

Category	Feature	Code/ Tag	Examples	Operationalisation	Normalisa tion unit	As coded by
Features for which there are no tags in the tagged texts						
General text properties	Total number of words	Words	<i>It's a shame that you'd have to pay to get that quality.</i> (= 14)	The number of tokens as tokenised by the Stanford Tagger, but excluding punctuation marks, brackets, symbols, genitive 's (POS), and filled pauses and interjections (FPUH). Contractions are treated as separate words, i.e., it's is tokenised as it and 's. Note that this variable is only used to normalise the frequencies of other linguistic features.	NA	Le Foll
General text properties	Average word length	AWL	<i>It's a shame that you'd have to pay to get that quality.</i> (42/12 = 3.50)	Total number of characters in a text divided by the number of words in that same text (as operationalised in the Words variable above, hence excluding filled pauses and interjections, cf. FPUH).	Words	Le Foll
General text properties	Lexical diversity	TTR	<i>It's a shame that you'd have to pay to get that quality.</i> (12/14 = 0.85)	Following Biber (1988), this feature is a type-token ratio measured on the basis of, by default, the first 400 words of each text only. It is thus the number of unique word forms within the first 400 words of each text divided by 400. This number of words can be adjusted in the command used to run the script (see instructions at the top of the MFTE script).	Words (by default first 400)	Le Foll
General text properties	Lexical density	LDE	<i>It's a shame that you'd have to pay to get that quality.</i> (3/14 = 0.21)	For this feature, tokens which are not on the list of the 352 function words from the {qdapDictionaries} R package, nor individual letters, or any of the fillers listed in FPUH are identified as content words. Lexical density is calculated as the ratio of these content words to the total number of words in a text.	Words	Le Foll
General text properties	Finite verbs	FV	<i>He discovered that the method involved imbibing copious amounts of tea. Ants can survive by joining together to morph into living rafts. Always wanted to experience the winter wonderland that Queen Elsa created?</i>	This feature is not directly listed in the MFTE output tables; however, it is used as a normalisation basis for many other linguistics features (see Normalisation column). It is calculated by tallying the number of occurrences of the following features: VPRT, VBD, VIMP, MDCA, MDCO, MDMM, MDNE, MDWO and MDWS.	NA	Le Foll
Features for which there are tags in tagged version of the texts processed by the MFTE						
Adjectives	Attributive adjectives	JJAT	<i>I've got a fantastic idea! I didn't sleep at all last night. Cheap, quick and easy fix!</i>	Whereas the Biber Tagger and the MAT first identify predicative adjectives and then consider all remaining J.* tags from the Stanford Tagger to be attributive adjectives, the MFTE proceeds the other way around because it is considerably easier to reliably identify attributive adjectives than it is predicative adjectives. Thus, all adjectives (J.*, as tagged by the Stanford Tagger) followed by another adjective, a noun or a cardinal number, or preceded by a determiner are tagged as attributive adjectives. Once these first attributive adjectives have been identified, an additional loop is run to capture any additional attributive adjectives found in lists of attributive adjectives.	Nouns	Le Foll

Adjectives	Predicative adjectives	JJPR	<i>That's right . One of the main advantages of being famous ... It must be absolutely wonderful .</i>	Once attributive adjectives have been identified (see JJAT) and tagged as JJAT, all remaining JJ, JJS and JJR tags are overwritten as JJPR. In addition, <i>ok</i> and <i>okay</i> in the construction <i>BE ok(ay)</i> are also tagged as JJPR. These words are otherwise identified as foreign words (FW) by the Stanford Tagger.	Finite verbs	Le Foll
Adverbials	Frequency references	FREQ	<i>We should always wear a mask. But he had found his voice again.</i>	Assigned to all occurrences of the frequency adverbs listed in the COBUILD (Sinclair et al. 1900: 270): <i>usually, always, mainly, often, generally, normally, traditionally, again, constantly, continually, frequently, ever, never, infrequently, intermittently, occasionally, often, periodically, rarely, regularly, repeatedly, seldom, sometimes</i> and <i>sporadically</i> .	Finite verbs	Le Foll
Adverbials	Place references	PLACE	<i>It's not far to go. I'll get it from upstairs . It's downhill all the way. It's there not here.</i>	Biber's (1988: 224) list of place adverbials was taken from Quirk et al. (1985:514ff) but inexplicably excludes many from this list. Those that do not fulfil other major functions were therefore added: downwind, eastward(s), westward(s), northward(s), southward(s), upwards, downwards, elsewhere, everywhere, here, offshore, nowhere, somewhere, thereabout(s) and there (but occurrences of there tagged as existential there (EX) by the Stanford Taggers were ignored). Only occurrences of far which have not previously identified as TIME references (e.g., so far, thus far) or emphatics (e.g., far better, far more) are tagged as PLACE references.	Finite verbs	Le Foll, adapted from Biber (1988)
Adverbials	Time references	TIME	<i>It will soon be possible. Now is the time. I haven't come across any issues yet .</i>	All occurrences of <i>afterwards, again, earlier, early, eventually, forever, formerly, immediately, initially, instantly, late, lately, later, momentarily, now, nowadays, once, originally, presently, previously, recently, shortly, simultaneously, subsequently, today, to-day, tomorrow, to-morrow, tonight, to-night, yesterday</i> . Following Nini (2014: 18), the word <i>soon</i> was not tagged as a time adverbial when followed by the word <i>as</i> . <i>Ago, already, beforehand, prior to</i> , and <i>far</i> (the latter only when proceeded by <i>so</i> or <i>thus</i> and not followed by an adjective or adverb), and <i>am</i> and <i>pm</i> as adverbs were added to the list, as well as <i>yet</i> tokens that have not previously been identified as concessives (CONC).	Finite verbs	Le Foll, adapted from Nini (2014)
Adverbials	Other adverbs	RB	<i>Unfortunately that's the case. Exactly two weeks. He could so easily but he knows better . He's still gonna come back .</i>	Corresponds to all the tokens tagged as RB, RBS, RBR or WRB by the Stanford Tagger apart from those identified as adverbs of frequency (FREQ), place (PLACE) or time (TIME), amplifiers (AMP), emphatics (EMPH), hedges (HDG) and downtoners (DWNT).	Words	Le Foll
Determinatives	s-genitives	POS	<i>the world 's two most populous country, my parents ' house</i>	As identified by the Stanford Tagger: the possessive endings on nouns ending in 's or '. Note that these tokens are not counted as Word in the computation of the lexical diversity (TTR) and average word length variables (AWL) features.	Nouns	Le Foll
Determinatives	Determiners	DT	<i>Is that a new top? The first line has to be interesting. Are they both Spice Girls? On either side of the page. To another room. They're five pounds each .</i>	As tagged by the Stanford Tagger (DT) (Santorini 1990: 2), with the exception of <i>that, this, these</i> and <i>those</i> which are counted as demonstratives (DEMO). Note that this Stanford Tagger category also includes pronouns such as <i>another</i> in <i>Shall I choose <u>another</u> ?</i>	Nouns	Le Foll

