Category	Feature	Code/Tag	Examples	Operationalisation	Normalisation unit	As coded by
				Features for which there are no tags in the tagged texts		
General text properties	Total number of words	**Words**	*It's a shame that you'd have to pay to get that quality.* (= 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	NA	Le Foll
General text properties	Average word length	**AWL**	*It's a shame that you'd have to pay to get that quality.* (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**	***It's a shame that you'd have to pay*** *to **get** that **quality** .* (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	Words (by default first 400)	Le Foll
General text properties	Lexical density	**LDE**	*It's a **shame** that you'd have to **pay** to get that **quality** .* (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**	*He discovered that the method involved imbiding 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*	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've got a **fantastic** idea! I didn't sleep at all **last** night. **Cheap** , **quick** and **easy** fix!*	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,	Nouns	Le Foll
Adjectives	Predicative adjectives	**JJPR**	*That's **right** . One of the main advantages of being **famous** ... It must be absolutely **wonderful** .*	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, *ok* and *okay* in the construction *BE ok(ay)* 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**	*We should **always** wear a mask. But he had found his voice **again.***	Assigned to all occurrences of the frequency adverbs listed in the COBUILD (Sinclair et al. 1900: 270): *usually, always, mainly, often, generally, normally, traditionally, again, constantly, continually, frequently, ever, never, infrequently, intermittently, occasionally, often, periodically, rarely, regularly, repeatedly, seldom, sometimes*	Finite verbs	Le Foll
Adverbials	Place references	**PLACE**	*It's not **far** to go. I'll get it from **upstairs**. It's **downhill** all the way. It's **there** not here.*	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.,	Finite verbs	Le Foll, adapted from Biber (1988)
Adverbials	Time references	**TIME**	*It will **soon** be possible. **Now** is the time. I haven't come across any issues **yet**.*	All occurrences of *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.* Following Nini (2014: 18), the word *soon* was not tagged as a time adverbial when followed by the word *as*. *Ago, already, beforehand, prior to,* and *far* (the latter only when proceeded by *so* or *thus* and not followed by an adjective or adverb).	Finite verbs	Le Foll, adapted from Nini (2014)
Adverbials	Other adverbs	**RB**	***Unfortunately** that's the case. **Exactly** two weeks. He could **so easily** but he knows **better**. He's **still** gonna come **back**.*	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**	*the world**'s** two most populous country, my parents**'** house*	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**	*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**.*	As tagged by the Stanford Tagger (DT) (Santorini 1990: 2), with the exception of *that, this, these* and *those* which are counted as demonstratives (DEMO). Note that this Stanford Tagger category also includes pronouns such as *another* in *Shall I choose* <u>another</u> *?*	Nouns	Le Foll
Determinatives	Quantifiers	**QUAN**	***Such** a good time in like **half** an hour. She's got **all** these great ideas. It happens **each** and **every** time.*	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): *nary, quite, rather* and *such* (e.g., quite a mess, rather a nuisance, many a moon), as well as all instances of all (unless immediately followed by *right*, cf. DMA), *any, a bit, both, each, every, few, half, many, much, several, some, lots, a lot (of), load(s) of, heaps of, wee, less* and *more* (as adjectives only).	Nouns	Le Foll

Determinatives	Numbers	**CD**	*That's her number **one** secret. **Two** eyes glowed just above the surface. It happened on **7** February, **2019**.*	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 *\b[0-9]+th_|\b[0-9]+nd_|\b[0-9]+rd_* are also tagged as numbers (CD).	Words	Le Foll
Determinatives	Demonstratives	**DEMO**	*What are you doing **this** weekend? I love **that** film. Whoever did **that** should admit it.*	Assigned to all occurrences of *that, this, these* and *those* identified by the Stanford Tagger as determiners (DT).	Words	Le Foll
Discourse organisation	Elaborating conjunctions	**ELAB**	***Similarly**, you may, **for example**, write bullet points **insomuch as** it helps you to focus your ideas.*	*Assigned to such that (not followed by a determiner), 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*	Finite verbs	Le Foll
Discourse organisation	Coordinating conjunctions	**CC**	***Instead** of listening to us, he **also** told John **and** Jill **but** at least his parents don't know yet.*	*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*	Finite verbs	Le Foll
Discourse organisation	Causal conjunctions	**CUZ**	*He was scared **because** of the costume. Yeah **coz** he hated it.*	Assigned to all occurrences of *because*, *'cause, cos, cuz* and *coz*. The latter four were not included in Biber's (1988) original variable. According to Biber (1988: 236) *because* "is the only subordinator to function unambiguously as a causative adverbial". Whilst it is true that many subordinators, e.g., *as, for,* and *since*, 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: *as a result, on account of, for that/this*	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 *although, though, tho, despite, except that, in spite of, albeit, granted that, nevertheless, nonetheless, notwithstanding, whereas, no matter +* WH-word, *(ir)regardless of*, and *granted.* Also assigned to *still* and *yet* when preceded by any punctuation mark or followed by a comma. Multi-word	Finite verbs	Le Foll
Discourse organisation	Conditional conjunctions	**COND**	***If** I were you... Even **if** the treatment works...*	Assigned to all occurrences of *if, as long as, unless, lest, in that case, otherwise, even when* and *whether.*	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): *actually, all right, anyway, God, goodness, gosh, OK, okay, right* (if tagged as an interjection by the Stanford Tagger), *well* (only if identified by the Stanford Tagger as an adverb or adjective and not if preceded by *as, how, very, really, quite,* a verb, an adjective or an adverb), *yes, yeah, yep, sure* (unless it is preceded by the verb *MAKE, for, not* or *you*). Verbal phrases such as *you know* and *I mean* 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: *now* (TIME), *please* (POLITE), *really* (EMPH), *quite* and *sort of* (HDG). The following items were	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 ah+, aw+, oh+, eh+, er+, erm+, mm+, ow+, um+, huh+, uhu+, uhuh, mhm+, hm+ (but not HM), oo+ps woo+ps, hi, hey, and interjections identified by the Stanford Tagger and not assigned to another category. The plus sign (+) signifies that that the preceding letter can appear	Words	Le Foll
Discourse organisation	*Like*	**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 like tagged as a preposition (IN) or adjective (JJ) by the Stanford Tagger are assigned this tag because, in spoken English, like typically fulfils a range of different functions, e.g.,fillers and softeners, and attempts to disambiguate like as a preposition or conjunct proved too error-prone. This category excludes occurrences of like identified as the quotative BE + like (QLIKE) if the QLIKE feature	Words	Le Foll
Discourse organisation	*So*	**SO**	*She had spent* ***so*** *many summers there.* ***So*** *there you go.*	Occurrences of *so* tagged as IN by the Stanford Tagger and not previously identified as either an emphatic (*so +* J.*/*much/many/little* ; EMPH) or an adverbial subordinator (*so that* + 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 *what, where, when, how, why, who, whom, whose* and *which* followed by a question mark within 15 tokens.	Finite verbs	Le Foll
Discourse organisation	Question tags	**QUTAG**	*Do they* ***?*** *Were you* ***?*** *It's just it's repetitive, isn't it* ***?***	Assigned to question marks preceded by (1) *innit, init;* (2) a modal verb (MD) or *did* or *had,* and a personal pronoun (P.+); (3) a modal verb or *did* or *had,* a negation (XX0), and a personal pronoun; (4) *is, does, was* or *has* , followed by *it, she* or *he;* (5) *is, does, was* or *has* , followed by a negation, and *it, she* or *he* ; (6) *do, were, are* or *have* , followed by *you, we* or *they* ; (7) *do, were, are* or *have* , followed by a negation, and *you, we* or *they* . In addition, the above patterns are not considered question tags if a question word occurs within six words to the left	Finite verbs	Le Foll
Discourse organisation	Yes/no questions	**YNQU**	***Have*** *you thought about giving up?* ***May*** *I take a seat?* ***Do*** *you mind?*	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	Finite verbs	Le Foll

