīṭu* “house” > *mara-vīṭu* “wooden house”
2. Assimilation: *kaṇ-* “eye” + *-ku* dative case > *kaṇ-ku* “to the eye”
3. Consonant insertion: *kal* + *āl* > *kal.l-āl* “with stone”; *pacu* “fresh” + *īṛa* “shrimp” > *pac.c-īṛa* “fresh shrimp”
4. Glide insertion: *katti* + *āl* > *katti.y-āl* “with a knife”

This chapter adopts Lehmann’s (1994) convention of placing a period before a segment that is automatically inserted by phonological rule in order to clarify morphemic identity. Such processes are obligatory with a bound morpheme, less frequent between members of a compound and least frequent elsewhere.

### 3.4 Phonotaxis

All vowels and the diphthong /au/ may occur in word-initial position. All vowels and diphthongs occur after all consonants except /n/ and *ḳ*. There is only a single occurrence of /a/ after /ŋ/: /yaŋŋa<sub>u</sub>am/ “in which way” (*Aka* 27.12). All vowels and /au/ may appear in word-final position. Only the nine consonants /p/, /t/, /c/, /k/, /m/, /n/, /ñ/, /y/, and /v/ appear in word-initial position. Word-finally, ten consonants, all of them nonobstruents, are permitted: /m/, /n/, /ṇ/, /ṇ/, /l/, /ḷ/, /y/, /w/, /r/, and /z/. Consonant clusters are limited in scope.

### 3.5 Prosody

Old Tamil is a quantitative language: quantitative units called *acai* “morae” fall at regular intervals. These units are combined into feet, and the feet into meters. Rajam 1992 outlines the prosodic system, particularly as it involves poetic composition.

## 4. MORPHOLOGY

### 4.1 Word formation and word classes

Old Tamil morphology is predominantly agglutinating, with a one-to-one correspondence between morpheme and morph. Despite what appear to be exceptions, it is exclusively suffixal. There also occur some instances of fusion.

Old Tamil has two major, formally distinct parts of speech – noun and verb. Most lexical stems belong to one of these two classes; some stems have a double categorial status: for example, *col* can be the verb stem meaning “say,” or the noun stem meaning “word.” Beyond this, consensus as to the number and identity of parts of speech breaks down. A small number of words fail to exhibit all the properties nouns and verbs typically exhibit: some scholars assign them to two minor classes, adjectives and adverbs; others treat them as defective nouns and verbs.

Distinct from the parts of speech is a set of clitic particles which combine with their host to form a phonological word, but which may syntactically combine with an entire clause. Clitics are herein identified by the boundary marker =, and include quantifiers, discourse particles, and emphatic markers.

## 4.2 Nominal morphology

Nominals in Old Tamil include common nouns, numerals, proper names, pronouns, and certain other forms. Nominals are primarily inflected for case and number, and secondarily for gender and person. Nominal stems may be simple or complex. Complex stems include all derivatives. The complex noun *āṇ-mai* “manliness, strength” (*Pati* 70.20; a complete list of abbreviations and texts cited is to be found at the end of the chapter) consists of the noun *āḷ* “man” and the abstract suffix *-mai*; while the complex noun *kaḷv-i* “female thief” consists of the stem *kaḷ-* “theft” and the feminine suffix *-i*.

### 4.2.1 Gender

In Old Tamil gender is largely natural. There are two basic genders, *uyartiṇai* “animate” (lit. “high-class”) and *akriṇai* “inanimate” (lit. “non-class”), which determine, *inter alia*, the choice of plural marker and pronouns.

### 4.2.2 Number

In Old Tamil, singular number is unmarked. The plural has three basic markers: *-kaḷ*, *-ar*, *-ir*. The first occurs with inanimate nouns, the second and third, with animate. Examples include the following: (i) *kaṇ* “eye” ~ *kaṇ-kaḷ* “eyes” (*Kali* 39.42); *iyam* “musical instrument” ~ *iyaṇ-kaḷ* “musical instruments” (*Malai* 277); *vāzi* “path” ~ *vāzi-kaḷ* “paths” (*Aka* 8.1); (ii) *arivai* “woman” ~ *arivai.y-ar* “women” (*Pati* 68.19); *koṭicci* “young girl” ~ *koṭicci.y-ar* “young girls” (*Kali* 40.11); *kēḷ* “relative” ~ *kēḷ-ir* “relatives” (*Kali* 61.3); (iii) *peṇṭu* “woman” ~ *peṇṭ-ir* “women” (*Aiṅk* 271.3). Certain other plural suffixes, such as *-mār*, are also attested: for example, *tōzi* “girlfriend” ~ *tōzi-mār* “girlfriends” (*Aka* 15.9). Even when a finite verb bears a plural suffix, the subject need not appear marked as a plural: thus, *paṭāa em kaṇ* “my<sub>2</sub> eyes<sub>3</sub> do.not.sleep<sub>1</sub>” (*Aka* 218.9).

### 4.2.3 Stem-forms

Old Tamil nouns may have an oblique stem that differs from the nominative. There are two basic kinds of oblique stems. For neuter nouns ending in *-am*, the oblique replaces the final *-m* with *-ttu*: for example, the nominative form *paṇam* “fruit” has the oblique stem *paṇa-ttu-* (*Aka* 292.14). Nouns that end in *-tu* or *-ru* double the consonant in the oblique: thus, nominative *nāṭu* “country” has the oblique stem *nāṭ.tu-* (*Aiṅk* 203.2). The oblique is the form to which non-nominative case markers and postpositions are added (3A, B); the form that appears when a case ending is elided (3C, D); and the form that serves as an appositive attribute (3E, F). It is, in short, the combining form of the noun.

- (3) A. *maṇa-tt-ōṭu*  
mind-OBL.-SOC.  
“With the mind” (*Kali* 47.17)
- B. *kaḷiṇ-ṭ-ōṭu*  
elephant-OBL.-SOC.  
“With the elephant” (*Pati* 66.7)
- C. *aṭuka-ttu aruvi vāza*  
cliff-OBL. waterfall-NOM. fall-INF.  
“As the waterfall descends from the cliff” (*Kali* 44.2)

- D. nāṭṭu.c      cell-al  
country-OBL. go-NEG.-IMPV.  
“Don’t return to [your] country” (*Aiṅk* 233.4)
- E. vēṣa-ttu.k      kōṭu  
elephant-OBL. tusk  
“The elephant tusk” (*Kuru* 100.4)
- F. nāṭṭu.k      kuṇram  
country-OBL. hill-NOM.  
“The hills of the country” (*Kuru* 249.3–4)

With the oblique case forms, compare the euphonic suffixes *-in-* and *-an-*, which may appear between a noun stem and a case-marker. Traditional accounts suggest these forms are inserted for metrical purposes. Examples include the accusative form *kulai.y-in-ai* “a bunch” (*Kali* 45.3) and the dative *naṭp-ir-ku* “for friendship” (*Pura* 236.6)

#### 4.2.4 Case

The *Tolkāppiyam* identifies eight cases in all: nominative, accusative, dative, instrumental, equative, genitive, locative, and vocative. Case-markers are added to the singular or plural stem of the noun; however, only rarely are plural inanimate nouns marked for case. Postpositions extend the case system.

##### 4.2.4.1 Nominative

The nominative is the unmarked case, and has several functions. It serves as the subject of a clause (4) and as predicate nominative (5), among other functions.

- (4) A. yān      vantanen  
I-NOM. come-PST-1ST PER. SG.  
“I have come” (*Narr* 267.8)
- B. yān=um      nī.y=um      e.v-vaṣi      aṛitum  
I-NOM.=and you-NOM.=and what.path meet-PST-1ST PER. PL.  
“Where did you and I meet?” (*Kuru* 40.3)
- (5) A. ivar      pāri makaḷ-ir  
these.ones-NOM. Pari daughter-PL.-NOM.  
“These are the daughters of Pari” (*Pura* 202.14–5)
- B. yāt=um      ūr=ē,      yāvar=uñ  
which one-NOM.=and town-NOM.=and whoever-NOM.=and  
kēḷir  
relation-PL.-NOM.  
“Any [town] is [our] town, all people are [our] kinfolk” (*Pura* 192.1)

Owing to the common elision of case-markers, many nouns, particularly inanimates, appear in the nominative even though the semantics of the clause would require some other case-marker. In (5A), for example, the proper name *Pari* “Pari” functions as a genitive, but appears in the nominative.

##### 4.2.4.2 Accusative

The accusative typically marks an animate direct (6A, B) or indirect (6C) object. The Old Tamil corpus, however, contains some examples of inanimate objects marked as accusative (6D).

- (6) A. en tōḻi.y-ai nōkki  
my friend-ACC. look.at-CF  
“[You are] looking at my friend” (*Kali* 50.8)
- B. orūpa nin.n-ai  
shun-NPST-3RD PER. PL. you-ACC.  
“They shun you” (*Pati* 34.1)
- C. nin.n-ai ampuli kātṭal initu  
you-ACC. moon-NOM. show-VN sweet-NOM.  
“It is nice to show the moon to you” (*Kali* 80.18–19)
- D. upp-ai māṛi vennel tariiya  
salt-ACC. trade-CF white.paddy-NOM. bring-INF.  
“In order to obtain white paddy by trading salt” (*Kuru* 269.5)

#### 4.2.4.3 Dative

The dative typically marks indirect object, (7A, B), direction (7C), or causality (7D, E):

- (7) A. annai-kku moḻi.y-um vēlaṇ  
mother-DAT. speak-NPST-3RD PER. SG. priest-NOM.  
“The priest speaks to the mother” (*Aiṅk* 249.1–2)
- B. nin.a-kku onru kuṛuvam kēl iṇi  
you-DAT. one.thing tell-NPST-1ST PER. PL. listen-IMPV. now  
“Listen now to what I (lit. we) have to say to you” (*Kali* 55.5)
- C. ūru-kku.p pōvōy  
village-DAT go-NPST-2ND PER.SG.  
“You will go to your village” (*Narr* 200.7)
- D. poruṭ-ku iratti  
wealth-DAT. depart-CF  
“Departing for riches” (*Kali* 10.12)
- E. vaṇ peyar-ku aviṇṭa painkoṭai mullai  
heavy rain-DAT. unfold-PST-ADN. fresh.vine jasmine  
“The fresh jasmine on the vine which blossomed because of the heavy rain”  
(*Aka* 124.11)

Old Tamil also has some structures which, in the light of modern Dravidian structures, could be interpreted as dative-subject constructions:

- (8) A. nin.a-kk=ō ariyunal neñc=ē  
you-DAT.=INTERR. know-VN-3RD SG. FEM. heart=VOC.  
“Is she someone known to you, O my heart?” (*Narr* 44.5) meaning:  
“Do you know her, O my heart?”
- B. emakku il  
we-DAT. not.be  
“There is nothing for us” (*Pati* 39.2) meaning: “We have nothing”

#### 4.2.4.4 Instrumental

The instrumental case may be signaled by the morphs *āṇ* and *-āl*. The suffix expresses the relations of instrument, association and location. Due to sandhi, it is sometimes difficult to identify which morph is used (as in 9C).

- (9) A. nī                      munn-att-ān                      kātt-in-ai  
 you-SG.-NOM.    sign-OBL.-INSTR.    show-PST-2ND PER. SG.  
 “You showed with a sign” (*Kali* 61.7)
- B. āti.y-ai              talai.y-in-āl                      toṭttu  
 foot-ACC.    head-OBL.-INSTR.    touch-CF  
 “Touching [his] foot with [my] head” (*Kali* 1108.55–56)
- C. nin                      kaṇṇ-ār                      kaṇpēn                      yān  
 you-OBL.    eye-INSTR.    see-NPST-1ST PER. SG.    I-NOM.  
 “I see with your eyes” (*Kali* 39.43)

#### 4.2.4.5 Sociative

The sociative is marked by the suffixes *-oṭu* and *-ōṭu*. It signals accompaniment or instrumentality:

- (10) A. kaḷirru.t                      toṭuti.y-ōṭu    vantu  
 elephant-OBL. driver-SOC. come-CF  
 “[He] came with a mahout” (*Pati* 62.1–5)
- B. ivaḷ-ōṭu    vāṇiya  
 she-SOC. live-OPT.  
 “May you live/prosper with her” (*Pati* 21.37–8)
- C. vitt-ōṭu    cenra                      vatti    parpala    mīn-ōṭu                      peyarum  
 seed-SOC. go-PST-ADN. basket many    many fish-SOC. return-NPST-3RD PER.SG.  
 “The basket which left with seeds returns with many kinds of fish” (*Narr* 210.3–4)

#### 4.2.4.6 Equative

The equative case marks an object of comparison with the suffix *-in*. Its subsidiary nuances include locative, instrumental and causal. This case no longer exists in Modern Tamil, having been replaced by the postposition *viṭa* “than” (e.g. *avaṇ-ai viṭa* “than<sub>2</sub> that.man<sub>1</sub>”):

- (11) A. tokai                      aṇi                      maṭantai.y-in    tōnr-um  
 peacock-NOM. decoration-NOM. girl-EQ.    appear-NPST-3RD PER. NEUT.  
 “The peacock looks like a decorated young girl” (*Aiṅk* 294.1–2)
- B. cirril                      kāl-in                      cītai.y-a  
 small.house foot-EQ. trample-CF  
 “[We] trample the hut with [our] feet” (*Kali* 51.2)
- C. irav-in                      var-al  
 night-EQ. come-NEG.IMPV.  
 “Do not come at night” (*Kali* 49.23)

#### 4.2.4.7 Genitive

The genitive, signaled by *-atu* and *a*, is adnominal: it marks such relations as possession between two noun phrases:

- (12) A. en                      toṭi.y-atu                      kavin  
 I-OBL. friend-FEM.-GEN. beauty  
 “My girlfriend’s beauty” (*Kali* 50.24)
- B. avar-a                      kayam  
 those.people-GEN. pool  
 “Their pools [of water]” (*Pura* 15.9–10)

4.2.4.8 *Locative*

The locative case, one of the most unstable in Dravidian, is marked by a case formant and no less than nineteen postpositions (*Tolkāppiyam*, *collatikāram*, 82). The case-marker *-il*, the inanimate locative case-marker in Modern Tamil, is attested in Old Tamil (13A, B). However, locative postpositions with specific meanings are more commonly encountered: *uṣai* “place” (13C), *vayin* “area” (13D). These typically combine with the oblique stem of the preceding noun. The oblique stem itself, unmodified by any case-marker, often conveys the sense of the locative (13E):

- (13) A. *cilamp-il tuṇcum kavari*  
 hillside-LOC. sleep-NPST-ADN. antelope  
 “The antelope that sleeps on the hillside” (*Pati* 11.21)
- B. *eḷḷ-in-il peyartal*  
 dusk-LOC. leave-VN  
 “Leaving at dusk” (*Aka* 100.4)
- C. *NP[kēḷir N[uṣai.c]NP cenru*  
 relative-PL.-OBL. place go-CF  
 “going to one’s relatives” (*Kali* 61.3)
- D. *kīḷavi NP[nam N[vayin]NP vantaṇru*  
 word-NOM. we-OBL. place come-PST-3RD PER.NEUT.  
 “Word has come to us” (*Kuru* 106.3–4)
- E. *nal nāṭ.ṭu.c celkam*  
 good land-OBL. go-NPST-3RD SG.MASC.  
 “He is going to his beautiful country” (*Aiṅk* 236.4)

4.2.4.9 *Vocative*

The vocative, used in address, is formed in several ways. Nouns that end in *-an* delete the final nasal: for example, *marukan* “son,” *marukā* “O son” (*Pati* 63.16). Certain nouns form the vocative by lengthening the final vowel of the last syllable: for example, *nutal* “forehead,” *nutāl* “O [one with the broad] forehead” (*Kali* 37.12); *annai* “mother,” *annāy* “O mother” (*Aiṅk* 201.1). In other instances, usually with inanimates, the clitic *=ē* is added: thus, *nāṭu* “country,” *nāṭ = ē* “O (my) country” (*Aiṅk* 221.4).

- (14) *annāy vāzi ventannai*  
 mother-VOC. live-OPT. listen-PST-2ND PER.SG.  
 “Bless you, my friend [lit. mother]. You must listen” (*Aiṅk* 203.1)

4.2.4.10 *Absence of case-marking*

Lehmann (1994: 52ff.) observes that case-markers in Old Tamil are often omitted in contexts that one would expect to trigger their presence. For example, the transitive verb *uḷ-* “think” ordinarily requires direct objects in the accusative case. Note that in (15A) the object in the first conjunct appears in the nominative case, the object in the second in the oblique stem. The result of this elision is that Old Tamil contains many phrases that resemble large compounds consisting of nominal stems, as in (15B).

- (15) A. *curram=um em.m=um uḷḷ-al*  
 companions-NOM.=and we-OBL.=and think-NEG.-3RD PER.FEM.  
 “She doesn’t think of us (=me) and our companions” (*Aka* 17.6)

- B. [[kaya vāy] [peru.ñ kai] yāñai]  
 [[great mouth] [big trunk] elephant]  
 “The elephant with a great maul and a large trunk” (*Aka* 118.7–8)

#### 4.2.5 Person

Old Tamil nouns may also mark person by suffixing personal endings that agree with a subject; these are the so-called *appellative nouns* of the literature. Lehmann (1994: 61ff.) shows these forms to be the result of a syntactic, not a morphological, process. Appellative nouns in Old Tamil are generally predicate nominals (16A–E) or vocatives. In Medieval Tamil they may take such non-nominative case forms as the accusative (16F), while they are absent in the modern language

- (16) A. tōl-ēñ  
 shoulder-1ST PER.SG.  
 “I, with [broad] shoulders” (*Aka* 82.18)
- B. peṇṭ-ir-ēm all-ēm  
 woman-PL.-1ST PER.PL. become-NEG.-1ST PER.PL.  
 “We are not women” (*Pura* 246.10)
- C. nall-āy  
 good.one-2ND PER.SG.  
 “You, who are good” (*Kali* 39.30)
- D. eyir-r-āl  
 tooth-3RD SG.FEM.  
 “She, with [shining] teeth” (*Aiñk* 256.3)
- E. nīḷ-ōr  
 shadow-3RD PER.PL.  
 “Those who are in the shadow” (*Pati* 68.20)
- F. aṭi.y-ēñ-ai.k kaṇṭāñ  
 devotee-1ST PER.SG.-ACC. see-PST-3RD SG.MASC.  
 “He saw me, a devotee”

#### 4.2.6 Pronouns

Old Tamil has personal as well as demonstrative and interrogative pronouns.

##### 4.2.6.1 Personal pronouns

The personal pronouns of Old Tamil are shown in (17). The nominative forms are followed by their oblique forms in parentheses. Note that Old Tamil distinguishes between an *inclusive* and an *exclusive* plural. The dramatis personae of Old Tamil poems often use first-person plural inclusive where the first-person singular might be expected; this convention persists in Modern Tamil when a speaker engages in musing or soliloquy.

(17)	Singular	Plural
First	yāñ (eñ(n), eña)	Exclusive: yām (yām(m), yama) Inclusive: nām (nām(m), nama)
Second	nī (niñ(n), niña)	nīr, nīyir (num(m), num)
Third	tāñ (tañ(n), taña)	tām (tam(m), tama)



#### 4.2.6.2 Demonstrative and interrogative pronouns

Old Tamil also has a series of demonstrative and interrogative pronouns. Forms in *i-* are proximal, *u-* medial, *a-* distal, and *e-/ya-* interrogative. The medial series is not fully attested in the old language: the forms \**uktu* “this one in-between” and \**uvai* “these in-between” are not attested in the corpus. The series is lost in the modern continental dialects. Members of the distal series frequently serve as general-purpose third-person pronouns.

(18)			<i>Proximal</i>	<i>Medial</i>	<i>Distal</i>	<i>Interrogative</i>
	<i>Singular</i>	<i>Masculine</i>	ivan̄	uvan̄	avan̄	yāvan̄
		<i>Feminine</i>	ivaḷ	uvaḷ	avaḷ	yāvaḷ
		<i>Neuter</i>	itu/i <u>k</u> tu	utu/NA	atu/a <u>k</u> tu	yātu/yāvatu
	<i>Plural</i>	<i>Animate</i>	ivar	uvar	avar	yār
		<i>Inanimate</i>	ivai	NA	avai	yāvai/yā

### 4.3 Verbal morphology

Verbs in Old Tamil mark such categories as illocutionary force, tense, mood, negative polarity, and subject–verb agreement. Verbs are formally distinguished as finite verbs and nonfinite verbs.

#### 4.3.1 Stem-forms

A majority of Tamil verb-bases may form two related stems, one weak, the other strong. This morphophonemic distinction corresponds to a voice distinction, *affective* voice versus *effective* voice (Paramasivam 1979). Certain stems have distinct variants for the negative conjugation: for example, positive *kaṇ-* “see” versus negative *kāṇ-*. Two important verbs, *āka* “become” and *iru* “be (located),” have the suppletive variants *al-* “not become” and *il-* “not be,” respectively, in the negative conjugation.

#### 4.3.2 Verbal conjugations

Old Tamil has seven conjugation classes, based on the allomorphs of the suffixes they take. Once the voice and the phonological shape of the stem are taken into account, it may be possible to reduce the number of conjugations.

#### 4.3.2 Nonindicative moods

In addition to the indicative (see §4.3.3) Old Tamil finite verbs can occur in the imperative and optative moods. All finite verbs mark subject–verb agreement.

##### 4.3.2.1 Imperative

Imperatives convey an order, request, and so forth, and encode second-person agreement. The simple verb-stem often functions as the singular imperative (19), although there exist exceptions: thus, the verb-stem *taru-* “give to you or me” has the imperative form *tā* “give (to me).”

- (19) A. aṅku iṛa  
       there go-IMPV.  
       “Go there” (*Kali* 63.9)

- B. *kēl*                      *avan*                      *nilaiy=ē*  
 listen-IMPV. he-GEN. condition=PCL.  
 “Listen to his condition” (*Peru* 38)

The negative singular imperative has two forms. One adds the ending *-ati* to the verb stem (20A). The other is a compound verb: the verb appears in its stem-form and is followed by the auxiliary verb *al* “not become” inflected for the imperative (20B):

- (20) A. *varunt-ati*  
 grieve-NEG.-IMPV.  
 “Don’t grieve” (*Kali* 107.30)
- B. *anpin*    *v[VS[aṭ]VS AUX[al]AUX]V*  
 love-EQ cry-STEM    not.become-IMPV.  
 “Do not cry from love” (*Narr* 309.3–4)

The plural imperative consists of a verb-stem and one of several suffixes: *-mati*, *-min*, or *-m*:

- (21) A. *ivar-ai*                      *koṇ-mati*  
 this.man-ACC. take-IMPV.-PL.  
 “Take this man” (*Pura* 201.16)
- B. *avar-ku aṛiya*                      *urai-min*  
 he-DAT. understand-INE. speak-IMPV.-PL.  
 “Speak so that he understands” (*Narr* 376.9)
- C. *yāvar=um varu-ka ēṇōr=um tā-m*  
 who=and come-OPT. others=and bring-IMPV.-PL.  
 “Let everyone [of you] come. Bring others, too” (*Matu* 747)

The negative plural imperative is periphrastic, consisting of the verb-stem and the auxiliary verb *al* “not become” (*aṇ-* by sandhi) inflected for the imperative with *-min*:

- (22) *evvam v[VS[patar]VS AUX[aṇ-min]AUX]V*  
 distress suffer-STEM    not.become-IMPV.-2  
 “Don’t suffer in distress” (*Kali* 9.22)

#### 4.3.2.2 Optative

The optative is marked by several suffixes, the most common of which is *-k(k)a*. Lehmann’s (1994:76) examples show that unlike the imperative, which is restricted to second-person subjects, the optative occurs with all persons.

- (23) A. *peṇṭu*                      *yāṇ*                      *āku-ka*  
 woman-NOM. I-NOM. become-OPT.  
 “May I become a woman”; “Would that I were a woman” (*Aka* 203.18)
- B. *oṇ*                      *kuz-āy*                      *cel-ka*  
 shining earring-2ND PER. SG. go-OPT.  
 “May you, with the shining earrings, go” (*Kali* 37.21)
- C. *yāy*                      *aṛintu*                      *uṇar-ka*  
 my.mother know-CF understand-OPT.  
 “May my mother know and understand [it]” (*Aka* 203.2)

The negative optative is an auxiliary compound verb consisting of the verb-stem of the main verb and *al* “not become” (*aṇ-* by sandhi), inflected for the optative suffix:

- (24) aiyam <sub>v</sub>[<sub>vs</sub>[*koḷḷ*]]<sub>vs</sub> <sub>AUX.</sub>[*aṇ-min*]<sub>AUX.</sub> <sub>v</sub> ar-*aṇival-ir*  
 doubt hold-STEM not.become-OPT. full-knowledge-2ND PER.PL.  
 “May you, who are full of knowledge, have no doubts” (*Pura* 216.5)

Old Tamil has additional suffixes and constructions used to convey the imperative and optative modes. Second-person indicative verbs often serve as imperatives or optatives.

### 4.3.3 Indicative mood

Indicative finite verbs consist of a (i) verb-stem, (ii) tense marker, and (iii) personal ending (though see below). In the negative, they consist of a (i) verb-stem, (ii) negative marker, and (iii) personal ending. In the positive, there are two tenses – past and non-past. Steever (1993) suggests that what have been treated as allomorphs of the past tense are actually markers of a present perfect tense which is cognate with present perfect forms in other Dravidian languages and which survives in Śrī Lankan Tamil. There is, in the negative, a single paradigm corresponding to the past and non-past positive paradigms (see [30] below).

The positive indicative forms exhibit some variation in their morphological composition. The majority consist simply of a verb-stem, tense marker, and personal ending that marks subject–verb agreement (25A). Some insert a euphonic increment between the stem and tense marker, while others place the increment between the tense marker and the personal ending (25B). Such euphonic increments may represent the historical residue of earlier tense suffixes (25C). In still other indicative forms, the tense marker and agreement-marker have fused into a portmanteau morph incapable of segmentation (25D). Lehmann (1994:79) presents these four possibilities with the verb *cey-* “do, make”:

- (25) A. *cey-t-āṇ*  
 do-PST-3RD SG.MASC.  
 “(He) did” (*Kali* 51.16)  
 B. *cey-t-aṇ-ai*  
 do-PST-EUPH-2ND PER.SG.  
 “(You) did” (*Aiṅk* 294.3)  
 C. *cey-ku-v-am*  
 do-EUPH.-NPST-1ST PER.PL.  
 “(We) do” (*Aiṅk* 288.2)  
 D. *cey.y-um*  
 do-NPST+3RD PER.PL.  
 “(They) do” (*Aiṅk* 244.4)

#### 4.3.3.1 Tense markers

The past tense morpheme has the following allomorphs: *-t-*, *-nt-*, *-tt-*, *-i-*, *-iṇ-*, and gemination of the stem-final consonant (marked -CC- in [26]). The non-past morpheme has the allomorphs *-v-*, *-p-*, *-pp-*. They are distributed according to the seven conjugational classes of Old Tamil as follows.

## (26) Old Tamil conjugational classes

Class	Past		Non-past	
I	-t-	cey-t-ēṇ “I did” ( <i>Kali</i> 37.12)	-v-	cey-v-ēṇ “I do” ( <i>Kali</i> 62.12)
II	-nt-	aṛi-nt-ēṇ “I knew” ( <i>Kali</i> 47.3)	-v-	aṛi-v-ēṇ “I know” ( <i>Aiṅk</i> 247.1)
III	-in-/i-	aṇc-in-āṇ “he feared” ( <i>Kali</i> 65.20)	-v-	aṇcu-v-aḷ “she fears” ( <i>Kali</i> 48.22)
IV	-CC-	per-r-aṇ-ar “they got” ( <i>Pura</i> 10.4)	-v-	peru-v-ai “you get” ( <i>Kali</i> 49.25)
V	-t-	kaṇ-ṭ-ai “you saw” ( <i>Kali</i> 64.6)	-p-	kaṇ-p-ēṇ “I see” ( <i>Kali</i> 39.43)
VI	-tt-	urai-tt-al “she spoke” ( <i>Kali</i> 39.21)	-pp-	urai-pp-atu “it speaks” ( <i>Kali</i> 48.19)
VII	-nt-	ira-nt-aṇ-aṇ “he asked” ( <i>Aiṅk</i> 257.2)	-pp-	ira-pp-āṇ “he asks” ( <i>Kali</i> 62.12)

## 4.3.3.2 Personal endings

The personal endings illustrated in (27) may be added to the past stem, the non-past stem, or the negative stem of a verb. They are the most general in the language, and give rise to the personal endings of the modern language.

## (27) Old Tamil personal endings I

	Singular	Plural
First	-ēṇ, -eṇ, -al, -aṇ-	-ām, -am, -ēm, -em
Second	-ai, -āy, -ōy	-ir, -ir
Third		
Masculine	-ān, -an, -oṇ	Epicene -ār, -ar, -or
Feminine	-āl, -aḷ, -ōḷ	
Neuter	-tu, -ttu, -atu	-a

The personal endings of (28) are added directly to the verb-stem; these are portmanteau forms that encode not merely person, number, and gender but also non-past tense. This structure, presented in (29B), is contrasted with the more general structure in (29A). Only the third-person neuter singular form *-um* survives into Modern Tamil; the loss of the portmanteau forms represents a reassertion of the general agglutinative character of Tamil morphology which discourages fusional forms.

## (28) Old Tamil personal endings II

	Singular	Plural
First	-ku/-kku	-tum, -kum, -kam
Second	-ti, -tti	-tir
Third		
Neuter	-um	-um
Epicene		-pa, -mār

## (29) A. ira-pp-āṇ

ask-NPST-3RD SG.MASC.

“He asks” (*Kali* 62.12)

## B. ira-kku

ask-NPST+1ST PER.SG.

“I ask” (*Pati* 61.11)

Corresponding to the past and non-past indicative paradigms is a single negative indicative paradigm. It consists of the verb-stem, negative marker, and personal ending (with a long vowel), as already seen in the negative imperative and optative forms. The negative marker is often realized by a zero morph, although its operation can sometimes be inferred from a change of vowel quantity in the verb-stem: for example, *kaṇpēn* “I see” versus *kāṇāl* “she didn’t/doesn’t see.”

- (30) A. *cel.l-ēn*                      *cel.l-ēn*                      *piṛar*                      *mukam*  
           go-NEG.-1ST PER.SG.    go-NEG.-1ST PER.SG.    other-GEN.    face-NOM.  
           *nōk.k-ēn*  
           look.at-NEG.-1ST PER.SG.  
           “I won’t go. I won’t go. I won’t look at the faces of others” (*Pura* 399:14)
- B. *ivaṇ-ai.p*    *poy-ppa*                      *viṭ-ēem*  
           he-ACC.    tell-lies-INF.    let-NEG.-1ST PER.PL.  
           “We won’t let him tell lies” (*Kali* 89.13)
- C. *anru*    *nam*                      *aṛi.y-āy*  
           then    we-OBL.    know-NEG.-2ND PER.SG.  
           “You did not know us then” (*Aka* 33.18)
- D. *pāṇan*                      *cūṭ-ān*                      *pāṭini*                      *aṇi.y-aḷ*  
           bard-NOM.    wear-NEG.-3RD SG.MASC.    bard’s.wife-NOM.    decorate-NEG.-3RD SG.FEM.  
           “The bard doesn’t wear [the jasmine], his wife doesn’t decorate herself [with it]”  
           (*Pura* 139.1)

#### 4.3.4 Periphrastic constructions

Old Tamil also has several periphrastic forms that simultaneously express tense and negation. One variety uses a serial verb construction (Steever 1988) that combines the past or non-past affirmative form of the main verb with the negative auxiliary *al-* “not become”; both are inflected for congruent personal endings.

- (31) A. *cel-v-ēm*                      *all-ēm*  
           go-NPST-1ST PER. PL.    become.not-1ST PER. PL.  
           “We will not go” (*Pura* 31.11)
- B. *aṛi-nt-aṇ-aḷ*                      *all-aḷ*  
           know-PST-EUPH-3RD SG.FEM.    become.not-3RD SG.FEM.  
           “She did not know” (*Aka* 98.6)

Such constructions alternate with another in which the auxiliary verb *al-* “not.become” combines with the bare root of the main verb, rather than any inflected form. The compound verb *maṛavalēn* “I will not forget” in (26A) consists of the root of the main verb *maṛa* “forget” and *alēn* “I do not become.”

- (32) A. *maṛa.v-al-eṇ*  
           forget-not.become-1ST PER.SG.  
           “I will not forget” (*Pura* 395.32)
- B. *vāṇ-al-aḷ*  
           live-not.become-3RD SG.FEM.  
           “She will not live” (*Aka* 12.14)

Other auxiliary verbs occur in similar constructions: for example, when the auxiliary *tara* “give to you or me” combines with the bare root of the main verb, it indicates that the action is oriented toward the speaker or addressee in the speech event.

- (33) pō-tara  
 go-give.to.you.or.me  
 “to come” (*Kali* 56.31)

Such constructions fell into disuse by the medieval period, being replaced by auxiliary verb constructions in which the main verb appears in an inflected, nonfinite form. Thus, while the bare verb root could function as a free form in Old Tamil, it no longer does so in Modern Tamil.

#### 4.3.5 Nonfinite verbals

Old Tamil has three sets of nonfinite verbals: (i) primary forms, (ii) secondary forms, and (iii) verbal nouns. The primary forms directly add a suffix to the verb-stem or, rarely, to the tensed stem. Secondary nonfinite forms add a clitic to a primary form. Verbal nouns are nominalized forms which may be inflected for case. As the number and distribution of finite predicates is greatly limited in the Old Tamil sentence (see Steever 1988, Lehmann 1994), nonfinite forms figure prominently in complex syntactic structures.

##### 4.3.5.1 Primary forms

There are four primary nonfinite verb forms: (i) the conjunctive, (ii) the infinitive, (iii) the conditional and (iv) the adnominal. The suffixes for the conjunctive and the conditional have several allomorphs in free variation. The infinitive entails five subtypes with various semantic functions. The adnominal form has two tensed and two negative forms. Nonfinite verbals are illustrated in (34), (38), and (40) with forms of the verb *oḷir-* “shine.”

##### (34) Primary nonfinite verbals

<i>Conjunctive</i>	<i>oḷir-ā, oḷirū, oḷir-ntu, oḷir-pu</i>
<i>Negative conjunctive</i>	<i>oḷir-ā, oḷir-ā-tu, oḷir-ā-mal, oḷir-ā-mai</i>
<i>Infinitive</i>	<i>oḷir-a, oḷir-iyā, oḷir-iyar, oḷir-mār, oḷir-vāṇ</i>
<i>Conditional</i>	<i>oḷir-in, oḷir-nt-āl</i>
<i>Adnominal</i>	
<i>Past</i>	<i>oḷir-nt-a</i>
<i>Non-past</i>	<i>oḷir-um</i>
<i>Negative adnominal</i>	<i>oḷir-a, oḷir-a-ta</i>

The *infinitive* in its varieties is the most general of the nonfinite forms in Old Tamil, conveying such notions as circumstance, result, and purpose:

- (35) A. <sub>S</sub> [mēni nalam tolai.y-a]<sub>S</sub> tuyaram cey-t-ōṇ  
 body beauty lose-INF. distress-NOM. do-PST-3RD SG.MASC.  
 “He brought distress, and [her] body lost its beauty” (*Aka* 278.13–14)
- B. parū-ntu icai niṛ-ka.p pāṭ-in-aṇ  
 spread-CF renown-NOM. remain-INF. sing-PST-3RD SG.MASC.  
 “He sang to spread [your] renown and [make] it remain” (*Pura* 126.13)

The various *conditional* verb forms mark the protasis of a conditional sentence. The simple conditional verb forms do not differentiate all of the verbal categories that finite verbs do; they do not distinguish, for example, between past and non-past tense. A periphrastic construction may also be used in which the conditional form *āka* “become” combines with a finite verb to mark a protasis.

- (36) A. [ayar emar] ān-āl aytti.y-ēm  
 cowherd-NOM. our.kin-NOM. become-CND. cowherd+FEM.-NOM.-1ST PER.PL.  
 yām  
 1ST PER.PL.INC-NOM.  
 “If our kin are cowherds, we are cowherdesses” (*Kali* 108.9)
- B. [nin marpu muyaṅk-ēm āy-in] yām  
 you-OBL. breast-OBL. embrace-NEG.-1ST PER.PL. become-CND. we-NOM.  
 cāytum  
 SWOON-NPST-1ST PER.PL.  
 “If we (=I) did not embrace your breast, we would swoon” (*Aka* 218.15–17)

The *adnominal* forms (called *adjectival* or *relative participles* in the literature) are nonfinite verbals that co-occur with a following nominal, with or without intervening material. Adnominal forms typically subordinate a clause to the following nominal: in (37A) it subordinates a relative clause to a head noun; in (37B) it subordinates a sentential complement to a noun; and in (37C) it helps form a complex adverbial expression.

- (37) A. <sub>NP</sub>[S [<sub>t<sub>1</sub></sub> nīram pāy-nt-a]<sub>s</sub> kaṇai<sub>1</sub>]<sub>NP</sub>  
 breast-NOM. pierce-PST-ADN. arrow  
 “The arrow that pierced [his] breast” (*Kali* 57.14)
- B. <sub>s</sub>[tiram=um vaiyai.y=um cēr-kinr-a]<sub>s</sub> kaṇ kavin  
 riverbank-NOM.=and Vaiyai-NOM.=and join-NPST-ADN. eye captivation-NOM.  
 “The eye-captivating beauty of the [river] Vaiyai joining the riverbank”  
 (*Pari* 22.35)
- C. <sub>s</sub>[nī iravu va-nta.k]<sub>s</sub> kāl  
 you-NOM. night-NOM. come-PST-ADN. time  
 “The time that you came by night” (*Kali* 38.14)

#### 4.3.5.2 Secondary forms

The four secondary nonfinite forms combine a primary form with an independent word or a suffix: (i) the causal (< conjunctive in *-ntu* + *ēna*); (ii) the equative (< conjunctive in *-ntu* + *āṅku*); (iii) the concessive conditional (< conditional + *=um*); and (iv) the factive concessive (< infinitive + *=um*).

#### (38) Secondary nonfinite verbals

<i>Causal</i>	oḷir-nt-ēna
<i>Equative</i>	oḷir-nt-āṅku
<i>Concessive conditional</i>	oḷir-in-um, oḷir-nt-āl-um
<i>Factive concessive</i>	oḷir-a.v-um

Consider the factive concessive as an example of a secondary finite verb. It consists of a verb form in the infinitive and the clitic *=um* “and, even,” and is translated as “even though V.”

- (39) [nāṭaṇ var-a.v=um] ivaḷ mēṇi paca-pp-atu evaṇ  
 land-3RD SG.MASC. come-INF.=and she-GEN. body-NOM. be.pale-VN why  
 “Why is it that, even though her chief had come, her body is pale?” (*Aiṅk* 217.3–4)

#### 4.3.5.3 Verbal nouns

Nominalized verb forms or verbal nouns are divided into tensed and tenseless verbal nouns. The latter group has five variant forms. Some verbal nouns, such as those with *-pu*, never mark case.

#### (40) Verbal nouns

<i>Tenseless verbal noun</i>	oḷir-al, oḷir-kai, oḷir-tal, oḷir-pu, oḷir-vu
<i>Tensed verbal noun</i>	
<i>Past</i>	oḷir-nt-ā-mai, oḷir-nt-atu
<i>Non-past</i>	oḷir-v-atu
<i>Negative verbal noun</i>	oḷir-ā-mai

At one level of generalization, some verbal nouns may be analyzed as combinations of an adnominal form and an abstract pronominal head noun. This construction is transparent in some instances, but is obscured by fusional morphology in others.

- (41) A. [nir kaṇ-ṭ-ar-kku] aṇanku ākum  
 you-OBL. see-PST+ADN.-3RD PER.PL.-DAT. awe-NOM. become-NPST-3RD SG.NEUT.  
 “Those who saw you were struck with awe” (*Kali* 56.21)
- B. [kuvalai malar-tal] ari-tu  
 kuvalai-flower bloom-VN rare.thing-3RD SG.NEUT.  
 “It is rare for the kuvalai flower to bloom” (*Aiṅk* 299.2–4)

#### 4.3.5.4 Nonfinite verbal constructs

Nonfinite verb forms dominate the formation of complex structures: compound verbs and complex clauses. The *conjunctive* form (also, *adverbial participle*) *vantu* “coming” functions as a main verb in the compound verb construction of (42A), and as a form that conjoins two clauses in (42B). Note also that the tenseless verbal noun *ayar-tal* “accomplish” in (42B) embeds a complement beneath a verb of wishing.

- (42) A. va-ntu ilar  
 come-CF not.be-3RD PER.PL.  
 “They did not come” (*Pari* 9.25)
- B. vaikaṛai va-ntu vatuvai ayartal vēṇṭu-v-al  
 daybreak come-CF marriage accomplish-VN wish-NPST-1ST PER. SG.  
 “I want you to come at daybreak and marry [me]” (*Kali* 52.22–23)

### 4.4 Adjectives and adverbs

Old Tamil has two “minor parts of speech”: adjectives and adverbs. These differ grammatically from nouns and verbs; each set has few members.

#### 4.4.1 Adjectives

Adjectives include *aru* “difficult,” *nal* “good,” *putu* “new,” and *peru* “big,” as well as words denoting color. They are morphologically invariant, lacking inflections: the comparative and superlative degrees are marked syntactically. Adjectives do not behave like nouns to the extent that they occur neither as subject nor as object; they do not behave like verbs in that they neither subcategorize verbal arguments nor assign case. They occur only as adnominal



attributes with an adjectival function. Adjectives such as *nal* “good” could, however, be treated as defective nouns, specifically ones that lack case inflection. They can therefore occur in compound nouns, but never as the head. They may also participate in noun derivation: like nouns, *nal* may take the abstract derivative suffix *-mai*, yielding *naṇ-mai* “goodness” or *-tu*, yielding *naṇ-ru* “that which is good, a good thing.”

#### 4.4.2 Adverbs

Adverbs constitute an even smaller set of uninflected words. They occur only as attributes of verbs, nouns, and adjectives. Words such as *uru*, *naṇi*, and *tava*, all meaning “much,” are examples of adverbs. These appear to be verb roots that have become frozen in an idiomatic function. Some inflected word forms are grammaticalized as adverbs with a particular lexical meaning: for example, the conjunctive form *azi-ttu* from *azi* “finish” idiomatically means “again.” The paucity of adjectives and adverbs in Old Tamil reflects the transparent, agglutinative morphological character of the language, which tends to discourage extensive morphophonemic variation and defective morphology alike.

## 5. SYNTAX

### 5.1 Word order

Like other South Dravidian languages, Old Tamil is a head-final, SOV language. In simple clauses, the unmarked order of the major constituents is Subject–Object–Verb. While explicit case-marking allows for the permutation of noun phrases within the clause, the verb firmly remains at the right clause boundary and is displaced from that position only under marked circumstances. The direct object tends to occur just before the verb, while other oblique arguments occur before the direct object but after the subject. The texts include departures from this expected SOV template, leading some (Rajam 1992) to doubt that the language is indeed SOV. In (43), for example, the object follows the verb rather than precedes it.

- (43) *orūpa*                                      *nin.n-ai*  
       shun-NPST-3RD PER.PL. you-ACC.  
       “‘They shun you” (*Pati* 34.1)

It must be borne in mind, however, that the Old Tamil corpus consists almost exclusively of poetic discourse. To accommodate their poetic designs, the bards permuted constituents and so departed from canonic SOV patterns. However, just below the surface lie robust SOV syntactic patterns. In harmony with an overall SOV framework, genitives in Old Tamil always precede their heads, main verbs always precede auxiliaries, and relative clauses always precede their head nominals.

### 5.2 Sentence structure

The simple sentence in Old Tamil consists of a subject and predicate. While the great majority of texts cast the subject in the nominative case, a few examples appear to cast it in the dative (44E), a phenomenon well documented in other South Dravidian languages. The predicate of a simple sentence may be a finite verb (44A, C, E) or a predicate nominal (44B, D) without any copula:

- (44) A. *nī munn-att-ān kṛtt-in-ai*  
 you sign-OBL.-INSTR. show-PST-2ND PER.SG.  
 “You showed with a sign” (*Kali* 61.7)
- B. *entai.y=um nuntai=um emmurai kṛlir*  
 my.father-NOM.=and your.father-NOM.=and what.degree kin-PL.-NOM.  
 “What kin are your father and mine?” (*Kuru* 40.2)
- C. *yān=um nī.y=um evvāṇi aritum*  
 I-NOM.=and you-NOM.=and what.path-NOM. know-NPST+1ST PER.PL.  
 “On what path would you and I meet?” (*Kuru* 40.3)
- D. *tōl-ēn*  
 shoulder-1ST PER. SG.  
 “I, with [broad] shoulders” (*Aka* 82.18)
- E. *nīn.a-kk=ō ariyunaḷ neñc=ē*  
 you-DAT.=INTERR. know-VN-3RD SG.FEM. heart=VOC.  
 “Is she someone known to you, O my heart?” (*Narr* 44.5)  
 meaning: “Do you know her, O my heart?”

### 5.3 Agreement

In simple sentences, predicates agree in person, number, and gender with subjects in the nominative case; agreement is marked by personal endings on the predicate. While predicate nominals in Old Tamil carried personal endings to mark agreement with their subjects (44B, E), their counterparts in Modern Tamil no longer do so.

### 5.4 Pro-drop

The use of personal endings on finite predicates allows for the omission of a subject noun phrase; consequently, Old Tamil is a pro-drop language. However, subject pronouns are seldom dropped when they occur in an extended usage (Steever 1981:80ff.). When, for example, the second-person plural pronoun is used honorifically for a singular referent, it is rarely dropped; nor is the first-person inclusive plural pronoun *yām* “we and you” absent when it is used to denote the speaker in soliloquy. Other arguments may be omitted as well. No South Dravidian language, including Old Tamil, has a verb phrase constituent, with the result that verbs need not overtly mark their objects to show their valence.

### 5.5 Clitics

Old Tamil has several clitics which may be added to noun and verbal forms, but not to adjectives. Although they combine morphologically with a noun or verb, their scope is the entire phrase or clause, whose head is that noun or verb. The clitic *=um* “and” coordinates noun phrases and nonfinite clauses; *=ō* and *=kol* mark a clause as being interrogative; and *=ē* “even” indicates emphasis.

### 5.6 Compound and complex clauses

The examples of (44B) and (44C) reveal that the subject of a simple sentence may be a coordinate structure; the quantifier *=um* “and, all,” morphologically a clitic, is added to each constituent of the conjunct. Predicates, however, may not be baldly conjoined in this

way to form complex sentence structures; instead, a variety of morphological and lexical devices are used to create complex structures with multiple clauses.

Lacking a distinct morphological category of *conjunction*, Old Tamil deploys its nonfinite verb forms (see §4.3.5) to join one clause to the next in the formation of subordinate and coordinate clauses. As a rule, there can be only one finite predicate per sentence; it generally occurs rightmost in the sentence and c-commands all other verbs in the sentence. All other predicates are nonfinite (see Steever 1988, Lehmann 1994). In (45A) the nonfinite conjunctive form *aṛintu* “knowing” conjoins the two clauses “my mother knows” and “my mother understands” to form a coordinate sentence; in (45B), the nonfinite infinitive form *tolaiya* “lose” subordinates a result clause to the following main clause.

- (45) A.  $s_0[s_1[yāy_1 \text{ aṛintu}]_{s_1} \text{ } s_2[t_1 \text{ uṇar-ka}]_{s_2}]_{s_0}$   
 my.mother know-CF understand-OPT.  
 “May my mother know and understand [it]” (*Aka* 203.2)
- B.  $s_0[s_1[mēni \text{ nalam tolai.y-a}]_{s_1} \text{ } \text{tuyarm} \text{ } \text{cey-t-ōn}]_{s_0}$   
 body beauty lose-INF. distress-NOM. do-PST-3RD SG.MASC.  
 “He brought distress, and/so that [her] body lost its beauty” (*Aka* 278.13–14)

The constraint against multiple finite predicates prevents any predicate nominal from appearing in a subordinate clause or a nonfinal conjunct *tout court*. In ruling out finite predicates in such contexts, it would also prevent the possibility of direct discourse. While Old Tamil texts do contain some examples of asyndetic parataxis, the language has three verbs that permit a finite predicate to be embedded in a complex sentence. The verbs *āka* “become,” *ēna* “say” and *pōla* “resemble” may take as their objects expressions of any category without imposing any morphological change on those objects; they may consequently embed finite predicates. In this capacity, these verbs contribute no lexical meaning to the structure. But as verbs, they may occur in nonfinite forms and thus be embedded in the larger structure. The infinitival form *ēna* “say” in (46A) embeds the finite verb *uḷan* “he is” – it acts as a complementizer to mark direct discourse. The conditional form *ānāl* “if become” in (46B) embeds the predicate nominal *emar* “our kin” in the protasis of a conditional. On occasion, other verbs of communication or perception, such as *kēḷ* “hear” (46C), may also embed finite predicates.

- (46) A.  $s_0[s_1[nin \text{ makan} \text{ } yāṇṭ(u) \text{ } uḷan=ō}]_{s_1} \text{ } \text{ēna}$   
 your son-NOM. where-NOM. be-NPST-3RD.SG.MASC.=INTERR. say-INF.  
 $\text{viṇavuti}]_{s_0}$   
 ask-NPST+2ND PER.SG.  
 “You ask (saying), ‘Where is your son?’” (*Pura* 40.1–2)
- B.  $[ayar \text{ } emar] \text{ } \text{ān-āl} \text{ } \text{aytti.y-ēm}$   
 cowherd-NOM. our.kin-NOM. become-CND. cowherd+FEM.-NOM.-1ST PER.PL.  
 $yām$   
 $\text{we}^{\text{inc}}\text{-NOM.}$   
 “If [it is the case that] our kin are cowherds, we are cowherdresses” (*Kali* 108.9)
- C.  $s_0[s_1[nin.a-kku \text{ } onru \text{ } kuṟuvam]_{s_1} \text{ } \text{kēḷ} \text{ } \text{iṇi}]_{s_0}$   
 you-DAT. one.thing tell-NPST-1ST PER.PL. listen-IMPV. now  
 “Listen now to what I have to say to you” (*Kali* 55.5)

Finally, Old Tamil permits finite verb forms to occur where nonfinite forms might otherwise be expected. The language has a structure called *murreccam* in the traditional

grammatical literature – a finite form that functions as a nonfinite form (Steever 1988: 45–52). The negative compound verbs in (31) above provide examples. Consider (47), where the finite past tense forms *āṭiṇir pāṭiṇir* “you celebrated” (lit. “you sang and danced”) occur where nonfinite conjunctive forms *āṭ-i pāṭ-i* “celebrating” (lit. “singing and dancing”) are expected:

- (47) *āṭi-ṇir*                      *pāṭ-ṇir*                      *cel-ṇ=ē*  
 dance-PST-3RD PER.PL. sing-PST-3RD PER.PL. go-CND.=EMP.-PCL.  
 “If you go celebrating (lit. singing and dancing)” (*Pura* 198.10)

This construction occurs in many other Dravidian languages (Steever 1988), but has dropped out of Modern Tamil.

## 6. LEXICON

The vocabulary of Old Tamil significantly reflects its Dravidian lexical heritage; however, even in the earliest stages of the language we find borrowings from Indo-Aryan languages, principally Sanskrit and the Prakrits. Sanskrit proper names appear in early texts: for example, in *Pura* 161.6, *Kaṅkai* “Ganges” renders the Sanskrit *Gaṅga*. Further examples of borrowings include (i) Prakrit *pāhuḍa* “gift,” which was borrowed into Old Tamil as *pākuṭam* with appropriate adjustments and rephonemicization to the Tamil sound system; (ii) Sanskrit *nagara* “town,” which became Tamil *nakar* (*Pura* 6.18); (iii) Prakrit *khavaṇa* “eject.” which gave Tamil *kavaṇai* “sling” (*Kali* 23.2).

In Old Tamil, the set of verbal bases was closed so that all borrowed words were nouns (note *kavaṇai* “sling” above). In the medieval language, the set of verbal bases was slightly expanded through borrowing; for example, *yōci-kka* “to think, ponder” comes from the Sanskrit root *yuj-*. However, this set became closed again with the advent of the modern language. In general, verbs from other languages are borrowed into Tamil as nouns that may then be compounded with native Tamil verbal bases.

Despite travel abroad and contacts with traders from the classical Mediterranean world in the Tamil coastal emporia (where spices and other luxury goods were traded for Roman gold coins), Tamilians appear not to have borrowed words from Western sources in antiquity. Attempts to assign particular words to Greek or Latin sources are uncertain – based perhaps more on fancy than on careful philological and etymological analysis. In the medieval and modern periods, however, Tamil has borrowed from a wide array of source languages – Indo-Aryan, Persian, Arabic, Portuguese, and English, among others.

## 7. DISCOURSE

Two poems are presented to illustrate connected discourse in Old Tamil. The language is represented by a fixed corpus; specifically, Middle Old Tamil is attested in two anthologies, which consist of 2,381 poems that range in length from 3 to 782 lines, totaling approximately 32,000 lines. It is only during the medieval period that this corpus came to be known as the *caṅkam* “academy, community” literature or *cāṇṇor ceyyul* “poetry of the nobles.”

The Old Tamil corpus consists primarily of poetic compositions. Ramanujan (1985:xi) observes that the poems of the two major anthologies are both “classical,” i.e., ancient, early,

and “classics”, i.e., they have withstood the test of time. As many as 475 bards composed these poems for their patrons, who were usually kings or chieftains. The poetry of the two anthologies treats two principal themes: love and war. While Tamil literature later grows to include didactic verse, epics and other literary forms, the poems of the two anthologies inform much of the later Tamil canon, as evidenced in medieval devotional (*bhakti*) literature.

Kailasapathy (1968) characterizes this corpus as bardic: the poetry was orally composed and transmitted, as various figures and tropes attest. Further, the poems form a highly coherent literary body: the many poets appealed to a shared set of conventions treating composition, leading to the name *caṅkam* “academy, community.” According to tradition, the first book in Old Tamil is the grammar, *Tolkāppiyam* “On ancient composition.” Its first two books treat phonology and morphology but the third, *Poruḷatikāram* “chapter on content,” discusses what constitutes a well-formed poem.

The *Poruḷatikāram* and later commentaries enumerate such poetic elements as appropriate themes, characters, landscapes, figures and meter. Poems belong to one of two main groups: *akam* “interior” and *puṛam* “exterior” The *akam* poems treat the different phases of love. These poems are general; the characters in them are not historical persons, but actors in an interior drama. The *puṛam* poems detail various acts of heroism such as valor in war or daring in cattle raids.

These poems are dramatic in two senses. First, they distill an experience – that is, an insight gained from action – and give it shape through poetic figures. Second, the colophons of the poems observe that each poem is spoken by a particular character in a drama: one might stipulate that a heroine is addressing her confidante over the delayed return of her lover. The poet thus stands at a remove, reinforcing the often anonymous nature of the composition.

The first poem, number 86 in the anthology *puṛaṇānūru* “Four hundred verses on heroism,” is a *puṛam* poem. The colophon observes that it is uttered by the mother of a warrior.

*What his mother said:*

<i>ciṛ.r-il</i>	<i>narrūṇ</i>	<i>parr-i</i>	<i>nin</i>	<i>makaṇ</i>		
small.house	pillar-NOM.	grasp-cf	you-OBL.	son-NOM.		
<i>yāṇṭ(u)</i>	<i>uḷaṇ=ō</i>		<i>ena</i>	<i>viṇavuti</i>	<i>en</i>	<i>makaṇ</i>
where-NOM.	be-NPST-3LD MASC.=INT		say-INF.	ask-NPST+2ND SG.	I-OBL.	son-NOM.
<i>yāṇṭ(u)</i>	<i>uḷaṇ</i>		<i>ā.y-in=um</i>	<i>ari.y-ēṇ</i>		<i>ōrum</i>
where-NOM.	be-NPST-3RD SG.MASC.		become-CND.=AND	know-NEG.+1ST SG.		once
<i>puli</i>	<i>cērntu</i>	<i>pōkiya</i>	<i>kal</i>	<i>alai</i>	<i>pōla</i>	
tiger-NOM.	join-CF	go-ADN.	stone-NOM.	cave-NOM.	resemble-INF.	
<i>īṇra</i>	<i>vayir=ō</i>		<i>itu.v=ē</i>			
give.birth-PST-ADN.	womb-NOM.=INTERR.		this.one=EVEN			
<i>tōṇravaṇ</i>		<i>mātō</i>	<i>pōrkalla.t</i>	<i>tāṇ=ē</i>		
appear-VN-3RD SG.. NEUT.		indeed	battlefield-OBL.	indeed=EVEN		

“You grasp the pillar of my hut and ask:  
 ‘Where is your son?’ Wherever my son might be,  
 I don’t know.

Though this womb, that gave him birth,  
 was once a den for that tiger,  
 Now he appears only on battlefields.”

The second poem belongs to the *akam* genre, and comes from the anthology *kuruntokai* “Collection of short poems” (40).

*What he said to her:*

<i>yāy=um</i>	<i>ñāy=um</i>	<i>yār</i>	<i>āk-ir=ō</i>
my.mother-NOM.=AND	your.mother-NOM.=AND	who-NOM.	become-NPST-3RD PL=INTERR.
<i>entai.y=um</i>	<i>nuntai=um</i>	<i>emmurai</i>	<i>kēḷir</i>
my.father-NOM.=AND	your.father-NOM.=AND	what.degree	kin-NOM.
<i>yān=um</i>	<i>nī.y=um</i>	<i>evvazī</i>	<i>aṛitum</i>
I-NOM.=AND	you-NOM.=AND	what.path-OBL.	know-NPST+1ST PL.
<i>cem-pola.p-peya-nīr</i>	<i>pōla</i>		
red-earth-rain-water	resemble-INF.		
<i>aṇp-uṭai</i>	<i>neñcam</i>	<i>tāṇ</i>	<i>kala-nt-aṇa.v=ē</i>
love-have	heart-NOM.	indeed	mix-PST-3RD PL.=EVEN

‘What is my mother to yours?  
 What kin are your father and mine?  
 And on what path could you and I have met?  
 ‘But, with love, our hearts have mingled  
 Like the red earth and pouring rain.’

### Abbreviations

<i>adn.</i>	adnominal form
<i>cf</i>	conjunctive form
<i>cnd.</i>	conditional
<i>euph.</i>	euphonic particle
<i>npst</i>	non-past tense
<i>pst</i>	past tense
<i>vn</i>	verbal noun
<i>vs</i>	verb-stem
=	clitic boundary
—	simple morpheme boundary
+	portmanteau form

### Texts cited

<i>Aka</i>	Akanānūru	<i>Pati</i>	Patirrupattu
<i>Aiṅk</i>	Aiṅkurunūru	<i>Peru</i>	Perupaṇāruppatai
<i>Kali</i>	Kalittokai	<i>Pura</i>	Puraṇānūru
<i>Kuru</i>	Kuruntokai	<i>Malai</i>	Malaipaṭukaṭam
<i>Narr</i>	Narriṇai	<i>Matu</i>	Maturaikkañci
<i>Pari</i>	Paripāṭal		

### Bibliography

- Akattiyalīṅkam, C. 1983. *Caṅkat Tamiz I, II, III*. Annamalainagar: Aṇaitintiyat Tamiz Moziyiar Kazakam.
- Andronov, M. 1969. *A Standard Reference Grammar of Modern and Classical Tamil*. Madras: New Century Book House.
- Caminataiyar, U. 1950. *En Carittiram*. Madras: Daktar U. Ve. Caminataiyar Nūl Nilaiyam.

- Daniels, P. and W. Bright (eds.). 1996. *The World's Writing Systems*. Oxford: Oxford University Press.
- Kailasapathy, K. 1968. *Tamil Heroic Poetry*. London: Oxford University Press.
- Lehmann, T. 1994. *Grammatik des Alttamil*. Stuttgart: Steiner.
- Paramasivam, K. 1979. *Effectivity and Causativity in Tamil*. Trivandrum: Dravidian Linguistics Association.
- Rajam, V. 1992. *A Reference Grammar of Classical Tamil Poetry*. Philadelphia: American Philosophical Society.
- Ramanujan, A. 1985. *Poems of Love and War*. New York: Columbia University Press.
- Steever, S. 1981. *Selected Papers in Tamil and Dravidian Linguistics*. Maturai: Muttu Patippakam.
- . 1987. "Tamil and the Dravidian languages." In B. Comrie (ed.), *The World's Major Languages*, pp. 725–746. London: Croom Helm.
- . 1988. *The Serial Verb Formation in the Dravidian Languages*. Delhi: Motilal Banarsidas.
- . 1993. *Analysis to Synthesis: The Development of Complex Verb Morphology in the Dravidian Languages*. New York: Oxford University Press.
- . 1996. "Tamil writing." In Daniels and Bright 1996, pp. 426–428.
- Zvelebil, K. 1967. (The language of Perunkunrur Kilar). In K. Zvelebil, J. Glazov, and M. Andronov (eds.), *Introduction to the Historical Grammar of the Tamil Language*. Moscow: Nauka.
- . 1973. *The Smile of Murugan: On Tamil Literature of South Asia*. Leiden: Brill.
- . 1974. *Tamil Literature*. Wiesbaden: Otto Harrassowitz.

# Mayan

VICTORIA R. BRICKER

## 1. HISTORICAL AND CULTURAL CONTEXTS

### 1.1 Linguistic prehistory and history

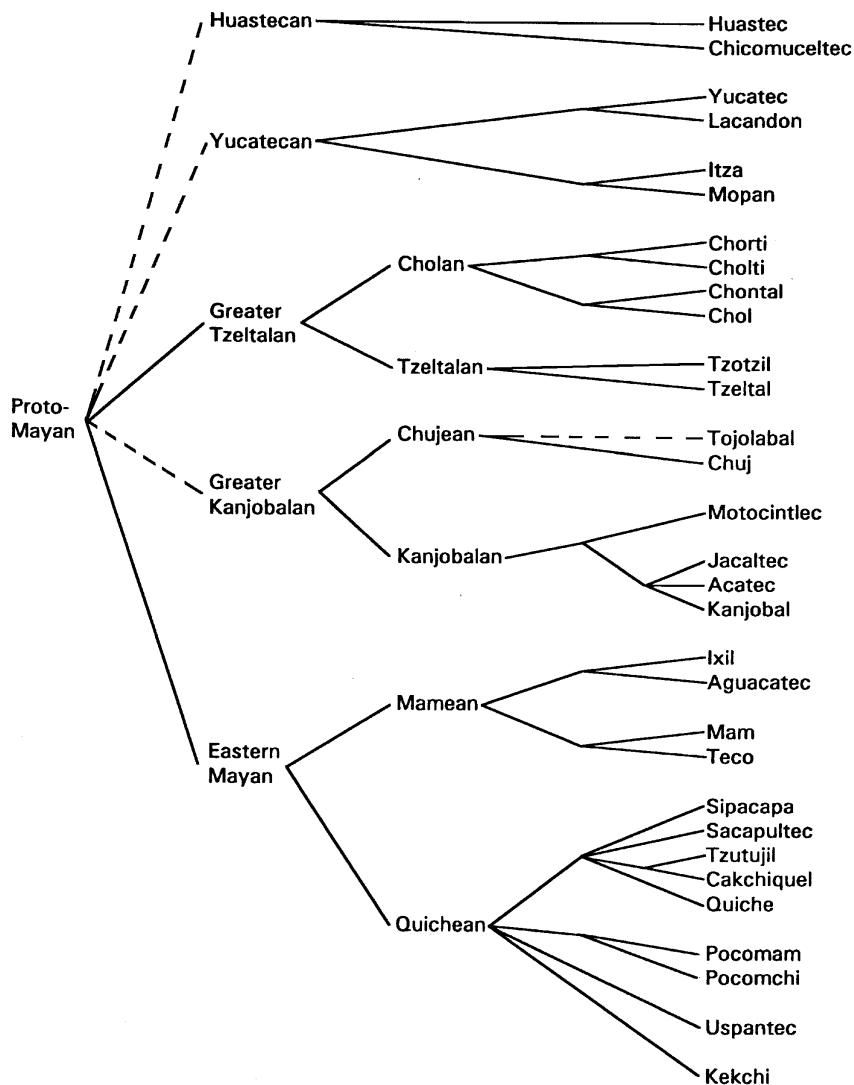
The language described herein as *Mayan* is known from a hieroglyphic script that was employed in a large region in Mesoamerica, encompassing much of what is today southeastern Mexico, northern and eastern Guatemala, all of Belize, and the western part of Honduras. The earliest securely dated and geographically provenienced hieroglyphic text from this region is Stela 29 from Tikal in northeastern Guatemala, which bears a date of 6 July AD 292. The script was in continuous use in this region until the second half of the sixteenth century, when it was replaced by the Latin-based alphabet introduced by the Spaniards.