Determinatives	Quantifiers	QUAN	<i>Such a good time in like half an hour. She's got all these great ideas. It happens each and every time.</i>	All occurrences of pre-determiners as tagged by the Stanford Tagger, which includes the following "determiner-like elements when they precede an article or possessive pronoun" (Santorini 1990: 4): <i>nary, quite, rather</i> and <i>such</i> (e.g., quite a mess, rather a nuisance, many a moon), as well as all instances of all (unless immediately followed by <i>right</i> , cf. DMA), <i>any, a bit, both, each, every, few, half, many, much, several, some, lots, a lot (of), load(s) of, heaps of, wee, less</i> and <i>more</i> (as adjectives only).	Nouns	Le Foll
Determinatives	Numbers	CD	<i>That's her number one secret. Two eyes glowed just above the surface. It happened on 7 February, 2019.</i>	All cardinal numbers as identified by the Stanford Tagger. This includes dates written in numbers, e.g., 1994. In addition, numbers listed as list markers (LS) by the Stanford are overwritten as CD and strings of the type <code>\b[0-9]+th_/\b[0-9]+nd_/\b[0-9]+rd_</code> are also tagged as numbers (CD).	Words	Le Foll
Determinatives	Demonstratives	DEMO	<i>What are you doing this weekend? I love that film. Whoever did that should admit it.</i>	Assigned to all occurrences of <i>that, this, these</i> and <i>those</i> identified by the Stanford Tagger as determiners (DT).	Words	Le Foll
Discourse organisation	Elaborating conjunctions	ELAB	<i>Similarly, you may, for example, write bullet points insomuch as it helps you to focus your ideas.</i>	Assigned to <i>such that</i> (not followed by a determiner), <i>such as, inasmuch as, insofar as, insomuch as, in that, to the extent that, in particular, in conclusion, in sum, in summary, to summarise, to summarize, for example, for instance, in fact, in brief, in any event, in any case, in other words, e(.)g(.), in summary, viz(.), cf(.), i. e., namely, etc(.), likewise, namely, as well as similarly and accordingly when followed by a comma.</i>	Finite verbs	Le Foll
Discourse organisation	Coordinating conjunctions	CC	<i>Instead of listening to us, he also told John and Jill but at least his parents don't know yet.</i>	<i>This category takes the coordinating conjunctions (CC) tagged by the Stanford Tagger as its basis which include and, but, nor, or, yet, "as well as the mathematical operators plus, minus, less, times (in the sense of 'multiplied by') and over (in the sense of 'divided by'), when they are spelled out" (Santorini 1990: 2). However, conjunctions already captured by other variables are excluded from this count: yet is assigned to concessive (CONC). In addition, the following (multi-word) conjunctions are also included in this category: also, besides, moreover, further (when tagged as an adverb), furthermore, in addition, additionally, as well (as) (except when preceded by least), however (provided it is preceded or followed by a punctuation mark), ibid, on the one hand, on the other hand, instead, besides, conversely, by/in contrast, on the contrary, in/by comparison, whereas, whereby, whilst.</i>	Finite verbs	Le Foll
Discourse organisation	Causal conjunctions	CUZ	<i>He was scared because of the costume. Yeah coz he hated it.</i>	Assigned to all occurrences of <i>because</i> , <i>'cause, cos, cuz</i> and <i>coz</i> . The latter four were not included in Biber's (1988) original variable. According to Biber (1988: 236) <i>because</i> "is the only subordinator to function unambiguously as a causative adverbial". Whilst it is true that many subordinators, e.g., <i>as, for</i> , and <i>since</i> , can fulfil a range of functions, including causative, and were therefore not included in this category, the following adverbs and multi-word conjunctions were added since they mostly fulfil a causative function: <i>as a result, on account of, for that/this purpose, thanks to, to that/this end, consequently, in consequence, hence, so that, therefore, thus</i> .	Finite verbs	Le Foll, adapted from Biber (1988)

Discourse organisation	Concessive conjunctions	CONC	Even though the antigens are normally hidden...	Assigned to all occurrences of <i>although, though, tho, despite, except that, in spite of, albeit, granted that, nevertheless, nonetheless, notwithstanding, whereas, no matter + WH-word, (ir)regardless of</i> , and <i>granted</i> . Also assigned to <i>still</i> and <i>yet</i> when preceded by any punctuation mark or followed by a comma. Multi-word units are only counted as one occurrence of CONC.	Finite verbs	Le Foll
Discourse organisation	Conditional conjunctions	COND	<i>If I were you... Even if the treatment works...</i>	Assigned to all occurrences of <i>if, as long as, unless, lest, in that case, otherwise, whether</i> .	Finite verbs	Le Foll
Discourse organisation	Discourse/pragmatic markers	DMA	Well no they didn't say actually . Okay I guess we'll see how things go right ?	Assigned to "interactional signals and discourse markers" (as listed in Stenström 1994: 59 and cited in Aijmer 2002: 2): <i>actually, all right, anyway, God, goodness, gosh, OK, okay, right</i> (if tagged as an interjection by the Stanford Tagger), <i>well</i> (only if identified by the Stanford Tagger as an adverb or adjective and not if preceded by <i>as, how, very, really, quite</i> , a verb, an adjective or an adverb), <i>yes, yeah, yep, sure</i> (unless it is preceded by the verb <i>MAKE, for, not</i> or <i>you</i>). Verbal phrases such as <i>you know</i> and <i>I mean</i> were excluded from this variable since literal occurrences could not be automatically disambiguated occurrences as discourse markers. A number of markers from Stenström's list are also not assigned this tag because they are captured by other variables: <i>now</i> (TIME), <i>please</i> (POLITE), <i>really</i> (EMPH), <i>quite</i> and <i>sort of</i> (HDG). The following items were added: <i>lol, IMO, omg, wtf, nope, mind you, of course, whatever</i> and <i>damn</i> (unless tagged as a verb, or followed by an adjective; in the latter case it is an emphatic, cf. EMPH).	Words	Le Foll
Discourse organisation	Filled pauses and interjections	FPUH	Oh noooooo, Tiger's furious! Wow! Hey Tom! Er I don't know. Hmm .	Assigned to all occurrences of <i>ah+, aw+, oh+, eh+, er+, erm+, mm+, ow+, um+, huh+, uhu+, uhuh, mhm+, hm+</i> (but not <i>HM</i>), <i>oo+ps woo+ps, hi, hey</i> , and interjections identified by the Stanford Tagger and not assigned to another category. The plus sign (+) signifies that that the preceding letter can appear multiple times, i.e., <i>ahh</i> and <i>errrr</i> are also assigned this tag.	Words	Le Foll
Discourse organisation	<i>Like</i>	LIKE	Sounds like me. And just like his father. And he was like this isn't true. I wasn't gonna like do it.	Occurrences of <i>like</i> tagged as a preposition (IN) or adjective (JJ) by the Stanford Tagger are assigned this tag because, in spoken English, <i>like</i> typically fulfils a range of different functions, e.g., fillers and softeners, and attempts to disambiguate <i>like</i> as a preposition or conjunct proved too error-prone. This category excludes occurrences of <i>like</i> identified as the quotative BE + <i>like</i> (QLIKE) if the QLIKE feature is included (which, by default, it is not, cf. tagger evaluation).	Words	Le Foll
Discourse organisation	<i>So</i>	SO	She had spent so many summers there. So there you go.	Occurrences of <i>so</i> tagged as IN by the Stanford Tagger and not previously identified as either an emphatic (<i>so</i> + J.* / <i>much/many/little</i> ; EMPH) or an adverbial subordinator (<i>so that</i> + NN.* / J.* ; OSUB) are assigned this tag.	Words	Le Foll
Discourse organisation	Direct WH-questions	WHQU	What's happening? Why don't we call the game off? How? And who is Dinah, if I might venture to ask the question?	Assigned to <i>what, where, when, how, why, who, whom, whose</i> and <i>which</i> followed by a question mark within 15 tokens.	Finite verbs	Le Foll