Discourse organisation	*that* relative clauses	**THRC**	*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.*	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 that_WDT tokens are typically mistagged demonstratives and are thus assigned to the DEMO category, e.g., *I don't think that's a problem that is.*	Finite verbs	Le Foll
Discourse organisation	*that* subordinate clauses (other than relatives)	**THSC**	*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.*	Assigned to that tokens which have been tagged as IN by the Stanford Tagger and are not immediately followed by a punctuation mark. Remaining that_IN 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., *Who was that?*	Finite verbs	Le Foll
Discourse organisation	Subordinator *that* omission	**THATD**	*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.*	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	Finite verbs	Le Foll, adapted from Biber (1988)
Discourse organisation	WH subordinate clauses	**WHSC**	*I'm thinking of someone **who** is not here today. Do you know **whether** the banks are open?*	Assigned when the words *what, where, when, how, whether, why, whoever, whomever, whichever, wherever* and *whenever* 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	Finite verbs	Le Foll
Lexis	Total nouns (including proper nouns)	**NN**	*a **cut**, my **coat**, the **findings,** cruelty **, comprehension,** on **Monday** 6 **Aug,** the **U.S.** , on the **High Street***	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**	*Surely this stone must be the last one to cover the dungeon **entrance** ! Experts say that the rare winter **phenomenon** is a natural occurrence.*	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 *Monday afternoon* and *Hollywood stars* but not *Barack Obama* and *Los Angeles* . 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	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	Words	Le Foll
Lexis	Hashtags	**HST**	*#phdlife #Buy1Get1Free*	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**	*www.faz.net https://twitter.com elefoll@uos.de*	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**	*Why do **n't** you believe me? There is **no** way that's happening any time soon. **Nor** am 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 *not_RB, n't* _RB, all occurrences of the words *nor* and *neither* , and *no* when followed by an adjective	Finite verbs	Le Foll
Prepositions	Prepositions	**IN**	*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.*	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 don't know. It isn't **my** problem.*	All occurrences of *me, myself* and *mine* and *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**	***We** were told to deal with it **ourselves** .*	All occurrences of us, we, our, ourselves and ours, as well as the contracted form of us (e.g., in let's). All these terms are case insensitive but an exception for US was added as this usually refers to the United States of America.	Finite verbs	Le Foll
Pronouns	Reference to the addressee	**SPP2**	*If **your** model was good enough, **you** 'd be able to work it out.*	Following Biber (1988), all occurrences of *you, your, yourself, yourselves* . Following Nini (2014: 18), also includes *thy, thee* and *thyself* . In addition, the forms *u, ur, ye, y'all, ya* , *thine* and the nominal possessive pronoun *yours* were also added.	Finite verbs	Le Foll, adapted from Nini (2014)

Pronouns	*it* pronoun reference	**PIT**	***It*** *fell and broke. I implemented **it**. **Its** impact has not yet been researched.*	All occurrences of the pronoun *it*. An exception was added for the all capital form *IT* which most frequently refers to *Information Technology*. Following Nini (2014: 18), also includes all occurrences of *itself* and *its*.	Finite verbs	Le Foll, adapted from Nini (2014)
Pronouns	Any personal pronoun not included in the other categories	**PPother**	***One*** *would hardly suppose that your eye was as steady as ever.*	This tag consists of the remaining personal pronouns not yet tagged as either first (PR1), second (PR2) or third (PR3) person pronouns. In practice, this should only leave *one* and any misspelt, historical or dialectal forms not captured by the other categories.	Finite verbs	Le Foll
Pronouns	Reference to one non-interactant	**TPP3S**	***He*** *is beginning to form **his** own opinions. **She** does tend to keep to **herself**.*	Following Biber (1988), all occurrences of *she, he, her, him, his, himself, herself* and *themself*. Note that the singular *they* form can only be accounted for with the possessive pronoun: *themself.*	Finite verbs	Le Foll
Pronouns	Reference to more than one non-interactant	**TPP3P**	*The text allows readers to grapple with **their** own conclusions. I wouldn't trust **them**.*	All occurrences of *they, them, themselves, theirs* and *em* when tagged by the Stanford Tagger as a pronoun.	Finite verbs	Le Foll
Pronouns	Quantifying pronouns	**QUPR**	*said Alice aloud, addressing **nobody** in particular.*	All occurrences of *anybody, anyone, anything, each other, everybody, everyone, everything, nobody, none, no one, nothing, somebody, someone* and *something*.	Finite verbs	Nini (2014)
Stance-taking devices	Politeness markers	**POLITE**	*Can you open the window, **please**? Would you **mind** giving me a hand? I was **wondering** whether you could help.*	Assigned to all occurrences of *thanks, thank you, cheers, ta* (unless it is preceded by *got* to avoid the confusion with *gotta*), *please, sorry, apology, apologies,* all forms of the verbs *excuse, I/we wonder, I/we + BE + wondering,* and the multi-word units *you mind* and *don't mind*. No exception was made for *please* as a verb because the Stanford Tagger frequently misidentifies *please* as a verb, e.g., *I was*	Words	Le Foll
Stance-taking devices	Amplifiers	**AMP**	*I am **very** tired. They were both **thoroughly** frightened.*	Assigned to the amplifiers from Biber's (1988) list: *absolutely, altogether, completely, enormously, entirely, extremely, fully, greatly, highly, intensely, perfectly, strongly, thoroughly, totally, utterly, very. Especially* was added.	Words	Le Foll, adapted from Biber (1988)
Stance-taking devices	Downtoners	**DWNT**	*These tickets were **only** 45 pounds. It's **almost** time to go.*	Assigned to all occurrences of *almost, barely, hardly, merely, mildly, nearly, only, partially, partly, practically, scarcely, slightly, somewhat*. In Biber (1988) *almost* is listed as both a hedge and a downtoner. Following Nini (2014), it is only considered	Words	Nini (2014)
Stance-taking devices	Emphatics	**EMPH**	*I **do** wish I hadn't drunk quite so much. Oh **really**? I **just** can't get my head around it.*	Following Biber (1988), assigned to all occurrences of *just, really, most, more, real* + ADJ, *so* + ADJ, *for sure, such a*. The algorithm was improved by adding *so + much/little/many, such a/* an (whilst excluding *such a/an* if proceeded by *of*), and ensuring that only *DO* + verb in base form (VB) are tagged. *Least* and *far* + J.*/RB were added (the latter only when not proceeded by *so* or *thus*). To account for recent language change (Aijmer 2018), *bloody, dead* + ADJ, *fucking* and *super*	Words	Le Foll, adapted from Biber (1988)

Stance-taking devices	Hedges	**HDG**	*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.*	Following Biber (1988: 240) assigned to all occurrences of *maybe, at about, something like,* and *more or less,* as well as *sort of* and *kind of* 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 *kind* must have been tagged as a noun (NN) by the Stanford Tagger was added to exclude phrases such as *it's very kind of you*. *Kinda* and *sorta* was added as colloquial alternatives to *kind of* and *sort of* and the adverbs *apparently*,	Words	Le Foll, adapted from Biber (1988)
Stative forms	Existential *there*	**EX**	***There*** *are students. And **there** is now a scholarship scheme.*	As tagged by the Stanford Tagger: "Existential *there* is the unstressed *there* 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**	*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.*	Following Biber (1988), this tag is assigned to the all forms of the verb be 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 BE and the rest of the pattern. Furthermore, the algorithm was slightly modified and improved: (a) the problem of a double-coding of any Existential there followed by a form of BE 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 BE." This latter improvement by Nini, however, resulted in tag questions also being assigned to BEMA. The present algorithm therefore further excludes any occurrences of BE found one or two to the left of a	Finite verbs	Le Foll, adapted from Nini (2014)
Syntax	Split auxiliaries and infinitives	**SPLIT**	*I would actually **drive**. You can just so **tell**. I can't ever **imagine** arguing with Jill.*	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	Finite verbs	Le Foll, adapted from Nini (2014)
Syntax	Stranded prepositions	**STPR**	*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**.*	As in Biber (1988), assigned to the prepositions *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* and *without* followed by any punctuation mark. Following Nini (2014: 30), *besides* 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. *Don't take it away	Tie her up	He roared out*;	Finite verbs	Le Foll, adapted from Nini (2014)

Verb features	Verbal contractions	**CONT**	*I do **n't** know. It is **n't** my problem. You **'ll** have to deal with it.*	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 *n't* _XX0.	Finite verbs	Nini (2014)
Verb features	Particles	**RP**	*I'll look it **up.** It's coming **down** . When will you come **over** ? Some of the birds hurried **off** at once.*	As tagged by the Stanford Tagger (RP) (Santorini 1990: 9-10).	Finite verbs	Le Foll
Verb features	*BE* -passives	**PASS**	*He must have been **burgled** . They need to be **informed** . He was **found** out. When were they **arrested** ?*	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	Finite verbs	Le Foll
Verb features	*GET-* passives	**PGET**	*He's gonna get **sacked.** She'll get me **executed** . It gets **done** all the time.*	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.*)	Finite verbs	Le Foll
Verb features	*Going to* constructions	**GTO**	*I'm not **gonna** go. You're **going** to absolutely love it there! **Gonna** come along?*	Assigned to all occurrences of *going to* and *gonna* followed by a base form verb (VB), allowing for up to one intervening word between *going to* or *gonna* and the infinitive. GTO constructions are excluded from the progressive (PROG) count.	Finite verbs	Le Foll
Verb features	Past tense	**VBD**	*It **fell** and **broke** . I **implemented** it. If I **were** rich.*	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**	*He texted me **saying** no. He just started **laughing** . I remember **thinking** about that.*	All verb forms ending in *-ing* as tagged by the Stanford Tagger, except those identified as progressives (PROG) or *going to* constructions (GTO). This category also includes "putative prepositions" ending in *-ing* 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**	*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?*	As tagged by the Stanford Tagger except for the exclusion of tokens identified as instances of the perfect aspect (PEAS), passives (PASS, PGET) and *used to* constructions (USEDTO). Note that according to the Stanford Tagger rules, this category includes "putative prepositions" ending in *-ed* such as <u>granted</u> that and <u>provided</u> that (Santorini 1990: 11).	Finite verbs	Le Foll