Lyle Campbell and Terrence Kaufman (1985) have classified the thirty or so Mayan languages in terms of five branches, each of which is further divided into groups and subgroups (Fig. 43.1). At the time of the Spanish Conquest, the inhabitants of the region where hieroglyphic texts have been found spoke languages representing the Yucatecan and Greater Tzeltalan groups (Fig. 43.2). The Yucatecan languages were confined to the Yucatan peninsula in the north. South of them was a broad band of Cholan languages, running from Chontal and Chol in the west to Cholti and Chorti in the east. Tzeltal was the only language in Tzeltalan Proper that was spoken in the region under consideration.

Kaufman's (1976) glottochronological estimates suggest that by AD 292 Yucatecan had already separated from Huastecan, and that Cholan and Tzeltalan Proper had already differentiated from each other. This means that the language which is herein called "Mayan" may have represented three quite distinct languages – Yucatecan, Cholan, and Tzeltalan – not dialects of a single language. By AD 600, the Cholan languages had differentiated into Chorti, Chol, and Chontal, and Tzeltal had separated from Tzotzil. The Yucatecan languages, Yucatec, Lacandon, Itza, and Mopan, probably did not emerge as separate languages until after AD 950 (Justeson *et al.* 1985:14–16), and so are outside the scope of this study.

The region inhabited by speakers of "Mayan" can, for the most part, be classified as lowland (which I have defined as land lying below 600 meters; see Bricker 1977). The exceptions include the adjacent highlands of eastern Chiapas and western Honduras, which were inhabited by speakers of Tzeltal and Chorti, respectively, when the Spaniards arrived (Fig. 43.2). The remaining Cholan languages and all the Yucatecan languages were limited to the lowlands at that time (Bricker 1977).

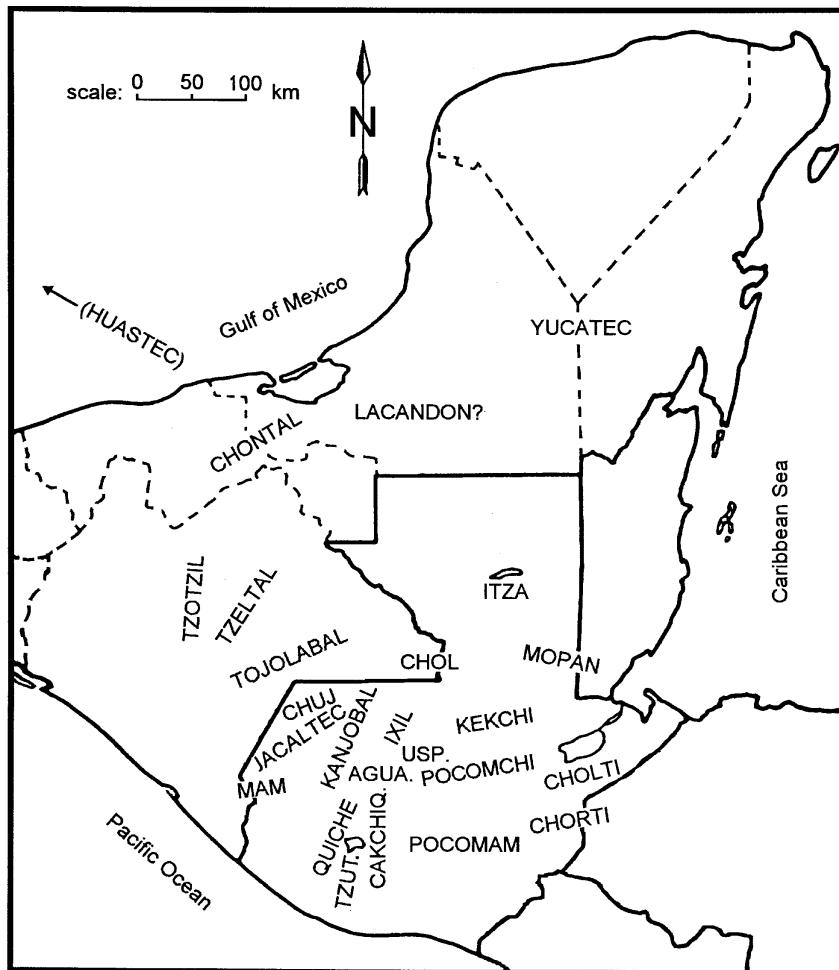




**Figure 43.1** Classification of Mayan languages. Dotted lines represent less secure groupings. After Campbell and Kaufman (1985:fig. 1)

## 1.2 The people and their culture

The ancestors of the people who spoke the languages encoded in the script did not appear in the region until about 1000 BC, which marks the beginning of the period archeologists call the Middle Preclassic (see Sharer 1994:Table 2.1). Historical linguists believe that they came out of the highlands of Guatemala in the south (e.g., Kaufman 1976:106–109). These first Maya settlers in the lowlands were maize farmers who lived in villages and larger, nucleated settlements dominated by terraced platforms and public buildings arranged in clusters connected by causeways (Sharer 1994:80–83). The Late Preclassic that followed (400 BC–250 AD) was characterized by “a rapid growth in population and in the development of stratified organizations, as demonstrated by elaborate funerary remains, massive ceremonial structures housing the artifacts of a variety of ritual activities, and the crystallization of a sophisticated art style, all recognized as typically Maya” (Sharer 1994:85). However, these defining traits of lowland Maya civilization did not yet include writing, which seems



**Figure 43.2**  
Approximate locations of  
Mayan languages in AD  
1550

to have been invented by speakers of non-Mayan languages outside the region in question during the Late Preclassic period (Ch. 44, §2.2) and was not adapted for use with Mayan languages until the very end of this period.

The Late Preclassic period was followed by the Early Classic (AD 250–600), a period characterized by Sharer (1994:138) as “the era when state-level political organizations developed and expanded in the Maya area, especially in the southern and central lowlands.” The settlements were larger, with a well-defined central core surrounded by residential areas. The centers contained various specialized structures faced with stone, such as palatial residences for the rulers and their families, ballcourts, and temples on stepped platforms (Sharer 1994:475). Stone was gradually replaced by perishable pole and thatch as building materials for residences, moving out from the center to the periphery of these cities.

### 1.3 The documents and their content

A strikingly new feature of Early Classic cities was the number of public monuments with inscribed hieroglyphic texts. Texts could be found carved on the walls and in the doorways

of buildings, including the lintels, jambs, and steps. They also appeared on the stone rings of ballcourts, through which the rubber ball had to pass in order for a team to score, and on free-standing slabs of stone, called stelae, scattered around the center of a site (Fig. 43.3).

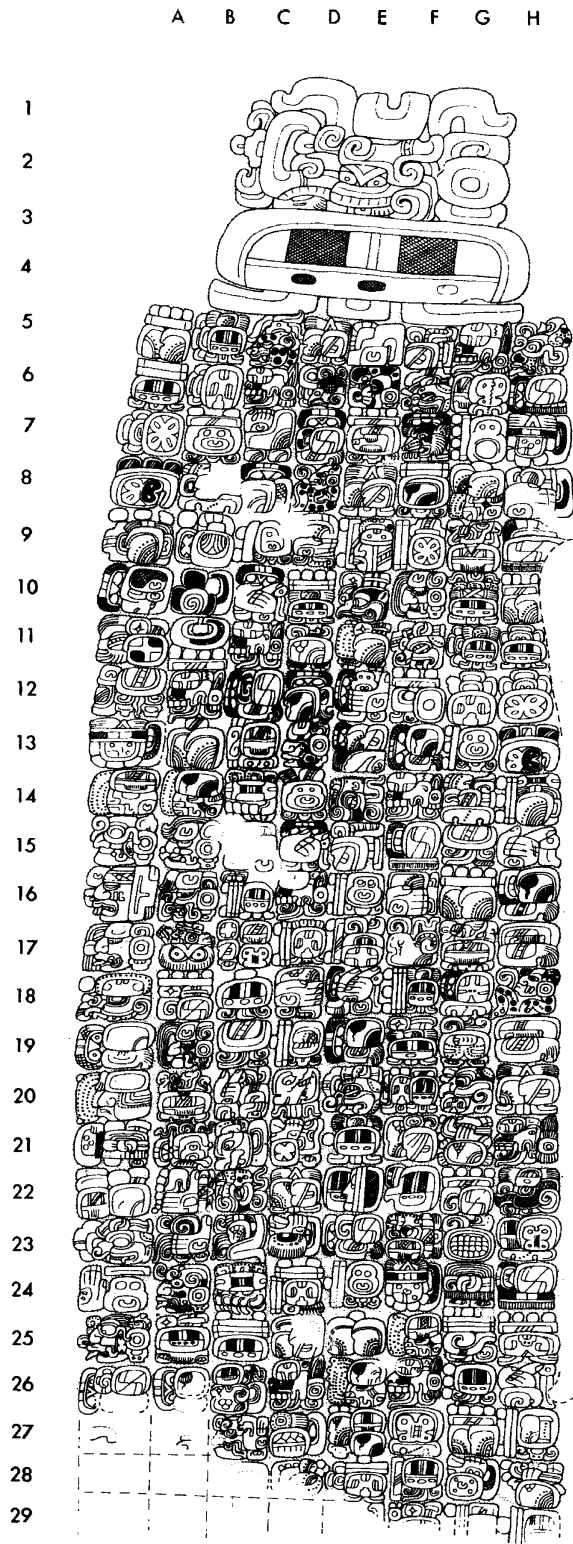
Ceramic vessels had bands of hieroglyphs around their rims, designating the ritual substances for which they served as containers (e.g., chocolate [Stuart 1988]), describing them as plates, cups, or bowls (Houston and Taube 1987), and naming the owners of the vessels and the artists who had painted or carved the texts on them (Stuart 1987:1–11). Pendants and earspools and flares of jade, shell trumpets, animal bones, and even sting-ray spines used in bloodletting often contained short hieroglyphic texts. Elaborate painted murals composed of both text and pictures covered the walls of rooms and tombs. Everything we know about the Early Classic form of the language of Mayan hieroglyphs comes from these sources. There may also have been, as there were in later centuries, screenfold books, or codices, made of animal hide and fig-bark paper (Sharer 1994:272); if so, the humid, tropical climate of the lowlands did not favor their survival into modern times.

The hieroglyphic texts carved on stelae and on the surfaces of buildings were primarily historical in content, containing the biographies of the rulers of the cities, which highlighted the dates of their birth, marriage, accession to office, raids on other cities, and death or burial. They also record genealogical information about the rulers and the rituals they performed at the end of major time periods and on anniversaries of the dates when they took office. Some of the longer texts refer to a succession of rulers, resembling, in this respect, the king lists of the ancient Near East.

The texts relating to a single ruler are often distributed rather widely over a site, with different “chapters” inscribed on separate buildings and stelae in several locations. For this reason, much of the research carried out by epigraphers has involved determining how the texts are related to each other historically and piecing together the biographies of the individual rulers (e.g., Proskouriakoff 1960; Jones and Satterthwaite 1982:124–131). Fortunately, the ancient Maya had a sophisticated calendar that permitted them to specify the chronological position of events in a cycle of more than five thousand years, and they were rather compulsive about dating their texts. Therefore, the histories of the major Early Classic cities are known in considerable detail (see, *inter alios*, Jones and Satterthwaite 1982).

The calendar employed by the lowland Maya was probably borrowed from the people who invented the Epi-Olmec script during the Late Preclassic period. The base of this calendar was a period of 360 days known as the *tun*. The *tun* was divided into eighteen smaller periods called *winals*, each containing 20 days (*k'ins*). Twenty *tuns* formed a larger unit called a *k'atun*, and 20 *k'atuns* were grouped into a *pik*. The complete cycle consisted of 13 (not 20) *piks*, which was known as the Long Count. The beginning of the Long Count was arbitrarily set to coincide with 11 August 3114 (Gregorian) BC, according to the correlation of the Maya and Western calendars that agrees best with ethnohistorical sources from the sixteenth century, and it will end on AD 21 December 2012. Obviously all of Maya history recorded in hieroglyphs falls within this period.

In addition to the Long Count, the Maya calendar contains two other cycles which also have their roots in the earlier Epi-Olmec culture. One is a ritual, or divinatory, cycle of 260 days, composed of two subsidiary cycles based on a sequence of 20 named days (*Pimiš*, *Pik'*, *ʔak'b'al*, *k'an*, *čikč'an*, *kimi*, *manik'*, *lamat*, *muluk*, *ʔok*, *čuwen*, *ʔeb'*, *b'en*, *hiš*, *men*, *kib'*, *kab'an*, *ʔets'nab'*, *kawak*, and *ʔahaw*) and the numbers from 1 to 13, which serve as coefficients of the days. The other represents the vague year of 365 days, which is divided into 18 named months of 20 days each (*k'an-halab'*, *Pik'-k'at*, *čak-k'at*, *sots'*, *katsew*, *tsikin*, *yaš-k'in*, *mol*, *č'en*, *yaš*, *sak*, *keh*, *mak*, *ʔuniw*, *muwan*, *paš*, *k'anasiy*, and *kumk'u*) and a five-day intercalary “month,” *wayeb'*, that ended the year. The least common multiple of these two cycles is the



**Figure 43.3** Tikal, stela 31, back. After Jones and Satterthwaite (1982:fig. 52b)

so-called Calendar Round of 18,980 days or fifty-two years, which was the Maya counterpart of the European century.

With a firm grasp on the passage of time, the Maya had the tools for recording, and later predicting, astronomical events. By the end of the fourth century AD, they were relating Long Count dates to a lunar calendar, recording several kinds of information: (i) the age of the Moon on the date in question; (ii) the position of the current lunar month in a six-month semester; and (iii) the length of the month as either twenty-nine or thirty days. Eventually, they produced books of tables for predicting dates of solar and lunar eclipses, equinoxes and solstices, heliacal risings of Venus as morning star, and retrograde periods of Mars, examples of which have survived only from much later times (Bricker and Bricker 1983; Bricker and Bricker 1986, 1988).

Both the content of the texts and the media on which they were written suggest that their principal function was to glorify the elite. The focus is on dynastic history, ritual, and on designating the owners and makers of highly valued objects, such as elaborately painted and carved vases and jade and shell ornaments. And although tribute items seem to be mentioned in some Late Classic texts (Stuart 1993), no records of mundane commercial transactions have been preserved in Mayan script.

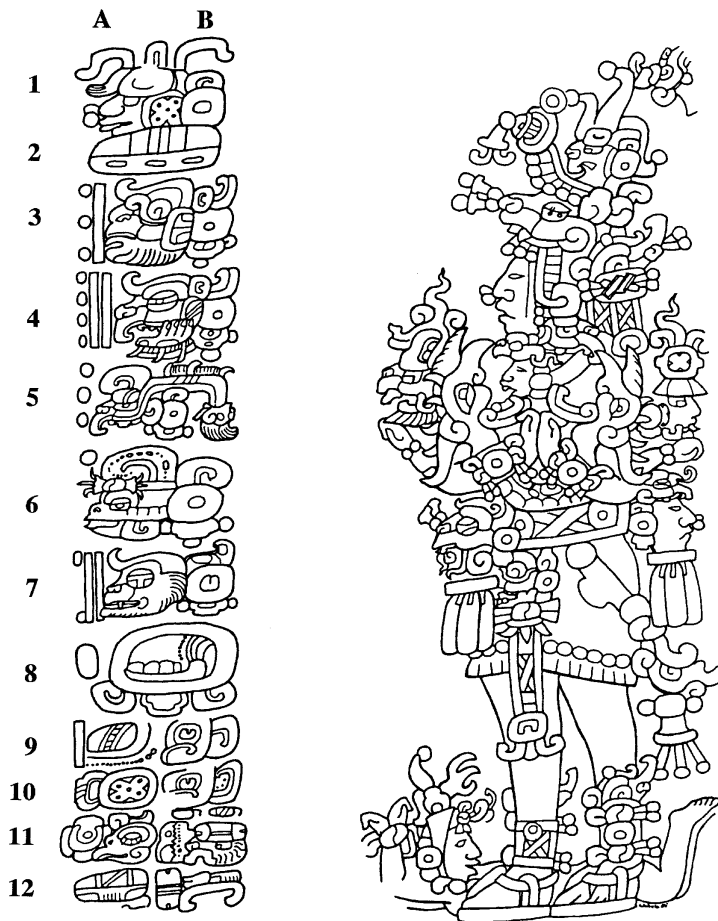
## 2. WRITING SYSTEM

### 2.1 Principles of Mayan writing

The Maya had a mixed writing system, consisting of both logographic and syllabic signs. The total number of different signs that have been identified ranges between 650 and 700 for the corpus as a whole, but the number of signs used at any one time apparently never exceeded 400 (Grube 1994:177). These figures are consistent with the logosyllabic nature of Mayan writing.

The reading order of a Mayan text is from top to bottom and from left to right in paired columns. The columns are labeled by scholars with capital letters and the rows with numbers. A glyph block is normally designated by a combination of a letter and a number, for example A5. In Figure 43.3, for example, after a large introductory glyph that accounts for four rows, the text begins at A5 and moves on to B5, then A6 and B6, A7 and B7, until the end of the first two columns. The reader then moves on to C5 and D5, C6 and D6, and so on through the inscription. The individual glyph blocks are also read from left to right and from top to bottom: prefixes and superfixes appear before the main sign, which in turn is read before postfixes and subfixes. The following transcription conventions are used in this chapter: phonetic transcriptions of the glyphs appear in boldface type, whereas morphemic transcriptions are italicized.

Phoneticism appears quite early in the history of Mayan writing. By AD 320, there is already evidence of the use of phonetic complementation, in which a word is represented by a logogram, but another sign is added to it as a prefix or a suffix to indicate how part of it is pronounced (Justeson and Mathews 1990:117). The first evidence of its use is in a text on a jade plaque bearing a Long Count date and Calendar Round corresponding to 15 September AD 320 in the Gregorian calendar (Fig. 43.4 left). The text records the accession of the ruler who is pictured on the other side of the plaque (Fig. 43.4 right). The collocation in question (at B9 in Fig. 43.4 left; see Fig. 43.5a) consists of the logogram for a verb meaning “sit” (*čum* in Cholan and *kum* in Yucatecan) and the sign for **mu**, which indicates that the final consonant of the word represented by the logogram is /m/ (Ringle 1985:153–154). Note that the vowel in **mu** also echoes the vowel /u/ in *čum/kum*.



**Figure 43.4** The Leyden Plaque (left = back; right = front). Drawing by Linda Schele

The Early Classic inscriptions contain a number of other examples of phonetic complementation, including the words for “day” (*k’in*, spelled as **k’in[ni]** in Fig. 43.5d), “20-day month” (*winal*, spelled as **winal[la]** in Fig. 43.5f), “sky” (*čan* in Cholan and *kaʔan* in Yucatecan, labeled as **čan[na]** in Fig. 43.5h), and “yellow” (*k’an*, spelled as **k’an[na]** in Fig. 43.5j). Examples of the same logograms without phonetic complements (Fig. 43.5c, e, and g) suggest that complementation was optional in Mayan writing.

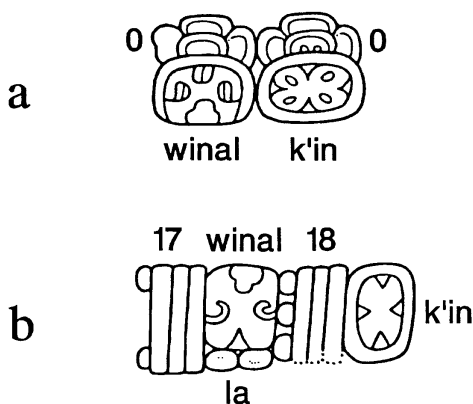
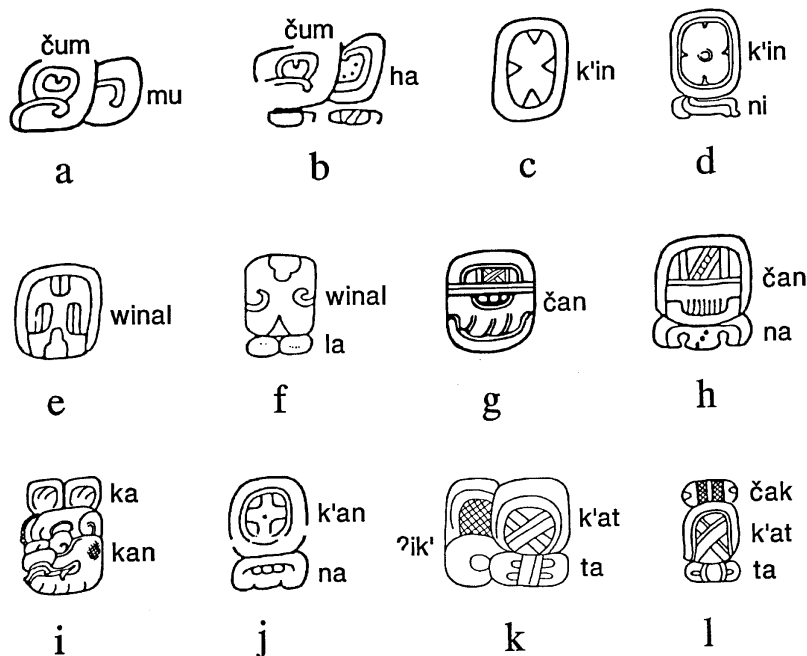
The original function of phonetic complementation was probably to disambiguate logograms that had several possible readings, but over time it was extended to logograms which had pronunciations that were never in doubt. Justeson and Mathews (1990:118–119) have found evidence that “extensions of orthographic practices are often promoted by similar practices in similar contexts.” For example, neither the **winal** nor the **k’in** logograms are polyvalent, so they can be read unambiguously without phonetic complements. These logograms frequently appear side by side in Long Count expressions (e.g., Fig. 43.6a). By AD 425, the **winal** logogram had acquired a phonetic complement (Fig. 43.6b), and less than a century later, in AD 514, the word *k’in* was also being spelled with a logogram plus phonetic complement in contexts where the **winal** logogram employed the same convention (Fig. 43.6c).

The first examples of syllabic writing can also be found in texts dating to the Early Classic period. Two kinds of syllables have been recognized in the script. One consists of a single vowel (V), the other of a consonant followed by a vowel (CV). Mayan words have two basic



**Figure 43.5** Phonetic complementation

(a, Leyden Plaque B9. b, Leyden Plaque B10. c, Balakbal, Stela 5 (Justeson and Mathews 1990:fig. 12). d, Palenque, Foliated Cross Tablet, D17 (Maudslay 1889–1902:IV, plate 82). e, Tikal, Stela 31, B6 (Jones and Satterthwaite 1982:fig. 52b). f, Balakbal, Stela 5 (Justeson and Mathews 1990:fig. 12). g, Tikal, Temple IV, Lintel 3, E3b (Jones and Satterthwaite 1982:fig. 74). h, Tikal, Stela 31, C13 (Jones and Satterthwaite 1982:fig. 52b). i, Pomona Panel, L1 (Schele and Miller 1986:fig. III. 12). j, Yaxchilan, Lintel 46, G3 (Graham 1979:101). k, Yaxchilan, Lintel 31, I4 (Graham 1979:71). l, Tila A, B5). a and b from a drawing by Linda Schele. l from *Maya Hieroglyphic Writing: An Introduction*, by J. Eric S. Thompson, fig. 16, 30. New edition copyright © 1960, 1971 by the University of Oklahoma Press

**Figure 43.6**

Phonetization (a, Tikal, Stela 31, G12–H12. b, Balakbal, Stela 5. c, Caracol, Stela 1, A3–B3). After Justeson and Mathews (1990:fig. 12)

**Figure 43.7** Syllabic and logosyllabic spellings (a, Rio Azul, pot, D. (Stuart 1988:fig. 2). b, Piedras Negras Lintel 3, P2. c, Rio Azul, pot, L. (Stuart 1988:fig. 2)

d, Tikal, Stela 31, M3 (Jones and Satterthwaite 1982:fig. 52a).

e, Caracol, Stela 16, B18 (Beetz and Satterthwaite 1981:fig. 15).

f, Naranjo, Initial Series pot, B' (Coe 1973:103). g, Tikal, Stela 31, L1 (Jones and Satterthwaite 1982:fig. 51a).

h, Tikal, Stela 31, P2 (Jones and Satterthwaite 1982:fig. 52a).

i, Rio Azul, Tomb 12, C. (Graham and Mobley 1986:456)

j, Rio Azul, pot, B (Stuart 1988:fig. 2). k, Yaxchilan, Lintel 2, Q (Graham and von Euw 1977:15).

l, Yaxchilan, Lintel 43, D4 (Graham 1979:95). m, Kabah, Structure 1

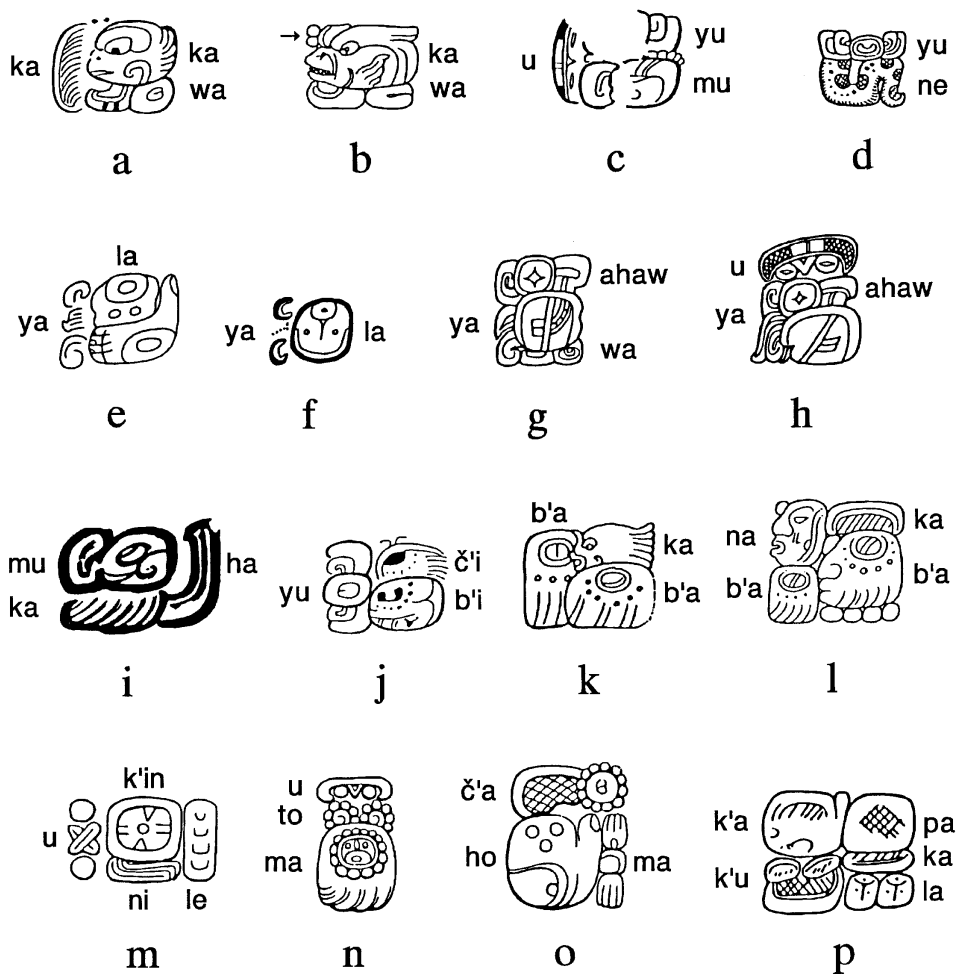
(Proskouriakoff and Thompson 1947:fig. 1g).

n, Copan, Altar S, glyph I (Maudsley 1889–1902:I, plate 94).

o, Copan, Stela A, F4 (Maudsley 1889–1902:I, plate 30).

p, Chichen Itza, Monjas, Lintel 2A, B1 (Graham 1977:269).

b after a drawing by John Montgomery



shapes: CVC and CVCVC. All words end in consonants, but only syllables ending in vowels have been attested in the script. This means that, in order to adapt such syllables to Mayan words, it was necessary to insert an extra vowel, which was never pronounced, at the end of the word. For example, the word *kakaw* “chocolate” was usually spelled syllabically as **ka-ka-wa** (Fig. 43.7a; the second **ka** sign is a variant of the first). This extra vowel is written in parentheses in transcriptions of syllabic spellings of Mayan words: **ka-ka-w(a)**. Another early example of the principle of vowel-insertion is the syllabic spelling of *u-yum* “his father” as **u-yu-m(u)** (Fig. 43.7c). The verb-stem, *muk-ah* “be buried,” is spelled as **mu-ka-h(a)** in Figure 43.7i. Note that the syllabic spelling overrides the morphemic boundary between *muk* and *-ah*. The same is true of the syllabic spelling of *y-al* “her child” as **ya-l(a)** in Figure 43.7e and f.

Occasionally a different spelling principle, consonant-deletion, was invoked, as in the rendition of *y-unen* “his, her child” as **yu-ne** (Fig. 43.7d). In words containing more than one instance of the same syllable, one of them might be omitted; in such cases, the scribe sometimes added two small dots beside the upper left corner of the sign, indicating that the syllable should be repeated when the word was pronounced (Stuart and Houston 1994:46 and Fig. 57; compare Fig. 43.7a and b).



	V	ʔ	b'	ts	ts'	č	č'	h	k	k'	l	m	n	p	p'	s	š	t	t'	w	y
a																					
e																					
i																					
o																					
u																					

**Figure 43.8** Early Classic syllables

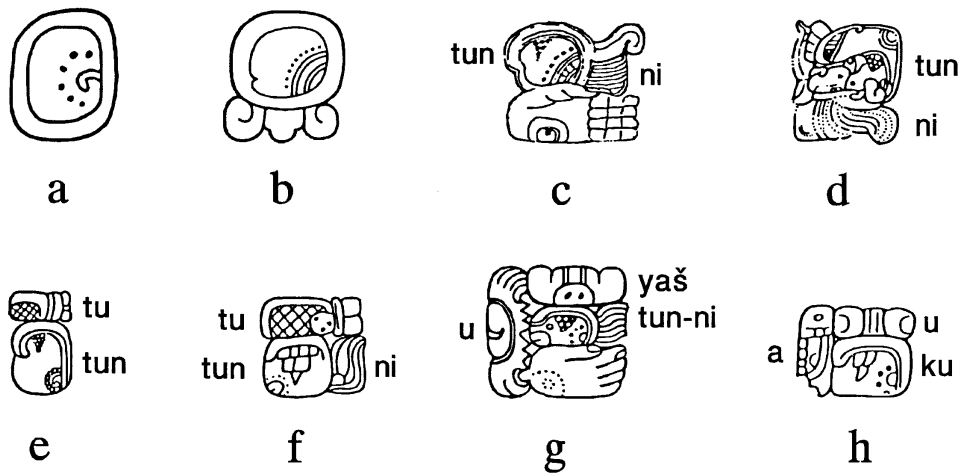
The syllabic signs that are known to have been in use during the Early Classic period (AD 250–534) are arranged in the grid shown in Figure 43.8.

For spelling words containing grammatical affixes, a mixture of logographic and syllabic principles was often employed. In Figure 43.7g, the word *y-ahaw* “his ruler” is spelled by prefixing the syllable **ya** to the logogram for **ahaw**. The possessive pronoun, *y-*, is represented by the consonant in **ya**, and the vowel complements the first /a/ in **ahaw**. A similar strategy was later adopted for the representation of grammatical suffixes. Figure 43.7m illustrates the logosyllabic spelling of *u-k'in-il* “on the day.” It contains one logogram (**k'in**) and three syllabic signs (**u**, **ni**, and **le**). The first syllabic sign (**u**) represents the clitic pronoun *u-*. The second syllabic sign (**ni**) has two functions: it complements the final consonant of **k'in** and also provides the vowel in the *-il* suffix. The third syllabic sign completes the spelling of the *-il* suffix by adding an /l/ and inserting an unpronounced /e/.

In addition to the phonetic complements that were used for clarifying which of several alternative readings for a logogram were intended, there were also semantic determinatives that served a similar purpose. The sign most commonly used as a determinative was a frame or cartouche that enclosed the glyphs for the days of the Maya week. An example of such a cartouche appears in Figure 43.9a, where it signals that the main sign enclosed by it refers to *kawak*, the nineteenth day of the twenty-day week.

In many cases, the cartouche rested on a pedestal, which served as a second semantic determinative for identifying day signs (Fig. 43.9b). When the main sign appeared without either the cartouche or the pedestal, it could be read in two different ways: as the syllable **ku** or the logogram **tun**. The logographic reading was usually signaled by the phonetic complement, **ni**, which was either postfixed or subfixed to the main sign (Fig. 43.9c, d, and g). During the Late Classic period, the **tun** reading was sometimes indicated by prefixing **tu** to the main sign (Fig. 43.9e), and occasionally both **tu** and **ni** served as complements for this sign (Fig. 43.9f). Finally, Figure 43.9h illustrates the syllabic use of this sign in the personal

**Figure 43.9** Polyvalence. (a, Piedras Negras, Lintel 3, D'6. b, Yaxchilan, Lintel 37, C6a (Graham 1979:83). c, Houston Lintel, E3. d, Piedras Negras, Lintel 12, E1. e, Piedras Negras, Throne 1, L. f, Yula, Lintel 1, B3. g, Piedras Negras, Altar 2 support, G3. h, Yaxchilan, Lintel 42, G3 (Graham 1979:93). a, d, and g after drawings by John Montgomery. c after a field sketch by Ian Graham. e–f from *Maya Hieroglyphic Writing: An Introduction*, by J. Eric S. Thompson, figs. 4.22 and 33.31. New edition copyright © 1960, 1971 by the University of Oklahoma Press



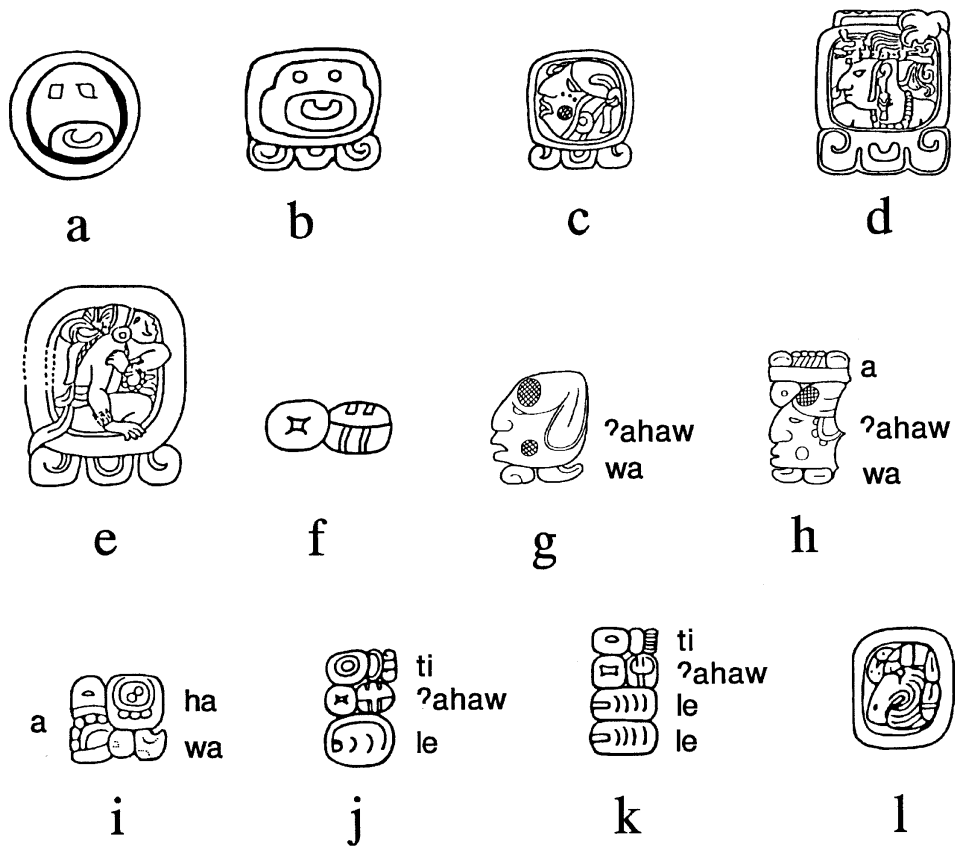
name, *?ah-?uk* (spelled as **a-u-k[u]**). These examples illustrate the polyvalent nature of many signs (cf. Fox and Justeson 1984). In this case, a single sign had three potential readings – **kawak**, **tun**, and **ku** – that were disambiguated by the presence or absence of two semantic determinatives and two phonetic complements. Other possible semantic determinatives include the hands shown supporting the glyphs for *y-al* “her child” and *tun* “stone, 360-day year” in Figures 43.7e and 43.9c, g, respectively. These examples contrast with other spellings of the same words that appear without the hand in Figures 43.7f and 43.9d.

The economy achieved by using one sign for several words and a syllable was outweighed by the great number of homophonous signs in the script. For example, the word *?ahaw* “lord, ruler” can be represented by a number of signs that differ markedly from one another. When it refers to the twentieth day of the Maya week, it can appear in a cartouche, with or without a pedestal (Fig. 43.10a and b). In that context, it is often shown as a simian face in frontal view, with two eyes, a nose, and a mouth (Fig. 43.10a and b). It can also be represented by the profile head of a young man, who is often depicted with a black dot in his cheek (Fig. 43.10c). In some cases, the head and the shoulders, or even the entire body of the young man, are shown (Fig. 43.10d and e). There is also a zoomorphic variant of this day sign as the profile head of a vulture (Fig. 43.10l). Still another variant, sometimes called “symbolic” or “geometric,” is never enclosed in a cartouche and therefore refers to a human ruler, not a day (Fig. 43.10f). The profile head variant without the cartouche and pedestal sometimes appears with phonetic complements (Fig. 43.10g and h), and there is a syllabic spelling of the same word that also lacks calendrical associations (Fig. 43.10i). The highly pictorial nature of the script has encouraged a multiplicity of sign forms, encompassing geometric, human, and zoomorphic head variants, and, occasionally, even full-figure depictions, that have greatly complicated the task of decipherment and the development of a usable font.

## 2.2 Evolution of Mayan writing

As the writing system developed, the inventory of syllabic signs shown in Figure 43.8 expanded in two ways: (i) the total number of syllables represented in the grid increased by

**Figure 43.10** Alternative spellings of *?ahaw* (a, Uaxactun, fresco, glyph 60. b, Tikal, Stela 31, D14 (Jones and Satterthwaite 1982:fig. 52b). c, Copan, Stela C, A2b. d, Quirigua, Stela D, D14. e, Copan, Stela D, A4b. f, geometric variant. g, Yaxchilan, Lintel 23, O5b (Graham 1982:136). h, Yaxchilan, Hieroglyphic Stairway 3, step IV, B3a (Graham 1982:170). i, Yaxchilan, Lintel 3, J1 (Graham and von Euw 1977:17). j, Piedras Negras, Stela 3, F5a (Marcus 1976:fig.12). k, Piedras Negras, Throne 1, H'3 (Morley 1937–1938:fig.111). l, Piedras Negras, Lintel 3, V4). a, c–e from *Maya Hieroglyphic Writing: An Introduction*, by J. Eric S. Thompson, figs. 10 47, 11 24, 33, 34. New edition copyright © 1960, 1971 by the University of Oklahoma Press. I after a drawing by John Montgomery



one-third from 49 to 66; (ii) the average number of signs per syllable doubled. Not all of this homophony was universal. Many signs were limited to a single site or region. But the general pattern was one of increasing variation and artistic elaboration, rather than simplification (Grube 1994:179–184).

Another trend that can be seen over time is an intensification in the use of phonetic complements with logograms to spell polysyllabic words. Whereas during the Early Classic period, it was usually sufficient to spell such words with a single phonetic complement, over time some logograms came to be accompanied by two and even three phonetic complements until the word was written both logographically and syllabically. This can be seen in the different spellings of the word *?uniw*, the name for the fourteenth month of the solar year. Figure 43.11a shows the original spelling as **uniw(wa)**. Later it came to be spelled as **uniw(ni-wa)** (Fig. 43.11b). By AD 713, it had acquired a third complement and was spelled as **(u)uniw(ni-wa)** (Fig. 43.11c). Finally, there is a slightly earlier example of the complete replacement of the logogram by the syllabic spelling, **u-ni-w(a)** (Fig. 43.11d).

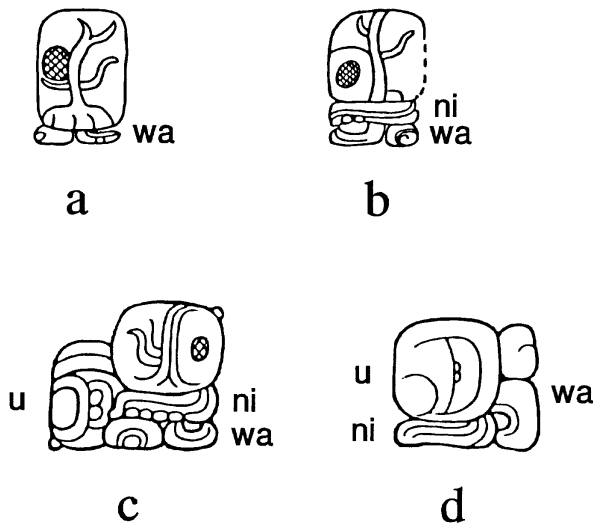
Nikolai Grube (1994:185) has pointed out that, even though phoneticism increased during the Late Classic period, it did not involve the replacement of logographic with syllabic writing. Both types of writing continued to exist side by side, increasing the possibilities for scribal virtuosity. He attributes the slow development of Mayan writing to the conservatism of a small scribal elite that had little interest in making writing more accessible to the masses.

**Figure 43.11** Alternative spellings of *ʔuniw*

(a, Palenque, Sun Tablet, H2. b, Seibal, Hieroglyphic Stairway, K1a. c, Dos Pilas, Stela 8, 113.

d, Yaxchilan, Hieroglyphic Stairway 3, step 1, D1a). (Graham 1982:166). a and b from *Maya*

*Hieroglyphic Writing: An Introduction*, by J. Eric S. Thompson, figs. 18, 23, 27. New edition copyright © 1960, 1971 by the University of Oklahoma Press. c after a drawing by Ian Graham



## 2.3 Origins of Mayan writing

The Maya were not the first people in Mesoamerica to use writing, and their script contains evidence of borrowing from earlier scripts that were invented in the region that lies to the west of their highland homeland. Two scripts emerged during the Late Preclassic period, one for an early form of Zapotecan (Whittaker 1992) and the other for an early form of Zoquean (see Ch. 44, §2). The former, and earlier, of the scripts was used in what is today the state of Oaxaca from c. 500 BC until c. AD 950 (Whittaker 1992:6). The Epi-Olmec script of the Isthmus of Tehuantepec first appeared in c. 150 BC and lasted only until c. AD 450 (Justeson and Kaufman 1993:1703). Thus, both scripts were contemporaneous with Mayan writing during some portion of their existence.

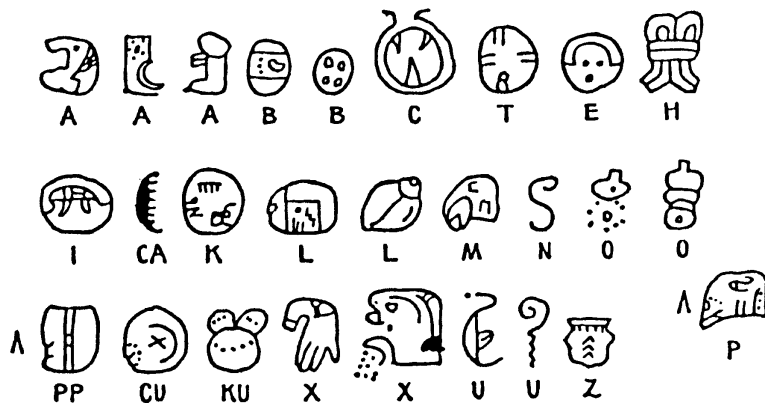
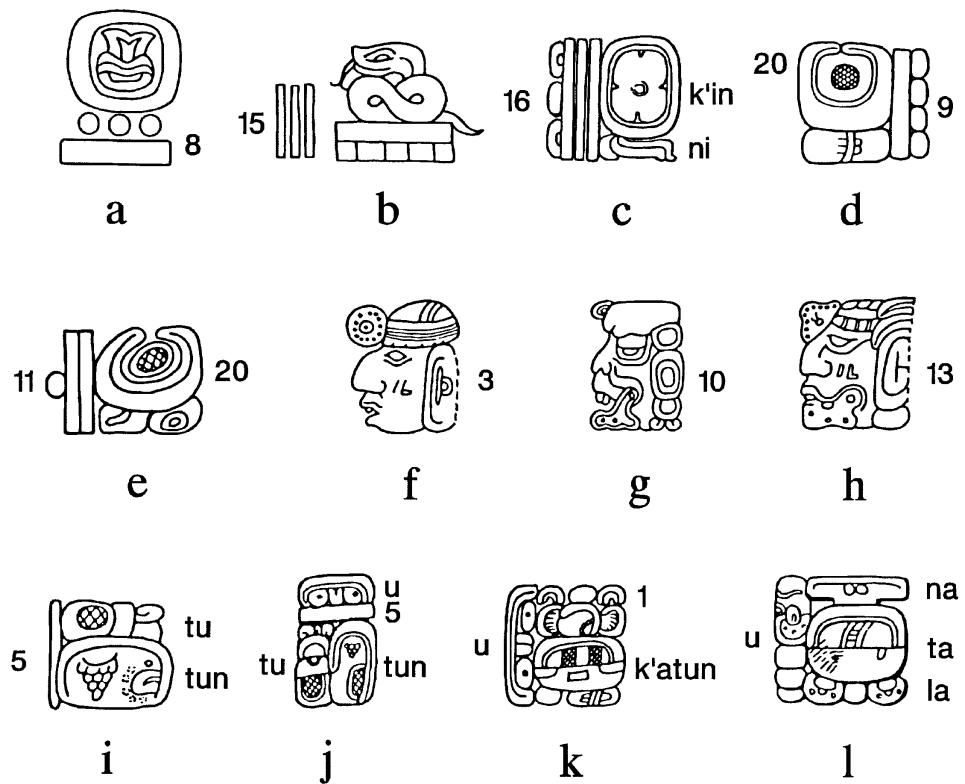
One feature shared by the Zapotecan and Epi-Olmec scripts was the use of a quinary notation for numbers, with 1 represented by a dot and 5 by a bar. The number 2 was written as two dots, 3 by three dots, and 4 by four dots. For numbers between 5 and 10, a single bar was combined with one to four dots. Two bars were used for 10, two bars and one dot for 11, three bars for 15, and so on. In the Zapotecan writing system, the bar-and-dot numbers were suffixed to main signs, following the order of nouns and their quantifiers in the spoken language (Fig. 43.12a). The reverse was true in Epi-Olmec writing, where numbers were prefixed to main signs (Fig. 43.12b). The Maya used the Epi-Olmec convention for bar-and-dot numbers (Fig. 43.12c), because their languages placed numbers before, not after, the nouns that they quantified, with one exception: in the lunar notations that follow Long Count dates, the bar-and-dot number representing 9 or 10 is postfixed or subfixed to the main sign in the collocation that refers to the length of the lunar month (the main sign is the glyph for 20; Fig. 43.12d). This convention must have been borrowed from the Zapotecan script.

The Long Count and the positional notation for recording it seems to have been invented by the Epi-Olmec people, and it diffused from there into Mayan writing. None of the other peoples of Mesoamerica had the Long Count.

## 2.4 Decipherment of Mayan writing

A gap of more than three hundred years separates the last practitioners of Mayan writing from the first serious efforts to decipher the script in the late nineteenth century. The closest

**Figure 43.12** Numbers (a, Monte Alban, Stela 12 (Marcus 1976:fig.3). b, La Mojarra, Stela 1 (Capitaine 1988:15). c, Palenque, Foliated Cross Tablet, D17 (Maudsley 1889–1902:IV, plate 82). d, Piedras Negras, Stela 3, A7. e, Balakbal, Stela 5, D10. f, Palenque, Sun Tablet, A6. g, Quirigua, F, C8b. h, Palenque, Cross Tablet, A5. i, Copan, Temple 11, north door, west panel. j, Piedras Negras, Stela 3, F4. k, Piedras Negras, Lintel 3, G1. l, Palenque, Inscriptions Temple, east panel, S9. (Maudsley 1889–1902: IV, plate 60). e-i and k from *Maya Hieroglyphic Writing: An Introduction*, by J. Eric S. Thompson, figs. 4 16, 24 12, 60, 25 11, 33 21, 29. New edition copyright © 1960, 1971 by the University of Oklahoma Press. d and j after a drawing by John Montgomery



**Figure 43.13** Landa's "alphabet". After Tozzer (1941:170)

thing to a Rosetta stone Mayan epigraphers have had to work with is the putative alphabet recorded by the Franciscan priest, Diego de Landa, in the middle of the sixteenth century (see Fig. 43.13).

For the most part, the signs elicited by Landa represented the closest pronunciation equivalents of the names of the letters of the Spanish alphabet (*a*, *be*, *ce*, etc.), which, of course, have a syllabic structure. However, Landa did not include signs for the Spanish letters, *d*, *f*, and *g*, which were not part of the Mayan phonemic inventory, and he included signs for the glottalized consonants, *k'* (written as *k*) and *p'* (written as *pp*), which do not

occur in Spanish. Therefore, it seems that he was not simply matching Mayan glyphs to Spanish letters (Durbin 1969). In eliciting different signs for *ca* (= **ka**) and *cu* (= **ku**) (and for *k* (= **k'a**) and *ku* (= **k'u**)), Landa intended only to mark the distinction between *c* and *q* in the Spanish alphabet, but in so doing he was providing a clue to the syllabic nature of a significant portion of the Mayan script. Limited as it was, Landa's "alphabet," together with his hieroglyphic spellings of the names of the twenty days and the nineteen months and a few other Mayan words, have been the key to hieroglyphic decipherment. The brilliant insights of Yuri Knorosov (1963) in the 1950s and the discoveries of more recent scholars (e.g., Lounsbury 1973; Fox and Justeson 1984; Bricker 1986; Stuart 1987) have all taken as their point of departure Landa's efforts to relate Mayan hieroglyphs to Spanish letters.

### 3. PHONOLOGY

#### 3.1 Consonants

Nineteen consonant phonemes can be distinguished in Mayan:

Table 43.1 The consonantal phonemes of Mayan						
Manner of articulation	Place of articulation					
	Bilabial	Alveolar	Palato-alveolar	Palatal	Velar	Glottal
Stop						
Voiceless	p	t			k	ʔ
Glottalized	p'	(t')			k'	
Voiced	b'					
Affricate						
Voiceless			ts	č		
Glottalized			ts'	č'		
Fricative						
Voiceless		s		š		h
Nasal	m	n				
Liquid		l				
Glide	w			y		

One of them, /p'/, is attested only in Landa's "alphabet" (as *pp* in Fig. 43.13), which is a very late source. However, the contrast between /b'/ and /p'/ is an innovation shared by Greater Tzeltalan and Yucatecan (Kaufman and Norman 1984:85), suggesting that the absence of /p'/ in hieroglyphic texts is probably accidental. The glottal stop is not overtly represented in the script, but the epenthetic /y/ that usually replaces it in ʔ-initial noun and verb roots when they are inflected with third-person pronominal clitics (see §3.4.1 and Fig. 43.7g and h) is demonstration that it was part of the phonemic inventory of the language, even though the script does not record it.

It is likely that Mayan also distinguished between /t/ and /t'/, as Greater Tzeltalan and Yucatecan have this contrast (Kaufman and Norman 1984:Table 4), but there is no evidence of /t'/ in the script – apparently another accidental gap.

### 3.2 Vowels

Five vowel phonemes can be distinguished in the hieroglyphic script:

#### (1) Mayan vowels

	Front	Central	Back
High	i		u
Mid		e	o
Low		a	

Greater Tzeltalan also had only five vowel qualities, but distinguished between long and short vowels. On the other hand, the contrast between long and short vowels was not retained in Proto-Cholan, except for *\*ā* and *\*a*. The long vowel *\*ā* became *\*a* while short *\*a* became *\*ə*. Thus, there are six vowels, not five, in Proto-Cholan (see Kaufman and Norman 1984:85). Proto-Yucatecan had only five vowel qualities and distinguished between long and short vowels.

The hieroglyphic script contains no evidence of more than five vowels. In fact, it used graphemes representing **Ca** syllables indiscriminately for spelling both *\*a*-medial and *\*ə*-medial roots. For example, a grapheme representing a gopher (Greater Tzeltalan *b'ah*) was used in spelling both Cholan *u-b'ə* “himself” (Fig. 43.14a) and Tzeltalan *u-b'ah* “he was going” (Fig. 43.16a). Similarly, the **ta** grapheme in Figure 43.8 was used as a syllable for spelling both *təl* “come” (Fig. 43.12l) and *tah* “torch” (Fig. 43.14i) in Proto-Cholan. The lack of a sixth vowel in the script suggests that Mayan distinguished between long and short vowels, but there is no direct evidence for such a contrast in the script.

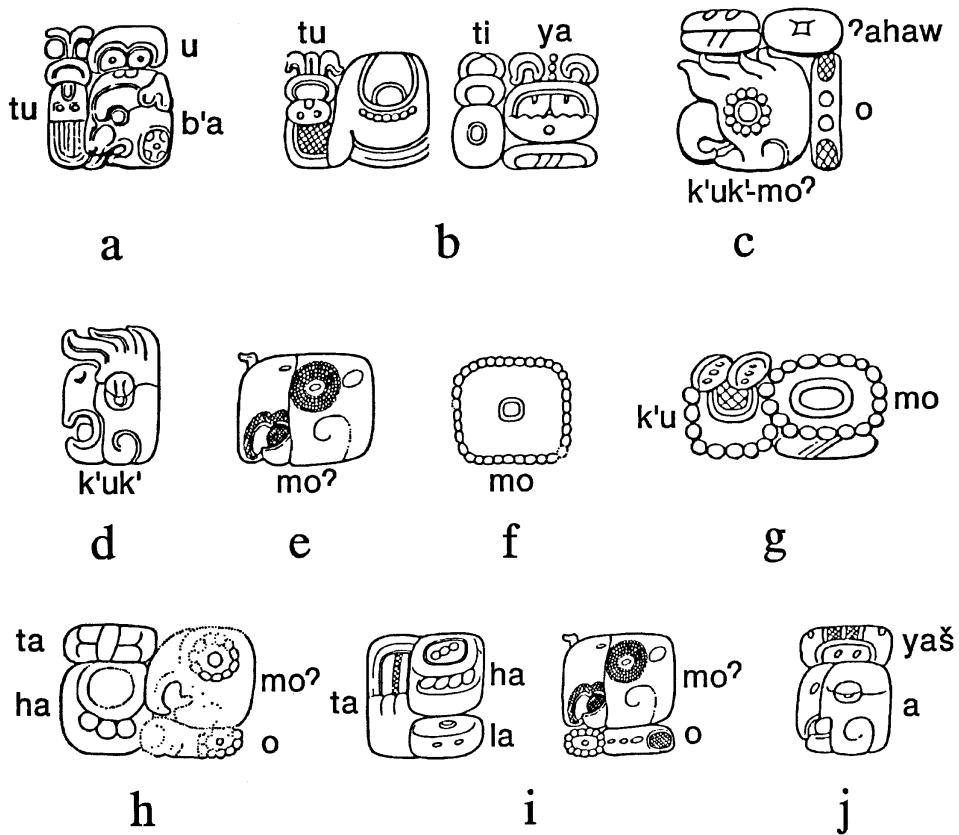
Houston, Stuart, and Robertson (1998) have argued that roots containing short vowels, CVC or CVCVC, are usually represented by synharmonic spellings, in which the inserted (but silent, i.e., purely orthographic) vowel in the last syllable or the phonetic complement echoes the vowel in the root (e.g., **la-k[a]** = *lak* “plate”; **k'u-k'[u]** = *k'uk'* “quetzal”; **k'an[na]** = *k'an* “yellow”), whereas roots with more complex vowels, either CV:C, CVCV:C, CV?C, or CVhC, are usually represented by disharmonic spellings, where the inserted vowel does not echo the vowel in the root (e.g., **b'a-k[i]** = *b'a:k* “bone”; **otot[ti]** = *?oto:t* “home”; **a-k[u]** = *?ahk* “turtle”). Their data set does not show a statistically significant pattern of synharmonic spellings for roots with short vowels nor of disharmonic spellings for roots with complex vowels, except for the neutral vowel /a/. Therefore, at present, there is not even indirect evidence for a general contrast between long and short vowels in the hieroglyphic script.

### 3.3 Syllable structure and phonotactic constraints

Mayan morphemes can have the following syllabic shapes: CVC, CV, VC, and V. Of these, CVC is by far the most common type in terms of either lexical or textual frequency. Many noun roots and all verb roots have this shape, as do a few inflectional suffixes. Examples of CVC roots include *k'in* “day,” *?al* “woman’s child,” *k'an* “yellow,” and *tun* “stone, 360-day year.” CVC suffixes are represented by *-lah* (positional) and *-lel* (abstractive). Most suffixes, such as *-ah* (thematic), *-Vw* (transitive), and *-il* (nominal), have a VC shape. CV is documented by the reflexive base *-b'a* and the preposition *ti* (or *ta*) and V by the third-person bound pronoun *u-*.

The language does not permit sequences of vowels or sequences of consonants within words, so that when such representations occur, phonological processes are applied to eliminate them. This is the cause of the allomorphic variation in the form of third-person





**Figure 43.14** Contraction and cluster reduction  
 (a, Palenque, Cross Tablet, S16 (Maudsley 1889–1902:IV, plate 77). b, Tikal, Temple I, Lintel 3, D3–C4 (Maudsley 1889–1902:III, plate 74). c, Copan, Altar Q, A3 (Maudsley 1889–1902:I, plate 93). d, Palenque, Sun Tablet, O9 (Maudsley 1889–1902:IV, plate 89). e, Machaquila, Structure 4, V3 (Graham 1967:fig. 39). f, Uaxactun, Stela 13, A5b (Graham 1986:163). g, Copan, Temple 11, step, glyph 22 (Maudsley 1889–1902:I, plate 8). h, Yaxchilan, Hieroglyphic Stairway 5, glyph 84 (Graham 1982:179). i, Machaquila, Structure 4, V2–V3 (Graham 1967:fig. 39). j, Tikal, Temple IV, Lintel 3, B4 (Jones and Satterthwaite 1982:fig. 74)

pronoun, appearing as *u-*, *uy-*, or *y-*, depending on whether the following noun or verb begins with a glottal stop or some other consonant. However, because the hieroglyphic script does not record glottal stops, vowel sequences appear in constructions involving vowel-final followed by glottal-stop-initial morphemes (e.g., Fig. 43.10j and k).

All the documented consonants in Mayan can begin and end syllables. In CVC syllables, however, there are few restrictions on which consonants can co-occur in initial and final position. If the first consonant in such a syllable is a glottalized stop or affricate, its plain counterpart cannot appear at the end of that syllable, and vice versa (e.g., *\*k'–k*, *\*k–k'*). Affricates also exemplify a principle of consonant harmony, a syllable-conditioned process that prevents them from co-occurring in the same syllable if they do not share the same point of articulation (*\*ts–č*, *\*ts–č'*, *\*ts'–č*, *\*ts'–č'*, *\*č–ts*, *\*č–ts'*, *\*č'–ts*, *\*č'–ts'*).

### 3.4 Morphophonemic processes

The following morphophonemic processes can be documented in Mayan: external sandhi, contraction, and cluster reduction.

#### 3.4.1 External sandhi

In Yucatecan and Greater Tzeltalan, most *ʔ*-initial roots have sandhi forms in which */ʔ/* is replaced by */y/* after the third-person pronoun *u-*. This process is reflected in the syllabic



spellings of *y-unen* “his, her child,” *y-al* “her child,” and *y-ahaw* “his lord, ruler” (Fig. 43.7d–h). The contracted form, *y-*, is much more common than *uy-* in Mayan texts.

### 3.4.2 Contraction

Contraction also occurs when the bound pronoun *u-* follows a preposition (either *ta* or *ti*). In Figure 43.14a, the dative reflexive construction *t-u-b'a* represents a contraction of *ti* (or *ta*) *u-b'a* “by himself.” This type of contraction is limited to contexts with the *u-* or *uy-* allomorphs of the bound pronoun. When the preposition precedes *y-*, it is spelled as *ta* or *ti* (e.g., Fig. 43.14b).

### 3.4.3 Consonant cluster reduction (sandhi)

Mayan sometimes eliminates consonant clusters across word boundaries (on the prohibition of word-internal clusters, see §3.3). Figure 43.14c contains an example of the personal name, *k'uk'-mo?* “quetzal-macaw,” in which the logogram for *mo?* (Fig. 43.14e) is infixed in the logogram for *k'uk'* (Fig. 43.14d), and the phonetic complement refers to the vowel in *mo?*. Note that the zoomorphic head in Figure 43.14c has both the distinctive feathers on the head of the quetzal in Figure 43.14d and the characteristic eye of the macaw in Figure 43.14e. The consonant cluster */-k'm-/* is eliminated in Figure 43.14g, where the name is spelled **k'u-mo(o)** (using the symbolic variant of **mo**; Fig. 43.14f). In Figure 43.14i, the name Torch-Macaw (*tahal-mo?*) is written in full; in 14h it is abbreviated to *taha-mo?* or, perhaps, *tah-mo?*, thereby eliminating the consonant cluster */-lm-/* (cf. Fig. 43.14i). Finally, *yaš-ha?* “green water,” the name of a large lake in northern Guatemala, has been reduced to *yaš-a?* in Figure 43.14j (Stuart 1985). The motivation for this abbreviated spelling may have been to eliminate the consonant cluster */-šh-/*.

## 3.5 Diachronic developments

The contribution of highland languages can be ruled out by the absence of a distinction between *\*k* and *\*q* (and *\*k'* and *\*q'*) in the script. This distinction was a characteristic of Proto-Mayan that has been preserved in Eastern Mayan and Kanjobalan, but was lost in Yucatecan and Greater Tzeltalan (Kaufman and Norman 1984:83). In this merger, *\*q* shifted to *\*k* and *\*q'* to *\*k'*. This change is reflected in the Mayan script, where **ka**, **ki**, and **ku** serve as complements and syllables in spellings that are cognate with highland Mayan words containing either *\*k* or *\*q*. For example, **ka** complements */k/* in the logogram for *kan* “snake” (Fig. 43.5i), which is cognate with Proto-Mayan *\*kān*; it also provides the */k/* in the syllabic spelling of *muk-ah* (Fig. 43.7i), which has a root that is cognate with Proto-Mayan *\*muq* “bury” (Fox and Justeson n.d.:20).

The shift of Proto-Mayan back velars to front velars in Greater Tzeltalan apparently triggered a concomitant forward shift of front velars to the affricates *č* and *č'*. “In Greater Tzeltalan, all instances of Proto-Mayan *\*k* and *\*k'* undergo this shift, except where the shift is blocked by particular phonological environments” (Kaufman and Norman 1984:83–84). The best evidence for this change in Mayan is the grapheme for the syllable **či** (Fig. 43.8). When enclosed by a cartouche it represents the seventh day of the Maya week, which corresponds to days named Deer in other Mesoamerican calendars: *\*cih* was the word for “deer” in Proto-Cholan, compare Proto-Mayan *\*kehx* (Kaufman and Norman 1984:118).

Yucatecan did not undergo the shift of front velars to affricates. Therefore, it is sometimes possible to identify Yucatecan spellings containing /k, k'/ instead of /č, č'/ . A case in point is the use of the phonetic complement **ka** in the collocation shown in Figure 43.5i, which indicates that the main sign refers to the Yucatecan spelling of *kān* “snake,” not its Cholan cognate, *čan*.

A phonological change that affected Yucatecan, but not Greater Tzeltalan, was the shift of \**t* to \**č* in a number of words (Fox and Justeson n.d.). As a result of this change, the Yucatecan word for “house” became \**čotoč*, whereas the Proto-Cholan word with this meaning remained \**totot* (Kaufman and Norman 1984:127). Although there are numerous examples of the *otot* spelling in hieroglyphic texts, the change to *otoč* cannot be documented before AD 950, so it cannot be used for distinguishing among languages using the script during the Early Classic period.