Discourse organisation	Question tags	QUTAG	<i>Do they ? Were you ? It's just it's repetitive, isn't it ?</i>	Assigned to question marks preceded by (1) <i>innit, init</i> ; (2) a modal verb (MD) or <i>did</i> or <i>had</i> , and a personal pronoun (P.+); (3) a modal verb or <i>did</i> or <i>had</i> , a negation (XX0), and a personal pronoun; (4) <i>is, does, was</i> or <i>has</i> , followed by <i>it, she</i> or <i>he</i> ; (5) <i>is, does, was</i> or <i>has</i> , followed by a negation, and <i>it, she</i> or <i>he</i> ; (6) <i>do, were, are</i> or <i>have</i> , followed by <i>you, we</i> or <i>they</i> ; (7) <i>do, were, are</i> or <i>have</i> , followed by a negation, and <i>you, we</i> or <i>they</i> . In addition, the above patterns are not considered question tags if a question word occurs within six words to the left of the question mark; consequently, <i>Why did you do it?</i> is not assigned this tag but rather WHQU.	Finite verbs	Le Foll
Discourse organisation	Yes/no questions	YNQU	<i>Have you thought about giving up? May I take a seat? Do you mind?</i>	Assigned to any form of the verbs BE, HAVE, DO or a modal verb (MD) followed by a personal pronoun (P.+), a noun (NN.*), a negation (XX0) or determiner (DT) and then a question mark within three to 15 tokens, as long as no WH-question (WHQU) or yes/no question tag (YNQU) is present one or two tokens before the auxiliary verb. Note that this variable should not overlap with question tags (QUTAG).	Finite verbs	Le Foll
Discourse organisation	<i>that</i> relative clauses	THRC	<i>You must be very clever to find a use for something that costs nothing. I'll just run a cable that goes from here to there.</i>	Assigned to that identified as introducing a relative clause by the Stanford Tagger (WDT), unless it is immediately followed by a punctuation mark. Any remaining <i>that_WDT</i> tokens are typically mistagged demonstratives and are thus assigned to the DEMO category, e.g., <i>I don't think that's a problem that is.</i>	Finite verbs	Le Foll
Discourse organisation	<i>that</i> subordinate clauses (other than relatives)	THSC	<i>Did you know that the calendar we use today was started by Julius Caesar? She resented being told constantly that she was ignorant and stupid.</i>	Assigned to that tokens which have been tagged as IN by the Stanford Tagger and are not immediately followed by a punctuation mark. Remaining <i>that_IN</i> tokens are assigned to the demonstrative category (DEMO): these are end-of-sentences/utterances tokens which are typically misidentified by the Stanford Tagger, e.g., <i>Who was that?</i>	Finite verbs	Le Foll
Discourse organisation	Subordinator <i>that</i> omission	THATD	<i>I mean [THATD] you'll do everything. I thought [THATD] he just meant our side. You don't think [THATD] he's a drug dealer? I know [THATD] that's not his thing.</i>	The THATD tag is assigned to the following patterns: (1) a public, private or suasive verb followed by a demonstrative pronoun (DEMO) or I, we, he, she, it, they and then a verb (V.* or MD); (2) a public, private or suasive verb followed by I, we, he, she, it, they or a noun (N.*), and then by a verb (V.* or MD); (3) a public, private or suasive verb followed by an adjective (J.*), an adverb (RB), a determiner (DT, QUAN, CD) or a possessive pronoun (PRPS), and then a noun (N.*), and then a verb (V.* or MD), with the possibility of an intervening adjective (J.*) between the noun and its preceding word. This tag corresponds to Biber's (1988: 244) category but its operationalisation has been improved to avoid the algorithm erroneously tagging constructions such as <i>Why would I know that?</i> and <i>He didn't hear me thank God.</i>	Finite verbs	Le Foll, adapted from Biber (1988)
Discourse organisation	WH subordinate clauses	WHSC	<i>I'm thinking of someone who is not here today. Do you know whether the banks are open?</i>	Assigned when the words <i>what, where, when, how, whether, why, whoever, whomever, whichever, wherever</i> and <i>whenever</i> have not been previously identified as part of a WH question (WHQU). Though many attempts were made, it proved impossible to reliably disambiguate between relative and other subordinate WH-clauses, which is why they are pooled together in this category.	Finite verbs	Le Foll