Verb features	Imperatives	**VIMP**	*Let me know! **Read** the website and **write** the names of the characters. In groups, **share** your opinion. Always **do** as you're told!*	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	Finite verbs	Le Foll
Verb features	Present tense	**VPRT**	*It'**s** ours. Who **does** n't love it? I **know** .*	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 dunno* and *there's*.	Finite verbs	Le Foll, adapted from Nini (2014)
Verb features	Perfect aspect	**PEAS**	*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** ?*	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 *been, had, done*	Finite verbs	Le Foll
Verb features	Progressive aspect	**PROG**	*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.*	Assigned to any form of BE followed by an -ing 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). *Going to* constructions are excluded from this category and are tagged	Finite verbs	Le Foll
Verb features	*HAVE got* constructions	**HGOT**	*He's **got** some. I haven't **got** any.*	Assigned to the word *got* 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	Finite verbs	Le Foll

			Example	Description		Source
Verb semantics	DO auxiliary	**DOAUX**	*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.*	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., *did it work?, didn't hurt?*); (b) when the next but two token (+3) is a base form verb (VB) (e.g., *didn't it work*); (c) when it is immediately followed by an end-of-sentence punctuation mark (e.g., *you did?*); (d) when it is followed by a personal pronoun (PRP) or *not* or *n't* (XX0) and an end-of-sentence punctuation mark (e.g., *do you? He didn't!*); (e) when it is followed by *not* or *n't* (XX0) and a personal pronoun (PRP) (e.g., *didn't you?*); (f) when it is followed by a personal pronoun followed by any token and then a question mark (e.g., *did you really? did you not?*); (g) when it is preceded by a WH	Finite verbs	Le Foll
Verb semantics	Activity verbs	**ACT**	*I got up and ran out. Bring your CV. Where have you worked before? I go to school.*	Assigned to all forms of the verbs: *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* and *reduce* (cf. Biber 2006: 246, based on the LGSWE, pp. 361–362, 367–368, 370). *Do* is only included when it has not previously been tagged as an auxiliary (DOAUX). *Get* and *go* were removed from Biber's (2006) list	Finite verbs	Le Foll, based on Biber (2006)
Verb semantics	Aspectual verbs	**ASPECT**	*You should just keep talking. I started early today.*	Following Biber (2006: 247, based on the LGSWE, pp. 364, 369, 371), assigned to all forms of the verbs: *start, keep, stop, begin, complete, end, finish, cease* and	Finite verbs	Biber 2006
Verb semantics	Facilitation and causative verbs	**CAUSE**	*He helped her escape. I pleaded with her to let me go.*	Following Biber (2006: 247, based on the LGSWE, pp. 363, 369, 370), assigned to all forms of the verbs: *help, let, allow, affect, cause, enable, ensure, force, prevent, assist, guarantee, influence, permit* and *require.*	Finite verbs	Biber 2006
Verb semantics	Communication verbs	**COMM**	*Describe it to your partner and say why. Write a list. Say what these words mean.*	Following Biber (2006: 247, based on the LGSWE, pp. 362, 368, 370), assigned to all forms of the verbs: *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* and *suggest.* British spellings and the verbs *agree, assert, beg, confide, command, disagree, object, pledge, pronounce, plead, report, testify, vow,* and *mean,* were added. The latter was on Biber's (2006) list for	Finite verbs	Le Foll, based on Biber (2006)
Verb semantics	Existential or relationship verbs	**EXIST**	*Weren't they representing Jamaica? It encouraged young athletes to stay.*	Following Biber (2006: 247, based on the LGSWE, pp. 364, 369, 370–371), assigned to all forms of the verbs: *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* and *represent*. This variable does not include the copular *be*. *Look* was removed from Biber's original list because it frequently acts as an activity verb, too, e.g., *I*	Finite verbs	Le Foll, based on Biber (2006)

Category	Feature	Tag	Example	Operationalisation	Scope	Source
Verb semantics	Mental verbs	**MENTAL**	*We **want** to **see** you tomorrow. Did you never **hear** back? I don't **recognize** any.*	Following Biber (2006: 246-247, based on the LGSWE, pp. 362–363, 368–369, 370), assigned to all forms of the verbs: *see, know, think, want, need* (unless identified as a necessity modal; cf. MDNE)*, 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* and *trust.* British spellings were added. *Afford* and *find* which can be found on Biber's original list, were removed for being too	Finite verbs	Le Foll, based on Biber (2006)
Verb semantics	Occurrence verbs	**OCCUR**	*Couldn't have **happened** at a busier time! The cricket **lasts** all day.*	Following Biber (2006: 247, based on the LGSWE pp. 364, 369, 370), assigned to all forms of the verbs: *become, happen, change, die, grow, develop, arise, emerge, fall, increase, last, rise, disappear, flow, shine, sink, slip* and *occur* .	Finite verbs	Biber 2006
Verb semantics	Necessity modals	**MDNE**	*I really **must** go. **Should** n't you be going now? You **need** not have worried. Everybody **needed** to be needed.*	As in Biber (1988), all occurrences of *ought, should* and *must*. Contrary to Nini's operationalisation (2014: 27), only occurrences tagged as modals (MD) by the Stanford Tagger were included. In addition, *need* when tagged as a modal by the Stanford Tagger (mostly when followed by *not* or *n't*) or when immediately followed by *to* not tagged as a preposition (IN) was also added to this variable.	Finite verbs	Le Foll, adapted from Biber (1988)
Verb semantics	Modal *can*	**MDCA**	***Can** I give him a hint? You **can** not. I **ca** n't believe it!*	All occurrences of *can* and *ca* tagged as modals by the Stanford Tagger (MD). *Ca* was included because the Stanford Tagger parses *can't* as *ca + n't.*	Finite verbs	Le Foll
Verb semantics	Modal *could*	**MDCO**	*Do you think someone **could** have killed her? Well, that **could** be the problem. **Could** you do it by Friday?*	All occurrences of *could* tagged as a modal by the Stanford Tagger (MD).	Finite verbs	Le Foll
Verb semantics	Modals *may* and *might*	**MDMM**	***May** I have a word with you? But it **might** not be enough.*	All occurrences of *may* and *might* tagged as modals by the Stanford Tagger (MD).	Finite verbs	Le Foll
Verb semantics	*will* and *shall* modals	**MDWS**	*It **wo** n't do. Yes it **will**. **Shall** we see?*	The tokens *will* and *shall* and their contractions *'ll, wo* and *sha* when tagged as modals by the Stanford Tagger (MD).	Finite verbs	Le Foll
Verb semantics	modal *would*	**MDWO**	***Would** n't you like to know? If I could afford to buy it I **would** . I **'d** like to think it works.*	The tokens *will* and *shall* and their contractions *'ll, wo* and *sha* when tagged as modals by the Stanford Tagger (MD).	Finite verbs	Le Foll
Verb semantics	*be able to*	**ABLE**	*It should be **able** to speak back to you. Would you be **able** to?*	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 chose turkish delight and **panna cotta . Merrry christmasss! Yo im** gonna love it!*	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**	*â 2 € a go. I hope so **†** . That's ***** all ***** they said!*	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**	*They were trying **to** find a solution. We like **to** think it's doable. I went in there **to** kinda like celebrate.*	Following Nini (2014: 21), all occurrences of *to* 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 *to* are counted as prepositions (IN) in	NA	Nini (2014)
Verb features	Verb base form	**VB**	*She would **sit** and **read** most afternoons. What do you **use** it for? Ask your parents to **drive** you to your friend's house.*	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 don't **think** this should be **assumed** . I **suspect** he can't even **remember** it.*	As in Biber (1988, based on 1985: 1181), all forms of the verbs 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,*	NA	Biber 1988
Verb semantics	Public verbs	NA	*She **promised** she'd **write** back.*	As in Biber (1988, based on 1985: 1181), all forms of the verbs *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*	NA	Le Foll, adapted from Biber (1988)
Verb semantics	Suasive verbs	NA	*They were **determined** to make this work. I'd **prefer** to do it that way.*	As in Biber (1988, based on 1985: 1182–3), all forms of the verbs *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* and *vote* . Note that this category is only used to identify	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 + *like*	**QLIKE**	*I was **like** oh this is really good. And everyone is **like** let's do this.*	Assigned to any form of *BE* followed by *like* 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	*Used to* constructions	**USEDTO**	*You'll get **used** to it in time. It works but not like it **used** to .*	Assigned to all occurrences of the bigram *used to* . These occurrences of *used* 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**	**read *and*** write, positive **or** negative, nouns **and** adjectives	All occurrences of *and, &, or* and *nor* in the following patterns: adverb + *and/or/nor* + adverb, adjective + *and/or/nor* + adjective, verb + *and/or/nor* + verb, noun + *and/or/nor* + noun. This feature was removed post-evaluation and these occurrences of *and* and *or* are now all included in coordinating conjunctions	NA	Biber 1988
Extended features added in the MFTE python version 1.0						
Lexis	Superlatives	**SUPER**	*Its **latest** will the **boldest** yet.*	All adjectives and adverbs ending in "*est*" except *honest, test, west, pest, lest* and *guest* , as well as *worst* and *best* followed by a hypthen. Also constructions of the type: *the least/most* + ADJ/ADV.	Words	Le Foll
Lexis	Comparatives	**COMPAR**	*A **larger** producer. This approach might be **harder** to pull off now.*	All adjectives and adverbs ending in "er" except *after, never, rather, other, ever, together, proper, super, clever, eager, queer, hyper, inner, utter, bitter, premier, sinister, sober, order* and *over,* as well as *worse* and *better* followed by a	Words	Le Foll
Syntax	*that* subordinate clauses (other than relatives) preceded by attitutidinal adjectives	**ThJATT**		*That* complement clause (THSC) tag preceded by any of the following words tagged as adjective: *afraid, amazed, (un)aware, concerned, disappointed, encouraged, glad, happy, hopeful, pleased, shocked, surprised, worried* **.**	Finite verbs	**Biber 2006, adapted by Author**
Syntax	*that* subordinate clauses (other than relatives) preceded by adjectives of evaluation	**ThJEVL**		*That* complement clause (THSC) tag preceded by any of the following words tagged as adjective: *apparent, certain, clear, confident, convinced, correct, evident, false, impossible, inevitable, obvious, positive, right, sure, true, well-known.*	Finite verbs	**Biber 2006, adapted by Author**