## 4. MORPHOLOGY

### 4.1 Word structure

The core of the Mayan word is the root, which is usually monosyllabic and composed of a consonant, a vowel, and a second consonant. Polysyllabic roots have a CVCVC structure and are limited to nouns. Inflectional and derivational processes are signaled by prefixing or suffixing grammatical morphemes with the following shapes to the root: V, VC, CV, and CVC. The language can be characterized as belonging to the agglutinating type because morpheme boundaries in word stems are clear, and words are easily segmented into their constituent morphemes.

Seven root classes have been identified in Mayan: nouns, adjectives, transitive verbs, intransitive verbs, positionals, numerals, and particles. The classification of transitives, intransitives, and positionals as separate form-classes is a characteristic that Mayan shares with Greater Tzeltalan and Yucatecan.

### 4.2 Nominal morphology

#### 4.2.1 Noun uses

Nouns occur in four morphological environments in Mayan, though the language, like Greater Tzeltalan and Yucatecan, does not inflect nouns for case. There does occur, however, a distinct, but limited, marking of possession; see §4.2.2.

1. In some contexts, they appear without affixes, indicating that they are neither possessed, nor quantified, nor marked for gender: for example, *kakaw* “chocolate”; *tun* “360-day year”; *čahaw* “lord, ruler” (see Figs. 43.7a, 43.9d, and 43.10).
2. In others, they are marked for possession, with the possessive pronominal prefixes *u-*, *uy-*, or *y-* (see §4.2.3): thus, *u-kan* “his captor”; *uy-ahaw* “his lord, ruler”; *y-al* “her child” (see Figs. 43.7e, f, h, and 43.17e). For the possessive “declension” see §4.2.2.
3. Nouns can also be quantified in compound expressions with prefixed numerals: for example, *waklahun-k'in* “16 days”; *ho-tun* “five 360-day years”; and *wuklahun-winal* “seventeen 20-day periods” (see Figs. 43.6b and 43.12c, i).
4. The agentive prefixes, *ah-* (male) and *naʔ-* (female) mark some nouns for gender. In Figure 43.9h, the *a(h)-* prefix derives an agentive noun, *ah-ʔuk* “Mr. Uk,” from a

#### 4.3.3.1 Transitive verbs

Active root transitives are marked by the suffix *-Vw*. Derived transitives take a suffix *-ah*, which resembles the thematic suffix that is obligatory with derived transitives in the Eastern Cholan languages (Cholti and Chorti; see Kaufman and Norman 1984:98). The third-person clitic pronouns, *u-*, *uy-*, and *y-*, mark agreement with the subjects of transitive verbs, and their objects are cross-referenced on the verb with the third-person suffix, *-Ø*, which, of course, has no graphemic representation.

Mayan examples of root transitives with third-person subjects and objects include *u-čuk-uw-Ø* “he seized it” (Fig. 43.15e) and *y-ak-aw-Ø* “he offered it” (Fig. 43.15d). Derived transitives with third-person subjects and objects are illustrated by *y-il-ah-Ø* “he saw it” (Fig. 43.15a and b) and *y-al-ah-Ø* “he said it” (Fig. 43.15c).

Passive stems were derived from root transitives by suffixing *-ah* to the root (see, e.g., the syllabic spelling of *čuk-ah* in Fig. 43.15f). The rules for inflection are described in §4.3.3.2.

#### 4.3.3.2 Intransitive verbs

During the Early Classic period, Mayan had an ergative verb system, in which the intransitive subject had the same form as the transitive object. The only examples in the glyphs are third-person intransitive subjects and third-person transitive objects, both zero (*Ø*) forms. The root took no stem suffix, and, because the subject pronoun was always *-Ø* in hieroglyphic examples, the inflected intransitive stem was identical to its root form, as in *?ut* “happen” and *?ut-Ø* “it happened” (see Fig. 43.15h).

Derived intransitives formed by passivizing root transitives were marked by the thematic suffix, *-ah*, a pattern that is found only in the Eastern Cholan languages (Cholti and Chorti; see Lacadena, forthcoming). Common examples of passives derived from root transitives occurring in hieroglyphic texts are provided by *muk-ah-Ø* “he was buried” and *čuk-ah-Ø* “he was captured” (see Figs. 43.7i, 43.15f).

Passives derived from nouns are also exemplified in Mayan. The passivizing suffix *-n* and the thematic suffix *-ah* follow the nominal root, as in *ts'ib'* “writing,” *ts'ib'-n-ah-Ø* “it was written” (see Fig. 43.15k). This pattern is also restricted to the Eastern Cholan languages (Lacadena, n.d.).

The above-described system of pronominal inflection underwent certain changes during the Late Classic period. By the middle of the eighth century AD, there were complement constructions such as *u-b'ah ti ?ak'ot* “he was going to dance” (see §5.2) in the inscriptions of three cities in the region, two in the west (Yaxchilan and Bonampak) and one in the east (Copan), in which the subject of the main verb, the root intransitive *b'ah* “go,” was marked by the ergative clitic *u-*, not the absolutive suffix *-Ø*, indicating a shift to a split-ergative pattern of pronominal inflection (see, e.g., Fig. 43.16a).

The form *u-b'ah* also contrasts with *b'ah-iy-Ø* (in identical contexts) during the same period at Copan (compare Fig. 43.16b and c; on the function of *-iy* see §5.3). There are also examples of *u-ts'ib'-n-ah-al* “it was being written” (Fig. 43.15l) contrasting in aspect and pronominal inflection with *ts'ib'-n-ah-Ø* “it was written” (Fig. 43.15k), suggesting that the ergative split corresponded to a distinction between imperfective and perfective aspects, with the former represented by *-al* and the latter by no suffix (i.e., *-Ø*). Clearly the pattern of split ergativity that has characterized the Cholan and Yucatecan languages since the sixteenth century must have had its roots in the Late Classic period. A third aspectual stem-suffix, *-om* “future,” occurs with the root intransitive, *?ut* “happen,” as *?ut-om-Ø* “it will happen” (Fig. 43.7n; Houston 1989) and with the absolutive form of the subject pronoun.

personal name. Figure 43.7k illustrates a syllabic spelling of *b'akab'* "sky bearer," a title often used by male rulers. When this title appears in the name phrases of women, it is usually written as *na?-b'akab'* "lady sky bearer" (Fig. 43.7l). The absence of *ah-* with *b'akab'* for men and the presence of *na?-* with *b'akab'* for women implies that females represented the marked category in Mayan.

A few words should be said concerning plural marking. In Greater Tzeltalan and Yucatecan, number can be marked by plural suffixes, but they are frequently not present. In Cholan and Yucatecan, the third-person plural suffix is *-ob'*, used on both nouns and verbs, though there appear to be no examples of that suffix in Mayan hieroglyphic texts.

#### 4.2.2 Possessive morphology

There is one "declension" in Mayan, for possession, which is represented only by the third-person singular forms in the script. Mayan nouns take either *-Ø* or *-il* when they are inflected for possession. Kinship terms comprise a semantic class that is marked by *-Ø* in possessive constructions: for example, *u-yum-Ø* "his father"; *y-unen-Ø* "his child," and *y-al-Ø* "her child" (see Fig. 43.7c–f). The form *u-kin-il* "the day" represents the inflection of *kin* "day" with suffix *-il* (see Fig. 43.7m).

#### 4.2.3 Pronouns

No examples of independent pronouns have been identified in Mayan writing. The only pronouns observable in the Early Classic script are the clitics and suffixes that refer to third-person subjects, objects, and possessors in their singular forms. As in both Greater Tzeltalan and Yucatecan, the marking of plural number was not obligatory in the third person (though it is for first and second) because it could be inferred from references to more than one individual or from other contextual clues.

The third-person transitive subject is usually represented by *(u)y-* before ?-initial roots and by *u-* before roots beginning with other consonants.

Direct objects are marked by suffixes, of which only that for the third person, which is a zero form (*-Ø*) in Greater Tzeltalan and Yucatecan, can be inferred in the Mayan script. The subject of intransitive verbs was identified with the direct object of transitive verbs during the Early Classic period (Houston 1997) and was therefore also *-Ø*. This ergative pattern of pronominal inflection began to change during the Late Classic, resulting in a split-ergative type of system based on aspect (see §4.3.3.2).

The third-person possessive pronoun also appears as *(u)y-* and *u-*, as in Proto-Cholan and Yucatecan. On the possessive construction, see also §4.2.1, 2 and §4.2.2.

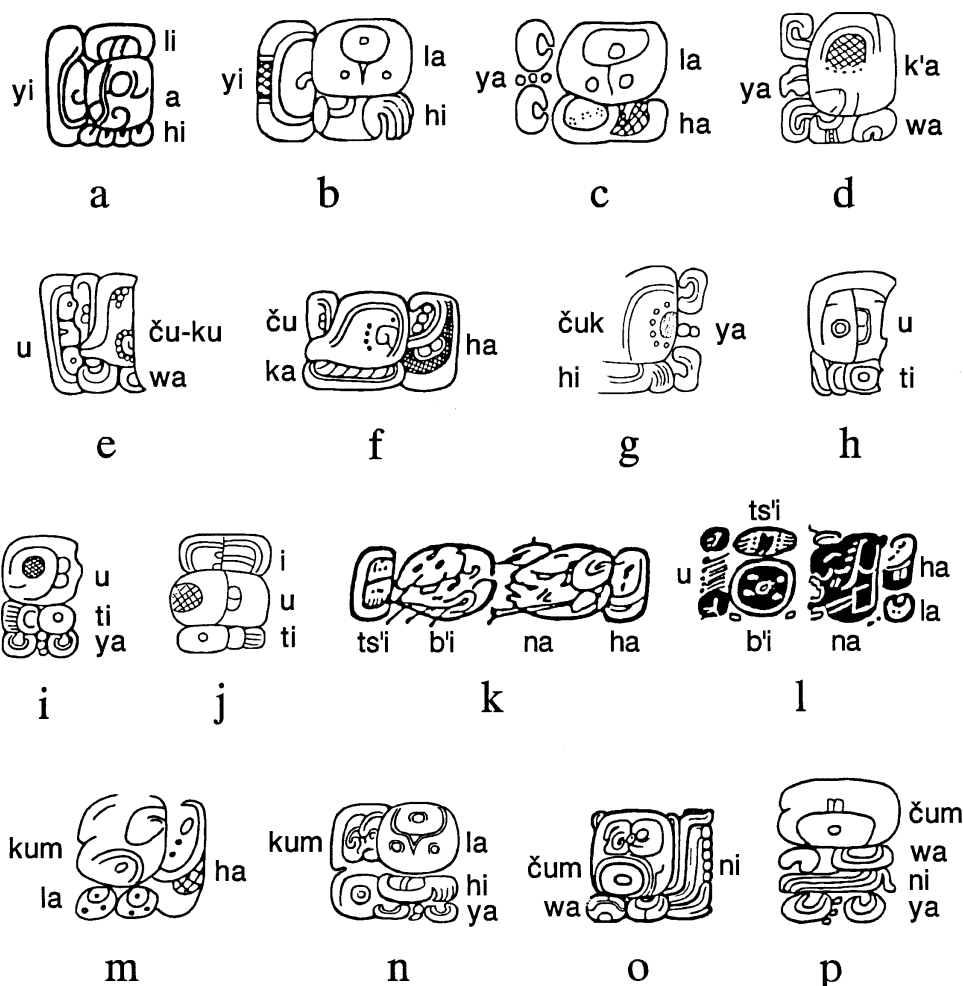
### 4.3 Verbal morphology

#### 4.3.1 Tense-aspect and mood

Most of the verbs that appear in Mayan hieroglyphic texts refer to events that are located in the past. Mayan used Calendar Round dates instead of tense or aspectual particles for placing events in time (cf. Bricker 1981:91–95) and two clitic particles for marking them as earlier (*-iy*) or later (*-i-*) in a sequence (e.g., Fig. 43.15i and j; see Wald 2000). On a perfective versus imperfective aspectual distinction, see §4.3.3.2.

Greater Tzeltalan and Yucatecan clearly have a grammatical category *mood* that includes the imperative and optative and is marked by suffixes. Transitive and intransitive verbs have

**Figure 43.15** Verbal inflection (a, Palenque, Inscriptions Temple, middle panel, H2 (Stuart 1987:fig. 36d). b, Piedras Negras, Lintel 3, J1. c, Piedras Negras, Altar support, A2 (Bricker 1986:fig. 148b). d, Palenque, Inscriptions Temple, middle panel, C8 (Robertson 1983:fig. 96). e, Piedras Negras, Throne 1, A'1. f, Yaxchilan, Hieroglyphic Stairway 3, step 1, tread, D1b (Graham 1982:166). g, Yaxchilan, Hieroglyphic Stairway 3, step 1, tread, A2 (Graham 1982:166). h, Palenque, Temple 18, jambs, D18a (Sáenz 1956:fig. 5). i, Piedras Negras, Stela 12, A16a. j, Yaxchilan, Lintel 31, 13b (Graham 1979:71). l, Yaxchilan, Lintel 31, 13b (Graham 1979:71). k, unprovenienced ceramic vessel. l, unprovenienced ceramic vessel (Grube 1991:figs. 5i and 9a). m, Deletaille panel, C 2 (Ringle 1985:fig. 2). n, Palenque, Foliated Cross Tablet, N7 (Maudslay 1889–1902:IV, plate 82). o, Dos Pilas, Stela 8, F14. p, Copan, Altar U, K2). (Maudslay 1889–1902:I, plate 98). b, e, and i after drawings by John Montgomery. k after a drawing by David Stuart. o after a drawing by Ian Graham



different mood suffixes. Such suffixes, however, are not represented in Mayan hieroglyphic texts.

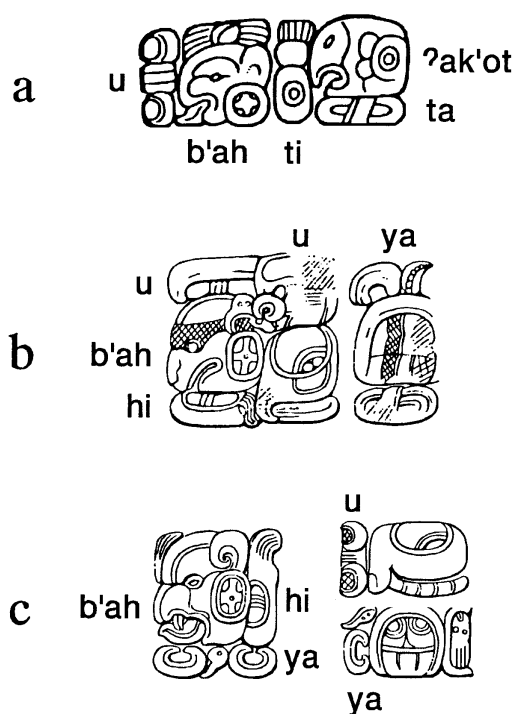
Later, there is some evidence of a future participial suffix (-*om*) in association with intransitive stems (e.g., Fig. 43.7n).

### 4.3.2 Voice

The script contains some information on active versus passive voice distinctions in Mayan during Early Classic times. On passivization see §§4.3.3.1 and 4.3.3.2.

### 4.3.3 Verb classes

Mayan has three verbal form classes: root transitives, root intransitives, and positionals.



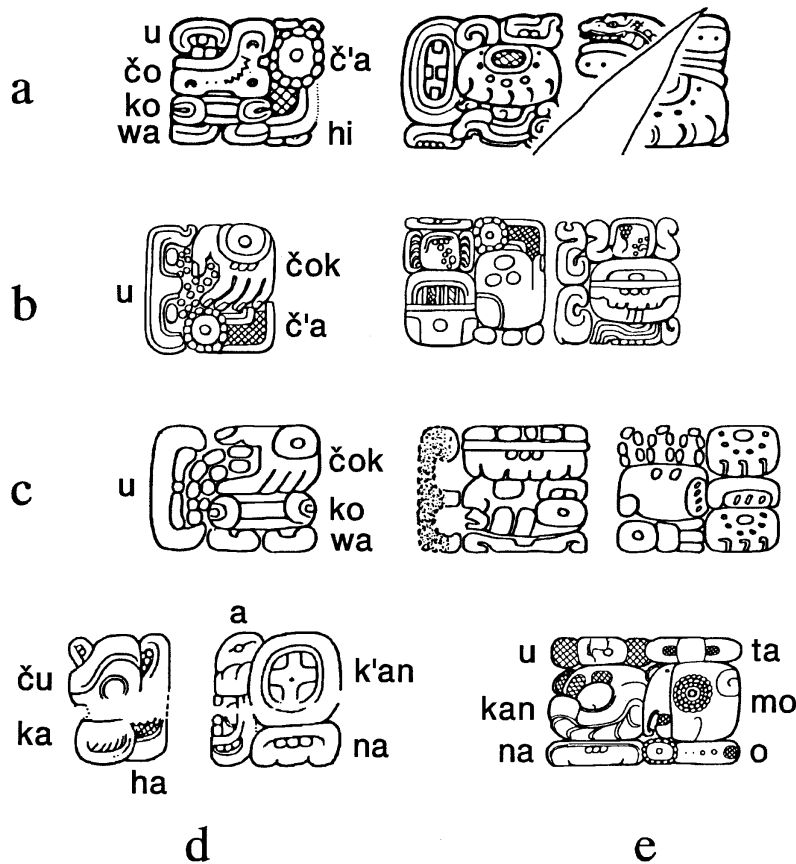
**Figure 43.16** Inflection of *b'ah* (a, Yaxchilan, Lintel 2, F–G (Graham and von Euw 1977:15). b, Copan, Stela 10, H3–G4a (Maudslay 1889–1902:I, plate 111). c, Copan, Stela 6, D8–D9). (Maudslay 1889–1902:I, plate 107)

#### 4.3.3.3 Positional verbs

Positional verbs can be distinguished from other verbs in terms of both formal and semantic criteria. They refer to physical states or positions, such as standing, sitting, kneeling, hanging, lying down, leaning, bending, and bowing, that human beings, animals, and inanimate objects can assume. Only one positional verb is known from the Early Classic period, *čum* (Cholan) or *kum* (Yucatecan), with the meaning “sit” (Fig. 43.5a–b), and it occurred with what are today the Yucatecan positional suffixes *-l-ah* (Fig. 43.5b). A new positional suffix, *-wan*, replaced *-l-ah* at many sites during the Late Classic period (Fig. 43.15o and p). Stems with these suffixes take the absolutive subject pronoun ( $-\emptyset$ ) in Mayan.

## 4.4 Derivational processes

Mayan derivational processes included not only the formation of agentive nouns described in §4.2.1, 4, but also the conversion of common nouns into abstract nouns and of transitive roots into instrumental nouns. The abstract *ʔahaw-lel* “rulership, reign” was derived from *ʔahaw* “lord, ruler” by suffixing *-lel* (frequently abbreviated as *-le*) to the noun root (Fig. 43.10j and k). The instrumental suffix *-ib'* was attached to the transitive root *ʔuč'* “drink,” and the resulting noun was inflected for possession as *y-uč'-ib'* “his cup” (in Fig. 43.7j; see MacLeod and Stross 1990). Gender-neutral agentive nouns were sometimes derived from nominal or verbal roots by suffixing *-om*, as in *č'ah-om* “caster of incense” (< *č'ah* “drop; incense”; Fig. 43.7o; Schele 1989). These are the only documented types of nominal derivation in Mayan writing.



**Figure 43.17** Word order (a, Dos Pilas 8, 15–16. b, Quirigua, Stela A, B16–B17 (Maudslay 1889–1902:II, plate 7). c, Quirigua, Structure 1, G–I (Morley 1937–1938: V, plate 175). d, Yaxchilan, Lintel 46, F3–G3 (Graham 1979:101). e, Machaquila, Structure 4, F1–F2) (Graham 1967:fig. 39). a after a drawing by Ian Graham

## 4.5 Compounds

Evidence for compounding in Mayan is limited to a few examples of noun incorporation involving the verb *čok* “throw, cast” and the noun *č’ah* “incense.” The verb can be represented by the syllables *čo* and *ko* (Fig. 43.17a) or by a logogram depicting a hand casting droplets or granules (Fig. 43.17b and c).

In syllabic spellings of *čok-ow*, the transitive suffix is produced by combining *ko* with *wa* (Fig. 43.17a). The same syllables are often suffixed to the *čok* logogram in logosyllabic spellings (Fig. 17c), or the logogram may appear only with *wa*. In Figure 17b, the *čok* logogram is followed by the syllable *č’a*, which is an abbreviation for *č’ah* “incense,” and there is neither a *ko* nor a *wa* suffix. This collocation cannot represent a transitive verb because there is no *-Vw* suffix. It may, however, be an example of a compound verb-stem with an incorporated direct object. If so, the verb is formally intransitive, and the use of the ergative pronoun *u-* makes it another example of ergative splitting.

## 4.6 Numerals

Although the bar-and-dot numbers used in writing Mayan had a quinary structure (see §2.3), the number words themselves did not. This can be seen in the head variants of the numbers, which used separate forms for the numbers from 1 through 12 (occasionally 13), but formed the numbers from 13 through 19 by combining the glyphs for 3 through 9



with the glyph for 10 (e.g., Fig. 43.12f–h). There was a separate glyph for 20, and numbers between 20 and 40 were constructed by prefixing the bar-and-dot numbers for 1 through 19 to this sign (e.g., Fig. 43.12e), except for the last collocation in lunar notations, which suffixed the bar-and-dot number to the sign for 20 (Fig. 43.12d). This means that, although the Mayan number system was fundamentally vigesimal in structure, the numbers below 20 had a decimal component. The numbers between 20 and 40 exemplify a principle of “overcounting” based on the previous score. Numbers above 39 usually had calendrical referents and were written in the positional notation employed for Long Count dates and Distance Numbers. For that purpose, there was also a sign for zero that served as a place holder (e.g., Fig. 43.6a and c).

The compounds formed by simply prefixing numbers to nouns are cardinal expressions. Ordinal numbers were formed by prefixing one of the allomorphs of the third-person pronoun to the compound (e.g., *u-ho-tun* “the fifth 360-day year”; Fig. 43.12j). However, there seem to have been three different ways of referring to “the first” in the script: (i) with a *u-* possessive clitic and a single dot for 1 (Fig. 43.12k); (ii) with *yax* replacing the dot (Fig. 43.9g); and (iii) with *na* also replacing the dot (Fig. 43.12l). Alternative words for “first” in the Cholan and Yucatecan languages are *yaš* and *nah* (Schele 1990).

## 5. SYNTAX

### 5.1 Word order

The basic word orders in Mayan are Verb–Object–Subject (VOS) in transitive clauses and Verb–Subject (VS) in intransitive and positional clauses. An example of VOS order appears in Figure 43.17a, in which the verb, *u-čok-ow-Ø* “he was casting [it],” is followed by the direct object, *č’ah* “incense,” and two collocations that refer to the subject, a ruler of Dos Pilas. The VS order is exemplified by the passive clause, *čuk-ah-Ø ah-k’an* “Mr. Kan was captured,” in Figure 43.17d.

Mayan also has verbless, or equational, clauses composed of two nouns, the second of which is inflected for possession with one of the clitic pronouns, *u-*, *uy-*, or *y-*. A case in point is the epithet, *k’ak’ u-pakal* “fire [is] his shield,” shown in Figure 43.7p, where the possessed noun, *u-pakal* “his shield,” functions as a stative verb. There is no verb having the meaning “to be” in Mayan.

At the phrase level, nouns follow their modifying adjectives, and the possessor noun follows the noun that refers to the thing possessed. Phrases such as *?ik’-k’at* “black cross” and *čak-k’at* “red cross” (< *?ik’* “black,” *čak* “red,” and *k’at* “cross”), which referred to the second and third months of the 365-day year, illustrate the syntax of nouns qualified by adjectives (Fig. 43.5k and l). The form *u-kan tah-mo?* “Torch-Macaw’s captor” (lit., “his captor Torch-Macaw”) provides an example of a possessor phrase, in which the noun representing the thing possessed (*kan* “captor”) is marked by the clitic pronoun *u-* and precedes the noun for the possessor (*tah-mo?* “Torch-Macaw”; Fig. 43.17e). This word order for possessor phrases is common to most Mayan languages.

### 5.2 Coordinate and subordinate clauses

Mayan clauses typically begin with a Calendar Round date such as *3 ik’ 15 yax-k’in*, which is followed by the verb, the direct object (if there is one), and the subject. The clauses are often linked by Distance Numbers, which express the interval separating the first date from the second in terms of the number of days, 20-day “months,” 360-day “years,” and so forth, that



lie between them. A Calendar Round date may have several clauses associated with it. If the subjects of both clauses are identical, one of them may be deleted, either the one referring to the first verb or the one referring to the second.

Another kind of subject deletion occurs with respect to clauses associated with different Calendar Round dates, but sharing the same subject. In such cases, neither event is introduced by a Calendar Round date; rather, the Distance Number that refers to the interval between them directly precedes the verbs for both events, and the date for the later of the two events appears after the subject at the end of the second clause (Lounsbury 1980):

(2) Distance Number–Verb<sub>1</sub>–Verb<sub>2</sub>–Subject–Date

The function of this word order is to focus on the events, rather than the dates that anchor them in time (Josserand 1991). In such cases, the verb that refers to the later of the two events is marked with the clitic particle *i-* (Lounsbury 1980).

Evidence of subordination can be found in complement constructions, in which only the main verb is inflected for subject. The root intransitive, *b'ah* “go,” serves as the main verb in such contexts. It is inflected with the ergative pronoun *u-* and is followed by the complementizer *ti* “to” and a verbal noun such as *?ak'ot* “dance” (Josserand *et al.* 1985). An example of *u-b'ah ti ?ak'ot* “he was going to dance” (Grube 1992) appears in Figure 43.16a.

### 5.3 Clitics

Mayan employs three clitic particles – *u*, *i*, and *iy* – each with referential functions. The clitic *u* serves as the third-person ergative subject pronoun in transitive stems (and occasionally as the nominative subject pronoun in intransitive stems; e.g., Figs. 43.15e, 43.16a and b) and as the possessive pronoun in nominal stems (e.g., Figs. 43.7c and p and 43.17e). The particle *i* is a focus marker, highlighting or drawing attention to the event in the narrative that follows (Josserand 1991:14). And *iy* is a temporal deictic enclitic that refers to previously reported events (Wald 2000). Thus, *čuk-ah-iy* in Figure 43.15g can be translated as “after he was captured,” whereas *čuk-ah* in Figure 43.15f means only “he was captured.” Similarly, *?ut-iy* in Figure 43.15i can be glossed as “after it happened,” whereas *?ut* in Figure 43.15h can only mean “it happened.” The forms *kum-lah-iy* and *čum-wan-iy* (in Fig. 43.15m and o) and *kum-lah* and *čum-wan* (in Fig. 43.15n and p) express the same contrast between already reported and not previously mentioned accession events. On the other hand, *i-?ut* (in Fig. 43.15j) highlights the event that follows – “and then it happened” – contrasting with both *?ut* “it happened” (in Fig. 43.15h) and *?ut-iy* “after it happened” (in Fig. 43.15j; see Wald 2000).

## 6. LEXICON

### 6.1 The inherited element

Mayan lexemes represent a number of semantic domains that can be grouped into three broad categories: (i) the natural world, including terms for animals, plants, colors and directions, astronomical bodies, and meteorological phenomena; (ii) the supernatural world, with terms for gods and spirits and the rituals used in propitiating them; and (iii) the human world, including terms for social and political relationships.

There is no general term for animal, but the names for eight mammals are known: armadillo (*?ib'ač*), bat (*sots'*), deer (*čih*), dog (*tsul* and *ts'i?*), gopher (*b'ah*), jaguar (*b'alam*),

mouse (*č'oh*), and spider monkey (*maš*). Mayan words for birds include the cotinga (*yašun*), hawk (*?i?*), heron (*b'ak*), macaw (*mo?*), owl (*kuy*), screech owl (*muwan*), great horned owl (*?ikim*), quetzal (*k'uk'*), turkey (*?ulum*), ocellated turkey (*kuts*), and vulture (*k'uč*). Terms for reptiles refer to iguana (*huh*), snake (*kan*), tortoise (*mak*), and turtle (*?ak*). There are general terms for fish (*kay*) and crab (*b'aw*). Three arthropods are referred to in hieroglyphic texts: ant (*sinik*), bee (*kab'*), and scorpion (*sinan*).

Mayan documents preserve only a few terms for flora. The words for tree (*te?* and *če?*), leaf (*le?*), seed (*hinah*), and flower (*nik*) are known. The blossom of the maize plant is *hanab'* and of *Pseudobombax ellipticum* (HBK) Dugan is *k'uy-nik*. There are also terms for the gumbo-limbo tree (*Bursera simaruba* [L.] Sargent; *čikah*), the kapok tree (*Ceiba pentandra* [L.] Gaertn.; *yaš-te?*), and *Pithecellobium dulce* (Roxb.) Benth. (*ts'iw-te?*).

Five primary colors are recognized in Mayan: red (*čak*), black (*?ik'*), white (*sak*), yellow (*k'an*), and blue/green (*yaš*). The first four colors are associated with the four cardinal directions, the names of only three of which have been deciphered: east (*lak'in*), west (*oč-k'in* and *čik'in*), and north (*šaman*). There are also terms for earth (*kab'*), sky (*čan* or *kan*), day (*k'in*), night (*?ak'ab'*), sun (*k'in*), and Venus (*k'an*). Among words for natural features of the landscape and meteorological phenomena are the following: water (*ha?*), stone (*tun* or *tunič*), mountain (*wits*), flint (*tok'*), obsidian (*tah*), rain (*čak*), cloud (*muyal*), rainbow (*čel*), smoke (*b'uts'*), and fire (*k'ak'*).

Words associated with the supernatural world and religious concepts are god (*č'uh* or *k'uh*), demon (*kisin*), hell (*šib'ah*), and alter ego (*way*). The names of several gods and one goddess are known: *čak* (the rain god), *k'awil* (the god of lightning), *?itsamna* (the creator god), *?ahaw k'in* (the sun god), and *čak čel* (the goddess of childbirth).

Rituals involve the casting (*čok*) of incense (*č'ah*) into censers, dancing (*?ak'ot*), and autosacrifice by perforating (*b'ah*) the tongue (*?ak'*) or penis (*?at* or *ton*) with a pointed object. The gods are offered (*?ak'*) pieces of paper (*hun*) spattered with blood. Some offerings are made in elaborate painted or carved cylindrical vessels (*?uč'ib'*), others on plates (*lak*). There is also a ritual ballgame (*pits*). Hieroglyphic texts refer to two kinds of musical instruments that were used in rituals: the upright drum (*paš*) and the horizontal drum (*tunk'uy* or *tunk'ul*).

Some kinship terms have been identified in Mayan writing: woman's child (*?al*), wife (*?atan*), father (*yum*), maternal grandfather (*mam*), maternal grandmother (*mim*), older brother (*sukun*), and younger sibling (*?its'in*). The head of a lineage is known as the *hol pop* "head of the mat." A number of lineage names that have been attested in Mayan are still in use today in the Maya area (e.g., *b'alam*, *b'atun*, *kokom*, *kupul*, *haw*, *k'awil*, *nik*, and *?uk*).

The ruler of a city or polity is called *?ahaw*. His immediate subordinate, who governed a smaller community, is known as *sahal*. These men are frequently involved in warfare, and there are accordingly words for warrior (*b'ate?*), shield (*pakal*), capture (*čuk*), captor (*kan*), captive (*b'ak*), and die (*kim*). Other significant roles in Maya society include priest (*?ah-k'in* or *čak*), scribe (*?ah-ts'ib'*), sculptor (*?ah-pol*), and wiseman (*miyats* or *?its'at*). There are also words for writing and painting (*ts'ib'*), hieroglyph (*woh*), and paper or book (*hun*). Other intellectual achievements are related to mathematics and calendrics, the terms of which are listed in §1.3 and §4.6.

Finally, there are terms for buildings and their components: house (*?otot* or *?otoč* and *na*), lintel (*pakab'*), and sweatbath (*pib'-na*). There are also words for body parts: bone (*b'ak*), tooth (*koh*), hand (*k'ab'*), foot (*?ok*), fingernail or claw (*?ič'ak*), and penis (*?at* or *ton*). The many other words that were part of the inherited lexicon are shown in Figures 43.5–7, 43.9–12, and 43.14–17 and in the Plates in Davoust (1995).

## 6.2 Influence of other languages

Only a few loans from other languages have been documented in Mayan. Of these, Mixe-Zoquean has made the largest contribution, including words for chocolate (*kakaw*), child (*?unen*), dog (*?ok*), jaguar (*hiš*), incense (*pom*), and monkey (*čowen* or *čuwen*). Loans from Zapotecan seem to be limited to the day names, *b'en*, *lamat*, and *manik'*. There is one loan each from Totonac (*pak'* "plant") and Nahuatl (*kot* "eagle"; Justeson *et al.* 1985:21–28).

## 7. READING LIST

The most comprehensive and authoritative single work on ancient Maya cultural history is *The Ancient Maya* by Robert J. Sharer (1994). *Breaking the Maya Code* by Michael D. Coe (1992) is an engaging account of the history of decipherment. The methodology of decipherment is clearly presented in two influential publications, *Ten Phonetic Syllables* by David Stuart (1987) and *Classic Maya Place Names* by David Stuart and Stephen Houston (1994). *L'écriture maya et son déchiffrement* by Michel Davoust (1995) is the best single volume source on Maya epigraphy – encyclopedic, up to date, and profusely illustrated.

### Bibliography

- Beetz, C. and L. Satterthwaite. 1981. *The Monuments and Inscriptions of Caracol, Belize*. University Museum Monograph 45. Philadelphia: University of Pennsylvania.
- Bricker, H. and V. Bricker. 1983. "Classic Maya prediction of solar eclipses." *Current Anthropology* 24:1–23.
- Bricker, V. 1977. *Pronominal Inflection in the Mayan Languages*. Tulane University Middle American Research Institute, Occasional Papers, 1. New Orleans.
- . 1981. "The source of the ergative split in Yucatec Maya." *Journal of Mayan Linguistics* 2:83–127.
- . 1986. *A Grammar of Mayan Hieroglyphs*. Tulane University Middle American Research Institute Publication 56. New Orleans.
- Bricker, V. and H. Bricker. 1986. "The Mars table in the Dresden Codex." In E. Andrews (ed.), *Research and Reflections in Archaeology and History: Essays in Honor of Doris Stone*, pp. 51–80. Tulane University Middle American Research Institute Publication, 57. New Orleans.
- . 1988. "The seasonal table and related almanacs in the Dresden codex." *Archaeoastronomy* (Supplement to *Journal for the History of Astronomy* 19) 12:S1–S62.
- Campbell, L. and T. Kaufman. 1985. "Mayan linguistics: where are we now?" *Annual Review of Anthropology* 14:197–198.
- Capitaine, F. 1988. *La estela 1 de La Mojarra, Veracruz, México*. Research Reports on Ancient Maya Writing 16. Washington, DC: Center for Maya Research.
- Coe, M. 1973. "The Maya scribe and his world." New York: Grolier Club.
- . 1992. *Breaking the Maya Code*. London: Thames and Hudson.
- Davoust, M. 1995. *L'écriture maya et son déchiffrement*. Paris: CNRS Éditions.
- Durbin, M. 1969. "An interpretation of Bishop Diego de Landa's Maya alphabet." In *Philological and Documentary Studies*, vol. II, pp. 169–179. Tulane University Middle American Research Institute Publication, 12. New Orleans.
- Fox, J. and J. Justeson. n.d. "Hieroglyphic evidence for the languages of the lowland Maya." Stanford University (ms).
- . 1984. "Polyvalence in Mayan hieroglyphic writing." In Justeson and Campbell 1984, pp. 17–76.
- Graham, I. 1967. *Archaeological Explorations in El Peten, Guatemala*. Tulane University Middle American Research Institute Publication 33. New Orleans.
- . 1977. "Illustrations of lintels." In J. Bolles (ed.), *Las Monjas: A Major Pre-Mexican Architectural Complex at Chichén Itzá*, pp. 267–275. Norman: University of Oklahoma Press.

- . 1979. *Corpus of Maya Hieroglyphic Inscriptions*, vol. III, part 2: *Yaxchilan*. Cambridge, MA: Peabody Museum, Harvard University.
- . 1982. *Corpus of Maya Hieroglyphic Inscriptions*, vol. III, part 3: *Yaxchilan*. Cambridge, MA: Peabody Museum, Harvard University.
- . 1986. *Corpus of Maya Hieroglyphic Inscriptions*, vol. V, part 3: *Uaxactun*. Cambridge, MA: Peabody Museum, Harvard University.
- Graham, I. and G. Mobley. 1986. "Looters rob graves and history." *National Geographic* 169:453–461.
- Graham, I. and E. von Euw. 1977. *Corpus of Maya Hieroglyphic Inscriptions*, vol. III, part 1: *Yaxchilan*. Cambridge, MA: Peabody Museum, Harvard University.
- Grube, N. 1991. "An investigation of the primary standard sequence on Classic Maya ceramics." In M. Robertson and V. Fields (eds.), *Sixth Palenque Round Table*, 1986, pp. 223–232. Norman: University of Oklahoma Press.
- . 1992. "Classic Maya dance: Evidence from hieroglyphs and iconography." *Ancient Mesoamerica* 3:201–218.
- . 1994. "Observations on the history of Maya hieroglyphic writing." In M. Robertson and V. Fields (eds.), *Seventh Palenque Round Table*, 1989, pp. 177–186. San Francisco: Pre-Columbian Art Research Institute.
- Houston, S. 1989. *Maya Glyphs*. Berkeley: University of California Press.
- . 1997. "The shifting now: aspect, deixis, and narrative in Classic Maya texts." *American Anthropologist* 99:291–305.
- Houston, S., D. Stuart, and J. Robertson. 1998. "Disharmony in Maya hieroglyphic writing: linguistic change and continuity in Classic society." In A. Ciudad Ruiz, Y. Fernández Marquín, J. García Campillo, M. Iglesias Ponce de León, A. Lacadena García-Gallo, and L. Sanz Castro (eds.), *Anatomía de una civilización. Aproximaciones interdisciplinarias a la cultura maya*, pp. 275–296. Sociedad Española de Estudios Mayas 4. Madrid.
- Houston, S. and K. Taube. 1987. "'Name-tagging' in Classic Mayan script." *Mexicon* 9:38–41.
- Jones, C. and L. Satterthwaite. 1982. *The Monuments and Inscriptions of Tikal: The Carved Monuments*. University Museum Monograph 44. Tikal Report 33, part A. Philadelphia.
- Josserand, J. 1991. "The narrative structure of hieroglyphic texts at Palenque." In M. Robertson and V. Fields (eds.), *Sixth Palenque Round Table*, 1986, pp. 12–31. Norman: University of Oklahoma Press.
- Josserand, J., L. Schele, and N. Hopkins. 1985. "Linguistic data on Mayan inscriptions: the *ti* constructions." In M. Robertson and E. Benson (eds.), *Fourth Palenque Round Table*, 1980, pp. 87–102. San Francisco: Pre-Columbian Art Research Institute.
- Justeson, J. and L. Campbell (eds.). 1984. *Phoneticism in Mayan Hieroglyphic Writing*. Institute for Mesoamerican Studies, State University of New York. Publication 9. Albany.
- Justeson, J. and T. Kaufman. 1993. "A decipherment of Epi-Olmec hieroglyphic writing." *Science* 259:1703–1711.
- Justeson, J. and P. Mathews. 1990. "Evolutionary trends in Mesoamerican hieroglyphic writing." *Visible Language* 29:89–132.
- Justeson, J., W. Norman, L. Campbell, and T. Kaufman. 1985. *The Foreign Impact on Lowland Mayan Language and Script*. Tulane University Middle American Research Institute Publication 53. New Orleans.
- Kaufman, T. 1976. "Archaeological and linguistic correlations in Mayaland and associated areas of Meso-America." *World Archaeology* 8:101–118.
- Kaufman, T. and W. Norman. 1984. "An outline of Proto-Cholan phonology, morphology, and vocabulary." In Justeson and Campbell 1984, pp. 77–166.
- Knorosov, Y. 1963. *Pisménnoš' Indeitsev Maiia*. Moscow: Academy of Sciences of the USSR.
- Lacadena, A. n.d. *Passive Voice in Classic Maya Texts: CV-jC-aj and -n-aj Constructions*.
- Lounsbury, F. 1973. "On the derivation and reading of the 'ben-ich' prefix." In E. Benson (ed.), *Mesoamerican Writing Systems*, pp. 99–143. Washington, DC: Dumbarton Oaks.
- . 1980. "Some problems in the interpretation of the mythological portion of the hieroglyphic text of the Temple of the Cross at Palenque." In M. Robertson (ed.), *Third Palenque Round Table*, 1978, vol. V, part 2, pp. 99–115. Austin: University of Texas Press.
- MacLeod, B. 1984. "Cholan and Yucatec verb morphology and glyphic verbal affixes in the inscriptions." In Justeson and Campbell 1984, pp. 233–262.

- MacLeod, B. and B. Stross. 1990. "The wing-quincunx." *Journal of Mayan Linguistics* 7:14–32.
- Marcus, J. 1976. "The origins of Mesoamerican writing." *Annual Review of Anthropology* 5:35–67.
- Maudslay, A. 1889–1902. *Biologia Centrali-Americana: Archaeology* (1 vol. text, 4 vols. plates). London: R. H. Porter and Dulau.
- Morley, S. 1937–1938. *The Inscriptions of Peten* (5 vols.). Carnegie Institution of Washington Publication 437. Washington, DC.
- Proskouriakoff, T. 1960. "Historical implications of a pattern of dates at Piedras Negras, Guatemala." *American Antiquity* 25:454–475.
- Proskouriakoff, T. and J. Thompson. 1947. "Maya calendar round dates such as 9 Ahau 17 Mol." In *Notes on Middle American Archaeology and Ethnology* No. 79, pp. 143–150. Carnegie Institution of Washington. Washington, DC.
- Ringle, W. 1985. "Notes on two tablets of unknown provenance." In M. Robertson and V. Fields (eds.), *Fifth Palenque Round Table*, 1983, vol. VII, pp. 151–158. San Francisco: Pre-Columbian Art Research Institute.
- Robertson, M. 1983. *The Sculpture of Palenque*, vol. 1: *The Temple of the Inscriptions*. Princeton: Princeton University Press.
- Sáenz, C. 1956. *Exploraciones en la pirámide de la Cruz Foliada*. Instituto Nacional de Antropología e Historia, Informe 5. Mexico City.
- Schele, L. 1989. "A new glyph for 'five' on stela E." *Copán Note 53*. Copán Mosaics Project and the Instituto Hondureño de Antropología e Historia. Copán, Honduras.
- . 1990. "Ba as 'first' in Classic period titles." *Texas Notes on Precolumbian Art, Writing, and Culture* 5. Austin: University of Texas.
- Schele, L. and M. Miller. 1986. *The Blood of Kings: Dynasty and Ritual in Maya Art*. Fort Worth: Kimbell Art Museum.
- Sharer, R. 1994. *The Ancient Maya* (5th edition). Stanford: Stanford University Press.
- Stuart, D. 1985. *The Yaxha Emblem Glyph as yax-ha*. Research Reports on Ancient Maya Writing 1. Washington, DC: Center for Maya Research.
- . 1987. *Ten Phonetic Syllables*. Research Reports on Ancient Maya Writing 14. Washington, DC: Center for Maya Research.
- . 1988. "The Rio Azul cacao pot: epigraphic observations on the function of a Maya ceramic vessel." *Antiquity* 62:153–157.
- . 1990. "The decipherment of 'directional count glyphs' in Maya inscriptions." *Ancient Mesoamerica* 1:213–224.
- . 1993. "Breaking the code: Rabbit story." In G. Stuart and G. Stuart (eds.), *Lost Kingdoms of the Maya*, pp. 170–171. Washington, DC: National Geographic Society.
- Stuart D. and S. Houston. 1994. *Classic Maya Place Names*. Dumbarton Oaks Studies in Pre-Columbian Art and Archaeology, 33. Washington DC.
- Thompson, J. 1960. *Maya Hieroglyphic Writing: An Introduction*. Norman: University of Oklahoma Press.
- Thompson, J. 1971. *Maya Hieroglyphic Writing: An Introduction* (revised edition). Norman: University of Oklahoma Press.
- Tozzer, A. 1941. *Landa's Relación de las cosas de Yucatan*. Papers of the Peabody Museum of Archaeology and Ethnology, Harvard University, 18. Cambridge, MA.
- Wald, R. 2000. "Temporal deixis in Colonial Chontal and Maya hieroglyphic narrative." *Written Language and Literacy* 3:123–153.
- Whittaker, G. 1992. "The Zapotec writing system." In V. Bricker (ed.), *Supplement to the Handbook of Middle American Indians*, vol. V, pp. 5–19. Austin: University of Texas Press.

# Epi-Olmec

TERRENCE KAUFMAN AND JOHN JUSTESON

## 1. HISTORICAL AND CULTURAL CONTEXTS

The Epi-Olmec language is the most ancient attested member of the Mije-Sokean family (c. 300 BCE to at least 533 CE). This family and its internal relationships are presented in Figure 44.1. In the sixteenth century, this family occupied a continuous area of southern Mesoamerica, extending from southern Veracruz to the border of present-day Guatemala, and with no trace of other native languages apart from islands of Nawa. The region seems to have been exclusively Mije-Sokean until the invasion of these Nawas, sometime between 600 and 900 CE; they influenced the vocabularies of individual Mije-Sokean languages but not that of Proto-Sokean or Proto-Mijejan.

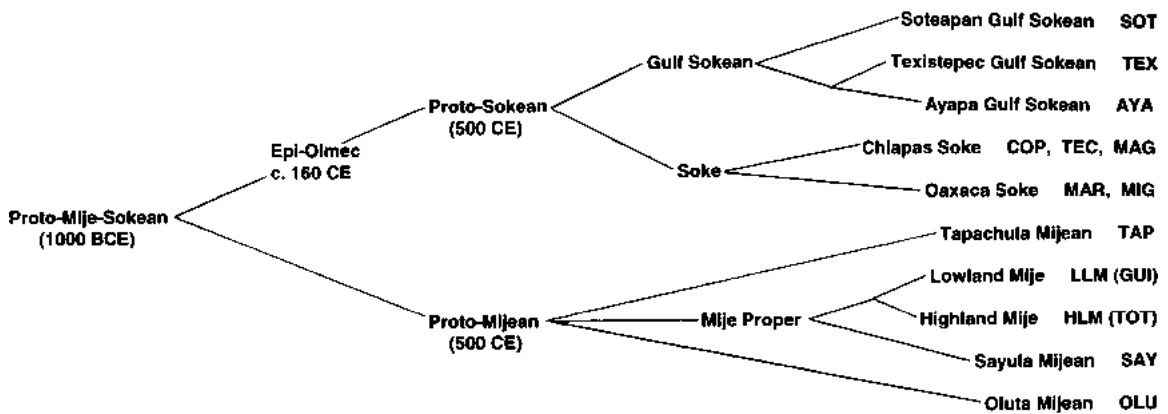
This region included the entire heartland of the Olmec civilization (1500–500 BCE), and the above circumstances provide a *prima facie* case that the Olmecs were Mije-Sokeans. Olmecs had widespread influence, diffusing innovations that would become distinctively Mesoamerican cultural characteristics. Evidence from associated linguistic diffusion confirms that at least some Olmecs spoke an early Mije-Sokean language (Campbell and Kaufman 1976); further analysis suggests to us that they were Sokean specifically.

Olmec civilization was originally defined by its distinctive art style, which developed *in situ* in the Olmec heartland, and whose classic form ended after the abandonment of La Venta (c. 500 BCE). The *Epi-Olmec* tradition is the Olmec tradition in its later manifestations. Later material remains from the region developed gradually from Olmec canons, a development observable especially at Olmec sites, like Tres Zapotes, that were occupied continuously from Olmec through Epi-Olmec times (Pool 2000).

Linguistic geography had already left little doubt that the Epi-Olmec population spoke Mije-Sokean at least until the breakup of Proto-Mijejan and/or Proto-Sokean – c. 500 CE according to glottochronology – when our decipherment of Epi-Olmec writing showed that in language, too, the Epi-Olmec tradition was a direct inheritance from the Olmecs. Only ten to twelve Epi-Olmec texts are now known to scholarship (see Table 44.1), and only seven have legible, diagnostically Epi-Olmec signs. Yet these texts spanned the greater part of Mije-Sokean territory. Among Mije-Sokean languages, only Mije lies outside the general area of Epi-Olmec writing.

At this writing, the Epi-Olmec language is known from just four legible Epi-Olmec texts. Several features of its morphology and syntax are specific, in Mesoamerica, to Mije-Sokean languages, and its vocabulary is specifically Sokean. The texts (except the shortest one) include phonological and/or grammatical features that Sokean lost by the Proto-Sokean stage and that today survive only in Mijejan (see §7). Two of these three pre-Proto-Sokean texts are dated, to 157 and 162 CE; since the fourth text is centuries older (c. 300 BCE),





**Figure 44.1** The Mije-Sokean language family. Parenthesized dates are Kaufman's current glottochronological estimates for subgroup diversification. The Mije-Sokean languages of Veracruz – those of Sayula, Oluta, Texistepec, and Soteapan – are popularly known as *Popoloca*, and Soteapan Gulf Sokean is specifically known as *Sierra Popoloca*. COP is Copainalá, MAG Magdalena (Francisco León), MAR Santa María Chimalapa, MIG San Miguel Chimalapa, TOT Totontepec, GUI San Juan Guichicovi.

it too must be pre-Proto-Sokean. The latest Epi-Olmec texts (468–533 CE) are almost totally illegible, so their language cannot be identified; if Sokean, as seems likely, they could be Proto-Sokean but hardly much later. Our discussion takes no account of a fifth text that reportedly came to light during or before the summer of 2002, because drawings and photographs of it became publicly available only as the present study was going to press.

## 2. WRITING SYSTEM

### 2.1 Decipherment

The Epi-Olmec language has only recently been recovered. Its script was deciphered by the authors in joint work, conducted largely from 1991 through 1994. Just four Epi-Olmec texts were legible enough to provide an empirical basis for establishing the pronunciations or meanings of its signs, or the rules for using them to represent the Epi-Olmec language. The decipherment was initially based only on the data from the two longest texts known at the time, La Mojarra Stela 1 (see Figure 44.2) and the Tuxtla Statuette (see Table 44.1), because available drawings of the other two relatively complete texts appeared to be unreliable. We eventually examined and redrew all of the Epi-Olmec texts; those not previously used were straightforwardly interpretable in terms of the previously established grammatical results and phonetic sign readings, providing independent evidence for the decipherment. The decipherment was supported by another independent test when a previously unseen column of text was discovered on the side of La Mojarra Stela 1 (Justeson and Kaufman 1997).

Summaries of our methods are provided elsewhere (Justeson and Kaufman 1993, 1996 [1992], 1997; Kaufman and Justeson 2001; Kelley 1993). Although cultural and chronological data and inferences played important roles, the decipherment hinged on an understanding of Mije-Sokean grammatical structure and vocabulary as previously worked out by Kaufman (1963); it was facilitated by Wichmann's (1995) expanded list of lexical reconstructions, produced using lexical data unavailable in 1963.

Many grammatical affixes were easily recognized, because of their high text frequency and because most of them were represented by CV syllables that corresponded to a single

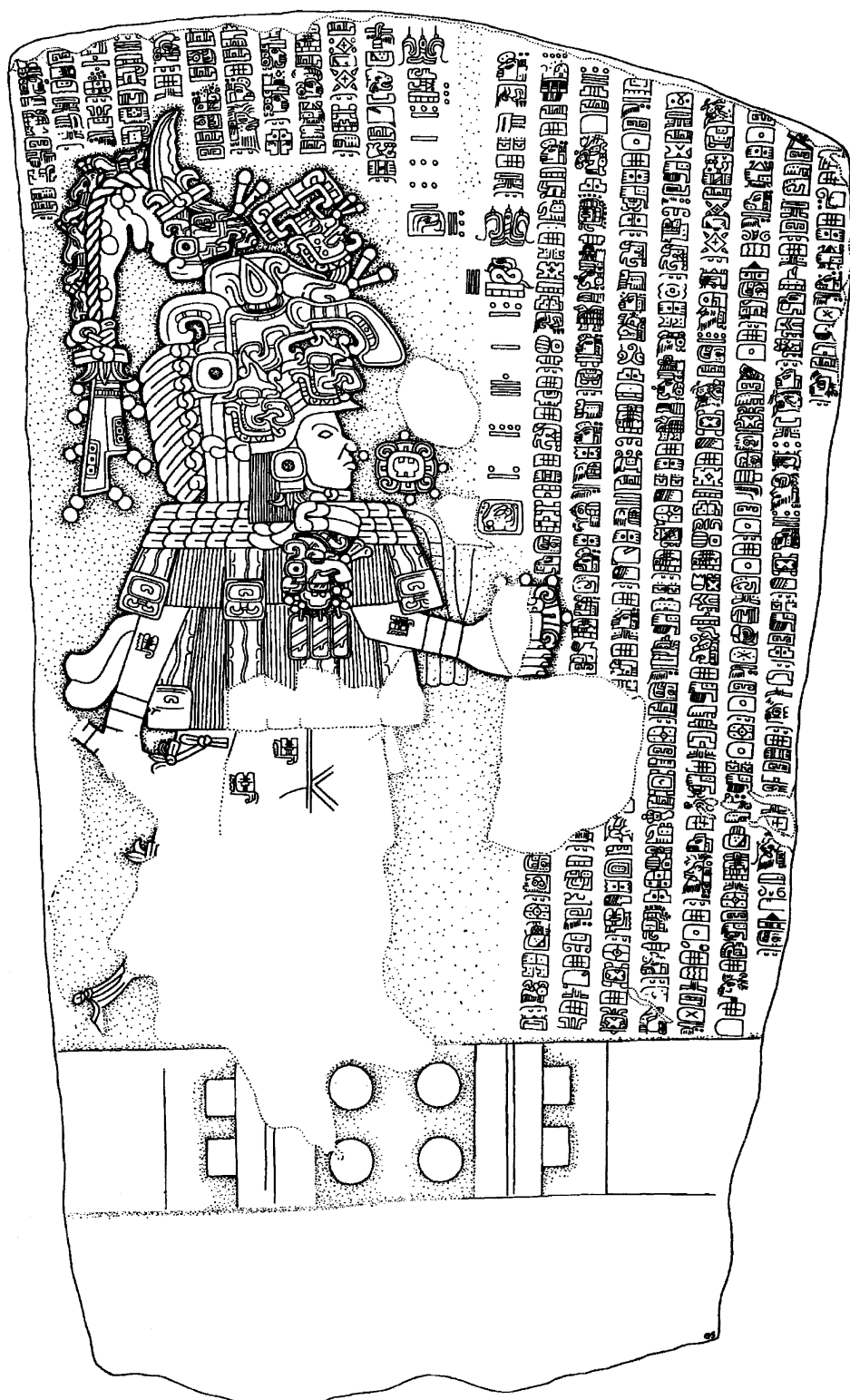


Figure 44.2 La Mojarra Stela 1. Drawing by George Stuart.



**Table 44.1 Sources and characteristics of Epi-Olmec and related texts**

	Abbr.	Date	Diagnostic Epi-Olmec signs	Total text length	Legible non-num. signs
<i>Chiapa de Corzo</i>					
<b>sherd</b>	CHP-sh	c. 300 BCE	6	16+	12
wall panel	CHP-2	36 BCE	(1)	9+	1
<i>Tres Zapotes</i>					
Stela C	3ZP-C	32 BCE	5–7	28	8–10
<i>La Mojarra</i>					
<b>Stela 1</b>	MOJ	157 CE	all	c. 544	490
<i>Tuxtla Mountains</i>					
<b>Tuxtla Statuette</b> 79	TUX	162 CE	all	87	79
<i>Cerro de las Mesas</i>					
Stela 5	MES-5	528 CE	0	c. 16	1
Stela 6	MES-6	468 CE	3–4	18	8
Stela 8	MES-8	533 CE	1	c. 40	3
Stela 15	MES-15	?468 CE	(1)	4	2
<i>provenience unknown</i>					
<b>O'Boyle mask</b>	OBM	unknown	16	27	27
Teotihuacan-style mask	TEO	??	24–25	104	99
<i>Alvarado</i>					
Stela 1	ALV	?	1–3	12–14	6
<i>El Sitio</i>					
celt	SIT	Late Precl	?	10–12	10–12
<i>Izapa</i>					
various		Late Precl	3		

Legible texts used for decipherment are boldfaced. The standard designation in the literature for the Chiapa de Corzo wall panel is the misnomer “Stela 2”; the label “CHP-2” is used to avoid confusion.

“Diagnostic Epi-Olmec signs” are signs occurring on La Mojarra Stela 1 (MOJ) and the Tuxtla Statuette (TUX) that are distinct from known signs of other Mesamerican scripts. “Text length” is the number of signs originally present in the text, sometimes estimated. CHP-2 and MES-15 are presumed to be Epi-Olmec because, in addition to being from sites yielding demonstrably Epi-Olmec texts, they share a sign form for the day Reed that is distinct from that of neighboring Mayan and Zapotec traditions. MES-8 bears the Epi-Olmec sign (mi).

sign; accordingly, most of the verb and noun morphology was worked out in the first few months of our collaboration. This enabled us to distinguish nouns from verbs, transitive verbs from intransitive, and subordinate from main clauses, and thereby to begin exploring syntactic patterns. As syntactic regularities were identified, they permitted us to refine our analysis of the morphology and vocabulary of ambiguous cases.

We now have almost complete translations and, more importantly, grammatical analyses, for all of the readable Epi-Olmec texts. They conform in all major features, and in almost all details, with what was already known of the grammatical structure of Mije-Sokean languages from the results of comparative reconstruction. A few features were recognized in Epi-Olmec texts before they were known from extant Mije-Sokean languages,

and other features of Epi-Olmec texts provided data that have contributed to comparative reconstruction.

Some professional epigraphers who do not know our evidence have expressed doubt about the reliability of the decipherment, but the essentials of the decipherment as it relates to Mije-Sokean linguistic structure are accepted by the leading authorities who do know the evidence (Grube, Kelley, Lounsbury, Mathews, Schele, Urcid). It is not believable that the model presented in this study for the phonological and grammatical structure of the Epi-Olmec language could fit both the comparative Mije-Sokean data and the Epi-Olmec epigraphic data in the detail that it does were it not fundamentally correct (in contrast, no such fit is feasible with a language model based, for example, on Mayan or Oto-Manguean). In particular, the picture we have uncovered of the system of person and aspect/mood marking we consider unassailable. The decipherment is further supported by linguistic features that were not initially known to be reconstructible, but which we found, on gathering and examining more extensive data from the extant Mije-Sokean languages, could be and should be reconstructed (see §§4.3.5, 4.3.6, 4.4.5, 5.1, 5.2.3).

## 2.2 The Epi-Olmec script






















































The historical position of the Epi-Olmec script among Mesoamerican traditions is not entirely clear. A text on the El Sitio celt, from the southeastern tip of Mije-Sokean territory, is in either a stylistically divergent form of the Epi-Olmec script or an otherwise unattested script that is its nearest relative. The iconography accompanying this text is considered early post-Olmec. We believe the El Sitio and Epi-Olmec scripts descend from some Olmec script; bare traces of one such script survive, from the end of the Olmec era, on La Venta Monument 13. Several Epi-Olmec signs derive from Olmec iconographic elements. Three signs belonging to the Epi-Olmec script are used as labels on iconography at Izapa, a few kilometers from El Sitio, but no evidence of textual writing survives from the site.

Otherwise, the closest relative of Epi-Olmec writing is Mayan writing. We believe that it arose from an ancestor or sister of Epi-Olmec writing. Mayan writing seems to have emerged from the (set of?) script(s) in Guatemala's southern highlands and adjacent Pacific slopes. That zone's only long, legible text, on Kaminaljuyu "Stela" 10, has some signs otherwise known only from Mayan and/or from Epi-Olmec writing, reflecting some currently un-specifiable historical relation to both Epi-Olmec and Mayan. Complicating the historical picture, some Epi-Olmec signs and their values were adopted by Mayans, and vice versa; other notational and cultural practices also passed between the two groups, like the "long count" calendar and positional numerical notation. Script traditions in the rest of Mesoamerica seem to descend ultimately from Zapotec writing; it is unclear whether Zapotec writing arose from an Olmec (or Olmec-derived) script, but its earliest sure attestation is some 200 years after the Olmec era.

Epi-Olmec signs are arranged in columns, read from top to bottom. Columns were normally read from left to right, and asymmetrical signs faced leftward, but text associated with rightward-facing iconography was reversed in orientation. Successive signs that are part of the same word or phrase often abut, or are set side by side within the space normally occupied by a single sign.

The script is "hieroglyphic" (i.e., its signs have a pictorial quality) and logosyllabic (i.e., it has both logographic and syllabic symbols). Often, sign form relates iconically to sign value. A logogram's form usually relates directly to its meaning – for example, the logogram for

**Table 44.2** Epi-Olmec syllabary

	<i>i</i>	<i>e</i>	<i>ɬ</i>	<i>a</i>	<i>u</i>	<i>o</i>
<i>p</i>						
<i>t</i>						
<i>tz</i>						
<i>k</i>						
<i>ʔ</i>						
<i>s</i>						
<i>j</i>						
<i>m</i>						
<i>n</i>						
<i>w</i>						
<i>y</i>						

piercing depicts a shaft passing through a rectangular field – but sometimes the relationship is more complex, and often it is unknown. The pictorial referent of syllabograms is usually unknown; when this referent is clear, the sign's value is the initial CV(C) of the pre-Proto-Sokean word for the depicted entity: for example, ⟨po⟩ from *pomɬ7* “incense”; ⟨nɬ⟩ from *nɬ7* “water”; ⟨na⟩ from *nas* “earth.”

In text frequency, somewhat over half of all signs represent a simple open (CV) syllable, and this is easily the most common type of phonetic sign; for the currently known instances, see Table 44.2. A few signs represent closed (CVC) syllables. No sign represents a simple vowel or VC syllable, because all syllables in Mije-Sokean languages begin with a consonant. Logograms are the most numerous sign type, although they are textually less frequent than syllabograms.

Words and stems of all grammatical classes are spelled by logograms or syllabograms or both. As almost always in all other writing systems, grammatical affixes are spelled using phonetic signs only and, with two restricted exceptions to be discussed, all Epi-Olmec grammatical morphemes are explicitly spelled out.

Spelling with CV signs is straightforward in the case of simple open syllables. In fully phonetic spellings, there is a mismatch between the structure of the Epi-Olmec language and that of the CV syllabary that was used to write it: in the case of syllable codas ending in a consonant, where a consonant in the language is not pre-vocalic, using a CV sign “inserts” an unpronounced vowel, and failing to do so suppresses a consonant. CVC signs are sometimes used, when appropriate ones exist, but they are seemingly rare in Epi-Olmec. In this script, “weak” consonants (/w/, /y/, /j/, /ɭ/) that are not followed by a vowel are never spelled by

CV signs; for example, /j/ is spelled in ⟨ja-ma⟩ for *jama* ‘day,’ but not in ⟨we-pa⟩ for *wej-pa* ‘he shouts.’

In fully phonetic spellings of words or morphemes, CV signs spell almost all coda instances of the remaining, “strong” consonants. The syllabogram’s vowel is always the last preceding vowel in the word, as in the following examples:

- |     |                            |     |                      |                            |
|-----|----------------------------|-----|----------------------|----------------------------|
| (1) | ⟨ta- <u>ma</u> ⟩           | for | +ta7 <u>m</u>        | ‘animate plural’           |
|     | ⟨t <u>u</u> - <u>nu</u> -⟩ | for | t <u>u</u> n+        | ‘inclusive ergative’       |
|     | ⟨7i- <u>si</u> ⟩           | for | 7i <u>s</u>          | ‘behold’                   |
|     | ⟨na-tze- <u>tze</u> -ji⟩   | for | na+tzet <u>z</u> ji  | ‘when I chopped it’        |
|     | ⟨7i-ki- <u>pi</u> -wu⟩     | for | 7i+kip <u>w</u>      | ‘they fought against them’ |
|     | ⟨na- <u>sa</u> -wu⟩        | for | na <u>s</u> wu       | ‘it passed’                |
|     | ⟨mi- <u>si</u> -na-wu⟩     | for | mi7k <u>s</u> nay7wu | ‘it had quivered’          |

This principle for choosing the vowel of a CV sign that spells a coda consonant is called *synharmony*. (Our conventions for spelling Mije-Sokean – including Epi-Olmec – words are given at the end of this study.)