Lexis	Total nouns (including proper nouns)	NN	<i>a cut, my coat, the findings, cruelty, comprehension, on Monday 6 Aug, the U.S., on the High Street</i>	Assigned to all singular (NN) and plural nouns (NNS) identified by the Stanford Tagger including proper nouns (NNP and NNPS). This variable differs from the Biber Tagger in that it includes nominalisations.	Words	Le Foll
Lexis	Noun compounds	NCOMP	<i>Surely this stone must be the last one to cover the dungeon entrance ! Experts say that the rare winter phenomenon is a natural occurrence.</i>	Assigned when two or more nouns follow each other without any intervening punctuation. The algorithm allows for the first noun to be a proper noun but not the second thus allowing for <i>Monday afternoon</i> and <i>Hollywood stars</i> but not <i>Barack Obama</i> and <i>Los Angeles</i> . It is also restricted to nouns with a minimum of two letters to avoid OCR errors (dots and images identified as individual letters and which are usually tagged as nouns by the Stanford Tagger) producing too many erroneous NCOMP's. Note that this feature works best with fully punctuated texts (see per-register recall and precision rates in the tagger documentation).	Nouns	Le Foll
Lexis	Emoji and emoticons	EMO	😍 🥰 🌈 :-) :DD XD <3 :/	Assigned to all emojis as of December 2018 (cf. https://unicode.org/emoji/charts/full-emoji-list.html) and to a range of emoticons, in particular three-character emoticons such as :-) . The source code also includes three lines which are by default commented out but can be uncommented for texts where short emoticons are expected. It is not recommended to use these lines for general English because they lead to a sharp decrease in precision: many of the shorter emoticons, e.g., :(:D :3 , are too easy to confuse with poorly scanned texts that are missing spaces, or with the punctuation styles of specific academic journals.	Words	Le Foll
Lexis	Hashtags	HST	<i>#phdlife #Buy1Get1Free</i>	Assigned to any string starting with a hashtag followed by at least three letters, digits or underscores.	Words	Le Foll
Lexis	URL and e-mail addresses	URL	<i>www.faz.net https://twitter.com elefoll@uos.de</i>	Assigned to all strings resembling a URL or an e-mail address (without claiming to only include valid URLs or e-mail addresses since this is not the aim). Regex for this feature was inspired by: https://mathiasbynens.be/demo/url-regex	Words	Le Foll
Negation	Negation	XX0	<i>Why do n't you believe me? There is no way that's happening any time soon. Nor am I.</i>	Biber's (1988) analytic and synthetic negation features were merged into one negation variable since the latter is too infrequent to be of use in the context of this study. This unique negation tag is assigned to the tokens <i>not_RB</i> , <i>n't_RB</i> , all occurrences of the words <i>nor</i> and <i>neither</i> , and <i>no</i> when followed by an adjective (J.*) or noun (NN.*).	Finite verbs	Le Foll
Prepositions	Prepositions	IN	<i>The Great Wall of China is the longest wall in the world. There are towers along the wall. I prefer to go to an art gallery. The objects on display are from all over the world.</i>	All items tagged as IN by the Stanford Tagger other than those assigned to CAUS, CONC, COND, OSUB, SO and LIKE.	Words	Le Foll

Pronouns	Reference to the speaker/writer	FPP1S	<i>I don't know. It isn't my problem.</i>	All occurrences of <i>me, myself</i> and <i>mine</i> and <i>I</i> if tagged by the Stanford Tagger as a pronoun, a list symbol (LS) or a foreign word (FW).	Finite verbs	Le Foll
Pronouns	Reference to the speaker/writer and other(s)	FPP1P	<i>We were told to deal with it ourselves.</i>	All occurrences of <i>us, we, our, ourselves</i> and <i>ours</i> , as well as the contracted form of <i>us</i> (e.g., in <i>let's</i>). All these terms are case insensitive but an exception for <i>US</i> was added as this usually refers to the United States of America.	Finite verbs	Le Foll
Pronouns	Reference to the addressee	SPP2	<i>If your model was good enough, you 'd be able to work it out.</i>	Following Biber (1988), all occurrences of <i>you, your, yourself, yourselves</i> . Following Nini (2014: 18), also includes <i>thy, thee</i> and <i>thysself</i> . In addition, the forms <i>ur, ye, y'all, ya, thine</i> and the nominal possessive pronoun <i>yours</i> were also added.	Finite verbs	Le Foll, adapted from Nini (2014)
Pronouns	<i>it</i> pronoun reference	PIT	<i>It fell and broke. I implemented it. Its impact has not yet been researched.</i>	All occurrences of the pronoun <i>it</i> . An exception was added for the all capital form <i>IT</i> which most frequently refers to <i>Information Technology</i> . Following Nini (2014: 18), also includes all occurrences of <i>itself</i> and <i>its</i> .	Finite verbs	Le Foll, adapted from Nini (2014)
Pronouns	<i>One</i> as a personal pronoun	PRP	<i>One would hardly suppose that your eye was as steady as ever.</i>	This tag consists of the remaining personal pronouns not yet tagged as either first (FPP1S and FPP1P), second (SPP2) or third (TPP3) person pronouns. In practice, this should only leave <i>one</i> .	Finite verbs	Le Foll
Pronouns	Reference to one non-interactant	TPP3S	<i>He is beginning to form his own opinions. She does tend to keep to herself.</i>	Following Biber (1988), all occurrences of <i>she, he, her, him, his, himself, herself</i> and <i>themselves</i> . Note that the singular <i>they</i> form can only be accounted for with the possessive pronoun: <i>themselves</i> .	Finite verbs	Le Foll
Pronouns	Reference to more than one non-interactant	TPP3P	<i>The text allows readers to grapple with their own conclusions. I wouldn't trust them.</i>	All occurrences of <i>they, them, themselves, theirs</i> and <i>em</i> when tagged by the Stanford Tagger as a pronoun.	Finite verbs	Le Foll
Pronouns	Quantifying pronouns	QUPR	<i>said Alice aloud, addressing nobody in particular.</i>	All occurrences of <i>anybody, anyone, anything, each other, everybody, everyone, everything, nobody, none, no one, nothing, somebody, someone</i> and <i>something</i> .	Finite verbs	Nini (2014)
Stance-taking devices	Politeness markers	POLITE	<i>Can you open the window, please? Would you mind giving me a hand? I was wondering whether you could help.</i>	Assigned to all occurrences of <i>thanks, thank you, cheers, ta</i> (unless it is preceded by <i>got</i> to avoid the confusion with <i>gotta</i>), <i>please, sorry, apology, apologies</i> , all forms of the verbs <i>excuse, I/we wonder, I/we + BE + wondering</i> , and the multi-word units <i>you mind</i> and <i>don't mind</i> . No exception was made for <i>please</i> as a verb because the Stanford Tagger frequently misidentifies <i>please</i> as a verb, e.g., <i>I was like please_VPRT just please_VB just get there</i> .	Words	Le Foll

