BOOK REVIEWS

Čermáková, Anna & Markéta Malá (Eds.) (2021). *Variation in Time and Space. Observing the World through Corpora*. De Gruyter. vi + 405 pp.

Reviewed by Arja Nurmi (Tampere University)

The volume at hand is based on papers presented at the ICAME conference in Prague in 2017. As is usual for such compilations, the theme is quite broad and vague. The three parts of the volume are not balanced, with three chapters in the first ("Meaning in Time and Space"), four in the second ("Variation in Time") and seven in the last ("Variation in Space"). The volume's introduction does not define these themes very clearly. Nevertheless, there is clear cohesion of the volume, arising from the ICAME community and the consistent ways in which authors tackle linguistic aspects of different varieties, genres and topic domains of English past and present, develop new methodologies for answering questions old and new, combine corpus methods with varied theoretical approaches and discuss corpus compilation. Most chapters also provide explicit explanations of how the data was retrieved and analysed, making the volume valuable both for those wishing to replicate the study with different data and those wishing to learn the best practices of corpus-based research.

The initial chapter in the volume, by Wolfgang Teubert, is perhaps the one least concerned with actual corpus-based research. It is more a call for further research on digital communication and particularly its impact on linguistic interaction as a whole (p. 12). Teubert's particular concern is the way in which we are now expected to formulate our exchanges in many contexts in accordance with the requirements of computers and databases, especially in official encounters. Even before computers, bureaucratic communication has often taken place in very structured ways, and, for example, filling in forms was not always suited to individual situations. Therefore, it would seem that the only new aspect here is communication to suit specifically electronic databases. This could be seen as one part of the larger topic of asymmetric interaction (see, e.g., Wrede et al. 2010), as it relates not only to human-machine communication but also to various contexts where different communicative needs meet, and questions of accessibility arise. Teubert highlights the power imbalances

typical in such situations, but is perhaps somewhat alarmist in his assessment of the influence of human-machine communication and online communication on the overall trends in human linguistic practices. Painting with a broad brush, Teubert does not acknowledge that a great deal of work has already been done under the umbrella of Computer-Mediated Communication (see, e.g., Kelsey & St.Amant 2008 for the more technical side of things or Herring et al. 2013 for a pragmatic approach). Overall, Teubert's chapter seems only implicitly relevant to the volume, although he does call for corpus-based research on the topic.

In the second chapter, Michael Stubbs approaches the old problem of intertextuality and meaning and proposes new ways of finding answers through the use of corpora. He shows the necessity of a deep understanding of cultural context when studying intertextuality and the ways it develops meanings in texts. As he uses Sherlock Holmes stories as his data, the recognition of intertextuality in the texts relies on familiarity with the period. His discussion is well-grounded in the discourse about science of the day, showing a useful way to employ Google Ngram for the purpose. Stubbs's suggestion for an iterative approach for studying intertextuality seems like a useful starting point for future studies. The chapter's focus is on developing a method suited to the purpose, and the chapter is rounded up with a good discussion and evaluation of it. The relevance of cultural context is highlighted, as Stubbs mentions (p. 60) that the interpretation of intertextuality is not fixed, but rather may change over time, even if the text itself does not change.

The third chapter, by Gaëtanelle Gilquin, is framed as exploration. Looking back at the history of learner corpus research, she shows how the field has developed over the decades. She brings up many relevant questions, and also presents the gaps in existing corpus-based research, pointing out little studied areas. The plan for studying learner language at different proficiency levels both as panel studies and trend studies seems like a useful way of charting new territory. Considering learners' exposure to English, which is not the same everywhere nor at different points in time, is linked to the problems of different corpora and corpus parts collected at different times. The diachronic trends of EFL are a promising approach, as e.g. changes in ENL varieties, used as models, are varyingly taken into account in teaching and the impact of learners being more and more familiar with the informal online modes of communication will probably lead to increasingly informal production on their part. Furthermore, as British English is gradually influenced by American English, perhaps it would be necessary to also evaluate the traditional evidence for e.g. lexical differences of the *flat/apartment* type in both varieties. As an exploratory study, the

comparison of learner varieties is obviously useful, but leaves a great deal of room for further, more in-depth studies.