One systematic exception is that *k* (and probably *p*) is never spelled before *s*. Four or five words with *k* or *p* before *s* have fully phonetic spellings. In each with *k* before *s*, the *s* is spelled but the preceding *k* is not:

- |     |                                |     |                                   |                             |
|-----|--------------------------------|-----|-----------------------------------|-----------------------------|
| (2) | ⟨7i-BLOOD-mi+si <sub>2</sub> ⟩ | for | 7i+n <u>u</u> 7pinmi7k <u>s</u> i | ‘when he quivered bloodily’ |
|     | ⟨mi-si-na-wu⟩                  | for | mi7k <u>s</u> nay7wu              | ‘it had quivered’           |
|     | ⟨7i-nu-si⟩                     | for | 7i+n <u>u</u> k <u>s</u> i        | ‘when it goes’              |
|     | ⟨?su+?su⟩                      | for | su7k <u>s</u> u7                  | ‘hummingbird’               |
|     | ⟨7o-wa-?ju-si⟩                 | for | 7owaju7p <u>s</u> i               | ‘macaw lashing’             |

In the one possible example with *p* before *s*, what we read as a syllabogram ⟨ju⟩ may be a logogram for LASH, pronounced /ju7ps/.

Besides being used in fully phonetic spellings, syllabograms could be used in a *phonetic complement*, which partially spells the beginning or ending of a word represented (in whole or in part) by a logogram. Vowel choice and consonant representation follow the same principles as in fully phonetic spellings – CV signs for coda consonants contain the last preceding vowel, and weak consonants are only represented prevocally.

- |     |                      |                      |  |                                     |
|-----|----------------------|----------------------|--|-------------------------------------|
| (3) |                      |                      | <i>Purely<br/>logographic<br/>spelling</i> | <i>With phonetic<br/>complement</i> |
|     | <i>Word</i>          | <i>Gloss</i>         |  |                                     |
|     | 7ame7                | ‘year’               | YEAR                                       | ⟨YEAR-me⟩                           |
|     | tzap                 | ‘sky’                | SKY  | ⟨SKY-pa⟩                            |
|     | 7i7ps                | ‘twenty’             |  | ⟨TWENTY-si⟩                         |
|     | ni7.jup.7            | ‘body-covering’      |  | ⟨LOINCLOTH-pu⟩                      |
|     | na+tzet <u>z</u> -ji | ‘when he chopped it’ |  | ⟨na-tze-tze-CHOP-ji⟩                |
|     | matza7               | ‘star’               |  | ⟨ma-STAR-tza⟩                       |

All of the above spelling conventions are followed without exception in Epi-Olmec texts. Two further conventions apply optionally.

First, the sequence /i7i/ can be spelled as if it were /i/ – otherwise said, the syllable /7i/ is not spelled out in some instances when it immediately follows the vowel /i/. This can be recognized because the syllable /7i/ represents the third-person ergative pronominal prefix, which is grammatically required in some contexts.

We do not believe that this orthographic reduction reflects a phonological process. Although somewhat similar phonological reductions of /ʔ/ are found in some Sokean languages, the pattern is not reconstructible back beyond the separate existing languages, and could easily be recent. As a fast-speech phenomenon, such a phonological reduction is unlikely to be recorded in formal writing.

Second, whenever a verb is spelled with a logogram that represents a verb stem, that sign alone may be read as spelling the verb root plus a suffix of the shape //E(7)// or //A(7)// (where capital-letter symbols are used to indicate that the vowel height alternates: //E// is realized as /i/ after a preceding high vowel, otherwise /e/; //A// is /a/ after a preceding mid vowel, otherwise /ʌ/).

The preceding convention has a plausible source in facts specific to Mije-Sokean grammatical structure. In these languages, no lexical verb can occur without either an aspect/mood suffix or a nominalizing suffix, the two most common of which are {-E} (either the dependent incompletive or the homophonous passive nominalization) and {-A7} (either the imperative or the homophonous active nominalization). A logogram that spelled a verb was probably cited as a nominalization that corresponded to that verb. For example, the sign ⟨PIERCE⟩, which probably spells the verb stem /wu7tz/ 'pierce,' depicts an empty area pierced, and so might have been pronounced /wu7tz.i/ 'pierced (thing).' As a result, such logograms would also be able to represent certain nominalizations of the verb. But these nominalizations were also homophonous with the stem and suffix in certain inflected forms of the same verb, so the citation form could be used for these verb forms as well. The most commonly used of these inflexions is a dependent incompletive form; thus ⟨7i-PIERCE⟩ alone can spell /7i+wu7tz-i/ 'when it gets pierced; when he pierces it,' as it does at MOJ:O\*34–35.

When explicit consonant-initial verbal suffixes were spelled out, the verb's logographic (or CVC) representation corresponded to the verb stem without its suffixes.

### 3. PHONOLOGY

#### 3.1 Phoneme inventory

Proto-Mije-Sokean had eleven consonantal phonemes and six vowel phonemes, and phonemic vowel length. (We write Mije-Sokean forms in a practical, Spanish-based orthography. Most letters have their usual Spanish pronunciation, but *j* represents [h]. 7 represents a glottal stop, *tz* represents a sibilant affricate [c], and *ʌ* represents a high, central-to-back unrounded vowel [i]. *a* is a low, central-to-back unrounded vowel. IPA equivalents are provided in the lexicon [§6].)

<u>consonants</u>					<u>vowels</u>		
p	t	tz	k	7	i	ʌ	u
		s		j	e		o
m	n				a		
		y	w				

The Epi-Olmec syllabary agrees with the phonological system of reconstructed Proto-Sokean, Proto-Mijejan, and Proto-Mije-Sokean, in contrasting eleven segmental consonants and six segmental vowels. The following sign series exemplify the contrasts (each syllable listed is the value we have assigned to an Epi-Olmec sign; parenthesized values we do not consider secure):



pSo \**koya7* ⇐ PMS \**ko:y7a7* 'tomato'  
 ⇒ Tzutujil /xko:ya:7/, Awakateko /xko:ya7/ ~ /xko:yi7/  
 pSo \**wetu7* 'fox' ⇐ pre-Proto-Sokean \**we:tu7* ⇒ Xinka /we:to/  
 pSo \**yumi* 'high-status person' ⇐ pre-Proto-Sokean \**yu:mi*  
 ⇒ Yukatekan /yu:m/  
 pSo \**7amu* 'spider' ⇐ pre-Proto-Sokean \**7a:mu* ⇒ Xinka /7a:mu/  
 pSo \**pomɯ7* ⇒ pMS \**po:mɯ7* 'incense' ⇒ *po:m* in various Mayan languages

Thus, it is possible that Epi-Olmec had vowel length, but the evidence for reconstructing it is slim, and the orthography of Epi-Olmec does not represent it.

### 3.3 Phonotaxis and diachronic developments

Proto-Mije-Sokean and Proto-Sokean syllable shapes include CV, CVC(s), CV7C(s) and CVC7. Disyllabic and trisyllabic words could end in V, V7, and Vj. Only *k* and *p* occurred before *s*.

In addition, Proto-Mije-Sokean and Proto-Mijejean have syllable shapes of the type CV:, CV:C(s), CV:7C(s), and CV:C7; and Proto-Mije-Sokean (but not Mijejean) syllables could have the shape CVC7. Proto-Mije-Sokean di- and trisyllabic words could probably end in V: and V:7 as well, but Mijejean points to only V versus V7 or V:.

In the evolution of Proto-Mije-Sokean to Proto-Sokean, the following simplifications occurred:

- [a] vowel length was lost
- [b] /7/ was deleted between C and V, unless /7/ began a suffix
- [c] /7/ was deleted word-finally after C, except when C was a resonant

There is no orthographic evidence that vowel length was preserved in Epi-Olmec (see §3.2), but /7/ was preserved between C and V. The Epi-Olmec words ⟨po-7a⟩ (/poy7a/) 'moon, month', ⟨HEAD.WRAP-7a⟩ (/ko7=mon7.a/) 'headgear', ⟨PLANT-7i⟩ (/nip7.i/) 'planting', and ⟨SPAN-7ɯ⟩ (/tsat7.ɯ) 'hand-span measure' are evidence for the preservation of /7/ in this environment.

The spelling ⟨kak-SCORPIUS-pe⟩ for 'Scorpius' shows that the Epi-Olmec pronunciation of 'scorpion' was /kakpe7/ as in Proto-Mije-Sokean, and not /kakwe(7)/ as now universal in Sokean. This shows that Epi-Olmec had not undergone the shift of Proto-Mije-Sokean \**kp* to Proto-Sokean \**kw*.

## 4. MORPHOLOGY

Mije-Sokean morphology (and syntax) is right-headed or left-branching: modifiers precede heads. This principle is not totally obvious morphologically, since in their inflection and derivation verbs take certain suffixes which are recruited from lexical verbs but have dependent functions grammatically. But with regard to word order, right-headedness is pervasive and obvious, and accounts for SOV, A N, G N, R N, and N Po orders; see §5.1.

In morphologically explicit representations of Mije-Sokean words, inflexional affixes are marked by -, clitics by +, derivational affixes by ., class-changers by >, and compounding, prepounds, and postpounds by =.



## 4.1 Word classes

Epi-Olmec (like Mije-Sokean languages generally) has the following root and lexeme classes: nominal (noun, adjective, quantifier), verb (transitive, intransitive, positional), and particle (of various functions).

## 4.2 Person and Number Marking

### 4.2.1 Person

Mije-Sokean languages distinguish four person categories: exclusive, inclusive, second, and third. While the meaning of inclusive entails at least two persons, all of these categories are subject to optional pluralization.

Proto-Mije-Sokean had an ergative morphology, reflected in two sets of person markers:

1. *The absolutive set*: This set of person markers agrees with (i) the object of a transitive verb (see [7B], [10A], [14], [17], [20AB], [21], [27]); (ii) the subject of an independent intransitive verb (see [6]–[9], [11], [12], [15], [18], [26A], [31]–[34]); and (iii) the subject of a predicate noun or adjective (see [10BC], [13], [16], [19], [22]–[25], [26B], [28], [29]). It forms the basis of independent non-third-person personal pronouns
2. *The ergative set*: This set marks (i) the subject (agent) of a transitive verb (see [7B], [10A], [14], [17], [20AB], [21], [27]); (ii) the subject of a dependent verb (see [9], [17], [19], [34]; and (iii) the possessor of a noun (see [9], [10ABC], [17], [18], [29]).

Person markers are proclitics; the ergative markers are arguably affixes, and the absolutive markers are arguably words. When both an absolutive marker and an ergative marker precede a lexical item, the absolutive marker precedes the ergative.

This Proto-Mije-Sokean system was maintained intact in Soteapan and Texistepec Gulf Sokean (and perhaps in Ayapa Gulf Sokean) and Epi-Olmec, but has been partially and differentially changed in the other individual Sokean languages. The person markers reconstructed for Proto-Mije-Sokean and Proto-Sokean are as follows (affixes that are actually attested in Epi-Olmec texts are in boldface type):

	<i>Absolutive</i>	<i>Ergative</i>
<i>First exclusive (X)</i>	<b>7u+</b> [22]	<b>na+</b> [9], [10A], [29]
<i>First inclusive (I)</i>	<b>tu+</b>	<b>tun+</b> [14]
<i>Second (2)</i>	mi+	<b>7in+</b> [10BC]
<i>Third (3)</i>	Ø ( <i>passim</i> )	<b>7i+</b> [9], [10A], [17]–[19], [20B], [21], [27] [29],[34]

The inclusive absolutive **tu+** is not found in our texts, and the exclusive absolutive marker **7u+** is found only in nominal predicates. Because the third-person absolutive marker is Ø (zero marking), independent intransitive verbs in these texts are spelled without any overt person marker, while transitives and dependent verbs are spelled with overt marking (except for the cases discussed in §2.2 where /i7i/ is spelled like /i/). Inclusive ergative **tun+** occurs on just one verb (twice), and second-person ergative **7in+** marks a god as possessor of two different nouns. Exclusive ergative **na+** is well attested and third-person ergative **7i+** is frequent.

### 4.2.2 Pluralization

The unmarked number is singular. Plural marking has several loci and subdivisions in Mije-Sokean languages; “plurality” refers to a noun or to a person agreement category. It can be marked on a noun, pluralizing the noun or its possessor. It can be marked on a verb,

pluralizing a subject or an object. A distinction may be made between third and non-third persons, and between animate and inanimate nouns.

The marking of plurality is optional, even avoided, once plurality for a noun phrase or pronominal category has been established. Complete data on plural marking are not yet on hand for all Mije-Sokean languages, but are known for Soke (Copainalá, Magdalena, Santa María Chimalapa, San Miguel Chimalapa), Soteapan Gulf Sokean, Oluta Mijean, Sayula Mijean, and Lowland Mije.

Originally, the plural marker on nouns was probably Proto-Mije-Sokean  $\{+tuk\}$ , found in Mijean and Oaxaca Soke, with no animacy distinction.

The pluralizer for third-person subjects, objects, and possessors in *every* Mije-Sokean language is identical to that language's lexical verb root meaning 'to be finished': Mijean  $\{kux\}$ ; Santa María Chimalapa Soke, San Miguel Chimalapa Soke  $\{suk\}$ ; other Sokean  $\{yaj\}$ . It was probably  $\{-yaj\}$  in Proto-Sokean; this is also found in Epi-Olmec (see [14], cf. [7A]).

The pluralizer for non-third-person subjects, objects, and possessors was Proto-Mije-Sokean  $\{-ta7m\}$ ; no affix with this function happens to be attested in Epi-Olmec texts.

Proto-Mije-Sokean  $\{+tuk\}$  is displaced by certain affixes that are probably to be seen as elite innovations. Gulf Sokean  $\{+yaj\}$  arises as an extension of  $\{-yaj\}$  to serve as an inanimate or nonhuman noun pluralizer; in Gulf Sokean and Chiapas Soke,  $\{+ta7m\}$  arises as an extension of  $\{-ta7m\}$  to serve as an animate or human noun pluralizer. Both extensions are found in Epi-Olmec as well (see [10C], [20B]).

### 4.2.3 Gender

Gender is not a grammatical category in Mije-Sokean languages.

## 4.3 Verb morphology

Verbs begin with (i) obligatory pronominal agreement markers (§4.2.1), optionally followed by (ii) incorporated modifiers (§4.3.3) and various derivational prefixes (§4.3.4). Next comes (iii) the verb root, then (iv) a variety of optional derivational suffixes and class changers (§4.3.5), followed by (v) a variety of optional inflexional suffixes, and finally (vi) a single obligatory aspect/mood marker (§4.3.2). Essential distinctions to be made include that between ergative and absolutive, transitive and intransitive, various aspects and moods, and dependent versus independent status.

### 4.3.1 Verb classes

#### 4.3.1.1 Transitivity

Any lexical verb is either transitive or intransitive, though a certain percentage are bivalent in that they can be inflected as either transitive or intransitive with no overt intransitivizers or transitivity markers.

In the Epi-Olmec texts, the verb stems /ko.wik/ 'sprinkle elsewhere/for others' and /saj/ 'to share' each occur both with and without an ergative prefix. In addition, several independent or optative verbs that are transitive in Mije-Sokean generally are found inflected intransitively in Epi-Olmec, with their single argument being a patient. The reason for this is that approximately one out of every six Mije-Sokean verbs is *bivalent*, occurring sometimes as a transitive and sometimes as an intransitive verb with no transitivity or detransitivizing suffix. In Sokean, the subject of some of the intransitive forms is an agent, in most a patient;

when the subject is a patient, the verb has a mediopassive interpretation. In Epi-Olmec, only patient subjects occur in the available texts: for example, *puw-wu* ‘it got scattered.’ Precisely which verbs are bivalent is a lexically specific fact that differs from language to language.

#### 4.3.1.2 Positional roots

Positional roots in Sokean are defined by their occurrence with three suffixes: (i) the suffix *\*{.nay7}*, which forms an *assumptive* (‘to get into X position/state’) intransitive stem; (ii) the suffix *\*{-wu7y}*, which forms a *depositive* (‘to leave something that is in X position/state’) transitive verb; and (iii) the suffix *\*{.na7}*, which forms a *stative* adjective/participle (‘that is in X position/state’). The positional root with no derivational suffix can normally be used as a transitive verb with causative function (‘to make something be in X position/state’). In Epi-Olmec only the suffix *{.na7}* (stative) is attested thus far.

#### (6) MOJ T24–28

T    te-ne-na-kak-wu  
 R    Ø-te7n.na7=kak-wu  
 G    3A-tip.toe-STAT-replace-1C  
 FT   It got replaced upright.

(Here and in other example sentences, the first line presents a transcription of the relevant portion of the cited inscription; the second line offers a pre-Proto-Sokean reading of this; next follows a morpheme-by-morpheme gloss; the last line presents a free translation; a full list of grammatical codes precedes the bibliography.)

### 4.3.2 Aspect and mood

Each Mije-Sokean verb carries an obligatory aspect/mood suffix as its final morpheme. In Proto-Sokean, there are six to eight such affixes. These are not distinguished by the transitivity of the verb to which they are attached; instead, aspect markers are distinguished on the basis of their dependent versus independent status.

There are apparently at least six categories of aspect and mood in Proto-Mije-Sokean: *incomplete*, *completive*, *imperative*, *vetative*, *optative*, and *irrealis*. Verbs form matched pairs differing for dependent versus independent function:

	<i>Independent</i>	<i>Dependent</i>
<i>Incomplete</i>	*-pa	*-e ( > *-i after V <sub>high</sub> )
<i>Completive</i>	*-wu	*-ji
<i>Imperative</i>	*-u7 ( > *-a7 after V <sub>mid</sub> )	
<i>Vetative</i>		*-wu <sub>2</sub>
<i>Optative</i>	*-7in (Proto-Sokean)	
<i>Irrealis</i>		*-u ( > *-a after V <sub>mid</sub> ; Proto-Soke )

The *vetative* is negative (and dependent) imperative, and has other functions in Soke languages that may not be original. Though homophonous with the independent completive, its functions are quite different, and it probably should be considered a separate morpheme. The *optative* is found in Sokean but is not known from Mijejan, where its function may be filled by the descendant of the imperative. Another (dependent) category, *irrealis* or subjunctive, that is pointed to by Soke languages, would have been phonologically eroded in Gulf Sokean and is therefore not directly reconstructible from them. The irrealis occurs with certain subordinating “conjunctions,” and in other so far poorly characterized

contexts. The *dependent completive* did not survive into present-day Sokean languages, but was still attested at the Epi-Olmec stage. Intransitive verbs with dependent incomplete and dependent completive suffixes use ergative rather than absolutive person agreement markers. This phenomenon is called *ergative shift*.

In Epi-Olmec we have identified independent incomplete *-pa* (see [21], [26AB], [29], [31], [33]); incomplete dependent *-e ~ -i* (see [9], [29], [34]); independent completive *-wu* (see [6], [9], [10A], [11], [12], [15], [18], [20B], [27], [32]–[34]); and dependent completive *-ji* (see [9], [14], [17]).

We have not identified any likely imperative, vetative, or irrealis verb forms in the texts. Three verbs appear as optatives, one of them twice (see also [30]):

- (7) A. *MOJ U1–3*  
 T yaj-7i “SACRIFICE”  
 R Ø-yaj -7i SACRIFICE  
 G 3E-finish-OPT ??  
 FT The “dripping sacrifice” was supposed to be finished/used.up.
- B. *MOJ R9–17* (with two possible readings given)  
 T AFTER-?su NINE ja-ma JAGUAR puk-ku-7i  
 R jus maktas=tujtu jama Ø-kajaw=puk-7i  
 kajaw Ø-puk-7i  
 G back four-past.five day 3A-jaguar-take-OPT  
 jaguar 3A-take-OPT  
 FT: Nine days later he was supposed to take a jaguar [with an incorporated direct object; see §4.3.3]  
 OR: Nine days later [once again] a [tenth] jaguar was supposed to get taken.
- C. *MOJ P\*40–Q2*  
 T NOW puk-7i 7o-wa-ju/LASH-si  
 R ADV<sub>1</sub>-ti Ø-puk-7i 7owa=ju7ps.i  
 G now 3A-take-OPT macaw-lash-PN  
 FT Now a macaw-lashing/?band was supposed to get taken.

As the preceding examples show, the optative suffix is spelled <7i> in Epi-Olmec texts rather than <7i-ni>, the expected spelling of /7in/; this discrepancy is discussed below (see §7.2.1).

### 4.3.3 Incorporation

One common feature of Mije-Sokean languages is the incorporation of adjective, noun, and verb stems as modifiers of a verb (so-called incorporees). Their incorporated status is signaled, for example, by the occurrence of the pronominal agreement markers of the verb *before* the incorporee. These texts provide several examples of noun incorporation. Most are intransitive, with third-person subject, whose agreement marker is Ø-, so that no pronominal is explicitly spelled out (also [6], [7B], [12]):

- (8) *MOJ Q27–30*  
 T SING-ne-DO-pa<sub>2</sub>  
 R Ø-wan.e=tzuk-pa  
 G 3A-sing-PN-do-II  
 FT He sings a song.

But there are dependent intransitive cases showing ergative shift (see §4.3.2) in which the sign for BLOOD occurs after the ergative pronominal 7i+ and before the verb (see also [29]):

(9) *MOJ S44–T6*

T	NOW	na-LOSE-ye	7i-saj	7i-BLOOD-SET-ji	mi-si-na-wu
R	ADV <sub>1</sub> -ti	na+yak>tokoy.e	7i+saj	7i+ <u>nu</u> 7pin=tu7p-ji	Ø-mi7ks-nay7-wu
G	now	XE-CAU-lose-PN	3E-wing	3E-blood-set-DC	3A-quiver-PRF-IC
FT	“Then when my overthrown [rival]’s wing/shoulder came to rest bloodily, he/it had been quivering/flapping.”				

#### 4.3.4 Derivational prefixes

{ko.} ‘elsewhere’: ‘in another’s place, on someone else’s thing, for someone else’

Proto-Mije-Sokean has a prefix \*ko:- that can be preposed both to verbs (and their nominalizations) and to nouns (i.e., those that are not nominalizations of verbs). Epi-Olmec has examples of both these functions; the following illustrate its attachment to verbs (10A), to nominalizations (10B), and to nouns (10C):

(10) A. *MOJ P3–9*

T	na-BLOOD	7i-ko-LOSE-pu-wu
R	na+nu7pin	Ø-7i+ <u>ko</u> .tokoy-pu7-wu
G	XE-blood	3A-3E-ELSE-lose-ENTIRELY-IC
FT	“He spilled/hid my blood in another’s place.”	

B. *TUX C9–D6*

T	to+ke	wu	7i <sub>2</sub> +ni <sub>2</sub> -ko-SPAN-7u <sub>2</sub>	TURTLE-ki	wu
R	Ø-tok.e	+wu7	7in+ <u>ko</u> .tzat7.u7	tuki	+wu7
G	3A-stain-PN	REL	2E-ELSE-measure-AN	turtle	REL
FT	Stained [with blood?] is your elsewhere [otherworldly] handspan measure which is made of turtle[-shell].				

C. *TUX C4–8*

T	FOUR	7i <sub>2</sub> +ni <sub>2</sub> -ko-SKY+PILLAR	ya <sub>2</sub>
R	Ø-maktas	7in+ <u>ko</u> .tzap=komp	+ya <sub>2</sub>
G	3A-four	2E-ELSE-sky-pillar	IP
FT	Four are your elsewhere [otherworldly] sky pillars.		

{ku.} ‘away’

Sokean languages have three derivational prefixes beginning with /k/ that are all pronounced /ku+/ in Soteapan Gulf Sokean:

<i>Soke</i>	<i>Soteapan Gulf</i>	
(MAR, MIG, COP)	<i>Sokean</i>	
ko7=	ku+	‘with respect to the head’ (incorporated by verb, prepounded to noun)
ko.	ku+	‘elsewhere’ (preposed to verb) ‘someone else’s, step=’ (preposed to noun)
ku.	ku+	‘away,’ ‘dispersed,’ ‘separate’

We would be inclined to reconstruct each of these prefixes as it is found in Soke, since Soteapan Gulf Sokean has radically reduced them to a single shape, but while Epi-Olmec

has ⟨ko⟩ for the second function (as expected), it has ⟨ku⟩ for the third function, in a so far single example on OBM:C1–3. Here, apparently, Soteapan Gulf Sokean preserves the original phonological shape.

(11) OBM C1–3

T	BEANS × ARRAY	ku=CROSS-wu
R	suk=      wu=tzúk.i	Ø- <u>ku</u> .jak-wu
G	bean-good-do-PN	3A-AWAY-cut-IC
FT	The bean-bedecked one [Jome7 Suk] crossed over.	

For the second function Sayula Mijean has {ku+} preposed to verbs and {ko:+} preposed to nouns; and for the third function it has {ku+}, which tends to support \**ku+* as the Proto-Mije-Sokean shape for a verb prefix meaning ‘away’, ‘dispersed’, ‘separate’. Oluta Mijean has {ko7=} ‘head,’ {ko:.} ‘benefactive,’ but no correspondent to Soteapan Gulf Sokean and Sayula Mijean {ku+} = Soke {kũ.} ‘away.’

{7aw=} ‘with respect to the mouth’

Though the Proto-Mije-Sokean prepond \*{7aw=} is clearly the same as the Proto-Mije-Sokean noun root \*7aw ‘mouth,’ the meaning of the prepond is not at all clear, and only occasionally indicates ‘mouth’ in any meaningful sense.

The nominalization 7aw=*ki7m.ɯ7* ‘rulership’ appears twice on La Mojarra Stela 1. /7aw=*ki7m*/ ‘to give orders’ is reconstructibly Proto-Sokean. Proto-Sokean \**ki7m* means ‘to go up,’ and in some Sokean languages is also a transitive verb ‘to mount.’ In the case of ‘to give orders’ there is in fact a plausible reason for mentioning the mouth, even though in Sokean \*7aw no longer means ‘mouth,’ having been replaced by \**jup*, a cognate of the word for ‘nose’ in Mijean.

{ni7.} ‘on the body’

This prefix goes back to Proto-Mije-Sokean \*{ni:7.}.

(12) MOJ V25–30

T	SHAPESHIFTER <sub>2</sub>	ma-sa-ni-APPEAR-wu
R	jama	Ø-masa= <u>ni7</u> .APPEAR-wu
G	shapeshifter	3A-god-BODY-appear-IC
FT	A shape-shifter appeared divinely on his body.	

{nũ.} ‘associative’

The meaning is ‘to VERB along with someone’.

(13) MOJ O25–26

T	HALLOW	nũ <sub>2</sub> =SPAN+EARTH
R	(Ø-)ko.nu7ks.i	<u>nũ</u> .tzat7.e=na
G	(3A-)ELSE-greet	ASSOC-measure-PN-earth
FT	... (At) the hallowed ground jointly measured by hand spans...	
OR	... The ground jointly measured by handspans had been hallowed.	

(Here the associative occurs in a passive nominalization.)

### 4.3.5 Non-aspect/mood suffixes on verbs

#### {.pu7} ‘entirely, completely, all of it’

A suffix *\*{.pu7}* can be reconstructed to Proto-Sokean, and on Sayula Mijeán evidence possibly to Proto-Mije-Sokean, with the meaning ‘entirely, completely, all of it’.

The glyph MS47 occurs three times, in two distinct contexts, in the Epi-Olmec corpus (see [10A], [14]). It follows MS149+50, which we interpret as LOSE/TOKOY, at MOJ:P8 in the string (na-BLOOD 7i-ko-LOSE-MS47-wu). If MS47 is a syllabic sign spelling a verb suffix, there are not many possibilities for reading it, given the set of syllables that are established as readings of other signs. The only suffix likely corresponding to an otherwise unread syllable is *\*{.pu7}* ‘entirely,’ which suggests that MS47 spells (pu). In some Sokean languages /tokoy.pu7/ means ‘to spill,’ which would fit nicely in the context. This reading is consistent with its other context, as a phonetic complement to a verb referring to how the king’s allies dealt with his enemy, or as another instance of the suffix *\*.pu7* ‘entirely’

#### {>jay7} ‘indirective’

The valency-changing suffix {>jay7}, which goes back to Proto-Mije-Sokean, adds an indirect object argument to a verb. When added to an intransitive verb it yields a transitive verb; when added to a transitive verb, it creates a double-object construction in which the verb can agree in person with one object and in number with another object:

#### (14) MOJ S7–12

T	tu-nu-“DEAL.WITH”-pu-ja-yaj-wu
R	Ø-tun-?tup(.pu7)>jay7-yaj-wu
G	3A-IE-?spear-?ENTIRELY-NDIR-3P-IC
FT	We ?spear- /dismembered him/them for him/them.

This is the only definite example in Epi-Olmec texts of a verb with multiple optional suffixes. It conforms to general Mije-Sokean order restrictions, with {>jay7} before a pluralizer and perhaps after {.pu7} ‘completely.’

#### {-nay7} ‘perfect/progressive’

In the Epi-Olmec texts, the verb (mi-si-) /mi7ks/ ‘to quiver’ occurs twice, once with the suffix {-na} /-nay7/ (*perfect/progressive*). From comparative evidence we see that there is a Proto-Mije-Sokean suffix *\*{.na:y7} ⇒ *\*{.nay7}* in Sokean. In the present-day Mije-Sokean languages this suffix has three standard functions: (i) to form an assumptive intransitive verb from a positional root (see §4.3.1.2; the root may itself occur as a transitive stem); (ii) to form a perfect or back-shifted completeive, and a progressive as well; and (iii) to form an iterative intransitive verb from a verb root or symbolic root – in this last function the root is reduplicated.*

Since iterative function without reduplication is unknown in surviving languages, if /mi7ks-nay7/ meant ‘to quiver repeatedly’ it would be anomalous; example (9) shows that /mi7ks-nay7/ is not reduplicated. Example (29) shows that /mi7ks/ is an intransitive stem, so a form with /nay7/ would not be an assumptive based on a positional. Function (ii), the back-shifted completeive, is so far known only in Soteapan Gulf Sokean, where it can be glossed as ‘has VERBEN,’ ‘had VERBEN,’ ‘had been VERBING.’

The suffixes {.pu7}, {>jay7}, and {-nay7} occur in this order in present-day Sokean languages; in Epi-Olmec, only the relative order of {.pu7} and {>jay7} may be attested.



### 4.3.6 Nominalizers

#### {.na7} ‘stative’

From positional roots Sokean languages form a *stative adjective* or *adverbial* with the suffix {.na7}; see (6). Such a derivation has not yet been found in any Mijeian language. Two such formations are seen in Epi-Olmec texts: /kuw.na7/ ‘(having been) set aside’; /te7n.na7/ ‘standing (upright) on tiptoe.’ Both serve as incorporated modifiers, and this function is common, though not universal or necessary, in the present-day languages.

#### {.kuy7} ‘instrument’

Sokean languages have a suffix \*{.kuy7} that forms instrument nouns from verbs. (There are relic forms in Mijeian that support reconstructing this to proto-Mije-Sokean.) These nominalizations sometimes have readings that suggest they are not only instruments, but may also be nouns naming the action of the verb. One such nominalization is found twice in Epi-Olmec, at MOJ:B5–C4 and MOJ:R4–8, ⟨pak-ku⟩ /pak.kuy7/ “beating instrument” = ‘bludgeon.’

#### (15) MOJ B5–C4

T	PIERCE	ma	pak-ku	wu	ma-STAR+tza	SHINE-wu
R	wu7tz.w7	ma	pak.kuy7	+wu7	matza7	Ø-kij-wu
G	pierce-AN	earlier	beat-NSTR	REL	star	3A-shine-IC
FT	Piercingly the bludgeon star [Venus] had shone earlier.					

It appears that in some Sokean languages, nouns in {.kuy7} must be formed on intransitive active themes. The result is that for many transitives, an antipassive theme in {.7oy} is the basis of the nominalization in {.kuy7}.

#### {.7} ‘instrument’

In San Miguel Chimalapa Soke, {.7} forms instrument nouns out of particular verbs that contain certain of the possible derivational prefixes. This word-final /7/ is actually pronounced if the verb ends in a resonant (resonants are the consonants after which word-final /7/ is preserved in this language). In all other Mije-Sokean languages, this final /7/ is uniformly lost; this correlates with the fact that, in all Mije-Sokean languages, there is a handful of instrument nominalizations that are not different from the verb itself. Epi-Olmec texts have yielded one instance, ⟨?LOINCLOTH-pu⟩ for /ni7.jup.7/ ‘body covering.’ The synharmonic spelling of the stem-final /p/ shows that the nominalizer must be {.7} rather than {.A7} or {.E(7)}.

#### Active and passive nominalizers

The suffixes {.E}, {.E7}, and {.A7} are used to form agent-focus and patient-focus nominalizations in Mije-Sokean languages. Their uses in Epi-Olmec are interpreted as follows:

From a transitive verb a patient-focus (passive) nominalization can be derived with suffixes {.E} and {.E7}, and an agent-focus (active) nominalization can be derived with a suffix {.A7}. Usually, these suffixes are represented in phonetic spellings, whether partial (in final complements) or full, but they are frequently implicit in logograms for the verb stem. Even the same word can be spelled either way: for example, ⟨PLANT⟩ and ⟨PLANT-7i⟩ for *nip7.i* ‘plant(ing).’

Passive nominalization with {.E} is abundantly attested: for example, ⟨ne-ke⟩ for /ne7k.e/ ‘set aside’; ⟨tu-si⟩ for /tus.i/ ‘bristling, prickling’; ⟨tze-tze⟩ for /tzet7.e/ ‘chopped off (thing)’; ⟨LOSE-ye⟩ for /yak>tokoy.e/ ‘overthrown one’; and ⟨SET.IN.ORDER⟩ for /ne7w.e/ ‘set in

order (stones)’ (see also [7C], [10B], [11], [13], [17], [18], [20B], [22], [26AB], [28], [31]). There is less evidence for active nominalization with {.A7}, but there are enough examples in our small corpus to show that it was a productive, if not entirely regular, pattern: for example, (KNOT+GOVERNOR-7a) for /ko7=mon7.a7/ ‘head-wrap’; (?wo-ma) for /wo7m.a7/ ‘sprout’; and (LOSE-ya) for /yak>tokoy.a7/ ‘overthrower’ (see also [10A], [15], [16], [20A], [21], [27], [32], [34]). Examples of {.E7} – such as (PLAY-tai) for /mutz.i7/ ‘impersonator’ (see also [21], [23]) – are rarer still.

The evidence from present-day Mije-Sokean languages shows a fair degree of uniformity concerning passive nominalizations in {.E} and {.E7}, but no unity regarding agent nominalizations. However, in present-day Mije-Sokean languages there are numerous non-productive nominalizations with the suffixes {.A} and {.A7}, and many of these have active readings. Thus, our best current interpretation is that in Epi-Olmec there was an active nominalization in {.A7} that later lost its productivity.

## 4.4 Nouns and Noun Phrases

Nouns in Mije-Sokean languages do not need to be inflected at all. They are subject to several inflexional and morphosyntactic processes. They may be possessed, with ergative pronominal markers (see [9], [10ABC], [17], [18], [29]). They may be predicate nouns, which have subjects that are expressed by absolutive pronominal markers, as in the following example:

### (16) *MOJ P19-22*

T	TIME <sub>2</sub> +SKY.GOD <sub>d</sub> =RAIN	ma-TEN	SKY.GOD <sub>d</sub> =SKY
R	<u>Ø-tuj7=7aw=suw=jej.a7</u>	mak	tzap
G	3A-rain-MOUTH-festival-live-AN	Ten	Sky
FT	[The god] Ten Sky is/was a/the rainy season god.		

Both nouns and their possessors may be pluralized. (Plural marking is discussed further in §4.2.2 and §7.2.2.)

A formula that allows for all these markers would look as follows:

Absolutive+	Ergative+	NOUN	+NounPL.	+PossPL.
SUBJECT	POSSESSOR			

### 4.4.1 Possessive constructions

When a possessor is expressed by a noun or noun phrase, and is thus in the third person, the possessing noun (phrase) precedes the possessed noun, and the possessed noun bears the third-person ergative possession marker *7i*+, as in the following (see also [9]):

### (17) *MOJ R28-40*

T	7i-ne-ji	ja <sub>2</sub> -SYMBOL-si	7i-LOSE-ya	7i-ki-pi-wu
R	Ø-7i+ne7w-ji	<u>jay7=ki7ps.i</u>	<u>7i+yak&gt;tokoy.a7</u>	Ø-7i+kip-wu
G	3A-3E-put.stones-DC	write-think-PN	3E-CAU-lose-AN	3A-3E-fight-IC
FT	When he placed stones in order he fought against the overthrow(er)(s) of inscribed monuments.			

Adjectives and quantifiers can also occur with possessive markers. Possessed adjectives are interpreted as the corresponding abstract noun of quality (‘my/your/its X-ness’), and possessed quantifiers are interpreted as collectives (‘the number of us/y’all/them’; ‘all of me/us/you/y’all/it/them’).

#### 4.4.2 Markers of noun phrase function: core arguments, peripheral arguments, locatives, and manner

Except in Soke languages, Mije-Sokean nouns bear no case markers for such functions as ergative, absolutive, accompaniment, or instrument, and their absence from Epi-Olmec suggests that the postposed enclitic case markers of Soke are an innovation.

For locative (and manner) functions, however, Mije-Sokean languages have a variety of postposed elements with both generic and specific function.

*Generic locative markers*, roughly glossable as ‘at,’ include Proto-Mije-Sokean \*{+mʉ7}, Proto-Soke \*{+ji}, and Proto-Mijejan \*{+pi}.

*Specific locative markers* are recruited from the class of nouns, and are called “relational nouns.” These include such elements as Proto-Mije-Sokean \**kuk* ‘middle,’ Proto-Mije-Sokean \**ku7* ‘down, under,’ Proto-Mije-Sokean \**ku*s ‘up, top,’ Proto-Mije-Sokean \**yuk* ‘up, on,’ and Proto-Sokean \**joj* ‘inside.’ As nouns these items may be found both alone, and compounded with both nouns and verbs, as prepositional phrases. As locative relators they are followed by one of the generic locators and preceded by the noun they govern, or bear a possessive ergative marker.

Epi-Olmec shows *kuk*= as a prepositional phrase (see [18]) and =*joj* as a locator followed by +*mʉ7*. The combination =*joj* +*mʉ7* is in turn followed by +*k* ‘from’ (see [19]), which is so far known only from Sokean. The relational nouns are treated as prepositional phrases, and the locative markers are treated as enclitics.

{*kuk*=} ‘middle’

(18) *MOJ N18–29*

T	pe-wʉ	7i-“MACAW.POWER <sub>1</sub> ”	7i-“ECCENTRIC.FLINT”
R	Ø- <i>pey</i> -wʉ	7i+MACAW.POWER	7i+ECCENTRIC.FLINT
G	3A-brandish-1C	3E-??	3E-??

T 7i-ku-MIDDLE-tza<sub>2</sub>-ja-me

R 7i+ *kuk*=tza<sub>2</sub>=jam.e

G 3E-middle-stone-remember-PN

FT His “Macaw.power”, his “eccentric.flint”, and his pectoral stone memento got brandished.

{=joj} ‘inside’; {+mʉ7} ‘at’; {+k} ‘from’

(19) *MOJ P31–\*39*

T	(TITLE <sub>2</sub> ) x x x x x	?7i-LET.BLOOD	PENIS-jo	mʉ	kʉ
R	ancestor(?)	7i+LET.BLOOD.E	kan= <i>joj</i>	+mʉ7	+k
G	??	3E-let.blood-DI	penis-inside	LOC	FROM

T PRINCE+BRACE wʉ

R Ø-PRINCE +wʉ7

G 3A-prince REL

FT (Ancestral(?)) ... when he was blood-letting from inside the penis, he was a prince-type.

OR [When] the ancestral[?]. ... was blood-letting from inside the penis, he was a prince-type.

#### 4.4.3 Enclitic relativizer

Mije-Sokean languages have a *relativizer*, an enclitic particle that can be added to any word that can serve as the head of a phrase or constituent. The relativizer may have allomorphs

distributed according to either phonological or grammatical conditions. The meanings of relativizations vary. When added to a verb, the relativizer produces a relative clause, which always precedes the noun it modifies, according to the right-headedness principle. But a relativized (usually only intransitive) verb can also stand alone as a nominalization, such as “one who kills people” = ‘murderer,’ “one who teaches” = ‘teacher,’ “one who has studied” = ‘teacher,’ “one who studies” = ‘student,’ “one who gives orders” = ‘ruler.’

In Epi-Olmec, the only known relativizer is *+w#7*, which is found added only to nouns (no relative clauses based on verbs are attested; see §5.3.1). It is used in three ways, each according with Mije-Sokean usage generally.

When an adjective or a noun modifies a noun, the modifier *may* be followed by the relativizer (type 1: prenominal modifier; see also [15]):

(20) A. *MOJ K5–L3*

T	<i>tz#-si w#</i>	“COMMANDING.GENERAL”	<i>tuk</i>		MOUNTAIN+LORD
R	<i>tz#si +w#7</i>	COMMANDING.GENERAL	<i>tuk.#7</i>	<i>kotz#k</i>	<i>ko.yumi</i>
G	child REL ??		harvest-AN	mountain	ELSE-leader
FT	... [said] the youthful “commanding general” Harvester Mountain-Lord.				

B. *MOJ D1–F6*

T	KNOT+HALLOW	<i>ma-sa-SPRINKLE</i>		<i>ta-ma</i>
R	<i>X+ko.nu7ks.i</i>	<i>masa=wik.i</i>		<i>+ta7m</i>
G	?-ELSE-greet-PN	god-sprinkle-PSN		AP
T	NOBLE	“WARLEADER”	<i>w#</i>	<i>kak+SUPPORT ta-ma</i>
R	<i>sa7sa7(=p#n)</i>	<u>WARLEADER</u>	<i>+w#7</i>	<i>kak.e=SUPPORT.A7 +ta7m</i>
G	noble(+person)	??	REL	replace-PN-support-AN AP
T	<i>7i-ki-pi-w#</i>			
R	<i>Ø-7i+kip-w#</i>			
G	3A-3E-fight-IC			
FT	Coronated ones hallowed by sprinkling fought against noble (and) “war-leader”-type succession-supporters [i.e., would-be successors/usurpers].			

An adjective or a noun (even when used as an adverb) may be followed by the relativizer but not modify a following noun; in such cases the combination means ‘a NOUN/ADJECTIVE-type one.’ In this construction, the noun or noun phrase with *+w#7* is in apposition to another noun or noun phrase, and that one having *+w#7* comes second (type 2: postnominal apposition; see also [10B]):

(21) *TUX F1–10*

T	GOD-ja	<i>ji2-tzi</i>	TITLE <sub>3</sub> -ne	<i>tu+CLOTH w#</i>	<i>7i2-sa2+pa2</i>
R	<i>jej.a7</i>	<i>ji7tz.i7</i>	TITLE <sub>3</sub>	<u><i>tuku7</i></u>	<i>+w#7 Ø-7i+saj-pa</i>
G	breathe-AN	wrinkle-PN	??	cloth	REL 3A-3E-share-II
FT	The god Longlip <sub>2</sub> was sharing out the “Macaw-Slantbar” cloth things.				

This is the only way a noun with *+w#7* can follow another noun phrase that it pertains to.

There is also an isolated or independent (type 3) use of the relativizer. The noun or noun phrase to which *+w#7* applies is independent of other nouns or noun phrases – it is not a modifier, and it is not in apposition (see also [10B], [19], [26AB], [29]):

## (22) TUX B4–C3

T	7u <sub>2</sub> -DEEDSMAN+ki	BEARD.MASK	sa7-NOBLE	wu
R	7u+tzuk.i=pun	<u>BEARD.MASK</u>	sa7sa7(=pun)	+wu7
G	xA-do-PN-person	??	noble-person	REL
FT	I am a “deedsman,” a beard-mask (wearer), a noble one.			

## (23) CHP-sh C1–5

T	LONGLIP <sub>2</sub>	CLOTH	wu	tu-ku
R	<u>ji7tz.i7</u>	<u>tuku7</u>	+wu7	Ø-tuk.u7
G	wrinkle-PN	cloth	REL	3A-cut-PN
FT	The thing that is made of pleated cloth has been cut.			

## 4.4.4 Numerals

Numerals are common in the Epi-Olmec texts but, as is characteristic in Mesoamerica, they occur mainly in calendric expressions. There are a few noncalendrical uses, and only in these instances are the numerals sometimes spelled out phonetically, partly or completely – *metz*= ‘2’ spelled ⟨me-tze⟩; *7i7ps* ‘20’ spelled ⟨TWENTY-si⟩; *mak* ‘10’ spelled ⟨ma-TEN⟩. The uses of numerals that have been observed in the Epi-Olmec texts are the following: (i) as an enumerator/counter (preceding the word it modifies, as in six months, thirteen years, one year); (ii) as an incorporated number of times of an action (as in *metz*= ‘2’; see [26B]); (iii) as an ordinal numeral (following the word it modifies, as in *7i7ps* ‘20’; see [32]); (iv) as a coefficient to a day name or month name (preceding the word it is associated with; cf. *mak* ‘10’ in [16]).

## 4.4.5 Demonstratives

There are three demonstrative roots that can be reconstructed to the Proto-Mije-Sokean stage, and all three are attested in Epi-Olmec: \**y<sub>u</sub>7* ‘proximal / near speaker / near time of event’ (24); \**je7* ‘distal / far from speaker / far from time of event’ (25); \**te7* ‘near hearer; aforementioned’ (26AB):

{y<sub>u</sub>7} ‘this’

## (24) MOJ R18–22

T	SKIN-DRUM	?su+?su	y <sub>u</sub>	“GOVERNOR”
R	Ø-naka=kowa	su7ksu7	<u>y<sub>u</sub>7</u>	GOVERNOR
G	3A-skin-drum	hummingbird	this	“governor”
FT	This “governor”(’s headdress) was a skin-drum (and a) hummingbird.			

{je7} ‘that’

## (25) MOJ N13–17

T	SKIN-DRUM	?su+?su	je	“GOVERNOR”
R	Ø-naka=kowa	su7ksu7	<u>je7</u>	GOVERNOR
G	3A-skin-drum	hummingbird	yon	“governor”
FT	That “governor”(’s headdress) was a skin-drum (and a) hummingbird.			

{te7} ‘the aforementioned’ (also [27])

(26) A. *MOJ Q26–33*

T te SING-ne-DO-pa<sub>2</sub> ja-ma wu  
 R te7 Ø-wan.e=tzük-pa jama +wu7  
 G that 3A-sing-PN-do-II shapeshifter REL  
 FT The aforementioned one sings a song which is about/of a  
 shape-shifter/day(’s length?).

B. *OBM F1–G4*

T me-tze=UPROOT-si te PLANT  
 R Ø-metz=wis.i te7 nip7.i  
 G 3A-two-uproot-PN that plant-PN  
 FT That planting has/had been uprooted by twos  
 (i.e., two stalks at a time, one in each hand).

Demonstrative roots are used as noun modifiers/identifiers and as noun substitutes, and (among other things) are the basis of adverbs of manner, time, and place (in Epi-Olmec, the bare demonstrative roots are known to occur as noun identifiers and noun substitutes).

/je7.tzu/ or /je.tzu/ ‘thus, that way’

In addition, a manner adverb (je-tzu) (/je7tzu/ or /jetzu/) ‘thus, in that way’ is attested:

(27) *MOJ M1–7*

T GO.UP je-tzu te 7i-si-wu  
 R ki7m.w7 je7.tzu te7 Ø-(7i+)7is-wu  
 G go.up-AN yon-manner that 3A-3E-see-IC  
 FT That was how the latter/aforementioned saw/witnessed the ascent/  
 installation/accession.

{7is} ‘lo, behold’

The word 7is ‘lo, behold’ (<Proto-Mije-Sokean \*7is) has an indexical function and is attested, alone (in San Miguel Chimalapa Soke) and combined with demonstrative roots, in some Mije-Sokean languages. It is the unadorned verb root ‘to see’. At the time that 7is was recognized as an Epi-Olmec deictic adverbial, it had not yet been reconstructed for any pre-modern form of Mije-Sokean.

(28) *MOJ H3–I4*

T 7i-si<sub>2</sub> THIRTEEN YEAR BUNDLE-ti  
 R 7is mak=tukw 7ame7 Ø-pit.i  
 G see ten-three year 3A-tie-PN  
 FT [When] behold, there was a prisoner for thirteen years.

#### 4.4.6 Interrogative-indefinites

The words that serve as interrogatives, ‘who?’ {7i} and ‘what?’ {ti}, are also used in indirect or indefinite function to mean ‘the one who; whoever’ and ‘that which; whatever’ respectively. They are not necessarily used in relative function, and any such use in present-day Mije-Sokean languages plausibly results from the influence of Spanish.

{ti} 'what'(29) *MOJ* S25–34

T	7i-BLOOD-mi+si <sub>2</sub>	ti	MACAW	we-pa	na-BLOOD	wɸ
R	7i+nɸ7pin=mi7ks-i	<u>ti</u>	7owa	Ø-wej-pa	Ø-na+nɸ7pin	+wɸ7
G	3E-blood-quiver-DI	what	macaw	3A-shout-II	3A-XE-blood	REL
FT	When he quivers/flaps bloodily, what Macaw shouts is: “It/He is my bloody thing/one.”					

Though in this usage {ti} is the logical object of a morphologically intransitive verb, this usage is regular and not a grammatical violation in Soteapan Gulf Sokean and San Miguel Chimalapa Soke. Crosslinguistically, it is common (without necessarily being regular or predominant) for intransitive verbs of speaking to have what is spoken mentionable without thereby becoming transitive.

{7i} 'who'(30) *TUX* B1–3

T	7i	7o-7i
R	<u>7i</u>	Ø-7oy-7i
G	who	3A-take.trip-OPT
FT	“Who should go on a trip?”	

## 4.5 Time words

Adverbials are predicate modifiers that specify time, frequency, manner, place, extent, purpose, reason, etc. In Mije-Sokean languages there is no lexical category of adverbs, nor any standard inflexional device that produces adverbials. Many of the words that function as adverbials in these languages are invariant forms of nouns or adjectives; others, while morphologically complex, are not subject to inflexion, and have been formed by non-productive patterns of compounding and suffixation.

In Epi-Olmec texts, the following words function as or act like temporal adverbs:

{ma} 'earlier' (see also [15])(31) *MOJ* Q3–8

T	ma	ke-ne	FOLD-pa	CLOTH
R	<u>ma</u>	ken.e	Ø-paks-pa	tuku7
G	earlier	see-PN	3A-fold-II	cloth
FT	Earlier, (a) garment/cloth(s) was/were getting folded in plain sight.			

{jus} 'after'

MS89 is a logogram that occurs at *MOJ* N\*38, R9, and T17; in the last-named two it is followed by MS178. In each instance MS89 or MS89 + MS178 is followed by a numeral expression and a spelling of the word *jama* ‘day’ (see also 23B). Epi-Olmec does not begin clauses with subordinators, and there are no coordinators known in Epi-Olmec or reconstructible to Proto-(Mije-)Sokean. The most likely non-nominal in sentence-initial position is an adverbial of time or manner. MS89 seems to mean ‘after.’ In Mije-Sokean languages this would be an adverbial formed on the noun \**jus* ‘back,’ but plausibly not consisting of the bare root. A unique Proto-Sokean or Proto-Mije-Sokean word ‘after’ cannot at this time be reconstructed. MS89 and MS89 + MS178 seem to have exactly the same function. MS178, if



a syllabogram, may represent /sʉ/, or whatever syllable or morpheme ended the Epi-Olmec word meaning ‘after.’ For the moment, pending possible improvement in our knowledge, we take the Epi-Olmec word for ‘after’ to have been or begun with /jus/:

(32) *MOJ T13–23*

T	GO.UP-JAGUAR	TWENTY+si-THREE		
R	ki7mʉ7=kajaw	7i7ps ko tukʉ		
G	go.up-AN-jaguar	twenty-and-three		
T	AFTER-?sʉ	THIRTEEN	SHAPESHIFTER <sub>1</sub>	pʉk-kʉ-wʉ
R	jʉs	mak=tukʉ	jama	Ø-pʉk-wʉ
G	back	ten-three	day	3A-take-1C
FT	After thirteen days ascent jaguar [number] twenty-three got taken.			

{win} ‘in front’

MS8’ is a logogram that we interpret as meaning ‘in front,’ partly because the preceding glyph (MS90) is plausibly a syllabogram ⟨wi⟩ at OBM:G1, and ‘in front’ in Mije-Sokean languages is probably based on Proto-Mije-Sokean \**win* ‘face, eye, surface, front.’ We transcribe MS8’ as BEFORE, IN.FRONT/WIN... However, we do not know of a Mije-Sokean language where the word for ‘in front’ is simply /win/: it always has at least one suffix attached, but the attested suffixes differ across the various languages. We take MS8’ to be a logogram for ‘in front,’ but we cannot at this point specify how the word ended:

(33) *MOJ O\*27–33*

T	POUND-wʉ	DRUM	?wi-BEFORE	FOLD+pa <sub>2</sub>	tu+CLOTH
R	Ø-naks-wʉ	kow.a	<u>win..</u>	Ø-paks-pa	tuku7
G	3A-pound-1C	drum	face	3A-fold-1I	cloth
FT	The drum got pounded; [then] the garments were getting folded in front.				

## 5. SYNTAX

### 5.1 Word order

Before 1990, Kaufman had realized that Proto-Mije-Sokean must have had Subject–(Object)–Verb (S(O)V) word order, although no modern Mije-Sokean language was known to preserve this order, as all other word order characteristics are consistent with it, and the verb-initial orders of the modern languages are interpretable as a diffused feature. The Epi-Olmec data indicate that this ancestral word order was preserved, in transitive and in active intransitive clauses, while the position of the subject relative to the verb was variable in nonactive intransitive clauses. Since the Epi-Olmec pattern was worked out, SOV was found to be the basic word order in Santa María Chimalapa Soke, confirming SOV as the basic word order in Proto-Sokean and throughout the pre-Proto-Sokean era.

On La Mojarra Stela 1, within a clause, *every* object encoded as a full noun or noun phrase immediately precedes the verb that governs it (see [10A], [17], [20B], [21]). Similarly, *every subject* of a transitive verb ([20B], [21], [27]) and *every agentive* subject of an *intransitive* verb (see [26A], [29]) precedes the verb. In transitive clauses, subjects precede objects if both are present; otherwise, subjects immediately precede verbs. In addition, subjects and objects can be fronted to outside the clause (see [27], [32]).

Nonagentive subjects usually follow, but sometimes precede, intransitive verbs. Among thirty-six nonactive predicates, the predicate is more likely to precede the subject by a 3:2

margin – whether that predicate is an intransitive stem, a mediopassive, or a nonverbal predicate.

Interrogative-indefinite constituents, no matter what their syntactic function, are fronted to clause-initial position (see [29]).

## 5.2 Subordination

Eleven clauses in Epi-Olmec texts contain a verb marked with a dependent suffix ( $\{-E\}$  *dependent incomplete* or  $\{-ji\}$  *dependent complete*). Word order is completely regular: each dependent clause precedes its associated independent clause.

There are four or five cases of a dependent clause followed by a clause containing a predicate noun or adjective (see [19]), and six or seven cases followed by a clause containing a verb with an independent aspect suffix ( $\{-pa\}$  *independent incomplete*, see [29], or  $\{-w\}$  *independent complete*, see [9], [17], [34]). Four or five of the independent clauses are complete and two are incomplete; as expected on pragmatic grounds, most (all but one) agree in aspect.

In each case, the contextual meaning of this construction has to be interpreted as either (i) ‘when  $SUBJ_1$   $VERB_1$ -s/-ed,  $SUBJ_2$   $VERB_2$ -s/-ed’ (where  $SUBJ_2$  *can* have the same referent as  $SUBJ_1$ , but need not); or (ii) ‘when  $SUBJ_1$   $VERB$ -s/-ed,  $SUBJ_2$  is/was NOUN/ADJ.’ This meaning for the construction was not known from extant Mije-Sokean languages at the time it was found and identified in the Epi-Olmec texts. Since then it has been found to be a living construction in Santa María Chimalapa Soke and in Totontepec Mije. It has not yet been determined whether such a construction exists in other Mije-Sokean languages.

$\{+7k\}$  ‘when; temporal subordinator with past tense reference’

This element is found just once in known Epi-Olmec texts. It occurs attached to a verb with the dependent incomplete suffix, which already has to be interpreted as temporally subordinated. Thus, its use here is presumably optional. (This happens to be the one dependent clause that disagrees in aspect with the independent clause.) A cognate suffix is actively used in Soke (Copainalá Soke  $\{+7k\}$ , ‘temporal subordinator [possibly with past tense reference on verbs]’; Magdalena Soke  $\{+7(\text{u})k\}$ , ‘temporal subordinator [with past tense reference on verbs]’; Santa María Chimalapa Soke  $\{+7k\}$ , ‘temporal subordinator in predicates with past tense reference’, and  $\{+k(\text{u})\}$ , ‘marker of subordination on preposed dependent clause’ and ‘marker of fronted verb phrase within a clause’). In addition, in Soteapan Gulf Sokean there are frozen instances of  $\{-k\}$  with the apparent meaning ‘temporal subordination’ on some temporal adverbs and preposed subordinators. In present-day Soke, the temporal subordinating clitic is used on verbs bearing independent (rather than dependent) aspect suffixes, but still showing ergative shift on intransitive verbs. The usage in present-day Soke suggests that in Epi-Olmec an additional meaning of ‘past time’ may have been present. The optional  $(\text{u})$  in Magdalena Soke is plausibly epenthetic and harmonic in origin, is not compatible with the Epi-Olmec spelling, and probably should not be projected back into the Proto-Sokean reconstruction, which we propose to be  $*\{+7k\}$ .

### (34) *MOJ Q18–25*

T ...	7i-ko-te	$k\text{u}$	PIERCE $\times$ NOW
R	$\emptyset$ -7i+kot-e	<u><math>+7k\text{u}</math></u>	wu7tz. $\text{u}7$ ADV <sub>1</sub> -ti
G	3A-3E-put.away-DI	WHEN	pierce-AN now
T	“STAR.WARRIOR”	HALLOW- $w\text{u}$	
R	STAR.WARRIOR	$\emptyset$ -ko.nu7ks- $w\text{u}$	
G	??	3A-ELSE-greet-IC	
FT	When he was putting it away, piercingly now the “star-warrior” got hallowed.		

### 5.3 Unattested traits of Epi-Olmec grammar

Some basic aspects of Mije-Sokean grammar are not yet documented in Epi-Olmec texts. None of the ten grammatical traits discussed below have been found in any of the known texts.

#### 5.3.1 Relativized VPs

In Mije-Sokean languages, relative clauses can be formed by postposing the relativizer to an inflected verb; often a relativized intransitive verb is lexicalized as the name of a type of person who does or undergoes some action. No relative clauses built out of an inflected verb have been identified in Epi-Olmec texts, and they are probably not present in them. The morphemes that mark the relativizing function are various; some languages show as many as three shapes, including /ʔ/, /wʉʔ/, /pʉʔk/, /pʉʔ/, /ʔpV/, /pʉ/, and /p/. Proto-Mije-Sokean probably had \*wʉʔ and a shape something like \*pʉʔ, since both pronunciations are found in languages of both major branches. The form /wʉʔ/ is known in Epi-Olmec following nouns; we are not entitled to predict the shape of the relativizer with verbs without explicit spellings, although something that would be spelled ⟨pʉ⟩ is a highly likely one.

#### 5.3.2 Negatives

A predication negator \*{ya} (~ \*{yak}) is known from most Sokean languages; most forms of Soke have other or additional predication negators which are higher predicates or auxiliaries. The negator \*{ʔuy} (~ \*{ʔu}) is found in negative imperatives (vetatives) and some quantifiers. Both \*{ya} and \*{ʔuy} are reconstructible to Proto-Sokean. Mijean has a negator of the approximate shape \*{ka}, and another, \*{ni}, which may be pre-Columbian.

#### 5.3.3 Causative

All Mije-Sokean languages make use of a reflex of the Proto-Mije-Sokean causative morpheme \*{yak>} to causativize both intransitive and transitive verbs.

#### 5.3.4 Antipassive

Sokean languages have an antipassivizing suffix {>ʔoy}, which would be spelled ⟨ʔo⟩ in Epi-Olmec texts. It has not been observed in our texts, and is probably not present. Mijean languages have no antipassivizing suffix and no cognate to Proto-Sokean \*{>ʔoy}.

#### 5.3.5 Passive and reflexive

All Sokean languages have at least one morphological device (a suffix) for removing the agent from a transitive verb, producing what can be called an agentless passive; the same suffix also has a reflexive reading in most languages (this form is intransitive, having a single argument). However, the Sokean languages do not agree on a single shape for this function, and some languages have more than one suffix that seem to have the same function. A suffix shape \*{>Atʉj} may be reconstructible to Proto-Sokean, but nothing like it has been found in Epi-Olmec texts. Mijean languages use a prefix \*{yak>} (homophonous with the causative) with a passive reading.

### 5.3.6 Syntactic reflexive

Besides the morphological agentless passive that may have a reflexive reading, Sokean languages may form a reflexive construction with the noun *\*win* ‘face; surface; front; body.’ As a direct object, *\*win* marks the reflexive person; it is possessed with the same ergative marker that marks the subject of the reflexive transitive verb. This construction has not been observed in Epi-Olmec texts.

### 5.3.7 Reciprocal

In Mije-Sokean languages generally, a reciprocal form of a verb is formed by preposing *\*{nay+}* to a *passive/reflexive* verb before the lexical stem and after the person markers.

### 5.3.8 Indefinite subject

In Sokean languages, from any intransitive verb, an indefinite subject form can be made by suffixing *\*{>Anum}* to the lexical stem before adding aspect and mood markers. The meaning is ‘people VERB,’ ‘there is VERBING.’

### 5.3.9 Auxiliaries

All Mije-Sokean languages have constructions in which a small set of *auxiliary* verbs act as the syntactic heads but semantic modifiers of a *main* lexical verb. These auxiliaries encode such meanings as ‘go to VERB,’ ‘come to VERB,’ ‘want to VERB,’ ‘finish VERBING,’ and a number of other similar notions. From existing languages, it would seem that the most archaic pattern would be

ABS+ (ERG+)MAIN.VERB-suffix # AUX-ASPECT/MOOD

The *suffix* in the above construction would be either a subordinating (dependent) suffix or a nominalizer. Oluta (Mije) has /-E/ and Santa María Chimalapa (Sokean) has /-A/. Since there are both subordinators and nominalizers with both of these shapes, no clear answer to the identity of these suffixes is likely to be soon forthcoming, although subordinators often induce ergative shift, and the auxiliary construction does not. In general, Mije-Sokean languages do not use non-finite verb forms as the heads/nuclei of verb phrases, so the “subordinator” interpretation is more plausible.

In any case, no auxiliary constructions have been noted in known Epi-Olmec texts. In all Sokean languages the main verb takes a suffix *{-w#}* (probably the vetative suffix *{-w#<sub>2</sub>}*) when preceded by an auxiliary, but we suspect that auxiliaries were clause-final in Epi-Olmec times and that this Proto-Sokean structure dates from a post-Epi-Olmec period.

Sokean languages use different suffixes after auxiliaries or negatives than elsewhere. These suffixes can be roughly characterized as dependent markers. The *words* that encode negatives are themselves distinguished for complete/incomplete aspect, but do not mark it explicitly. While negative markers are preverbal in all Mije-Sokean languages, and have probably always been so, auxiliaries were probably originally postverbal and their largely or entirely preverbal distribution in most languages is an innovation. The suffix *\*-w#<sub>2</sub>* occurs in post-auxiliary and post-negative constructions in all Sokean languages (in these constructions, *\*-w#<sub>2</sub>* never encodes complete aspect).

### 5.3.10 Independent personal pronouns

Non-third person independent pronouns can be reconstructed for Proto-Mije-Sokean, and Proto-Sokean, as follows: *7#-tz* ‘I’; *tu-tz* ‘we (inclusive)’; *mi-tz* ‘you’. All of these

could be pluralized by combination with a pluralizing enclitic that does not have a single reconstructible form for Proto-Mije-Sokean or Proto-Sokean. *7u-tz* + pluralizer means ‘we (exclusive)’; *tu-tz* + pluralizer means ‘we (inclusive)’; *mi-tz* + pluralizer means ‘you all’. Not all Mije-Sokean languages have a third-person independent personal pronoun. Languages lacking a third-person pronoun use the demonstrative words in this function. Those languages that do have a third-person pronoun typically base it on the demonstrative *\*je7*.

## 6. LEXICON

The currently recognized Epi-Olmec vocabulary is almost exclusively Mije-Sokean in origin. This is not surprising. Mesoamerican languages generally are relatively resistant to lexical borrowing, and to date no Proto-Sokean, Proto-Mijejan, or Proto-Mije-Sokean words have a recognized foreign origin. In contrast, many other Mesoamerican language groups show reconstructible borrowings from Mije-Sokean. This reflects both cultural attitudes and the culturally prominent roles of Mije-Sokeans over other Mesoamericans in early intercultural interactions. There is one non-Mije-Sokean word identified in Epi-Olmec: /nup/ (spelled ⟨nu-pu⟩) ‘counterpart; the other member of a pair’, which is borrowed from Greater Lowland Mayan *\*nuhp* of the same meaning.