Stance-taking devices	Amplifiers	AMP	<i>I am very tired. They were both thoroughly frightened.</i>	Assigned to the amplifiers from Biber's (1988) list: <i>absolutely, altogether, completely, enormously, entirely, extremely, fully, greatly, highly, intensely, perfectly, strongly, thoroughly, totally, utterly, very. Especially</i> was added.	Words	Le Foll, adapted from Biber (1988)
Stance-taking devices	Downtoners	DWNT	<i>These tickets were only 45 pounds. It's almost time to go.</i>	Assigned to all occurrences of <i>almost, barely, hardly, merely, mildly, nearly, only, partially, partly, practically, scarcely, slightly, somewhat</i> . In Biber (1988) <i>almost</i> is listed as both a hedge and a downtoner. Following Nini (2014), it is only considered a downtoner here.	Words	Nini (2014)
Stance-taking devices	Emphatics	EMPH	<i>I do wish I hadn't drunk quite so much. Oh really? I just can't get my head around it.</i>	Following Biber (1988), assigned to all occurrences of <i>just, really, most, more, real</i> + ADJ, <i>so</i> + ADJ, <i>for sure, such a</i> . The algorithm was improved by adding <i>so</i> + <i>much/little/many, such a/ an</i> (whilst excluding <i>such a/an</i> if proceeded by <i>of</i>), and ensuring that only <i>DO</i> + verb in base form (VB) are tagged. <i>Least</i> and <i>far</i> + J.*/RB were added (the latter only when not proceeded by <i>so</i> or <i>thus</i>). To account for recent language change (Aijmer 2018), <i>bloody, dead</i> + ADJ, <i>fucking</i> and <i>super</i> were also added. Multi-word units are counted as one EMPH tag but several Words.	Words	Le Foll, adapted from Biber (1988)
Stance-taking devices	Hedges	HDG	<i>There seemed to be no sort of chance of getting out. I wish that kind of thing never happened. She's maybe gonna do it.</i>	Following Biber (1988: 240) assigned to all occurrences of <i>maybe, at about, something like, and more or less</i> , as well as <i>sort of</i> and <i>kind of</i> as long as they are not preceded by a determiner (DT), quantifier (QUAN), cardinal number (CD), adjective (J.*), possessive pronoun (PRPS) or WH word. The condition that <i>kind</i> must have been tagged as a noun (NN) by the Stanford Tagger was added to exclude phrases such as <i>it's very kind of you</i> . <i>Kinda</i> and <i>sorta</i> was added as colloquial alternatives to <i>kind of</i> and <i>sort of</i> and the adverbs <i>apparently, conceivably, perhaps, possibly, presumably, probably, roughly</i> and <i>somewhat</i> were also added to the list.	Words	Le Foll, adapted from Biber (1988)
Stative forms	Existential <i>there</i>	EX	<i>There are students. And there is now a scholarship scheme.</i>	As tagged by the Stanford Tagger: "Existential <i>there</i> is the unstressed <i>there</i> that triggers inversion of the inflected verb and the logical subject of a sentence" (p. 3).	Finite verbs	Le Foll
Stative forms	Be as main verb	BEMA	<i>It was nice to just be at home. She's irreplaceable. It's best I think. How was your mum on Sunday? It's not long.</i>	Following Biber (1988), this tag is assigned to the all forms of the verb <i>be</i> when followed by a determiner (DT), a possessive pronoun (PRPS) a preposition (IN), or an adjective (JJ). In addition, Nini (2014: 20) improved the Biber Tagger "by taking into account that adverbs or negations can appear between the verb <i>BE</i> and the rest of the pattern. Furthermore, the algorithm was slightly modified and improved: (a) the problem of a double-coding of any Existential <i>there</i> followed by a form of <i>BE</i> as a BEMA was solved by imposing the condition that there should not appear immediately before or two before the pattern; (b) the cardinal numbers (CD) tag and the personal pronoun (PRP) tag were added to the list of items that can follow the form of <i>BE</i> ." This latter improvement by Nini, however, resulted in tag questions also being assigned to BEMA. The present algorithm therefore further excludes any occurrences of <i>BE</i> found one or two to the left of a question tag (QUTAG), as well as <i>BE</i> occurrences one or two to the left of a present participle form tagged as PROG or past participle form tagged as PASS.	Finite verbs	Le Foll, adapted from Nini (2014)

Syntax	Split auxiliaries and infinitives	SPLIT	<i>I would actually drive . You can just so tell . I can't ever imagine arguing with Jill.</i>	This category merges Biber's (1988) split auxiliaries and split infinitive categories and follows Nini's (2014: 30) operationalisations. Hence, this tag is assigned every time the infinitive marker to (TO) is followed by one or two adverbs and a verb base form, and every time an auxiliary (any modal verb MD, or any form of DOAUX, or any form of BE, or any form of HAVE) is followed by one or two adverbs and a verb form. Nini's algorithm was improved to ensure that negated split auxiliaries would also be identified, e.g., They have not yet developed cancer.	Finite verbs	Le Foll, adapted from Nini (2014)
Syntax	Stranded prepositions	STPR	<i>We've got more than can be accounted for . Open the door and let them in . Where is it from ? It's not the sort of music we're into .</i>	As in Biber (1988), assigned to the prepositions <i>against, amid, amidst, among, amongst, at, between, by, despite, during, except, for, from, in, into, minus, of, off, on, onto, opposite, out, per, plus, pro, than, through, throughout, thru, toward, towards, upon, versus, via, with, within</i> and <i>without</i> followed by any punctuation mark. Following Nini (2014: 30), <i>besides</i> was removed from Biber's original list since it also frequently serves as a conjunct and, in this function, is usually followed by a punctuation mark. Note that Nini's (2014:30) operationalisation tagged all occurrences of these word forms as prepositions regardless of how they were tagged by the Stanford Tagger. Here, it was decided to improve accuracy by restricting the query to tokens tagged as IN by the Stanford Tagger (thus excluding many RB and RP tokens, e.g., <i>Don't take it <u>away</u> ! Tie her <u>up</u> ! He roared <u>out</u> : "Come <u>away</u> !")</i>).	Finite verbs	Le Foll, adapted from Nini (2014)
Verb features	Verbal contractions	CONT	<i>I do n't know. It is n't my problem. You 'll have to deal with it.</i>	Following (Nini 2014: 29), all occurrences of an apostrophe followed by a word identified as a verb (V. *, MD) by the Stanford Tagger and all occurrences of the token <i>n't _XX0</i> .	Finite verbs	Nini (2014)
Verb features	Particles	RP	<i>I'll look it up . It's coming down . When will you come over ? Some of the birds hurried off at once.</i>	As tagged by the Stanford Tagger (RP) (Santorini 1990: 9-10).	Finite verbs	Le Foll
Verb features	BE -passives	PASS	<i>He must have been burgled . They need to be informed . He was found out. When were they arrested ?</i>	Assigned to past participles (here: VBN or VBD) preceded by the following patterns: 1) any form of the verb BE; 2) BE followed by one or two adverb(s) (RB) and/or a negation (XX0); 3) BE followed by a noun (NN. *) or personal pronoun (PRP); 4) BE followed by a noun (NN. *) or personal pronoun, and an adverb (RB) or negation (XX0). Unlike Biber (1988), no subdivision is made for by-passives and agentless passives. This choice is a) theoretically motivated because passives are too infrequent to be robustly measured at this level of granularity in most texts and b) for practical reasons because the algorithm proposed to identify by-passives resulted in too many false positives (e.g., looking for things that have been made by hand).	Finite verbs	Le Foll
Verb features	GET- passives	PGET	<i>He's gonna get sacked . She'll get me executed . It gets done all the time.</i>	Assigned to past participles (here: VBN or VBD) preceded by the following patterns: 1) any form of the verb GET; 2) GET followed by a noun (NN. *) or personal pronoun (PRP); 3) GET followed by a determiner (DT) or a noun (NN. *) plus a noun (NN. *).	Finite verbs	Le Foll

