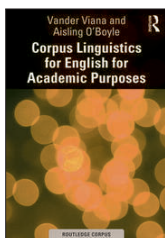# Book Reviews

Corpus Linguistics for English for Academic Purposes.

Vander Viana & Aisling O'Boyle (Eds.)
London/New York: Routledge, 2022. 260 pages. ISBN 978-1-138-74208-6.

Studies central to English for academic purposes (EAP) have highlighted methodological innovations over the past decade. In particular, the development of corpus linguistics (CL) has provided researchers in this area with novel approaches to look at, or reconsider, the core issues in EAP that were addressed based on limited data, thus breathing new life into the field of English language education. In corpus-based/corpus-driven EAP studies, language data plays an integral part in attesting and expounding theoretical frameworks/models, in which corpora are viewed as 'assets' possessing academic value instead of merely some sort of resources. In this respect, data processing and data management appear to be the prerequisites for methodological innovations in EAP studies, revealing a paradigmatic shift from 'knowledge-based exploration' towards scientific discovery (Szalay & Gray, 2006; Hey et al., 2009). It should be noted, nonetheless, that scholarship pertaining to CL-EAP discussions is commonly seen in academic journals, thus showing a lack of monographs, or edited volumes, which are specifically designed to bring together such methodological innovations from a holistic perspective.

As researchers with more than 20 years of experience in EAP, ESP (English for Specific Purposes) and EFL (English as a Foreign Language) teaching in New Zealand and China, we believe that Viana and O'Boyle's recent work entitled *Corpus Linguistics for English for Academic Purposes* can be seen as a much-welcomed publication that attempts to address the above issue. The book is of particular interest to audiences such as novice language researchers, teachers, educators and students of all tertiary levels, because it elucidates for readers a panoramic review of the key notions, practices and research applications of corpus approaches for the EAP community, and it offers, with real-world examples, hands-on data processing techniques to

those wishing to combine quantitative and qualitative analyses for their EAP studies.

Funded by the British Academy under Vander Viana's Skills Innovator Award, the book under review contributes to presenting theoretical underpinnings of CL, and foregrounds the interplay between corpus investigations and EAP research by mining the 'treasure-trove' of empirical data. With regard to its contents, the book consists of ten chapters, and exhibits a variety of academic merits by motivating readers to address challenges in the changing educational landscapes. First of all, a range of fundamental concepts, such as corpus, context and genre/register, are neatly identified and clarified with an extensive account of their connotations and expansions, and therefore researchers can have a clear clue of how to manage terminology appropriately in their studies. For instance, when discussing EAP in its context (Chapter 2), Viana and O'Boyle delineate the idiosyncrasy of academic discourse in terms of epistemology and taxonomy (spoken, written and computer-mediated), emphasising that it represents a highly 'controlled' language type which is associated with the way language is used in particular contexts to disseminate ideas and knowledge, to encourage critical thinking, and to facilitate social and institutional communication. Readers would be aware of the context of an utterance, and understand academic discourse as "the way that individuals collaborate and compete with others, to create knowledge, to educate neophytes, to reveal learning and define academic allegiances" (Hyland, 2009, p. 2).

Second, the book helps bridge an important gap in the literature by investigating the CL-EAP interface (Chapters 3 to 6) with regard to discourse taxonomy. In the current study, Viana and O'Boyle define a corpus as "an electronic collection of [naturally occurring] texts" that are "wide encompassing" and can "fulfil a specific communicative purpose in a given context" (p. 23), and they argue that "new configurations in the methodological paradigm and/or knowledge source may bring about professional changes" (p. 51). This would mean that CL can bring a number of advantages to the EAP community, such as facilitating data retrieval and processing, implementing situated learning and teaching, providing tools for contrastive inter-/intralingual analyses, reconfiguring EAP practices, as well as developing numerous skills for EAP research. With regard to corpus diversity and corpus compilation, Viana and O'Boyle highlight the variations (spoken and written academic English, diachronic change in academic discourse, and academic genre comparisons) relevant to the EAP