Much of the morpheme stock of Epi-Olmec is specifically Sokean. We have identified forty Sokean roots and nine Sokean grammatical morphemes spelled phonetically (in part or entirely): for example, *jama* ‘day’ spelled ⟨ja-ma⟩, *.na7* ‘stative’ spelled ⟨na⟩.

It is to be expected that some words that are now specifically Mijejan were lost in Sokean in the last few centuries before the Proto-Sokean stage, and therefore that a few such words might be attested in Epi-Olmec texts. We have securely read only two words that are at present restricted to Mijejan, *s#w* ‘sun’ and *w#7m.i* ‘nodding’. It was already known that *\*s#w* was the Proto-Mije-Sokean word for ‘sun.’ Commonly in Mesoamerica, a single word means ‘sun,’ ‘day,’ and ‘festival,’ which is the range of meanings of Proto-Mijejan *\*s#w*; Proto-Sokean has it with the meaning ‘festival,’ alongside innovated *\*jama* for ‘sun, day.’

The following vocabulary lists all Epi-Olmec lexical items that are spelled phonetically, whether by a fully phonetic spelling or by a logogram with a phonetic complement, as we had identified them by the fall of 2001. The forms are cited in pre-Proto-Sokean phonological garb, first in our practical orthography and afterward in IPA. The stage at which the lexeme is reconstructible within Mije-Sokean is also specified: pMS is Proto-Mije-Sokean; pS is Proto-Sokean. All reconstructions are in the form determined by Kaufman. TK marks cognate sets put together by Kaufman, along with the year it was recognized; SW marks cognate sets put together by Wichmann, all published in 1995.

Grammatical class information is specified as follows: adj = adjective; adv = adverbial; num = numeral; s = substantive; sr = relational noun; sv = verbal noun; vi = intransitive verb; vt = transitive verb; med = mediopassive; pc = participle; unacc = unaccusative.

**7ame7** (< pMS *\*7a:me7*) s year. Spelled ⟨DRUM/YEAR, DRUM/YEAR-me⟩. IPA: ?ame?. (TK 1963)

**7i** (> pS) pron:interr-indef who?. Spelled ⟨7i⟩. IPA: ?i. (TK 1963)

**7i7ps** (< pMS *\*7i:7ps*) num twenty. Spelled ⟨MOON/TWENTY-si₂⟩. IPA: ?i?ps. (TK 1963)

**7is** (< pMS *\*7is*) vt/unerg to see. Spelled ⟨7i-si-w#⟩. IPA: ?is. (TK 1963)

**7is** (< pMS *\*7is*) expl lo!, behold! Spelled ⟨7i-si, 7i-si₂⟩. IPA: ?is. (TK 1991)

**7otuw** (> pS) vi to speak. Spelled ⟨7o-tu-pa⟩. IPA: ?otuw. (SW 1991)

- 7owa** (< pMS \*7owa) s macaw, parrot. Spelled ⟨7o+wa, MACAW⟩. IPA: ?owa. (TK 1992)
- 7oy** (< pMS \*7oy) vi to go (and return), to take a trip. Spelled ⟨7o-7i⟩. IPA: ?oy. (TK 1963)
- jak** (> pS) vt/vi to cut. IPA: hak. (TK 1963)
- ku.jak** (> pS) vi to cross over. Spelled ⟨ku-CROSS-wu⟩. IPA: ku.hak. (TK 1993)
- jam** (> pS \*jam.uj) vt to remember. IPA: ham. (TK 1963)
- kuk=tza7=jam.e** s pectoral stone memento. Spelled ⟨7i-ku-MIDDLE-tza<sub>2</sub>-ja-me⟩. IPA: kuk=tsa?=ham.e. (TK 1963 + TK 1991)
- jama** (> pS) s day; shape-shifter's animal guise. Spelled ⟨ja-ma, ja<sub>2</sub>-ma, GUISE<sub>1</sub>, GUISE<sub>2</sub>⟩. IPA: hama. (TK 1963 + TK 1991)
- jay7** (< pMS \*ja:y7) vt/vi to write. IPA: hay?. (TK 1963)
- jay7=ki7ps.i** sv < v+v inscribed monument. Spelled ⟨ja<sub>2</sub>-SYMBOL-si<sub>2</sub>⟩. IPA: hay?=ki?ps.i.
- je7** (< pMS \*je7) dem that. Spelled ⟨je⟩. IPA: he?. (TK 1963)
- je7.tzu** (> pS) dem thus, like that. Spelled ⟨je-tzu⟩. IPA: he?.tsi. (TK 1992)
- jej** (> pS) vi to live, breathe. IPA: heh. (TK 1963)
- jej.a7** (> pS) sv < vi god: 'living one'. Spelled ⟨GOD-ja⟩. IPA: heh. a?. (TK 1994)
- ji7tz** (> pS) vt/vi to (get) wrinkle(d). IPA: hi?ts. (TK 1994)
- ji7tz.i7** (> pS) sv < vt wrinkled, pleated; earthly Longlip god ("wrinkled one"). Spelled ⟨ji-LONGLIP<sub>2</sub>, ji-tzi⟩. IPA: hi?ts.i?. (TK 1994)
- joj** (< pMS \*jo:t? > pS \*joj) sr inside. Spelled ⟨jo⟩. IPA: hoh. (TK 1963)
- jome7** (< pMS \*jome7) adj new. Spelled ⟨jo-me-NEW⟩. IPA: home?. (TK 1963)
- jup** (< pMS \*jup) vt to cover. IPA: hup. (TK 1992 + SW 1991)
- ni7.jup.7** sv body-covering. Spelled ⟨?LOINCLOTH-pu⟩. IPA: ni?.hup.?.
- ju7ps** (< pMS \*ju7ps) vt to lash, to tie something onto something else. IPA: hu?ps. (SW 1991)
- 7owa-ju7ps.i** sv macaw-lashing. Spelled ⟨7o+wa=ju/LASH-si⟩. IPA: ?owa=hu?ps.i.
- jus** (< pMS \*jus) adv after[ward]. Spelled ⟨AFTER, AFTER-?su⟩. IPA: his. (SW 1991 + TK 1993)
- kak** (> pS) vt/unacc to (get) replace(d). Spelled ⟨kak-wu⟩. (SW 1991)
- kak.e** sv < vt exchange, replacement. Spelled ⟨kak⟩.
- kak.u7** (> pSoke \*kak.u7) sv < vt replacer, successor. Spelled ⟨kak⟩. IPA: kak-i?. (TK 1992)
- kakpe7** (< pMS \*kakpe7) s scorpion; Scorpius. Spelled ⟨kak-SCORPIUS-pe⟩. IPA: kakpe? (TK 1963)
- ken** (> pS) vt to see. (TK 1993)
- ken.e** (> pSoke) pcp < vt seen: visible, public. Spelled ⟨ke-ne⟩. (TK 1993)
- kij** (< pMS \*kij) vi to give light, shine. Spelled ⟨SHINE-wu, ki-wu⟩. IPA: kih. (SW 1991)
- ki7m** (> pS) vi to go up, accede. Spelled ⟨7i-GO.UP⟩. IPA: ki?m. (TK 1963)
- ki7m.u7** (> pS) sv < vi accession, rising, ascent, installation. Spelled ⟨GO.UP⟩. IPA: ki?m.i? (SW 1991)
- 7aw=ki7m** (> pS) vi to rule. IPA: ?aw=ki?m. (SW 1991)
- 7aw=ki7m.u7** (> pS) sv < vi rule (rship). Spelled ⟨7aw-GO.UP, 7aw-GO.UP-mu⟩. IPA: 7aw=ki?m.i?. (TK 1991)
- ko.ki7m.i(7)** sv < vi accession, or one.who.accedes on.behalf.of.others/elsewhere. Spelled ⟨ko-ki-mi-GO.UP⟩. IPA: ko.ki?m.i(?).
- kip** (< pMS \*kip) vt to fight against. Spelled ⟨7i-ki-pi-wu⟩. (SW 1991)
- ki7ps** vt to try, test, think (TK 1963)



- ki7ps.i** (< pMS \*ki7ps.i) sv < vt symbol(stone) [celt, figurine, stela; badge, token, memento, souvenir]. Spelled ⟨SYMBOL-si<sub>2</sub>⟩. IPA: ki?ps.i. (TK 1963 + TK 1992)
- kom** (< pMS \*kojm) s notched house-post, pillar. (TK 1963)
- tzap(=?win)=kom** s type s sky(-?face) pillar. Spelled ⟨ko-SKY-?FACE-PILLAR⟩. IPA: tsap(=?win)=kom.
- ko.nu7ks** (< pMS \*ko.nu:7ks) vt to bless, hallow. Spelled ⟨HALLOW-wu⟩. (SW 1991)
- ko.nu7ks.i** (< pMS \*ko.nu7ks.i) pcp < vt blessed, hallowed. Spelled ⟨HALLOW-si<sub>2</sub>⟩. (SW 1991 + TK 1992)
- ko.yumi** (> pS) s lord. Spelled ⟨ko-LORD-mi⟩. (TK 1963 + TK 1994)
- kot** (> pS) vt to put away. Spelled ⟨7i-ko-te⟩. (TK 1963)
- kuk** (< pMS \*kuk) s middle. (TK 1993)
- kuk=tza7=jam.e** s pectoral stone memento. Spelled ⟨7i-ku-MIDDLE-tza<sub>2</sub>-ja-me⟩. IPA: kuk=tsa?=ham.e.
- kuw** (> pSoke) vt to raise; to put up/away. (SW 1991 + TK 1994)
- kuw.na7** adv < vt raised; put up/away. Spelled ⟨ku-na⟩. IPA: kuw.na?. (TK 1992)
- kū7** (< pMS \*kū7) s hand, arm. Spelled ⟨na-kū⟩. IPA: ki?. (TK 1963)
- kūw7** (< pMS \*kū:w7) vt/unacc to (get) dye(d). Spelled ⟨kū-wu⟩. IPA: kiw?. (SW 1991 + TK 1993)
- kuy7** (< pMS \*kū:y7) vt/unacc to (get) cover(ed). IPA: kiy?. (SW 1991)
- ko.wu7tz=kuy7** vt + vt/unacc to get pierced and covered for others. Spelled ⟨ko-PIERCE-kū-wu⟩. IPA: ko.wu?ts=kiy?. (SW 1991)
- ma** (< pMS \*ma) adv earlier. Spelled ⟨ma⟩. (TK 1963)
- mak** (< pMS \*mak) num ten. (TK 1963)
- mak tzap** num+s Ten Sky (a god). Spelled ⟨ma-FIVE-FIVE-SKY⟩. IPA: mak tsap.
- masa(n)** (< pMS \*ma:san > pS \*masan ~ masa=) s/adj holy (thing), god. (TK 1963 + TK 1992)
- masa=wik.i** sv something/someone hallowed by sprinkling. Spelled ⟨ma-sa-SPRINKLE-ta-ma⟩.
- masa=ni7.APPEAR** vi to appear divinely on the body. Spelled ⟨ma-sa-ni-APPEAR-wu⟩.
- matza7** (< pMS \*ma:tza7) s star. Spelled ⟨ma-STAR-tza⟩. IPA: matsa?. (TK 1963)
- may** (< pMS) vt/unacc to count. Spelled ⟨ma-wu, ma-wu<sub>2</sub>⟩. (TK 1963)
- metz=** (< pMS \*metz=) num by twos. Spelled ⟨me-tze⟩. IPA: mets=. (TK 1963)
- mi7ks** (< pMS \*mi7ks) vi to quiver. Spelled ⟨7i-BLOOD-mi-si<sub>2</sub>, mi-si-na-wu⟩. IPA: mi?ks. (SW 1991)
- mon7** (< pMS \*mon7) vt to wrap. IPA: mon?. (TK 1963)
- RULER=ko7=mon7.a7** sv ruler's head-wrap. Spelled ⟨“KNOT”+“GOVERNOR”-7a⟩.
- mutz** (> pS) vi to play. IPA: mits. (TK 1963)
- mutz.i7** (> pS) sv < v impersonator. Spelled ⟨PLAY-tzi⟩. IPA: mitz.i7. (TK 1994)
- nas** (< pMS \*nas) vi to pass. Spelled ⟨na-sa-wu⟩. (TK 1963)
- ne7k** (> pS) vt to set aside. IPA: ne?k. (SW 1991)
- ne7k.e** (> pS) pcp < vt set aside. Spelled ⟨ne-ke⟩. IPA: ne?k.e. (SW 1991 + TK 1993)
- ne7w** (< pMS \*ne7w) vt to set stones in a row/wall/circle. Spelled ⟨7i-ne-ji⟩. IPA: ne?w. (SW 1991)
- ne7w.e** sv/pcp (stones) set in order. Spelled ⟨ORDER.STONES = ne⟩. IPA: ne?w.e.
- nip7** (< pMS \*ni:p7) vt/unacc to plant, sow; bury. Spelled ⟨PLANT-pi-wu⟩. IPA: nip?. (TK 1963)
- nip7.i** (< pMS \*ni:p7.i) sv < vt planting, planted (thing). Spelled ⟨PLANT, PLANT-7i⟩. IPA: nip?.i. (SW 1991)



- nup** (< Greater Lowland Mayan *nuhp*) s counterpart, one of two members of a pair. (TK 2001)
- nuks** (< pMS \**nuks*) vi to go along. Spelled ⟨7i-nu-si⟩. IPA: niks. (TK 1963)
- pak** vt to beat (TK 1992)
- pak.kuy7** (> pS) sv bludgeon. Spelled ⟨pak-ku⟩. IPA: pak.kuy?.
- paki7** (> pS) adj hard, strong, powerful. Spelled ⟨pa-ki⟩. IPA: paki?. (SW 1991)
- pey** (> pS \**pey.e7*) vt unacc to get waved/swung. Spelled ⟨pe-wu⟩. (SW 1991 + TK 1993)
- pini7** (> pSoke \**pini7*) s brother-in-law of man. Spelled ⟨pi-ni⟩. IPA: pini?. (TK 1963)
- pit** (< pMS \**pit*) vt to tie (in a bundle). (TK 1963)
- pit.i** sv < vt bundle; prisoner. Spelled ⟨TIE, TIE-ti⟩.
- poy7a** (< pMS \**poy7a*) s moon; month, veintena. Spelled ⟨po-7a, MOON, MONTH⟩. IPA: poy7a. (TK 1963)
- puw** (> pS) vt/unacc to (get) scatter(ed). Spelled ⟨pu-wu⟩. (TK 1991)
- puk** (< pMS \**puk*) vt/unacc to (get) take(n)/acquire(d)/achieve(d). Spelled ⟨puk-7i, puk-ku-7i, puk-ku-wu⟩. IPA: pik. (TK 1963)
- sa7.sa7** (> pS) adj noble, healthy
- sa7.sa7** (=p~~un~~) (> pSoke) s noble, aristocrat. Spelled ⟨sa<sub>2</sub>-NOBLE<sub>2</sub>⟩. IPA: sa7.sa7 (=pin). (SW 1991 + TK 1993)
- saj** (< pMS \**saj*) vt/unacc to (get) share(d) out. Spelled ⟨7i<sub>2</sub>-sa<sub>2</sub>+pa<sub>2</sub>, saj-wu⟩. IPA: sah. (SW 1991)
- saj** (< pMS \**saj*) s wing; shoulder. Spelled ⟨7i-sa⟩. IPA: sah. (TK 1963)
- si7i7** (> pS) s backside, butt. Spelled ⟨si<sub>2</sub>-7i⟩. IPA: si7i7?. (TK 1992–1994)
- su7ksu7** (< pMS \**su7ksu7*) s hummingbird. Spelled ⟨?su×?su⟩. IPA: su7ksu7?. (SW 1991)
- te7** (< pMS \**te7*) dem the aforementioned; the latter; it; that. Spelled ⟨te⟩. IPA: te7?. (TK 1992)
- te7n** (< pMS \**te:7n*) vi to stand (on tip-toe), to step (on). IPA: te7n. (TK 1963)
- te7n.na7** (> pS) adv upright(ly), on tip-toe. Spelled ⟨te-ne-na⟩. IPA: te7n.na7?. (TK 1963 + TK 1992)
- tok** (> pSoke) vt to stain. (TK 1993)
- tok.e** (> pS) pcpc stained. Spelled ⟨to-ke⟩. (TK 1993)
- tokoy** (< pMS \**tokoy*) vi to be lost. (TK 1963)
- ko.tokoy-pu7** vt to cover up/over OR to spill.on.behalf.of.others/elsewhere. Spelled ⟨7i-ko-LOSE-pu-wu⟩. IPA: ko.tokoy-pi7?. (TK1963 + TK 1992)
- (yak>)tokoy.a7** sv < v overturning/upsetting/dumping OR overturner/upsetter/dumper etc. Spelled ⟨7i-LOSE-ya⟩. IPA: (yak>)tokoy.a7?. (TK 1963 + TK 1992)
- (yak>)tokoy.e** sv < v overturned/upset/dumped one. Spelled ⟨na-LOSE-ye⟩. (TK 1963 + TK 1992)
- tuk** (> pS) vi to happen. Spelled ⟨tuk×pa⟩. (TK 1963)
- tuk** (< pMS \**tuk*) vt to cut, harvest. (TK 1963)
- tuk.u7** sv < vt harvester. Spelled ⟨tuk⟩. IPA: tuk.i7?
- tuk.u7** pcpc < vt having been cut. Spelled ⟨tu-ku⟩. IPA: tuk.i7?
- wu=tuk.i** sv well-harvested (thing). Spelled ⟨wu-tuk?⟩. IPA: wi =tuk.i.
- tuki** (> pS) s water turtle. Spelled ⟨TURTLE-ki⟩. (TK 1963)
- tu7ki** (> pGulf Sokean \**tu:7ki*) s trogon, quetzal. Spelled ⟨TROGON⟩. [Final /i/ implied by omission of following ⟨7i⟩.] IPA: tu7ki. (TK 1997)
- tuku7** (< pMS \**tuku7*) s cloth, garment. Spelled ⟨CLOTH, tu+CLOTH⟩. IPA: tuku7?. (TK 1992)
- tus** (> pS) vt to prick, sting. (TK 1963)
- tus.i** (> pS) adj < vt with hair standing on end. Spelled ⟨tu-si⟩. (TK 1992)

**tup** (< pMS \**tup*) vt to pierce with a shafted or shaft-shaped piercer. Spelled ⟨tu-nu-“DEAL.WITH”-pu-ja-wu⟩, ⟨tu-nu-“DEAL.WITH”-pu-ja-yaj-wu⟩. IPA: tip. (TK 1963)

**tza7** (< pMS \**tza:7*) s stone. IPA: tsaʔ. (TK 1963)

**kuk=tza7=jam.e** s pectoral stone memento. Spelled ⟨7i-ku-MIDDLE-tza<sub>2</sub>-ja-me⟩. IPA: kuk=tsaʔ=ham.e.

**tza7yji** (> pSoke) adv late in the day. Spelled ⟨tza<sub>2</sub>-ji⟩. IPA: tsaʔyhi. (SW 1991 + TK 1992)

**tzap** (< pMS \**tzap*) s sky. Spelled ⟨SKY, SKY-pa⟩. IPA: tsap. (TK 1963)

**mak tzap** num+s Ten Sky (a god). Spelled ⟨ma-FIVE-FIVE-SKY⟩. IPA: mak tsap.

**tzat7** (< pMS \**tzat7*) vt to measure by hand-spans. IPA: tsatʔ. (SW 1991)

**tzat7.u7** sv hand-span measuring device. Spelled ⟨7i<sub>2</sub>-ni<sub>2</sub>-ko-SPAN-7u<sub>2</sub>⟩. IPA: tsatʔ.iʔ.

**nu.tzat7.e=nas** adj <vt + s ground jointly measured by hand-spans. Spelled ⟨nu<sub>2</sub>-SPAN=EARTH⟩. IPA: ni .tsatʔ.e.

**tzetz** (< pMS \**tzetz*) vt to chop (off). Spelled ⟨na-tze+tze-CHOP-ji⟩. IPA: tsets. (SW 1991)

**tzetz.e** sv <vt chopped off (thing). Spelled ⟨na-tze+tze⟩. IPA: tsets.e.

**tzuk** (> pSoke) vt to do (< ?to touch). Spelled ⟨DO-pa⟩. IPA: tsik. (TK 1963)

**tzuk.i=pun** sv deedsman. Spelled ⟨DO<sub>2</sub>×ki⟩. IPA: tsik.i=pin.

**tzusi** (> pS) s child under 12. Spelled ⟨tzu-si<sub>2</sub>⟩. IPA: tsisi. (TK 1963)

**wan** (> pS) vt/vi to sing. (TK 1963)

**wan.e** (> pS) sv <vi song, chant. Spelled ⟨SING-ne⟩. (TK 1963)

**wan.e=tzuk** (> pSoke) vi:incorp to perform a chant. IPA: wan.e=tsik. (TK 1963 + TK 1992)

**wej** (> pS) vi to shout. Spelled ⟨we-pa⟩. IPA: weh. (TK 1963)

**wen.e7** (> pS) sv <vt (something) broken, piece. Spelled ⟨we-ne⟩. IPA: wen.eʔ. (SW 1991 + TK 1992)

OR **we7n.e** (< pMS \**we:7n.e*) sv <vt a few, some. Spelled ⟨we-ne⟩. IPA: weʔn.e. (SW 1991 + TK 1992)

**wik** (> pS) vt/unacc to (get) sprinkle(d). (TK 1992)

**ko.wik** (> pS) vt/unacc to (get) sprinkle(d) for.others/elsewhere. Spelled ⟨7i-ko-SPRINKLE-ki-pa, ko-SPRINKLE-ki-pa⟩. (SW 1991 + TK 1992)

**wik.i** (> pS) sv <vt result of sprinkling. Spelled ⟨SPRINKLE⟩. (SW 1991 + TK 1992)

**masa=wik.i** sv something/someone hallowed by sprinkling. Spelled ⟨ma-sa-SPRINKLE ta-ma⟩. (TK 1992)

**win** (< pMS \**win*) sr in front. Spelled ⟨wi-BEFORE⟩. (TK 1963)

**wis** (< pMS \**wis*) vt to uproot. (SW 1991)

**wis.i** pcpc <vt uprooted. Spelled ⟨UPROOT-si<sub>2</sub>⟩ or ⟨wi<sub>2</sub>-si<sub>2</sub>⟩.

**wo7m** (> pS) vi to sprout. (SW 1991)

**wo7m.a7** (> pS) s sprout. Spelled ⟨?wo-ma⟩. IPA: woʔm.aʔ. (SW 1991 + TK 1993)

**wu** (> pS) adj good. IPA: wi. (TK 1963)

**wu=tuk.i** sv <vt well-harvested (thing). Spelled ⟨wu-tuk⟩. IPA: wi=tuk.i.

**wu7m** (< pMS \**wu7m*) vi to nod (SW 1991 + TK 2000)

**wu7m.i** pcpc nodding. Spelled ⟨wu-mi⟩. IPA: wiʔm.i.

**yaj** (> pS [elite]) vi:med to be finished. Spelled ⟨yaj-7i⟩. IPA: yah. (TK 1963)

**yu7** (< pMS \**yu7*) dem this. Spelled ⟨yu⟩. IPA: yiʔ. (TK 1963)

## 7. THE PLACE OF EPI-OLMEC IN THE MIJE-SOKEAN FAMILY

### 7.1 Epi-Olmec and innovations in the diversification of Sokean

The etymological sources of Epi-Olmec vocabulary (see §6), along with a number of grammatical traits, show that the Epi-Olmec language belongs to the Sokean branch of the Mije-Sokean family.

In some ways Epi-Olmec is less evolved than Proto-Sokean as it can be reconstructed from its surviving daughters, as attested (i) *lexically* by *suw* ‘sun’ (see §6); (ii) *morphologically* by {-ji} ‘dependent completive’ and by the apparently productive use of {.A7} for agentive nominalization; and (iii) *phonologically* by /7/ between C and V, by /p/ after /k/, and by the maintenance of original vowel length (if real). (In fact, the phonological system of Epi-Olmec cannot be distinguished from that of Proto-Mije-Sokean or Proto-Mijeian.) Accordingly, when Epi-Olmec agrees specifically with one subgroup of Sokean and diverges from the other, the straightforward conclusion is that the divergent branch is innovative.

The features that are most telling are grammatical; the straightforward cases are provided in Table 44.3 :

**Table 44.3 Grammatical innovations in Sokean subgroups**

Features	Epi-Olmec	Proto-Sokean	Gulf Sokean	Soke
<b>GULF SOKEAN IS CONSERVATIVE</b>				
1. who?	<7i>	*7i	SOT 7iH	MAR 7i-wu
2. what?	<ti>	*ti	SOT tyiH	MAR ti-yu
3. perfect	<na>	*-nay7	SOT {-ne7}	—
4. that	<je>	*je7	SOT je7 ‘s/he’	—
5. like	<tzu>	*+tzu	SOT {-tzu}	—
			frozen	
6. away	<ku>	*ku.	SOT {ku+}	MAR {ku.}
<b>SOKE IS CONSERVATIVE</b>				
7. this	<yu>	*yu7	(SOT yu7p)	MAR yu7
8. the	<te>	*te7	—	MAR te7
9. from	<ku>	*+k.	SOT {-k}	COP {+k}
			frozen	
10. when (REL.)	<ku>	*+7ku	SOT {-k}	COP {+7k}
			frozen	MAG {+7uk}
				MAR {+k(u)}
11. relativizer	<wu>	*+wu7	—	COP {+wu7}
				MIG {+V7k}
12. completely	<pu>	*-pu7	—	COP {-pu7}
13. on the body	<ni>	*ni7.	—	MAR {ni7.}
14. stative	<na>	*.na7	—	MAR {.na7}
15. word order	SOV	*SOV	VSO	MAR SOV vs. COP VOS

In Sotepan Gulf Sokean, the symbol H is used to transcribe a morphophoneme that sometimes is realized as vowel length (see §3.2), sometimes as vowel length followed by /j/, and sometimes as nothing. It occurs in a fair number of morphemes, both roots and suffixes; the particular phonological realization is determined by phonological context. The designation “frozen” for Gulf Sokean means that the morpheme occurs in a small number of lexical items, and is not productive, whereas in the corresponding Soke forms the morpheme occurs productively.

## 7.2 Elite innovations

Certain Epi-Olmec words – for example, *kak* ‘to replace,’ *kuw* ‘to raise, to put up/away’ – are now attested only in Chiapas Soke, even though our documentation of other Sokean languages is even more extensive. This suggests the hypothesis that Chiapas Soke maintains some elite forms that have been lost elsewhere. Evidence from three grammatical morphemes, the optative suffix and two pluralizers, supports this hypothesis.

### 7.2.1 Sokean optative

The optative suffix is spelled ⟨7i⟩ in Epi-Olmec texts (see §4.3.2). This complicates our picture of Sokean morphology. The optative is *-7i* in Chiapas Soke, but the Proto-Sokean form was apparently *\*-7in*; the presence of the final /n/, for which there is otherwise no straightforward source, is indicated by Santa María Chimalapa Soke *-7in* and Soteapan Gulf Sokean *-7iny*.

One logical possibility is that /n/ was somehow lost in Chiapas Soke. But Epi-Olmec predated the Proto-Sokean phase, and it is implausible that ⟨7i⟩ is spelling /7in/; Epi-Olmecs never failed to spell /n/ (or any other non-weak consonant, except /p, k/ before /s/). The only straightforward solution is that Proto-Sokean had both *\*-7in* and *\*-7i*, in conditioned or free variation. If the elite variety (represented earlier in Epi-Olmec texts) preferred /-7i/, while the lower class variety preferred /-7in/ (which may have been more archaic), this would allow both variants and would agree roughly with the hypothesis, for which there is other evidence, that Oaxaca Soke is conservative and Chiapas Soke is at least partly descended from an elite variety of Sokean.

### 7.2.2 Mije-Sokean plural markers

In two grammatical features, Epi-Olmec agrees with innovations in Gulf Sokean and Chiapas Soke and diverges from conservative forms in Oaxaca Soke.

As noted in §4.2.2, third-person plural markers are always recruited from the particular language’s verb root meaning ‘to be finished’: Mijean *\*kux*, Oaxaca Soke *\*suk*, Epi-Olmec and other Sokean *\*yaj*. The remaining plural markers – which are not also known to be roots – are as follows (without their precise functions, which vary from language to language):

- {\*-ta7m} (Soke, SOT, OLU, LLM)
- {\*+tuk} (OLU, LLM, MAR, MIG)
- {\*jate7} (OLU, SAY, MIG)

Among the Sokean languages, Santa María Chimalapa Soke and San Miguel Chimalapa Soke often differ from the rest and agree instead with Mijean. This might be due either to contact with Mijean, or to conservatism, since certain features common to most Sokean languages might be elite Epi-Olmec innovations preserved in surviving Gulf Sokean (Soteapan Gulf Sokean, Texistepec Gulf Sokean, Ayapa Gulf Sokean) and Chiapas Soke but not in Oaxaca Soke (see the preceding discussions).

On the latter assumption, we would reconstruct the following for Proto-Mije-Sokean:

- {\*-ta7m} first- and second-person plural agreement: S/O/P
- ‘to be finished’ third-person plural agreement: S/O/P
- {\*+tuk} noun plural (survives in MAR, MIG, LLM, OLU)
- {\*jate7} ‘each, several’ (used to pluralize nouns and adjectives in SAY, pronouns in MIG, first- and second-person subjects and objects in OLU)

A set of elite Epi-Olmec innovations would then be:

animate noun plural  $\Rightarrow \{+ta7m\}$   
 inanimate noun plural  $\Rightarrow \{+yaj\}$

Epi-Olmec attests these Proto-Mije-Sokean traits plus the postulated elite innovations, except that no first or second person plural agreement-marker is attested.

### *Acknowledgments*

Our research on Epi-Olmec decipherment, and on the documentation and comparative analysis of present-day Mije-Sokean languages, has been supported mainly by the National Science Foundation (grants SBR-9411247, SBR-9511713, and SBR-9809985), the Salus Mundi Foundation, and the National Geographic Society (4190–92 and 5319–94). We thank Peter Keeler, director of the Maya Meetings at Texas, for supporting our participation and research on Epi-Olmec writing at these annual workshops. Our work was also supported by the Faculty Research Awards Program (grant 320-9753P) at the State University of New York at Albany, and Justeson received research support from the Department of Artificial Intelligence at IBM's T. J. Watson Research Center. We have received smaller grants from the University of Pittsburgh and SUNY at Albany. We thank Roberto Zavala Maldonado for useful comments on this study.

### **Phonological transcriptions**

This paper uses a practical orthography with the following IPA equivalents:

	p	t	tz	k	ʔ	j	s	m	n	nh	w	x	y
IPA:	p	t	ts	k	ʔ	h	s	m	n	ŋ	w	š	y
	i	e	ɛ	a	u	o							
IPA:	i	e	ɛ	a	u	o							

### **Morphological transcriptions**

In morphologically explicit representations of Mije-Sokean words, inflexional affixes are marked by -, clitics by +, derivational affixes by ., class-changers by >, and compounding morphemes by =.

### **Example sentence format**

T	Transcription
	uncertainty that a sign's identification is correct is indicated by a postposed question mark
	less secure readings or interpretations of signs are indicated by a preposed question mark
R	pre-Proto-Sokean Reading
	questions and doubts are marked in line T, not here
G	morpheme-by-morpheme Gloss (morphemes are separated by hyphens)
FT	Free (but still fairly literal) Translation

## Grammatical codes

### *Absolutive person markers*

- XA exclusive absolutive: {7u+}  
 3A third person absolutive: {Ø}

### *Ergative person markers*

- XE exclusive ergative: {na+}  
 IE inclusive ergative: {tun+}  
 2E second person ergative: {7in+}  
 3E third person ergative: {7i+}

### *Derivational prefixes on verbs*

- AWAY away: {ku.}  
 MOUTH with the mouth: {7aw=}  
 BODY on the body: {ni7.}  
 ELSE in someone else's place; elsewhere, for someone else: {ko.}  
 ASSOC together, jointly: {nu.}  
 CAU causative {yak>}

### *Verb suffixes*

- ENTIRELY entirely: {.pu7}  
 NDIR indirective: {>jay7}  
 PRF perfect: {-nay7}

### *Plural person marking suffixes*

- AP animate plural: {+ta7m}  
 IP inanimate plural: {+yaj}  
 3P third person plural: {-yaj}

### *Aspect/mood suffixes*

- II independent incomplete: {-pa}  
 DI dependent incomplete: {-e} ~ {-i}  
 IC independent complete: {-wu}  
 DC dependent complete: {-ji}  
 OPT optative: {-7i}

### *Stative-deriving suffix*

- STAT stative: {.na7}

### *Noun-deriving suffixes*

- PN passive nominalization: {.E}, {.E7}, {.A7}  
 AN active nominalization: {.A7}, {.E7}  
 NSTR instrument noun: {.kuy7}, {.7}

### *Locative enclitics*

- LOC locative: {+mu7}  
 FROM from: {+k}

*Subordinating enclitics*REL        relativizer: {+wu7}WHEN    when: {+7ku}**Bibliography**

For our discussion of comparative Mije-Sokean linguistics and the features of individual Mije-Sokean languages, we have drawn upon the mostly unpublished results of field work by linguists working with the Project for the Documentation of the Languages of Meso-America. For Oluta Mijean and San Miguel Chimalapa Soke, some of this data is published online at <http://www.albany.edu/pdlma/>. The Project has lexical databases for all of the Mije-Sokean languages, as follows (language abbreviation, linguist(s) responsible): MIG (Heidi Johnson and Terrence Kaufman), MAR (Terrence Kaufman), COP (Clifton Pye and Terrence Kaufman), TEC (Roberto Zavala), AYA (James Fox and Giulia Oliverio), TEX (Catherine Berezna and Ehren Reilly), SOT (Terrence Kaufman and Valerie Himes), OLU (Roberto Zavala), SAY (Richard Rhodes and Dennis Holt), GUI (Giulia Oliverio), TOT (Daniel Suslak).

- Campbell, L. and T. Kaufman. 1976. "A linguistic look at the Olmecs." *American Antiquity* 41:80–89.
- Justeson, J. and T. Kaufman. 1993. "A decipherment of epi-Olmec hieroglyphic writing." *Science* 259:1703–1711.
- . 1996 [1992]. "Un desciframiento de la escritura epi-olmeca: métodos y resultados." *Arqueología* 8:15–25.
- . 1997. "A newly-discovered column in the hieroglyphic text on La Mojarra Stela 1: a test of the epi-Olmec decipherment." *Science* 277:207–210. (An extended version appears in *Science Online* at [www.sciencemag.org/features/data/justeson.shl/](http://www.sciencemag.org/features/data/justeson.shl/).)
- Kaufman, T. 1963. "Mixe-Zoque diachronic studies." Unpublished monograph.
- Kaufman, T. and J. Justeson. 2001. *Epi-Olmec Hieroglyphic Writing and Texts*. Workshop of the XXV Maya Meetings at Texas. Online at <http://www.albany.edu/pdlma/EOTEXTS.pdf>.
- Kelley, D. 1993. "The decipherment of epi-Olmec writing as Zoquean by Justeson and Kaufman." *The Review of Archaeology* 14(1):29–32.
- Macri, M. and L. Stark. 1993. *A Sign Catalog of the La Mojarra Script*. Pre-Columbian Art Research Institute, Monograph 5. San Francisco: PARI.
- Pool, C. 2000. "From Olmec to epi-Olmec at Tres Zapotes." In J. Clark and M. Pye, eds., *Olmec Art and Archaeology in Mesoamerica*. Studies in the History of Art, 58. Washington, D.C.: National Gallery of Art.
- Wichmann, S. 1995. *The Relationship among the Mixe-Zoquean Languages of Mexico*. Provo: University of Utah Press.



# Zapotec

Most of what is widely understood about Zapotec (Sapoteko) writing is not linguistic. This knowledge consists of identifications of the signs for numerals from 1 to 19, which are in the same bar-and-dot system as in the rest of Mesoamerica (changed to dots only, later, in central Mexico), and the signs for names of most of the twenty days in the ritual calendar, and, based on it, the fifty-two-year cycle of named years. The current reliable results on the calendar are due mostly to Javier Urcid (1992, 2001), building on work by Alfonso Caso (1928, 1947). Using Urcid's results, Justeson and Kaufman (1996) worked out that the numerical coefficients of a glyph that Caso called *Glyph W* indicate the position of an associated ritual calendar date within a lunation.

Zapotec texts were generally written in columns from top to bottom. Columns were read from left to right, and the signs in these cases face leftward. Above rightward-facing animate beings in accompanying iconography, texts were read from right to left and their signs face rightward; this deviation also usually occurs when columnar text is adjacent to such figures (Kaufman and Justeson 1993–2003).

Almost everything else thus far worked out in the script has taken calendrics as a starting point. People were named for the ritual calendar date on which they were born, so when figures in scenes are accompanied by what seem to be dates, these dates are often to be taken as names of those people (the same practice is found in later Mexican pictorial books, which are famous for their so-called narrative pictography; they have no connected glyphic texts, only glyphic captions for people, places, and dates).

One exception is provided by a set of signs and sign groups occurring above a logogram representing 'hill' on a set of wall panels known as "conquest slabs." Caso (1947) proposed that these sign groups were place names, because their context and arrangement resembled those of Nahuatl place name glyphs. Marcus (1976) and Whittaker (1980) built on this proposal and attempted to identify specific sign groups with specific places. Justeson and Kaufman (1996) found, to the contrary, that none of these sign groups name places; instead, they give noncalendrical personal names and titles of individuals involved in warfare. Urcid (1992) has identified other glyphic place names, outside the conquest slabs; these names are spelled by glyph groups that are overlaid on the logogram for 'hill.' We agree with Urcid's interpretation of these data and find the same pattern with additional glyphs besides (HILL) that represent places.

Most researchers believe these texts are written in Zapotec because the distribution of the script generally (though not exactly) matches the distribution of Zapotec speech at the time of the Spanish invasion of Mesoamerica. In fact, the very earliest texts in the script are located near the center of the Zapotec region, in the Valley of Oaxaca. Day names occur before their numeral coefficients, an order documented only in the Zapotec calendar (Whittaker 1980). Urcid (1992) observed that two of the day signs relate specifically to Zapotec: Zapotecs had

a day named *Knot*, agreeing with the form of a day sign that Caso (1928) recognized as depicting a knotted cloth, and a day named *Corn*, agreeing with the form of a day sign that Urcid recognized as depicting an ear of corn.

Our joint research on the decipherment of Zapotec writing, begun in 1992, is the first to use a systematic linguistic approach based on detailed documentation and reconstruction of Zapotec and other Mesoamerican languages. This has helped us to advance the interpretation of the calendrics, and enabled us to make the first reliable readings of phonetic signs – that is, readings that are supported by distinct semantic and grammatical contexts. We have also identified about half of the most frequent grammatical morphemes that would be expected in a text (Kaufman and Justeson 1993–2003), though nothing as comprehensive or decisive as in the Epi-Olmec case.

These results from the analysis of the grammar of Zapotec inscriptions have been based on Kaufman's models for Proto-Zapotec and Proto-Zapotecan grammar and vocabulary (Kaufman 1988, 1995). Our language documentation project (PDLMA, Project for the Documentation of the Languages of Meso-America; <http://www.albany.edu/pdlma/>) has gathered data on ten Zapotecan languages for the purpose of reconstructing the vocabulary and grammar of Proto-Zapotecan and Proto-Zapotec. As we now know it, some of the main features of the Proto-Zapotec and Proto-Zapotecan languages are as follows:

*Phonology.* All syllables in these languages are of the shape CV (except that *\*k* – in some cases originally probably some kind of deictic enclitic – occurred at the end of some stems; see Kaufman 2000). The vowel can be short, long, broken/squeezed, or checked. Syllable-initial consonants can be single or geminate (except that *\*y* and *\*w* are never geminate, and the marginal loaned phoneme *\*m* is always geminate). Tone is phonemic.

*Morphology.* Verbs obligatorily take one of several aspect-mood proclitics, none of which has a zero shape or allomorph. Most verb roots are consonant-initial, but some are vowel-initial; unpredictable allomorphs of the completive and potential proclitics define four verb classes. There are no pronominal agreement-markers; when personal pronouns occur, they do so in the same syntactic slot as any other noun. Pronouns distinguish gender and social status.

*Syntax.* Word order is VS(O). Sense permitting, any transitive verb that can occur in a VSO frame can also occur without an object. Nuclear case/role categories are unmarked, except by position with respect to the predicate: thus, there is no definitive syntactic evidence for accusativity, ergativity, or agentivity.

Using chronological data to divide up texts into small chunks for analysis, Kaufman and Justeson (1993–2003) have analyzed the structure of about 50 Zapotec texts, exploiting the reconstructed vocabulary and grammar of Proto-Zapotec and Proto-Zapotecan, and have found that the structure of the texts conforms to the structure of these languages. This is clearest in the matter of word order and verb morphology, but also in a number of other details.

The most systematic and compelling results of this analysis involve the identification of a number of grammatical morphemes: three third-person pronouns *\*k<sup>w</sup>e* 's/he: adult nonforeigner', *\*ne* 's/he: god, high-status human', and *\*ni* 'thing, bad person, foreigner'; the first-person pronoun *\*na* 'I'; and a series of aspect/mood markers including the allomorphs *\*ko-* and *k<sup>w</sup>e-* of the completive proclitic. In conformity with Zapotec verb classes, we find *\*ko-* on verbs represented by logograms for 'to speak' and 'to stand'; and *\*k<sup>w</sup>e-* on a verb logogram seemingly representing 'to punish,' or something similar, which would have occurred with that proclitic in early Zapotec. These morphemes are spelled by CV syllabograms, the values of which were probably based on the acrophonic principle: for example, the sign ⟨na⟩ was based on (Central and Northern Zapotec) *\*na7* 'hand'; ⟨k<sup>w</sup>e⟩ was based on *\*na7 k<sup>w</sup>e* 'right

(hand)' (= \**na7* 'hand' + *k<sup>w</sup>e* 'straight'); and (ne) was based on \**nesa* 'road,' using the pan-Mesoamerican icon of paired footprints to indicate a road or path (the identification of the contextual grammatical functions of the signs spelling these grammatical morphemes preceded their phonetic readings).

Apart from signs used for grammatical elements, which are spelled by syllabograms, most words seem to be spelled by logograms. Phonetic complements occur but seem to be rare. It is unclear whether any words are spelled out in full with syllabograms alone. One logogram, representing Zapotec \**ko+* *kke* 'lord, lady,' is followed by the string (ko-ke) as a phonetic complement. This form provides evidence that consonant gemination, a contrastive feature of Zapotec, need not be written – and there is no known evidence for explicit marking of gemination. Possible further support for this spelling convention comes from a place name (HILL-ko-ti), which can be interpreted as /*\*tani kotti*/ 'hill of the dead'.

The structure of texts fits with what we know about the structure of Proto-Zapotec(an); the acrophonic origins of known syllabograms agree with Zapotec vocabulary; and a few lexical items spelled with syllabograms, the values of which are known through their grammatical uses, are readable as Zapotec words. The earliest texts including complete sentences go back to about 300 BCE, which, according to glottochronology, would be at or a little after the Proto-Zapotecan stage – when the Chatino branch of Zapotecan and the Zapotec branch separated. The latest inscriptions were produced around 900 CE, which must have been centuries after the Proto-Zapotec stage, when the westernmost subgroup of Zapotec broke off from the remaining dialects (c. 500 CE). The linguistic markers or traits that we find in the texts have in general not yet been seen to correlate with any of the phonological or grammatical or lexical differentiation that occurred in the history of Zapotecan and Zapotec – with two exceptions: (na) is based on North-Central Zapotec \**na7* 'hand' (other Zapotecan \**ya7*); \**ko+* *kke* 'lord, lady' is not known outside Central-Northern Zapotec.

## Bibliography

- Caso, A. 1928. *Las estelas zapotecas*. Monografías del Museo Nacional de Arqueología, Historia, y Etnografía, Mexico City.
- . 1947. *Calendario y escritura de las antiguas culturas de Monte Albán*. Talleres Gráficos de la Nación.
- Justeson, J. and T. Kaufman. 1996. "Warfare in the Preclassic inscriptions of Monte Alban." Unpublished manuscript.
- Kaufman, T. 1988. "Oto-Manguean tense/aspect/mood, voice, and nominalization markers." Unpublished manuscript.
- . 1995. "Proto-Zapotec reconstructions." Unpublished manuscript.
- . 2000. "A typologically-odd reconstruction for proto-Zapotecan: stem-final \*k." Paper presented at the Eighth Spring Workshop on Theory and Method in Linguistic Reconstruction, March 2000, University of Pittsburgh.
- Kaufman, T. and J. Justeson. 1993–2003. "The decipherment of Zapotec hieroglyphic writing." Unpublished manuscript.
- Marcus, J. 1976. "The iconography of militarism at Monte Alban and neighboring sites in the Valley of Oaxaca." In H. Nicholson (ed.), *Origins of Religious Art and Iconography in Preclassic Mesoamerica*, pp. 123–139. UCLA Latin American Studies Series 31. Los Angeles: UCLA Latin American Center.
- Urcid, J. 1992. "Zapotec hieroglyphic writing." (2 vols.). Ph.D. dissertation, Yale University.
- . 2001. *Zapotec Hieroglyphic Writing*. Studies in Pre-Columbian Art and Archaeology. 34. Dumbarton Oaks.
- Whittaker, G. 1980. "The hieroglyphics of Monte Alban." Ph.D. dissertation, Yale University.

# Reconstructed ancient languages

DON RINGE

## 1. INTRODUCTION

This chapter will necessarily be rather different from most of those that precede, since it deals with languages of which no direct record survives – languages which are, by definition, *prehistoric*. Prehistoric languages can only be studied inferentially, and the only sound basis for our inferences is the well-known *uniformitarian principle* (UP). As applied in historical linguistics, the UP states the following:

- (1) If the conditions of language use and acquisition cannot be demonstrated to have undergone any relevant alteration between the prehistoric and historical periods, nor between recorded history and the present, we must assume that the same types of language structures and language changes that we can observe today also underlie our historical records and were present in prehistory as well.

Since the only alternative is unconstrained guesswork, all scientific historical linguists must take the UP seriously. It follows that we must not interpret what we find in the written historical record in any way that is inconsistent with the range of structures and changes-in-progress that we can observe in languages currently spoken, nor must we posit for a prehistoric language any type of structure or change that is not actually attested somewhere among the known languages of the world. This very general principle has remarkably specific consequences, especially with regard to phonological change (see §3 below), which constrain and guide our hypotheses with a precision often surprising to interested observers outside the field. Of course the conditional clause in the UP is by no means automatically satisfied, since archeological evidence, modern anthropological work on stone-age populations (e.g., neolithic farmers in highland New Guinea) and the known principles of demography do enable us to judge the conditions of language use in prehistory to a surprising extent; for an eye-opening illustration of the consequences of taking this seriously, see Nichols 1990.

The remainder of this chapter is based squarely on the application of the uniformitarian principle by the whole community of mainstream historical linguists for well over a century.

Since no language system ever remains static, prehistoric languages have all met one of two fates. Some, perhaps even most, have died out without leaving traces of any sort; probably in most cases “language death” occurred because the language’s speakers were absorbed by another population or for some other reason abandoned their old language for a different one (cf., e.g., Foley 1986:24–25), though perhaps a few languages died out because of the biological extinction of their speaking populations. Other prehistoric languages survived into the historical period; but since all languages change continually, a language system as it emerges into the historical record is inevitably quite different from what it was sixty or

eighty generations earlier – at least as different as, say, Italian is from Classical Latin, and perhaps as different as French is from Latin (to cite two cases which have histories that are known in great detail).

If a prehistoric language has left only one historical descendant – that is, if it has evolved into only one surviving historical language – we may never be able to infer much about its structure at any time substantially before it emerged into history; too much will have changed. Some information about its prehistory will probably be recoverable by the techniques of *internal reconstruction*; but there are many linguistic structures to which they cannot be applied, or from which they lead to incorrect inferences (full discussion of this matter is well beyond the scope of this chapter; see Hoenigswald 1960:68–69, 99–111; Fox 1995:145–216; and Ringe, forthcoming)

But single speech communities often split into two or more new communities, which gradually lose touch and go their own ways, linguistically and otherwise. In each of these new communities language change will continue to occur, *but the specific changes will largely be different in each*, provided that contact between them is minimal or nonexistent. The result will be a family of related languages; in fact, that is the definition of the term *language family*. If two or more such descendants of a prehistoric language survive to be recorded, we can *reconstruct* at least some of the structure of their ancestor by a simple but rigorous mathematical procedure called the *comparative method*. Explication of the comparative method is beyond the scope of this chapter; the definitive codification of the method is Hoenigswald 1960, and a good practical introduction can be found in Fox 1995:17–144.

The ancient languages discussed in this chapter, and all other prehistoric languages about which we have substantial information, have each been reconstructed from multiple historically attested descendants by specialists using the comparative method. That single fact, more than anything else, determines what can be known about them. In the remainder of this chapter I will refer to reconstructed parent languages as *protolanguages*, the standard technical term for such inferred entities.

Most of the examples herein will be drawn from the Indo-European (IE) family, chiefly because that is the family of languages with which the author is most familiar. But there is nothing special about Indo-European; the principles discussed here, and the general statements made, apply equally to all protolanguages.

If any kind of language change proceeded at a constant rate, we could compare a protolanguage with its descendants of known date and calculate when the protolanguage must have been spoken. Unfortunately, rates of linguistic change appear to vary considerably, so that fixing the date of any protolanguage is a matter of informed guesswork. We can attempt to narrow the range of our estimates by seeking to correlate our results with the findings of archeologists, but some uncertainty inevitably remains. To a linguist this does not matter much; the *relative* chronology of important linguistic changes is often recoverable, and *absolute* chronology has little to do with the internal history of a language's structure.

More worrisome is the fact that the linguistic features reconstructed for a single protolanguage can actually be of slightly different dates, so that the reconstruction is temporally “out of focus.” This occurs chiefly because the members of a diversifying language family can undergo identical changes even after they have parted company, especially if the changes are natural and easily repeatable; if *all* the daughter languages undergo a particular change early in their separate careers, the effects of that change will naturally – but incorrectly – be projected into the protolanguage.

To consider an example: in most Indo-European languages the so-called laryngeal consonants of Proto-Indo-European (PIE) have been lost when preceded by a vowel, and if that vowel belonged to the same syllable as the laryngeal, it has been lengthened. Even in

the Anatolian subgroup, which preserves some laryngeals in some positions, a number of these “contractions” of vowels and laryngeals have occurred (Melchert 1994:67–69, 73). For Proto-Indo-European we must rely on phonological alternations and/or morphological evidence to distinguish, for example, between *\*ā* (as in *\*swād-* “sweet,” Stang 1974) and *\*eh<sub>2</sub>* (as in *\*weh<sub>2</sub>g-* “break,” Kimball 1988:245, Rix *et al.* 1998:605–606) before stop consonants. In reconstructing *\*eh<sub>2</sub>* we are certainly recovering the underlying form, but we cannot be certain that a surface contraction to *\*[ā]* had not already occurred in the last common ancestor of the Indo-European languages. In the absence of relevant alternations or morphologically related forms, we do not always know whether we are dealing with a vowel-plus-laryngeal sequence or an original long vowel. Thus, is PIE “arm” *\*b<sup>h</sup>āg<sup>h</sup>u-* or *\*b<sup>h</sup>eh<sub>2</sub>g<sup>h</sup>u-*? Even more uncertain is the historical status of laryngeal “coloring,” by which short *\*e* became *\*[a]* next to *\*h<sub>2</sub>* and *\*[o]* next to *\*h<sub>3</sub>*. All Indo-European languages show the results of this change, but can we be sure that it had already happened in Proto-Indo-European? Such problems have led the best historical linguists to be rather cautious about trying to identify the communities that spoke particular protolanguages, recognizing that to a certain extent any protolanguage is an idealized construct which is likely to have a complex relation to “real history.”

Strictly speaking, a reconstructed language has no speaking population; yet something very like each competently reconstructed protolanguage must have been spoken by some group of human beings. We can learn something about their society by examining the vocabulary that can be reconstructed for the protolanguage. For example, we know that the speakers of Proto-Indo-European – more exactly, of the actual language that most closely resembled our reconstructed Proto-Indo-European – wore clothes, since we can reconstruct not only such forms as *\*wéstor* “(s)he’s wearing,” *\*woséyeti* “(s)he’s dressing [someone else]” (Rix *et al.* 1998:633–634), *\*yeh<sub>3</sub>s-* “wear a belt” (Rix *et al.* 1998:275–276), and *\*h<sub>2</sub>w<sup>l</sup>h<sub>1</sub>no-* “wool” (Peters 1980:23–26, fn. 18), but also *\*neg<sup>w</sup>nós* “naked,” which implies that people customarily wear clothes.

There are, however, obvious limits to what we can learn in this way. In particular, we can argue only from the presence of reconstructible words for a particular article or concept, not from their absence; the replacement of inherited words by completely different words is a universal and very common type of linguistic change, and that alone can easily account for the fact that there are so many gaps in our reconstructible lexica. This is clearest from a consideration of body-part terms. For example, no Proto-Indo-European word for “finger” is reconstructible; yet surely speakers of Proto-Indo-European had fingers, and (like every other human community) they must have had a word for them!

Apparent breaches of this principle of limitation are just that – apparent rather than real. For example, each major subgroup of the Indo-European family has its own word for “iron” (except that Proto-Germanic *\*īsarnā* appears to have been borrowed from Proto-Celtic *\*īsarnom*; for a full discussion see Birkhan 1970:128–137), so that no word for that metal can be reconstructed for Proto-Indo-European; and it’s true that virtually all serious scholars believe iron to have been unknown to the Proto-Indo-European speech community. But that belief is *not* based on the fact that no relevant words can be reconstructed; it follows instead from well-known archeological findings about the geographical and chronological distribution of iron among ancient cultures.

The necessary methodology of comparative reconstruction imposes further limitations on what can be known about the prehistory of even the most solidly reconstructible protolanguages. For example, we are mathematically constrained to reconstruct a more or less unitary dialect as the ancestor of each attested family. Yet experience with living languages leads us to infer that most of these reconstructed dialects must have been members of dialect

networks – all the other dialects of each network having more or less completely died out. Our ability to reconstruct the relative chronology of the changes that occurred as each protolanguage diversified into a language family is likewise limited: we can recover the relative chronology of those changes that interacted (one change producing the conditions under which another could then take place, or removing examples which would otherwise have undergone a later change), but changes that had nothing to do with one another cannot be ordered chronologically.

Finally, an unpleasant fact of language change imposes the most drastic limitation on what can be known. All languages gradually replace their inherited vocabulary with completely different and unrelated vocabulary items, and also replace, lose, and restructure the affixes with which full words are formed. “Basic” vocabulary is, of course, replaced at a relatively slow rate, and inflectional affixes are also resistant to change; but in the long run every word will be replaced, and inherited inflectional patterns will be transformed beyond recognition. When the vast majority of even the most tenacious items have disappeared, the few remaining cognates shared by genuinely related languages will be indistinguishable from chance resemblances – so that the relationship will be undiscoverable, and reconstruction of a protolanguage will be impossible. There is, therefore, a temporal limit beyond which we will probably never be able to penetrate prehistory; and though estimates of that limit differ, it seems clear that a threshold even ten millennia before the earliest attested documents of a language family is beyond our reach for all practical purposes.

## 2. PHONOLOGY

### 2.1 Regularity of sound change

It is the phonology of protolanguages that is most solidly reconstructible, for a simple reason: in any line of linguistic development, *sound changes* (changes in pronunciation) are overwhelmingly regular. That is, in a given line of development – say, from Latin into some specific dialect of Italian – within a given span of time, either a given sound  $x$  always develops into  $x'$  (which may or may not be phonetically identical with  $x$ ), or else the conditions under which  $x$  becomes  $x'$  are statable entirely in terms of other sounds in the same word or phrase. Since the sound changes that took place in the development of a given language  $L_1$  from its ancestor  $P$  are regular, and the sound changes that took place in the development of another given language  $L_2$  from the same ancestor are likewise regular (though different from those that occurred in  $L_1$ ), we will find *regular correspondences* between the sounds of  $L_1$  and the sounds of  $L_2$  to the extent that both languages have preserved words and forms inherited from their common ancestor  $P$ ; and we can exploit those correspondences to “triangulate” back to  $P$  by the comparative method. The regularity of sound change operates on contrastive units of sound – that is, on phonemes, whether “classical,” lexical, or underlying, depending on one’s analysis – as Hoenigswald 1960 demonstrates; thus the comparative method recovers a protolanguage’s phonemic system and the phonemic shapes of its words and affixes (insofar as their reflexes survive in two or more daughters).

Modern work in sociolinguistics has shown that the scenario just summarized is slightly oversimplified; most importantly, sound changes pass through a variable phase before “going to completion,” and occasionally the progress of a sound change is arrested in the variable phase, giving rise to irregularities (see, e.g., Labov 1994 for discussion). But the statistical preponderance of regular sound changes remains impressively massive, and it is almost always methodologically advisable to treat explanations involving irregular sound changes



with suspicion (essentially for the same reason that one always hesitates to draw to an inside straight). The correctness of this view has been confirmed repeatedly. For example, late in the nineteenth century a strict application of the comparative method – assuming the absolute regularity of sound change – led Karl Brugmann and others to reconstruct for Proto-Indo-European three sets of dorsal stops, conventionally called palatals, velars, and labiovelars. Such a reconstruction was repeatedly denied by some linguists, partly on grounds of sheer implausibility (though see below) and partly because no single Indo-European language then known clearly preserved different reflexes of all three sets. Yet Melchert 1987 showed that the three Proto-Indo-European voiceless dorsal stops do have different reflexes in Cuneiform Luvian, a language whose records were discovered only in this century (namely, PIE  $*k >$  Luvian  $z$ ,  $*k > k$ , and  $*k^w > ku$ ), thus vindicating the traditional reconstruction of Proto-Indo-European and the comparative method. (For a striking vindication of the regularity of sound change from a quite different empirical perspective, using a wide range of modern data, see Labov 1994:419–543.)

## 2.2 Proto-phonetics

The *phonetics* of proto-phonemes, on the other hand, are not fully determined by the distribution of sound correspondences; inferences about proto-phonetics are probabilistic, and the best we can do is to maximize our chances of arriving at the correct solution by using all potentially relevant information at our disposal. The most solid basis for inferences about proto-phonetics is (naturally) the phonetics of a proto-phoneme's reflexes in the attested daughter languages; a hypothesis about the phonetic identity of any proto-phoneme will be plausible to the extent that its attested phonetic reflexes can be derived from the posited proto-sound by natural, plausible sound changes. Unfortunately, our judgments of the naturalness of sound changes rest on experience, and even the collective experience of the whole community of historical linguists is insufficient to solve some puzzles; for example, most Indo-Europeanists remain noncommittal about the phonetics of the so-called laryngeals reconstructed for Proto-Indo-European (see Ch. 17 §2.1.3) simply because the historical record does not provide us with many examples of sounds that become dorsal fricatives in some daughter languages but vowels in others. For similar reasons it is less than clear whether the emphatic stops of Proto-Afro-Asiatic (see Ch. 6 §1.3.1) were glottalized or pharyngealized.

It has become customary to bring typological arguments to bear on problems of proto-phonetics; but both our knowledge of the full range of typological possibilities in human language and the usefulness of typology in general have been at times abused. A notorious case is the *glottalic hypothesis* regarding Proto-Indo-European stop consonants, according to which the traditional voiced stops should be reconstructed as glottalized, the traditional voiceless stops as aspirated, and the traditional voiced aspirated (i.e., breathy-voiced) stops as voiced (with or without aspiration; see e.g., Gamkrelidze and Ivanov 1973, Hopper 1973). One motivation for this revision was the supposed typological impossibility of the system of Proto-Indo-European stops as usually reconstructed; yet a remarkably similar system is found in present-day Madurese (Stevens 1968:16, 38) and some other Indonesian languages, as Hock (1986:625–626) observes. Moreover, adopting the glottalic hypothesis would force us to posit serious implausibilities in the phonological development of early Iranian loanwords in Armenian (Meid 1987:9–11). In other words, what is known about the phonetics of the reflexes of Proto-Indo-European stops actually furnishes hard evidence *against* the glottalic hypothesis. Brugmann's reconstruction of three sets of Proto-Indo-European dorsal stops was also rejected in part because of its supposed implausibility; yet similar systems are well

attested in the languages of the Caucasus and of the northwest coast of North America, and new evidence has vindicated Brugmann directly (see §2.1). Debacles of this kind reveal that typological arguments should be used with considerably more caution than they have been used with by some investigators.

### 2.3 Limitations on reconstruction

In view of these facts it is not surprising that the prosodic phenomena of protolanguages may or may not be recoverable, and that any interesting allophony of proto-phonemes is likely to be recoverable only under favorable circumstances. Moreover, the reconstruction of phonological rules, phonotactic constraints, syllable structure, and the like depends largely on being able to reconstruct sufficient vocabulary and paradigms.

Even above the level of the phoneme, some types of linguistic development can make reconstruction difficult or impossible. A well-known example is the phonological rule called *Bartholomae's Law*, by which the “breathy voice” of a voiced aspirated stop spreads rightward to an adjacent apical stop or fricative. The rule still affects stop + stop clusters in Sanskrit (e.g., *bud<sup>h</sup>* - “awaken” + *-tá-* (part.) → *budd<sup>h</sup> á-* “awake, enlightened”; see Ch. 26 §3.4.2.1); and though voiced aspirated obstruents have merged with voiced obstruents in Avestan, the most archaic dialect of that language (Gāthic Avestan) still contains numerous stop + stop and stop + fricative clusters the voicing of which shows that they had been affected by Bartholomae's Law before the merger. Bartholomae's Law can then be reconstructed for Proto-Indo-Iranian (the latest common ancestor of the Indic languages, including Sanskrit, and the Iranian languages, including Avestan) with certainty (see Schindler 1976).

It seems clear, however, that Bartholomae's Law is easily lost from a language's grammar. In Sanskrit it has ceased to apply to stop + fricative clusters, so that corresponding to Avestan *diβža-* “deceive” (from Proto-Indo-Iranian *\*d<sup>h</sup>)ib<sup>h</sup>z<sup>h</sup>a-*) we find Vedic Sanskrit *dipsa-* “want to deceive.” In Younger Avestan (later than Gāthic) it has largely been eliminated in all environments, so that in place of Gāthic *aogdā* “(he) proclaimed” (<*\*aug<sup>h</sup> d<sup>h</sup> a* < *\*aug<sup>h</sup> - + \*-ta*) we find *aoxta* (note that underlying /k/, or inherited *\*k*, normally appears as *x* before consonants in Avestan).

This raises an obvious question: is it possible that Bartholomae's Law was a phonological rule of Proto-Indo-European, and that it has been lost in all branches of the family other than Indo-Iranian? Such a hypothesis is by no means implausible, but proof would depend on finding unarguable relic forms in at least one other branch. Unfortunately the evidence is equivocal, and among specialists no consensus has emerged (see Mayrhofer 1986:115–117 with references).

### 2.4 Reconstruction of phonemic systems

The degree of similarity that obtains between the phonological system of a reconstructable protolanguage and those of its historically attested daughters is not a constant; it depends on what sound changes have occurred in each of the daughters. At one end of the spectrum, the phonemic system of Proto-Semitic is almost identical with that of Classical Arabic (Bergsträsser 1928:3–5, 134), simply because few sound changes occurred in the development of the latter. Similarly, the sound system of Proto-Algonkian strongly resembles those of many of its daughters, the most notable deviation being the presence of a consonant, conventionally symbolized as *\*θ*, which has merged with other consonants in most daughter languages (Bloomfield 1946:85–90). On the other hand, the phonemic system of Proto-Indo-European is fairly different from that of any daughter language:

## (2) The Proto-Indo-European phonemic system

<i>Obstruents</i>					<i>Sonorants</i>		
p	t	$\hat{k}$	k	k <sup>w</sup>	m	$\bar{m}$	
b	d	$\hat{g}$	g	g <sup>w</sup>	n	$\bar{n}$	
b <sup>h</sup>	d <sup>h</sup>	$\hat{g}^h$	g <sup>h</sup>	g <sup>wh</sup>	l	$\bar{l}$	
	s	h <sub>1</sub>	h <sub>2</sub>	h <sub>3</sub>	r	$\bar{r}$	
					y	i	$\bar{i}$
					w	u	$\bar{u}$
<i>Vowels</i>							
e	a	o					
$\bar{e}$	$\bar{a}$	$\bar{o}$					

The position of the laryngeals (\**h*<sub>1</sub>, \**h*<sub>2</sub>, \**h*<sub>3</sub>) is speculative; it does not affect the discussion that follows. There was also some sort of pitch-accent system.

