

A CONCISE INTRODUCTION TO THE IPA

The Need for the IPA

English spelling is quite confusing today, but there was a time that English pronunciation (and that probably holds true for virtually any language) was highly predictable from spelling; words came to be spelled a certain way for a reason, after all. Pronunciation changes, however, and it changes at a rate with which it is almost impossible for spelling to keep up. In fact, it changes so fast that, in the time it would take for everyone to grow acquainted with virtually any spelling system, it would be more likely than not that that system would no longer accurately represent the underlying sounds.

Now, decades of research have it under good authority that native and non-native speakers learn a language in vastly different ways. One of the key differences is that the former group learns (or, better yet, acquires) the language first by listening, and just later goes into the specifics of representing in writing the sounds they hear and have learned to produce. Non-natives, on the other hand, by and large have their first contact with the language in its written form, and just later figure out how to listen for and produce the sounds supposedly represented by the writing system they have learned.

This means that the lack of correspondence between writing and sound poses a problem of particular significance for non-native speakers. After all, how can they be expected to produce sounds that bear no sensible correspondence to the writing system they have learned?

Of course, one answer is that learners should try to be exposed to spoken language as much and as often as possible, by trying to listen to and have frequent conversations with native speakers, for example. But this approach suffers from a fundamental limitation, which is another well-established conclusion of decades of research: non-native speakers hear sounds filtered by the sound systems of their native languages.

Contrary perhaps to common sense, the sounds that are harder to learn are not those that sound foreign, but rather those that sound – to non-natives – just like some other sound that exists in their native languages. Except that they do not: non-native speakers hear them as though they did, but in reality they are a different sound, close enough to a sound that exists in the non-native learners' own native inventory that they cannot really notice the difference. The inescapable conclusion is thus that listening and repetition can only get learners so far, unless they know exactly *what* to listen for and repeat.

That is all the more reason why an accurate system of representation is needed, and the *International Phonetic Alphabet* (IPA) came to fulfill this need; due to its usefulness, it rapidly gained widespread acclaim. We will now be detailing the key aspects of this transcription system, as well as the fundamental linguistic concepts on which it is based.

The IPA Transcription System

Two sounds are never exactly the same in an absolute physical sense, but there is a range of variation below which the human ear can no longer possibly distinguish them – sounds that fall within that range are said to be the same phone. Less restrictively, there is a range of variation below which a speaker of a particular language can no longer distinguish them – roughly stated, sounds that fall within that range can be considered the same phoneme in that language.

It follows that the definition of phoneme is inherently language-specific. As is, it is also rather subjective: which "speaker" are we talking about? Bilinguals from birth, for example, are likely to be able to make many more distinctions, since they would have learned the sound systems from both their native languages. But even if we restrict our analysis to monolinguals, it is still quite subjective. Some people may have had more exposure to foreign languages than others, or to different dialects of the same language than others, or can be simply naturally better at recognizing this kind of variation.

To get rid of the subjectivity, linguists have developed a litmus test to determine whether two sounds have different phonemic status, which is to inspect the existence of contrasting variation, that is, the existence of minimal pairs: two sounds are said to belong to the same phoneme if two different words exist that are distinguished just by those two sounds and no others. In English, for example, the sound commonly represented by the letter "t" is different from the sound commonly represented by the letter "sh," and the proof is the existence of minimal pairs such as "tip" and "ship" or "fit" and "fish."

It follows from the aforementioned definitions that some phonemes can have more than one possible realization, or, in other words, that multiple phones can correspond to the same phoneme in a given language. This phenomenon is called allophonic variation, and each such phone is called an allophone of the underlying phoneme.

With that in mind, we can now explain the principles informing the IPA transcription system, which employs two sets of symbols, called letters and diacritics. New IPA letters are assigned whenever necessary to represent contrasting phones, that is, phones that are known to correspond to different phonemes in at least one language. When two phones never show phonemic contrast, that is, when they always allophone in all languages in which they both appear, the same letter is used to represent them, modified by diacritics as necessary.

It has become customary to use slashes for a broad transcription representing only the underlying phonemes, and square brackets for a narrow notation representing particular phones. That is the convention we will be using in this work, using tildes to separate possible allophones in narrow transcription. So, in the previous example, we would broadly transcribe the initial consonants the words "tip" and "ship" as /t/ and /ʃ/, respectively, using slashes to indicate they represent different phonemes in English.

The Phonemes of English

The precise number of phonemes of a given language may of course vary across different dialects, but it also depends on some judgement calls as to which sound combinations should be counted merely as combinations of phonemes or, rather, as new phonemes entirely. Should the sound commonly represented by "ch" in words such as "**chip**" or "**chat**," for example, be taken as a new phoneme, or as simply a combination of the two phonemes /t/ and /ʃ/ we previously saw?

The answer is largely a matter of how we choose to analyze the language. The IPA indeed reserves no separate symbol for this sound, directing it to be transcribed just as /tʃ/. Native speakers, however, generally do not feel like they are making two separate sounds: they think of /tʃ/ as a sound of its own, so if we ask a native English speaker (not trained in linguistics) how many different sounds there are in the word "chip," we are much more likely to get three (rather than four) as a reply. Since different analyses are possible, we will then show this sound enclosed in parentheses in the table below, which details the consonantal phonemes of American English.