community, with a particular focus on specialised corpora. In order to 'marry' CL and EAP in a more feasible way, they suggest the need to build a variety of specialised corpora, such as discipline-specific corpora, EFL corpora, corpora of extended written genres by postgraduate students, academic popularisation corpora, EAP teacher corpora, and corpora of computer-mediated academic discourse. They also summarise the differences between illustrative levels of corpus specialisation (i.e., educational level, major, and type of essay), and set 24 criteria for specialised corpus compilation (e.g., medium, genre, field, language, time, availability, text length, text production circumstances, and level of interactivity). In addition, useful techniques for corpus analysis for EAP are introduced with the aid of *AntConc* toolkits, ranging from wordlist and concordance to collocation and n-grams. Researchers will benefit greatly from reading this because language data obtained from corpus analysis will either provide them with empirical evidence to examine existing theories/hypotheses, or 'drive' them to observe the invisible rules behind such data.

Third, the book familiarises readers with data-analytic techniques for analysing different types of academic discourse (Chapters 7 to 9), which specifically include descriptive analysis, predictive analysis and prescriptive analytics. With descriptive analysis, researchers can describe the features of data in static statistics, such as frequency. A typical example can be seen when Viana and O'Boyle analyse the turns between students and teachers drawing on data from the UNITALK corpus, in which they found that the average length of the student turn is 18 words whereas that of the tutor turn is 45 words. Predictive analysis involves inferential statistics to generalise the features of language use (academic discourse in this context), with which researchers can draw evidence-based conclusions to predict potential patterns, universal features (see Feng, 2020) and future tendencies. Prescriptive analytics enables researchers to take influential factors into consideration, such as resources and modelling, and implement plans for decision-making based on corpus data. In this respect, Viana and O'Boyle re-emphasise McCarthy's (2021) claim that "[c]ollocability disrupts the belief that lexical items can simply be dropped into grammatical slots to produce well-formed sentences" (p. 10) by exemplifying genre-specific collocations (e.g., *calibrate system* and *adopt system*).

In terms of the monograph's contents, the text produced is logically organised with precise and concise academic language, outlining a detailed framework of integrating CL into EAP research. As aforesaid, all the key

notions are defined clearly with rich background information, and therefore even novice researchers and students without the expertise in corpus-based/corpus-driven EAP studies will not find the book too hard to follow. Additionally, visualisation techniques and well-crafted examples are incorporated into the explanations of theories and principles, making readers self-motivated to explore, in breadth and depth, more potential scaffolding skills and progressive methods in CL-EAP discussions. Another notable strength of the monograph lies in Viana and O'Boyle's offer of three types of in-text tasks, that is, 'activate your knowledge', 'check your understanding', and 'apply it to your context', all of which make the text not merely a fascinating read for its own sake, but more of a practice-based guidebook to 'situate' readers in context-specific scenarios. Readers would feel like sitting *vis-à-vis* the authors and engage with them throughout these chapters. Furthermore, a total of more than 300 references provides readers with a collection of valuable resources pertinent to the interplay between CL and EAP. With such references, readers are encouraged to return to the book's index and extensive resources to look at similar arguments, mapping their thoughts with regard to the state-of-the-art in this research field and coming up with novel ideas for methodological innovations in EAP studies.

By and large, Viana and O'Boyle's present work offers an excellent foundation for either in-class instruction or self-study. For us EAP and ESP teachers and practitioners, it appears that many cutting-edge discussions in the monograph help respond to the query on how CL can support the EAP community in the changing educational landscapes, and therefore it deserves our enthusiastic recommendation.

Reviewed by **Haoda Feng**
Centre for Intercultural Maritime Translation, Dalian Maritime University (China)
haoda_feng@163.com

Acknowledgements

References

Feng, H. (2020). *Form, meaning and function in collocation: A corpus study on commercial Chinese-to-English translation*. Routledge.

Hey, T., Tansley, S., & Tolle, K. (2009). *The fourth paradigm: Data-intensive scientific discovery*. Microsoft Research.

Hyland, K. (2009). *Academic discourse*. Continuum.

McCarthy, M. (2021). *Innovations and challenges in grammar*. Routledge.

Szalay, A., & Gray, J. (2006). Science in an exponential world. *Nature, 440*, 413-414.