Of the attested daughters, Sanskrit best preserves the stop system but has merged the laryngeals with other sounds. Hittite best preserves the laryngeals but has undergone mergers in the stop system (Melchert 1994:60–62, 117–120). The labiovelar stops are especially prone to change; in only a few attested daughters can they be shown to survive as distinctive unit phonemes. The very distinctive Proto-Indo-European system of sonorants, in which the nonsyllabic sonorants alternated with their syllabic counterparts under well-defined conditions, has nowhere been preserved without alteration. As might be expected, the vocalic portion of the system (\**y* ~ \**i* and \**w* ~ \**u*) survives longest (in Indic and Germanic). However, it does not follow that the above reconstruction is in any way doubtful; the comparative method reconstructs with confidence the system of phonological contrasts reported here, and the fact that all attested systems of Indo-European languages are more or less different is a straightforward consequence of the regular sound changes they have each undergone.

## 2.5 The uniqueness of sound change

Since sound change is the only type of language change that exhibits statistically overwhelming regularity giving rise to clear patterns that pervade comparative data, it is the only type of change that can be exploited directly to reconstruct protolanguages. The recovery of all other aspects of protolanguage structure depends on our ability to reconstruct individual words and affixes by exploiting regular sound change. The methodological observations of the following sections therefore will not amount to a complete and coherent procedure comparable to the (phonological) comparative method.

## 3. MORPHOLOGY

### 3.1 The nature of morphological reconstruction

Much of what has traditionally been subsumed under morphological change appears in the light of current theory to be essentially phonological or syntactic. For example, the leveling of morphophonemic alternations is sometimes described as the restriction or loss of a phonological rule; and the complete loss of grammatical categories, such as the loss of the instrumental case within the history of Old English, is clearly a syntactic phenomenon, though it obviously has morphological consequences. The irreducible residue of morphological

changes does not seem to be governed by any regular “rules”; one repeatedly finds that although some particular type of change is quite common, its converse is not rare. For instance, systems of nominal case-marking are frequently simplified over the course of a language’s development (as in many European languages), but we also find that new case-markers evolve by the accretion of postpositions (most spectacularly in Tocharian, but also in Lithuanian and Armenian); aspect-based verb systems often evolve into tense-based systems (e.g., in Latin), but new systems of aspect can and do arise (e.g., in Russian; see further below).

For these reasons the reconstruction of a protolanguage’s morphology simply cannot be pursued on the basis of matching functional categories in the daughter languages without regard to their formal expression. On the contrary, reliable reconstruction of morphological categories depends almost entirely on reconstructing the morphemes that instantiate them by exploiting the regularity of sound change in the usual way. This can be illustrated by considering the reconstruction of two details of the Proto-Indo-European verb, namely the system of voices and the tense and aspect system.

### 3.2 Reconstructing morphology

The morphological expression of voice in the verb systems of the oldest, well-attested representatives of each major branch of the Indo-European family is summarized in (3):

#### (3) Morphological voice in selected Indo-European languages

	<i>System of oppositions</i>	<i>Deponent verbs?</i>
<i>Hittite</i>	active vs. mediopassive	yes
<i>Sanskrit</i>	active vs. middle vs. passive	yes (middle only)
<i>Avestan</i>	active vs. middle vs. passive	yes (middle only)
<i>Greek</i>	active vs. middle vs. passive	yes (both middle and passive)
<i>Latin</i>	active vs. passive	yes
<i>Armenian</i>	active vs. mediopassive	yes
<i>Gothic</i>	active vs. passive	no
<i>Old Irish</i>	active vs. passive	yes (distinct from passives)
<i>Tocharian</i>	active vs. mediopassive	yes
<i>Old Church Slavonic</i>	(active only)	—
<i>Lithuanian</i>	(active only)	—
<i>Albanian</i>	active vs. mediopassive	yes

A large majority of these languages exhibit an opposition of at least two voices in their verb morphologies; it would, however, be rash to project such a situation back into Proto-Indo-European without further investigation, as it is conceivable that the nonactive voices evolved independently in these languages. Such a development can be demonstrated for Russian and Icelandic, which have created new middle-voice forms by the accretion of a reflexive clitic to active forms. In fact, the Albanian mediopassive appears to be a new formation as well: its only nonperiphrastic forms (made to the present stem) are constructed with a suffix *-(h)e-* and ordinary, active-looking endings, and there are no clear cognates in any other language. A closer look at the three-voice systems of Greek and the Indo-Iranian languages (Sanskrit and Avestan) shows that their morphological opposition between middle and passive is likewise the result of secondary developments: in all three languages those two voices are distinguished only when there is a separate passive stem (Greek aorist *-(t<sup>h</sup>)ē-* and

future  $-(t^h)\bar{e}-se/o-$ ; Sanskrit and Avestan present  $-ya-$  and aorist [third singular only]  $-i-$ , and though some of those stems have cognate formations in other languages, they do not exhibit passive function in those languages.

To demonstrate that Proto-Indo-European had a morphological opposition between an active and a nonactive voice we must be able to reconstruct at least some clear and distinctive morphological markers for the latter, and in fact that can be done. For example, in past tenses we find the following cognate set for the third singular *nonactive* ending: Sanskrit, Avestan  $-ta$  = Greek  $-to$  = Tocharian B  $-te$  = Tocharian A  $-t$  < PIE  $*-to$ . Moreover, the same nonactive  $*-to$  appears with further suffixes in Hittite past third singular  $-(t)tat(i)$  <  $*-to-ti$  (cf. Melchert 1994:60) and in Latin third singular  $-tur$  <  $*-to-r$  (Latin does not systematically distinguish past from non-past endings; there is much more to be said about  $-tur$  and its cognates, but that is beyond the domain of this discussion [see further below]). For the *active* past third singular ending we have a quite different cognate set: Hittite, Sanskrit, Avestan  $-t$  = Greek, Tocharian  $\emptyset$  = Old Latin  $-d$  < PIE  $*-t$ . The opposition between these two reconstructible endings, and between the members of other, similar pairs, establishes an active: nonactive contrast for the Proto-Indo-European verb.

In the case just sketched, enough of the daughter languages have preserved the same morphological markers in comparable functions to enable us to reconstruct part of the proto-system. We are not always so lucky, as the following case will show.

The Hittite verb system is remarkably simple. Except for a tiny handful of anomalies, each verb has only one stem. Conjugation is effected by means of endings; each finite ending expresses person, number (singular vs. plural), voice (active vs. mediopassive), and tense-mood (present vs. past vs. imperative). The only notable complication is the fact that there are two lexical classes, called the *hi*-conjugation and the *mi*-conjugation after the shape of their first singular present active endings (see Ch. 18 §4.4.7).

In the other archaic Indo-European languages, the situation is vastly different. Simplifying somewhat, we can say that Greek exhibits a rich and elaborate inflectional system based on aspect: most verbs exhibit an imperfective stem (called the “present stem”) and a perfective stem (called the “aorist stem”), and many also have a stem which in Homeric Greek is still usually stative (called the “perfect stem”), though verbs lacking one or more of these stems are not rare. On each stem are built a non-past and a past indicative tense (except that there is no perfective non-past), as well as subjunctive, optative, and imperative mood forms, participles, and infinitives; as noted above, there are three voices. In addition, there are future tense stems with a defective set of forms. The system of the Indo-Iranian languages is strikingly similar; in purely formal terms, a large proportion of its morphology corresponds point for point with that of Greek, though some functional shifts have occurred. The Armenian system also shows a fundamental contrast between perfective and imperfective stems, though there is nothing corresponding to the Greek and Indo-Iranian “perfect” (i.e., stative); and though the Latin system has been more extensively restructured, there are numerous clear indications that it closely resembled the Greek system at no very great remove in its prehistory. The same can be said of the Old Church Slavonic verb, which still shows, for example, a clear contrast between “present” and “aorist” stems. Even the Baltic and Germanic verb systems, which are based on a simpler opposition of tenses (present vs. past), exhibit *two different stems* per verb: in both groups the present stem is cognate with the “present” (imperfective) stem of Greek, etc.; the Germanic past of strong verbs is cognate with the “perfect” (stative) of the more southerly languages, while the Germanic weak past and the Baltic past appear to be innovations. The verb systems of Tocharian, Old Irish, and Albanian show more unique features and present many more puzzles; but in those languages too it is normal to find at least two stems per verb, and a large proportion of each

of those systems can be explained as late developments of a Greek-type system without great difficulty.

Should we reconstruct for Proto-Indo-European, then, a complex verb system much like that of Greek and Indo-Iranian, and suppose that the Hittite verb system, in spite of its very early date of attestation (mid-second millennium BC), has undergone radical simplification? Some scholars have concluded that that is correct, but there is a further fact that gives cause for doubt. Roughly speaking, the endings of the Hittite *hi*-conjugation are cognate with those of the “perfect” (stative) of the other archaic Indo-European languages, in spite of the fact that most *hi*-conjugation verbs are not stative in meaning, while the endings of the *mi*-conjugation are cognate with those of the “present” and “aorist” (perfective and imperfective) of the other languages. There is, thus, a complete and irreducible lack of “fit” between the functional categories of the two systems:

(4) Anatolian and the other Indo-European languages: cognations in the verb system

<i>Hittite</i>		<i>Greek, Sanskrit, etc.</i>
<i>hi</i> -class vs. <i>mi</i> -class (lexical classes)	≈	“perfect” (stative) vs. “present” and “aorist” (imperfective and perfective)
one stem per verb	:	more than one (aspect-)stem per verb

It is not easy to imagine a system of Proto-Indo-European verb categories that could have given rise to two systems as different as these by uncontroversially natural changes. Not surprisingly, specialists have been arguing about the matter for decades, and there is still no consensus.

This shows clearly that, even in a thoroughly researched family of languages for which we have abundant early evidence of high quality, the very structure of the data can frustrate our attempts to arrive at a plausible reconstruction of parts of the morphological system. The basic reason for this difficulty is that there seem to be no general “laws” of morphological change comparable to the regularity of sound change; morphological systems, idiosyncratic by nature, seem to change in idiosyncratic ways, and when we are presented with a pattern of data that has no close parallels in the attested history of languages, we can do no better than make informed guesses about the changes that might have produced it.

Sometimes, however, distributional factors can be exploited to tell us more than we could otherwise have figured out. Let us return to Latin passive third singular *-tur* (see above). Though Latin has largely eliminated an inherited contrast between past and non-past endings, many related languages have preserved it, and it is clear that *-tur* is cognate with a set of non-past third singular nonactive endings. But there are two cognate sets in that function:

(5) Indo-European non-past mediopassive third singular endings

- Set 1: Hittite *-(t)ta* ~ *-(t)tari* < *\*-tor* (± *\*-i*; on the complex development of this ending see Yoshida 1990)  
 Phrygian *-tor*  
 Tocharian A and B *-tär* < *\*-tor* (Ringe 1996b:86–87)  
 Latin *-tur* < *\*-tor*  
 Old Irish (conjunct) *-thar* < *\*-tor*
- Set 2: Greek *-toi* (Arkadian dialect) / *-tai* (most dialects)  
 Sanskrit *-tē*, Avestan *-taē* < *\*-toi* or *\*-tai*  
 Gothic *-da* < *\*-toi* or *\*-tai*



Which set reflects the Proto-Indo-European ending? Two distributional facts suggest that Set 1 does, and that the correct Proto-Indo-European reconstruction is *\*-tor*.

In the first place, note the formal relations between the full set of Proto-Indo-European (nonimperative) third singular endings if *\*-tor* is the correct reconstruction:

(6) Proto-Indo-European third singular endings

	<i>Past</i>	<i>Non-past</i>
<i>Active</i>	<i>*-t</i>	<i>*-ti</i>
<i>Mediopassive</i>	<i>*-to</i>	<i>*-tor</i>

It seems clear that the consonant *\*-t(-)* is what marks the third-person singular, and that an additional suffix *\*-o(-)* marks the mediopassive. Moreover, the non-past endings are distinguished from the past endings by yet a further suffix, *\*-i* in the active and *\*-r* in the mediopassive. If these reconstructions are correct, we have a plausible source for *\*-toi*, the most probable reconstruction for Set 2 in (5): the system has been simplified by extending the scope of *\*-i* from the active to the mediopassive, replacing *\*-r* so as to create a unitary non-past marker (and *\*-o-i* > *\*-oi* is not problematic; cf. the similar summary of Yoshida 1990:117). On the other hand, if the correct Proto-Indo-European reconstruction for the lower right-hand cell of (6) were *\*-toi* or *\*-tai*, we would have no plausible source for the *\*-r* of Set 1 *\*-tor*. This is a strong distributional argument (though not sufficiently strong to convince all specialists).

The other distribution that favors a reconstruction of *\*-tor* for Proto-Indo-European is geographical. Of the subgroups that show reflexes of *\*-tor*, nearly all are unarguably geographically peripheral: Celtic and Italic are at the western edge of the (known) Indo-European speech area, Tocharian at the eastern edge; and since the Anatolian languages are found in what is now Turkey in the second millennium BC, when nearly all other Indo-European-speaking groups that we know about seem to have been living further north and northwest, it is overwhelmingly likely that they were then at the southern margin of the Indo-European speech area (the geographical position of Phrygian at that date is completely obscure). By contrast, the groups that exhibit Set 2 endings are all more centrally located. We are accustomed to thinking of Indo-Iranian as the far southeastern corner of the Indo-European speech area, but it is abundantly clear that Indo-Iranian spread east (and later south) from the Eurasian steppe, where speakers of Iranian languages continued to be an important part of the population in the last few centuries BC (cf., e.g., Schmitt 1989:92–93 with bibliography). Speakers of Balto-Slavic must have been to the northwest throughout the last two millennia BC, with speakers of Germanic still further northwest (and all three in contact, for which there is abundant linguistic evidence; cf., e.g., Porzig 1954:139–147, 164–166; Stang 1972; Hock 1986:442–444, 451–455, 667 with bibliography). While the position of Greek at an early date is less clear, it seems at least to have been less peripheral than Anatolian. We can therefore suggest that the Set 2 endings reflect an innovation that occurred in the central part of the Indo-European speech area but did not spread to the margins; and that is a reasonable hypothesis whether we regard the central languages as a valid subgroup exclusively sharing an immediate parent (in which the change could have occurred) or as a group of diversified dialects still in contact (in which case the change could have spread from dialect to dialect). In the former case we must suppose that Balto-Slavic, and perhaps also Armenian and Albanian, shared the change before losing the mediopassive voice; in the latter it is at least thinkable that some of those groups had already lost the mediopassive before the change in the endings occurred. Of course this line of reasoning too can be challenged; in particular, it is not completely clear that the innovation that produced

the Set 2 endings could not have occurred more than once independently (Jay Jasanoff, personal communication).

It can be seen from these examples that reconstructing the morphological system of a protolanguage involves procedures much less mechanical and much less rigorous than the reconstruction of a protolanguage's phonology and of the shape of its words and affixes. In effect, we use every bit of information that we have in morphological reconstruction – including everything that linguistic theory, the description of modern languages, and the historical record can tell us about morphological structure and change. Nevertheless, the results are often less clinching than in phonological reconstruction, and morphology is both the area in which the most intensive and interesting work on many language families is being pursued and the area in which disagreements between specialists are most prevalent.

#### 4. SYNTAX

Theoretically well-informed work on the syntax of protolanguages is still in its infancy. Most Neogrammarian work (e.g., Delbrück 1893–1900) treated syntax as an adjunct of morphology, so that syntactic constructions were discussed in terms of the morphological categories that mark them; and though the rich body of data amassed by our predecessors continues to be useful, modern work in syntax shows all too clearly that a theory which treats syntax as an extension of morphology cannot be correct. It should be noted, however, that at least one line of Neogrammarian research, Wackernagel's pioneering work on the placement of "second-position" clitics (Wackernagel 1892), has proved to be very fruitful in a modern theoretical context (see further below).

Within Indo-European studies, an approach to syntactic change that was developed before the mid-1980s concentrated on the basic order of major constituents in the clause (cf., e.g., Lehmann 1974); that approach has been disappointing, since the *surface order* of constituents is not a primitive of Proto-Indo-European syntax. A promising beginning to syntactic reconstruction had been made in the 1960s (see Watkins 1963, 1964, Kiparsky 1968); but syntactic theory has developed so rapidly that those early studies need to be thoroughly reevaluated in a more modern syntactic framework, and so far that has not been attempted. Indeed, there are good reasons why the attempt has been postponed: cogent reconstruction of the syntax of protolanguages will become possible only when the processes of syntactic change are much better understood, and syntactic change can be investigated in the requisite detail only in the context of a highly articulated theory. Chomskyan Government and Binding theory and a number of its competitors reached a suitable level of sophistication and precision in the mid-1980s, and since then a steadily increasing amount of useful and detailed work on syntactic change has appeared. (A thorough and up-to-date discussion of what is known and what remains to be done is Kroch 2001; for discussion of a large quantity of interesting crosslinguistic data on syntactic change from a theoretically detached perspective, see Harris and Campbell 1995.)

#### 5. LEXICON

As noted above, the reconstruction of a protolanguage's lexicon by the comparative method must always remain incomplete because of the replacement of lexical items in the daughter languages; the inferences that can be drawn from our lexical reconstructions are therefore



limited. A particular problem is the fact that changes in the meanings of words seem to be very idiosyncratic indeed; for that reason we are occasionally unable to specify the meaning of a proto-lexeme with any precision even when its form is uncontroversially reconstructable. A case in point is the protoform of Greek *stóma* “mouth,” Avestan (acc.) *stamanəm* “jaws (of a dog),” and Hittite *ištaman*, Luvian *tūmmant-* “ear.” From these reflexes we can reconstruct a Proto-Indo-European noun *\*stéh<sub>3</sub>m̥r̥* ~ *\*sth<sub>3</sub>m̥r̥* (Melchert 1994:73–74), and it is reasonable to infer that it referred to some orifice of the head; but given that we can also reconstruct Proto-Indo-European *\*(h<sub>1</sub>)éh<sub>3</sub>s* “mouth” (Melchert 1994:115–116) and *\*h<sub>2</sub>éusos* “ear” (Szemerényi 1967), the exact referent of *\*stéh<sub>3</sub>m̥r̥* remains indeterminable. An even more spectacular case involves the Proto-Indo-European words for “head” and “horn”; see Nussbaum 1986 for a full and fascinating discussion which does achieve firm results.

Occasionally the evidence for a proto-lexeme’s meaning unanimously supports a reconstruction which is demonstrably anachronistic. For example, a Proto-Algonkian word *\*paaškesikani* can be reconstructed from cognates in several daughter languages, all of which mean “gun, firearm” (Fox *paaškesikani*, Ojibwa *paaškəšikan*, Cree *paaskisikan*, Bloomfield 1946:106). Yet it is clear that the Proto-Algonkian speech community, which must have existed roughly two millennia ago, cannot possibly have known and used firearms. In fact, the attested words are all instrument nouns derived from a verb meaning “shoot with a gun” (Bloomfield, loc. cit.), and that verb is actually a complex formation which literally means “make things burst by means of fire” (Bloomfield 1946:114, 120). Either the verb and its derived instrument noun have been created independently in the languages in which they occur, or (more probably) the word was coined in one language and translated morpheme-by-morpheme into the others as firearms spread from community to community. This case, in which the anachronism of a *prima facie* reconstruction is obvious, should alert us to the possibility of similar pitfalls in cases over which we have fewer or less reliable means of external (dis)confirmation.

A possibly relevant case is provided by the Proto-Indo-European terms for “wheel.” Two etymological groups of words are well attested throughout the family. On the one hand, Vedic Sanskrit *ráthas* “chariot” reflects substantivization of an adjective *\*(H)roth<sub>2</sub>-ó-s* “having a set of wheels.” That adjective is evidently derived from *\*(H)roté-h<sub>2</sub>* “set of wheels,” the source of Latin *rota* “wheel”; and *\*(H)roté-h<sub>2</sub>* is in turn the collective of *\*(H)rót-o-s* “wheel,” the ultimate source of German *Rad* and Lithuanian *rātas* “wheel” (Rix *et al.* 1998:459; note that in the preceding Proto-Indo-European reconstructions the identity of the laryngeal is uncertain [hence noted as H], and, moreover, parentheses indicate that the very presence of the laryngeal is uncertain – the data *underdetermine* the reconstructions of these words, a fairly common problem). On the other hand, we can also reconstruct a term *\*k<sup>w</sup>ék<sup>w</sup>lo-s* “wheel,” collective *\*k<sup>w</sup>ék<sup>w</sup>lé-h<sub>2</sub>*, which is the source of English *wheel*, Homeric Greek *kúklos* (pl. *kúkla*), Sanskrit *cakráṃ* (pl. *cakráṅ*), and so on.

Both of these words are transparently descriptive. *\*(H)rót-o-s* is literally “(act of) running,” an unproblematic action noun derived from *\*(H)ret-* “run” (cf. Old Irish *rethid* “(s)he runs”), while *\*k<sup>w</sup>ék<sup>w</sup>lo-s* is a reduplicated derivative of *\*k<sup>w</sup>el-* “turn” (cf. Sanskrit *cáratī* “(s)he wanders,” Homeric Greek *peri-tellómenos* “going around, revolving,” etc.). Thus, neither of these words by itself is watertight evidence for wheeled vehicles in the Proto-Indo-European speech community, as opposed to its immediate daughters.

In contrast, the reconstructibility of Proto-Indo-European *\*h<sub>2</sub>iHséh<sub>2</sub>* or *\*h<sub>3</sub>iHséh<sub>2</sub>* “thill” (cf. Hittite *hiššaš*, Sanskrit *īṣā*; Melchert 1994:78, 152), which is completely opaque, and of a basic verb *\*wég<sup>h</sup>eti* “(s)he transports . . . in a vehicle,” is better testimony for Proto-

Indo-European wheeled-vehicle technology; and the fact that an entire technical vocabulary for wheeled vehicles (including, of course, the two terms for “wheel”) can be reconstructed adds further weight to the argument (Mallory 1989:275–276, fn. 25; Anthony 1995:556–558).

It is sometimes possible to find patterns among proto-lexemes from which we can make inferences about the deeper prehistory of a protolanguage. The reconstructible Proto-Indo-European lexicon is especially amenable to this procedure because of a quirk of Proto-Indo-European word structure.

Both the inflectional and the derivational morphology of Proto-Indo-European exhibit a relatively simple but pervasive system of vowel alternation (*ablaut*). Literally hundreds of reconstructible paradigms and word-families, including the vast majority of reconstructible items, participate in the system, which is consequently well understood. One general principle of Proto-Indo-European ablaut is that we expect to find only one full vowel (\**e*, \**o*, or \**a*, either long or short; on PIE vowel gradation see Ch. 17 §2.2) per form, except for a closed list of derivational and inflectional morphemes (such as the *thematic vowel*, a stem-final vowel which always appears as full \**-e-* or \**-o-* no matter what ablaut grade appears in the root-syllable; see Ch. 17 §3.4). It is therefore rather startling to find that almost every word for “ax” that might be reconstructible for Proto-Indo-European violates that principle, or fails to conform to the usual *sound laws* that reflect regular sound change (see §2.1), or both. Thus, \**peleku-s* (Gk. *pélekus*, Skt. *paraśús*) exhibits two full-grade \**e*’s in its root, whereas we expect to find only one (and note that the attested reflexes do not agree on the position of the accent). The Greek word *aksínē* seems to be related to Germanic forms such as Gothic *aqizi* and Old High German *acchus*; but the first consonant of the Germanic words should reflect Proto-Indo-European \**g<sup>w</sup>*, which in Greek ought to appear as *-p-*, not *-k-*, before an immediately following *-s-*. The Germanic words also point to a protoform with two full vowels in the root (\**ag<sup>w</sup>esiH-*). Most stunning of all is the apparent connection between Hittite *ates* “ax” and Old English *adesa* “adze” – there are no other cognates – which, if it is not a mirage, can only reflect a preform \**ad<sup>h</sup>es-*, again with two full vowels (see Puhvel 1984:227–228). The only “ax” word that makes sense in Proto-Indo-European terms is the one reflected in Latin *secūris* and Old Church Slavonic *sěkyra*, both derived from \**sek-* “cut” (though the Slavic form reflects a long \**ē* in the root) with a suffix containing the sequence \**-ūr-* or \**-uHr-* (though the final stem-vowel differs). It would be reasonable to infer from this pattern of data that all the “ax” words but the last were borrowed into Proto-Indo-European (or some of its immediate daughters) from languages of quite different structure – and that axes were important trade items in early Indo-European communities.

Patterns like these can be exploited to identify probable loanwords in proto-lexica, provided that the reconstructible structure of the protolanguage is idiosyncratic enough and well enough understood. Sometimes we are even luckier: occasionally two or more protolanguages between which no genetic relationship can be demonstrated exhibit words so similar in form and meaning that some sort of historical relationship can reasonably be inferred, and in such cases borrowing is by far the best hypothesis to account for the similarities. Proto-Indo-European \**peleku-s*, for example, might be connected with Semitic Akkadian *pilaqu* by a chain of lexical borrowings between geographically intermediate languages, and a similar situation probably accounts for the similarity between Proto-Indo-European \**tauros* and Proto-Semitic \**ṭauru* “bull.” (The archeology of Proto-Indo-European groups puts direct borrowing out of the question in both cases; see in general Mallory 1989.) But it should be clear that questions like these must be addressed on a case-by-case basis.

## 6. CONCLUSIONS

The foregoing discussion can be summarized as follows.

1. Because protolanguages can only be reconstructed inferentially from the data of their historical descendants, the methodology according to which the inferences are drawn is a matter of the utmost importance. Over the past century and a half, historical linguists have evolved a thoroughly reliable methodology for reconstructing protolanguages, based on what is known about language structure and language change from contemporary and historical records.
2. Reconstruction of the phonological systems of protolanguages and the phonological shapes of their words and affixes is based on the observation that sound change is overwhelmingly regular; that basis guarantees both the realism and the rigor of phonological reconstruction. Proposed alternative methodologies (such as reliance on the phonetic similarity of words) are unrealistic and lack rigor; for both reasons they are unacceptable.
3. Reconstruction of the phonological rules and morphological systems of protolanguages (and of their syntax, when that becomes feasible) depends crucially on a thorough knowledge of the daughter languages and of relevant areas of linguistic theory and description, since in these components of the grammar we have no regular patterns of change that can be exploited directly.
4. Reconstruction of proto-lexica, especially with regard to the meanings of the lexemes, presents us with the greatest range of possible pitfalls; we can best deal with this situation by adhering rigorously to the regularity of sound change and treating all other aspects of the work with caution.

At this point it should be clear to the reader that rigor, caution, and a general knowledge of linguistics that is as wide as possible are crucial to the reconstruction of protolanguages. Those considerations alone refute the claims of some scholars to have established so-called long-range genetic groupings of languages that include several recognized families (such as “Nostratic” and “Amerind”), because *without exception* their work fails to meet the best standards of mainstream historical linguistics (see refutations in, *inter alios*, Campbell 1988, Vine 1991). It is also true that simple, robust statistical tests reveal such claims to be untenable (see Ringe 1995, 1996a, 1999; Nichols and Peterson 1996 with references).

The prospects for further reconstruction of protolanguages are therefore clear: we must continue as we have begun, paying close attention to the data, employing the comparative method with the greatest rigor we can muster, and bringing to bear on our reconstructions of morphology and lexica all the tools of linguistics at our disposal.

## Bibliography

- Anthony, D. 1995. “Horse, wagon and chariot: Indo-European languages and archaeology.” *Antiquity* 69:554–565.
- Baltin, M., and C. Collins (eds.). 2001. *The Handbook of Contemporary Syntactic Theory*. Oxford: Blackwell.
- Bammesberger, A. (ed.). 1988. *Die Laryngaltheorie*. Heidelberg: Carl Winter.
- Bergmann, R., H. Tiefenbach, and L. Voetz (eds.). 1987. *Althochdeutsch, Band I: Grammatik, Glossen und Texte*. Heidelberg: Carl Winter.
- Bergsträsser, G. 1928. *Einführung in die semitischen Sprachen*. Munich: Hueber.
- Birkhan, H. 1970. *Germanen und Kelten bis zum Ausgang der Römerzeit*. Vienna: Österreichische Akademie der Wissenschaften.

- Bloomfield, L. 1946. "Algonquian." In Osgood 1946, pp. 85–129.
- Campbell, L. 1988. "Review of Greenberg 1987." *Language* 64:591–615.
- Cowgill, W. and M. Mayrhofer. 1986. *Indogermanische Grammatik*, vol. I, 1–2. Heidelberg: Carl Winter.
- Delbrück, B. 1893–1900. *Vergleichende Syntax der indogermanischen Sprachen*. Strasburg: Trübner.
- Foley, W. 1986. *The Papuan Languages of New Guinea*. Cambridge: Cambridge University Press.
- Fox, A. 1995. *Linguistic Reconstruction*. Oxford: Oxford University Press.
- Gamkrelidze, T. and V. Ivanov. 1973. "Sprachtypologie und die Rekonstruktion der gemeinindogermanischen Verschlüsse." *Phonetica* 27:150–156.
- Greenberg, J. 1987. *Language in the Americas*. Stanford: Stanford University Press.
- Harris, A. and L. Campbell. 1995. *Historical Syntax in Cross-Linguistic Perspective*. Cambridge: Cambridge University Press.
- Hock, H. 1986. *Principles of Historical Linguistics*. Berlin: Mouton de Gruyter.
- Hoenigswald, H. 1960. *Language Change and Linguistic Reconstruction*. Chicago: University of Chicago Press.
- Hopper, P. 1973. "Glottalized and murmured occlusives in Indo-European." *Glossa* 7:141–166.
- Janda, R. and B. Joseph. Forthcoming. *A Handbook of Historical Linguistics*. Oxford: Blackwell.
- Kimball, S. 1988. "Analogy, secondary ablaut, and \*OH<sub>2</sub> in Common Greek." In Bammesberger 1988, pp. 241–256.
- Kiparsky, P. 1968. "Tense and mood in Indo-European syntax." *Foundations of Language* 4:30–57.
- Kroch, A. 2001. "Syntactic change." In Baltin and Collins 2001, pp. 699–729.
- Labov, W. 1994. *Principles of Linguistic Change. Volume 1: Internal Factors*. Oxford: Blackwell.
- Lehmann, W. 1974. *Proto-Indo-European Syntax*. Austin: University of Texas Press.
- Lunt, H. (ed.). 1964. *Proceedings of the Ninth International Congress of Linguists*. The Hague: Mouton.
- Mallory, J. 1989. *In Search of the Indo-Europeans*. London: Thames and Hudson.
- Mayrhofer, M. 1986. "Lautlehre (segmentale Phonologie des Indogermanischen)." In Cowgill and Mayrhofer 1986, pp. 73–216.
- Meid, W. 1987. "Germanische oder indogermanische Lautverschiebung?" In Bergmann *et al.* 1987, pp. 3–11.
- Melchert, H. 1987. "Proto-Indo-European velars in Luvian." In Watkins 1987, pp. 182–204.
- . 1994. *Anatolian Historical Phonology*. Amsterdam: Rodopi.
- Nichols, J. 1990. "Linguistic diversity and the first settlement of the New World." *Language* 66:475–521.
- Nichols, J. and D. Peterson. 1996. "The Amerind personal pronouns." *Language* 72:336–71.
- Nussbaum, A. 1986. *Head and Horn in Indo-European*. Berlin: de Gruyter.
- Osgood, C. (ed.). 1946. *Linguistic Structures of Native America*. New York: Viking Fund.
- Peters, M. 1980. *Untersuchungen zur Vertretung der indogermanischen Laryngale im Griechischen*. Vienna: Österreichische Akademie der Wissenschaften.
- Porzig, W. 1954. *Die Gliederung des indogermanischen Sprachgebiets*. Heidelberg: Carl Winter.
- Puhvel, J. 1984. *Hittite Etymological Dictionary*, vols. I–II. Berlin: Mouton.
- Ringe, D. 1995. "'Nostratic' and the factor of chance." *Diachronica* 12:55–74.
- . 1996a. "The mathematics of 'Amerind'." *Diachronica* 13:133–154.
- . 1996b. *On the Chronology of Sound Changes in Tocharian*, vol. 1. New Haven: American Oriental Society.
- . 1999. "How hard is it to match CVC-roots?" *Transactions of the Philological Society* 97:213–244.
- . Forthcoming. "Internal reconstruction." To appear in Janda and Joseph (in press).
- Rix, H., M. Kümmel, T. Zehnder *et al.* 2001. *Lexikon der indogermanischen Verben*. Wiesbaden: Dr. Ludwig Reichert.
- Schindler, J. 1976. "Diachronic and synchronic remarks on Bartholomae's and Grassmann's Laws." *Linguistic Inquiry* 7:622–637.
- Schmitt, R. 1989. "Andere altiranische Dialekte." In R. Schmitt (ed.), *Compendium Linguarum Iranicarum*, pp. 86–94. Wiesbaden: Dr. Ludwig Reichert.
- Stang, C. 1972. *Lexikalische Sonderübereinstimmungen zwischen dem Slavischen, Baltischen und Germanischen*. Oslo: Universitetsforlaget.
- . 1974. "Ieur. \*swad-/ \*swād-." *Norsk Tidsskrift for Sprogvidenskap* 28:99–101.

- Stevens, A. 1968. *Madurese Phonology and Morphology*. New Haven: American Oriental Society.
- Szemerényi, O. 1967. "The history of Attic οὔς and some of its compounds." *Studi micenei ed egeo-anatolici* 3:47–88.
- Vine, B. 1991. "Indo-European and Nostratic." *Indogermanische Forschungen* 96:9–35.
- Wackernagel, J. 1892. "Über ein Gesetz der indogermanischen Wortstellung." *Indogermanische Forschungen* 1:333–436.
- Watkins, C. 1963. "Preliminaries to a historical and comparative analysis of the syntax of the Old Irish verb." *Celtica* 6:1–49.
- . 1964. "Preliminaries to the reconstruction of Indo-European sentence structure." In Lunt 1964, pp. 1035–1045.
- Watkins, C. (ed.). 1987. *Studies in Memory of Warren Cowgill*. Berlin: de Gruyter.
- Yoshida, K. 1990. *The Hittite Mediopassive Endings in -ri*. Berlin: de Gruyter.

# Index of general subjects

- Abecedaria 494, 500, 517,  
791, 816
- Abella 813
- Abū Ṣalābīkh 20, 218, 220
- Acco 365
- Achaemenids (Achaemenid)  
60–61, 63, 64, 65, 717,  
718, 722, 740
- Acrophonic principle 164,  
367, 946, 1110, 1111
- Adab 20
- Adiego, Ignacio-Javier  
609
- Adria 840
- Adriatic Sea 11, 840
- Aegean Sea 10, 12, 943, 944
- Aeolians (Aeolian) 650
- Aeolus 650
- Aeschylus 740
- Aetolian Confederacy 651
- Afghanistan 60
- Africa 138, 139, 141, 160,  
366, 383, 427, 789
- Aḥiram 367
- Ahura Mazdā 762
- Åkerblad, Johan David 169
- Akhmim 162
- Akkad 19, 21
- Akkade 62, 95, 218
- Akkadia (Akkadian) 96, 226
- Akkadograms 66, 90, 553
- Aksum (Aksumites) 427,  
430, 431, 433, 435, 449,  
451
- Alalakh 97, 220
- Albania 12
- Alcaeus 651
- Aleppo 391, 392, 396
- Alexander the Great 12, 13,  
14, 22, 362, 779
- Alexandria 162
- Al-Ġawf 490
- Algeria 138, 139, 366
- Alphabet (*See also* Epichoric  
alphabets) 5, 7, 8, 67, 97,  
162, 163, 367, 558, 592,  
599, 602, 609, 610, 615,  
648, 653, 674, 744, 766,  
779, 781, 791, 813,  
814–817, 818, 820, 821,  
840, 844–845, 846, 847,  
848, 855, 882, 883, 884,  
886, 908, 909, 910, 911,  
912, 923, 924, 945–946,  
968, 969, 1055
- Alps 9, 845, 883, 944
- Al-ʿulā 455, 490
- Amarna *See* Tell el-Amarna
- Amathus 6
- Amenhopis III 96
- Americas 39
- Amorites 288
- Analects of Confucius* 989,  
990
- Anatolia 12, 22, 63, 96, 117,  
163, 219, 220, 383, 386,  
387, 551, 553, 558, 573,  
574, 576, 591, 601, 609,  
648, 650, 651, 655, 658,  
777, 779, 780, 923
- Ancient Lybian 7
- Anglo-Saxon 908
- Anittas 551
- Anshan 60–61, 62, 63
- Antioch 97, 389
- Antiquity 1
- Apennines 10, 789
- Appian of Alexandria 11
- Apulia 14, 812
- Arabia (Arabian) 427, 431,  
454, 455, 485, 488, 489,  
490, 492, 494, 498, 500,  
514, 520, 522, 529
- Arabic Vulgate Bible 430
- Arabs (Arab) 13, 392, 490,  
493, 500, 780
- Arameans 367, 391, 392,  
395
- Arcadia 650
- Arcarnanian 13
- Archaic Cypro-Minoan 5
- Arezzo 10
- Argišti I 119
- Ariaramnes 718
- Aristophanes 657
- Armavir Blur 60, 63
- Armenia (Armenian) 60, 63,  
119, 922, 923, 940
- Arno River 943
- Arsames 718
- Artaxerxes I 717
- Artaxerxes II 718
- Artaxerxes III 717, 718
- Arzawa 574
- Asia (Asian) 2, 21, 39, 138,  
141, 160, 188, 386, 674,  
702, 715, 717
- Asia Minor 365, 378, 381,  
551, 573, 777, 779, 857,  
967
- Aśoka 674, 700, 1015
- Asomtavruli 968, 969
- Assur 777
- Assurnaširpal II 120
- Assyria (Assyrian) 22, 60, 61,  
62, 63, 95, 96, 119, 120,  
218, 219, 320, 391, 454,  
573, 777
- Assyrian Empire 63, 362
- Aṣṭādhyāyī* 673
- Aśvaghoṣa 702
- Asyut 162
- Athenaeus 13
- Athens 13, 609, 615, 650,  
655
- Athos 777

- Atiedian brotherhood 813  
 Atlas Mountains 138  
 Attica 614, 615, 650, 655  
 Augustus 789, 790  
*Avesta* 743, 762  
 Awām-temple 454  
 Awan 61  
 Azerbaijan 61, 119  
  
 Bab-al-Mandab 427, 455  
 Babylon 22, 62, 64, 219, 387, 427  
 Babylonia (Babylonian) 20, 22, 27, 61, 62, 63, 96, 98, 218, 219, 220, 320, 321, 322, 499  
 Babylonian Empire 392, 492  
 Babylonian Exile 319, 321  
 Babylonian Talmud *See* Talmud  
 Babylonian vocalization *See* Vocalization  
 Baghdad 19  
 Balearic Islands 365, 366  
 Balikh River 391  
 Balkan peninsula 614, 648, 651, 656  
 Balkans 11, 12, 15, 551, 777, 779  
 Baluchistān 61  
 Bantia 813, 819  
 Baratella 840  
 Barth, Jacob 335  
 Barthélemy, Jean 169  
 Bartholomae, Christian 759  
 Bastam 133  
 Bay of Attaleia 591  
 Beeston, A. F. L. 454  
 Behistun (= Bisitūn) 223  
 Beja 139  
 Belize 1041  
 Belluno 840  
 Beqa' Valley 366  
 Beth-Shemesh 290  
 Bhāsa 702  
 Bible 289, 319, 320, 322, 323, 333, 430, 433, 809, 881, 882, 907, 923, 933, 936, 940  
 Biconsonantal symbols 163  
 Bilecik 778  
 Bisitūn (= Behistun) 63, 91, 717, 720, 721, 722, 923  
  
 Bithynia 778  
 Black Sea 535, 882, 883, 967  
 Blaëne 585  
 Boeotia 651  
 Boğazköy 552, 555, 578, 585, 777, 778  
 Bonampak 1062  
 Bopp, Franz 2  
 Bosphorus 551  
 Bossert, Helmut 579  
 Botorrita 858  
 Boustrophedon 366, 367, 369, 455, 495, 781, 817, 845, 910  
 Brāhmī 674, 700, 702, 1015, 1016–1018  
 Briges 777  
 Britain 8, 789, 857  
 Broken writings 67, 228  
 Bronze Age 670, 777  
     Early Bronze Age 96  
     Late Bronze Age 95, 117, 163, 288, 319  
     Middle Bronze Age 95  
 Bronze script 990, 991  
 Bruges 777  
 Brugmann, Karl 2, 1116  
 Brugoi 777  
 Brukes 777  
 Bruttium 812, 816  
 Buddhism (Buddhist) 701, 715, 1016  
 Burgundians 882  
 Burnouf, E. 723  
 Büyükkale 777  
 Byblos 365, 366, 376, 378, 380, 388  
 Byzantine Greek 11  
 Byzantium (Byzantine) 430, 882  
  
 Cadore 843, 845  
 Caesar 806  
 Cairo 498  
 Calabri 14  
 Calabria 14  
 Calymna 651  
 Cambyzes 721  
 Cameroon 139  
 Caminata Aiyar, U. V. 1016  
 Campania 812, 813, 816, 943  
 Canaan (Canaanite) 365, 386, 387, 388  
  
 Canary Islands 139  
 Cappadocia 778, 779  
 Capua 812, 813, 943  
 Caria 609, 779  
 Carinthia 840  
 Carthage 366, 367  
 Cartouche 169, 1050, 1051, 1058  
 Caslar Situla 9  
 Celtoligurians 11  
 Celts (Celtic) 11, 944  
 Central African Republic 139  
 Ceylon 431  
 Chad 139  
 Champollion, Jean-Francois 169  
 Charles V 883  
 Chiapas 1041  
 China (Chinese) 427, 988, 989, 990  
 Chinese Turkestan 536  
 Chios 650  
 Choghā Zambil 62, 91  
 Christianity (Christian) 167, 188, 394, 430, 449, 779, 790, 809, 882, 907, 923, 940, 967, 969, 986  
 Cicero 790  
 Cilicia 923  
 Cimmerians 779  
 Classifier (*See also* Determinatives) 20, 25  
 Cleopatra 13  
 Clerical script 992  
 Cnidos 658  
 Colchis 967, 986  
 Colfiorito 816  
 Confucius 988  
 Consonantal principle 164  
 Consonantal scripts 98, 289–291, 321, 367, 368, 388, 393, 431, 460, 494, 615, 653, 718  
 Copan 1062  
 Coptic script 167  
 Corsica 943  
 Cortona 816, 943  
 Çorum 97  
 Cos 650, 651, 658, 670  
 Cretan Hieroglyphic 4  
 Cretan Pictographic 4  
 Crete (Cretan) 4–5, 614, 651, 655  
 Crimea 883



- Crimean Goths 883  
 Croatia 11  
 Cryptographic 167  
 Cumae 812, 944  
 Cuneiform (*See also*  
     Peripheral cuneiform;  
     Proto-Cuneiform) 15,  
     19–20, 21, 22, 24–25, 26,  
     27–28, 31, 48, 53, 65,  
     67–69, 71, 96, 97–98, 99,  
     120, 122, 141, 163, 165,  
     220, 222, 223, 224–226,  
     281–7, 289–290, 387,  
     388, 552, 553, 554, 555,  
     557, 558, 559, 569, 573,  
     574, 576, 578, 585, 586,  
     591, 594, 717, 718, 719,  
     721, 722, 724  
 Cyblos 6  
 Cycladic Islands 650, 651,  
     655  
 Cyme 670  
 Cypriot Syllabary 615, 653  
 Cypro-Minoan 5, 289, 653  
     Cypro-Minoan 1 6  
     Cypro-Minoan 2 6  
     Cypro-Minoan 3 6  
 Cyprus (Cypriot) 5–6, 219,  
     365, 367, 376, 614–635,  
     650, 653, 658, 670  
 Cyrus II (the Great) 63, 722  
  
 Dadan 490, 517  
 Damascus 391, 392, 498  
 Danube 11, 12, 882  
 Darius I 63, 91, 223, 717, 718,  
     721, 722, 724, 725, 726,  
     923  
 Daskyleion 778  
 Daunuii 14  
 Dead Sea Scrolls 320, 392  
 Decipherment 3, 4, 25, 167,  
     169, 222–223, 290, 388,  
     609, 670, 700, 722, 782,  
     859, 1051, 1053, 1068,  
     1072–1075, 1110  
 Deir ‘Alla 392  
 de Landa, Diego 1054, 1055  
 Delbrück, Berthold 546, 547,  
     548  
 Delos 455  
 Delphi 651  
 Demosthenes 13  
 Demotic 7, 163, 166, 141, 991,  
     992  
 Denmark 882, 909, 910  
 de Saussure, Ferdinand 681  
 Determinatives 25, 66–67, 69,  
     164, 166, 224, 553, 578,  
     1050  
 Determiners *See*  
     Determinatives  
 Deucalion 650  
 Devanāgarī 674, 702  
 Dharmapada 701  
 Dhorme, Edouard 6  
 Diacritics 224, 226, 322, 323,  
     326, 394–395, 674, 816,  
     1018  
 Digambara Jains 702  
 Digraphs 791, 792, 845, 860,  
     924, 969  
 Dinar 779  
 Diodorus Siculus 9, 10  
 Dionysius of Halicarnassus  
     943, 944  
 Dionysius Thrax 926  
 Diyala 259  
 Djibouti 160  
 Dniester 882  
 Dokimeion document  
     779  
 Dorians (Dorian) 650  
 Dorus 650  
 Ductus 554  
 Dūmā 490  
 Dunand, Maurice 6  
  
 Early Song period 989  
 Ebla 20, 30, 218, 223  
 Egypt (Egyptian) 2, 13, 22,  
     96, 138, 139, 140, 163,  
     169, 179, 187, 219, 222,  
     365, 371, 383, 386, 387,  
     389, 392, 430, 455, 490,  
     499, 500, 517, 609, 656,  
     657, 659, 789, 943  
     Lower Egypt 162  
     Middle Egypt 162  
     Upper Egypt 162  
 Elam (Elamite) 60–63, 64, 65,  
     89, 90, 718  
 Elmalı 778  
 El Sitio 1075  
 Elymais 90  
 Emar 97, 220  
 Emilia 943  
 England 908  
 English Channel 857  
*Enūma Eliš* 219  
 Epichoric alphabets 615, 653,  
     658, 659  
     Blue alphabets 655  
     Dark blue 655  
     Light blue 655  
     Fayum alphabet 656  
     Green alphabets 655  
     Red alphabets 655  
 Epi-Olmecs (Epi-Olmec) 15,  
     1044, 1053, 1071, 1105  
 Epirus 11  
 Eritrea 139, 141, 428  
 Eshnunna 21  
 Eskişehir 779  
 Eskişehir-Dorylaion 778  
 Estrangelo 393  
 Este 840, 841, 843, 845, 846,  
     852  
 Eteo-Cypriots 6  
 Ethiopia (Ethiopian) 7, 13,  
     139, 140, 141, 160, 427,  
     428, 430, 433, 435  
 Ethiopian Orthodox Church  
     428, 430  
 Etruria 791, 844, 944, 945, 946  
 Etruscans 10, 656, 816, 845,  
     943  
 Euboea 650  
 Euphrates 96, 97, 119  
 Eurasia (Eurasian) 22, 967  
 Eurasian steppe 1122  
 Europe (European) 2, 7–9, 11,  
     789, 857, 882, 883, 907,  
     988  
 Eusebius 371  
 Ezana I 430  
 Ezana II 430, 431  
  
*Fangyan* 989  
 Fara 20, 218  
 Fārs 60, 61, 62, 63, 65, 717  
 Fayum 656  
 Fertile Crescent 95, 96, 117  
 First Intermediate Period 161  
 Forrer, Emil 577, 578, 585  
 France 366, 789  
 Futhark 909–911  
     Older futhark 908  
     Younger futhark 908  
  
 Gail River 840  
 Galatia (Galatian) 778, 779,  
     780, 787  
 Galilee (Galilean) 320, 321,  
     330

- Gāthās* 673, 742, 743, 747, 762  
 Gaul 858, 877, 882  
     Transalpine Gaul 860  
 Gebal 6  
 Gebleh Plain 289  
 Gediz River 601  
 Gelb, Ignace 578  
 Georgia 967, 968, 986  
 Gepids 882  
 Germanic 882, 907  
 Germany (German) 882, 907  
     High German 911  
 Getty Museum 656  
*Gilgamesh* 219  
 Ginsberg, H. L. 335  
 Girnar 674  
 Girsu 21  
 Glossenkeil 387, 552  
 Glyphs 1046, 1050, 1055, 1109  
 Gómez-Moreno, M. 859  
 Gordion 777, 778, 779  
 Gospel of John 882  
 Goths (*See also* Crimean Goths) 882, 907, 911  
 Grammarians  
     Arab 488, 493, 497, 499, 525  
     Georgian 976  
     Greek 645, 926  
     Indian 728  
     Latin 616  
     Medieval Hebrew 346  
     Prākṛit 701, 715  
     Sanskṛit 702  
 Graphemes 290, 291  
 Great Phrygia 778  
 Great Rift Valley 427  
 Greece (Greek) 6, 9, 10, 11, 12, 13, 14, 64, 161, 365, 367, 368, 601, 614, 615, 616, 621, 648, 650, 651, 653, 777, 779, 787, 789–804, 809, 812, 860, 923, 943, 944, 967, 986  
 Greek choral lyric 651  
 Greek epic 650, 651  
 Grimm, Jacob 2, 898  
 Grotefend, G. F. 722  
 Guatemala 1041, 1042, 1071, 1075  
 Gubbio 813  
 Gulf of Antalya 591  
 Gulf of Corinth 651  
 Gulf of Fethiye 591  
 Gulf of Telmessos 591  
 Gurina 840  
 Hadramawt 455  
 Hadrian 789  
 Haft Tepe 62  
 Ḥaldi 120  
 Halicarnassus 650  
 Halys River 585  
 Hamadān 60  
 Hammurabi 62, 64, 220, 221, 224, 228, 254  
 Han period 988, 992, 996, 1010  
 Haram 519, 521  
 Harrapan 3  
 Hatti (Hattic) 219, 585  
 Hattuša 96, 98, 99, 107, 108, 111, 113, 114, 220, 387, 552, 574, 576, 577, 578, 585, 777  
 Hattusilis I 551  
 Hawkins, David 579  
 Ḥawr Rūrī 455  
 Hebrews 13, 367, 391  
 Hegrā 492  
 Hellenicus 943  
 Hellen 650  
 Hellenism (Hellenistic) 64, 163, 319, 371  
 Hellespont 650  
 Hermann, Gottfried 618  
 Hermopolis 392  
 Hermos River 601  
 Herodotus 7, 12, 13, 500, 517, 650, 740, 777, 922, 943, 944  
 Heruls 882  
 Hesiod 614  
 Hesychius 14, 780  
 Hetzron, Robert 429  
*Hexapla* 320, 342, 350  
 Hexaplaric vocalization *See* Vocalization  
 Hieratic 163, 166  
 Hieroglyphs (Hieroglyphic) 2, 6, 7, 121, 161, 163, 165, 166, 167, 169, 553, 554, 576, 577, 578, 1041, 1043, 1044, 1055, 1056, 1057, 1060, 1061, 1062, 1067, 1075  
 Himyar 454, 455  
 Hincks, Edward 223  
 Hinduism (Hindu) 1016  
 Hippocrates 650  
 Ḥismā 492  
 “Hittite Hieroglyphs” 163, 578  
 Hittites (Hittite) 96, 114, 289, 387, 391, 551, 553, 573, 574, 576, 585, 777  
 Hoffmann, Karl 762  
 Homer 11, 614, 779  
 Homophony 66, 226, 1051, 1052  
 Honduras 1041  
 Horapollo 167  
 Horn of Africa 427  
 Hrozný, Bedřich 555, 578  
 Hübschmann, Heinrich 717  
 Hurria (Hurrian) 95, 391, 574  
 Hystapes 718  
 Iapodes 11  
 Iapygia 14  
 Iberia (Iberian) 7, 858, 859, 864, 967  
     Northeast Iberian 7, 8  
     South Iberian 7  
 Iberian peninsula 857  
 Ideograms 165  
 Idria 845  
 Iguvine Tablets 813, 829, 833  
 Iguvium 813  
 Illyrian 11–12, 15  
 India 13, 22, 431, 535, 673, 674, 698, 700, 702, 1015  
 Indian Ocean 431, 455  
 Indo-Europeans 614  
 Indo-Iranians 717  
 Indus Valley 3–4  
 International Phonetic Alphabet 16, 924, 970  
 Ionia (Ionian) 616, 650, 655, 779, 781  
 Iran (Iranian) 2, 3, 22, 60, 61, 62–63, 64, 65, 70, 89, 223, 940  
 Iraq 19, 60, 95, 96, 97, 119, 218, 391, 498  
 Ireland 8  
 Iron Age 11, 12, 163, 319, 320, 321, 330, 361, 362, 365, 777, 840, 920  
 Ishan Miziya 21  
 Isin 21  
 Islam (Islamic) 90, 138, 427, 430, 488, 494, 509, 529

- Israel 141, 319, 320, 321, 365  
 Iṣṭahrī 64  
 Ister 11  
 Isthmus of Tehuantepec 1053  
 Italian peninsula 9, 840, 857  
 Italy 9, 10, 13, 14, 15, 656,  
     789, 812, 813, 814, 816,  
     837, 840, 882, 943, 944  
 ʿIzbet Ṣarṭaḥ 388  
  
 Jabal al-ʿAwd 455  
 Jacobite vocalization *See*  
     Vocalization  
 Jainism (Jain) 701, 702, 715,  
     1016  
 Jebel al-Aqraʿ 288  
 Jebel Ansariyeh 288  
 Jemdet Nasr 20  
 Jerusalem 320, 321, 334, 498,  
     924  
 Jerusalem Talmud *See* Talmud  
 Jewish script 321  
 Jordan 391, 490, 492  
 Josephus 371  
 Judaea 320, 362  
 Judah 319, 320, 321, 362  
 Judaism (Jewish) 319, 394  
 Jones, Sir William 2, 535  
  
 Kaishu 992  
 Kālidāsa 702  
 Kandahar 60  
 Kaneš 219, 573  
 Karatepe 378, 381, 382, 383,  
     579  
 Karlgren, B. 992  
 Kaska (Kaskeans) 551, 585  
 Kaunos Bilingual 609, 610,  
     612  
 Kelišen stele 120  
 Kenya 139, 160  
 Kerala 1015  
 Kermān 65  
 Kermanshah 63  
 Khabur River 96, 391  
 Kharoṣṭhī 674, 700, 701, 702  
 Khuzestān 60, 61, 62–63, 64,  
     70  
 Kircher, Athanasius 169  
 Kirkuk 97  
 Kish 23  
 Kızıl Irmak 585  
 Kizzuwatna 96, 574  
 Knossos 651, 660  
 Konya 779  
  
 Koriwn 923  
 Krahe, Hans 11  
 Kültepe 219, 551  
 Kulturwörter 90  
 Kuşaklı 552  
 Kūtahya 779  
  
 Lachish Ewer 388  
 Lagash 24  
 Lágole di Calalzo 840, 845,  
     847, 853  
 La Graufesenque 878  
 Lake Chad 160  
 Lake Turkana 160  
 Lake Van 119  
 Large seal script (*See also* Seal  
     script) 991  
 Larissa 651, 670  
 Laroche, Emmanuel 577  
 Larsa 20  
 Larsam 96  
 Lassen, C. 723  
 Latini 789  
 Latium 789, 790  
 La Venta 1071  
 Lebanon 6, 365, 366, 386,  
     391  
 Lecce 15  
 Lemnos 10  
 Lesbos 651  
 Létóon Trilingual 591, 592,  
     594, 599, 601  
 Levant 22  
 Libya 139, 366  
 Ligatures 674  
 Ligurians 10  
 Linear A 4, 5, 653  
 Linear B 4, 615, 650, 651, 652,  
     653, 657, 663, 670  
 Linear Elamite 3, 61, 65  
 Lipit-Eshtar 21  
 Livy 10, 13, 14, 808, 840  
 Locris 651  
 Logograms (Logographic) 4,  
     25, 27, 66–67, 69, 97,  
     120, 164, 165, 166, 221,  
     223–224, 226, 229, 245,  
     251, 262, 387, 553, 577,  
     578, 582, 585, 719, 765,  
     766, 1046, 1047, 1050,  
     1052, 1058, 1064, 1075,  
     1076, 1077, 1078, 1099,  
     1109, 1110, 1111  
 Logographic spelling 25–26,  
     54, 1050, 1052  
  
 Lucania 812, 813, 816  
 Lu Fayān 992  
 Lugano 859  
 Lurestān 61  
 Luwiya 585  
 Luxor 367  
 Lycia 591, 609, 778  
 Lydia (Lydian) 601, 609, 779,  
     944  
  
 Macedon 13  
 Macedonia (Macedonian) 12,  
     13, 14, 651, 777, 778, 779  
 Madāʾin Šālīḥ 492  
 Magna Graecia 656  
 Mahābhārata 673  
 Mahāsāmghika-  
     Lokottaravāda school  
     702  
 Mali 139, 160  
 Malta 365, 366  
 Mamertini 812  
 Mani 427  
 Manichaeism 427  
 Mari 20, 96, 240, 253, 288,  
     386  
 Marib 450, 454, 455, 459  
 Marseilles 10  
 Masat 552  
 Masoretes (*See also* Masoretic  
     text) 331, 332, 345  
 Masoretic text 323, 326, 327,  
     329, 331 332, 333, 334,  
     343, 350  
 Matres lectionis 165, 292, 301,  
     321, 323, 346, 368, 381,  
     388, 394, 395, 397, 458,  
     495  
 Mauritania 139  
 Mauryan Empire 700  
 Mayans (Maya) 15, 1042,  
     1043, 1044, 1050, 1051,  
     1053, 1068, 1075  
 Medes 13, 717  
 Media 717  
 Mediterranean 13, 95, 96, 97,  
     160, 188, 365, 366, 382,  
     386, 391, 431, 490, 648,  
     789, 864, 910, 1037  
 Meher kapısı 120  
 Melos 651, 655, 658  
 Memphis 162  
 Mencius 989, 990  
 Meriggi, Piero 578  
 Meroë 7

- Mesoamerica 3, 15, 1041,  
     1053, 1071–1083, 1092,  
     1099, 1109, 1110, 1111  
 Mesoamerican chronology  
     Early Classic Period 1043,  
     1044, 1047, 1050, 1052,  
     1059, 1060, 1061, 1062,  
     1063  
     Late Classic Period 1046,  
     1050, 1060, 1062, 1063  
     Late Preclassic Period 1042,  
     1043, 1044, 1053  
     Middle Preclassic Period  
     1042  
 Mesopotamia  
     (Mesopotamian) 2, 15,  
     19, 21, 53, 61–63, 64,  
     65–67, 68, 69–70, 71, 89,  
     95–96, 163, 165, 218,  
     219, 220, 223, 271, 277,  
     386, 387, 490, 552  
 Messana 812  
 Messapii 14  
 Mexico (Mexican) 1041, 1109  
 Midas 777, 780, 786  
 Midas II 777  
 Midas City 778  
 Middle Ages 161, 222, 882  
 Middle Kingdom 161, 166  
 Mije-Sokeans (Mije-Sokean)  
     1071, 1099  
 Miletus 650  
 Minaeans 492  
 Minoan 4, 5, 653  
 Mishnah 320  
 Mita 777, 779  
 Mitanni 95, 96, 117, 188, 219,  
     551  
 Moldova 882  
 Monoconsonantal symbols  
     163  
 Montebelluna 840  
 Moors (Moorish) 882  
 Morocco 138, 139  
 Morphographemic writings  
     229  
 Morpurgo Davies, Anna 579  
 Mount Amanus 391  
 Mount Bermion 777  
 Mount Haemus 12  
 Mount Odra 11  
 Mount Vesuvius 812  
 Mrglovani 968  
 Mtac'mideli, Giorgi 968  
 Muḥammad 64  
 Mukish 289  
 Muqaddasī 64  
 Mursilis I 551  
 Mursilis III 551  
 Muşafir 120  
 Mushkis 777, 779  
 Mxedruli 969  
 Mycenaean (Mycenaean) 4,  
     614, 615, 650, 653  
 Mysia 778  
 Mytilene 670, 671  
 Nabataeans (Nabataean) 492,  
     498  
 Nāgarī 674  
 Nahal Hever 320  
 Napoleon 169  
 Naram-Sim 62  
 Nawas 1071  
 Naxos 650  
 Near East 13, 15, 21, 22, 24,  
     90, 95, 96, 163, 219, 223,  
     391, 392, 424, 430, 449,  
     529, 573, 718, 740, 924,  
     967, 1044  
 Neo-Assyrian 62  
 Neo-Babylonian Empire 321  
 Neo-Elamite 62  
 Neogrammarians 1123  
 Neolithic period 288  
 Nestorian vocalization *See*  
     Vocalization  
 Neumann, Günter 579  
 New Kingdom 161, 162,  
     188  
 New Testament 882  
 Niebuhr, K. 223, 722  
 Niger 139, 160  
 Nigeria 139  
 Nile Valley 162  
 Nilgiri Mountains 1015  
 Nineveh 60, 63  
 Nippur 19, 20, 21, 23, 44,  
     233  
 Niya Documents 701  
 Nola 813  
 Nora 366  
 Normalization 224  
 North Africa 7  
 Norway 907, 909  
 Nubia (Nubian) 7, 163  
 Numerals 66, 719  
 Numidian 7  
 Nusxa-xucuri 969  
 Nuzi 66, 97  
 Oasis of Dedān 455  
 Oasis of Fayyum 162  
 Oasis of Siwa 160  
 Oaxaca 1053, 1109  
 Occam's Razor 722  
 Oderzo 840, 849  
 Ötztir 910  
 Ogham 8, 9  
 Old Akkadian 61–62  
 Old Kandahar 63  
 Old Kingdom 161, 162, 163,  
     166  
 Old Testament 320  
 Olmecs (Olmec) 1071, 1075  
 Olympia 13, 651  
 Oman 141, 455  
 Omo River 139, 160  
 Oppert, J. 223, 723  
 Oracle bone script 990, 991  
 Origen of Caesarea 320  
 Orontes 97  
 Ortaköy 97, 552  
 Orvieto 816  
 Osthoff, Hermann 2  
 Ostrogoths 882  
 Otten, Heinrich 577, 585  
 Padova 840, 841, 847, 849  
 Padua 840  
 Paestum 812  
 Palā 574, 585  
 Palestine 290, 320, 321, 322,  
     366, 367  
 Palestinian vocalization *See*  
     Vocalization  
 Pamphylia 651  
 Pāṇini 673, 680  
 Paphlagonia 574, 585  
 Papyri 657, 659  
 Paros 650  
 Parthia (Parthian) 13, 64, 90,  
     922, 940  
 Pelagonia 777  
 Pella 14  
 Peloponnese 650, 651, 655  
 Pentateuch (*See also*  
     Samaritan Pentateuch)  
     345  
 Peripheral cuneiform 66  
 Persepolis 60, 63, 91, 740  
 Persia (Persian) 61, 63, 219,  
     319, 386, 427, 601, 717,  
     723, 779, 967, 986  
 Persian Empire 321, 362, 392,  
     718, 739, 967