The chapters focusing on time in the volume are all research from a diachronic or historical perspective. Sirkku Ruokkeinen brings insights from corpus-based translation studies, focussing on Renaissance texts exploring the world in the literal sense, even if mediated through texts rather than personal experience. Her study shows best corpus linguistic practices in comparing the topic domains studied to other domains. Ruokkeinen's focus is on the way paratext can be seen as offering an interpretation of the work it is attached to. The results are not particularly surprising in all instances, as promotional paratext can be expected to be positively evaluative, but the approach presented in the study shows useful ways of approaching these less studied forms of text.

Also dealing with historical data are Maura Ratia and Carla Suhr. They explore the evolving style of medical writing in Early and Late Modern English, and compare it to Present-day English. The varied nature of texts included in their corpora makes the comparisons somewhat hard to follow at times, and it is not always obvious that they are in fact comparing texts that are sufficiently similar in nature – a perennial problem in long diachrony, as genres evolve. The third historically oriented chapter in the volume, by Beatrix Busse, Kirsten Gather, and Ingo Kleiber, is also diachronically oriented, showing how grammarians' references changed over the nineteenth century. This is another chapter with a methodological focus, combining corpora and networks. The authors show sophisticated methods to deal with a corpus still in a 'messier' stage of construction, and the further results of their project will, no doubt, provide fascinating insights.

The section dealing with time is brought to a close by Cinzia Bevitori's chapter on discourses of justice. Her starting point is philosophical, as she is interested in the ways justice is construed linguistically and discursively. Bevitori approaches her corpus through Critical Discourse Analysis and appraisal theory. The strength of the chapter lies in the detailed analysis of examples, bringing to light how different views of justice are encoded. The discussion of political views, expected from the parliamentary corpus, is successfully integrated into the corpus analysis. Bevitori highlights the importance of corpus-based research being more than counting or statistical methods (p. 195): the results should always be interpreted in their context to increase our understanding.

The last part of the volume, space, deals with different present-day varieties of English. It begins with the inner circle and ends with the expanding circle, to use Kachruan

terminology. Daniela Kolbe-Hanna discusses the variation in the use of *if* and *whether* in a selection of ICE corpora. Through a comprehensive statistical analysis she is able to tease out differences between the varieties studied. Another chapter making good use of the ICE corpora is Martin Schweinberger's on intensifiers. Again, the focus is very much linguistic and the explanation of what was done in retrieving and classifying data is very meticulous and clear. Schweinberger shows a good understanding of his data, as he takes into account the potential diachrony in the ICE corpora he uses, as they represent different points in time.

Moving on to the Outer Circle, Guyanne Wilson discusses existential constructions in Nigerian English. Again, the ICE corpus is in use, allowing comparisons to other varieties as well. Wilson brings up the fact that in the case of outer circle Englishes speaker variables such as gender and age are only rarely used. Another speaker variable to take into account is the speakers' first language or other languages, as this might well have an impact on the English of multilingual nations (cf. Lange 2017; Hansen 2018), which Wilson also discusses to some extent. Furthermore, she brings up the potential problem of English being restricted in many outer circle countries to the more formal contexts with specific associations to status and education. Wilson's discussion of the ICE corpora is detailed, and she makes the wide range of text types and potentially somewhat different text types for the different varieties visible in her discussion.

Carrying on with comparisons of the different ICE corpora, Vander Viana and John Kirk take up academic writing in ten countries from both the inner and the outer circles. This is once again corpus linguistics at its best: not just studying a phenomenon, but also evaluating widely used corpora. The authors also highlight the importance of documenting the texts included in any corpus and making this documentation available to corpus users. They describe a complex array of potential reference corpora for their key word analysis (p. 290) before explaining the reasoning for the choices in their study. This is a useful tool for anyone planning a similar study. One of Viana and Kirk's key findings, although hardly unexpected, is that academic texts are not as uniform as could be expected.

The last chapters of the volume discuss expanding circle varieties. Krittaya Ngampradit's topic is also academic English, as the chapter discusses the language of linguistics dissertations in Thailand and the United States. The corpus has been compiled carefully, and the methodological choice of the author analysing the data twice and comparing the results is an interesting solution for scholars working without a research group to assist them. While cultural differences seem to account for the variances found between

the varieties studied, an interesting question not really touched upon would be to see how academic writing is taught in both contexts.

