Book reviews

From Corpus to Classroom. *Language use and language teaching*. Anne O'Keeffe, Michael McCarthy and Ronald Carter. Cambridge University Press, 2007. 315 Pages. ISBN-13: 978-0-521-61686-7 (pbk). ISBN-13: 978-0-521-85146-6 (hbk). Price: £19.50 (pbk), £50.00 (hbk).

Drawing on their knowledge and expertise in the domain of corpus linguistics and pedagogy, the authors of this brand new volume have taken up the highly challenging task of explaining and illustrating the importance of mediation between research findings in corpus linguistics and classroom pedagogy. In the introduction to the book, they refer to the 'frequent mismatch between CL research and what goes into materials and resources, and what goes on in the language classroom' (p.xi). To prevent this mismatch, the authors have decided to explain and demonstrate how corpora can inform teachers and teaching in the classroom rather than tell teachers what to teach or how to teach. The book draws primarily on spoken language corpora, which constitutes a welcome bias given the numerous calls for more focus on speech in educational environments, especially in instructed settings where lack of exposure to speech often turns out to be detrimental to the learners' communicative competence.

The book is divided into eleven sections. The introductory section addresses a number of key issues in corpus linguistics, such as corpus building, types of corpora, corpus tools and analyses. The next three sections focus on vocabulary. Section two deals with single words, frequency aspects, depth and breadth of vocabulary knowledge, and links between vocabulary, frequency and levels of competence. Section three shifts from a single word to a multiword perspective and tells us not only how corpora can be used to reveal patterns in the language, but also how these patterns can be interpreted. Section four complements the previous one and focuses on the role of idioms: their retrieval, classification, and role in teaching and learning, more particularly from an EFL perspective. Section five expands further on the patterned nature of language and presents the grammar and lexis interface from a Sinclairian perspective. Section six addresses variability issues in the language and does so by focussing on three structures (i.e. non-restrictive which clauses, if clauses and wh- cleft clauses) and by demonstrating how grammatical patterns are related to specific contexts of use. Sections seven to nine deal with interactional issues: section seven focuses on listener response tokens (such as 'yeah' used to acknowledge incoming talk, or 'mm' used as a continuer)

and reveals the importance of such tokens in spoken interaction; section eight, entitled 'relational language', addresses issues such as conversational routines, small talk, hedging, vagueness and approximation in language; and section nine deals with the role of creativity and suggests ways in which creativity (beyond the normative aspects often associated with it) can be used to create links from the corpus to the classroom. The aim of section ten, entitled 'specialising: academic and business corpora', is to argue that smaller corpora can be used to reveal lexis and structures representative of a specific genre. As for the final section of the book, it looks at what corpora can tell us about our own teaching by analysing corpora of classroom interactions and teacher talk. In this section, the authors focus on the local context and demonstrate the potential of such corpora "as a tool for reflective practice within pre-service teacher education and (…) career development" (p.220). The book also contains a number of appendices, one of which provides an excellent overview of existing corpora (a note on whether or not these corpora are freely available might however have been added to inform the readers), and an impressive bibliography.

Throughout the book, the authors never fail to adopt the teachers' perspective: they provide clear examples or case studies to illustrate their point and regularly suggest options for pedagogical applications. One aspect which could however still be slightly improved is the heading of some sections in the illustrations. I will refer here to one specific example on p.138 where the authors illustrate the case study of wh- cleft clauses by showing how the results can be implemented in the grammar sections of textbooks. The heading of the grammar section is: 'What clauses and long noun phrases as subjects'. Although the content of the section clearly reflects the links between context of use, functions of language and grammatical choices, a heading such as 'Introducing important information' might be more appealing to learners and, perhaps more importantly, stress from the start the fact that structural choices as strategic acts have important interpersonal consequences, as stated by the authors themselves on p.138.

Whilst, according to he authors, no prior knowledge or experience in corpus linguistics is required to read the book, I would suggest that computer literacy and a relatively good linguistic background are necessary to fully appreciate the content of the book. As for readership, although both pre- and in-service teachers are targeted, I believe that the book will perhaps be most useful in teacher education programmes when pre-service teachers can still benefit both from support from lecturers and from the computational environment usually at their disposal (the authors themselves, at the end of the book, stress the importance of teachers' needs in terms of training, level of expertise and wider access to corpora).

O'Keefe, McCarthy and Carter's book is a must for teachers as it will definitely contribute to a better understanding of what corpora can bring to language pedagogy. The authors also stress the importance of what teachers can bring to corpus studies and repeatedly argue that feedback from teachers on applications of corpus-based material is necessary. At the end of the book, the authors argue that on one level they have tried to show the application and importance of corpus findings for language teaching but, on another level, have sought to raise teachers' interest in using language corpora themselves to pursue their own enquiries and enhance their professional development. I am confident that the *From Corpus to Classroom* volume will play a major role in stimulating the synergy between corpora and language teaching.

Fanny Meunier, Centre for English Corpus Linguistics, UCL, Belgium.