Verb features	Going to constructions	GTO	<i>I'm not gonna go. You're going to absolutely love it there! Gonna come along?</i>	Assigned to all occurrences of <i>going to</i> and <i>gonna</i> followed by a base form verb (VB), allowing for up to one intervening word between <i>going to</i> or <i>gonna</i> and the infinitive. GTO constructions are excluded from the progressive (PROG) count.	Finite verbs	Le Foll
Verb features	Past tense	VBD	<i>It fell and broke. I implemented it. If I were rich.</i>	As tagged by the Stanford Tagger, except where VBD tags are assumed to have been misassigned by the Stanford Tagger and are instead attributed to the perfect aspect (PEAS), passives (PASS, PGET) or USEDTO categories.	Finite verbs	Le Foll
Verb features	Non-finite verb -ing forms	VBG	<i>He texted me saying no. He just started laughing. I remember thinking about that.</i>	All verb forms ending in <i>-ing</i> as tagged by the Stanford Tagger, except those identified as progressives (PROG) or <i>going to</i> constructions (GTO). This category also includes "putative prepositions" ending in <i>-ing</i> such as <u>according</u> to and <u>concerning</u> your request (Santorini 1990: 11).	Finite verbs	Le Foll
Verb features	Non-finite -ed verb forms	VBN	<i>These include cancers caused by viruses. Our content is grouped into sections called topics. Have you read any of the books mentioned in the blog?</i>	As tagged by the Stanford Tagger except for the exclusion of tokens identified as instances of the perfect aspect (PEAS), passives (PASS, PGET) and <i>used to</i> constructions (USEDTO). Note that according to the Stanford Tagger rules, this category includes "putative prepositions" ending in <i>-ed</i> such as <u>granted</u> that and <u>provided</u> that (Santorini 1990: 11).	Finite verbs	Le Foll
Verb features	Imperatives	VIMP	<i>Let me know! Read the website and write the names of the characters. In groups, share your opinion. Always do as you're told!</i>	This tag is first assigned to any verb in base form (VB) occurring 1) immediately after a punctuation mark except a comma (e.g., Okay: do it!), an emoji or emoticon (EMO), a symbol (SYM), hashtag (HST), foreign word (FW) or a list marker (LS), or 2) after a punctuation mark and an adverb (e.g., 1A. Then practice the dialogue), unless the VB token is please or thank or has previously been identified as a DO auxiliary (DOAUX). In a second loop, the VIMP tag is assigned to VB verb tokens (except thank or please) when preceded by an imperative as identified above, with up to two optional intervening tokens, and the tokens and or or (e.g., Describe or draw, Listen carefully and repeat, Read the text and answer the questions). In addition, a number of verbs frequently found in instructions are listed as exceptions (e.g., Complete, Choose, Check) and are always assigned to this category when they are found at the beginning of a sentence regardless of their tag because these were found to be frequently erroneously identified by the Stanford Tagger as nouns (NN).	Finite verbs	Le Foll
Verb features	Present tense	VPRT	<i>It's ours. Who does n't love it? I know.</i>	Subsumes the VBP (present tense other than third-person singular) and VBZ (third-person singular present tense) tags assigned by the Stanford Tagger. The MFTE also corrects systematic errors in the Stanford Tagger output by adding VPRT tags in strings such as <i>I dunno</i> and <i>there's</i> .	Finite verbs	Le Foll, adapted from Nini (2014)

Verb features	Perfect aspect	PEAS	<i>Have you been on a student exchange? She'd already seen it. He has been told before. Is this the last novel you've read?</i>	Assigned to past participles (VBN, VBD) preceded by the following patterns: 1) any form of the verb HAVE; 2) HAVE followed by one or two adverb(s) (RB) and/or a negation (XX0); 3) HAVE followed by a noun (NN. *) or personal pronoun (PRP); 4) HAVE followed by a noun (NN. *) or personal pronoun, and an adverb (RB) or negation (XX0); 5) HAVE followed by a participle tagged as a passive (PASS); 6) HAVE followed by one or two adverb(s) (RB) and/or a negation (XX0), and a passive participle (PASS); 7) HAVE followed by a noun (NN. *) or personal pronoun (PRP), and a passive participle (PASS); 8) 's as a verb (VBZ) followed by <i>been, had, done</i> or a stative verb; 9) 's as a verb (VBZ) followed by an adverb (RB) or negation (XX0), and <i>been, had, done</i> or a stative verb (as listed under JJPR).	Finite verbs	Le Foll
Verb features	Progressive aspect	PROG	<i>He wasn't paying attention. I'm going to the market. I'm guessing you're not going to be alone. I must be getting home.</i>	Assigned to any form of BE followed by an <i>-ing</i> form of any verb (VBG). The algorithm allows for an intervening adverb (RB), emphatic (EMPH) and/or negation (XX0). The interrogative form is captured as BE followed by a noun (N. *) or personal pronoun (PRP) followed by the VBG token. As for the affirmative version, the latter algorithm also accounts for an intervening adverb (RB) and/or negation (XX0). <i>Going to</i> constructions are excluded from this category and are tagged separately (GTO).	Finite verbs	Le Foll
Verb features	HAVE got constructions	HGOT	<i>He's got some. I haven't got any.</i>	Assigned to the word <i>got</i> preceded by the following patterns: 1) any form of the verb HAVE; 2) HAVE followed by one or two adverb(s) (RB) and/or a negation (XX0); 3) HAVE followed by a noun (NN, NNP) or personal pronoun (PRP); 4) HAVE followed by a noun (NNP, NNP) or personal pronoun, and an adverb (RB) or negation (XX0). Note that this algorithm overwrites the perfect aspect (PEAS) and passive (PASS) tag.	Finite verbs	Le Foll
Verb semantics	DO auxiliary	DOAUX	<i>Should take longer than it does. Ah you did. She needed that house, did n't she? You don't really pay much attention, do you? Who did not already love him.</i>	Assigned to do, does and did as verbs in the following patterns: (a) when the next but one token is a base form verb (VB) (e.g., <i>did it work?</i> , <i>didn't hurt?</i>); (b) when the next but two token (+3) is a base form verb (VB) (e.g., <i>didn't it work?</i>); (c) when it is immediately followed by an end-of-sentence punctuation mark (e.g., <i>you did?</i>); (d) when it is followed by a personal pronoun (PRP) or <i>not</i> or <i>n't</i> (XX0) and an end-of-sentence punctuation mark (e.g., <i>do you? He didn't!</i>); (e) when it is followed by <i>not</i> or <i>n't</i> (XX0) and a personal pronoun (PRP) (e.g., <i>didn't you?</i>); (f) when it is followed by a personal pronoun followed by any token and then a question mark (e.g., <i>did you really? did you not?</i>); (g) when it is preceded by a WH question word. Additionally, all instances of DO immediately preceded by <i>to</i> as an infinitive marker (TO) are excluded from this tag.	Finite verbs	Le Foll
Verb semantics	Activity verbs	ACT	<i>I got up and ran out. Bring your CV. Where have you worked before? I go to school.</i>	Assigned to all forms of the verbs: <i>buy, make, give, take, come, use, leave, show, try, work, move, follow, put, pay, bring, meet, play, run, hold, turn, send, sit, wait, walk, carry, lose, eat, watch, reach, add, produce, provide, pick, wear, open, win, catch, pass, shake, smile, stare, sell, spend, apply, form, obtain, arrange, beat, check, cover, divide, earn, extend, fix, hang, join, lie, obtain, pull, repeat, receive, save, share, smile, throw, visit, accompany, acquire, advance, behave, borrow, burn, clean, climb, combine, control, defend, deliver, dig, encounter, engage, exercise, expand, explore</i> and <i>reduce</i> (cf. Biber 2006: 246, based on the LGSWE, pp. 361–362, 367–368, 370). <i>Do</i> is only included when it has not previously been tagged as an auxiliary (DOAUX). <i>Get</i> and <i>go</i> were removed from Biber's (2006) list due to their high polysemy. Like Biber (2006), for practical reasons, no phrasal verbs were included in this variable.	Finite verbs	Le Foll, based on Biber (2006)

