Tables

1.1	The contents of the Brown Corpus	4
1.2	English-language corpora	11
1.3	BYU English corpora	16
1.4	ICE corpora completed by 2017	17
2.1	The fifty most frequent words in the BNC	28
2.2	The fifty most frequent lemmas in the BNC	30
2.3	The fifty most frequent noun lemmas in the BNC	31
2.4	The fifty most frequent verb lemmas in the BNC	33
2.5	The fifty most frequent adjective lemmas in the BNC	35
2.6	Word forms in the BNC	36
2.7	Frequency of *but* and *however* in COCA, 1990–1994, compared with 2005–2007	38
2.8	Observed and expected frequency of *but* and *however* in COCA, 1990–1994, compared with 2005–2007	39
2.9	Chi square critical values of significance	40
2.10	Distribution of a selection of adjectives over subcorpora in COCA	41
2.11	Distribution of *out* and *out of* in spoken and written British material	41
2.12	Frequency of noun-*to*-noun patterns with identical body-part nouns	42
2.13	The C5 tagset for the CLAWS tagger	46
3.1	Frequency of *maybe* and *perhaps* in COCA, 1990–2015	62
3.2	Frequency of *may* and *may be* in COCA, 1990–2015	63
3.3	Corpora used in the investigation of *fatigue*	64
3.4	Frequency of *fatigue* compounds in *The Independent* and *The New York Times*, 1990–2000	66
3.5	Vague pronouns: per 1,000 words	68
3.6	Top ten positive keywords for six characters in *Romeo and Juliet*	69
4.1	Verb collocates of *dog* in the BNC: frequency	75

4.2	Verb collocates of *dog* in the BNC: MI score	75
4.3	Noun collocates of *dog* in the BNC: frequency	76
4.4	Collocates of *educational* in the BNC (window −5 to +5): by absolute frequency	77
4.5	Collocates of *educational* in the BNC (window −5 to +5): by MI score	78
4.6	Collocates of *educational* in the BNC (window −5 to +5): by z-score	78
4.7	Collocates of *educational* in the BNC (window −5 to +5): by log-likelihood	79
4.8	Right collocates of *go to* in the BNC	80
4.9	Three-word combinations ending in *bed* in the BNC	81
4.10	Left collocates of the noun *hand* in the BNC	82
4.11	Items described as being *at hand*, *in hand*, *on hand* or *to hand*	83
4.12	Abstract items referred to by *at hand*, *in hand*, *on hand* or *to hand*	86
5.1	Conventionalised actions described by formulaic 6-grams containing *hand* in the BNC	91
5.2	Variations in the idiom *storm in a teacup*	94
5.3	Canonical forms of idioms in global Englishes	95
5.4	Some common types of single clause elements with a length of three words or more in the LLC	97
5.5	Some frequent incomplete phrases with a length of three words or more in the LLC	98
5.6	Right collocates of some frequent incomplete phrases in the BNC	99
5.7	The most frequent 5-grams in the BNC	100
5.8	The most frequent tag sequences (POS-grams) in the BNC	102
5.9	*Good* + common noun in the BNC	103
5.10	*On X's toes* in the BNC	106
5.11	The 8-grams in Dickens that occur five times or more	107
6.1	Examples of conceptual metaphors	115
6.2	Part of speech of conventional animal metaphors	117
6.3	Types of metonymy	118
6.4	Subject complements of *his/her/my mind was a/an X* in COCA	122
6.5	Lexical verbs most frequently occurring immediately after *his mind* in COCA	123
6.6	Metaphorical mappings for the concept FEAR posited in the literature and the frequency of these in a sample of 1,000 tokens of *fear* in the BNC	124

