

Alexandra U. Esimaje, Ulrike Gut and Bassey E. Antia (eds.). *Corpus linguistics and African Englishes* (Studies in Corpus Linguistics 88). Amsterdam / Philadelphia: John Benjamins Publishing Company. 2019. ix, 403 pp. ISBN 9789027202192 (HB). Reviewed by **Frederic Zähres**, Bielefeld University.

The volume *Corpus linguistics and African Englishes* presents a diverse collection of corpus analyses as well as a variety of recent specialized corpus compilation projects in the context of African Englishes. With this interface of corpus linguistics and World Englishes research, the ambitious volume gives a voice to a wide range of both junior and senior researchers dealing with under-researched varieties of English across Sub-Saharan Africa and their individual sociolinguistic contexts, which makes it a valuable resource and contribution to both fields of (variational) linguistic inquiry. The book, according to the editors' introduction, addresses both readers who are "new to, and experienced with, corpus research" (p. 4) and, additionally, aims to be "a reference work for all international researchers investigating the linguistic properties of African Englishes" (back cover). The 15 chapters found in the volume certainly achieve that aim with their accessible presentation of multifarious approaches to corpus compilation and analysis. Unfortunately, the variety of approaches and contents leads to a slightly unfocussed volume: the framing of the book as a whole remains somewhat vague and its internal structure is not entirely plausible, as will be discussed below.

After the preface by **Biber** and the brief introduction by the editors **Esimaje, Gut**, and **Antia**, the 15 individual chapters of the book are divided into three parts: Part 1 is titled "Corpus linguistics and corpus building", Part 2 is named "Corpus-based analysis of African Englishes", and Part 3 is called "Applications to corpora in English language teaching and learning".

Part 1 of the volume aims at introducing the general topics of corpus linguistics and compilation to the reader by presenting seven different African perspectives on the matter. Chapter 1.1 by **Esimaje** and **Hunston** serves as a comprehensive introduction to the discipline of corpus linguistics and deals with

the history of corpora, important concepts and terminology as well as applications of corpus research. This is followed by Chapter 1.2 by **Fuchs, van Rooy, and Gut** who present a practical introduction to corpus-based research by recommending useful tools and resources for researchers and illustrate different approaches to corpus analysis via three case studies using data from the GloWbE corpus and the Nigerian component of the ICE corpora. These two handbook-like chapters have a rather general scope and are aimed at readers new to the field of corpus linguistics – and both chapters fulfill this role well with their very accessible, transparent, and illustrative style. Since both introductory chapters of Part 1 have a similar scope, they also contain a certain degree of overlap, which could have been avoided in lieu of clearer descriptions and explanations of some of the technical aspects of Chapter 1.2, e.g. the use and usage of concordance lines and concordancers on pp. 46–47, which is illustrated with a screenshot (Figure 2) but not well enough explained for inexperienced readers. Apart from that, the introductory chapters of Part 1 set the stage for the remaining five chapters, which all deal with individual specialized corpora and their creation processes.

The specialized corpora presented include an error-annotated learner corpus of Nigerian and Cameroonian English (Chapter 1.3 by **Esimaje**), a corpus of spoken English by Black Namibians for morpho-syntactic analysis (Chapter 1.4 by **Steigertahl**), a historical corpus of written Ghanaian English covering the period from 1966 to 1975 allowing for real-time diachronic analysis when complemented with the Ghanaian ICE component (Chapter 1.5 by **Brato**), a tagged corpus of spoken Cameroon Pidgin English giving comparatively rare insights into properties of creole language (Chapter 1.6 by **Ozón, FitzGerald, and Green**), and a corpus of educated Cameroon English as a resource for English language teaching and comparative analyses in the Cameroonian and international contexts (Chapter 1.7 by **Nkemleke**).

While there is a large amount of variety in corpus design and contents, the focus of the majority of the chapters is relatively homogeneous in that the corpus compilation process is described very extensively and transparently, especially regarding challenges African contexts potentially pose. Examples for this can be found in Chapter 1.4, where Steigertahl thoroughly reflects her elicitation instruments and methodology in the northern Namibian context, where she faced literacy issues, or Chapter 1.5, where Brato meticulously describes the challenges of collecting suitable historical data from archives and the subsequent digitization process. However, with the volume's aim of being a reference work and virtually presenting best practice cases for corpus compilation in this first part of the book, some methodological aspects of the individual chapters

would have profited from more problematization, as exemplified by the practice of data processing in Chapter 1.3, p. 77: the author uses an automatic online file conversion service to create PDF files from the scanned handwritten student essays that are part of the learner corpus, which poses potential ethical problems, especially with this particular type of data. Nonetheless, the inclusion of a learner corpus incorporating data from two neighboring West African countries is a valuable addition to the volume. The same goes for the inclusion of a corpus on Cameroon Pidgin English, which is also the only corpus of the whole volume (with the exception of the included ICE and BYU corpora) that is explicitly made publicly available on the Internet. While the notion for making corpora available for other researchers is strongly present in the introductory chapter 1.2 (p. 62), unfortunately only a minority of the contributors position themselves and their work in this regard.