Verb semantics	Aspectual verbs	ASPECT	<i>You should just keep talking. I started early today.</i>	Following Biber (2006: 247, based on the LGSWE, pp. 364, 369, 371), assigned to all forms of the verbs: <i>start, keep, stop, begin, complete, end, finish, cease</i> and <i>continue</i> .	Finite verbs	Biber 2006
Verb semantics	Facilitation and causative verbs	CAUSE	<i>He helped her escape. I pleaded with her to let me go.</i>	Following Biber (2006: 247, based on the LGSWE, pp. 363, 369, 370), assigned to all forms of the verbs: <i>help, let, allow, affect, cause, enable, ensure, force, prevent, assist, guarantee, influence, permit</i> and <i>require</i> .	Finite verbs	Biber 2006
Verb semantics	Communication verbs	COMM	<i>Describe it to your partner and say why. Write a list. Say what these words mean.</i>	Following Biber (2006: 247, based on the LGSWE, pp. 362, 368, 370), assigned to all forms of the verbs: <i>say, tell, call, ask, write, talk, speak, thank, describe, claim, offer, admit, announce, answer, argue, deny, discuss, encourage, explain, express, insist, mention, offer, propose, quote, reply, shout, sign, sing, state, teach, warn, accuse, acknowledge, address, advise, appeal, assure, challenge, complain, consult, convince, declare, demand, emphasize, excuse, inform, invite, persuade, phone, pray, promise, question, recommend, remark, respond, specify, swear, threaten, urge, welcome, whisper</i> and <i>suggest</i> . British spellings and the verbs <i>agree, assert, beg, confide, command, disagree, object, pledge, pronounce, plead, report, testify, vow</i> and <i>mean</i> were added. The latter was on Biber's (2006) list for mental verbs but, in most contexts encountered in the present study, it was found to be more likely to be a communication verb.	Finite verbs	Le Foll, based on Biber (2006)
Verb semantics	Existential or relationship verbs	EXIST	<i>Weren't they representing Jamaica? It encouraged young athletes to stay.</i>	Following Biber (2006: 247, based on the LGSWE, pp. 364, 369, 370–371), assigned to all forms of the verbs: <i>seem, stand, stay, live, appear, include, involve, contain, exist, indicate, concern, constitute, define, derive, illustrate, imply, lack, owe, own, possess, suit, vary, deserve, fit, matter, reflect, relate, remain, reveal, sound, tend</i> and <i>represent</i> . This variable does not include the copular <i>be</i> . <i>Look</i> was removed from Biber's original list because it frequently acts as an activity verb, too, e.g., <i>I was looking for my glasses</i> .	Finite verbs	Le Foll, based on Biber (2006)
Verb semantics	Mental verbs	MENTAL	<i>We want to see you tomorrow. Did you never hear back? I don't recognize any.</i>	Following Biber (2006: 246–247, based on the LGSWE, pp. 362–363, 368–369, 370), assigned to all forms of the verbs: <i>see, know, think, want, need</i> (unless identified as a necessity modal; cf. MDNE), <i>feel, like, hear, remember, believe, read, consider, suppose, listen, love, wonder, understand, expect, hope, assume, determine, agree, bear, care, choose, compare, decide, discover, doubt, enjoy, examine, face, forget, hate, identify, imagine, intend, learn, mind, miss, notice, plan, prefer, prove, realize, recall, recognize, regard, suffer, wish, worry, accept, appreciate, approve, assess, blame, bother, calculate, conclude, celebrate, confirm, count, dare, detect, dismiss, distinguish, experience, fear, forgive, guess, ignore, impress, interpret, judge, justify, observe, perceive, predict, pretend, reckon, remind, satisfy, solve, study, suspect</i> and <i>trust</i> . British spellings were added. <i>Afford</i> and <i>find</i> which can be found on Biber's original list, were removed for being too polysemous. Note that the phrase <i>dunno</i> , which is incorrectly parsed by the Stanford Tagger, was also retagged as <i>du_VPRT n_XX0 no_VB</i> and that <i>no_VB</i> tokens are also assigned to this category.	Finite verbs	Le Foll, based on Biber (2006)

