

Aijmer, Karin (ed.). **Corpora and Language Teaching**. Amsterdam/
Philadelphia: John Benjamins, 2009. 232 pp. ISBN 978 90 272 2307 4. Price
(hardback): £ 84.

Given the growing interest in corpus-based language learning and teaching in recent years, it is only natural to find that there has been a corresponding increase in the amount of scholarly work and publication on this topic (cf. e.g. Römer 2008). One exponent of this trend is the book *Corpora and Language Teaching* (2009), edited by Karin Aijmer. This 232-page volume is based on a selection of papers presented at a symposium on the role of corpora and corpus-linguistic techniques in language teaching organized at the University of Gothenburg in 2005, and has later been complemented by further papers solicited from leading scholars with an interest in this field. Apart from the introduction, the book contains eleven separate papers divided into four sections: “Corpora and second-language acquisition” (pp. 13–44), “The direct corpus approach” (pp. 47–98), “The indirect corpus approach” (101–154) and “New types of corpora” (pp. 157–230).

In her introduction to the volume, the editor gives an account of the current relationship between corpus-based research and pedagogical practice. In particular, she notes that while corpus linguistics is generally seen as having a great deal to offer the teaching profession, the use of corpora in the EFL classroom still seems to be a fairly rare and unexploited phenomenon. This is a situation which may have several different causes: for example, teachers may know too little about the existence of corpora, lack the skill to use them, or be unused to their inductive methodology, or, equally possible, corpus linguists may have failed to reach teachers, to respond properly to their pedagogical needs, or to supply a workable interface between the two disciplines. Given this situation, the complex mission of the book emerges, namely (i) to unravel the mysteries of corpus application in a pedagogical perspective, (ii) to provide practical examples of how corpora can be put to use in the language classroom, (iii) to clarify the differential perspectives of language teaching typically maintained by corpus linguists and EFL teachers, and (iv) to account for the use of corpora in applied linguistics research, notably by analysing learner corpora in order to chart the different forms of (inter)language produced by advanced learners.

The first section of the book, “Corpora and second-language acquisition”, accommodates two papers which are concerned with the significance of corpus data in processes of second language acquisition (SLA). In her opening paper, Sylviane Granger gives a survey of how learner corpus research has been influenced by neighbouring disciplines – corpus linguistics, linguistic theory, second language acquisition and foreign language teaching – all of which are seen as important for optimal exploitation of authentic language data. She discusses

several controversial issues in this context, for example how research on such corpora can join forces with SLA, how practitioners within this field can profit from the use of corpus data, and how such data relate to native speaker norms. She then turns to pedagogical applications of learner corpora, notably by discussing the distinction between corpora for immediate (teacher-oriented) use and corpora for delayed (researcher-oriented) use. This applied treatment also includes an assessment of learner-corpus-driven pedagogical tools as a function of different types of proficiency parameters, notably accuracy, complexity and fluency.

Stig Johansson continues the discussion by considering the empirical foundation of using corpora in second language acquisition. Based on a review of comparative evidence from previous studies on direct and indirect teaching models, he takes up the cudgels for performing controlled studies of the effects of different pedagogical approaches in general, and of corpus-based teaching methods in particular. Since earlier findings seem to suggest that the average language learner is engaged in both hypothesis formation and testing, there is just a short step, he claims, to the use of corpus data in current language teaching. According to this view, the usage-based model is clearly of relevance to SLA, and different uses of corpora may well serve its purpose, especially those relating to textbook corpora, learner corpora and multilingual corpora. Yet, sounding a note of caution, the author points out that there is still a need for more systematic studies on the effectiveness of corpora in the classroom, and that their use is only vindicated to the extent that it corresponds to our knowledge of language and its learning processes.

