REVIEW

Paweł Szudarski:
Collocations, corpora and language learning

(Cambridge: Cambridge University Press, 2023. 89 pages)

Xiaoyi Yang

Hangzhou Dianzi University
Orcid: 0000-0003-4791-112X
China
lucyyang@hdu.edu.cn

Yuan Ping

Hangzhou Dianzi University
Orcid: 0000-0003-2962-6464
China
y.ping@hdu.edu.cn

ONOMÁZEIN 66 (December 2024): 220 - 225
Xiaoyi Yang and Yuan Ping
Paweł Szudarski: *Collocations, corpora and language learning*

221

Collocation has long been of significant interest across various fields, including phrase-ology, corpus linguistics and second language acquisition (SLA), due to its implications for them in diverse ways. The field of corpus linguistics, in particular, has seen a surge of interest in collocation studies, spurred by recent advancements. A substantial body of research (e.g., Nesselhauf, 2005; Durrant, 2014; Gablasova and others, 2017) has ex-plored collocations used by second language (L2) learners through corpus analysis. Among them, key studies have underscored the pivotal role of collocations and corpora in English language instruction and acquisition. However, systematic appreciation of the crucial use of corpora as a tool for teaching English vocabulary remains underdeveloped (Szudarski, 2023).

Paweł Szudarski's *Collocations, corpora and language learning* is a noteworthy attempt to address the use of corpora in collocation teaching and learning. This book is the latest addition to Cambridge University Press's Elements in Corpus Linguistics series edited by Susan Hunston. Despite its brevity at 89 pages, the book provides a nonetheless extensive overview of corpus research into the issues of learning collocations for L2 learners.

The book has an introduction, three main chapters, and a conclusion. The introductory chapter delineates the book's aims to present various corpus methods and types of analysis related to collocations to give readers "a thorough theoretical understanding of collocations as a key concept in corpus linguistics" (2) so that they become "well versed in the mechanics and methodologies associated with corpus-based analyses of collocations" (ibid.). The au-thor sets out seven research questions that readers should be able to answer after reading the book.

Chapter 2 delves into how collocations are conceptualized as a key type of formulaic lan-guage. The author adopts Sinclair's (1991) idiom principle which views language texts as composed of phrases and segments and highlights the terminological challenges due to multiple terms employed in the literature and different types of phrases in defining these concepts. He defines collocations as graded phenomena with various degrees of probability and identifies three main traditions in collocation research: phraseology-based, frequen-cy-based, and a hybrid of the two. The author discusses the relationship between the various ways of defining collocations, providing examples of studies that use varying criteria to do so. The chapter explores the research on collocations as a textual phenomenon but also reviews the studies that examine the psycholinguistic reality of collocations, such as semantic preferences, "tendencies of words to cluster together as units of language that share semantic features" (23), and semantic prosody, "evaluative or attitudinal meanings that result from the co-occurrence of words within collocations" (ibid.).

Chapter 3 focuses on the types of corpus analysis conducted, measures used, and di-mensions identified in corpus-based collocation research, covering various types, such

ONOMÁZEIN 66 (December 2024): 220 - 225
Xiaoyi Yang and Yuan Ping
Paweł Szudarski: *Collocations, corpora and language learning*

222

as frequency analysis, collocation analysis, and keyword analysis, both quantitative and qualitative. The author discusses several corpus-derived measures of collocations, such as t-score, mutual information (MI), Delta P, and log dice in detail. The t-score, for example, "tests the level of confidence in evidence accumulated in a given corpus in terms of the representation or distribution of word co-occurrence in language at large" (27). MI measures "the strength, tightness and exclusivity of a relationship between two words" (28) and is often used to determine the exclusivity of collocations. Delta P is a measure of the directionality and asymmetry of collocations. Log dice also measures the exclusivity of collocations and their dispersion. The chapter also discusses other statistical procedures used in such studies, such as Cohen's d, chi-square, and log-likelihood. The author emphasizes that each measure has its limitations and recommends the use of several. The chapter concludes with a discussion of the synergy of quantitative and qualitative analysis in collocation research. Statistics in corpus analysis yield quantitative data, while qualitative insights are primarily derived from concordancing through the use of keywords in contexts (KWIC).

