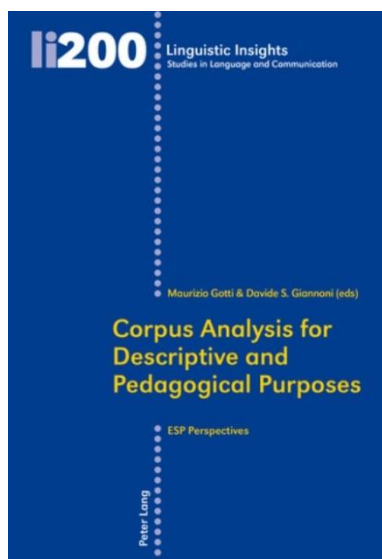# BOOK REVIEW

EXPLORING THE PRACTICAL
APPLICATIONS OF CORPUS
LINGUISTICS

Maurizio Gotti and Davide S. Giannoni (Eds.). CORPUS ANALYSIS FOR DESCRIPTIVE AND PEDAGOGICAL PURPOSES. ESP PERSPECTIVES (2014), Bern: Peter Lang. 432 pp., ISBN 978-3-0343-1516-6 (PB).

Corpus analysis for descriptive and pedagogical purposes. ESP perspectives is the 200th volume in the landmark series "Linguistic Insights" that is so familiar to all of us working in the field of applied linguistics. As the title suggests, this volume deals with the interface between corpus linguistics, language study and language pedagogy, within the wider context of English for specific (and academic) purposes. It is structured in three sections focusing rather loosely on corpus research methodology, corpus-based language description and pedagogical applications, each of which will be discussed below.

The five chapters on corpus research methods present a variety of perspectives on the way that corpus linguistics can be used to gain a deeper understanding of specialised discourse. In the opening chapter, Lynne Flowerdew presents a survey of possible starting points for ESP corpus research, with a useful overview of the units for linguistic analysis that form the basis for bottom-up corpus work. Next, taking a methodological approach which applies corpus tools to investigate genre, Marina Bondi discusses her research into authorial voice in academic and popular history articles, showing how corpus linguistics can be used to reveal the different ways in which the reader is constructed in these two genres. Her exploration of the uses of "you" and "we" provides new insights into the issue

of interpersonality in specialised texts, which has been extensively researched elsewhere (Breeze, Gotti, & Sancho Guinda, 2014). Her corpus data illustrate how readers of academic journals are primarily constructed as partners in the development of coherent arguments, whereas readers of popular history are typically addressed as "people today": the author guides them through the account of events, suggesting ways that they may feel or react, and providing ready-made interpretations. Winnie Cheng's chapter on using concgrams to investigate research article sections explores the kind of two-word co-occurrences found in different article sections. Her results suggest that such co-selections are more likely to be generic in nature in the discussion, literature review and abstract, whereas they are more subject-specific in the other sections of the article. Hilary Nesi's investigation of citation practices in BAWE provides a useful model of the way Corpus Query Language can be used to identify different citation practices. Her results provide further evidence of interdisciplinary differences in this area, while also indicating ways in which student coursework tends to differ from more advanced academic practices. Carmen Pérez-Llantada then closes this section with a comprehensive review of tendencies in analysing academic genres through multilingual corpora. Her chapter negotiates the complex territory between contrastive rhetoric and emergent international norms, taking the view that native-centred discourse practices are likely to be displaced, and raises a number of fundamental questions concerning the attitude that researchers should take to the differences they may find. She concludes by suggesting some ways ahead for research into academic writing in global contexts where native speaker models are increasingly being called into question. In general, this section provides a concise introduction to some interesting approaches to ESP and EAP corpora, which should inspire larger-scale future research.

The second section is devoted to corpus-based language description, covering a wide range of subject matter ranging from law blogs to newsroom discourse. The first chapter, by Shelley Staples and Douglas Biber, looks at stance in nurse-patient interactions, and identifies various interesting features of nurses' language, including a high frequency of prediction, possibility and likelihood, and frequent use of adjective + "that" and adjective + "to". The authors also draw our attention to the way that corpus tools can be used to research the relative asymmetry of nurse-patient interactions when compared with a corpus of more general conversation. Aside from its obvious applications to language training for nurses, this chapter also provides a table of lexico-grammatical features used for stance analysis, similar to those developed by Biber elsewhere (Biber, 2006), which could serve as a starting point for future research in other genres and contexts. Next, Alan Partington's chapter examines how speakers emphasise importance in TED talks. As Partington points out, importance is an under-researched category and worthy of investigation in different genres. Although more attention needs paying to the borderline between importance and necessity, or between importance markers as such and boosters and emphatics, Partington's