Syntax	*that* subordinate clauses (other than relatives) preceded by factive adjectives	ThJFCT	*That* complement clause (THSC) tag preceded by any of the following words tagged as adjective: *afraid, amazed, (un)aware, concerned, disappointed, encouraged, glad, happy, hopeful, pleased, shocked, surprised, worried* .	Finite verbs	**Biber 2006, adapted by Author**
Syntax	*that* subordinate clauses (other than relatives) preceded by likelihood adjectives	ThJLIK	*That* complement clause (THSC) tag preceded by any of the following words tagged as adjective: *doubtful, likely, possible, probable, unlikely* .	Finite verbs	**Biber 2006, adapted by Author**
Syntax	*that* subordinate clauses (relative clauses) preceded by attitudinal nouns	ThNATT	*That* relative clause (THRC) tag preceded by any of the following words tagged as noun: *grounds, ground, hope, hopes, reason, reasons, view, views, thought, thoughts* .	Nouns	**Biber 2006, adapted by Author**
Syntax	*that* subordinate clauses (relative clauses) preceded by factive nouns	ThNFCT	*That* relative clause (THRC) tag preceded by any of the following words tagged as noun: *assertion, assertions, conclusion, conclusions, conviction, convictions, discovery, discoveries, doubt, doubts, fact, facts, knowledge, knowledges, observation, observations, principle, principles, realization, realizations, result, results, statement, statements* .	Nouns	**Biber 2006, adapted by Author**
Syntax	*that* subordinate clauses (relative clauses) preceded by likelihood nouns	ThNLIK	*That* relative clause (THRC) tag preceded by any of the following words tagged as noun: *assumption, assumptions, belief, beliefs, claim, claims, contention, contentions, feeling, feelings, hypothesis, hypotheses, idea, ideas, implication, implications, impression, impressions, notion, notions, opinion, opinions, possibility, possibilities, presumption, presumptions, suggestion, suggestions* .	Nouns	**Biber 2006, adapted by Author**

Syntax	*that* subordinate clauses (relative clauses) preceded by non factive nouns	**ThNNFCT**	*That* relative clause (THRC) tag preceded by any of the following words tagged as noun: *comment, comments, news, news, proposal, proposals, proposition, propositions, remark, remarks, report, reports, requirement, requirements.*	Nouns	**Biber 2006, adapted by Author**
Syntax	*that* subordinate clauses (other than relatives) preceded by attitutidinal verbs	**ThVATT**	*That* complement clause (THSC) tag preceded by any of the following words tagged as verb: *agreeing, agreed, agree, agrees, anticipates, anticipated, anticipate, anticipating, complain, complained, complaining, complains, conceded, concede, concedes, conceding, ensure, expecting, expect, expects, expected, fears, feared, fear, fearing, feel, feels, feeling, felt, forgetting, forgets, forgotten, forgot, forget, hoped, hope, hopes, hoping, minding, minded, minds, mind, preferred, prefer, preferring, prefers, pretending, pretend, pretended, pretends, requiring, required, requires, require, wishes, wished, wish, wishing, worry, worrying, worries,*	Finite verbs	**Biber 2006, adapted by Author**

Syntax	*that* subordinate clauses (other than relatives) preceded by communicative verbs	**ThVCOMM**	*That* complement clause (THSC) tag preceded by any of the following words tagged as verb: *say, says, saying, said, tell, tells, telling, told, call, calls, calling, called, ask, asks, asking, asked, write, writes, writing, wrote, written, talk, talks, talking, talked, speak, speaks, spoke, spoken, speaking, thank, thanks, thanking, thanked, describe, describing, describes, described, claim, claims, claiming, claimed, offer, offers, offering, offered, admit, admits, admitting, admitted, announce, announces, announcing, announced, answer, answers, answering, answered, argue, argues, arguing, argued, deny, denies, denying, denied, discuss, discusses, discussing, discussed, encourage, encourages, encouraging, encouraged, explain, explains, explaining, explained, express, expresses, expressing, expressed, insist, insists, insisting, insisted, mention, mentions, mentioning, mentioned, offer, offers, offering, offered, propose, proposes, proposing, proposed, quote, quotes, quoting, quoted, reply, replies, replying, replied, shout, shouts, shouting, shouted, sign, signs, signing, signed, sing, sings, singing, sang, sung, state, states, stating, stated, teach, teaches, teaching, taught, warn, warns, warning, warned, accuse, accuses, accusing, accused, acknowledge, acknowledges, acknowledging, acknowledged, address, addresses, addressing, addressed, advise, advises, advising, advised, appeal, appeals, appealing, appealed, assure, assures, assuring, assured, challenge, challenges, challenging, challenged, complain, complains, complaining, complained, consult, consults, consulting, consulted, convince, convinces, convincing, convinced, declare, declares, declaring, declared, demand, demands, demanding, demanded, emphasize, emphasizes, emphasizing, emphasized, emphasise, emphasises, emphasising, emphasised, excuse, excuses, excusing, excused, inform, informs, informing, informed, invite, invites, inviting, invited, persuade, persuades, persuading, persuaded, phone, phones, phoning, phoned, pray, prays, praying, prayed, promise, promises, promising, promised, question, questions, questioning, questioned, recommend, recommends, recommending, recommended, remark, remarks, remarking, remarked, respond, responds, responding, responded, specify*	Finite verbs	**Biber 2006, adapted by Author**
Syntax	*that* subordinate clauses (other than relatives) preceded by factive verbs	**ThVFCT**	*That* complement clause (THSC) tag preceded by any of the following words tagged as verb: *concluding, conclude, concluded, concludes, demonstrates, demonstrating, demonstrated, demonstrate, determining, determines, determine, determined, discovered, discovers, discover, discovering, finds, finding, found, find, knows, known, knowing, know, knew, learn, learns, learning, learnt, means, meaning, meant, mean, notifies, notices, notice, noticed, notify, notifying, noticing, notified, observed, observes, observing, observe, proven, prove, proving, proved, proves, realized, realised, realizes, realises, realize, realise, realizing, realising, recognizes, recognises, recognize, recognise, recognized, recognised, recognizing, recognising, remembered, remember, remembers, remembering*	Finite verbs	**Biber 2006, adapted by Author**

Syntax	*that* subordinate clauses (other than relatives) preceded by likelihood verbs	**ThVLIK**	*That* complement clause (THSC) tag preceded by any of the following words tagged as verb: *assumes, assumed, assuming, assume, believe, believing, believes, believed, doubting, doubted, doubts, doubt, gathers, gathering, gathered, gather, guessed, guess, guessing, guesses, hypothesised, hypothesise, hypothesises, hypothesising, hypothesizing, hypothesized, hypothesize, hypothesizes, imagine, imagining, imagines, imagined, predict, predicted, predicting, predicts, presupposing, presupposes, presuppose, presupposed, presumes, presuming, presumed, presume, reckon, reckoning, reckoned, reckons, seemed, seems, seem, seeming, speculated, speculate, speculating, speculates, suppose, supposes*	Finite verbs	**Biber 2006, adapted by Author**
Verb semantics	mental/attitudinal verbs in other contexts	**VATToth er**	Any word in the following list tagged as verb and (1) not followed by a *_WHSC* or *_THSC* tag or *to_* (2) not having an additional tag apart from the usual underscore separated tag *_TAG* : *agreeing, agreed, agree, agrees, anticipates, anticipated, anticipate, anticipating, complain, complained, complaining, complains, conceded, concede, concedes, conceding, ensure, expecting, expect, expects, expected, fears, feared, fear, fearing, feel, feels, feeling, felt, forgetting, forgets, forgotten, forgot, forget, hoped, hope, hopes, hoping, minding, minded, minds, mind, preferred, prefer, preferring, prefers, pretending, pretend, pretended, pretends, requiring,*	Finite verbs	**Biber 2006, adapted by Author**