Chapter 4 addresses corpus-based research on learning and teaching L2 collocations. It explores corpora, collocations, and usage-based approaches to language learning, as well as learner corpora, interlanguage, and SLA. Granger (2012: 7) defines learner corpora as "computerized databases of foreign or second learner language". The chapter discusses learner-corpus research and collocations as indicators of L2 proficiency. It examines aspects of L2 collocational learning across three dimensions: amount of use, accuracy/appropriacy of use, and goodness/automaticity of intuitions of formulaic language (Schmitt, 2010). The chapter covers contrastive interlanguage analysis and assessment of L2 lexical richness and sophistication and discusses applied corpus linguistics and the use of corpora for teaching L2 collocations. The chapter highlights the indirect and direct impact of corpora on teaching L2 collocations. Indirect impacts include "the use of corpus findings to select pedagogically relevant vocabulary, create lexically oriented teaching materials and raise L2 learners' awareness of single-word and phrasal vocabulary" (59). Direct impacts involve treating corpus analysis "as a teaching technique involving the presentation of corpus data" (ibid.). The chapter mentions several corpus-based lists of collocations, such as the Academic Collocation List (Ackermann and Chen, 2013) and the Academic English Collocation List (Lei and Liu, 2018), and concludes with a discussion on corpus- and technology-enhanced L2 teaching resources, computer-assisted language learning, and data-driven learning, providing examples, such as IdiomTube (Lin, 2022) and ColloCaid (Frankenberg-Garcia and others, 2019).

Chapter 5 provides conclusions and outlines avenues for future research. It discusses the "corpus revolution" and recent developments in collocation research, including crowdsourcing and data crunching. The chapter suggests new avenues for collocation research, such as adopting group versus individual perspectives on collocational development, and advo-

ONOMÁZEIN 66 (December 2024): 220 - 225
Xiaoyi Yang and Yuan Ping
Paweł Szudarski: *Collocations, corpora and language learning*

223

cates more mixed-method and interdisciplinary research, replication studies (Szudarski and Mikołajczak, 2023), and work that bridges the gap between research and practice.

This book covers the main concepts and theorizations of corpus-based collocation research adequately and adds to our understanding of collocations as used by L2 learners, which stands out from others on the subject (e.g., Nesselhauf, 2005; Durrant, 2014; McCallum and Durrant, 2022). For instance, it takes a systematic approach to corpus-based collocation research compared to Nesselhauf (2005), which is based on a study on collocations in a corpus of English by L2 advanced learners. It also differs from McCallum and Durrant (2022), another book in the series that explores the linguistic and non-linguistic factors influencing students' grades in two writing modules of the first-year composition program at the University of South Florida. However, it bears similarities to Durrant (2014) in its function as a meta-analysis of corpus-based collocation research and as a resource for corpus research into L2 collocations.

The book has several distinctive features, including 15 insightful quotes that are placed alongside the main text that underscore key points and provide expert perspectives that enrich the discussion. It also features 10 "study boxes" with examples and summaries of studies in a format similar to that in Mackey and Gass (2012) to help readers get the main ideas of these studies. Each summary provides a brief overview of the study, covering aspects, such as its background and aims, the research question(s) it seeks to answer, the methodology employed, and the results and discussion arising from it. The book is generally coherent, clearly and well written and organized though, as a minor point, some of the figures (such as figure 4) are not of high enough resolution to be easily readable, particularly with regard to the concordance lines presented in KWIC format.

