



Book Review

Mikhailov, M. & Cooper, R. (2016). *Corpus Linguistics for translation and contrastive studies: A guide for research*. Routledge, 234 pages. ISBN: 9781138944046

Reviewed by: Mehrdad Vasheghani Farahani
Leipzig University, Germany
mehrdadfarahani1365@gmail.com

DOI: 10.12807/ti.113201.2021.r01

‘Corpus’ refers to “a large collection of authentic texts that have been gathered in electronic form according to a specific set of criteria” (Bowker & Pearson, 2002, p. 9). Since their advent in the 1960s, corpora have been very useful and applicable in various research areas, including Contrastive Analysis, Second Language Teaching and Learning, Critical Discourse Analysis, Stylistics and Translation Studies. The application of corpora in translation has resulted in many studies, as parallel corpora enable researchers to conduct different types of analysis of source and target texts. Notwithstanding this proliferation of research in corpus linguistics, there are by no means many books in this field of inquiry, especially in corpus-based translation studies. This paucity of resource books in this research field has motivated Mikhail Mikhailov and Robert Cooper to publish *Corpus Linguistics for translation and contrastive studies: A guide for research*. This book is comprised of 7 chapters in which the authors coherently and comprehensively elucidate the different concepts and techniques related to Corpus Linguistics and Translation Studies, as well as steps necessary for doing practical research in corpus-based translation studies.

Chapter 1, entitled “Parallel text corpora: a general overview”, functions as an opening chapter which outlines at length the concept of ‘corpus’ and its variations. Indeed, the authors first introduce different types of text corpora so that the novice reader is able to gain an initial insight into the notion of ‘corpus’. Then, they allocate some pages to problems which are inevitable in exploiting corpora. The authors also shed light on the multifarious benefits translators can gain from corpora. They conclude this chapter by explaining text corpora and their advantages for translating, justifying the use of corpora for both researchers and translators.

Creating a parallel corpus is by nature an arduous, convoluted and time-consuming task which necessitates various considerations (Zanettin, 2014). Considering this critical issue, Chapter 2 on “Designing and compiling parallel corpora” introduces the basic and principal preconditions for designing and creating parallel corpora. Step by step, the authors introduce the stages necessary for creating parallel corpora for translation and research purposes.

They explain the required corpus software, planning the corpus, inputting the texts into the software, aligning the texts, annotating the texts, storing the corpus and copyright issues. It is worth mentioning that this chapter is salient in that it gives a broad introduction to the practical aspects of designing and creating a corpus for translators and translation researchers.

Every corpus study, whether in the domain of translation or another area of inquiry, inevitably requires some basic and common techniques (McEnery, Hardie & Andrew, 2011). Without knowing these techniques and applying them, where necessary, the corpus-based study will be flawed and incomplete. Taking this into account, in Chapter 3 on “Using parallel corpora: basic search procedures” the authors delineate such basic techniques as conducting searches, concordance lines, frequency lists, and collocations. Subsequently, readers are introduced to basic terms, including collocations, N-grams as well as KWIC (keywords in context). All of these concepts are introduced with authentic examples from various corpora, such as the British National Corpus (BNC) and the TamBiC Corpus. As the authors roundly argue, unless the competent user learns how to run these orders in corpus software, s/he will be unable to claim that s/he has gained a deep understanding of the corpus usages.

Chapter 4 is the first practical chapter in which corpus is studied in relation to translation by applying concepts already addressed and explained in the previous chapter. It sets forth to offer a clear explanation of the key concepts and techniques for doing practical research in the domain of corpus linguistics. For this purpose, the authors allocate the whole chapter “Processing search results” to translation research and corpora, and shed light on equivalent and parallel corpora, processing frequency lists, and studying collocations and parallel corpora in monolingual research projects. The examples the authors give are from ParRus, a Finnish-Russian corpus compiled by the authors themselves. The authors rely on this parallel corpus as it is related to translation and clarify the concepts for translation researchers.

Any corpus-based research entails mastery of computer tools as well as quantitative/statistical measurements as two preconditions. As a matter of fact, understanding the software prerequisite is eminently enlightening for students interested in doing intensive corpus-based research. Taking this into account, in Chapter 5 “Using parallel corpora: more advanced search procedures”, the authors introduce far more advanced procedures, particularly software tools for quantitative analysis of corpora. For this purpose, this chapter sets forth to clarify the reliability of data for corpus research and measuring the central tendency of data as well as quantitative analysis of concordance as statistical analyses for quantitative corpus-based research. As in the previous chapter(s), to make these concepts more tangible, the authors provide examples from the ParRus corpus as well as some case studies. This chapter ends with collocation research and keywords analysis.

Although chapters 4 and 5 are quite practical, Chapter 6, entitled “Applications of parallel corpora”, is the most practical chapter of the book when it comes to translation. In this chapter, Mikhailov and Cooper introduce a wide range of applications of parallel corpora, including as a resource for online dictionaries, lexicography, terminology as well as morphology, syntax, pragmatics and translation studies. For each section, they give some authentic examples to clarify their point. This chapter is constructive and useful in that it fluently and succinctly delineates these applications with real examples from the Finnish-Russian language pair. As an example, for the purpose of parallel corpora as a resource for online dictionaries, Mikhailov and Cooper give examples from the word *keskittyä*, which is most often translated into English

as ‘concentrate’ or ‘centralize’, but in other contexts it has been translated by other equivalents such as ‘be confined to’, ‘be limited to’, or ‘be restricted to’. These real examples taken from authentic contexts give the reader an opportunity to familiarize him/herself with the way parallel corpora can be used to find appropriate equivalents in the context.

With the advances made in computer sciences, there has been a major progress in creating various types of corpora for different purposes in relation to translation research. This has encouraged the authors to dedicate Chapter 7 of the book, entitled “A survey of available parallel corpora”, to this variety of corpora and purposes. They introduce three types of corpora: multilingual parallel corpora, bilingual parallel corpora and interpretation corpora (spoken corpora). For this reason, they suggest some of the well-known and freely available corpora such as ELRA, LDC, and CLARIN. Further, they suggest some parallel corpora in various language pairs such as the Amsterdam Slavic Parallel Aligned Corpus, The CLUVI Parallel Corpus, and CRATER & CRATER II. For each corpus, there comes a short description so as to familiarize the potential user with it.

Overall, this book can be used as an introductory self-study resource for researchers in corpus-based translation studies. The book has been written in an understandable language and coherent way: the chapters are introduced in a logical and well-structured manner so that the readers can follow the content of the book step by step. Another strength of this book is that at the end there is a glossary of terms so that the reader can have an understanding of the terms used in the book. In addition, the book has two appendices. Appendix 1 is about “Corpus-based M.A. theses at the University of Tampere” and Appendix 2 is about “Sample programs”. By reading these two appendices, the reader will gain a clear view of the topics of this area of inquiry as develop their research skills in corpus-based translation studies.

One limitation is that, although the book includes a wide range of examples, mostly extracted from the studies done by the researchers themselves, the examples are in the Finnish-English or Russian-English language pairs, which may be a challenge for those who have no knowledge of Finnish and Russian.

References

- Bowker, L., & Pearson, J. (2002). *Working with specialized language: A practical guide to using corpora*. London: Routledge.
- McEnery, T. & Hardie, A. (2011). *Corpus linguistics: Method, theory and practice*. Cambridge: Cambridge University Press. doi:10.1017/CBO9780511981395.
- Zanettin, F. (2014). *Translation-driven corpora: Corpus resources for descriptive and applied translation studies*. London: Routledge.