- Persian Gulf 62  
 Perusia 816  
 Petra 492  
 Peucetii 14  
 Phaistos Disk 4  
 Pharsalus 651  
 Philip II 12  
 Philip V 13  
 Philip of Macedon 13, 14, 614  
 Philistines 365  
 Phoenicia (Phoenician) 6,  
     365, 368, 382, 383, 388,  
     391, 393, 615  
 Phonetic borrowing principle  
     991  
 Phonetic complements 163,  
     224, 553, 554, 578, 1046,  
     1047, 1050, 1051, 1052,  
     1056, 1058, 1059, 1077,  
     1099, 1111  
 Phonetic compounding 991  
 Phonetic signs 223  
 Phonograms 25, 163–164, 166  
 Phrygia (Phrygian) 12, 777,  
     778, 779, 780, 786, 922  
 Piave River 840, 843  
 Picenum 812  
 Pictograms (*See also*  
     Pictographs) 221, 367  
 Pictographs (Pictographic)  
     15, 121, 163, 165, 221,  
     990, 991  
 Picts 9  
 Pinara 591  
 Pinyin 988  
 Pisa 10  
 Plato 662  
 Plautus 366, 802  
 Plene spelling (*See also*  
     Scriptio plena) 68, 98,  
     121, 124, 228, 234, 237,  
     327, 355, 460  
 Pliny the Elder 10, 15  
 Plutarch 13  
 Po River 840  
 Po Valley 855  
 Poland 882  
 Polybius 10, 13, 14  
 Polyphony 66, 226  
 Pompeii 812, 813  
 Pontine marshes 789  
 Portugal 7  
 Praenesta 790  
 Prometheus 650  
 Propontide 779  
 Proto-Byblic 6  
 Proto-Cuneiform 65  
 Proto-Elamite 2, 3, 61, 65  
 Provence 943  
 Pseudo-Hieroglyphic 6  
 Pteria 778  
 Ptolemies 13, 163, 166, 167,  
     169  
 Punctuation (*See also* Syllabic  
     punctuation) 221, 455  
 Puzrish-Dagan 21  
 Puzur-Inšušinak 61, 65  
 Pylos 651, 660  
 Pyrgi 943  
*Qieyun* 992  
 Qin dynasty 988, 992, 1007  
 Qumran 320, 322, 324, 325,  
     330, 342, 343, 350, 392  
 Raeti 9, 10  
*Rāmāyaṇa* 673  
 Ramesside 161  
 Ramlat as-Sabʿatayn 454  
 Ras Ibn Hani 288  
 Ras Shamra 288  
 Rask, Rasmus 2, 723  
 Rawlinson, H. C. 223, 723  
 Ray, John 609  
 Rebus principle 164, 221, 553,  
     991  
 Red Sea 140, 427, 428, 431,  
     455  
*Ṛgveda* (*See also* *Rig-veda*)  
     673, 691  
 Rhodes 365, 651  
 Rif Mountains 138  
*Rig-veda* (*See also* *Ṛgveda*)  
     673, 682, 688, 705  
 Roman Empire 1, 789, 812,  
     813, 882  
 Roman Republic 789, 812,  
     813, 846  
 Rome (Roman) 9, 10, 11, 13,  
     161, 163, 166, 167, 321,  
     362, 377, 427, 488, 495,  
     500, 651, 656, 766, 779,  
     784, 789, 790, 806, 809,  
     812, 813, 816, 837, 840,  
     855, 857, 858, 859, 860,  
     862, 864, 867, 877, 882,  
     907, 923, 944, 946,  
     1037  
 Rosenkranz, Bernhard 577  
 Rosetta Stone 169, 1054  
 Rubʿ al-ḥālī 454  
 Rudradāman 674  
 Rules of Pachomius 430  
 Rumania 882  
 Runes 882, 883, 907, 908, 911,  
     912, 913, 918  
 Sabaʿ 455  
 Sabeans 454  
 Sabellian 812  
 Šabwa 455  
 Sahara 138  
 Sainovics 535  
 St. Augustine 366  
 St. Jerome 330  
 St. Mesrop 923, 924  
 Sakākā 490  
 Salāla 455  
 Salentini 14  
 Sallentine peninsula 14  
 Samʿal 392  
 Samaria 320  
 Samaritan Pentateuch 319,  
     321, 323, 330, 342, 349  
 Samārum 455  
 Samnites (Samnite) 812  
 Samnium 812, 816  
 Samos 650  
 Samukha 96  
 Sapinuwa 96, 97  
 Sappho 651  
 Sardinia 365, 366  
 Sardis 601, 601  
 Sardure II 119  
 Sargon II 777, 779  
 Sargon of Akkade 218  
 Sasanians (Sasanian) 764, 940  
 Satricum 789  
 Saudi Arabia 455, 490, 492,  
     514  
 Sayce, A. H. 609  
 Šayhad 454  
 Scandinavia 907, 908, 909  
 Schleicher, August 2  
 Schürr, Diether 609  
 Scotland (Scots) 9  
 Scriptio plena (*See also* Plene  
     spelling) 122, 556, 566,  
     577, 585  
 Sea Peoples 365, 552, 574  
 Seal script 992  
 Segmental characters 858, 859  
 Seistān 65  
 Selucids (Selucid) 219  
 Semagram 164, 165

- Semitic 391, 573  
 Septuagint 325  
 Serabit al-Khadem 367  
 Shang dynasty 990, 991  
 Sheba 431  
 Shen Yue 993  
 Shevoroshkin, V. 609  
*Shi jing* 992, 994  
 Shulgi 64  
*Shuo wen jie zi* 992  
 Shuruppak (= Suruppak) 20  
 Sicani 9  
 Sicily 9, 365, 366, 656, 812  
 Sidon 378  
 Sinai peninsula 367, 387, 498  
 Sippar 23  
 Şirwāh 454  
 Siyannu 289  
 Skeireins 882  
 Slovenia 11  
 Small seal script (*See also* Seal script) 992  
 Smyrna 650  
 Socrates 662  
 Solomon 431  
 Somalia 139, 160  
 Sorrentine peninsula 812  
 Spain (Spanish) 7, 365, 366, 789, 882, 1109  
 Spring and Autumn period 990, 991  
 Śri Lanka 701, 1015  
 Strabo 7, 11, 14  
 Strahlenberg 535  
 Subir 96  
 Sudan 7, 138, 140, 160  
 Suleiman the Magnificent 883  
 Sumer 19, 20, 21, 26  
 Sumeria (Sumerian) 2, 61, 64, 96, 220  
 Sumerograms 33, 66, 69, 97, 103, 120, 223, 553, 554, 570  
 Suppiluliumas I 551  
 Suppiluliumas II 552, 574  
 Suruppak (= Shuruppak) 218  
 Susa 60, 61–63, 65, 91  
 Susiana 60, 61  
 Śvetāmbara Jains 702  
 Sweden 882, 907, 909  
 Syllabary (Syllabic script) 6, 26, 70, 98, 120, 226, 227, 228, 229, 431, 494, 553, 555, 557, 577, 578, 579, 585, 609, 615, 621, 653, 658, 674, 946, 1016, 1017, 1078  
 Syllabic punctuation 845–846, 946  
 Syllabic spelling 26–27, 29, 54, 66, 69, 165, 224, 335, 1047, 1049, 1050, 1051, 1057, 1058, 1064  
 Syllabic symbols (*See also* Syllabograms) 67, 71, 97, 291, 578, 859, 1046, 1050, 1051, 1052, 1075  
 Syllabograms 4, 25, 26, 27, 66, 72, 120, 223, 229, 1076–1077, 1110, 1111  
 Syracuse 668  
 Syria (Syrian) 6, 13, 20, 26, 29, 30, 62, 95, 96, 163, 218, 220, 223, 391, 393, 488, 490, 492, 498, 519, 551, 552, 553, 573, 576  
 Syria-Palestine 6, 430  
 Tabula Bantina 813  
 Talbot, W. H. Fox 223  
 Tall-i Malyān 60  
 Talmud 330  
     Babylonian 320  
     Jerusalem 320  
 Tamil Nadu 1015  
 Tamilians 1037  
 Tamiẓ Eẓuttu. 1016–1018  
 Tang period 990, 1010  
     Late Tang period 989  
 Tanzania 139, 160  
 Targums 416  
 Taurini 11  
 Tavşanlı 778  
 Taymā' 490  
 Tekor 924  
 Tel Fekheriye 392, 396  
 Tell al-Maskhūṭah 517  
 Tell Asmar 20  
 Tell Beydar 20  
 Tell Brak 20, 96  
 Tell el-Amarna 96, 219, 288, 334, 371, 387, 388  
 Tell Mardikh 223  
 Tell Sukas 365  
 Thasos 658  
 Thebes (Theban; Egypt) 161, 162  
 Thebes (Greece) 651  
 Theocritus 668  
 Theodoric 882  
 Thera 651, 655  
 Thessaly 651  
 Thrace (Thracians) 11, 12, 777  
 Thracian Sea 650  
 Thucydides 657  
 Tiber River 789, 812, 943–946  
 Tiberian vocalization *See* Vocalization  
 Tiberias 323  
 Tibet 22  
 Tifinagh 139  
 Tifinigh 139  
 Tiglathpilesar I 777  
 Tigris 95, 777  
 Tigonanu 96  
 Tikal 1041  
 Timna' 455  
 Tišatal 96  
 Tlapanecs 15  
 Tlos 591  
 Tofseta 320  
*Tolkappiyam* 1016  
 Trajan 789  
 Transcaucasia  
     (Transcaucasian) 967, 969  
 Transcription 23, 69, 70, 72, 79, 90, 99, 100, 122, 221, 224, 230, 232, 233, 325, 369, 371, 373, 381, 395, 500, 724, 765, 766, 786, 813, 947, 962, 1046, 1049  
     Broad transcription 553  
     Conventional transcription 69  
     Morphological transcription 69  
     Narrow transcription 553  
 Transjordan 319, 320, 361, 386  
 Transliteration 25, 29, 30, 69, 70, 99, 100, 122, 169, 223, 224, 323–324, 396, 399, 403, 433, 495, 497, 500, 501, 604, 765, 766, 767, 909  
 Tres Zapotes 1071  
 Triconsonantal symbols 163  
 Trieste 840  
 Troglodytes 13

- Trojan War 840  
 Tuareg 139  
 Tunisia 139, 366, 943  
 Turdetan 7  
 Turgutlu 601  
 Turkey 60, 95, 96, 218, 391, 392, 585, 967, 1122  
 Turks (Turkish) 883  
 Tuşpa 119  
 Tušratta 96  
 Tuttul 96  
 Tuz Gölü 779  
 Tyana 777, 778  
 Tyrrhenian Sea 943  
  
 Ugarit 6, 66, 97, 98, 99, 220, 288, 289–290, 315, 319, 431, 552  
 Ukraine 882  
 Umbria 812  
 Umma 21  
 Uppsala 882  
 Uqair 20  
 Ur 20, 21, 23, 61, 64, 95, 376  
 Urartian 60, 119, 391  
 Urkeš 96  
 Uruk 19, 20, 21, 22  
 Uşak 778  
 Üyücek 778  
  
 Van 60  
 van Busbecq, Ogier 883  
 Vandals 882  
 Varro 9, 789  
 Vedaś 673  
 Veii 845  
 Veneti 840, 855  
 Venetia 840, 843, 844, 845, 846  
 Ventris, Michael 650  
 Veracruz 1071  
 Vergil 805  
 Verner, Karl 2  
 Vezirhan 778  
 Vicenza 840  
*Vidēvdāt* 743  
 Virāma 674  
 Visigoths 882  
  
 Vocalization 455–456, 490, 494–497, 552–555, 577–579, 585–586, 592, 602, 609–610, 615–616, 653–657, 674–675, 702, 718–723, 744–745, 764–766, 781, 791, 814, 818, 842, 843, 844–846, 847, 858–860, 883–884, 909–911, 923–924, 945, 968, 990–992, 1016–1018, 1046–1055, 1072–1078  
     Babylonian vocalization 552–555, 577–579, 585–586, 592, 602, 609–610, 615–616, 653–657, 674–675, 702, 718–723, 744–745, 764–766, 781, 791, 814, 818, 842, 843, 844–846, 847, 858–860, 883–884, 909–911, 923–924, 945, 968, 990–992, 1016–1018, 1046–1055, 1072–1078  
     Hexaplaric vocalization 653–657, 674–675, 702, 718–723, 744–745, 764–766, 781, 791, 814, 818, 842, 843, 844–846, 847, 858–860, 883–884, 909–911, 923–924, 945, 968, 990–992, 1016–1018, 1046–1055, 1072–1078  
     Jacobite vocalization 818, 842, 843, 844–846, 847, 858–860, 883–884, 909–911, 923–924, 945, 968, 990–992, 1016–1018, 1046–1055, 1072–1078  
     Nestorian vocalization 909–911, 923–924, 945, 968, 990–992, 1016–1018, 1046–1055, 1072–1078  
     Palestinian vocalization 322 1016–1018, 1046–1055, 1072–1078  
     Tiberian vocalization 323–324, 326, 329, 331, 332, 340, 343, 346, 394–395, 399  
 von Humboldt, Wilhelm 535  
 Vowel points (*See also* Vocalization) 322  
  
 Wackernagel, Jacob 546, 759, 1123  
 Wadi Bayḥān 455  
 Wadi Hadramawt 455  
 Wadi Ḥarīb 455  
 Wadi Maḏāb 455  
 Wadi Murabbaʿat 320  
 Wadi Ramm 492  
 Warring States period 988, 989, 992, 1010, 1012  
 Western Ghats 1015  
 Western Zhou period 990  
 Winckler, Hugo 555  
 Woito River 140  
 Word division (dividers) 67, 121, 134, 221, 495, 508, 592, 602, 609, 719, 722, 817, 946  
 Writing 2  
 Writing systems 24–27, 65–69, 97–98, 120–121, 163–169, 220–229, 231, 237, 241, 245, 251, 269, 273, 288, 289–291, 292, 321–324, 367–369, 387–388, 391, 393–395, 396, 397, 399, 428, 431–432, 433, 436, 455–456, 490, 494–497, 552–555, 577–579, 585–586, 592, 602, 609–610, 615–616, 653–657, 674–675, 702, 718–723, 744–745, 764–766, 781, 791, 814, 818, 842, 843, 844–846, 847, 858–860, 883–884, 909–911, 923–924, 945, 968, 990–992, 1016–1018, 1046–1055, 1072–1078  
     Wulfila 882, 883, 884, 885, 886, 887, 904, 907, 911  
     Würmlach 840  
  
 Xanthos 591  
 Xanthos stele 592  
 Xanthus 650  
 Xerxes I 717, 724, 726  
 Xinjiang Uygur 536  
 Xu Shen 992  
  
 Yalburt 591  
 Yang Xiong 989  
*Yasnas* 743  
*Yašts* 743  
 Yaxchilan 1062  
 Yazilikaya 577  
 Yeha 427  
 Yemen (Yemeni) 141, 427, 454, 455, 485, 488, 519, 521  
 Young, Thomas 169  
 Yuan period 989  
 Yucatan peninsula 1041  
  
 Ẓafār 454, 455  
 Zab Valley 119  
 Zagreb mummy 943, 958, 961  
 Zagros Mountains 95, 120  
 Zagros valleys 62  
 Zapotecs (Zapotec) 1109  
 Zarathuštra 673, 742, 743  
 Zoroastrianism 742, 743, 762, 764



# Index of grammar and linguistics

- Ablaut 252, 301, 304, 307, 311, 335, 402, 409, 540, 541, 561, 567, 625, 626, 627, 637, 675, 679, 728, 753, 795, 868, 887, 893, 898, 902, 912, 915, 916, 929, 979, 980, 1125
- Absolute chronology 398, 1113
- Absolute state *See* State
- Accent 30, 372, 539–540, 556, 557, 587, 596, 604
  - Expiratory accent 949
  - Pitch accent 540, 557, 619, 680, 705, 725, 753, 794, 1118
  - Stress accent 73, 100, 123, 145, 173, 234, 332–333, 400, 439, 540, 557, 601, 659, 753, 793, 794, 798, 822, 848, 864, 927
  - Tonal accent 659, 725, 993, 995, 997, 1110
- Acrostic 567, 627
- Adjectivals (Chinese) 1005
- Adjectives (*See also*
  - Comparative adjectives; Superlative adjectives) 37, 74, 175, 247, 296–297, 314, 373, 418, 441–442, 505–506, 683, 1001, 1033, 1065, 1068, 1088
- Adverbials (Chinese) 1004, 1005
- Adverbs 37, 83, 114, 133, 260–261, 308–309, 415, 448, 546, 642, 683, 737, 903, 1002, 1005, 1034
- Agglutinating morphology 401, 557, 703, 704, 727, 819, 950
  - 22, 73, 101, 543, 912, 928, 951, 973, 975, 1019, 1029, 1059
- Agreement 47, 86, 116, 135, 226, 247, 262, 264, 295, 297, 300, 314, 345, 356, 360–361, 418, 419, 420, 421, 483–484, 505, 521, 523, 526, 561, 572, 581, 583, 589, 628, 647, 697, 714, 727, 738, 753, 808, 833, 853, 873, 874, 877, 918, 963, 986, 1035, 1062, 1087, 1096, 1110
- Aktionsart 565
- Allocutive 73
- Allophones 72, 233, 1079
- Allophonic variation 70, 98, 397, 863–864
- Alpha-thematic morphology 637, 637
- Amphikinetic 627
- Analogy 236, 299, 348, 353, 355, 380, 412, 534, 538, 539, 595, 620, 627, 628, 690, 708, 709, 749, 750, 751, 758, 773, 865, 866, 950, 951, 953, 1010
- Analytic morphology (*See also* Isolating morphology) 161, 162, 179, 180, 795, 995
- Anaphora (*See also* Pronouns [Anaphoric]) 581, 784
  - Long-distance anaphora 647
- Anaphoric pronouns *See* Anaphora; Pronouns (Anaphoric)
- Anaptyxis 100, 123, 235, 245, 328, 329, 330, 331, 340, 401, 557, 703, 704, 727, 819, 950
- Annexation 525
- Antecedents 248, 251, 264, 272–273, 300, 313, 314, 315, 343, 480, 481, 484, 528, 572, 630, 647, 697, 832, 833, 836, 963, 974
- Anusvāra 677, 704
- Aorist tense *See* Tense stems
- Aoristic aspect *See* Aspect
- Aphaeresis 596
- Apocope 29, 605, 867, 949, 951, 952, 953, 954, 957
- Apodosis 271, 310, 312, 313, 422, 477, 479
- Apophony 31, 540, 887
- Apposition 264, 918, 1091
- Archaisms 31, 573, 742, 761
- Archiphoneme 723
- Areal features 572
- Articles 179, 242, 346, 731, 784, 896, 934, 938, 962, 972
  - Definite articles 179, 301, 308, 361, 377–378, 381, 461, 488, 505, 507, 517–518, 525, 631, 666, 667, 686, 738, 897, 932, 973, 974, 985
  - Indefinite articles 179, 346, 403
- Aspect 38, 39–41, 80, 110, 128, 151, 174, 177, 303–304, 347–348, 411, 565, 581, 588, 632–633, 688, 710, 731, 732, 756, 802, 827, 869, 901, 932, 933, 934, 936, 972, 982, 1060, 1062, 1083–1084, 1096, 1098, 1110, 1120
- Aoristic 633

- Durative 347, 977  
 Imperfective 39, 40, 41, 47, 151, 178, 259, 304, 346, 347, 378, 411, 565, 581, 588, 633, 732, 976, 1060, 1062, 1120, 1121  
 Infectum 827, 828  
 Nonimperfective 303, 565  
 Perfective 39, 40, 41, 47, 151, 178, 259, 303, 304, 346, 347, 378, 411, 632, 732, 869, 976, 978, 1060, 1062, 1120, 1121  
 Perfectum 827, 828  
 Prospective 178  
 Punctual (Punctiliar) 347, 976, 977, 980  
 Stative 1120, 1121  
 Assimilation 579, 603, 666  
 Assimilation 100, 124, 143, 232, 233, 238–239, 248, 250, 253, 255, 258, 261, 269, 300, 301, 325, 351, 354, 370, 381, 400–401, 404, 409, 439, 459, 498, 501, 502, 513, 538, 557, 587, 595, 638, 677, 678, 704, 705, 706, 750, 752, 862, 951, 958, 1019  
 Reciprocal assimilation 239  
 Assyrian vowel harmony *See* Vowel harmony  
 Asyndeton 269, 271, 274, 312, 480, 481, 484, 589, 739, 873, 963, 1036  
 Athematic morphology 105, 541, 542, 544, 545, 561, 567, 626, 627, 635–636, 667, 683, 686, 689, 706, 733, 784  
 Attic declension 626, 628  
 Attic future *See* Tense stems (Future)  
 Attic reduplication 638  
 Augment 545, 682, 689, 707, 711, 732, 757, 780, 785, 922  
 Babylonian vowel harmony *See* Vowel harmony  
 Bahuvrīhi *See* Compounds  
 Basic conjugations *See* Basic verb stems  
 Basic verb stems (*See also* Derived verb stems) 511, 512  
 G stem 153, 252–255, 302, 349, 378, 379, 380  
 Qal stem 349  
 Bghadhkhpath 497  
 Binyānim (*See also* Derived verb stems) 346, 352, 355  
 Borrowings (*See also* Loanwords) 14, 54, 648, 697, 715, 739, 740, 749, 779, 780, 809, 837, 877, 922, 929, 964, 1011, 1015, 1037, 1099, 1125  
 Bound state *See* State  
 Breaking (Brechtung) 887, 892  
 Broken plurals 148, 440, 460  
 Caland System 543  
 Calques 37, 86, 88, 804, 809, 837, 940, 969  
 Case 31, 34–35, 74, 107–108, 126, 146, 148, 241, 295, 314–315, 337, 373, 374, 441, 442, 461, 502, 505, 543, 560–561, 581, 588, 596, 605, 624, 684, 695, 707, 727, 753, 784, 796, 822, 831, 832, 848, 852, 865, 892, 913, 929, 938, 951–953, 963, 971, 972–973, 1021–1025, 1059  
 Case attraction 647, 938  
 Casus obliquus 770  
 Casus rectus 770  
 C-command 1036  
 Centum 536, 614, 742–748, 842  
 Chiastic concord (*See also* Gender polarity) 262, 263, 298, 419  
 Classifiers 1010  
 Cleft sentences 176, 183, 268  
 Clitic doubling 876  
 Clitics 154, 268–269, 299–300, 452, 462, 540, 548, 573, 587, 590, 592, 598, 607, 645, 660, 696, 737, 759–761, 786, 798, 806, 868, 870, 872, 874, 875, 876, 888, 932, 938, 962, 978, 981, 985, 998, 1019, 1031, 1032, 1035, 1060, 1062, 1065, 1066, 1080, 1110, 1119  
 Enclitics (*See also* Enclitic chains) 49, 82, 84, 97, 100, 108, 110, 111, 115, 127, 129, 130, 134, 147, 239, 246, 268, 269, 299, 300, 301, 308, 310, 342, 374–376, 404, 406, 444, 450, 468, 472, 502, 506–507, 526, 545, 547, 555, 556, 560, 562, 563, 564, 570, 571, 572, 582, 587, 589, 590, 597, 598, 606, 612, 619, 629, 632, 645, 660, 686, 709, 730, 731, 736, 737, 752, 755, 771, 774, 775, 786, 799, 888, 904, 905, 914, 919, 949, 952, 954, 955, 962, 963, 1090–1092, 1110  
 Proclitics 76, 260, 268, 299, 338, 339, 345, 346, 357, 367, 381, 415, 416, 450, 598, 645, 696, 786, 1081, 1110  
 Cognates 14, 377, 1115, 1119, 1121, 1124, 1125  
 Cohortative 42, 113, 348–349  
 Comparative adjectives 74, 147, 248, 297, 402, 423, 543, 561, 629, 642, 685, 709, 730, 797, 825, 897, 1033  
 Comparative method of historical linguistics 2, 534, 944, 1113, 1115, 1116, 1118, 1123, 1126  
 Compensatory lengthening 236, 237, 330, 351, 622–623, 661–662, 663, 679, 769, 795, 887, 890  
 Complementizers 178, 548, 646, 759, 1066  
 Complements 1004–1005  
 Compounds 31, 38–39, 83, 261, 311, 542, 569, 582, 589, 598, 643, 693–694, 713, 735, 804–805, 831, 852, 916–917, 935–936, 996, 1064  
 Bahuvrīhi 694, 697, 713  
 Copulative 643, 693, 713, 735, 936  
 Dependent 693

- Compounds (*cont.*)  
 Descriptive 693  
 Determinative 643, 682, 713, 735  
 Dvandva 693, 713, 871  
 Endocentric 643, 936  
 Exocentric 542, 643, 804, 916, 936  
 Nominal 693  
 Paral 693  
 Possessive 682, 697, 713, 735  
 Tatpuruṣa 693, 713  
 Verbal 693  
 Verbal-governing 804  
 Concord 38, 547, 937–938  
 Conditional clauses 52, 271–272, 313, 359, 422, 477–479, 647, 669, 1031, 1036  
 Conditional mood *See* Mood  
 Connective particles *See* Particles  
 Consonantal length 228, 233  
 Consonants 28–29, 70–72, 98–99, 121–122, 169–172, 230–231, 235–239, 291–292, 324–325, 333–334, 369–370, 388, 395–397, 433–436, 457–458, 497–502, 535–537, 555, 579–580, 586–587, 592, 603–604, 611, 616, 657–659, 676–677, 703–705, 723–724, 725–726, 745, 747–751, 766–767, 782–783, 792–793, 818–819, 820–821, 846–847, 861, 862, 863, 884–885, 912, 924–926, 947–948, 970, 992–993, 994–995, 1018, 1055, 1078, 1110  
 Construct chains 311, 338, 339, 360, 361, 373, 419, 463, 483  
 Construct state *See* State  
 Contract verbs 634, 636, 667  
 Convergence 601, 790  
 Converted imperfect 347, 348, 358  
 Converted perfect 347, 348–349, 358  
 Coordinate clauses (*See also* Coordination) 117, 135, 312, 422, 1065–1066  
 Coordination (*See also* Coordinate clauses) 51, 87, 155, 184–185, 269–270, 357–359, 474, 589, 598, 645, 669, 738, 854, 873, 985  
 Copulative compounds *See* Compounds  
 Creolization 427  
 Deaspiration 620, 638  
 Declension 149  
 Declined state *See* State  
 Definite articles *See* Articles  
 Deictic pronouns *See* Deixis; Pronouns  
 Deixis 308, 581, 686, 784, 798, 932, 1066, 1110  
 Delocutive 73  
 Demonstrative adjectives *See* Pronouns  
 Demonstrative pronouns *See* Pronouns  
 Deontic modality *See* Mood  
 Depalatalization 949  
 Dependent compounds *See* Compounds  
 Deponent verbs 687, 803, 869, 933, 934  
 Derivation 542–543, 544–545  
 Primary 542  
 Secondary 542  
 Derivational classes (*See also* Derived verb stems) 445  
 Derivational morphology 32–33, 568–569, 830–831, 871, 935  
 Derived conjugations *See* Derived verb stems  
 Derived verb stems (*See also* Basic verb stems) 153, 302–303, 352–355, 380, 409–411, 466, 511–513  
 C stem 153, 378, 380  
 D stem 153, 256, 302, 378, 380  
 Ethpa‘al / ‘Ithpa‘al 409, 410  
 Ethpə‘el / ‘Ithpə‘el 409, 410  
 ‘Ettaph‘al / ‘Ittaph‘al 409–410  
 Haph‘el / ‘Aph‘el 409, 410  
 Hip‘il 354  
 Hitpa‘el 354  
 Hitpolel 355  
 Hop‘al 355  
 L stem 302  
 Minor stems 410–411  
 N stem 153, 255–256, 302, 378, 380  
 Nip‘al 353  
 Pa‘el 409, 410  
 Pə‘al 409, 410  
 Pi‘el 353  
 Polal 355  
 Polel 355  
 Pu‘al 353  
 R stem 153, 302  
 Rare stems 258  
 Š stem 257, 302  
 t stems 257, 302, 378, 380  
 Descriptive compounds *See* Compounds  
 Desiderative mood *See* Mood  
 Deśya 715  
 Determinate state *See* State  
 Determinative compounds *See* Compounds  
 Determinative-relative pronouns *See* Pronouns  
 Determiners 914, 917  
 Dialect geography 591, 652  
 Diglossia 161  
 Diphthongization 617  
 Direct discourse 89  
 Discontinuous morphemes 35, 51  
 Dissimilation 237–238, 329, 389, 401, 517, 538, 620, 638, 660, 948  
 Qatqat-qitqat dissimilation 334, 341  
 Doric future *See* Tense stems (Future)  
 Durative aspect *See* Aspect  
 Dvandva *See* Compounds  
 Emphatic consonants 16, 140, 143, 227, 230, 237, 324, 396, 435, 498, 500  
 Emphatic pronouns *See* Pronouns  
 Emphatic state *See* State  
 Empty words (Xuzi) 1001–1003  
 Enclitic chains 570, 572  
 Enclitics *See* Clitics

- Endocentric compounds *See* Compounds
- Epenthesis 727, 747, 748, 752–753, 911
- Epistemic modality *See* Mood
- Ergative shift 1084, 1085, 1098
- Ergativity 22, 35, 36, 37, 42, 47, 50, 107, 110, 111, 116, 126, 129, 131, 132, 133, 148, 547, 564, 774, 972, 974, 975, 976, 1060, 1062, 1066, 1081, 1084, 1085, 1089, 1090, 1098, 1110
- Split ergativity 22, 560, 572, 714, 972, 1060, 1062
- Exclusive pronouns *See* Pronouns
- Exocentric compounds *See* Compounds
- Expiratory accent *See* Accent
- Extraposition 183, 582, 606
- Family tree 535
- Flectional morphology (*See also* Fusional morphology) 174
- Focalization 176, 183
- Focus 44, 759, 760, 774, 806, 919, 1066, 1088
- Fractions *See* Numerals
- Fricativization 670
- Fronting (*See also* Left dislocation; Verb movement) 450, 582, 598, 606, 759, 760, 761, 774, 853, 919, 1096
- Full-grade 679
- Full words (Shizi) 1000–1001
- Fusional morphology 145, 174, 241, 293, 335, 401, 543, 559, 588, 624, 795, 822, 865, 928, 951, 1019, 1033
- Future tense *See* Tense stems
- G stem *See* Basic verb stems
- Gemination (*See also* Consonantal length) 438, 536, 538, 541, 556, 572, 595
- Generalization (*See also* Regularization) 636, 743, 770, 771
- Gender 31, 35, 73, 102, 124, 146, 147–148, 151, 174, 241, 242, 252, 295, 303, 314–315, 336, 346, 372, 374, 378, 401, 402, 408, 411, 417, 418, 419, 420, 421, 439, 441, 460, 462, 502–503, 505, 509, 516, 521, 523, 526, 542, 543, 559, 581, 588, 589, 596, 605, 624, 683, 707, 714, 727, 753, 784, 796, 822, 848, 865, 892, 913, 929, 951, 1020, 1059, 1082, 1110
- Gender polarity (*See also* Chiastic concord) 469
- Genitive chains 264
- Gerundives 568, 692, 713, 803, 804, 807, 830
- Gerunds 451, 683, 692, 696, 713, 714, 804, 807
- Glottalic consonants (*See also* Emphatic consonants) 324
- Glottalic hypothesis 536, 1116
- Glottalization 16, 140, 435
- Glottochronology 1071, 1111
- Grammaticalization 180, 875, 897, 916, 1003, 1007, 1008, 1010
- Grammatischer Wechsel 887, 902
- Guṇa 679, 686
- Hendiadys 382, 422
- Heteroclites 568, 628, 730, 731
- Hic et nunc particle *See* Particles
- Ḥi-conjugation 566–568, 573, 581, 589, 597, 1120, 1121
- Homonymy 224, 957
- Hypercorrection 29, 702, 783, 795
- Hyperlong vowels (*See also* Ultralong vowels) 892
- Hypotaxis 185, 357, 547
- Hysterokinetic 431, 893
- Imperatives 43, 81, 113, 131, 254, 305, 348, 378, 380, 414, 445, 466, 515–516, 639, 668, 690, 712, 958, 1026–1027, 1060, 1083, 1084
- Imperfective *See* Aspect
- I-mutation 597, 606
- Inchoatives 110
- Inclusive Pronouns *See* Pronouns
- Incorporation 1084–1085
- Indefinite articles *See* Articles
- Indefinite pronouns *See* Pronouns
- Indefinite relative pronouns *See* Pronouns
- Indicative mood *See* Mood
- Infectum *See* Aspect
- Infinitives 114, 152, 155, 178, 180, 181, 260, 265–266, 307, 380, 414–415, 422–423, 445, 451, 467, 472, 482–483, 484, 545, 568, 582, 589, 598, 606, 641, 668, 683, 692, 713, 734, 738, 758, 772, 773, 804, 807, 808, 813, 830, 834, 870, 897, 898, 899, 934, 935, 959, 961, 973, 984, 1031, 1032
- Infinitive absolute 307, 349, 353, 378, 379, 380, 382, 423
- Infinitive construct 349, 353, 378, 379, 380
- Injunctive mood *See* Mood
- Innovations 345, 534, 535, 540, 572, 573, 591, 731, 736, 813, 814, 817, 820, 864, 903, 1122
- Interjections 83, 115
- Internal derivation 542
- Internal reconstruction 534, 1113
- Interrogative pronouns *See* Pronouns
- Intonation 30, 234, 540
- Irregularities 1115
- Isogloss 135–136, 277, 288, 315, 658, 661, 669, 739, 780, 843
- Isolating morphology (*See also* Analytic morphology) 770
- Jussive 113, 131, 348–349, 382, 413, 445, 466, 515, 527

- Koines 908  
 Language death 1112  
 Language families 1113, 1115  
 Laryngeals 537, 558, 573, 587, 594, 597, 623, 636, 638, 680, 681, 747–748, 842, 889, 890, 1113, 1116, 1118, 1124  
 Left dislocation (*See also* Fronting) 267, 876  
 Length metathesis *See* Metathesis  
 Lengthened-grade 679  
 Lenition 558, 587, 749, 768, 864  
 Leveling 1118  
 Lexicalization 32, 45, 179, 180, 410, 445, 773, 996, 1097  
 Lingua franca 13, 14, 219, 289, 386, 391, 392, 424, 427, 717, 718, 740, 967  
 Linguistic continuum 814  
 Loanwords (*See also* Borrowings) 14, 19, 26, 29, 30, 33, 51, 54, 66, 90–91, 99, 117, 125, 136, 235, 240, 276, 277, 315, 336, 394, 411, 424–425, 428, 436, 499, 500, 569, 576, 583, 586, 599, 648, 715, 766, 791, 809, 837, 855, 884, 922, 923, 926, 939, 940, 943, 944, 947, 948, 965, 986, 1012, 1015, 1068, 1125  
 Locutive 73  
 Long-distance anaphora *See* Anaphora  
 Measure words 1010  
 Mergers 229, 231, 232, 235, 239, 290, 292, 321, 324, 333, 345, 370, 534, 555, 559, 560, 617, 620, 675, 676, 678, 679, 783, 793, 794, 825, 842, 843, 861, 863, 891, 910, 911, 912, 951, 1058, 1117, 1118  
 Metathesis 101, 124, 239, 354, 400, 411, 459, 512, 928  
 Length metathesis 795  
 Quantitative metathesis 626, 662, 665  
 Meter 540, 601, 605  
 Mi-conjugation 566–567, 581, 589, 597, 1120, 1121  
 Mimation 54, 232, 237, 242, 245, 246, 248, 461, 484  
 Monophthongization 293, 307, 329, 331, 501, 502–507, 617, 659, 676, 724, 726, 783, 793, 821, 825, 847, 861, 862, 892, 911, 951, 962  
 Mood (*See also* Imperatives; Jussive; Precative; Prohibitives; Vetitive) 38, 41–43, 80, 131–133, 304–305, 515–516, 544, 565–566, 581, 588, 597, 632, 690–691, 712, 732, 756, 771, 785, 799, 826, 834, 850, 869, 897, 932, 956, 958–959, 1000, 1026–1030, 1060, 1083, 1110  
 Conditional mood 114, 132  
 Deontic mood 41–42  
 Desiderative mood 114  
 Epistemic mood 41–42  
 Indicative mood 42, 112, 151, 515, 1028  
 Injunctive mood 639, 691  
 Necessitative mood 958  
 Optative mood 42, 43, 80, 113, 131, 516, 527, 544, 639, 691, 712, 1027–1028, 1060, 1083, 1084, 1105  
 Potential mood 113  
 Subjunctive mood 42, 43, 515, 527, 544, 639, 668, 691, 958, 1083  
 Mora 618, 619, 1019  
 Morpheme boundaries 232, 239, 677, 927, 950, 971, 1049, 1059  
 Morpheme slots 975, 978–981  
 Morphological restructuring 770  
 Morphophonemics 44, 46, 47, 223, 323, 327, 397, 556, 604, 605, 677, 683, 706, 794, 887, 929, 951, 952, 1018, 1026, 1057–1058, 1118  
 Morphophonologies 250, 912  
 Morphosyntactics 597, 669, 674, 1089  
 Multiplicatives *See* Numerals  
 Narten morphology 567  
 Negation 111, 178  
 Negative particles *See* Particles  
 Neutralization 438, 536, 538, 558, 586, 638, 679, 681, 782, 783, 926  
 Nisbation 175  
 Nominal compounds *See* Compounds  
 Nominal declensions *See* Nominal stem-classes  
 Nominal morphology 31–35, 73–74, 102–110, 124–127, 146–150, 175, 241–249, 294–297, 336–342, 372–373, 402–404, 439–442, 460–461, 470, 502–518, 542, 559–562, 580–581, 588, 596–597, 605–606, 611, 624, 663–665, 683–686, 707–709, 727–730, 753–754, 770–771, 784, 796–798, 822–824, 848–850, 865–867, 892–895, 913–914, 951–954, 971–975, 1020–1025, 1059–1060  
 Nominal stem-classes (*See also* Noun classes) 561–562, 589, 684–685, 708, 728–730, 797, 822, 848, 865–867, 893, 913–914, 929–931  
 Nominal stem formation 728, 1020  
 Nonfinite verbals 81, 178, 568, 1036  
 Nonimperpective aspect *See* Aspect  
 Noun chains 32–35  
 Noun classes (*See also* Nominal stem-classes), 624–628, 663  
 Noun endings 373, 685, 708–709, 895  
 Noun formation 125–126, 244, 294, 339–341

- Noun state *See* State  
 Nouns (*See also* Nominal morphology) 1000  
 Number 31, 33–34, 40, 73, 107–108, 126, 146, 148–149, 151, 174, 241, 243, 252, 295, 303, 314–315, 336–337, 346, 372, 374, 376, 378, 401, 403, 411, 418, 420, 421, 440, 441, 442, 460, 502, 503–505, 509, 516, 523, 526, 543, 544, 560, 564, 580, 581, 588, 589, 596, 597, 605, 606, 624, 632, 684, 687, 707, 710, 714, 727, 753, 771, 784, 796, 822, 826, 848, 850, 865, 913, 915, 929, 932, 953–954, 956, 963, 971, 1020, 1060, 1081–1082  
 Numerals 49, 83, 115, 133, 153–154, 182, 262, 297–298, 355–356, 417–418, 448–449, 468–470, 472, 521–523, 569–570, 582, 643–645, 669, 694–695, 697, 714, 735–736, 758, 774, 805, 831–832, 871, 904, 936, 937, 939, 961, 985, 1001, 1059, 1060, 1064, 1092, 1109  
 Cardinal numerals 49, 83, 115, 153, 154, 182, 246, 262, 297, 298, 355, 356, 417–418, 419, 448, 468–469, 521–523, 583, 643, 644, 694, 714, 735, 805, 831, 871, 936, 961, 1065  
 Fractions 469, 736  
 Multiplicatives 470  
 Ordinal numerals 49, 83, 115, 154, 182, 263, 297, 298, 356, 418, 419, 448, 469, 523, 569, 570, 644, 695, 714, 736, 774, 805, 831, 871, 936, 961, 1065, 1092  
 Nūn energicum 348  
 Nunation 242, 245, 248  
 Old Perfective (Egyptian) 178, 180  
 Optative mood *See* Mood  
 Overlong syllables 794  
 Palatalization (*See also* Depalatalization) 380, 660, 676, 678, 747, 816, 818, 947, 952, 953, 962  
 Paral compounds *See* Compounds  
 Parataxis 185, 357, 451, 548  
 Participles 48, 152, 178, 181, 259, 304, 308, 347, 349, 378, 379, 380, 411, 415, 424, 441–442, 448, 467, 503, 504, 516–517, 527–528, 545, 546, 568, 582, 589, 598, 606, 642, 683, 692, 696, 710, 713, 714, 735, 758, 773, 785, 800, 803, 804, 807, 808, 830, 851, 870, 897, 916, 935, 939, 959–960, 984, 1032, 1061  
 Particles 48, 51, 133, 155, 178–179, 180, 239, 251, 260–261, 267, 268, 269, 273, 301, 308–310, 311, 312, 313, 339, 345, 349, 357, 377, 381, 400, 409, 416, 419, 421, 422, 423, 467, 472, 508, 517–521, 559, 570, 581, 582, 590, 598, 607, 612, 666, 669, 680, 731, 732, 737, 759, 760, 761, 799, 806, 826, 876, 904, 937, 978, 981, 1003, 1006, 1007, 1009, 1019, 1060, 1066, 1090  
 Connective particles 570  
 Hic et nunc particle 545  
 Negative particles 77, 133, 416, 422, 468, 472, 527  
 Perfect tense *See* Tense stems  
 Perfective *See* Aspect  
 Perfective adjective 152  
 Perfectum *See* Aspect  
 Periphrastic constructions 565, 688, 689, 690, 712, 731, 735, 773, 800, 803, 804, 808, 828, 835, 897, 903, 939, 1027, 1030, 1031  
 Person 73, 151, 252, 303, 314, 346, 374, 376, 378, 411, 509, 544, 564, 581, 588, 597, 606, 632, 687, 710, 771, 826, 850, 915, 932, 956, 1025, 1081  
 Personal pronouns *see* Pronouns  
 Pharyngealization 16, 140, 324, 396, 401, 435  
 Phonotactic constraints (*See also* Phonotaxis) 331–332, 401, 432, 1056  
 Phonotaxis (*See also* Phonotactic constraints) 99, 122, 234–235, 437–439, 538, 580, 587, 596, 605, 618, 680, 705, 725, 927, 970, 1019, 1080  
 Piginization 427  
 Pitch accent *See* Accent  
 Portmanteau morpheme 687, 709, 1029  
 Possessive adjectives *See* Pronouns  
 Possessive pronouns *See* Pronouns  
 Postposed constituents 589, 1097  
 Postpositions 84, 116, 134, 582, 598, 606, 695, 806, 831, 832, 833, 872, 873, 919, 962, 999, 1021  
 Potential mood *See* Mood  
 Pragmatics 759, 760, 774, 775, 806, 872, 874, 937  
 Precative 80, 255, 466, 691  
 Predicative state *See* State  
 Prefix conjugation 140, 160, 379, 380, 444, 464, 465–466, 472, 514, 515, 527, 528  
 Prehistoric languages 1112–1113  
 Preposed constituents 598  
 Prepositions 84, 134, 155, 184, 260, 268, 309–310, 381, 415, 448, 467–468, 518, 519–520, 582, 598, 645, 695, 731, 780, 786, 796, 806, 833, 872, 875, 905, 918, 919, 935, 937, 938, 999, 1002, 1003, 1004, 1008, 1010  
 Present tense *See* Tense stems  
 Preterite tense *See* Tense stems  
 Preterito-presents 899–901, 903

- Primary derivation *See* Derivation
- Primary tenses *See* Tense
- Primary verb endings *See* Verb endings
- Proclitics *See* Clitics
- Pro-drop 562, 871, 874, 920, 1035
- Prohibitives 81, 255
- Pronominal adjectives *See* Pronouns
- Pronominal state *See* State
- Pronouns 35–37, 72, 77, 1060
- Anaphoric pronouns 85, 128, 249, 570, 582, 607, 683, 686, 695, 696, 697, 709, 731, 737, 738, 784, 785, 798, 799, 825, 826, 868, 932, 1002
- Deictic pronouns (*See also* Deixis) 109, 128, 443, 606, 736, 868, 931, 938, 962
- Demonstrative adjectives 249, 518, 525, 526
- Demonstrative pronouns 75, 140, 151, 176, 249, 301, 305, 308, 315, 345, 346, 357, 367, 376, 408, 418, 462, 463, 471, 505, 508–518, 544, 563, 581, 588, 590, 597, 612, 631, 666, 683, 686, 696, 709, 731, 737, 755, 784, 785, 798, 825, 826, 866, 867, 895, 896, 897, 932, 949, 955, 962, 972, 973, 974, 998, 999, 1002, 1007, 1010, 1026, 1092–1093
- Far demonstratives 376, 408
- Near demonstratives 376, 408
- Determinative-relative pronouns 151, 177, 251, 264, 272, 345, 377
- Emphatic pronouns 825, 868, 870
- Exclusive pronouns 1025
- Inclusive pronouns 1025
- Indefinite pronouns 76, 251, 273, 301, 315, 346, 377, 463, 471, 631, 666, 731, 785, 825, 826, 896, 932, 956, 974, 1093, 1096
- Indefinite relative pronouns 785, 799, 904, 1002
- Interrogative pronouns 37, 110, 151, 177, 251, 301, 315, 345–346, 376, 443, 508, 544, 563, 581, 588, 597, 631, 666, 686, 709, 731, 759, 799, 825, 895, 896, 904, 932, 955, 974, 985, 998, 999, 1002, 1026, 1093–1094, 1096
- Personal pronouns 35, 37, 75, 108–109, 127, 140, 150–151, 175–176, 249–250, 299–300, 314, 342–345, 357, 374–376, 404–408, 414, 415, 418, 420, 442, 462, 470, 481, 506–507, 526, 544, 562–563, 581, 597, 606, 629–630, 665–666, 683, 684, 686, 709, 730, 737, 755, 759, 771, 774, 798, 806, 813, 825, 850, 868, 896, 914, 931, 932, 955, 974, 1001, 1010, 1025, 1098, 1110
- Possessive adjectives 248
- Possessive pronouns 36, 75, 127, 409, 544, 868, 974, 1060
- Pronominal adjectives 544, 630, 686, 709, 731, 850
- Reciprocal pronouns 630, 631
- Reflexive pronouns 36, 77, 150, 251, 408, 544, 630, 631, 647, 666, 785, 808, 813, 825
- Long-distance reflexives 630
- Relative pronouns 32, 37, 76, 110, 128, 177, 180, 186, 251, 272–273, 300, 301, 308, 310, 313, 315, 314, 345, 346, 367, 377, 382, 443, 450, 463, 471, 480, 481, 507–508, 528, 544, 548, 563, 581, 588, 597, 612, 632, 647, 667, 686, 695, 696, 709, 731, 738, 751, 759, 780, 785, 799, 805, 825, 833, 836, 868, 873, 896, 955, 956, 964, 974, 985
- Resumptive pronouns 76, 87, 109, 264, 360, 473, 582, 589, 606, 738
- Prosodic flip 760
- Prosody 1019
- Protasis 271, 310, 422, 477, 1031, 1036
- Proterokinetic 627, 893, 895
- Prothetic vowels 367, 401, 522, 557, 705, 951
- Protolanguages 534, 1113, 1114–1115, 1117, 1118, 1119, 1123, 1125–1126
- Pseudoparticiple 178
- Psilotic 658
- Punctual (punctiliar) aspect *See* Aspect
- Purpose clauses 380, 381
- Qal stem *See* Basic verb stems
- Qualitatives (Egyptian) 180, 181
- Quantitative metathesis *See* Metathesis
- Qatqat-qitqat dissimilation *See* Dissimilation
- Reciprocal assimilation *See* Assimilation
- Reconstructed languages 1112
- Reconstruction 497, 534, 1075
- Reflexive pronouns *see* Pronouns
- Regularity of sound change 1115–1116, 1119, 1121, 1126
- Regularization (*See also* Generalization) 901, 902
- Relative chronology 398, 652, 1113, 1115
- Relative clauses 51, 52, 88, 117, 135, 186, 245, 251, 254, 272–273, 313, 314, 360, 382, 450, 463, 473, 479–481, 484, 528, 547, 548, 572, 582, 589, 598, 606, 647, 696, 697, 738, 806, 832, 836, 963, 974, 1005, 1032, 1091, 1097
- Relative pronouns *see* Pronouns



- Resumptive pronouns *see*  
     Pronouns  
 Retentions 534  
 Rheme 178, 187  
 Rhotacism 580  
 Right dislocation 572  
 Root constraints 541, 682  
 Root structure 541, 1059  
 Ruki 677, 747, 749, 750, 771  
  
 Sandhi 239, 677–678, 681,  
     743, 745, 1018, 1022,  
     1057, 1058  
 Satem 536, 614, 676, 742, 751,  
     780  
 Schwa secundum 538, 541  
 Secondary derivation *See*  
     Derivation  
 Secondary tenses *See* Tense  
 Secondary verb endings *See*  
     Verb endings  
 Segholates (Segholation) 331,  
     340, 440  
 Segmental loss 100, 236–237  
 Sentence-types (Chinese)  
     1005–1009  
     Complex 1008–1009  
     Copular 1007  
     Exclamatory 1006  
     Imperative 1006  
     Interrogative 1006  
     Passive 1007–1008  
 Sequence of tense 270, 807,  
     835  
 Serial verb constructions 1003,  
     1010  
 Sound change 534  
 Spiritus asper 658  
 Split ergativity *See* Ergativity  
 Splits (phonemic) 534, 621  
 State 146–147, 242, 295,  
     338–339, 373, 403, 418,  
     460–461, 505  
     Absolute state 242,  
         246–247, 295, 296, 338,  
         373, 403, 418, 419, 420,  
         421  
     Bound state (*See also*  
         Construct state) 244–245  
     Construct state (*See also*  
         Construct chains) 146,  
         242, 295, 296, 311, 338,  
         339, 344, 373, 403, 419,  
         441, 461, 480, 481, 484,  
         502  
     Declined state 242–244  
     Determinate state 461, 505  
     Emphatic state 403, 419,  
         420  
     Predicative state 242,  
         245–246  
     Pronominal state 295, 296  
     Stative (Egyptian) 178  
     Stative aspect *See* Aspect  
     Stress accent *See* Accent  
     Strong cases 684, 753  
     Strong declension 896, 897  
     Strong stems 1026  
     Strong verbs 252, 306,  
         446–447, 898–899,  
         902–903, 915  
     Subject-verb inversion 853  
     Subjunctive mood *See* Mood  
     Subordinate clauses (*See also*  
         Subordination) 117, 135,  
         272–276, 312–314,  
         359–360, 422, 834–836,  
         1065  
     Subordination (*See also*  
         Subordinate clauses)  
         51–53, 88, 155, 184–186,  
         258, 357, 450–452,  
         474–479, 646, 696–697,  
         738, 806–808, 873, 876,  
         985, 1096  
     Suffix conjugation 141, 160,  
         178, 181, 379, 380, 444,  
         464, 465, 471, 513,  
         514–515, 516, 526–527,  
         528  
     Superheavy syllables 794  
     Superlative adjectives 32, 74,  
         147, 248, 297, 402, 561,  
         629, 642, 685, 709, 730,  
         797, 825, 897, 1033  
     Supine 568, 804, 807, 835  
     Suppletion 40, 41, 255, 804,  
         868, 897, 974, 1026–1030  
     Syllable boundaries 228  
     Syllable structure 144–145,  
         173, 234, 331–332, 400,  
         437, 618, 680, 725, 794,  
         949, 1056–1057  
     Syncope 123, 144, 240, 247,  
         253, 254, 255, 595, 596,  
         605, 793, 794, 814, 822,  
         848, 912, 944, 949, 950,  
         952, 957, 971  
     Synthetic morphology 161,  
         179, 181, 727  
  
 Tadbhavas 715  
 Tatpuruṣa *See* Compounds  
 Tatsamas 715  
 Temporal clauses 359–360  
 Tense 39, 110, 111, 128, 174,  
     177, 303, 411–413, 445,  
     544, 565, 581, 588, 597,  
     606, 632, 688, 710, 732,  
     756, 771, 785, 799, 827,  
     828–829, 850, 868–869,  
     901, 932, 956, 957–958,  
     1028, 1060  
     Primary tenses 646  
     Secondary tenses 646  
 Tense stems  
     Aorist tense 636–637, 667,  
         690, 934  
     Future tense 636, 667  
         Attic future 636  
         Doric future 667  
     Perfect tense 637–639, 668,  
         688, 802–803  
     Present tense (*See also*  
         Preterito-presents)  
         633–636, 667, 689–690,  
         710–711, 800–802,  
         933–934, 957  
     Preterite tense 711, 957  
 Thematic morphology (*See*  
     *also* Alpha-thematic  
     morphology) 104, 126,  
     252, 541, 542, 544, 545,  
     561, 567, 625, 628,  
     633–635, 667, 668, 683,  
     686, 689, 706, 733, 753,  
     784, 786, 868, 895, 901,  
     902, 1125  
 Theme 178, 183, 187  
 Tmesis 682, 696, 707, 743,  
     904, 905, 978  
 Tonal accent *See* Accent  
 Topicalization 35, 50, 116,  
     134, 175, 183, 267–268,  
     312, 548, 571, 598, 696,  
     737, 759, 852, 962  
 Transitivity 39, 80  
 Triphthongization 328, 329,  
     331, 334  
 Typology 558, 947, 967,  
     1116  
  
 Ultralong vowels (*See also*  
     Hyperlong vowels) 240  
 Umlaut 596, 911, 912, 948,  
     951, 980

- Uniformitarian principle 1112–1116
- Valence 39, 110, 111, 129
- Velarization 16, 435
- Ventive 44, 258
- Verb classes (*See also* Verbal conjugations; Verbal stem-classes) 1061–1063, 1082
- Verb endings 545–546, 640–641, 668, 733–734, 827, 850–851, 869–870, 901, 957, 1029–1030
- Primary 545, 565, 640, 688, 712, 732, 756, 801, 802, 827, 850
- Secondary 545, 565, 640, 688, 711, 712, 732, 756, 801, 802, 827, 850, 901
- Verb inflection 349–352
- Verb movement (*See also* Fronting) 919
- Verbal adjectives 45, 259, 642, 735, 870, 935
- Verbal chains 41–48
- Verbal compounds *See* Compounds
- Verbal conjugations (*See also* Mi-conjugation; Hi-conjugation; Verbal stem-classes; Verb classes) 78, 513–515, 526–527, 566, 589, 800, 828, 850, 932–934, 975–976, 1026–1030
- Verbal-governing compounds *See* Compounds
- Verbal morphology 38–49, 78–82, 110–114, 128–133, 151–153, 177–178, 180–181, 252–260, 302–308, 346–355, 378–380, 409–415, 443–447, 463–467, 471–472, 509–517, 544–546, 563–568, 581–582, 588–589, 597–598, 606, 612, 632–642, 667, 686–692, 709–713, 731–735, 756–758, 771–774, 785, 799, 826–830, 850, 868–870, 897–903, 915–916, 956, 975–984, 1026–1033, 1060, 1082–1089
- Verbal nouns 568, 934, 956, 959–961, 984, 1031, 1033
- Verbal stem-classes (*See also* Basic verb stems; Derived verb stems; Tense stems; Verbal conjugations; Verb classes) 566–567, 869, 915–916
- Verbs (*See also* Verbal Morphology) 1000–1001
- Vetitive 255, 466
- Visarga 677
- Voice 80, 174, 177, 304, 410, 515, 516, 544, 564, 581, 588, 597, 632, 687, 710, 732, 785, 800, 826, 850, 869, 932, 933, 934, 956, 957, 1061, 1078
- Vowel contraction 72, 100, 123
- Vowel harmony 30, 124
- Assyrian vowel harmony 240
- Babylonian vowel harmony 232, 236, 240, 245, 253
- Vowel length 29, 72, 228, 233
- Vowel shift 371
- Vowel weakening 794, 795
- Vowels 29–30, 72, 99, 122, 172–173, 231–233, 240–241, 292–293, 326–330, 334–335, 370–372, 397–400, 437, 458, 502, 537, 556, 580, 587, 595–596, 604, 611, 616–618, 659, 676, 703, 724, 726–727, 746, 751–752, 767, 783, 793, 819, 821–822, 847, 861, 862, 863, 885, 911–912, 926–927, 948, 970, 992, 994, 1018, 1056, 1110
- Vṛddhi 541, 679, 690, 703, 706, 728
- Wave model 430
- Weak cases 684, 753
- Weak declension 896, 897
- Weak roots 174
- Weak stems 1026
- Weak verbs 252, 307, 350–352, 446–447, 460, 466, 898, 899, 903, 915, 916
- Wh-movement 696
- Word boundaries 539, 1058
- Word classes 293–294
- Word formation 30–31, 73, 174–175, 293, 402, 409–411, 540, 559, 682–683, 706–707, 865, 928–929, 950, 1019
- Word order 22, 50–51, 84, 116, 133–134, 154, 182–184, 263–264, 311–312, 356–357, 381, 422, 449–450, 473–474, 524–526, 540, 547, 548, 570–572, 582, 589, 598, 606, 645, 695–696, 736–737, 759, 774–775, 785–786, 805–806, 832–833, 852–853, 871–872, 873, 874–876, 904, 917, 919–920, 937, 961, 985, 997, 1006, 1034, 1065, 1066, 1095–1096, 1110
- Word structure 30, 101, 124, 145–146, 335–336, 372, 460, 971, 1059
- Zero-grade 679

# Index of named linguistic laws and principles

- |   |  |   |
|---|--|---|
| Barth-Ginsberg Law 304, 307,<br>329, 335, 389   | Grassmann's Law 620, 678,<br>682             | Stang's Law 559   |
| Bartholomae's Law 538, 678,<br>679, 750, 1117   | Grimm's Law 888–889                          | Sturtevant's Law 586  |
| Brugmann's Law 747                              | Law of three moras 892                       | Thurneysen's Law 887  |
| Canaanite Shift 293, 334, 371,<br>386, 389, 615 | Lindeman's Law 538                           | Two-Mora Rule 703, 706  |
| Čop's Law 580                                   | Osthoff's Law 623, 663                       | Vendryes' Restriction 875,<br>878   |
| Final Trochee Rule 619                          | Philippi's Law 328, 339                      | Verner's Law 540, 887, 889,<br>910  |
| Geer's Law of Dissimilation<br>237              | Phoenician Shift 371                         | Wackernagel's Law 540,<br>547, 570, 582, 590,<br>645, 696, 737–738,<br>759, 760–761, 775, 806,<br>875 |
| Germanic Consonant Shift<br>888, 889            | Sievers' Law 537, 538, 539,<br>556, 888, 893 |   |
|   | Sievers-Edgerton Law<br>680                  |   |

# Index of languages

- Achaemenid Elamite *See* Elamite 435, 441, 442, 444, 445, 446, 480, 497,  
Aeolic *See* Greek Dialects 500, 501, 552, 553,  
Aequian *See* Sabellian 555, 559, 569, 574,  
Aetolian (*See* Greek Dialects 586, 766, 1125  
[= Northwest Greek Assyrian 120, 125, 144,  
Koine]) 14 219, 220, 221, 226,  
Afar-Saho 139 231, 232, 240–241,  
Afro-Asiatic (*See also* 245, 248, 249, 252,  
Common 253, 255, 256, 257,  
Afro-Asiatic; 259, 260, 263, 268,  
Proto-Afro-Asiatic) 272, 273, 369, 371,  
26, 138–141, 152, 153, 717  
156, 160, 169, 172, Middle Assyrian 219,  
174, 187, 319, 335, 223, 228, 232, 236,  
372, 430, 440, 446, 237, 238, 239, 242,  
535, 967 254, 259, 277  
North Afro-Asiatic 140 Neo-Assyrian 66, 119,  
Agaw (Central Cushitic) 120, 122, 219, 232,  
*See* Cushitic 236, 237, 238, 239,  
Ainu 22 240, 242, 246, 255,  
Akhmimic *See* Coptic 258, 259, 260, 362  
Akkadian 15, 19, 20, 21, 22, Old Assyrian 219, 221,  
23, 26, 27–28, 29, 33, 227, 228, 232, 235,  
34, 35, 36, 37, 38, 39, 237, 238, 239, 240,  
45, 50, 54, 60–61, 62, 242, 243, 246, 249,  
65, 66, 67, 69–71, 72, 250, 254, 259, 260,  
82, 87, 89–90, 95, 96, 267, 268, 270, 273,  
97–98, 103, 104, 105, 277, 551, 553  
106, 115, 117, 119, Post-Old Assyrian 243  
120, 122, 126, 136, Babylonian 60, 143, 219,  
141, 143, 144, 145, 220, 221, 226, 231,  
147, 148, 149, 151, 232, 239, 240, 241,  
152, 154, 155, 169, 244, 249, 253, 255,  
188, 218–280, 246, 256, 257, 259, 263,  
288, 289, 291, 293, 272, 273, 717, 721,  
296, 303, 304, 310, 722  
297, 298, 334, 349, Late Babylonian 219,  
362, 366, 371, 377, 230, 233, 239, 242,  
386, 387, 388, 389, 265, 277, 278, 740  
391, 395, 401, 403, Middle Babylonian 219,  
411, 422, 424, 429, 228, 233, 236, 238,  
239, 242, 248, 249,  
253, 254, 260, 265,  
271, 273, 274, 276,  
277  
Neo-Babylonian 219,  
233, 239, 240, 242,  
246, 249, 265, 278,  
362  
Old Babylonian 62, 64,  
219–220, 222, 226,  
227, 228–229, 232,  
233, 235, 236, 237,  
238, 239, 240, 242,  
243, 246, 247, 248,  
249, 250, 252, 254,  
258, 262, 263, 267,  
270, 273, 276, 277,  
278, 552, 573, 577,  
585  
Post-Old Babylonian  
227, 238, 243, 249  
Standard Babylonian  
220, 233, 239, 242,  
243, 249, 258, 273,  
276, 277  
Old Akkadian 51, 54, 65,  
148, 155, 218, 219,  
221, 227, 228–229,  
231, 232, 235, 236,  
237, 238, 240, 242,  
243, 245, 246, 247,  
249, 250, 251, 252,  
253, 254, 255, 256,  
259, 260, 268, 272,  
273, 276, 277, 377,  
559  
Post-Old Akkadian  
232  
Peripheral Akkadian 220,  
226, 253, 278, 552,  
573  
Alagwa 139

- Albanian 11, 12, 535, 969,  
     1119, 1120, 1122  
 Algonquian (*See also*  
     Proto-Algonquian)  
     535  
 Altaic 1012  
 Amerind 1126  
 Amharic 141, 428, 431, 433,  
     434  
 Ammonite *See* Canaanite  
     Dialects  
 Amorite 61, 277, 288, 362  
 Anatolian (*See also* Common  
     Anatolian;  
     Proto-Anatolian) 4,  
     163, 535, 540, 551,  
     556, 558, 559, 560,  
     562, 563, 567, 569,  
     572, 573, 576, 581,  
     582, 587, 588, 589,  
     590, 591, 596, 597,  
     598, 601, 604, 605,  
     606, 607, 610, 612,  
     645, 787, 875, 922,  
     1114, 1122  
 Ancient Chinese *See*  
     Chinese  
 Ancient Nordic 907  
 Ancient North Arabian *See*  
     North Arabian  
 Ancient South Arabian *See*  
     South Arabian  
 Ancient Venetic *See* Venetic  
 Anzanite 60  
 Apabhramśa *See* Indic  
     (Middle Indic)  
 Arabic (*See also* Proto-Arabic)  
     138, 140, 141, 143,  
     144, 145, 147, 148,  
     153, 154, 155, 160,  
     162, 169, 188, 253,  
     277, 288, 292, 295,  
     296, 300, 305, 308,  
     310, 314, 315, 319,  
     324, 337, 343, 372,  
     373, 377, 381, 382,  
     387, 391, 392, 428,  
     429, 430, 431, 434,  
     435, 436, 441, 442,  
     444, 445, 447, 449,  
     457, 464, 465, 476,  
     484, 485, 488–490,  
     493, 497, 498, 499,  
     500, 501, 502, 503,  
     504, 505, 506, 507,  
     509, 510–515, 516,  
     517, 518–521, 523,  
     525, 526, 529, 1037  
 Classical Arabic 2, 488, 493,  
     497, 498, 500, 501,  
     503, 505, 506, 507,  
     508, 509, 510, 513,  
     514, 515, 516, 517,  
     521–523, 524, 526,  
     527, 528, 529, 530,  
     1117  
 Dialects 488, 498, 500,  
     501, 502, 503, 507,  
     508, 509, 518, 526,  
     529  
 Middle Arabic 488  
 Modern Standard Arabic  
     500  
 Old Arabic 488, 489, 490,  
     498, 502, 517, 518  
 Aramaic (*See also*  
     Proto-Aramaic) 6,  
     90, 141, 143, 144,  
     147, 148, 218, 219,  
     230, 253, 277, 288,  
     293, 299, 308, 315,  
     319, 320, 321, 322,  
     329, 330, 332, 337,  
     338, 342, 361, 362,  
     368, 370, 377, 381,  
     383, 386, 387, 389,  
     391–426, 429, 431,  
     435, 436, 444, 445,  
     485, 488, 489, 497,  
     498, 499, 500, 501,  
     517, 591, 601, 674,  
     700, 717, 718, 737,  
     740, 744, 764, 765,  
     766, 923, 967  
 Biblical Aramaic 416  
 Christian Palestinian  
     Aramaic 392, 393  
 Dialects 392  
 East Arabian Aramaic  
     490  
 Galilean Aramaic (*See*  
     Jewish Palestinian  
     Aramaic) 392  
 Hatran 392  
 Imperial Aramaic 392,  
     396–397, 397, 401,  
     402, 406–407, 408,  
     411, 412, 413, 414,  
     415, 417, 421, 422,  
     423, 424, 425, 500  
 Jewish Babylonian Aramaic  
     392, 405, 406–408,  
     412, 413, 414, 416, 421  
 Jewish Palestinian Aramaic  
     392, 405, 406, 407,  
     408, 412, 414, 415, 417  
 Late Aramaic 392, 393, 394,  
     395, 397, 399–400,  
     402, 403, 405, 406,  
     407, 408, 409, 410,  
     411, 412, 414, 415,  
     417, 418, 419, 420,  
     421, 423, 424  
 Eastern Late Aramaic  
     413  
 Western Late Aramaic  
     413, 425  
 Maʿlulan 392  
 Mandaic 392  
 Middle Aramaic 392, 394,  
     397, 398, 399, 402,  
     403, 408, 409, 411,  
     412, 413, 414, 415,  
     417, 418, 421, 423, 424  
 Modern Aramaic 392, 397,  
     399  
 Nabatean 392, 489, 498,  
     500, 509, 515  
 Old Aramaic 392, 396, 397,  
     401, 402, 403, 404,  
     405, 406–407, 408,  
     410, 411, 412, 413,  
     414, 415, 421, 423  
 Samʿal dialect 396, 403,  
     404, 406, 408, 413  
 Official Aramaic (*See*  
     Imperial Aramaic) 392  
 Palmyrene 392  
 Samaritan Aramaic 392,  
     405, 406, 407, 408,  
     412, 413, 414, 417  
 Standard Literary Aramaic  
     392  
 Syriac 392, 393, 403, 405,  
     406–408, 412, 413,  
     414, 415, 416, 417,  
     421, 422, 431, 518,  
     940, 967, 986  
 Turoyo 392  
 Arcado-Cypriot *See* Greek  
     Dialects  
 Archaic Biblical Hebrew *See*  
     Hebrew  
 Archaic Greek *See* Greek  
 Archaic Hebrew *See* Hebrew