The book has several pedagogical implications. The author highlights the ongoing challenge of bridging the gap between corpus research, L2 learning, and teaching practice. This challenge stems from the need to translate theoretical insights from corpus research into practical strategies for L2 learning and teaching. The empirical nature of corpus-based research methods not only describes language and provides qualitative functional explanations, but also enhances the teaching of English vocabulary. The corpus-based lists of collocations and corpus- and technology-enhanced L2 teaching resources discussed in the book are several examples. These lists can serve as valuable resources for both teachers and learners, and the resources can make language learning more engaging and relevant, helping L2 learners understand how collocations are used in real-world contexts.

In sum, this book not only contributes to the academic discourse on collocations and corpus linguistics but also provides practical insights that can enhance L2 teaching and learning and, as such, is a highly recommended read, especially for students, teachers, and researchers in the fields of SLA, corpus linguistics, phraseology, and collocation research.

ONOMÁZEIN 66 (December 2024): 220 - 225
Xiaoyi Yang and Yuan Ping
Paweł Szudarski: *Collocations, corpora and language learning*

224

Funding

References

Ackermann, Kirsten, and Yu-Hua Chen, 2013: "Developing the Academic Collocation List (ACL): A corpus-driven and expert-judged approach", *Journal of English for Academic Purposes* 12 (4), 235-247.

Durrant, Philip, 2014: "Corpus frequency and second language learners' knowledge of collocations: A meta-analysis", *International Journal of Corpus Linguistics* 19 (4), 443-477.

Frankenberg-Garcia, Ana, Robert Lew, Jonathan C. Roberts, Geraint Paul Rees and Nirwan Sharma, 2019: "Developing a writing assistant to help EAP writers with collocations in real time", *ReCALL* 31 (1), 23-39.

Gablasova, Dana, Vaclav Brezina and Tony McEnery, 2017: "Collocations in corpus-based language learning research: Identifying, comparing, and interpreting the evidence", *Language Learning* 67 (S1), 155-179.

Granger, Sylviane, 2012: "How to use foreign and second language learner corpora" in Alison MacKey and Susan M. Gass (eds.): *Research Methods in Second Language Acquisition: A Practical Guide*, Chichester: Blackwell, 7-29.

Lei, Lei, and Dilin Liu, 2018: "The academic English collocation list", *International Journal of Corpus Linguistics* 23 (2), 216-243.

Lin, Phoebe, 2022: "Developing an intelligent tool for computer-assisted formulaic language learning from YouTube videos", *ReCALL* 34 (2), 185-200.

MacKey, Alison, and Susan M. Gass (eds.), 2012: *Research Methods in Second Language Acquisition: A Practical Guide*, Chichester: Blackwell.

McCallum, Lee, and Philip Durrant, 2022: *Shaping Writing Grades: Collocation and Writing Context Effects*, Cambridge: Cambridge University Press.

Nesselhauf, Nadja, 2005: *Collocations in a Learner Corpus*, Amsterdam: John Benjamins.

Schmitt, Norbert, 2010: *Researching Vocabulary: A Vocabulary Research Manual*, Basingstoke: Palgrave Macmillan.

ONOMÁZEIN 66 (December 2024): 220 - 225
Xiaoyi Yang and Yuan Ping
Paweł Szudarski: *Collocations, corpora and language learning*

225

SINCLAIR, John, 1991: *Corpus, Concordance, Collocation*, Oxford: Oxford University Press.

SZUDARSKI, Paweł, 2023: "Corpora and teaching vocabulary and phraseology" in Reka R. JABLONKAI and Eniko CSOMAY (eds.): *The Routledge Handbook of Corpora and English Language Teaching and Learning*, Abingdon: Routledge, 41-55.

SZUDARSKI, Paweł, and Sylwia MIKOŁAJCZAK, 2023: "Proficiency, language of assessment, and attention to meaning and form during L2 comprehension: Methodological considerations in L2 replication research", *Studies in Second Language Acquisition* 45 (1), 276-288.