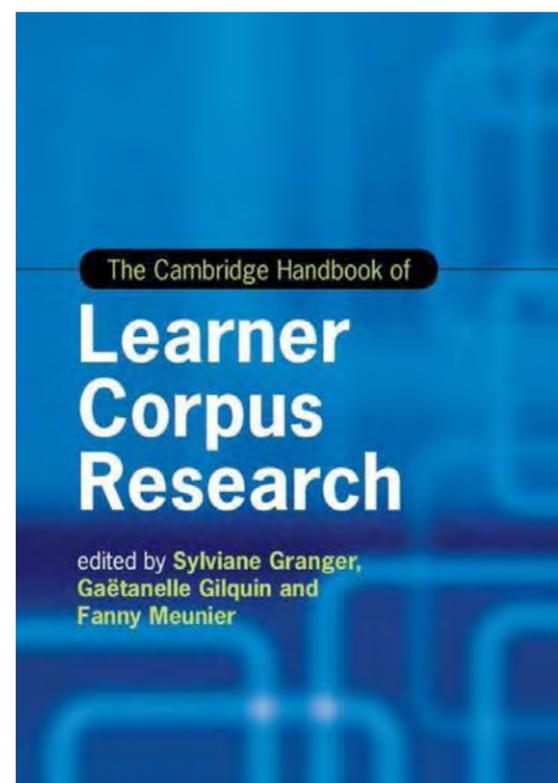


BOOK REVIEW: THE CAMBRIDGE HANDBOOK OF LEARNER CORPUS RESEARCH

Lee McCallum



Title: The Cambridge Handbook of Learner Corpus Research
Editors: Sylviane Granger, Gaëtanelle Gilquin, and Fanny Meunier.
Publisher: Cambridge University Press
Available: www.amazon.co.uk, £60.84 (Kindle edition)
Published: 2015
ISBN: 978-1-107-04119-6
Pages: 755

The Cambridge Handbook of Learner Corpus Research is part of the Cambridge Handbooks in Language and Linguistics series and is an overview of the latest research and theory in learner corpora. The book is aimed at students, teachers and researchers who are interested in gaining a basic overview of how collections of learner texts (known as corpora) can be used.

The book is divided into five parts with a bibliography as well as author, corpus and software indexes at the end of the book to consolidate the key authors and software that readers have access to. Each part has chapters based around a theme, and each chapter is written by an established authority on corpora including input from Sylviane Granger, Marcus Callies and Sandra Götz. The chapters all make reference to a wealth of historical and contemporary research in learner corpora, and the bibliography cites references from the 1800s to early 2014, giving readers access to historical and contemporary information.

Part 1 details the design and methodology of learner corpora and holds particular relevance to budding linguists and teachers who may want to experiment in creating and annotating their own corpus for learning or teaching purposes. Chapter 7 on error annotation may be useful for teachers who want to explore their students' errors in more detail. It includes information on available error tagging software and highlights the steps that error analysis requires. Part 2 builds on this by analysing learner language with a focus on lexis, grammar and discourse in chapters 9, 11 and 12 respectively. These chapters show the latest research and applications of corpora in language teaching and detail how widely available software and tools such as LexTutor and Sketch Engine can be used to encourage the book's intended audience to experiment with corpus tools and software. Part 3's strength is that it brings together corpora and second language acquisition theories. Part 3 successfully focuses, through

chapter 14, on second language theory and corpora, and chapter 17 explores learners' development patterns and how corpora can help support and inform traditional and contemporary theories. The corpora contain large amounts of authentic learner language, and, alongside the corpus tools and software that the handbook promotes, it is easy for readers to see how corpora lend support to these theories in an efficient and accurate manner. As the penultimate part, part 4 sets out how learner corpora can be used to inform different branches of language teaching with chapters on pedagogic corpora, EAP and ESP corpora, and using corpora to inform instructional materials and to validate language tests. The final part, part 5, focuses on Natural Language Processing and includes automatic grammar and spell checking tools for language learners in chapter 25 and automated essay scoring in chapter 26. These chapters are invaluable modern contributions to the field of learner corpus research and will likely be researched further in the future given the increasing influence of computer-aided language learning on traditional classroom practice.

The strengths of the handbook are its wide ranging and well-researched software and corpus indexes. The software index includes information on well-known taggers which annotate corpus texts and identify word classes and further help teachers, students and researchers become more aware of different language patterns. The book's readers will find it especially useful that the book promotes established software like the CLAWS tagger, Sketch Engine and WordSmith as well as newer or less well-known software including E-rater by the English Testing System (ETS), the Example Extractor Engine for Language Teaching (Exxelant), and also parsing and tagging software such as the Stanford parser and tagger.

Similarly, the corpus index lists

a wealth of existing corpora spanning different languages, genres and language varieties. For example, British and American English corpora are included as well as large general corpora that allow access and comparative investigation of different genres including newspaper English and English used in novels. These corpora can be used on their own for research and pedagogic purposes; however, large general corpora such as the British National Corpus (BNC), the Contemporary American English Corpus (COCA) and the specialized British Academic Writing English (BAWE) can be used as reference corpora. These reference corpora can be used with smaller learner corpora such as the Chinese Learner English Corpus (CLEC) to give comparative insights into the interlanguage of Chinese learners compared to the branches of English found in the reference corpora.

Despite the wide-ranging content of the handbook and the awareness that it raises about learner corpora, the handbook may benefit from focusing on using learner corpora for classroom application to gauge how learners respond to using this data. The handbook could also further explore longitudinal individual or group learners' progress which would provide a unique insight into development and may be a topic that learners show interest in.

Overall, this handbook makes a key contribution to the emerging value of learner corpora and gives those who want to learn about or use learner language all the support and tools they need to gain confidence in exploring its unique patterns. ○

**AUTHOR BIOGRAPHY**

Lee McCallum is an EdD candidate at the University of Exeter and currently teaches in China. She has wide-ranging teaching and assessment experience from Europe and the Middle East. Her research interests include testing, corpus linguistics and L2 writing.

EMAIL

leemccallum811@gmail.com