Corpora in the Foreign Language Classroom, Encarnación Hidalgo, Luis Quereda and Juan Santana (eds.) Rodopi, 2007. xiv + 362 Pages. ISBN: 978-90-420-2142-6. Price: €76.

What do George Bush, a pile of manure, and a porcupine called Karl have in common? Answer: they are all in this volume of selected papers from the sixth *Teaching and Language Corpora (TaLC) conference* held in Granada in 2004, ten years after the first *TaLC* conference. In the intervening decade, corpora and the tools for using them have become much more widely available, to the extent that there are now no major material obstacles facing those who wish to use corpora. So how far have we actually come, in ten years, in using corpora in language teaching? This question is directly addressed in the first paper in this volume (3-16), where Angela Chambers gives an overview of publications devoted to learner use of corpora since the early 1990s. Most of the studies are qualitative, focusing on evaluation of the activities and on students' reactions to them. Learners appear to appreciate the authenticity of corpus data, the abundance of examples and the autonomy afforded by corpus-based work. On the negative side, some learners report that they found working with corpora difficult, time-consuming and tedious. Chambers concludes that corpus use has a positive contribution to make, but that there remain obstacles to its spread into contexts beyond higher education and to independent use of corpora outside the classroom.

This is a crucial point, and many of the papers collected here are concerned, in one way or another, with the transfer between corpus-based research and language learning. Stig Johansson (17-28) discusses the advantages of using bilingual and learner corpora for examining points of convergence and divergence between students' L1 and L2, and between non-native and native-speaker usage. Johansson's main point is that there is a continuity between learning and research, since language learning involves a process of hypothesis forming and testing. He remarks however that there is little mention of the role of corpora in works on second language acquisition. This is partly because using corpora is seen as primarily a pedagogical concern, but there are indeed ways in which corpora are of interest to SLA, notably in relation to frequency effects and the acquisition of formulaic language.

Five other papers in the volume are devoted to accounts of classroom use of corpora. Stephan Coffey (161–173) proposes awareness-raising activities around lexical sets, Sara Gesuato (175–190) looks at the use of corpus data to help learners to distinguish between near-synonyms, and Julia Lavid (237–252) shows how two large corpora, the British National Corpus and the Corpus del Español, can be used for students to investigate the different transitivity patterns of English and Spanish. The two remaining papers, by Andy Cresswell (267–287) and Alejandro Curado Fuentes (309–326), not only describe the pedagogical use of corpus material (for teaching connector usage and English for tourism, respectively), but also try to evaluate its effect by comparing their learners' results with those of a control group following more conventional instruction. Both authors report positive effects for the corpus-based learning groups, but in different ways. For Cresswell's students, there were qualitative effects on metalinguistic ability, but no significant difference in actual use of connectors. Curado Fuentes, on the other hand, reports consistently better results on the post-test for his experimental group. This is encouraging, but there is a danger inherent in this kind of study, described by

Cobb (1997: 308) and cited by Chambers in the first paper of this volume, namely that it is extremely difficult to keep the learning conditions of experimental and control groups clearly distinct. Consequently, it is not always easy to know exactly what is being measured and how.

Returning to Johansson's observation that there is little mention of corpora in SLA publications, if one searches for the combined terms "corpora" and "SLA" with an Internet search engine, the vast majority of hits refer to the use of learner corpora in investigating second language acquisition. This is a thriving area of research, and is somewhat under-represented in the present volume. Agnieszka Leko-Szymaska (253–266) examines the use of past progressive forms in the written production of Polish-speaking learners of English. This is a detailed study testing the predictions of the Aspect Hypothesis (Andersen & Shirai, 1994); the conclusion is that this framework needs to be refined to take into account other factors, such as cross-linguistic influence, particularly in the case of two languages such as English and Polish, which encode aspect in quite different ways.

Xiaotian Guo's paper (91–104) provides an original study of naturally-occurring learner language, based on a small longitudinal corpus of email messages written by a young Chinese speaker to her English-speaking friends. Guo traces the appearance and gradual replacement of three types of non-target usage, such as verbal use of adjectival items (e.g. make sure you don't late). His analysis, inspired by Ellis's variable competence model (1994:364–5) aims to show that these forms should not be taken as "errors", but as a normal part of learning. This is not a new idea, to be sure, but Guo's joint use of a small longitudinal corpus and a larger cross-sectional corpus is an interesting way to examine systematic and non-systematic variation in L2 use. Szilvia Papp (207–220) also uses a Chinese learner corpus, but for a pedagogical purpose, by coupling it with a reference corpus to encourage learners to monitor their production. The paper is an outline for research into the viability of corpora and concordancing in the classroom rather than a report of results, but Papp gives a useful overview of existing corpora and research related to Chinese-speaking learners of English.