6.7 Additional FEAR metaphors found via metaphorical pattern analysis 124
6.8 Source domains for metaphors in the US Inaugural Corpus 125
6.9 Cognitive metaphors based on a number of source domains 126
6.10 Literal and metaphorical uses of *build* in a selection of 100 tokens from the Time Corpus, 2000s 127
6.11 Literal and metaphorical uses of *cathedral* in a selection of 100 tokens from the Time Corpus, 1990–2007 128
6.12 Lexical verbs most frequently occurring immediately after *White House* in COCA 131
6.13 Lexical verbs most frequently occurring immediately after *Washington* in COCA 132
6.14 Lexical verbs most frequently occurring immediately after *Downing Street* in the British subcorpus of GloWbE 133
7.1 Proportion of *whom* out of all *whom* + *who* in COHA, 1820s–2000s 139
7.2 *Get*-passives in LOB, FLOB, Brown and Frown 140
7.3 Frequency indices for selected verbs used in the *get*-passive in the spoken components of the BNC and COCA, 1990–2015 142
7.4 The form of sentential complements of *accustomed to* in a number of subcorpora in the Bank of English 144
7.5 The form of sentential complements of *accustomed* in the BNC and COCA, 1990–2015 144
7.6 Prepositional and directly linked gerunds with the verbs *find* and *get* dependent on and immediately following *he has difficulty/difficulties* on Google, 27 September 2007 146
7.7 The most frequent verbs in the construction *difficulty* (*in*) **ing* in COCA, 1990–2015 146
7.8 The most frequent verbs in the construction *difficulty* (*in*) **ing* in the BNC 147
7.9 The most frequent verbs in the construction *difficulties* (*in*) **ing* in COCA, 1990–2015 148
7.10 The most frequent verbs in the construction *difficulties* (*in*) **ing* in the BNC 148
7.11 Frequency of *facing difficulty/difficulties* (*in*) **ing* in COCA and the BNC 150
7.12 Prepositional and directly linked gerunds with a selection of verbs dependent on and immediately following *[have] difficulty/difficulties* in COCA, 1990–2015 151

7.13 Prepositional and directly linked gerunds with a selection of verbs dependent on and immediately following *[have] difficulty/difficulties* in the BNC 152
7.14 Active and passive clauses in six text categories in ICE-GB 153
8.1 Frequency of *firemen* in the COHA Corpus in the last 100 years 157
8.2 Frequency of *firefighters* in the COHA Corpus in the last 100 years 158
8.3 Frequency of *fire fighters* in the COHA Corpus in the last 100 years 158
8.4 Frequency of *firemen* and *fire fighters/firefighters* in the COHA Corpus in the last 100 years 159
8.5 Adjacent right collocates of *fire* in the COHA Corpus in the last 100 years 159
8.6 Frequency of *poet* and *poetess* in the Time Corpus 160
8.7 Frequency of *actor* and *actress* in the Time Corpus 161
8.8 Ratio of *actor* and *actress* in the Time Corpus 161
8.9 The most frequent adjectives associated with *man* and *woman* in the frame *a* + adjective + *man/woman* in the BNC 162
8.10 Frequency of *man*, *woman*, *men* and *women* in the BNC 163
8.11 Frequency of *single man*, *single woman*, *single men* and *single women* in the BNC 164
8.12 Number of *single man/men* and *single woman/women* per 10,000 mentions of *man/men* and *woman/women* in the BNC 164
8.13 Most frequent verbs in the frames *she* + verb + *him* and *he* + verb + *her* in the BNC 165
8.14 Male and female authors' use of colour terms in the BNC 166
8.15 Frequency of hedging *kind of/sort of* in different disciplines in MICASE 168
8.16 Distribution of *you know* by function and by sex of speaker in a corpus of spoken New Zealand English 169
8.17 *Kind of* in the spoken components of the BNC 170
8.18 *Sort of* in the spoken components of the BNC 170
9.1 Frequency of adverbial *likely* in different contexts in COCA and the BNC 177
9.2 Distribution of *will likely* in *The Independent*, 1990–2005 178
9.3 The noun *side* with animate and inanimate referents in the ME part of the HC 182

9.4 *Beside*(*s*) in the Early Modern English part of the HC 183
9.5 *Beside* and *besides* in the ARCHER Corpus 184
9.6 The *OED* quotation base: chronological breakdown 189
9.7 Third person singular indicative suffix *-s* vs. *-th* 192
10.1 Prepositions following *different* 197
10.2 Regional distribution of singular and plural agreement with *my*/*your*/*his*/*her*/*our*/*their* + *family* + *is*/*are* 202
10.3 Standard and non-standard past tense forms in FRED 207
10.4 Standard and non-standard past tense forms on the web 208
10.5 The ten most frequent 4-grams in the blogs and student theses components of the Corpus of English in Sweden (SWE-CE) 215