| Verb semantics | communication verbs in other contexts | **VCOMM other** | Any word in the following list tagged as verb and (1) not followed by a _WHSC_ or _THSC_ tag or to_ (2) not having an additional tag apart from the usual underscore separated tag _TAG_ : *say, says, saying, said, tell, tells, telling, told, call, calls, calling, called, ask, asks, asking, asked, write, writes, writing, wrote, written, talk, talks, talking, talked, speak, speaks, spoke, spoken, speaking, thank, thanks, thanking, thanked, describe, describing, describes, described, claim, claims, claiming, claimed, offer, offers, offering, offered, admit, admits, admitting, admitted, announce, announces, announcing, announced, answer, answers, answering, answered, argue, argues, arguing, argued, deny, denies, denying, denied, discuss, discusses, discussing, discussed, encourage, encourages, encouraging, encouraged, explain, explains, explaining, explained, express, expresses, expressing, expressed, insist, insists, insisting, insisted, mention, mentions, mentioning, mentioned, offer, offers, offering, offered, propose, proposes, proposing, proposed, quote, quotes, quoting, quoted, reply, replies, replying, replied, shout, shouts, shouting, shouted, sign, signs, signing, signed, sing, sings, singing, sang, sung, state, states, stating, stated, teach, teaches, teaching, taught, warn, warns, warning, warned, accuse, accuses, accusing, accused, acknowledge, acknowledges, acknowledging, acknowledged, address, addresses, addressing, addressed, advise, advises, advising, advised, appeal, appeals, appealing, appealed, assure, assures, assuring, assured, challenge, challenges, challenging, challenged, complain, complains, complaining, complained, consult, consults, consulting, consulted, convince, convinces, convincing, convinced, declare, declares, declaring, declared, demand, demands, demanding, demanded, emphasize, emphasizes, emphasizing, emphasized, emphasise, emphasises, emphasising, emphasised, excuse, excuses, excusing, excused, inform, informs, informing, informed, invite, invites, inviting, invited, persuade, persuades, persuading, persuaded, phone, phones, phoning, phoned, pray, prays, praying, prayed, promise, promises, promising, promised, question, questions, questioning, questioned, recommend, recommends, recommending, recommended, remark, remarks, remarking, remarked, respond, responds,* | Finite verbs | **Biber 2006, adapted by Author** |
| Verb semantics | factive verbs in other contexts | **VFCToth er** | Any word in the following list tagged as verb and (1) not followed by a _WHSC_ or _THSC_ tag or to_ (2) not having an additional tag apart from the usual underscore separated tag _TAG_ : *concluding, conclude, concluded, concludes, demonstrates, demonstrating, demonstrated, demonstrate, determining, determines, determine, determined, discovered, discovers, discover, discovering, finds, finding, found, find, knows, known, knowing, know, knew, learn, learns, learning, learnt, means, meaning, meant, mean, notifies, notices, notice, noticed, notify, notifying, noticing, notified, observed, observes, observing, observe, proven, prove, proving, proved, proves, realized, realizes, realize, realizing, recognizes, recognize, recognized, recognizing, realised, realises, realise, realising, recognises, recognise* | Finite verbs | **Biber 2006, adapted by Author** |

Verb semantics	likelihood verbs in other contexts	**VLIKother**	Any word in the following list tagged as verb and (1) not followed by a _WHSC_ or _THSC_ tag or *to_* (2) not having an additional tag apart from the usual underscore separated tag _TAG_ : *assumes, assumed, assuming, assume, believe, believing, believes, believed, doubting, doubted, doubts, doubt, gathers, gathering, gathered, gather, guessed, guess, guessing, guesses, hypothesising, hypothesised, hypothesise, hypothesises, hypothesizing, hypothesized, hypothesize, hypothesizes, imagine, imagining, imagines, imagined, predict, predicted, predicting, predicts, presupposing, presupposes, presuppose, presupposed, presumes, presuming, presumed, presume, reckon, reckoning, reckoned, reckons, seemed, seems, seem, seeming, speculated, speculate, speculating, speculates, suppose, supposes*	Finite verbs	**Biber 2006, adapted by Author**
Syntax	*to* clauses preceded by ability adjectives	**ToJABL**	*To* immediately followed by a verb and preceded by any of the following words tagged as adjective: *anxious, (un)able, careful, determined, eager, eligible, hesitant, inclined, obliged, prepared, ready, reluctant, (un)willing*	Words	**Biber 2006, adapted by Author**
Syntax	*to* clauses preceded by ability adjectives	**ToJCRTN**	*To immediately followed by a verb and* preceded by any of the following words tagged as adjective: *apt, certain, due, guaranteed, liable, likely, prone, unlikely, sure*	Words	**Biber 2006, adapted by Author**
Syntax	*to* clauses preceded by adjectives of ease	**ToJEASE**	To immediately followed by a verb and preceded by any of the following words tagged as adjective: *difficult, easier, easy, hard, (im)possible, tough*	Words	**Biber 2006, adapted by Author**
Syntax	*to* clauses preceded by factive adjectives	**ToJEFCT**	*To* immediately followed by a verb and preceded by any of the following words tagged as adjective: *afraid, ashamed, disappointed, embarrassed, glad, happy, pleased, proud, puzzled, relieved, sorry, surprised, worried*	Words	**Biber 2006, adapted by Author**
Syntax	*to* clauses preceded by evaluative adjectives	**ToJEVAL**	*To* immediately followed by a verb and preceded by any of the following words tagged as adjective: *bad, worse, (in)appropriate, good, better, best, convenient, essential, important, interesting, necessary, nice, reasonable, silly, smart, stupid, surprising, useful, useless, unreasonable, wise, wrong*	Words	**Biber 2006, adapted by Author**

Syntax	*to* clauses preceded by verbs of desire	**ToVDSR**	*To* and split infinitive clauses (cf. SPLIT above for explaination) preceded by any of the following words tagged as verb: *agreeing, agreed, agree, agrees, chooses, chosen, choose, choosing, chose, decide, deciding, decided, decides, hate, hates, hating, hated, hesitated, hesitates, hesitate, hesitating, hoped, hope, hopes, hoping, intended, intend, intending, intends, likes, liked, like, liking, loving, loves, love, loved, means, meaning, meant, mean, needs, need, needing, needed, planning, plan, planned, plans, preferred, prefer, preferring, prefers, prepares,*	Finite verbs	**Biber 2006, adapted by Author**
Syntax	*to* clauses preceded by verbs of effort	**ToVEFRT**	*To* and split infinitive clauses (cf. SPLIT above for explaination) preceded by any of the following words tagged as verb: *allowance, allowing, allowed, allowancing, allow, allowances, allows, allowanced, attempting, attempted, attempts, attempt, enables, enabled, enabling, enable, encourages, encouraging, encouraged, encourage, fails, fail, failing, failed, help, helping, helps, helped, instructs, instructed, instruct, instructing, managing, managed, manage, manages, oblige, obligate, obliged, obligates, obliging, obligating, obliges, obligated, order, ordering, orders, ordered, permitted, permits, permit, permitting, persuaded, persuades, persuade, persuading, prompts, prompting, prompted, prompt,*	Finite verbs	**Biber 2006, adapted by Author**
Syntax	*to* clauses preceded by mental verbs	**ToVMNTL**	*To* and split infinitive clauses (cf. SPLIT above for explaination) preceded by any of the following words tagged as verb: *assumed, assumes, assume, assuming, believing, believes, believe, believed, considered, considers, consider, considering, expecting, expects, expected, expect, find, found, finding, finds, forgetting, forget, forgets, forgot, forgotten, imagine, imagined, imagining, imagines, judge, adjudicates, adjudicate, judges, judged, knowing, knows, known, know, knew, learnt, learning, learns, learn, presumes, presuming, presumed, presume, pretend,*	Finite verbs	**Biber 2006, adapted by Author**
Syntax	*to* clauses preceded by verbs of probability	**ToVPROB**	*To* and split infinitive clauses (cf. SPLIT above for explaination) preceded by any of the following words tagged as verb: *appear, appeared, appears, appearing, happens, happened, happen, happening, seemed, seems, seem, seeming, tending, tends, tended, tend*	Finite verbs	**Biber 2006, adapted by Author**
Syntax	*to* clauses preceded by verbs of speech	**ToVSPCH**	*To* and split infinitive clauses (cf. SPLIT above for explaination) preceded by any of the following words tagged as verb: *asks, ask, asking, asked, claiming, claims, claim, claimed, invite, inviting, invited, invites, promising, promised, promise, promises, reminding, remind, reminded, reminds, requesting, request, requests, requested, saying, say, said, says, teaches, teaching, taught, teach, tell, tells,*	Finite verbs	**Biber 2006, adapted by Author**
Syntax	*WH* subordinate clauses preceded by attitudinal verbs	**WhVATT**	WHSC tag preceded by any of the following words tagged as verb: *agreeing, agreed, agree, agrees, anticipates, anticipated, anticipate, anticipating, complain, complained, complaining, complains, conceded, concede, concedes, conceding, ensure, expecting, expect, expects, expected, fears, feared, fear, fearing, feel, feels, feeling, felt, forgetting, forgets, forgotten, forgot, forget, hoped, hope, hopes, hoping, minding, minded, minds, mind, preferred, prefer, preferring, prefers, pretending, pretend, pretended, pretends, requiring, required, requires, require,*	Finite verbs	**Biber 2006, adapted by Author**