Featuring three papers, the second section of the book, “The direct corpus approach”, is concerned with more practical matters, in particular how corpora can be used in the classroom and how learners can assume the role of language researchers. In the first paper, Solveig Granath deals with the application of corpora in English grammar and proficiency courses. Arguing for the general pedagogical relevance of such data, she shows how they can be used practically to create language exercises, demonstrate grammatical variation, and throw light on patterns of synonymy and collocation. In this process, she also gives an overview of students’ attitudes towards corpus-based methodologies, and evaluates the pros and cons of introducing them in the EFL classroom. Yet, while concluding that the positive effects of corpus tools tend to prevail, she notes that many teachers are still reluctant to employ them in their teaching, apparently because they lack sufficient skills to do so. It is argued, accordingly, that learning how to use corpora needs to be part of future teacher training, all in order to make corpus data just as natural for both teachers and learners to consult as conventional dictionaries and grammar books.

Signe Oksefjell Ebeling takes the corpus-based methodology one step further by describing its application in an interactive learning environment, Oslo Interactive English (OIE), recently developed at the University of Oslo. This on-line platform provides a wide spectrum of corpus-driven exercises for undergraduate students of English, elementary as well as advanced, aiming primarily to encourage more flexible learning methods, and to make students reflect critically on authentic language data. The exercises relate to seven different grammatical categories, and can be dealt with in terms of either multiple choice, gap-filling, error correction or open answers. In order to do the exercises properly, it is necessary to consult a pre-defined corpus, specifically a trimmed version of the Longman-Lancaster English Language Corpus consisting of seven million words of running text. After two years of use, evaluation of the OIE environment indicates that it has been beneficial to many language learning initiatives, in particular those on a more advanced level; it has also been popular among the students, although far from all of them have actually done the exercises.

Ute Römer winds up the direct corpus approach by trying to assess the impact corpora have made so far on English language teaching. Given the existence of a number of useful corpora and related tools, she claims that we still do not know to what extent they provide the information that teachers and learners need or are on the lookout for. In order to throw light on this issue, she distributed a questionnaire among English language teachers at secondary schools in Germany, asking them to rate the extent to which teaching materials provided support for their work with grammar and vocabulary, and what their attitude was towards using corpora in the classroom. The results showed that while teachers often felt insecure with regard to issues of proper language usage, they did not consider corpus consultation a major alternative in this context, basically because they were unaware of, or did not master, the potential of this methodology. Thus, the author concludes that, given a greater focus on the interface between research and practice, corpora can actually have a significant impact on language teaching – it only takes some effort from the corpus-linguistic community.

The third section of the book, “The indirect corpus approach”, contains another three papers, and deals mainly with so-called delayed pedagogical research, i.e. studies describing specific types of interlanguage. Jennifer Herriman and Mia Boström Aronsson are first to embark on this approach through their study of non-native learners’ writing taken from the Swedish component of the International Corpus of Learner English (ICLE). Noting that such performance is often felt to be somehow deviant from native writing, they compared how Swedish advanced learners of English and native speakers of English organize information in argumentative essays, specifically with regard to theme selection and thematic variation. The resulting patterns, clearly difficult to pinpoint without large quantities of text, indicate that learners tend to use more interpersonal information than native speakers, and that they often overuse cleft constructions when themat-

tising new information and expressing personal attitude. With such choices made, learners' argumentative writing can be said to exhibit an emphatic and persuasive style which is more similar to native speakers' spoken language, thus showing the need for students to get further into the charmed circle of cultural norms and register variation in the target language.

Hilde Hasselgård follows suit by considering thematic distribution in the argumentative writing of Norwegian students as it appears in the corresponding part of the ICLE corpus. Specifically, she zooms in on the extent to which such students apply Norwegian patterns in their choice of thematic structure, and how this relates to their ability to adhere to relevant stylistic and grammatical norms of English. Whereas learners make rather few thematic errors on the whole, her findings show that they tend to overuse certain features as compared to native speakers of English, for example formal subjects, initial adverbials and thematically marked first person pronouns. These observations indicate that although Norwegian students typically have a good command of grammatical structure in English, they seem to have problems with some discourse conventions in the genres concerned. Again, then, the analysis reveals that many learners tend to use features in their argumentative writing which are more at home in the spoken language of native speakers.