chapter provides a stimulating read, and opens up a number of lines for future enquiry. Following this, Giuliana Garzone's chapter looks at the generic integrity of law blogs, centring mainly on whether the feature of individualistic self-expression, thought to be typical of blogs from their origins onwards, has been preserved in this specialised variant. She concludes that this is indeed the case, since law blogs resemble other blogs – and thereby differ from many other areas of legal language – in their frequent self-mentions, and in their use of first and second person pronouns. Begoña Crespo's chapter addresses the complex question of female authorial voice as it might be manifested in scientific research articles. She analyses the prefaces of scientific and scholarly books written by women between 1700 and 1900, finding fairly frequent use of personal pronouns, often combined with "private" verbs used to express thoughts and feelings, and concluding that this points to high author-reader involvement. However, since her analysis only extends to prefaces by women, the question as to how their discourse differs from that of their male counterparts remains unanswered. The same issue arises in the next chapter, which sets out from the same macro-corpus of female scientific writing. Here, Isabel Moskowich and Leida Maria Monaco contribute further to our understanding of women writers by searching for their use of abstraction, this time focusing on whole texts from different disciplinary areas. Basing their enquiry on Biber's (1988) observation that an abstract style is often instantiated through use of conjuncts, agentless and *by* passives, and adverbial subordination, they find that astronomy is written in a more abstract style than either life sciences or history in this particular corpus. After this, in a very different subject area, Roberta Facchinetti describes how corpus tools were used to build a bilingual glossary of newsroom language (English and Italian), starting from a list of headwords which were then researched in SketchEngine to find information about collocates, grammatical and syntactic behaviour, and so on, and to obtain illustrative examples. This chapter nicely illustrates the practical uses of corpora in contemporary professional contexts where standard dictionaries prove inadequate.

240

The last two chapters in this section provide evidence for positive cross-fertilisation between quantitative and qualitative approaches to analysing discourse. First, Rita Salvi explains how small corpora can be used to analyse institutional discourse concerning the financial crisis. Focusing on speeches by José Manuel Barroso and Mario Draghi, she uses keywords and semantic searches to establish "aboutness", and draws on frequent n-grams to profile the way words such as "crisis" are used in this context. Her analysis moves into the challenging area of corpus-informed discourse analysis, tackling issues such as "cultural keywords" and interpersonality. In a similar vein, Jane H. Johnson then examines the concept of "risk" in the context of news reporting on the family in the British press in the 1990s and 2000s, contrasted with academic sociological discourse from the second period. In the earlier news corpus, family-associated risk is found to have been associated mainly with medical and economic threats, whereas in the

later news corpus and the academic corpus, interpersonal risks such as violence or victimisation are prominent. Moreover, she also notes that the frequency of "risk" declines from the earlier to the later news corpus, whereas "threat" increases, which suggests that there may have been changes in the way the media present danger to their readership. These two chapters contribute to the small but growing volume of research that combines classic discourse analysis with quantitative corpus data in stimulating and innovative ways.

The final section on pedagogical applications opens with a more theoretical chapter by Averil Coxhead, then focuses mainly on different applications for developing students' academic writing. Coxhead's chapter reviews some of the advantages of working with ESP corpora and sets out some of the challenges that still exist. Cassi L. Liardét then examines use of grammatical metaphor in a learner corpus of Chinese students' written academic English, describing the specific features of their writing in this respect and suggesting how greater awareness of grammatical metaphor could be promoted in future pedagogic designs for academic writing courses. Josef Schmied's chapter turns to cultural and disciplinary differences in academic writing, centring on MA theses written by South African students. Turo Hiltunen and Martti Mäkinen revisit the use of formulaic expressions in non-native academic writing in the areas of business and economics, comparing it with published articles in these fields. Their research is interesting in its use of the Academic Formulas List (Simpson-Vlach & Ellis, 2010) to track the frequency of formulas in texts by different groups of writers, showing that both published articles and student writing contain around 10 core formulas per 1000 words. This approach opens up a promising field for research into formulaic language in different disciplines and genres. The final chapter, by Gillian Mansfield, tackles the use of corpora in the classroom, which she sees as bridging the gap between knowledge transmission and skill acquisition. In her experience, hands-on experience with corpora can prove extremely fruitful with some students, and she provides a description of several original and creative ways in which students can be introduced to corpus tools.

In general, this volume is another important contribution to the body of research on corpus linguistics and its intersection with specialised language. Highlights of the book include its description of promising new corpora, such as the Chinese Longitudinal Learners Corpus (Liardét) or the Hong Kong Corpus of Research Articles (Cheng), and its integrated bibliography (presented at the end, rather than after each individual chapter). Moreover, the concise but clear methodological descriptions provided by writers such as Nesi or Staples and Biber, on topics of importance like citation or stance, lend themselves to direct application in other areas of specialised discourse. As the editors themselves conclude (p. 21), the natural partnership between corpus linguistics and ESP/EAP, which are bonded through a common concern in the empirical dimension of clearly delineated samples of language, is likely to prove increasingly fruitful over the years to come. This volume serves as evidence of the cross-fertilisation between

241

these two sub-disciplines, but more importantly, also outlines some of the current debates and provides a number of empirical models that will inspire future research into specialised languages.

[Review submitted 3 Aug 2015]

Reviewed by **RUTH BREEZE**
Instituto Cultura y Sociedad
University of Navarra
Spain
rbreeze@unav.es

References

Biber, D. (1988). *Variation across speech and writing*. Cambridge: Cambridge University Press.

Biber, D. (2006). *University language: A corpus-based study of spoken and written registers*. Amsterdam: John Benjamins.

Breeze, R., Gotti, M., & Sancho Guinda, C. (2014). *Interpersonality in legal genres*. Bern: Peter Lang.

Simpson-Vlach, R., & Ellis, N. (2010). An academic formulas list: New methods in phraseology research. *Applied Linguistics*, *31*, 487-512.