Contents

List of Figures xii
List of Tables xiv
Acknowledgements xviii

1 Introduction 1
   1.1 Corpus Linguistics 1
   1.2 About this Book 3
   1.3 Terminology 5
   1.4 Commonly Used Resources 9
   1.5 Corpus Linguistics: A Personal View 10

2 Designing a Corpus 18
   2.1 A Historical Perspective on Corpus Design 18
   2.2 Types of Corpus 21
   2.3 Issues in Designing or Selecting a Corpus 30
   2.4 Corpus Source Material and Issues 38
   2.5 Conclusion 43
   Appendix: List of Some Commonly Used Corpora of English 44

3 Learning from a Corpus: Finding Pattern in Concordance Lines 47
   3.1 Introduction 47
   3.2 Phraseology and Meaning in Concordance Lines 51
   3.3 ‘Hidden Meaning’ in Concordance Lines 55
   3.4 Types of Patterning 60
   3.5 Conclusion 79

4 Foundational Quantitative Concepts in Corpus Linguistics 81
   4.1 Introduction 81
   4.2 Frequency, Relative Frequency and Comparisons 82
   4.3 Frequency and Normalisation 84
   4.4 Keywords 88
   4.5 Measuring Collocation 93
# Table of Contents

4.6 Lexical Bundles 102
4.7 Multidimensional Analysis 105
4.8 Semantic and Other Annotation 111
4.9 Conclusion 114

5 Exploring Corpora through Quantitative Measures 116
5.1 Introduction 116
5.2 More about Keyness 117
5.3 Collostructions 121
5.4 Clustering Texts 126
5.5 Clustering Words 135
5.6 Conclusion 140

6 The Language Learner and the Corpus 143
6.1 Introduction 143
6.2 Reference Materials for Learners: Dictionaries and Grammars 143
6.3 Learner Corpora 152
6.4 Language Learners Using Corpora 174
6.5 Conclusion 184

7 Corpora and the Analysis of Discourse 186
7.1 Introduction 186
7.2 Academic English: Comparing Disciplines 187
7.3 Corpus Linguistics and News Discourse: Being Critical 210
7.4 Stylistics 217
7.5 Further Applications of Corpus Studies of Discourse 221
7.6 Conclusion 224

8 Further Applications of Corpus Research 226
8.1 Introduction 226
8.2 Applications of Bilingual Corpora 226
8.3 Corpora in Forensic Linguistics 233
8.4 Information and Opinion Extraction 241
8.5 Identifying Fake News 245
8.6 Sociolinguistics: The Study of Variation in Language 247
8.7 Conclusion 253

9 Corpora and Language Theory 255
9.1 Introduction 255
9.2 Sinclair and Language Theory 256
9.3 The Mental Modelling of Language 258
## Table of Contents

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>9.4 Language as Social Semiotic</td>
<td>268</td>
</tr>
<tr>
<td>9.5 Conclusion</td>
<td>286</td>
</tr>
<tr>
<td>10 Conclusion</td>
<td>288</td>
</tr>
<tr>
<td>10.1 First Catch Your Corpus</td>
<td>289</td>
</tr>
<tr>
<td>10.2 Investigation Methods: Top-down</td>
<td>292</td>
</tr>
<tr>
<td>10.3 Investigation Methods: Bottom-up</td>
<td>295</td>
</tr>
<tr>
<td>10.4 And the Outcome Is?</td>
<td>303</td>
</tr>
<tr>
<td>10.5 Conclusion</td>
<td>304</td>
</tr>
<tr>
<td>References</td>
<td>307</td>
</tr>
<tr>
<td>Index</td>
<td>334</td>
</tr>
</tbody>
</table>