Susan Hunston, finally, takes a more general approach to the usefulness of corpus-based descriptions of English by considering the connection between corpus frequency and phraseology. Using the Bank of English (BoE) as her main source of reference, she argues that some of the most interesting information about word frequency patterns is often difficult to present to learners in a meaningful way, in particular when it comes to relative frequency that is dependent on complex variables. To illustrate her point, she examines a set of multi-word units, where overall frequency is not of primary relevance but rather the strength of collocation between the internal elements. This approach has an interesting potential because it is able to identify fairly long phrases, semantic sequences in the author's terminology, which provide information on what is relatively often said rather than what is regarded as grammatically correct per se. It is concluded that while the significance of such collocational data is probably limited in the language classroom, they pose as relevant, although somewhat intractable, in the context of devising teaching materials, primarily with regard to the inclusion of proper structural features.

The final section of the book, "New types of corpora", comprises three contributions on recent compilations of pedagogically relevant corpora in the fields of economics and ELT. In the first paper, Winnie Cheng outlines a corpus-driven approach to the learning and teaching of phraseology in English for Specific Purposes (ESP). Using text extracts from two corpora in the fields of economics and financial services, she applies the search engine ConcGram to

identify lexical patterns, or so-called concgrams, specific to each of the two text types, what is referred to as their textual aboutness. These data are then compared to the corresponding patterns of their source corpora as well as a reference corpus of general English usage, all in order to establish an initial indication of the aboutness of the texts. It is argued that this process has important implications for ESP learners and teachers in that it raises their language awareness, and provides a means of accessing key phraseologies in discipline-specific discourses.

Fanny Meunier and Céline Gouveneur continue this thread by discussing the relationship between corpora and pedagogical material in current ELT literature. Capitalizing on the fact that many textbooks still tend to shy away from empirical data, they report on their project of compiling a corpus of general-purpose textbook material, the TeMa Corpus, comprising 700 000 words of text collected from popular ELT books on the market. With tagging added for the different types and levels of exercise, the corpus is shown to provide some interesting data, notably in terms of novel search options which can be used to capture, for example, ELT-specific frequency patterns and typical metalanguage usage. The authors conclude that pedagogically annotated corpora of this type might be useful for future learners, teachers and material designers when trying to meet new educational challenges.

Joybrato Mukherjee completes this section by looking into the syntactic features of spontaneous conversation as it appears in the German part of the Louvain International Database of English Interlanguage (LINDSEI). He reports on three case studies of lexicogrammatical forms and structures recorded in the speech of advanced German learners, viz. verb-noun collocations, the discourse marker *you know*, and performance phenomena such as repetitions and contractions. The essence of these studies lies in the fact that they all serve to home in on distinctive properties of the grammar of conversation, showing a more limited use of verb-noun collocations, a generally high incidence of discourse markers, and a wide functional variety of repetitions and other disfluencies, respectively. Based on the outcome of these investigations, a pedagogical discussion is maintained of the possible avenues of teaching and learning spoken grammar in the EFL classroom.

Turning now to some critical remarks, it is clear that the edited volume *Corpora and Language Teaching* makes up an important contribution to the field of corpus-driven language pedagogy. With its broad coverage of topics and genuine efforts at trouble-shooting, it provides a good account of the possibilities and complexities in this field, highlighting both the effects of using corpora and related techniques in the language classroom, and the research potential provided by learner corpora consisting of spoken as well as written material. In the former case, the inclusion of concrete illustration of how authentic corpus data can be used for different learning activities is to be particularly applauded, not least because it offers good examples for practitioners-to-be in this area (cf. the second

section of the book); in the latter case, the same thing goes for the novel approaches to compiled learner corpus data, which give the reader a great deal of new insight into patterns of interlanguage that would otherwise be difficult to discern (cf. the third and fourth sections). As far as research-oriented work is concerned, several papers also bring to the fore the important observation that language pedagogy is now increasingly designing its own corpora to its own criteria (cf. Aston 2000:16).

True to its complex mission, the book also features an interesting theoretical discussion. One aspect of this is the effort to plot learner corpus research on the map of other linguistic disciplines, specifically by describing how it relates to linguistic theory and corpus linguistics proper, as well as the applied areas of second language acquisition and foreign language teaching. While establishing such bearings is important in itself, it is in fact crucial for the further development of corpus-driven language pedagogy, since open interfacing with these disciplines seems to be a prerequisite for full realization of the potential of corpus data, whether in the classroom or elsewhere. This aim is also something which the book contributes significantly to. Another virtue of the discussion is the healthy reminder that the same potential is only real to the extent that it can be supported by controlled studies of corpus-based teaching methods. Clearly, as quoted by Johansson in this volume, “there is thus a need for research to underpin the integration of corpora and concordancing in the language-learning environment” (Chambers 2005:111).