One important way of spreading corpus use outside a small circle of corpus-nerds is to make corpora pedagogically relevant and as accessible as possible. Sabine Braun (31-46) describes the rationale behind the ELISA English interview corpus, which will already be familiar to readers of *ReCALL* (see Braun, 2005). Kiyomi Chujo *et al*. (47-69) discuss indices (readability scores, average word length, etc.) for measuring the accessibility of English and Japanese text samples, and Carmen Pérez Basanta and María Elena Rodríguez Martín (141–158) describe a customised corpus of film transcripts for teaching conversation skills. As Braun puts it, a corpus is like a "text museum"; without preparation and background knowledge, visitors will not derive maximum benefit from the exhibits. Hence the need for pedagogical mediation to integrate corpora into a wider learning environment. Even with relatively advanced learners, this kind of pedagogical integration may be beneficial. Nele Olivier *et al*. (221–235) describe a unit from COLLEX, a corpus-based learning environment designed to facilitate the passage between proficiency work and research. The unit in question covers the grammaticalization of size nouns as quantifiers (e.g. a foot-high pile of manure vs. heaps of fun). The general principle is that by developing their ability to learn and recognise patterns, students also become better able to describe patterns

themselves and so can enrich the environment with their own contributions, thus participating in a virtuous circle as part of a learner-researcher community.

Another link in the relationship between research and teaching is the way in which language research informs the content of textbooks. There is now a considerable body of research to support the view that phraseology plays an important role both in L1 and L2. But to what extent have these research findings influenced the content of EFL textbooks? Fanny Meunier and Céline Gouverneur (119–139) analyse a corpus of textbook material for advanced EFL learners, to see how the textbooks in question incorporate formulaic aspects of language. Their conclusion is that ELT editors are aware of the importance of phraseology, but that this is not given sufficient salience in the material, and that there is a certain cacophony in the metalanguage used to refer to phraseological phenomena.

Tailor-made corpora can be used for analyses which are not based on specific classroom practice, but which have pedagogical implications, particularly in the teaching of English for Specific or Academic Purposes. Josta van Rij-Heyligers (105–118) argues for a corpus of World English for Academic Purposes, Peter Lam (71-89) analyses some keywords and their collocations in tourism industry texts, and David Minugh (191–205), starting from George Bush's "crusade" against terrorism, presents an example of how corpora can be used to examine historical or literary metaphors in present-day English. Chris Tribble's paper (289–308) looks at the distribution of certain features accompanying keywords in professional correspondence, such as use of a zero complementizer with the verb hope (e.g. I hope this is not the end of the world vs. I hope that the event goes well), a prime example of the kind of choice that may well strike the L2 learner as arbitrary. Tribble's results indicate that use of zero vs. that varies according to the relationship between reader and writer – a finding that is concordant with the idea that complementizers such as I hope […] have come to function as a chunk, a discourse formula rather than a true subordinator, in contrast to I hope that[…] which continues to mark the subordinate status of the following clause (for a discussion, see Tagliamonte & Smith, 2005). This is a very nice example of a corpus-based micro-analysis, using an easily accessible methodology to shed light on a seemingly arbitrary phenomenon, and whose results can be translated into guidelines for learners. Advice based on evidence, as Tribble observes, is better than advice based on intuition.

Linguistic intuition, argues Bill Louw in the last paper in this volume (329–362), does not necessarily lead us in the right direction. It is important for corpus linguists to accept the implications of John Sinclair's plea to "trust the text" (Sinclair, 2004:23). Trusting the text, says Louw, means allowing corpus evidence to shake our faith in intuition, introspection and established models. One could say that this has profound implications for the use of corpora in language teaching, were it not Louw's observation that the phrase "profound implications" is frequently used by those who wish to prevent institutional change. Corpora have the potential to be not just a useful tool for learning, but also to change the way that teachers and learners look at and work with language. For the moment, the area of change is circumscribed, and this volume perhaps does not have the sense of excitement and innovation of earlier work around corpora and language teaching. While there are no startling new insights, the volume shows a diversifying range of corpus applications in teaching and the study of learner language. TaLC conferences bring together people whose first interest may be in computational

linguistics, grammar, SLA or other fields, which complicates the editors' task of giving cohesion to a necessarily disparate collection of papers. There are, however, a number of unifying themes running through this volume, notably the interplay between evidence and analysis and between research and learning. The price of the book may discourage many individual purchasers, but this is a volume that all those interested in language learning and corpora should be able to consult. And if you want to know where the porcupine comes in, you will have to read it all the way to the end.

References

Andersen, R. and Shirai, Y. (1994) Discourse motivation for some cognitive acquisition principles. *Studies in Second Language Acquisition*, **16**(2): 133-156.

Braun, S. (2005) From pedagogically relevant corpora to authentic language learning contents. *ReCALL*, **17**(1): 47-64.

Cobb, T. (1997) Is there any measurable learning from hands on concordancing? *System,* **25**(3): 301-315.

Ellis, R. (1994) *The study of Second Language Acquisition*. Oxford: Oxford University Press.

Sinclair, J. (2004) *Trust the text: language, corpus and discourse*. London: Routledge.

Tagliamonte, S. and Smith, J. (2005) No momentary fancy! The zero 'complementizer' in English dialects. *English Language and Linguistics,* **9**(2): 289-309.

John Osborne
Université de Savoie, Chambéry