Lynne Flowerdew, *Corpora and Language Education*. 2012. Basingstoke: Palgrave Macmillan. ISBN 9781403998927 (HB), 9781403998934 (PB). 368 pp.

Reviewed by Paul Pauwels (KU Leuven/Thomas More)

This is a big book – with a slightly misleading title. What the reader gets is an introduction to corpus linguistics (CL) in its various guises and its various applications, and a wealth of references to sources and materials. It is part of the *Research and Practice in Applied Linguistics* series (eds. Candlin and Hall), which includes titles on diverse applied linguistics topics like vocabulary, community interpreting, business discourse, literature in language education and so on, and which focuses on “exploring the relationship between research and practice in applied linguistics” (o.c., xiv). The overall framework of this book, starting with key concepts and ending with an elaborate resource section, is a hallmark of the series, as are the ‘boxes’ with definitions of concepts, examples, and quotes by key practitioners. In this book, these boxes are used a bit too often, and their contents are often too lengthy, which makes the book less, rather than more, readable and accessible. It also leads to sections where the reader has to shift from text to box and back far too often. Section 6.4 is a case in point: in the span of 12 pages, there are no fewer than 15 boxes, and pages 156-157, for example, start with the second half of a box continued from the previous page, followed by 5 lines of text, a boxed quote, a new subtitle with 8 lines of text, a boxed concept, three lines of text, and the first half of a boxed example. The feeling is that the author wants to let the original sources speak for themselves too much of the time.

The book is divided into four major parts, starting with Key Concepts and Approaches. The first chapter provides a clear introduction, defining corpus, providing a critical assessment of different possible uses of corpora, and introducing the different types of applications. There are clear definitions (in boxes) of concepts like lemmatization, type/token, semantic prosody, and colligation. At the same time, there is also a full page concept box on “implications of semantic prosody and semantic preference for professional discourses” – which is clearly an aside and considerably slows down reading. The chapter ends with a survey of the limitations of the corpus linguistic approach, and an introduction to its major applications. The second chapter presents a brief historical overview, and situates CL as a reaction to Chomskyan linguistics. This allows Flowerdew to provide a view of the advantages of a corpus linguistic approach using a number of oppositions – competence vs performance, intuition vs attestation, grammaticality vs acceptability, creativity vs formulaicity – which she does with a clear sense of nuance. The third chapter provides an overview of the main approaches to CL. Five approaches are distinguished: the Neo-Firthian, the probabilistic, the systemic-functional, the multidimensional and the sociolinguistic. Some of these approaches are described as closely related to specific corpus projects, such as Sinclair’s Cobuild (probabilistic), Biber’s Longman Grammar Project (multidimensional) and the CANCODE project of the Nottingham School (sociolinguistic). The different views on the nature and role of corpora and the key concepts of these approaches are clearly outlined. A summary at the end of the chapter – maybe in table form, showing the main differences between the approaches (bottom up vs top down, text vs samples, etc.), would not have gone amiss, though.

In part two of the book, corpus linguistics is situated with respect to textlinguistics and sociolinguistics. Looking at the literature on discourse studies (Jaworsky and Coupland 1999; Bhatia et al. 2008), this division seems a bit arbitrary, and also in the book it does not seem to have been completely resolved, as witnessed by the discrepancy between the title of part 2 ('The nexus of corpus linguistics, text linguistics and sociolinguistics'), the title of chapter four ('How is corpus linguistics related to discourse analysis?') and the page headings in that chapter (corpus linguistics and textlinguistics). The chapter starts off by discussing the status of corpus linguistics, and concludes that it is best seen as a methodology rather than as a theory. The remainder of the chapter is a survey of corpus-based approaches to written, spoken and multimodal discourse analysis. The structure of the chapter, nesting the different discourse analytical approaches inside the major subdivision into written, spoken and multimodal has resulted in a very fragmented survey where the focus, at times, shifts to the explanation of different types of discourse analysis. A survey starting from different types of approaches (genre, rhetoric, etc.) or foci (on text features, text structure, the relation between text and context, etc.) might have been more transparent, and might have led to fewer (sub)sections. As it stands, some of the distinctions, like the 'problem-solution based approach' (4.3.2) or 'linguistic devices with discourse functions' (4.3.3), fit awkwardly with 'genre-based approaches' (4.3.1) or 'critical discourse-based approach' (4.3.4). Chapter five (on sociolinguistics) has a much clearer structure, since it only distinguishes two main approaches (interactional sociolinguistics and variationism). The chapter ends on a critical note, discussing the limitations of corpus-based approaches to sociolinguistics. On the whole, part two seems slightly out of balance, in so far as chapter 4 covers much more ground than chapter 5; dividing the subject matter in a different way over the chapters might have increased readability.