There is also a downside to the present book, however. One such aspect, for example, is its apparent limitation of the discussion to corpora and corpus-related tools from the traditional sphere, i.e. finite, stand-alone collections of authentic data from various texts and text types. Thus, while the teacher-oriented papers included tend to draw on findings from corpora such as the BNC and the Longman/Lancaster Corpus, as well as similar, purpose-built corpora and newspaper CDs, there is virtually no mention or consideration of information collected from web-based corpora and associated applications. This circumscription is clearly unfortunate. As the web forms the largest body of text in existence, in effect a huge, unlimited monitor corpus covering a variety of domains and text types, and being continuously added to and updated, it stands to reason that it has a great deal to offer also in this context, despite occasional problems of search methodology and interpretation (cf. Bergh & Zanchetta 2008). In fact, its wide range of (sometimes unique) linguistic information has the potential of contributing significantly to the work in the EFL classroom, in particular when it comes to providing up-to-date snapshots of current language usage, and highlighting patterns of expression in a particular domain or text type.

Another aspect that deserves critical mentioning is the treatment of the relationship between corpora and pedagogical material. Pace one paper dealing

with textual patterns derived from a corpus of current ELT books, and a few others making brief reference to this field, little is in fact said in the volume about the possibilities of using corpora to inform teaching materials and reference works, and the same situation applies to the design of language teaching syllabi, in particular those emphasizing communicative competence (cf. Römer 2008). Arguably, this is a perspective that would have profited from a more thorough and systematic discussion in the book, all to the purpose of putting the productive capacity of corpus-based resources in a more accurate perspective.

As is often the case with conference proceedings and similar publications, finally, the collection of papers in the book turns out to be a rather mixed bag of topics that do not lend themselves to straightforward formation. This is probably also why the editor's presentation of the different contributions to the volume does not always correspond to the actual organisation of the book: not only are the headlines occasionally different but also the grouping of the papers, a situation which may appear somewhat bewildering to the reader. Another peculiar circumstance is the fact that the punctuation of parts of the text leaves a great deal wanting. This applies, for example, to the 10-page introduction of the volume, where, somewhat surprisingly, there are virtually no commas around (except for a few cases of enumeration), despite the fact that it holds a fair amount of complex sentence structure. Clearly, such a reductive measure does not add to the overall readability of the text.

All in all, however, there is no doubt that the positive impression of the book *Corpora and Language Teaching* prevails. Offering a range of fruitful perspectives on the role and effectiveness of corpora in EFL teaching and in research on learner language, as well as proffering an open problem-solving approach to recognised complications, it will clearly contribute to the ongoing mission of trying to bridge the gap that still persists between corpus-based theory and pedagogical practice. In short, to tune in with Jan Svartvik on the back cover, "this volume will serve as a welcome guide and source of inspiration for language teachers who want a new tool for educational challenges".

Gunnar Bergh

References

- Aston, Guy. 2000. "Corpora and language teaching." In Burnard, Lou & Tony McEnery (eds.) *Rethinking Language Pedagogy from a Corpus Perspective*. Frankfurt: Peter Lang, pp. 7-17.
- Bergh, Gunnar & Eros Zanchetta. 2008. "Web linguistics." In Anke Lüdeling & Merja Kytö (eds.) *Corpus Linguistics. An International Handbook*. Volume 1. Berlin: Walter de Gruyter, pp. 309-327.
- Chambers, Angela. 2005. "Integrating corpus consultation in language studies." *Language Learning and Technology* 9:111-125.
- Römer, Ute. 2008. "Corpora and language teaching." In Anke Lüdeling & Merja Kytö (eds.) *Corpus Linguistics. An International Handbook*. Volume 1. Berlin: Walter de Gruyter, pp. 112-131.