Four corpus linguistic feature comparison

Rows are ranked by the number of X marks across the feature columns.

X = feature present \ = marginal case Blank = not attested Marginal cases are excluded from the X totals.

Corpus	Inconsistent introductory commas	Comma overuse	Serial or Oxford comma	Inconsistent sentence-final periods	Inconsistent contraction apostrophes	Colons	Semicolons	Inconsistent question marks	Sentence-initial And But Then Or	Quotation use*	P.S. use**	Em-dash use	etc. use***	Questions or rhetorical questions	Answering implied questions	shall use	must use	Misspellings	V NP Prep order	V Prep NP order	Frequent hyphenation or closed compounds	Underhyphenation or overspacing	Sentence-initial emphatic Yes	Sentence-initial So without comma	materials	fun	did meaning kill	promise	rather as degree adverb	comply or complied	about as approximator	Phrasal repetition	X total	Marginal total
Zodiac	X	X	\	X	X	X	X	X	X	\	X	\	X	X	X	X	X	X	X	X		X	X	X	X	X	X	X	X	X	X	X	28	3
Margolis	X	X	\			X	X		X	X	X	X	X	X	\	X	X	X	X	X	X			X					X	X	X	X	21	2
Bates	X		X		X				X	X					X	X	X	X	X	X		*	X	X							X	X	15	0
Dahlia	X	\		X				X		X				X		\	X		X						X	X	X	X					11	2

*Quotation use for Zodiac is not reported speech and does not include context for inside/outside punctuation.

**Punctuation for P.S.: For Margolis, P.S., for Zodiac, P.S., PS., PS

***Zodiac and Margolis differ on comma convention preceding the etc. as well as on use of single or multiple “etc.” Margolis does not use a comma after last list item while Zodiac does - a strong distinction and the largest piece of contradictory evidence of single authorship.

*Bates letter very limited in possible cases, but “brush offs” and “blue eyed” are unhyphenated (no overspacing)

Coding notes: The repeated sentence-initial conjunction category is combined. Serial/Oxford comma examples for Zodiac and Margolis not fully conventional context,

Distinguishes deontic must in Dahlia and Zodiac from epistemic must in Bates and Margolis.