- Archaic Venetic *See* Venetic  
 Ardha-Māgadhī *See* Indic  
     (Middle Indic)  
 Argolic *See* Greek Dialects  
 Armenian 122, 535, 614, 636,  
     657, 670, 785, 967,  
     969, 986, 1116, 1119,  
     1120, 1122  
     Classical Armenian 2,  
         922–942  
     Eastern Armenian 923, 925  
     Middle Armenian 923  
     Modern Armenian 925,  
         926, 928, 937  
     Old Armenian 658  
     Western Armenian 923, 925  
 Aryan (*See also* Proto-Aryan)  
     717, 735, 739  
 Asōkan *See* Indic (Middle  
     Indic)  
 Assyrian *See* Akkadian  
 Attic Greek *See* Greek  
     Dialects; *see also* Greek  
 Attic-Ionic *See* Greek Dialects  
 Austro-Asiatic 698  
 Austronesian 535, 989, 1012  
 Avestan 362, 536, 539, 569,  
     636, 697, 721, 723,  
     726–727, 728, 733,  
     735, 738, 739,  
     742–763, 764, 766,  
     767–770, 772, 773,  
     1117, 1119, 1120,  
     1124  
 Gāthic Avestan (*See also* Old  
     Avestan) 547, 742,  
     743, 747, 749, 755,  
     757, 758, 1117  
     Old Avestan (*See also*  
         Gāthic Avestan) 562,  
         673, 726, 736, 742,  
         743, 747, 749, 750,  
         752, 759, 761, 762  
     Young Avestan 731, 732,  
         736, 742, 743, 747,  
         748, 749, 750, 752,  
         753, 755, 758, 761,  
         762, 1117  
 Awgni 139  
 Ayapa Gulf Sokean *See*  
     Sokean  
 Aztec 15
- Babylonian *See* Akkadian  
 Badaga 1015
- Baltic 895, 1120  
 Balto-Slavic 535, 540, 677,  
     877, 891, 1122  
 Basque 22, 858, 877  
 Bauchi 139  
 Beja 140  
 Bengali 673  
 Berber 3, 7, 138–139, 140,  
     141, 152, 160, 188  
 Biblical Hebrew *See* Hebrew  
 Bilin 139  
 Birale 140  
 Birelle 140  
 Boeotian *See* Greek Dialects  
 Bohairic *See* Coptic  
 Book Pahlavi *See* Pahlavi  
 Breton 11  
 British *See* Celtic  
 Buddhist Hybrid Sanskrit *See*  
     Indic (Middle Indic)  
 Bura 139  
 Burgundian 882  
 Burunge 139  
 Burushaski 22  
 Byblic 6  
 Byzantine Greek *See* Greek
- Canaanite (*See also* Canaanite  
     Dialects;  
     Proto-Canaanite) 6,  
     141, 151, 288, 305,  
     319, 334, 361, 365,  
     367, 369, 371, 372,  
     377, 383, 391, 404,  
     429, 435, 444, 445, 489  
     North Canaanite 361, 386  
     Old Canaanite 346  
     South Canaanite 361, 386  
 Canaanite Dialects (*See also*  
     Canaanite) 334,  
     386–390, 396, 403, 408  
     Ammonite 319, 338, 345,  
         346, 361, 365, 377, 386  
     Edomite 319, 361, 365, 377,  
         386  
     Moabite 319, 338, 345, 361,  
         365, 377, 386, 391  
 Carian 551, 606, 609–613  
 Caucasian 605, 967–968, 970,  
     1117  
     North Caucasian 989  
 Celtiberian *See* Celtic  
 Celtic (*See also* Proto-Celtic)  
     8, 11, 535, 548, 779,  
     787, 809, 842, 847,  
     850, 852, 857, 875,  
     877, 891, 920, 1122  
     British 857, 878  
     Celtiberian 8, 857  
     Cisalpine Celtic 877  
     Continental Celtic 8,  
         857–880  
     Galatian 857, 858, 877,  
         878  
     Gallo-Brittonic 857  
     Gaulish 857, 858, 860,  
         862–863, 864,  
         866–867, 868, 869,  
         870, 871, 874–877,  
         878  
     Cisalpine Gaulish 857,  
         859, 866  
     Transalpine Gaulish 857  
     Goidelic 857  
     Hispano-Celtic 8, 857,  
         858–859, 861, 862,  
         863, 864, 865–866,  
         867, 868, 869, 870,  
         871–873, 874, 877, 878  
     Insular Celtic 8, 857, 861,  
         863, 865, 868, 869,  
         870, 875, 876  
     Lepontic 850, 857, 858,  
         859–860, 862, 866,  
         868, 873–874, 877  
     Central Chadic *See* Chadic  
     Central Cushitic *See* Cushitic  
     Central Semitic *See* Semitic  
     Chadic 3, 139, 140, 160  
         Central Chadic 139  
         Eastern Chadic 139  
         West Chadic 139  
     Chaldaean *See* North Arabian  
         (Ancient North  
         Arabian)  
     Chaldisch 119  
     Chiapas Soke *See* Soke  
     Chinese (*See also*  
         Proto-Chinese) 22,  
         165, 989  
     Ancient Chinese 988–1014  
     Archaic Chinese 989, 990,  
         996, 997, 998, 999  
     Early Archaic Chinese  
         989, 997, 998, 1008  
     Late Archaic Chinese  
         (*See also* Classical  
         Chinese) 989, 991,  
         997, 998, 1008,  
         1010–1011

- Pre-Archaic Chinese  
 989, 997, 1011  
 Classical Chinese 988, 989,  
 999, 1000, 1002, 1003,  
 1004, 1005, 1006,  
 1007, 1008, 1009,  
 1010–1011  
 Contemporary Chinese  
 988, 989, 995, 996,  
 997, 998, 999, 1011  
 Dialects 989–990  
 Eastern Lu 990  
 Southwestern Chu 990  
 Mandarin Chinese  
 Early Mandarin 989  
 Old Mandarin 989  
 Medieval Chinese 989, 996,  
 997, 999  
 Early Medieval Chinese  
 989, 1008, 1010  
 Late Medieval Chinese  
 989  
 Pre-Medieval Chinese  
 989, 1010, 1011  
 Middle Chinese 992–993,  
 994, 995, 995  
 Early Middle Chinese  
 989, 993  
 Late Middle Chinese 989  
 Modern Chinese 989, 997,  
 999  
 Pre-Modern Chinese  
 989  
 Old Chinese 989, 992,  
 994–995  
 Preclassic Chinese 988  
 Chol *See* Mayan  
 Cholan *See* Mayan  
 Cholti *See* Mayan  
 Chontal *See* Mayan  
 Chorti *See* Mayan  
 Christian Latin *See* Latin  
 Christian Palestinian Aramaic  
*See* Aramaic  
 Classical Arabic *See* Arabic  
 Classical Armenian *See*  
 Armenian  
 Classical Chinese *See* Chinese  
 Classical Egyptian *See*  
 Egyptian  
 Classical Greek *See* Greek  
 Classical Hebrew *See* Hebrew  
 Classical Latin *See* Latin  
 Classical Old Georgian *See*  
 Georgian  
 Classical Sanskrit *See* Sanskrit  
 Coan *See* Greek Dialects  
 Common Afro-Asiatic (*See*  
*also* Proto-Afro-  
 Asiatic) 146, 149, 151  
 Common Anatolian (*See also*  
 Proto-Anatolian) 557,  
 558, 560, 564, 570  
 Common Phoenician *See*  
 Phoenician  
 Common Semitic (*See also*  
 Proto-Semitic) 143,  
 144, 156, 273, 345,  
 435, 437, 440, 441,  
 442, 443, 445, 448  
 Continental Celtic *See* Celtic  
 Continental Tamil *See* Tamil  
 Copainalá *See* Soke  
 Coptic 160–191 (especially),  
 222, 431  
 Akhmimic 162  
 Bohairic 162, 169  
 Fayyumic 162  
 Lycopolitan 162  
 Old Coptic 167  
 Sahidic 162, 169, 171  
 Subakhmimic 162  
 Corinthian *See* Greek Dialects  
 Cretan *See* Greek Dialects  
 Cuneiform Luvian *See* Luvian  
 Cushitic 3, 139, 141, 152, 160,  
 263, 428, 435  
 Central Cushitic 139  
 East Cushitic 139  
 Highland East Cushtic  
 139  
 Lowland East Cushtic  
 139  
 North Cushitic 139, 140  
 South Cushitic 139  
 Cyrenaean *See* Greek Dialects  
 Daco-Mysian 12, 780  
 Dadanitic *See* North Arabian  
 (Ancient North  
 Arabian)  
 Danish 881  
 Dedanite *See* North Arabian  
 (Ancient North  
 Arabian)  
 Delphian *See* Greek Dialects  
 Demotic (Egyptian) *See*  
 Egyptian  
 Dispersed Oasis North  
 Arabian *See* North  
 Arabian (Ancient  
 North Arabian)  
 Doric *See* Greek Dialects  
 Dravidian (*See also*  
 Proto-Dravidian) 4,  
 29, 61, 70, 535, 673,  
 698, 715, 1015, 1022,  
 1024, 1028, 1037  
 South Dravidian 1034,  
 1035  
 Dumaitic *See* North Arabian  
 (Ancient North  
 Arabian)  
 Dutch 881  
 Earlier Egyptian *See* Egyptian  
 Early Iranian *See* Iranian  
 Early Latin *See* Latin  
 Early Middle Indic *See* Indic  
 (Middle Indic)  
 Early Old Georgian *See*  
 Georgian  
 Early Punic *See* Punic  
 East Arabian Aramaic *See*  
 Aramaic  
 East Cushitic *See* Cushitic  
 East Germanic *See* Germanic  
 East Greek *See* Greek Dialects  
 East Italic 8  
 East Semitic *See* Semitic  
 Eastern Chadic *See* Chadic  
 Eastern Iranian *See* Iranian  
 Eastern Lu *See* Chinese  
 (Dialects)  
 Eastern Mayan *See* Mayan  
 Eastern Middle Indic *See*  
 Indic (Middle Indic)  
 Eblaite 15, 20, 26, 51, 65, 141,  
 155, 218, 220, 223,  
 229, 231–232,  
 235–238, 241, 242,  
 243, 245, 246, 247,  
 249, 250, 251, 252,  
 253, 260, 263–264,  
 266, 267, 270, 271,  
 272, 273, 275, 277, 278  
 Edomite *See* Canaanite  
 Dialects  
 Egyptian 2, 3, 6, 7, 13, 25, 54,  
 138, 140, 141, 146,  
 152, 160–191, 230,  
 277, 289, 291, 300,  
 362, 367, 383, 424,  
 425, 609, 611, 740  
 Classical Egyptian 161

- Egyptian (*cont.*)  
   Coptic *See* Coptic  
   Demotic 162, 169, 188, 395  
   Dialects 162  
   Earlier Egyptian 161, 169, 172, 173, 175, 177, 179, 183, 186  
   Late Egyptian 161, 186  
   Later Egyptian 161–162, 179, 181  
   Middle Egyptian 161, 162, 170, 171  
   Old Egyptian 161, 171, 246  
   Traditional Egyptian 161  
 Elamite 3, 15, 16, 60–94, 223, 277, 717, 721, 722, 724, 735–736, 740  
   Achaemenid Elamite 63, 64–65, 66, 67–69, 70–72, 74–78, 79–84, 86–89, 90–91  
   Dialects 64  
   Later Elamite 64  
   Middle Elamite 62, 64–65, 67, 70, 71, 74, 75, 76–77, 78–79, 80–83, 84, 86–89, 90–91  
   Neo-Elamite 62–63, 64, 67, 69, 70, 71, 74, 76, 77, 78, 80, 81, 83, 84, 86–87, 88–89  
   Old Elamite 61, 67, 70, 71  
   Pre-Achaemenid Elamite 83, 84, 88, 89, 90, 91  
   Proto-Elamite 2, 3  
 Elean *See* Greek Dialects  
 Elymian 9  
 Emesal *See* Sumerian Dialects  
 English 48, 50, 254, 548, 558, 563, 568, 573, 644, 677, 758, 766, 890, 891, 897, 939, 973, 997, 1037, 1124  
   Middle English 557, 882  
   Modern English 881  
   Old English 2, 11, 12, 881, 882, 895, 902, 914, 1118, 1125  
 Epic Sanskrit *See* Sanskrit  
 Epigraphic South Arabian *See* South Arabian  
 Epigraphic Tamil *See* Tamil  
 Epi-Olmec 3, 15, 1044, 1053, 1071–1108  
 Eretrian (Greek) *See* Greek Dialects  
 Eteo-Cretan 5  
 Eteo-Cypriot 6  
 Ethiopian Semitic *See* Semitic  
 Ethiopic (*See also* Ge'ez; Proto-Ethiopic; Semitic [Ethio-Semitic]) 142, 143, 144, 147, 151, 154, 155, 253, 260, 263, 324, 445, 485, 488, 494  
 Ethio-Semitic *See* Semitic  
 Etruscan (*See also* Proto-Etruscan) 8, 9, 22, 791, 809, 816, 817, 837, 844, 855, 859, 867, 877, 943  
   Archaic Etruscan 944, 947, 949, 950, 952, 953, 954, 955, 957, 958, 962, 963  
   Late Etruscan 944, 947, 948, 949, 950, 952, 953, 954, 955, 957, 958, 962  
 Euboean *See* Greek Dialects  
 European languages 1119  
 Faliscan 789, 791, 798, 801, 809, 813  
 Faroese 881  
 Fayyumic *See* Coptic  
 Finno-Ugric 535  
 Flemish 883  
 French 864, 973, 1113  
   Norman French 939  
 Galatian *See* Celtic  
 Galilean Aramaic *See* Aramaic  
 Galla 139  
 Gallo-Brittonic *See* Celtic  
 Gāndhārī Prakrit *See* Indic (Middle Indic)  
 Gāthic Avestan *See* Avestan  
 Gaulish *See* Celtic  
 Gaunche 139  
 Ge'ez 2, 142, 427–453  
 Georgian (*See also* Proto-Georgian) 22, 922  
   Classical Old Georgian 968, 969, 975  
   Early Georgian 967–987  
   Early Old Georgian 2, 968  
   Haemet'i 968  
   Middle Georgian 968  
   Modern Georgian 968, 978  
   Old Georgian 923, 968, 969, 979, 985  
   Xanmet'i 968, 984  
 German 548, 743, 881, 884, 886, 897, 919, 1124  
   Middle High German 882  
   Old High German 560, 850, 881, 882, 891, 895, 903, 914, 1125  
 Germanic (*See also* Proto-Germanic) 11, 535, 540, 563, 648, 809, 877, 881, 882, 886, 887, 888, 889, 889–890, 892, 893, 895, 897, 901, 902, 903, 905, 907, 908, 910, 920, 928, 1118, 1120, 1122, 1125  
   East Germanic 881, 882, 883, 907, 908  
   North Germanic 881, 891, 907–908  
   Northwest Germanic 882, 907, 908, 914  
   Pre-Germanic 888, 889, 897, 903  
   West Germanic 881, 891, 907, 908, 913, 914  
 Goidelic *See* Celtic  
 Gothic 2, 657, 850, 881, 906, 907, 908, 910, 913, 919, 1125  
 Greater Tzeltalan *See* Mayan  
 Greek (*See also* Greek Dialects; Proto-Greek) 5, 6, 12, 13, 14, 23, 122, 156, 162, 163, 167, 169, 171, 172, 188, 230, 233, 277, 320, 321, 325, 330, 362, 366, 367, 369, 371, 383, 393, 395, 402, 416, 417, 422, 424, 425, 430, 431, 435, 449, 450, 485, 488, 489, 495, 497, 500, 535, 536, 538, 539, 540, 542, 543, 546, 547, 548, 556, 557, 558, 559, 564,



- 565, 567, 569, 573,  
 588, 591, 592, 593,  
 594, 599, 601, 602,  
 603, 604, 609, 610,  
 611, 614–649, 676,  
 684, 686, 697, 700,  
 715, 735, 740, 748,  
 750, 752, 760, 766,  
 779, 780, 781, 782,  
 784, 785, 786, 787,  
 791, 794, 798, 800,  
 804, 809, 810, 816,  
 817, 820, 837, 844,  
 850, 851, 854, 862,  
 864, 867, 877, 882,  
 884, 885, 886, 895,  
 897, 901, 904–905,  
 910, 911, 916, 922,  
 923, 924, 925, 929,  
 933, 938, 940, 943,  
 944, 945, 946,  
 947–948, 949, 950,  
 959, 964, 967, 969,  
 972, 986, 1037, 1119,  
 1120, 1121, 1122,  
 1124, 1125  
 Archaic Greek 614  
 Byzantine Greek 614, 619,  
 659  
 Classical Greek 614  
 Hellenistic Greek 188,  
 624  
 Homeric Greek 539, 547,  
 561, 621, 623, 631,  
 636, 638, 639, 647,  
 661, 664, 666, 670,  
 1120, 1124  
 Koine Greek 495, 614, 657,  
 659  
 Modern Greek 594, 614,  
 657, 659  
 New Testament Greek (*See*  
*also* Koine Greek) 904  
 Greek Dialects (*See also*  
 Greek) 650–672  
 Aeolic 650, 651, 652, 659,  
 660, 662, 663, 664,  
 666, 667, 668, 671, 780  
 Anatolian Aeolic 651,  
 652  
 Balkan Aeolic 651, 652,  
 667  
 Arcado-Cypriot 650, 652,  
 660, 661, 662, 663,  
 666, 667, 668, 669, 670  
 Arcadian 658, 660, 661,  
 662, 663, 664, 666,  
 668, 669, 670  
 Cypriot 658, 660, 664,  
 666, 669, 670  
 Argolic 651, 658, 661, 662,  
 663, 664  
 Attic (*See also* Greek) 13,  
 14, 536, 544, 614–649,  
 650, 651, 657, 658,  
 659, 660, 661, 662,  
 663, 664, 665, 666,  
 667, 668, 669, 779  
 Early Attic 624, 626, 664  
 Attic-Ionic (*See also*  
 Proto-Attic-Ionic)  
 539, 623, 637, 650,  
 652, 662, 663  
 Boeotian 651, 652, 658,  
 659, 660, 661, 666,  
 667, 668, 669  
 Coan 661, 670  
 Corinthian 651  
 Cretan 658, 661, 662, 663,  
 664, 666, 667, 668,  
 669, 670  
 Cypriot *See*  
 Arcado-Cypriot  
 Cyrenaean 662  
 Delphian 666  
 Doric 619, 650, 651, 652,  
 655, 657, 658, 661,  
 662, 663, 664, 665,  
 666, 667, 668, 669,  
 670, 780, 837  
 Doric Koine 14  
 East Greek 652  
 Elean 651, 658, 661, 662,  
 663, 664, 668  
 Eretrian 662  
 Euboean (*See also* West  
 Ionic) 791  
 Ionic (*See also* Attic-Ionic)  
 616, 635, 636, 650,  
 651, 652, 661, 662,  
 663, 664, 665, 666,  
 667, 668, 669  
 Anatolian Ionic *See* West  
 Ionic  
 Central Ionic 621, 650,  
 661  
 East Ionic 650, 651, 658,  
 661  
 West Ionic 621, 650, 662,  
 668  
 Laconian 614, 651, 657,  
 658, 661, 662, 663,  
 668, 670  
 Lesbian 651, 652, 658, 660,  
 661, 662, 663, 664,  
 665, 666, 667, 668,  
 669, 670  
 Locrian 670  
 East Locrian 651  
 West Locrian 651  
 Megarian 651  
 Messanian 651  
 Mycenaean 536, 547, 560,  
 616, 620, 652, 657,  
 658, 660, 661, 663,  
 664, 666, 668, 670,  
 671, 780  
 Mycenaean I 651, 652,  
 661  
 Mycenaean II 651, 652,  
 661  
 Normal Mycenaean 651  
 Special Mycenaean 651  
 North Greek 652  
 Northwest Greek 14, 650,  
 651, 652, 662, 664,  
 665, 666, 667, 668,  
 669, 670  
 Northwest Greek Koine 14,  
 651, 666  
 Pamphylian 651, 658, 664,  
 668  
 Pelasgiotis 651  
 Phocian 651, 670  
 Rhodian 651, 661, 668,  
 669  
 Sicilian 665, 668  
 South Greek 652  
 Thera 661, 662  
 Thessalian 651, 652, 660,  
 661, 662, 663, 664,  
 665, 666, 667, 669,  
 670  
 Thessaliotis 651  
 Tsaconian 614  
 West Greek 651, 652  
 Gujarati 673  
 Gulf Sokean *See* Sokean  
 Hadramitic *See* South  
 Arabian (Ancient  
 South Arabian)  
 Haemet'i *See* Georgian  
 Hamitic 513  
 Hamito-Semitic 138, 160

- Haramic *See* South Arabian  
(Ancient South Arabian)
- Hasaitic *See* North Arabian  
(Ancient North Arabian)
- Hatran *See* Aramaic
- Hattic 551, 552, 572, 573, 586, 590
- Hausa 139
- Hebrew (*See also* Proto-Hebrew) 60, 80, 141, 143, 144, 145, 147, 148, 151, 246, 253, 260, 270, 277, 292, 293, 294, 295, 296, 298, 299, 300, 301, 308, 309, 315, 319–364, 365, 366, 368, 369, 370, 372, 377, 381, 386, 387, 388, 389, 391, 393, 397, 408, 416, 422, 423, 424, 425, 429, 432, 434, 436, 440, 445, 447, 449, 457, 459, 495, 497, 498, 499, 500, 501, 517, 519, 740
- Archaic Biblical Hebrew 319
- Archaic Hebrew 319, 345
- Biblical Hebrew 288, 291, 293, 301, 303–304, 307, 309, 312, 313, 314, 319, 320, 325, 326, 336, 337, 338, 341, 342, 343, 345, 347, 348, 349, 361, 362, 372, 375, 376, 377, 378, 380, 381, 382
- Classical Hebrew 319
- Israeli Hebrew 320
- Israelite Hebrew 334, 345, 362
- Judahite Hebrew 334, 345, 377
- Late Biblical Hebrew 319, 320, 324, 337, 339, 342, 345, 348, 350, 352, 362
- Late Classical Hebrew 319
- Medieval Hebrew 320
- Middle Hebrew 320
- Modern Hebrew 320
- Northern Hebrew 334, 339
- Pre-Hebrew 341
- Rabbinic Hebrew 320, 324, 325, 330, 336, 337, 338, 339, 341, 342, 343, 345, 346, 347, 348, 349, 350, 352, 353, 354, 355, 362
- Southern Hebrew 334
- Hellenistic Greek *See* Greek
- Helleno-Armenian 922
- Hernican *See* Sabellian
- Herulian 882
- Hieroglyphic Luvian *See* Luvian
- Hindi 673
- Hismaic *See* North Arabian  
(Ancient North Arabian)
- Hispano-Celtic *See* Celtic
- Hittite 15, 54, 65, 66, 95, 97, 117, 156, 163, 277, 289, 315, 544, 547, 548, 551–575, 576, 577, 578, 579, 581, 582, 583, 585, 586, 587, 588, 589, 590, 591, 597, 598, 599, 601, 606, 607, 648, 766, 787, 809, 901, 1118, 1120, 1121, 1124, 1125
- Middle Hittite 551, 554, 571, 573, 574
- Neo-Hittite 551, 554, 560, 569, 571
- Old Hittite 551, 554, 560, 562, 563, 571, 576, 585, 873
- Homeric Greek *See* Greek
- Huastecan *See* Mayan
- Hungarian 22
- Hurrian (*See also* Proto-Urarto-Hurrian) 15, 61, 65, 80, 95–118, 119, 122, 124–125, 126–127, 128–129, 130, 131–132, 133, 134, 135–136, 277, 289, 291, 315, 552, 572, 574, 577
- Dialects 97
- Mitanni Letter Hurrian 97, 98, 99, 107, 108, 110, 111, 112, 113, 114, 116
- Old Hurrian 95, 97, 99, 103, 110, 111, 112, 113, 116, 119
- Iberian 877
- Icelandic 881, 1119
- Illyrian 11–12, 15
- Imperial Aramaic *See* Aramaic
- Indic (*See also* Indo-Aryan) 16, 677, 694, 742, 759, 1117, 1118
- Middle Indic (*See also* Indic [Prākṛit]) 673, 674, 678, 698, 700–716
- Apabhraṃśa 702
- Ardha-Māgadhī 702, 707, 711
- Aśokan 703, 707, 710, 711
- Buddhist Hybrid Sanskrit 702, 715
- Early Middle Indic 700
- Eastern Middle Indic 704
- Gāndhārī Prākṛit 701, 702
- Jain-Mahārāṣṭrī 702
- Jain-Śaurasenī 702
- Late Prākṛit 702
- Māgadhī 702
- Māhārāṣṭrī 702, 704, 707
- Northwest Prākṛit 715
- Pāli 701, 703, 704, 707, 708, 709, 711, 714, 715
- Prākṛits “proper” 701, 711
- Śaurasenī 702, 704
- Modern Indic 673
- Old Indic (*See also* Indo-Aryan [Old Indo-Aryan]; Sanskrit) 673, 674, 700, 703, 705, 710, 712, 715
- Prākṛit (= Middle Indic; *see* Indic [Middle Indic]) 673, 674, 700–716, 1037
- Sanskritized Prākṛit 702
- Indo-Aryan (*See also* Indic) 4, 96, 117, 673, 674, 688, 692, 700, 714, 715, 1015, 1037
- Old Indo-Aryan 673, 725, 728, 735, 739, 740

- Indo-European (*See also*  
 Proto-Indo-European) 2, 4, 8, 9,  
 10, 11, 12, 14, 15, 22,  
 54, 104, 140, 148, 156,  
 174, 188, 343, 377,  
 402, 424, 534–550,  
 551, 556, 557, 558,  
 559, 560, 561, 562,  
 563, 564, 565, 567,  
 568, 570, 572, 573,  
 574, 576, 580, 585,  
 588, 589, 596, 601,  
 604, 605, 614, 633,  
 642, 643, 645, 647,  
 648, 673–694, 679,  
 680, 692, 694, 695,  
 697, 700, 717, 727,  
 728, 732, 738, 742,  
 748, 749, 753, 755,  
 758, 759, 760, 761,  
 770, 771, 780, 784,  
 785, 786, 787, 789,  
 796, 798, 799, 800,  
 801, 809, 813, 836,  
 842, 848, 852, 854,  
 861, 865, 866,  
 868–869, 877, 890,  
 891, 892, 893, 895,  
 896, 901, 902, 904,  
 905, 907, 908, 911,  
 912, 913, 916, 917,  
 919, 920, 922, 927,  
 928, 929, 931, 937,  
 938, 939, 940, 951,  
 959, 967, 973, 989,  
 1113, 1114, 1116,  
 1118, 1119, 1120,  
 1121, 1122, 1123, 1125
- Indo-Iranian (*See also*  
 Proto-Indo-Iranian)  
 535, 538, 539, 565,  
 614, 636, 642, 673,  
 675, 676, 678, 679,  
 700, 717, 721, 732,  
 736, 742, 747, 749,  
 752, 759, 768, 773,  
 785, 809, 922, 1117,  
 1119, 1120, 1121, 1122
- Indonesian languages 1116
- Insular Celtic *See* Celtic
- Ionic *See* Greek Dialects
- Iranian (*See also*  
 Proto-Iranian) 61, 62,  
 64, 69, 71, 75, 90–91,  
 188, 599, 648, 676,  
 677, 697, 715, 717,  
 722, 726, 732, 736,  
 742, 749, 759, 765,  
 766, 768, 769, 770,  
 773, 922–923, 926,  
 939, 940, 1116, 1122
- Early Iranian 721
- Eastern Iranian 724, 739,  
 740, 742
- Middle Iranian 731, 764,  
 766, 770
- Western Middle Iranian  
 764
- Northwestern Iranian 717
- Old Iranian 64, 80, 87, 90,  
 717–736, 739, 740,  
 770, 771, 772, 774,  
 1012
- Southwestern Iranian 717,  
 739
- Iraqw 139
- Irish 8  
 Ogham Irish 8, 876  
 Old Irish 8, 867, 1120, 1124
- Irula 1015
- Israeli Hebrew *See* Hebrew
- Israelite Hebrew *See* Hebrew
- Italian 1113, 1115
- Italic (*See also* Proto-Italic) 8,  
 9, 15, 535, 790, 809,  
 813, 836, 837, 842,  
 950, 951, 965, 1122
- Itza *See* Mayan
- Jain-Mahārāṣṭrī *See* Indic  
 (Middle Indic)
- Jain-Śaurasenī *See* Indic  
 (Middle Indic)
- Jawfian *See* North Arabian  
 (Ancient North  
 Arabian)
- Jewish Babylonian Aramaic  
*See* Aramaic
- Jewish Palestinian Aramaic  
*See* Aramaic
- Jibbālī 141, 143, 503
- Judahite Hebrew *See* Hebrew
- Kabyle 138
- Kanjobalan *See* Mayan
- Kannada 1015
- Kartvelian (*See also*  
 Proto-Kartvelian) 922,  
 923, 967, 970, 971, 977
- Kassite 277
- Kemant 139
- Kera 139
- Khūzī 64
- Kodagu 1015
- Koine Greek *See* Greek
- Kota 1015
- Kotoko-Logone 139
- Kurdish 391
- Lacandon *See* Mayan
- Laconian *See* Greek Dialects
- Late Aramaic *See* Aramaic
- Late Egyptian *See* Egyptian
- Late Latin *See* Latin
- Late Prakrit *See* Indic (Middle  
 Indic)
- Late Punic *See* Punic
- Later Egyptian *See* Egyptian
- Later Elamite *See* Elamite
- Latin 8, 9, 10, 12, 13, 15, 25,  
 188, 220, 224, 335,  
 346, 362, 366, 367,  
 369, 371, 373, 377,  
 381, 383, 424, 425,  
 500, 535, 536, 543,  
 544, 548, 554, 558,  
 560, 563, 567, 568,  
 573, 582, 587, 588,  
 637, 668, 684, 687,  
 725, 726, 780, 781,  
 784, 785, 789–811,  
 812, 813, 816, 817,  
 818, 820, 821, 825,  
 827, 830, 831, 833,  
 837, 840, 842, 843,  
 846, 847, 848, 849,  
 851, 854, 855, 860,  
 868, 871, 877, 882,  
 884, 891, 895, 897,  
 908, 910, 911, 916,  
 939, 943, 944, 945,  
 946, 947, 949, 956,  
 962, 964–965, 977,  
 988, 1037, 1113, 1115,  
 1119, 1120, 1121,  
 1124, 1125
- Christian Latin 790
- Classical Latin 790, 791,  
 792, 793, 794, 796,  
 802, 804, 806, 807,  
 809, 1113
- Early Latin 789, 790, 792,  
 793, 794, 798, 799,  
 801, 802, 803

- Latin (*cont.*)  
   Late Latin 790, 803, 804  
   Old Latin (*See* Early Latin)  
     568, 569, 870, 1120  
   Post-Classical Latin 790,  
     794  
   Praenestine Latin 790, 830  
   Pre-Classical Latin 790  
   Preliterary Latin 790  
   Roman Latin 790  
   Silver Latin 790  
   Vulgar Latin 790  
 Latino-Faliscan 789, 836  
 Latino-Punic *See* Punic  
 Latino-Venetic *See* Venetic  
 Laz 967  
 Lemnian (*See also*  
   Proto-Lemnian) 9–10,  
   944, 949, 964  
 Lepontic *See* Celtic  
 Lesbian *See* Greek Dialects  
 Libyan 139, 160, 188  
 Libyco-Berber 139  
 Ligurian 10–11  
 Lihyanite *See* North Arabian  
   (Ancient North  
   Arabian)  
 Lithuanian 564, 895, 1119,  
   1124  
 Locrian *See* Greek Dialects  
 Lowland Mije *See* Mije  
 Luvian (Luwian) (*See also*  
   Proto-Luvian) 15,  
   117, 163, 381, 383,  
   552, 553, 558, 559,  
   560, 569–570, 574,  
   576–584, 588, 589,  
   590, 591, 592, 597,  
   598, 601, 606, 607,  
   610, 612, 787, 866,  
   1124  
   Cuneiform Luvian (Also  
   CLuvian) 551, 552,  
   576, 577, 579, 581,  
   583, 586, 1116  
   Hieroglyphic Luvian (Also  
   HLuvian) 551, 574,  
   576, 577–579, 580,  
   581, 582, 583, 591  
 Lycian 551, 558, 559, 579, 588,  
   591–600, 601, 603,  
   605, 606, 610, 611, 612  
   Lycian A 592, 593, 594, 596  
   Lycian B 592  
   Milyan 592, 594, 596  
 Lycopolitan *See* Coptic  
 Lydian 551, 558, 591, 592,  
   597, 221, 610, 611, 946  
 Macedonian 12–14  
 Madhābic *See* South Arabian  
   (Ancient South  
   Arabian)  
 Madurese 1116  
 Māgadhi *See* Indic (Middle  
   Indic)  
 Magdalena *See* Soke  
 Māhārāṣṭrī *See* Indic (Middle  
   Indic)  
 Mahrian *See* Semitic  
 Malayalam 1015  
 Maʿlulan *See* Aramaic  
 Mandaic *See* Aramaic  
 Mandarin *See* Chinese  
 Manichean Middle Persian  
   *See* Persian (Middle  
   Persian)  
 Margi 139  
 Marrucinian *See* Sabellian  
 Marsian *See* Sabellian  
 Masa 139  
 Mayan (*See also*  
   Proto-Mayan) 15,  
   1041–1070, 1075,  
   1099  
   Chol 1041  
   Cholan (*See also*  
     Proto-Cholan) 1041,  
     1056, 1059, 1060,  
     1062, 1063, 1065  
     Eastern Cholan 1062  
   Cholti 1041, 1062  
   Chontal 1041  
   Chorti 1041, 1062  
   Eastern Mayan 1058  
   Greater Tzeltalan 1041,  
     1055, 1056, 1057,  
     1058, 1059, 1060  
   Huastecan 1041  
   Itza 1041  
   Kanjobalan 1058  
   Lacandon 1041, 1062  
   Mopan 1041  
   Tzeltal 1041  
   Tzeltalan 1041, 1056  
   Tzotzil 1041  
   Yucatec 1041  
   Yucatecan (*See also*  
     Proto-Yucatecan)  
     1041, 1055, 1057,  
     1058, 1059, 1060,  
     1062, 1063, 1065  
   Median 60, 597, 739, 740  
   Medieval Hebrew *See* Hebrew  
   Medieval Tamil *See* Tamil  
   Mediterranean languages 809,  
     877  
   Megarian *See* Greek Dialects  
   Mehri 141, 503  
   Meroitic 7, 163  
   Mesopotamian languages 19  
   Messanian *See* Greek Dialects  
   Messapic 14–15, 535  
   Miao-Yao 1012  
   Middle Arabic *See* Arabic  
   Middle Aramaic *See* Aramaic  
   Middle Chinese *See* Chinese  
   Middle Egyptian *See* Egyptian  
   Middle Elamite *See* Elamite  
   Middle English *See* English  
   Middle Georgian *See*  
     Georgian  
   Middle Hebrew *See* Hebrew  
   Middle High German *See*  
     German  
   Middle Hittite *See* Hittite  
   Middle Indic *See* Indic  
   Middle Iranian *See* Iranian  
   Middle Persian *See* Persian  
   Middle Phrygian *See*  
     Phrygian  
   Migama 139  
   Mije (*See also* Mijejan) 1071  
     Lowland Mije 1082  
     Totontepec Mije 1096  
   Mijejan (*See also* Mije;  
     Proto-Mijejan) 1071,  
     1080, 1082, 1083,  
     1086, 1088, 1097,  
     1099, 1105  
     Oluta Mijejan 1082, 1086,  
     1098  
     Sayula Mijejan 1082, 1086,  
     1087  
   Mije-Sokean 1071, 1072,  
     1074–1075, 1076,  
     1077, 1078, 1079,  
     1080–1081, 1082,  
     1083, 1084, 1087,  
     1088, 1089, 1090,  
     1091, 1093,  
     1094–1095, 1096,  
     1097, 1098–1099,  
     1104, 1105–1106  
   Milyan *See* Lycian

- Minaic *See* South Arabian  
     (Ancient South Arabian)  
 Mingrelian 967  
 Mitanni 96  
 Mitanni Letter Hurrian *See* Hurrian  
 Mixe-Zoquean 1068  
 Mixtec 15  
 Moabite *See* Canaanite  
     Dialects  
 Modern Aramaic *See* Aramaic  
 Modern Chinese *See* Chinese  
 Modern Georgian *See* Georgian  
 Modern German *See* German  
 Modern Greek *See* Greek  
 Modern Hebrew *See* Hebrew  
 Modern Indic *See* Indic  
 Modern Persian *See* Persian  
 Modern Russian *See* Russian  
 Modern South Arabian *See* South Arabian  
 Modern Standard Arabic *See* Arabic  
 Modern Tamil *See* Tamil  
 Mopan *See* Mayan  
 Mubi 139  
 Muṇḍa languages 698  
 Mycenaean *See* Greek Dialects  
  
 Nabatean *See* Aramaic  
 Nahuatl 1068  
 Nawa 1071  
 Neo-Assyrian *See* Akkadian  
     (Assyrian)  
 Neo-Elamite *See* Elamite  
 Neo-Hittite *See* Hittite  
 Neo-Phrygian *See* Phrygian  
 Neopunic *See* Punic  
 New Testament Greek *See* Greek  
 Norman French *See* French  
 North Afro-Asiatic *See* Afro-Asiatic  
 North American languages 41, 1117  
 North Arabian 454, 455, 485, 488, 495, 517  
     Ancient North Arabian 488–533  
     Chaldaeian 490  
     Dadanitic 489, 490, 493, 494, 495–497, 498, 499, 500, 501, 502, 503, 504, 505, 506, 507, 508, 509–519, 520–522, 523–524, 525–526, 527, 529, 530  
     Dedanite 490  
     Dialects 490  
     Dispersed Oasis North Arabian 490  
     Dumaitic 490, 495, 500, 501, 519  
     Hasaitic 490, 492, 494, 495, 500, 501, 506, 507, 508, 517, 519, 521, 523, 529  
     Hismaic 490, 492, 493, 494–495, 497, 498, 500, 501, 504, 505, 506, 507, 508, 509, 517, 518–519, 520, 523, 527, 528, 529, 530  
     Jawfian 490  
     Lihyanite 490, 529, 530  
     Oasis North Arabian 490–492  
     Safaitic 490, 492, 493, 494–495, 497, 498, 499, 500–501, 502, 503, 504–505, 506, 507, 508, 509, 510–511, 513–514, 516, 517, 518–520, 521, 522–523, 524–525, 527, 528, 529–530  
     South Safaitic 492  
     Southern Thamudic 490, 492  
     Tabuki Thamudic 492  
     Taymanitic 490, 492, 493, 494, 495, 499, 500, 501, 502, 506, 507, 508, 517, 518–519, 520, 529, 530  
     Thamudic 492, 493, 508, 530  
     Thamudic A 492  
     Thamudic B 490, 492, 494, 495, 500, 501, 503, 506, 507, 508, 510, 516, 517, 518–519, 520, 527  
     Thamudic C 490, 494, 495, 500, 506, 507, 508, 517, 518–519  
     Thamudic D 490, 492, 494, 495, 500, 506, 507, 508, 517, 518–519  
     Thamudic E 492  
 North Canaanite *See* Canaanite  
 North Cushitic *See* Cushitic  
 North Germanic *See* Germanic  
 North Greek *See* Greek  
     Dialects  
 North Picene 8  
 Northeast Caucasian 95  
 Northern Hebrew *See* Hebrew  
 Northern Semitic *See* Semitic  
 Northwest Germanic *See* Germanic  
 Northwest Greek *See* Greek  
     Dialects  
 Northwest Prakrit *See* Indic  
     (Middle Indic)  
 Northwest Semitic *See* Semitic  
 Northwestern Iranian *See* Iranian  
 Norwegian 881  
 Nostratic 22, 140, 967, 1126  
 Numidian 139, 367, 383  
  
 Oasis North Arabian *See* North Arabian  
     (Ancient North Arabian)  
 Oaxaca Soke *See* Soke  
 Official Aramaic *See* Aramaic  
 Ogham Irish *See* Irish  
 Old Akkadian *See* Akkadian  
 Old Arabic *See* Arabic  
 Old Aramaic *See* Aramaic  
 Old Armenian *See* Armenian  
 Old Assyrian *See* Akkadian  
     (Assyrian)  
 Old Avestan *See* Avestan  
 Old Babylonian *See* Akkadian  
     (Babylonian)  
 Old Bulgarian 657  
 Old Byblian *See* Phoenician  
 Old Canaanite *See* Canaanite  
 Old Chinese *See* Chinese  
 Old Church Slav (Slavonic) 892, 895, 1120, 1125  
 Old Coptic *See* Coptic  
 Old Egyptian *See* Egyptian  
 Old Elamite *See* Elamite  
 Old English *See* English

- Old Frisian 882  
 Old Georgian *See* Georgian  
 Old High German *See* German  
 Old Hispanic 8  
 Old Hittite *See* Hittite  
 Old Hurrian *See* Hurrian  
 Old Icelandic 895  
 Old Indic *See* Indic  
 Old Indo-Aryan *See* Indo-Aryan  
 Old Iranian *See* Iranian  
 Old Irish *See* Irish  
 Old Kannada 2  
 Old Khmer 2  
 Old Latin *See* Latin  
 Old Low Franconian 881  
 Old Norse 12, 882, 910, 911, 919  
 Old Persian *See* Persian  
 Old Prussian 890  
 Old Sabellian 8  
 Old Saxon 881, 895  
 Old South Arabian *See* South Arabian  
 Old Sumerian *See* Sumerian  
 Old Tamil *See* Tamil  
 Old Telugu 2  
 Old Welsh *See* Welsh  
 Oluta Mijeane *See* Mijeane  
 Omotic languages 139, 160  
 Ongota 140  
 Oromo 139  
 Oscan *See* Sabellian  
 Oto-Manguean 1075  
  
 Paelignian *See* Sabellian  
 Pahlavi 16, 743, 744, 764–776  
     Book Pahlavi 764, 771, 772  
 Palaic 551, 552, 574, 577, 579, 585–590, 591, 592, 597, 598, 601, 606  
 Paleo-Phrygian *See* Phrygian  
 Pali *See* Indic (Middle Indic)  
 Palmyrene *See* Aramaic  
 Pamphylian *See* Greek Dialects  
 Parthian 726, 766, 775, 940  
 Pashto 22, 725  
 Pazend 766  
 Pelasgian 648  
 Pelasgiotis *See* Greek Dialects  
 Persian (*See also* Proto-Persian) 13, 90, 91, 277, 424, 594, 986, 1037  
     Middle Persian 718, 721, 724, 725, 735, 738, 740, 743  
     Manichean Middle Persian 766, 775  
     Modern Persian 721, 738, 740  
     Old Persian 15, 60, 63, 65, 74, 75–76, 79, 80–81, 83, 84, 86, 87–88, 90, 223, 362, 717–741, 752, 768–770, 923  
 Phocian *See* Greek Dialects  
 Phoenician (*See also* Punic) 5, 6, 141, 290, 293, 300, 308, 315, 319, 321, 332, 338, 345, 346, 361, 365–385, 386, 387, 388, 389, 391, 578, 653, 656, 671, 943, 946  
     Common Phoenician 366  
     Cypriot Phoenician 376  
     Old Byblian 366, 367, 374, 377, 382  
     Standard Phoenician 366, 367, 370, 374, 376  
 Phrygian 12, 535, 636, 777–788, 922, 1122  
     Middle Phrygian 778, 779, 783, 784, 786, 787  
     Neo-Phrygian 778, 779–780, 781, 782–783, 784, 785, 786, 787  
     Paleo-Phrygian 777, 778–779, 780, 781, 782, 783, 784, 785, 786  
 Pictish 9  
 Pisidian 551, 606  
 Portuguese 1037  
 Prakrit *See* Indic  
 Pre-Achaemenid Elamite *See* Elamite  
 Pre-Classical Latin *See* Latin  
 Pre-Germanic *See* Germanic  
 Pre-Hebrew *See* Hebrew  
 Preliterary Latin *See* Latin  
 Pre-Samnite *See* Sabellian  
 Proto-Afro-Asiatic (*See also* Common Afro-Asiatic) 3, 138, 140, 141, 143, 144, 1116  
 Proto-Algonquian 1117, 1124  
 Proto-Anatolian (*See also* Common Anatolian) 557, 558, 559, 591, 596, 601, 611  
 Proto-Arabic 488  
 Proto-Aramaic 392, 395, 397–398, 400, 401, 402, 410, 411, 417  
 Proto-Aryan (*See also* Proto-Indo-Iranian) 723, 726, 731, 732, 733, 735, 737, 740  
 Proto-Attic-Ionic 635  
 Proto-Canaanite 326, 379, 386, 387–388  
 Proto-Celtic 857, 865, 866, 1114  
 Proto-Central Semitic 369, 372  
 Proto-Chinese 997  
 Proto-Cholan 1056, 1058, 1059, 1060  
 Proto-Dravidian 61, 90  
 Proto-Elamite *See* Elamite  
 Proto-Elamo-Dravidian 61  
 Proto-Ethiopic 436  
 Proto-Ethio-Semitic 428, 437  
 Proto-Etruscan 944  
 Proto-Euphratic 54  
 Proto-Georgian 967  
 Proto-Germanic 886, 889, 890–891, 892, 895, 897, 901, 910, 911, 916, 1114  
 Proto-Greek 620, 623, 625, 634, 657, 658, 659, 660, 665, 667  
 Proto-Hebrew 301, 332, 334  
 Proto-Indo-European 2, 11, 14, 535, 548, 536, 537, 540, 541, 546–547, 557, 565, 568, 573, 579, 580, 581, 583, 587, 588, 589, 594, 596–597, 606, 612, 614, 619, 620–621, 622, 623, 624, 625, 626–628, 631, 632, 633, 634–635, 636–637, 638–639, 640, 643, 647, 652, 660, 661, 662, 663,

- 675–676, 679, 680,  
681, 688, 697, 726,  
727–728, 730, 733,  
735, 737, 739, 740,  
747, 748–749, 750,  
751, 753, 756, 758,  
773, 774, 777, 782,  
783, 784, 785, 792,  
814, 836, 842, 847,  
850, 854, 858, 866,  
869, 871, 873, 887,  
888, 889–890, 892,  
893, 895, 896, 897,  
901–902, 903, 919,  
922, 928, 939,  
1113–1114, 1116,  
1117, 1118,  
1119–1120, 1121,  
1122, 1123,  
1124–1125  
Proto-Indo-Iranian (*See also*  
Proto-Aryan) 747,  
748–749, 750, 752,  
767–770, 772, 774,  
1117, 1117  
Proto-Iranian 720, 722, 724,  
725, 726–727, 728,  
732, 739, 750,  
767–770, 771  
Proto-Italic 948  
Proto-Kartvelian 967  
Proto-Lemnian 944  
Proto-Luvian 591  
Proto-Mayan 1058  
Proto-Mijean 1071, 1078,  
1080, 1090, 1099, 1104  
Proto-Mije-Sokean 1078,  
1080, 1081, 1082,  
1083, 1085,  
1086–1087, 1088,  
1090, 1092,  
1094–1095, 1097,  
1098, 1099, 1104,  
1105, 1106  
Proto-Northwest Semitic  
333–335, 337, 338,  
345, 346, 347, 348,  
352, 369, 373  
Proto-Persian 726  
Proto-Rhaetic 944  
Proto-Sabellian 821  
Proto-Semitic (*See also*  
Common Semitic) 2,  
140, 142–156, 227,  
229, 230–232, 233,  
235–236, 237, 238,  
240, 241, 242, 244,  
252, 253, 276, 277,  
288, 293, 321, 324,  
326, 327, 329, 331,  
333–335, 337, 338,  
339, 340, 342, 343,  
345, 346, 350, 369,  
370, 371, 373, 376,  
386, 398, 429, 435,  
436, 440, 441, 444,  
445, 449, 457, 499,  
1117, 1125  
Proto-Sinaitic 7, 163,  
387–388  
Proto-Sino-Tibetan 997  
Proto-Soke 1090  
Proto-Sokean 1071, 1076,  
1078, 1079, 1080,  
1081, 1082, 1083,  
1086, 1087, 1094,  
1095, 1096, 1097,  
1098, 1099, 1104, 1105  
Proto-Tigridian 54  
Proto-Tyrsenic 944, 949, 964  
Proto-Ugaritic 294  
Proto-Urartian 124, 136  
Proto-Urarto-Hurrian 95, 97,  
127  
Proto-Yucatecan 1056  
Proto-Zapotec 1110, 1111  
Proto-Zapotecan 1110, 1111  
Punic 139, 293, 345, 366, 367,  
368, 370, 371, 373,  
375, 376, 377, 381,  
386  
Early Punic 375  
Late Punic 366, 368, 373,  
375, 380  
Latino-Punic 366, 373  
Neopunic 293, 366  
Qatabanian *See* South  
Arabian (Ancient  
South Arabian)  
Rabbinic Hebrew *See* Hebrew  
Recent Venetic *See* Venetic  
Rhaetic (*See also*  
Proto-Rhaetic) 9–10,  
877, 944, 949, 964  
Rhodian *See* Greek Dialects  
Rift Valley Cushitic (Highland  
East Cushitic) *See*  
Cushitic  
Roman Latin *See* Latin  
Romance languages 427, 793,  
794, 803, 804, 858  
Ron languages 139  
Runic Norse (*See also* Ancient  
Nordic) 892  
Russian 1119  
Modern Russian 920  
Sabaic *See* South Arabian  
(Ancient South  
Arabian)  
Sabeian = Sabaic  
Sabellian (*See also*  
Proto-Sabellian) 8, 15,  
789, 809, 812–839,  
855, 946  
Aequian 812  
Hernican 812  
Marrucinian 812, 828, 837,  
855  
Marsian 812, 837  
Oscan 789, 796, 812, 813,  
814, 816, 817,  
818–820, 821, 822,  
825–826, 828–831,  
832–834, 835, 836,  
837, 842, 854  
Paelignian 812, 816, 826,  
828, 829, 830, 837,  
855  
Pre-Samnite 812  
Sabine 809, 812  
South Picene 8, 812, 813,  
814, 817  
Umbrian 789, 796, 809,  
812, 813, 814, 816,  
817, 820–822, 824,  
825–826, 827,  
828–834, 835, 837,  
854, 869, 948, 965  
Vestinian 812, 829  
Volscian 812, 828, 830, 837  
Sabellic = Sabellian  
Sabine *See* Sabellian  
Safaitic *See* North Arabian  
(Ancient North  
Arabian)  
Sahidic *See* Coptic  
Sam'al *See* Aramaic (Old  
Aramaic)  
Samaritan Aramaic *See*  
Aramaic  
San Miguel Chimalapa *See*  
Soke

- Sanskrit 11, 117, 535, 536,  
     538, 539, 540, 541,  
     542, 548, 582, 588,  
     620, 637, 642, 643,  
     668, 673–699, 700,  
     701, 703–705,  
     706–709, 710–715,  
     742, 747, 748–749,  
     750, 751–753, 756,  
     760, 761, 762, 798,  
     851, 869, 873, 890,  
     895, 896, 901, 905,  
     1012, 1015, 1037,  
     1117, 1118, 1119,  
     1120, 1124  
 Classical Sanskrit 674, 678,  
     680, 682, 689, 690,  
     691, 692, 693, 696,  
     697, 698, 700, 701,  
     708, 711, 712, 713, 714  
 Epic Sanskrit 673  
 Vedic Sanskrit 539, 540,  
     546, 547, 548, 557,  
     560, 561, 564, 566,  
     567, 568, 569, 636,  
     673, 674, 678, 680,  
     682, 684, 685, 686,  
     688, 689, 690, 691,  
     692, 693, 696, 697,  
     698, 700, 705, 708,  
     711, 713, 721,  
     726–727, 731, 735,  
     736, 737, 739, 743,  
     761, 762, 850, 1117,  
     1124  
 Sanskritized Prākṛit *See* Indic  
     (Prākṛit)  
 Santa María Chimalapa *See*  
     Soke  
 Sapoteco = Zapotec  
 Śaurasenī *See* Indic (Middle  
     Indic)  
 Šayhadic 141, 143, 144, 148,  
     454  
 Sayula Mijejan *See* Mijejan  
 Scandinavian languages 881,  
     882  
     East Scandinavian 908  
     West Scandinavian 908  
 Scythian 60  
 Semitic (*See also* Common  
     Semitic;  
     Proto-Semitic) 2, 15,  
     19, 20, 21, 26, 27, 51,  
     54–55, 117, 138, 139,  
     140, 141–142, 155,  
     156, 160, 178, 188,  
     218, 230, 234, 241,  
     244, 246, 249, 252,  
     254, 260, 261, 262,  
     263, 264, 277, 288,  
     289, 290, 293, 295,  
     296, 297, 299, 300,  
     302, 304, 305, 310,  
     311, 315, 319, 337,  
     338, 346, 349, 350,  
     353, 356, 357, 361,  
     365, 368, 369, 370,  
     372, 373, 376, 377,  
     378, 381, 382, 383,  
     388, 391, 392, 402,  
     428, 430, 434, 435,  
     436, 439, 441, 442,  
     443, 444, 445, 446,  
     448, 449, 456, 457,  
     460, 462, 464, 469,  
     485, 494, 495, 500,  
     502, 503, 506, 513,  
     525, 529, 535, 555,  
     599, 615, 648, 656,  
     674, 740, 781, 877,  
     885, 967  
 Central Semitic (*See also*  
     Proto-Central  
     Semitic) 141, 144, 146,  
     151, 152, 155, 253,  
     319, 365, 429, 489  
 East Semitic 141, 154, 223,  
     288, 303, 429  
 Ethiopian Semitic 142, 144,  
     148, 154, 427, 434,  
     435, 437, 445, 447,  
     451, 452  
 Ethio-Semitic (*See also*  
     Proto-Ethio-Semitic;  
     Semitic [Ethiopian  
     Semitic]) 254, 428,  
     436, 438  
     Northern Ethio-Semitic  
         428  
     Southern Ethio-Semitic  
         428  
 Mahrian Semitic 142, 144,  
     148  
 Northern Semitic 445  
 Northwest Semitic (*See also*  
     Proto-Northwest  
     Semitic) 141, 143,  
     144, 148, 154, 163,  
     165, 288, 290–291,  
     292, 294, 295, 315,  
     319, 332, 338, 347,  
     361, 365, 372, 378,  
     379, 380, 386, 388,  
     389, 391, 396, 411,  
     429, 494, 495  
     South Semitic 290, 427,  
         429, 435, 454, 494,  
         495, 517  
     South West Semitic 489  
     West Semitic 69, 141, 143,  
         151, 152, 154, 155,  
         188, 223, 251, 252,  
         260, 270, 277, 288,  
         290, 291, 293, 297,  
         298, 302, 303, 304,  
         308, 309, 311, 315,  
         319, 377, 387–388,  
         429, 444  
 Semito-Hamitic 138, 160  
 Sicel 9  
 Sicilian (Greek) *See* Greek  
     Dialects  
 Siculan 9  
 Sidamo 139  
 Sidetic 551, 606  
 Silver Latin *See* Latin  
 Sinitic 988, 989  
 Sino-Caucasian 989  
 Sino-Tibetan (*See also* Proto-  
     Sino-Tibetan) 989,  
     1012  
 Sino-Tibetan-Austronesian  
     989  
 Slavic 12, 573, 657, 800, 809,  
     895, 978, 1125  
 Sogdian 739  
 Soke (*See also* Proto-Soke;  
     Sokean) 1082, 1083,  
     1085, 1090, 1096  
     Chiapas Soke 1082, 1105  
     Copainalá 1082, 1096  
     Magdalena 1082, 1096  
     Oaxaca Soke 1082, 1105  
     San Miguel Chimalapa  
         1082, 1088, 1093,  
         1094, 1105  
     Santa María Chimalapa  
         1082, 1095, 1096,  
         1098, 1105  
 Sokean (*See also*  
     Proto-Sokean; Soke)  
     1071, 1078, 1079,  
     1080, 1081, 1082,  
     1083, 1085, 1086,



- 1087, 1088, 1090,  
 1097–1098, 1099,  
 1104–1106  
 Ayapa Gulf Sokean 1105  
 Gulf Sokean 1082, 1083,  
 1105  
 Soteapan Gulf Sokean  
 1081, 1082, 1085,  
 1087, 1094, 1096, 1105  
 Textistepec Gulf Sokean  
 1105  
 Somali 139  
 Soqotri 141  
 Soteapan Gulf Sokean *See*  
 Sokean  
 South Arabian 141, 260, 290,  
 434, 435, 492, 495,  
 499, 508  
 Ancient South Arabian  
 454–487, 488, 489,  
 493, 494, 498, 499, 505  
 Hadramitic 142, 454,  
 455, 459, 470–472  
 Haramic 454, 476, 477,  
 478, 485  
 Madhābic 492  
 Minaic 454, 455, 459,  
 470–472, 484  
 Qatabanian 454, 455,  
 470–472, 484  
 Sabaic 427, 454, 455,  
 456, 459, 467, 471,  
 472, 480–481, 484,  
 485, 489, 493, 494,  
 509, 519, 521, 529  
 Early Sabaic 454, 455,  
 461, 465, 469, 470, 474  
 Late Sabaic 454, 455,  
 459, 462, 469, 474, 476  
 Middle Sabaic 454,  
 459, 465, 469, 474,  
 476, 483  
 Epigraphic South Arabian  
 428, 431  
 Modern South Arabian  
 142, 253, 428, 429,  
 435, 444, 445, 485,  
 488, 499  
 Old South Arabian 324,  
 427, 428, 429, 430,  
 435, 445  
 South Canaanite *See*  
 Canaanite  
 South Greek *See* Greek  
 Dialects
- South Picene *See* Sabellian  
 South Safaitic *See* North  
 Arabian (Ancient  
 North Arabian)  
 South Semitic *See* Semitic  
 South West Semitic *See*  
 Semitic  
 Southern Hebrew *See* Hebrew  
 Southern Thamudic *See*  
 North Arabian  
 (Ancient North  
 Arabian)  
 Southwestern Chu *See*  
 Chinese (Dialects)  
 Southwestern Iranian *See*  
 Iranian  
 Spanish 1054, 1078, 1093  
 Śri Lankan Tamil *See* Tamil  
 Standard Literary Aramaic  
*See* Aramaic  
 Standard Literary Sumerian  
*See* Sumerian  
 Standard Phoenician *See*  
 Phoenician  
 Subakhmimic *See* Coptic  
 Suberian 96, 277  
 Sumerian (*See also* Sumerian  
 Dialects) 15, 19–59,  
 60–61, 62, 65, 66, 70,  
 89–90, 97, 103, 117,  
 141, 152, 154, 155,  
 218, 220, 221, 223,  
 227, 229, 230, 233,  
 240, 251, 260, 262,  
 263, 264, 271, 276,  
 277, 278, 289, 387,  
 552, 553, 648, 766  
 Old Sumerian 45  
 Sumerian Dialects (*See also*  
 Sumerian) 23–24  
 Emesal 23–24  
 Standard Literary  
 Sumerian 23  
 Susian 60  
 Svan 967  
 Swedish 881  
 Syriac *See* Aramaic
- Tabuki Thamudic *See* North  
 Arabian (Ancient  
 North Arabian)  
 Tai 1012  
 Tamazight 138  
 Tamil 673  
 Continental Tamil 1015
- Epigraphic Tamil 1015  
 Medieval Tamil 1015,  
 1025, 1031, 1037  
 Mixed Tamil 1015  
 Modern Tamil 1015,  
 1023, 1024, 1025,  
 1029, 1031, 1035,  
 1037  
 Old Tamil 1015–1040  
 Early Old Tamil 1015,  
 1016  
 Late Old Tamil 1015,  
 1016  
 Middle Old Tamil 1016  
 Śri Lankan Tamil 1015,  
 1028  
 Tamil of the Academy  
 1016  
 Tarafit 138  
 Tselhit 138  
 Taymanitic *See* North Arabian  
 (Ancient North  
 Arabian)  
 Telegu 673  
 Textistepec Gulf Sokean *See*  
 Sokean  
 Thamudic (A, B, C, D, E) *See*  
 North Arabian  
 (Ancient North  
 Arabian)  
 Theran *See* Greek Dialects  
 Thessalian *See* Greek  
 Dialects  
 Thessaliotis *See* Greek  
 Dialects  
 Thracian 12, 780  
 Thraco-Dacian 780  
 Thraco-Phrygian 780  
 Tibetan 22  
 Tibeto-Burman 989, 997  
 Tigre 428, 434  
 Tigrinya 141, 428, 431, 433,  
 434  
 Tocharian 2, 535, 536, 548,  
 573, 809, 1119, 1120,  
 1122  
 Tocharian A 1120  
 Tocharian B 1120  
 Toda 1015  
 Totonac 1068  
 Totontepec Mije *See* Mije  
 Traditional Egyptian *See*  
 Egyptian  
 Tsaconian *See* Greek  
 Dialects

- 
- Tuareg 139  
 Tulu 1015  
 Turkish 12, 22  
 Turoyo *See* Aramaic  
 Tzeltal *See* Mayan  
 Tzeltalan *See* Mayan  
 Tzotzil *See* Mayan  
  
 Ugaritic (*See also*  
     Proto-Ugaritic) 7, 15,  
     97, 98, 105, 141, 148,  
     151, 270, 288–318,  
     329, 335, 337, 338,  
     343, 346, 349, 361,  
     382, 386, 388, 389,  
     396, 497  
 Umbrian *See* Sabellian  
 Urartian (*See also*  
     Proto-Urartian;  
     Proto-Urarto-  
     Hurrian) 15, 65, 95,  
     96, 97, 119–137, 277,  
     577, 923  
     Dialects 120  
  
 Vandal 882  
 Vannic 119  
 Vedic Sanskrit *See* Sanskrit  
 Venetic 535, 840–856, 866,  
     946  
     Ancient Venetic 842  
     Archaic Venetic 842  
     Latino-Venetic 842, 844,  
         847, 848, 851  
     Recent Venetic 842, 843,  
         844, 845, 847, 848  
 Vestinian *See* Sabellian  
 Volscian *See* Sabellian  
 Vulgar Latin *See* Latin  
  
 Welsh 11, 499  
     Old Welsh 8  
  
 West Chadic *See* Chadic  
 West Germanic *See* Germanic  
 West Greek *See* Greek Dialects  
 West Semitic *See* Semitic  
 Wolaytta 139  
  
 Xamir 139  
 Xanmet'i *See* Georgian  
  
 Yeniseian 989  
 Yucatec *See* Mayan  
 Yucatecan *See* Mayan  
  
 Zapotec (*See also*  
     Proto-Zapotec) 3, 15,  
     1075, 1109  
 Zapotecan (*See also*  
     Proto-Zapotecan)  
     1053, 1068, 1111  
     Dialects 1111  
 Zoquean 1053