In part three, the perspective is changed, and the focus is on different topics of research (rather than approaches) and applications in teaching. Chapter six surveys corpus based research on English as a lingua franca, language use in business and healthcare, forensic linguistics, stylistics, translation studies, learner corpora and second language acquisition, lexicography, and testing. It should be clear that the chapter again covers a lot of ground, and that the treatment of the different areas can only be exemplary. There are occasional gaps, such as the failure to mention the Europarl-corpus when discussing parallel corpora in translation studies. It is also striking to see that some, but not all, of the subsections end with an assessment and an indication of future directions of research. Chapter seven discusses the different ways in which corpora have been, and can be, used to inform teaching. The chapter starts with a cautionary note discussing the pedagogical relevance of corpora, and then moves on to discuss indirect applications (through reference and teaching materials) and direct data-driven learning. A caveat voiced in the discussion of legal discourse, viz. that "[it] is such an intricate discourse area [that] corpus based methodologies may not completely align themselves with legal writing tasks" (o.c., 202), implying that writing instruction cannot solely be based on corpus materials, could equally well have been voiced for other genres. As becomes clear from the section on the "potential impediments to DDL [data-driven learning]" we are still far from a situation where learners have the tools and skills to comfortably and confidently use corpora for learning. The structure of the remainder of the chapter is again a little muddled, covering a wide range of topics (7.4 underrepresented corpora, 7.5 corpora in teaching translation, 7.6 corpora in teacher

education, 7.7 corpora in teaching literary analysis). At this stage, the addition of 22 example boxes may have succeeded in providing more flavour, but at the same time it did not help readability. Chapter eight presents ten examples of corpus-based research projects, outlining context, aims, methodology and results, and providing a brief commentary and suggestions for further research. The projects range from large-scale phraseology research in comparable native and non-native corpora (LINDSEI and LOCNEC) to small-scale CDA-oriented studies like investigating evaluative styles. Throughout the survey, a number of corpus tools (Wordsmith, SketchEngine, ConcGram) are introduced. The final three research examples are of a different type, in that they investigate the effectiveness of the use of corpora or corpus-based materials for teaching purposes.

Part four of the book consists of 70-plus pages of reference materials. Chapter nine introduces a selection of 'Key Sources' (books, journals, associations, websites and mailing lists; 8 pages), then there are 46 pages of references proper, followed by a glossary, an author index and a subject index.

In sum, it is probably fair to say that the author took on a very big challenge, especially because so many different types of corpora are used in so many different contexts, and with so many different purposes. Although I cannot say that I have a much clearer picture of corpus linguistics after reading this book – it is far too sprawling for that – I have gained a clearer insight into many of the key issues, I have been given food for thought, and I know I have a useful resource for my teaching.

References

Bhatia, Vijay K., John Flowerdew and Rodney H. Jones (eds.) 2008. *Advances in Discourse Studies*. London and New York: Routledge.

Jaworsky, Adam and Nikolas Coupland (eds.) 1999. *The Discourse Reader*. London and New York: Routledge.