A further study approaching academic writing is that by Yuesen Yang and Naixing Wei. Their focus is on hedging and they compare Chinese and 'Western' scholars. The term Western seems somewhat ambiguous, as it might mean native speakers of English in inner circle countries or just speakers of English affiliated to universities in inner circle countries. Yang and Wei bring up the importance of academic writing pedagogy and hope their results could be used to improve it. At the same time, it would be interesting to study some of the things already taught and how the texts in the corpus align with the teaching. Here, as in the previous chapter, there are cultural assumptions which need to be taken into account in interpreting the results, such as the inclination to collectivism among Chinese scholars. Yang and Wei have decided to look at the big picture first, but also the potential differences between disciplines and how they construct information may well be revealing in future studies.

The final chapter of the volume is also engaged with academic English in an expanding circle context, but this time in the spoken mode. Michael P. Stevens, Yu-Hua Chen and Simon Harrison describe their corpus compilation process for a multimodal corpus of English. They characterise the variety represented in the corpus as an "emerging variety of L2 academic English as a Lingua Franca." Multimodal corpora bring new opportunities for scholars, but they also bring many research ethical problems, which the authors discuss thoroughly. The description of the different processes involved in corpus compilation, such as the practicalities of data collection and annotation, is also very explicit, which makes the corpus construction transparent and allows others to benefit from the experience of the authors.

As can be seen from the above discussion, there are many links between the chapters. One way of making these links more visible would have been allowing authors to read each other's chapters or adding editorial cross-referencing, but in the time pressures familiar in today's academic world, there is seldom time or opportunity for such measures. As to the subheading of the volume, the editors note (p. 3) that "observing the world through corpora" was changed from the topic of the conference itself, "interpreting the world through corpora". The suggested reason for the change is to allow for the larger scope of topics in the volume. This seems to downplay the merits of the chapters included. All chapters definitely observe the world, but they also interpret the world on some level, often tying the linguistic with the social, political and cultural. As Matti Rissanen used to say, "research begins where counting

ends" (see e.g. Rissanen & Tyrkkö 2013 for a discussion of this). This is echoed by many authors in the volume, as they all feel that observing the number of instances in each category is not enough but rather it is important to understand what lies behind the attested frequencies and patterns made visible. That is to say, we need to interpret the results gained from our corpora to really make sense of them, and we can only really make sense of even the most basic results, if we look at what they represent and how.

References

Hansen, Beke (2018). *Corpus Linguistics and Sociolinguistics. A Study of Variation and Change in the Modal Systems of World Englishes*. Brill.

Herring, S., Stein, D., & Virtanen, T. (Eds.) (2013). *Pragmatics of Computer-Mediated Communication* [Handbooks of Pragmatics 9]. De Gruyter.

Kelsey, S., & St. Amant, K. (Eds.) (2008). *Handbook of Research on Computer-Mediated Communication*. Information Science Reference.

Lange, C. (2017). Indian English or Indian Englishes? Accounting for speakers' multilingual repertoires in corpora of postcolonial English. In A. Nurmi, T. Rütten, & P. Pahta (Eds.), *Challenging the Myth of Monolingual Corpora* (pp. 16–38). Brill.

Rissanen, M., & Tyrkkö, J. (2013). The Helsinki Corpus of English Texts (HC). In A. Meurman-Solin & J. Tyrkkö (Eds.), *Principles and Practices for the Digital Editing and Annotation of Diachronic Data* [Studies in Variation, Contacts and Change in English 14]. VARIENG. https://varieng.helsinki.fi/series/volumes/14/rissanen_tyrkko/.

Scott, M. (2010). What can corpus software do? In A. O'Keeffe & M. McCarthy (Eds.), *The Routledge Handbook of Corpus Linguistics* (136–151). Routledge.

Wrede, B., Kopp, S., Rohlfing, K., Lohse, M., & Muhl, C. (2010). Appropriate feedback in asymmetric interactions. *Journal of Pragmatics*, 42, 2369–2384.