Syntax				Biber 2006, adapted by Author
	WH subordinate clauses preceded by communicative verbs **WhVCOM**	*saying, said, tell, tells, telling, told, call, calls, calling, called, ask, asks, asking, asked, write, writes, writing, wrote, written, talk, talks, talking, talked, speak, speaks, spoke, spoken, speaking, thank, thanks, thanking, thanked, describe, describing, describes, described, claim, claims, claiming, claimed, offer, offers, offering, offered, admit, admits, admitting, admitted, announce, announces, announcing, announced, answer, answers, answering, answered, argue, argues, arguing, argued, deny, denies, denying, denied, discuss, discusses, discussing, discussed, encourage, encourages, encouraging, encouraged, explain, explains, explaining, explained, express, expresses, expressing, expressed, insist, insists, insisting, insisted, mention, mentions, mentioning, mentioned, offer, offers, offering, offered, propose, proposes, proposing, proposed, quote, quotes, quoting, quoted, reply, replies, replying, replied, shout, shouts, shouting, shouted, sign, signs, signing, signed, sing, sings, singing, sang, sung, state, states, stating, stated, teach, teaches, teaching, taught, warn, warns, warning, warned, accuse, accuses, accusing, accused, acknowledge, acknowledges, acknowledging, acknowledged, address, addresses, addressing, addressed, advise, advises, advising, advised, appeal, appeals, appealing, appealed, assure, assures, assuring, assured, challenge, challenges, challenging, challenged, complain, complains, complaining, complained, consult, consults, consulting, consulted, convince, convinces, convincing, convinced, declare, declares, declaring, declared, demand, demands, demanding, demanded, emphasize, emphasizes, emphasizing, emphasized, emphasise, emphasises, emphasising, emphasised, excuse, excuses, excusing, excused, inform, informs, informing, informed, invite, invites, inviting, invited, persuade, persuades, persuading, persuaded, phone, phones, phoning, phoned, pray, prays, praying, prayed, promise, promises, promising, promised, question, questions, questioning, questioned, recommend, recommends, recommending, recommended, remark, remarks, remarking, remarked, respond, responds, responding, responded, specify, specifies, specifying, specified, swear, swears,*	Finite verbs	
Syntax				Biber 2006, adapted by Author
	WH subordinate clauses preceded by factive verbs **WhVFCT**	WHose tag preceded by any of the following words tagged as verb: *concluding, conclude, concluded, concludes, demonstrates, demonstrating, demonstrated, demonstrate, determining, determines, determine, determined, discovered, discovers, discover, discovering, finds, finding, found, find, knows, known, knowing, know, knew, learn, learns, learning, learnt, means, meaning, meant, mean, notifies, notices, notice, noticed, notify, notifying, noticing, notified, observed, observes, observing, observe, proven, prove, proving, proved, proves, realized, realised, realizes, realises, realize, realise, realizing, realising, recognizes, recognises, recognize, recognise, recognized, recognised, recognizing, recognising,*	Finite verbs	

Syntax	*WH* subordinate clauses preceded by likelihood verbs	**WhVLIK**	WhSC tag preceded by any of the following words tagged as verb: *assumes, assumed, assuming, assume, believe, believing, believes, believed, doubting, doubted, doubts, doubt, gathers, gathering, gathered, gather, guessed, guess, guessing, guesses, hypothesised, hypothesise, hypothesises, hypothesising, hypothesizing, hypothesized, hypothesize, hypothesizes, imagine, imagining, imagines, imagined, predict, predicted, predicting, predicts, presupposing, presupposes, presuppose, presupposed, presumes, presuming, presumed, presume, reckon, reckoning, reckoned, reckons, seemed, seems, seem, seeming,*	Finite verbs	**Biber 2006, adapted by Author**
Syntax	*To* clauses preceded by stance nouns	**ToNSTNC**	*To* immediately followed by a verb and preceded by any of the following words tagged as noun: *agreement, agreements, decision, decisions, desire, desires, failure, failures, inclination, inclinations, intention, intentions, obligation, obligations, opportunity, opportunities, plan, plans, promise, promises, proposal, proposals, reluctance, reluctances, responsibility, responsibilities, right, rights,*	Nouns	**Biber 2006, adapted by Author**
Syntax	Prepositions preceded by stance nouns	**PrepNSTNC**	*assertion, assertions, conclusion, conclusions, conviction, convictions, discovery, discoveries, doubt, doubts, fact, facts, knowledge, knowledges, observation, observations, principle, principles, realization, realizations, result, results, statement, statements, assumption, assumptions, belief, beliefs, claim, claims, contention, contentions, feeling, feelings, hypothesis, hypotheses, idea, ideas, implication, implications, impression, impressions, notion, notions, opinion, opinions, possibility, possibilities, presumption, presumptions, suggestion, suggestions, grounds, ground, hope, hopes, reason, reasons, view, views, thought, thoughts, comment, comments, news, news, proposal, proposals, proposition, propositions, remark, remarks, report, reports, requirement, requirements, agreement, agreements, decision, decisions, desire, desires, failure, failures, inclination, inclinations, intention, intentions, obligation, obligations, opportunity,*	Nouns	**Biber 2006, adapted by Author**
Adjective semantics	Attitudinal adjectives without a *that* clause after	**JJATDother**	Any of the following words tagged as adjective and not followed by a *to_* or *_THSC* (*that* complement clause) tag: *afraid, amazed, (un)aware, concerned, disappointed, encouraged, glad, happy, hopeful, pleased, shocked, surprised, worried*	Nouns	**Biber 2006, adapted by Author**
Adjective semantics	Adjectives related to color	**JJCOLR**	Any of the following words tagged as adjective and not having an additional tag apart from the underscore separated _TAG: *black, white, dark, bright, blue, browm, green, grey, gray, red*	Nouns	**Biber 2006, adapted by Author**

Adjective semantics	Epistemic adjectives without a *that* clause after	**JJEPSTother**	Any of the following words tagged as adjective and not followed by a *to_* or *_THSC* (*that* complement clause) tag: *apparent, certain, clear, confident, convinced, correct, evident, false, impossible, inevitable, obvious, positive, right, sure, true, well-known, doubtful, likely, possible, probable, unlikely*	Nouns	**Biber 2006, adapted by Author**
Adjective semantics	Evaluative adjectives	**JJEVAL**	Any of the following words tagged as adjective and not having an additional tag apart from the underscore separated _TAG: *bad, beautiful, best, fine, good, great, lovely, nice, poor*	Nouns	**Biber 2006, adapted by Author**
Adjective semantics	Relational adjectives	**JJREL**	Any of the following words tagged as adjective and not having an additional tag apart from the underscore separated _TAG: *additional, average, chief, complete, different, direct, entire, external, final, following, general, initial, internal, left, main, maximum, necessary, original, particular, previous, primary, public, similar, single, standard, top, various, same*	Nouns	**Biber 2006, adapted by Author**
Adjective semantics	Size related adjectives	**JJSIZE**	Any of the following words tagged as adjective and not having an additional tag apart from the underscore separated _TAG: *big, deep, heavy, huge, long, large, little, short, small, thin, wide, narrow*	Nouns	**Biber 2006, adapted by Author**
Adjective semantics	Time related adjectives	**JJTIME**	Any of the following words tagged as adjective and not having an additional tag apart from the underscore separated _TAG: *annual, daily, early, late, new, old, recent, young, weekly, monthly*	Nouns	**Biber 2006, adapted by Author**
Adjective semantics	Topical adjectives	**JJTOPIC**	Any of the following words tagged as adjective and not having an additional tag apart from the underscore separated _TAG: *chemical, commercial, environmental, human, industrial, legal, medical, mental, official, oral, phonetic, political, sexual, social, ventral, visual*	Nouns	**Biber 2006, adapted by Author**

| Noun semantics | Nouns abstract and process | **NNABSP ROC** | from the underscore separated _*TAG* : *action, actions, activity, activities, application, applications, argument, arguments, development, developments, education, educations, effect, effects, function, functions, method, methods, research, researches, result, results, process, processes, accounting, accountings, achievement, achievements, addition, additions, administration, administrations, approach, approaches, arrangement, arrangements, assignment, assignments, competition, competitions, construction, constructions, consumption, consumptions, contribution, contributions, counseling, counselings, criticism, criticisms, definition, definitions, discrimination, discriminations, description, descriptions, discussion, discussions, distribution, distributions, division, divisions, eruption, eruptions, evolution, evolutions, exchange, exchanges, exercise, exercises, experiment, experiments, explanation, explanations, expression, expressions, formation, formations, generation, generations, graduation, graduations, management, managements, marketing, marketings, marriage, marriages, mechanism, mechanisms, meeting, meetings, operation, operations, orientation, orientations, performance, performances, practice, practices, presentation, presentations, procedure, procedures, production, productions, progress, progresses, reaction, reactions, registration, registrations, regulation, regulations, revolution, revolutions, selection, selections, session, sessions, strategy, strategies, teaching, teachings, technique, techniques, tradition, traditions, training, trainings, transition, transitions, treatment, treatments, trial, trials, act, acts, agreement, agreements, attempt, attempts, attendance, attendances, birth, births, break, breaks, claim, claims, comment, comments, comparison, comparisons, conflict, conflicts, deal, deals, death, deaths, debate, debates, demand, demands, answer, answers, control, controls, flow, flows, service, services, work, works, test, tests, use, uses, war, wars, change, changes, question, questions, study, studies, talk, talks, task, tasks, trade, trades, transfer, transfers, admission, admissions, design, designs, detail, details, dimension, dimensions, direction, directions, disorder, disorders, diversity, diversities, economy, economies, emergency, emergencies, emphasis, emphases,* | Nouns | **Biber 2006, adapted by Author** |

