

Contents

Acknowledgements v

Preface xi

1 Introduction 1

- 1.1 Introduction: the basics 1
- 1.2 What is a corpus and how can we use it? 1
- 1.3 Which corpus, what for and what size? 3
- 1.4 How to make a basic corpus 5
- 1.5 Basic corpus linguistic techniques 8
- 1.6 Lexico-grammatical profiles 14
- 1.7 How have corpora been used? 17
- 1.8 How have corpora influenced language teaching? 21
- 1.9 Issues and debates in the use of corpora in language teaching 25

2 Establishing basic and advanced levels in vocabulary learning 31

- 2.1 Introduction 31
- 2.2 Frequency and native-speaker vocabulary size 31
- 2.3 The most frequent words and the core vocabulary 33
- 2.4 The broad categories of a basic vocabulary 37
- 2.5 Chunks at the basic level 46
- 2.6 The basic level: conclusion 46
- 2.7 The advanced level 47
- 2.8 Targets 48
- 2.9 The vocabulary curve 49
- 2.10 The 6,000 to 10,000 word band 50
- 2.11 Meanings and connotations 53
- 2.12 Breadth and depth 54

3 Lessons from the analysis of chunks 58

- 3.1 Introduction 58
- 3.2 The single word 58
- 3.3 Collocation 59
- 3.4 Strings of words in corpora 60
- 3.5 Phraseology and idiomaticity 62
- 3.6 Looking at corpus data 64
- 3.7 Interpreting the data: chunks and single words 69
- 3.8 Chunks and units of interaction 70
- 3.9 Conclusions and implications 75

4	Idioms in everyday use and in language teaching	80
4.1	Introduction	80
4.2	Finding and classifying idioms	82
4.3	Frequency	84
4.4	Meaning	86
4.5	Functions of idioms	87
4.6	Idioms in specialised contexts	90
4.7	Idioms in teaching and learning	94
5	Grammar and lexis and patterns	100
5.1	Introduction	100
5.2	The example of <i>border</i>	102
5.3	Grammar rules and patterns: deterministic and probabilistic	104
5.4	The <i>get</i> -passive: an extended case study	106
5.5	Previous studies of the <i>get</i> -passive	106
5.6	<i>Get</i> -passives and related forms	108
5.7	Core <i>get</i> -passive constructions in the CANCODE sub-corpus	109
5.8	Discussion	113
5.9	Grammar as structure and grammar as probabilities: the example of ellipsis	114
5.10	Conclusions and implications	115
6	Grammar, discourse and pragmatics	120
6.1	Introduction	120
6.2	Non-restrictive <i>which</i> -clauses	120
6.3	Previous studies of <i>which</i> -clauses	122
6.4	Concordance analysis of <i>which</i> -clauses	122
6.5	<i>If</i> -clauses	127
6.6	<i>Wh</i> -cleft clauses	130
6.7	Bringing the insights together	136
6.8	Corpus grammar and pedagogy	137
7	Listenership and response	140
7.1	Introduction	140
7.2	Forms of listenership	142
7.3	Response tokens across varieties of English	145
7.4	Functions of response tokens	148
7.5	Conclusions and implications	155
8	Relational language	159
8.1	Introduction	159
8.2	Conversational routines	163
8.3	Small talk	168
8.4	Discourse markers	171
8.5	Hedging	174

8.6 Vagueness and approximation 176

8.7 Conclusions and implications 181

9 Language and creativity: creating relationships 184

9.1 Introduction 184

9.2 Spoken language and creativity 184

9.3 Corpora and creativity 188

9.4 Creative speakers 190

9.5 Applications to pedagogy 191

9.6 Corpus to pedagogy: creating relationships 192

9.7 SUEs and creativity 192

9.8 Quantitative and qualitative 196

9.9 Conclusions 197

10 Specialising: academic and business corpora 198

10.1 Introduction 198

10.2 Written academic English 198

10.3 Written academic English: examples of frequency 200

10.4 Spoken academic corpora 203

10.5 Spoken academic English, conversation and spoken business English 204

10.6 The CANBEC business corpus 206

10.7 Chunks 210

10.8 *Problem* and its institutional construction in CANBEC 214

10.9 Summary 216

10.10 Pedagogical implications 216

11 Exploring teacher corpora 220

11.1 Introduction 220

11.2 Classroom discourse 222

11.3 Frameworks for the analysis of classroom language 222

11.4 Applying the frameworks to a corpus of classroom data 229

11.5 Looking at questioning in the classroom 233

11.6 Teacher corpora in professional development 240

11.7 Conclusions and considerations 243

Coda 246

References 249

Appendix 1 284

Appendix 2 297

Appendix 3 301

Author index 305

Subject index 310

Publisher's acknowledgements 314