Sounds produced with obstruction or narrowing of the vocal tract (aka "consonants"):

Place vs. Manner	Labial	Dental	Alveolar	Post-Alveolar	Palatal	Velar	Glottal
Plosives	/p/ as in pit /b/ as in bit		/t/ as in tip /d/ as in dip			/k/ as in cat /g/ as in get	
Affricates				(/tʃ/ as in chip) (/dʒ/ as in just)			
Fricatives	/f/ as in fit /v/ as in vet	/θ/ as in think /ð/ as in there	/s/ as in sip /z/ as in zip	/ʃ/ as in ship /ʒ/ as in genre			/h/ as in hi
Nasals	/m/ as in me		/n/ as in nap		/ŋ/ as in sing		
Approximants			/l/ as in law	/r/ as in raw	/j/ as in yes	/w/ as in we	

Sounds produced with no obstruction or narrowing of the vocal tract (aka "vowels"):

Openness v. Backness	Front	Centralized	Back
Close to Near-Close	/i/ as in FLEECE	/u/ as in GOOSE	
Near-Close to Close-Mid	/ɪ/ as in KIT	/ʊ/ as in FOOT	
Close-Mid to Open-Mid	/eɪ/ as in FACE	/oʊ/ as in GOAT	/ɔɪ/ as in CHOICE
Open-Mid to Near-Open	/ɛ/ as in DRESS	/ʌ/ as in STRUT	/ɔ/ as in THOUGHT
Near-Open to Open	/æ/ as in TRAP /aʊ/ as in MOUTH /aɪ/ as in PRICE /ɑ/ as in LOT		

Allophones

We mentioned before that phonemes may have more than one possible realization (that is, they can be pronounced as different phones) depending on the context in which they appear. Perhaps the most famous example lies with the letter /t/, which can be produced in at least four or five different ways in American English, that is to say, four or five different allophones are possible for the single phoneme /t/ (we'll be using slashes to represent phonemes, as before, and square brackets to represent phones).

The /t/ in words such as /tɪp/ "tip" is pronounced [t^h]: the superscript "h" is a type of diacritic, and is employed in the IPA to indicate aspiration, that is, that the sound is produced with a simultaneous puff of air, much like the an actual [h], though often with less intensity and duration. Now, an unaspirated version of /t/ shows up following an /s/ in words such as /step/ "step," in which the /t/ is realized – guess what – simply as [t]; as surprising as this may sound, however, it's actually the least common allophone of /t/ in English.

When intervocalic (between two vowels) following a stressed syllable, the /t/ in American English is often realized as [ɾ], an alveolar tap. Note also that, in this case, the allophone takes a whole different IPA letter, rather than merely a diacritic mark. As we have learned, the reason for that is that there are languages in which the alveolar tap is a separate phoneme on its own. It's often the sound of "r" in many languages worldwide, such as Spanish and Portuguese, though not so for most current varieties of English. It used to be the case that British English employed a similarly tapped "r" as an allophone of /ɹ/ between vowels; there is even a joke that this pronunciation sounds "vetty British," because it sounded like an intervocalic /t/ for Americans!

It's worth pointing out that, for great many Americans as well as Canadians, [ɾ] is also a possibly realization of the phoneme /d/ between vowels, which leads to an important corollary: different phonemes can end up being realized the same way in certain environments. That's called neutralization. An extra step would be if they started to be realized the same way in all environments; a merger would then be said to have occurred, and they would be then the same phoneme.

Going forward, word-finally, the /t/ is often unreleased or "held," an allophone represented with yet another diacritic: [t̚], in a word such as /bɪt/ "bit." It may also show, often coupled with what's called a glottal reinforcement – a combination transcribed as [t̚ʔ] – before a syllabic /n/ (that is, before the consonant /n/ standing alone as the nucleus of its syllable), as in /bʌt.n/ "button."

There's also a final possibility which is the null allophone [], meaning that /t/ can make no sound at all. This happens for words like /wɪntə/ "winter," which in many places in the U.S. can be pronounced just like /wɪnə/ "winner." It will also happen in connected ("fast") speech with words such as /fækt̚s/ "facts," which you'll hear more often than not pronounced as something like [fæks:] (the colon is a diacritic indicating that the IPA letter before it is pronounced extra long).

Of course, many other phonemes have allophones in English; if you'd like to learn more, please get in touch by filling out the form below for details on how to book a session. We also offer a free accent evaluation to all potential customers, so be sure to include a link to a 1-2min [Vocaroo.com](https://www.vocaroo.com) recording of you speaking freely or reading a couple of paragraphs from a book you like, and we'll reach back to you with a fully personalized lesson plan! Rest assured the evaluation is entirely complimentary, though, as does not oblige you to purchase any of our products or services.

Copyright notice: All pages of this material are the exclusive property of [Phonology.net](https://www.phonology.net) and may only be copied, shared, or redistributed to third parties for non-commercial purposes and with explicit indication of the source/ownership.