| Noun semantics | Nouns cognitive | NNCOG | ...from the underscore separated _TAG_ : _analysis, analyses, decision, decisions, experience, experiences, assessment, assessments, calculation, calculations, conclusion, conclusions, consequence, consequences, consideration, considerations, evaluation, evaluations, examination, examinations, expectation, expectations, observation, observations, recognition, recognitions, relation, relations, understanding, understandings, hypothesis, hypotheses, ability, abilities, assumption, assumptions, attention, attentions, attitude, attitudes, belief, beliefs, concentration, concentrations, concern, concerns, consciousness, consciousnesses, concept, concepts, fact, facts, idea, ideas, knowledge, knowledges, look, looks, need, needs, reason, reasons, sense, senses, view, views, theory, theories, desire, desires, emotion, emotions, feeling, feelings, judgement, judgements, memory,_ | Nouns | **Biber 2006, adapted by Author** |

| Noun semantics | Nouns concrete | **NNCONC** | from the underscore separated _TAG_ : tank, tanks, stick, sticks, target, targets, strata, stratas, telephone, telephones, string, strings, telescope, telescopes, sugar, sugars, ticket, tickets, syllabus, syllabuses, tip, tips, salt, salts, tissue, tissues, screen, screens, tooth, teeth, sculpture, sculptures, sphere, spheres, seawater, seawaters, spot, spots, ship, ships, steam, steams, silica, silicas, steel, steels, slide, slides, stem, stems, snow, snows, sodium, mud, muds, solid, solids, mushroom, mushrooms, gift, gifts, muscle, muscles, glacier, glaciers, tube, tubes, gun, guns, nail, nails, handbook, handbooks, newspaper, newspapers, handout, handouts, node, nodes, instrument, instruments, notice, notices, knot, knots, novel, novels, lava, lavas, page, pages, food, foods, transcript, transcripts, leg, legs, eye, eyes, lemon, lemons, brain, brains, magazine, magazines, device, devices, magnet, magnets, oak, oaks, manual, manuals, package, packages, marker, markers, peak, peaks, match, matches, pen, pens, metal, metals, pencil, pencils, block, blocks, pie, pies, board, boards, pipe, pipes, heart, hearts, load, loads, paper, papers, transistor, transistors, modem, modems, book, books, mole, moles, case, cases, motor, motors, computer, computers, mound, mounds, dollar, dollars, mouth, mouths, hand, hands, movie, movies, flower, flowers, object, objects, foot, feet, table, tables, frame, frames, water, waters, vessel, vessels, arm, arms, visa, visas, bar, bars, grain, grains, bed, beds, hair, hairs, body, bodies, head, heads, box, boxes, ice, ices, car, cars, item, items, card, cards, journal, journals, chain, chains, key, keys, chair, chairs, window, windows, vehicle, vehicles, leaf, leaves, copy, copies, machine, machines, document, documents, mail, mails, door, doors, map, maps, dot, dots, phone, phones, drug, drugs, picture, pictures, truck, trucks, piece, pieces, tape, tapes, note, notes, liquid, liquids, wire, wires, equipment, equipments, wood, woods, fiber, fibers, plant, plants, fig, figs, resistor, resistors, film, films, sand, sands, file, files, score, scores, seat, seats, belt, belts, sediment, sediments, boat, boats, seed, seeds, bone, bones, soil, soils, bubble, bubbles, bud, buds, water, waters, bulb, bulbs, portrait, portraits, bulletin, bulletins, step, steps, shell, shells, stone, stones, cake, cakes, tree, trees, camera, cameras, video, videos, face, faces, wall, walls, acid, acids, alcohol, alcohols, cap, caps, aluminium, aluminiums, clay, | Nouns | **Biber 2006, adapted by Author** |
| Noun semantics | Nouns group | **NNGRP** | Any of the following words tagged as noun and not having an additional tag apart from the underscore separated _TAG_ : airline, airlines, institute, institutes, colony, colonies, bank, banks, flight, flights, church, churches, hotel, hotels, firm, firms, hospital, hospitals, household, households, college, colleges, institution, institutions, house, houses, lab, labs, laboratory, laboratories, community, communities, company, companies, government, governments, university, | Nouns | **Biber 2006, adapted by Author** |

| Noun semantics | Nouns human | NNHUM AN | from the underscore separated _TAG : *family, families, guy, guys, individual, individuals, kid, kids, man, men, manager, managers, member, members, parent, parents, teacher, teachers, child, children, people, peoples, person, people, student, students, woman, women, animal, animals, applicant, applicants, author, authors, baby, babies, boy, boys, client, clients, consumer, consumers, critic, critics, customer, customers, doctor, doctors, employee, employees, employer, employers, father, fathers, female, females, friend, friends, girl, girls, god, gods, historian, historians, husband, husbands, American, Americans, Indian, Indians, instructor, instructors, king, kings, leader, leaders, male, males, mother, mothers, owner, owners, president, presidents, professor, professors, researcher, researchers, scholar, scholars, speaker, speakers, species, species, supplier, suppliers, undergraduate, undergraduates, user, users, wife, wives, worker, workers, writer, writers, accountant, accountants, adult, adults, adviser, advisers, agent, agents, aide, aides, ancestor, ancestors, anthropologist, anthropologists, archaeologist, archaeologists, artist, artists, artiste, artistes, assistant, assistants, associate, associates, attorney, attorneys, audience, audiences, auditor, auditors, bachelor, bachelors, bird, birds, boss, bosses, brother, brothers, buddha, buddhas, buyer, buyers, candidate, candidates, cat, cats, citizen, citizens, colleague, colleagues, collector, collectors, competitor, competitors, counselor, counselors, daughter, daughters, deer, deer, defendant, defendants, designer, designers, developer, developers, director, directors, driver, drivers, economist, economists, engineer, engineers, executive, executives, expert, experts, farmer, farmers, feminist, feminists, freshman, freshmen, eologist, eologists, hero, heroes, host, hosts, hunter, hunters, immigrant, immigrants, infant, infants, investor, investors, jew, jews, judge, judges, lady, ladies, lawyer, lawyers, learner, learners, listener, listeners, maker, makers, manufacturer, manufacturers, miller, millers, minister, ministers, mom, moms, monitor, monitors, monkey, monkeys, neighbor, neighbors, observer, observers, officer, officers, official, officials, participant,* | Nouns | **Biber 2006, additions and adapted by Author** |

Noun semantics	Nouns place	**NNPLACE**	from the underscore separated _TAG : apartment, apartments, interior, interiors, bathroom, bathrooms, moon, moons, bay, bays, museum, museums, bench, benches, neighborhood, neighborhoods, bookstore, bookstores, opposite, opposites, border, borders, orbit, orbits, cave, caves, orbital, orbitals, continent, continents, outside, outsides, delta, deltas, parallel, parallels, desert, deserts, passage, passages, estuary, estuaries, pool, pools, factory, factories, prison, prisons, farm, farms, restaurant, restaurants, forest, forests, sector, sectors, habitat, habitats, shaft, shafts, hell, hells, shop, shops, hemisphere, hemispheres, southwest, hill, hills, station, stations, hole, holes, territory, territories, horizon, horizons, road, roads, bottom, bottoms, store, stores, boundary, boundaries, stream, streams, building, buildings, top, tops, campus, campuses, valley, valleys, canyon, canyons, village, villages, coast, coasts, city, cities, county, counties, country, countries, court, courts, earth, earths, front, fronts, environment, environments, district, districts, field, fields, floor, floors, market, markets, lake, lakes, office, offices, land, lands, organization, organizations, lecture, lectures,	**Biber 2006, adapted by Author**	Nouns
Noun semantics	Nouns quantity	**NNQUANT**	Any of the following words tagged as noun and not having an additional tag apart from the underscore separated _TAG : cycle, cycles, rate, rates, date, dates, second, seconds, frequency, frequencies, section, sections, future, futures, semester, semesters, half, halves, temperature, temperatures, height, heights, today, todays, number, numbers, amount, amounts, week, weeks, age, ages, day, days, century, centuries, part, parts, energy, energies, lot, lots, heat, heats, term, terms, hour, hours, time, times, month, months, mile, miles, period, periods, moment, moments, morning, mornings, volume, volumes, per, weekend, weekends, percentage, percentages, weight, weights, portion, portions, minute,	**Biber 2006, adapted by Author**	Nouns
Noun semantics	Nouns technical	**NNTECH**	from the underscore separated _TAG : cell, cells, unit, units, gene, genes, wave, waves, ion, ions, bacteria, bacterias, electron, electrons, chromosome, chromosomes, element, elements, cloud, clouds, sample, samples, isotope, isotopes, schedule, schedules, neuron, neurons, software, softwares, nuclei, nucleus, solution, solutions, nucleus, nuclei, atom, atoms, ray, rays, margin, margins, virus, viruses, mark, marks, hydrogen, hydrogens, mineral, minerals, internet, internets, molecule, molecules, mineral, minerals, organism, organisms, message, messages, oxygen, oxygens, paragraph, paragraphs, particle, particles, sentence, sentences, play, plays, star, stars, poem, poems, thesis, theses, proton, protons, unit, units, web, webs, layer, layers, center, centers, matter, matters, chapter, chapters, square, squares, data, circle, circles, equation, equations, compound, compounds, exam, exams, letter, letters, bill, bills, page, pages,	**Biber 2006, adapted by Author**	Nouns

