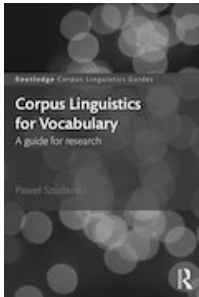


## Corpus Linguistics for Vocabulary: A Guide for Research

November 2020. Volume 24, Number 3

Corpus Linguistics for Vocabulary: A Guide for Research		
<b>Author:</b>	Pawel Szudarski (2018)	
<b>Publisher:</b>	London, New York: Routledge	
<b>Pages</b>	<b>ISBN</b>	<b>Price</b>
Pp. x + 228	978-1-138-18722-1 (paper)	\$148.00 U.S.



Corpus linguistics is typically defined as the analysis of language using corpora (McEnery & Hardie, 2012). Corpora, in turn, refer to large organized collections of authentic data (spoken, written, or both) stored electronically (Tognini-Bonelli, 2001). Corpus linguistics plays a key role in researching, teaching, and learning vocabulary because it demonstrates how language is used authentically and in different contexts. This, in turn, enables instructors to teach language more efficiently, material developers to develop more authentic materials, and students to learn and use authentic words, collocations, and patterns. *Corpus Linguistics for Vocabulary: A Guide for Research* by Pawel Szudarski is an introductory book written to familiarize readers with the interrelationship between corpus linguistics and vocabulary learning. This book is a helpful resource for students and teachers who wish to conduct corpus-based research on vocabulary-related questions.

*Corpus Linguistics for Vocabulary* contains ten chapters and begins by discussing general points concerning corpus linguistics (Chapters 1-2) before moving on to elucidate more specific uses of corpus linguistics for vocabulary purposes (Chapters 3-10). Most chapters deal with common issues of interest to researchers and teachers, such as spoken/written differences, formulaic language, academic and general vocabulary, and so on. The chapters are largely unstructured (except for a “Summary” section that exists in all of them).

Chapter 1 provides clear definitions and explanations of fundamental terms (e.g., corpus linguistics, corpora) and concepts including, but not limited to, common features of corpora, merits and demerits of corpus linguistics, and different types of corpora (e.g., technical corpora, written corpora, spoken corpora). In the second chapter, more specific corpora-related tools, such as frequency analysis, n-gram analysis, and word combinations (e.g., collocations) which might be utilized for searches in corpus linguistics, are explained.

Chapter 3 introduces the concept of vocabulary learning and pertinent terms, such as “vocabulary,” “lexis,” “lexeme,” “lemma,” and “word family.” The author intends to make readers familiar with these terms, as they are necessary in running a corpus analysis and using corpora to analyze various aspects of vocabulary (e.g., synonymy, polysemy, etc.).

The rest of the book (4-10) depicts how different facets of vocabulary, such as formulaic sequences, academic lexical items, etc. can be studied via corpus linguistics. Specifically, Chapter 4 discusses how corpora can be used for finding frequencies of lexical items. It explains how frequency lists might be used by researchers (e.g., adopting stimuli lists appropriate to participants’ proficiency levels) and teachers (e.g., selecting vocabulary items appropriate to participants’ proficiency levels). Chapter 5 explains how corpus linguistics might be employed to explore formulaic language. It discusses lexical bundles, collocations, and colligations, which are the most commonly-used types of formulaic language. It further explains that corpus analysis is useful to recognize and categorize different types of formulaic sequences as integral parts of natural language use. The use of corpus linguistics in teaching vocabulary is presented in Chapter 6 via two approaches: the indirect teaching approach and the direct teaching approach. An indirect teaching approach entails consideration of corpus analysis findings (such as frequency results) in selecting and teaching lexical items. A direct teaching approach, in contrast, foregrounds the direct manipulation of corpora by learners to inductively discover and learn word meanings.

Moving on in the book, Chapter 7 raises the issue of using learner corpora to examine learner language. It depicts the ways that learner corpora can facilitate research questions about learners’ vocabulary learning and development. For example, the chapter describes the English Profile Project (EPP), a research program based on the Cambridge learner corpus which brings together a number of language institutes. It aims to obtain corpus evidence that will serve as a basis for the description of the English learner.

Chapter 8 elaborates on the use of corpora for English for specific and academic purposes. It contains discussion on specialized kinds of corpora, such as the Corpus of Contemporary American English (COCA) and Corpus Linguistics in Cheshire (CLiC) by first illuminating the characteristic features of such corpora and then demonstrating their usefulness. The chapter presents examples of their applications in different areas, including English for Specific Purposes and English for Academic Purposes. Chapter 9 highlights how pragmatics and discourse can be used to examine vocabulary learning. The chapter first discusses the advantages of combining corpus and discourse approaches to analyze lexical items (e.g., reaching more robust and contextualized findings). It then elucidates the process of how corpus techniques can help researchers with pragmatic functions of vocabulary, such as speech acts. Finally, the tenth chapter provides a condensed but comprehensive summary of the book. It also offers numerous ideas and recommendations for researchers, including how to run a corpus-based study to explore the use of language in spoken and written contexts, or how to look into the use of phrasal and single-word verbs in spoken and written contexts.

*Corpus Linguistics for Vocabulary* presents a novel view of corpus linguistics for vocabulary practice, introducing an array of corpus tools and procedures that are increasingly important to examine intricacies and patterns of vocabulary in natural contexts. It is a practical guide for teachers, language learners, and vocabulary researchers who have little or no experience in corpus linguistics. Another advantage of the book is that it provides authentic examples and tasks (with answers and commentary) for almost

every topic raised in the chapters. It deals with real-life issues, such as authentic spoken and written conversations, speech acts, ambiguity in conversations, collocations in speech, etc.

The main shortcoming of the book is its rudimentary explanations and topics. The author could have devoted more chapters to advanced and professional issues regarding vocabulary and corpus linguistics, such as the use of different research methods (quantitative, qualitative, and mixed-methods) in running corpus-based studies to appeal to more experienced teachers and language professionals. Also, the author could have dedicated a chapter to testing and assessment-related issues of vocabulary that can be addressed through corpus analyses. For example, a chapter examining the way corpus analyses can be performed to select items for developing vocabulary tests could have been included.

Taken together, *Corpus Linguistics for Vocabulary: A Guide for Research* is strongly recommended to language learners, teachers, materials developers, and syllabus designers in that it is an up-to-date and insightful collection of information regarding corpus linguistics in vocabulary learning.

## References

McEnery, T., & Hardie, A. (2012). *Corpus Linguistics: Method, theory and practice*. Cambridge University Press.

Tognini-Bonelli, E. (2001). *Corpus linguistics at work*. John Benjamins.

## Reviewed by

Kamal Heidari

Shiraz University, Iran

<[K\\_86\\_tefl@yahoo.com](mailto:K_86_tefl@yahoo.com)>

© Copyright rests with authors. Please cite TESL-EJ appropriately.