Verb semantics	Occurrence verbs	OCCUR	<i>Couldn't have happened at a busier time! The cricket lasts all day.</i>	Following Biber (2006: 247, based on the LGSWE pp. 364, 369, 370), assigned to all forms of the verbs: <i>become, happen, change, die, grow, develop, arise, emerge, fall, increase, last, rise, disappear, flow, shine, sink, slip</i> and <i>occur</i> .	Finite verbs	Biber 2006
Verb semantics	Necessity modals	MDNE	<i>I really must go. Should n't you be going now? You need not have worried. Everybody needed to be needed.</i>	As in Biber (1988), all occurrences of <i>ought, should</i> and <i>must</i> . Contrary to Nini's operationalisation (2014: 27), only occurrences tagged as modals (MD) by the Stanford Tagger were included. In addition, <i>need</i> when tagged as a modal by the Stanford Tagger (mostly when followed by <i>not</i> or <i>n't</i>) or when immediately followed by <i>to</i> not tagged as a preposition (IN) was also added to this variable.	Finite verbs	Le Foll, adapted from Biber (1988)
Verb semantics	Modal <i>can</i>	MDCA	<i>Can I give him a hint? You can not. I ca n't believe it!</i>	All occurrences of <i>can</i> and <i>ca</i> tagged as modals by the Stanford Tagger (MD). <i>Ca</i> was included because the Stanford Tagger parses <i>can't</i> as <i>ca + n't</i> .	Finite verbs	Le Foll
Verb semantics	Modal <i>could</i>	MDCO	<i>Do you think someone could have killed her? Well, that could be the problem. Could you do it by Friday?</i>	All occurrences of <i>could</i> tagged as a modal by the Stanford Tagger (MD).	Finite verbs	Le Foll
Verb semantics	Modals <i>may</i> and <i>might</i>	MDM M	<i>May I have a word with you? But it might not be enough.</i>	All occurrences of <i>may</i> and <i>might</i> tagged as modals by the Stanford Tagger (MD).	Finite verbs	Le Foll
Verb semantics	<i>will</i> and <i>shall</i> modals	MDWS	<i>It wo n't do. Yes it will. Shall we see?</i>	The tokens <i>will</i> and <i>shall</i> and their contractions <i>'ll, wo</i> and <i>sha</i> when tagged as modals by the Stanford Tagger (MD).	Finite verbs	Le Foll
Verb semantics	modal <i>would</i>	MDWO	<i>Would n't you like to know? If I could afford to buy it I would. I 'd like to think it works.</i>	The tokens <i>will</i> and <i>shall</i> and their contractions <i>'ll, wo</i> and <i>sha</i> when tagged as modals by the Stanford Tagger (MD).	Finite verbs	Le Foll
Verb semantics	<i>be able to</i>	ABLE	<i>It should be able to speak back to you. Would you be able to?</i>	Assigned to occurrences of the bigram (un)able to, whenever (un)able has previously been identified as a predicative adjective (JJPR). These occurrences of (un)able are subsequently excluded from the JJPR count.	Finite verbs	Le Foll
Tags not counted by MFTE but important to understand the operationalisation of other features						
Lexis	Foreign words	FW	<i>I chose turkish delight and panna cotta. Merry christmasss! Yo im gonna love it!</i>	All remaining words tagged by the Stanford Tagger as foreign words and not identified as other variables by the MFTE. Frequently includes words spelt with non-standard spellings, missing apostrophes, and poorly OCR'ed due to unusual fonts. Note that this feature is not counted by the MFTE.	NA	Stanford Tagger
Lexis	Symbols	SYM	<i>â 2 € a go. I hope so †. That's *all* they said!</i>	All remaining non alphanumeric tokens tagged by the Stanford Tagger as symbols (SYM) or list markers (LS) and not identified as other variables by the MFTE. Also frequently includes words poorly OCR'ed due to unusual fonts or poorly encoded text. Note that this feature is not counted by the MFTE.	NA	Stanford Tagger

Verb features	to -infinitives	TO	<i>They were trying to find a solution. We like to think it's doable. I went in there to kinda like celebrate.</i>	Following Nini (2014: 21), all occurrences of <i>to</i> except when followed by another _IN token, a number (CD), determiner (DT), adjective (J.*), possessive pronoun (PRPS), WH-word (WPS, WDT, WP, WRB), pre-determiner (PDT), noun (N.*) or pronoun (PRP). Note that, unlike Nini (2014), this feature is only used to identify other linguistic features. All occurrences of <i>to</i> are counted as prepositions (IN) in the MFTE output tables.	NA	Nini (2014)
Verb features	Verb base form	VB	<i>She would sit and read most afternoons. What do you use it for? Ask your parents to drive you to your friend's house.</i>	As tagged by the Stanford Tagger, except those identified as imperatives (VIMP). This feature is not included in the tables of counts outputted by the MFTE because it overlaps with other features (e.g., all the modal verb features). However, it is used to identify many other linguistic features.	NA	Le Foll
Verb semantics	Private verbs	NA	<i>I don't think this should be assumed. I suspect he can't even remember it.</i>	As in Biber (1988, based on 1985: 1181), all forms of the verbs <i>accept, anticipate, ascertain, assume, believe, calculate, check, conclude, conjecture, consider, decide, deduce, deem, demonstrate, determine, discern, discover, doubt, dream, ensure, establish, estimate, expect, fancy, fear, feel, find, foresee, forget, gather, guess, hear, hold, hope, imagine, imply, indicate, infer, insure, judge, known, learn, mean, note, notice, observe, perceive, presume, presuppose, pretend, prove, realize, reason, recall, reckon, recognize, reflect, remember, reveal, see, sense, show, signify, suppose, suspect, think</i> and <i>understand</i> . Note that this category is only used to identify <i>that</i> -omissions (THATD).	NA	Biber 1988
Verb semantics	Public verbs	NA	<i>She promised she'd write back.</i>	As in Biber (1988, based on 1985: 1181), all forms of the verbs <i>acknowledge, add, admit, affirm, agree, allege, announce, argue, assert, bet, boast, certify, claim, comment, complain, concede, confess, confide, confirm, contend, convey, declare, deny, disclose, exclaim, explain, forecast, foretell, guarantee, hint, insist, maintain, mention, object, predict, proclaim, promise, pronounce, prophesy, protest, remark, repeat, reply, report, retort, say, state, submit, suggest, swear, testify, vow, warn</i> and <i>write</i> . Note that this category is only used to identify <i>that</i> -omissions (THATD).	NA	Le Foll, adapted from Biber (1988)
Verb semantics	Suasive verbs	NA	<i>They were determined to make this work. I'd prefer to do it that way.</i>	As in Biber (1988, based on 1985: 1182–3), all forms of the verbs <i>agree, allow, arrange, ask, beg, command, concede, decide, decree, demand, desire, determine, enjoin, ensure, entreat, grant, insist, instruct, intend, move, ordain, order, pledge, pray, prefer, pronounce, propose, recommend, request, require, resolve, rule, stipulate, suggest, urge</i> and <i>vote</i> . Note that this category is only used to identify <i>that</i> -omissions (THATD).	NA	Biber 1988
Features removed from the MFTE feature portfolio post-evaluation of v.2.9 (Note that the corresponding lines are commented out in v.3.0+ and may still be run, if wished)						
Verb semantics	Quotative BE + <i>like</i>	QLIKE	<i>I was like oh this is really good. And everyone is like let's do this.</i>	Assigned to any form of <i>BE</i> followed by <i>like</i> tagged as a preposition (IN) by the Stanford Tagger and not followed by a noun (NN.*), adjective (J.*), determiner (DT), preposition (IN) or a full stop, comma, exclamation or question mark. This feature is deactivated by default but can be uncommented in the script.	Finite verbs	Le Foll

Verb features	<i>Used to</i> constructions	USEDTO	<i>You'll get used to it in time. It works but not like it used to .</i>	Assigned to all occurrences of the bigram <i>used to</i> . These occurrences of <i>used</i> are excluded from the VBN/VBD counts. This feature is deactivated by default but can be uncommented in the script.	Finite verbs	Le Foll
Discourse organisation	Phrasal coordination	PHC	<i>read and write, positive or negative, nouns and adjectives</i>	All occurrences of <i>and</i> , <i>&</i> , <i>or</i> and <i>nor</i> in the following patterns: adverb + <i>and/or/nor</i> + adverb, adjective + <i>and/or/nor</i> + adjective, verb + <i>and/or/nor</i> + verb, noun + <i>and/or/nor</i> + noun. This feature was removed post-evaluation and these occurrences of <i>and</i> and <i>or</i> are now all included in coordinating conjunctions (CC).	NA	Biber 1988