					Biber 1988, Nini 2014 adapted by Author
Noun semantics	Nominalizations	**NOMZ**	Nouns ending at *tion(s), ment(s), ness, nesses, ity, ities* and having a length of greater than 5 characters and not having an additional tag aprt from the *_NN* tag.	Nouns	
Noun semantics	Proper nouns	**NNP**	Nouns tagged as NNP and NNPS by Stanford Tagger.	Nouns	**Le Foll and Author**
Noun semantics	Stance nouns without prepositions	**NSTNCother**	tag: *assertion, assertions, conclusion, conclusions, conviction, convictions, discovery, discoveries, doubt, doubts, fact, facts, knowledge, knowledges, observation, observations, principle, principles, realization, realizations, result, results, statement, statements, assumption, assumptions, belief, beliefs, claim, claims, contention, contentions, feeling, feelings, hypothesis, hypotheses, idea, ideas, implication, implications, impression, impressions, notion, notions, opinion, opinions, possibility, possibilities, presumption, presumptions, suggestion, suggestions, grounds, ground, hope, hopes, reason, reasons, view, views, thought, thoughts, comment, comments, news, news, proposal, proposals, proposition, propositions, remark, remarks, report, reports, requirement, requirements, agreement, agreements, decision, decisions, desire, desires, failure, failures, inclination, inclinations, intention, intentions, obligation, obligations, opportunity,*	Nouns	**Biber 2006, adapted by Author**
Adverb semantics	Attitudinal adverbs	**RATT**	Any of the following words tagged as adverb and not having an additional tag apart from the underscore separated *_TAG* : *amazingly, astonishingly, conveniently, curiously, hopefully, fortunately, importantly, ironically, rightly, sadly, surprisingly, unfortunately, even worse*	Words	**Biber 2006, adapted by Author**
Adverb semantics	Factive adverbs	**RFACT**	Any of the following words tagged as adverb and not having an additional tag apart from the underscore separated *_TAG* : *actually, always, certainly, definitely, indeed, inevitably, never, obviously, really, undoubtedly, nodoubt, ofcourse, infact, of course, in fact, without doubt, no doubt*	Words	**Biber 2006, adapted by Author**
Adverb semantics	Adverbs of likelihood	**RLIKELY**	Any of the following words tagged as adverb and not having an additional tag apart from the underscore separated *_TAG* : *apparently, evidently, perhaps, possibly, predictably, probably, roughly, maybe, in most cases*	Words	**Biber 2006, adapted by Author**

Adverb semantics	Non factive adverbs	**RNONFACT**	Any of the following words tagged as adverb and not having an additional tag apart from the underscore separated _*TAG* : *confidentially, frankly, generally, honestly, mainly, technically, truthfully, typically, reportedly, primarily, usually*	Words	**Biber 2006, adapted by Author**
		Composite tags	Important: The following features consist of aggregates of individual tags. To avoid redundant correlations, undue skew etc. in the data, they should not be used in combination with the respective individual tags. E.g., either use PP1 counts or the		
Pronouns	All 1st person referents	**FPPAll**	Any word tagged with PP1P, PP1S	Finite verbs	**Biber 1988, adapted by Author**
Pronouns	All 3rd person referents	**PP3**	Any word tagged with PP3P, PP3S	Finite verbs	**Biber 1988, adapted by Author**
Verb semantics	All modals of possibility	**MDPOSSCAll**	Any word tagged as MDCA, MDCO, MDMM	Finite verbs	**Author**
Verb semantics	All modals of prediction	**MDPREDAll**	Any word tagged as MDWS, MDWO, GTO	Finite verbs	**Author**
Verb features	All *be* and *get* passives	**PASSAll**	Any word tagged as PASS, PGET	Finite verbs	**Author**
Adverb semantics	All adverbs related to stance	**RSTNCAll**	Any word tagged as RATT, RNONFACT, RFACT, RLIKELY	Words	**Biber 2006, adapted by Author**
Syntax	All that complement clauses preceded by stance adjectives, nouns, and verbs	**ThJSTNCAll**	Any word tagged as ThJATT, ThJFCT, ThJLIK, ThJEVL	Finite verbs	**Biber 2006, adapted by Author**

Syntax	All *that* complement clauses preceded by stance nouns	**ThNSTNCAll**	Any word tagged as ThNNFCT, ThNATT, ThNFCT, ThNLIK	Finite verbs	**Biber 2006, adapted by Author**
Syntax	All *that* complement clauses preceded by stance adjectives, nouns, and verbs	**ThSTNCAll**	Any word tagged as ThVCOMM, ThVATT, ThVFCT, ThVLIK, ThJATT, ThJFCT, ThJLIK, ThJEVL, ThNNFCT, ThNATT, ThNFCT, ThNLIK	Words	**Biber 2006, adapted by Author**
Syntax	All *that* complement clauses preceded by stance verbs	**ThVSTNCAll**	Any word tagged as ThVCOMM, ThVATT, ThVFCT, ThVLIK	Finite verbs	**Biber 2006, adapted by Author**
Syntax	All *to* infinitive clauses preceded by stance adjectives	**ToJSTNCAll**	Any word tagged as ToJCRTN, ToJABL, ToJEFCT, ToJEASE, ToJEVAL	Words	**Biber 2006, adapted by Author**
Syntax	All *to* infinitive clauses preceded by stance adjectives, nouns, and verbs	**ToSTNCAll**	Any word tagged as ToVDSR, ToVEFRT, ToVPROB, ToVSPCH, ToVMNTL, ToJCRTN, ToJABL, ToJEFCT, ToJEASE, ToJEVAL, ToNSTNC	Words	**Biber 2006, adapted by Author**
Syntax	All *to* infinitive clauses preceded by stance verbs	**ToVSTNCAll**	Any word tagged as ToVDSR, ToVEFRT, ToVPROB, ToVSPCH, ToVMNTL	Finite verbs	**Biber 2006, adapted by Author**

					Biber 2006, adapted by Author
Syntax	All *WH* complement clauses preceded by stance verbs	**WhVSTNCAll**	Any word tagged as WhVATT, WhVFCT, WhVLIK, WhVCOM	Finite verbs	
Prepositions	Prepositions having no additional tag	**INother**	Any word tagged as _IN_ and having no additional semantic tag separated by a space. Use this tag when extended features are enabled, to avoid overlap with IN.	Words	**Author**
Adjectives	Attributive adjectives having no additional tag	**JJATother**	Any word tagged as _JJAT_ and having no additional semantic tag separated by a space. Use this tag when extended features are enabled, to avoid overlap with JJAT.	Nouns	**Author**
Adjectives	Predicative adjectives having no additional tag	**JJPRother**	Any word tagged as _JJPR_ and having no additional semantic tag separated by a space. Use this tag when extended features are enabled, to avoid overlap with JJPR.	Finite verbs	**Author**
Nouns	Nouns having no additional tag	**NNother**	Any word tagged as _NN_ and having no additional semantic tag separated by a space. Use this tag when extended features are enabled, to avoid overlap with NN.	Words	**Author**
Adverbs	Adverbs having no additional tag	**RBother**	Any word tagged as _RB_ and having no additional semantic tag separated by a space. Use this tag when extended features are enabled, to avoid overlap with RB.	Words	**Author**
Discourse organization	*that* relative clauses not preceded by a stance noun	**THRCother**	Any word tagged as _THRC_ and having no additional semantic tag separated by a space. Use this tag when extended features are enabled, to avoid overlap with THRC.	Finite verbs	**Author**
Discourse organization	*that* complement clauses not preceded by a stance adjective or verb	**THSCother**	Any word tagged as _THSC_ and having no additional semantic tag separated by a space. Use this tag when extended features are enabled, to avoid overlap with THSC.	Finite verbs	**Author**

| Discourse organization | *WH* complement clauses not preceded by a stance verb | **WHSCother** | Any word tagged as _*WHSC* and having no additional semantic tag separated by a space. Use this tag when extended features are enabled, to avoid overlap with WHSC. | Finite verbs | **Author** |