The book's second part gives readers an overview of scopes of and in-depth approaches to corpus analysis by presenting five individual studies. This section starts with a comparative morphological corpus analysis of spoken Nigerian and Black South African Englishes via data from ICE Nigeria and the Xhosa-English corpus, which methodologically also explores more complex statistical means for gaining insights from corpus data (Chapter 2.1 by **van Rooy**). Chapter 2.2 by **Gut** and **Unuabonah** compares the frequency and stylistic variation of stance markers across West African Englishes and British English by analyzing data from the Ghanaian, Nigerian, and British components of the ICE corpus, respectively. Chapter 2.3 by **Kautzsch** describes the compilation of a corpus of online newspapers in the Namibian context for the investigation of potentially nativized lexical and morpho-syntactic features of English in Namibia, especially in conjunction with the available NOW corpus. After his introduction of the Historical Corpus of English in Ghana in Chapter 1.5, **Brato** describes a real-time diachronic lexicon study in Chapter 2.4 comparing frequencies of word-formations potentially exclusive to the Ghanaian context across the historical corpus and the Ghanaian component of the ICE corpus. The final chapter of the volume's second part is Chapter 2.5 by **Isingoma** and **Meierkord**, which sheds light on language contact phenomena and semantic changes within the lexicon of Ugandan English according to data from the Ugandan ICE component. The authors complement the ICE corpus data with the compilation of a web-based corpus of Ugandan English and the conduction of acceptability studies as well as production studies with speakers of Ugandan English for the triangulation of data.

As obvious from these brief summaries, the second part of the volume contains a diverse selection of papers that address structural and sociolinguistic

issues in West, Southern and East Africa. The approaches to corpus analysis in this section of the book give inexperienced readers a good idea of how varied corpus linguistics can be, especially in conjunction with further (e.g. statistical, experimental, or qualitative) methodologies. The chapters of Part 2 also give an overview of relevant questions for linguistic analyses in the World Englishes context, e.g. by addressing placement of varieties of English in current World Englishes models (as shown in Chapters 2.3 and 2.4) as well as the issue of whether a researcher's focus should lie on national varieties instead of supranational ones (e.g. Chapter 2.2, but also Chapter 1.3).

Part 3 of this volume aims to investigate the relationship of African Englishes corpora and educational perspectives and accordingly presents three chapters: Chapter 3.1 by **Adeyemi** compares frequency and stylistic variation in use of conjunctions by Nigerian learners of English through a comparison of data from the Nigerian Learner English Corpus and the Louvain Corpus of Native English Student Essays. Chapter 3.2, by **Schmied**, critically reflects the notion and adequacy of native-speaker standards for the context of African Englishes learners, as illustrated by an analysis of preposition usage within a corpus of Cameroon English MA theses, which is compared to data from Kenyan and British ICE components as well as GloWbE and NOW. While the message of the chapter is quite a crucial one in the World Englishes context, the analysis partly loses its effectiveness through its somewhat cumbersome presentation in the form of numerous low-quality screenshots of the respective GloWbE and NOW corpus user interfaces, which are supplemented with only little description to assist readers with little previous exposure to the BYU corpora (pp. 364–369). The final Chapter 3.3 by **Antia** and **Hendricks** is a diachronic corpus study of social semiotics in the discourse of a South African political party, which presents linguistic evidence for the party's growing responsiveness to diversity on different levels. With a focus on political discourse analysis, this last chapter stands out methodologically compared to the other contributions of the volume, and its explicit focus on sociolinguistic in addition to structural phenomena showcases the potential of corpus linguistics in the World Englishes context. The placement in this third section of the volume seems not quite plausible, however, since the study does not relate to any educational context.

While the individual contributions give great insights into corpus linguistics methodology and approaches to African Englishes, the editors unfortunately have not been able to create an adequate framing and overall structure to the whole volume. What might already be deducible from the above summaries alone becomes evident after having studied the book's chapters in more detail: the three parts the volume is divided into are quite artificial. There does not

seem to be a plausible reason why, for example, Chapters 2.3 and 2.5 are placed in the second part of the book instead of the first one as they contain at least an equal amount of details regarding corpus compilation as some papers that have been placed in that section of the book. This applies to all three sections of the volume. The motivation for this structure does not become more transparent when referring to the editors' brief introduction to the volume. Also, the editors acknowledge criticism in recent corpus linguistics and World Englishes research (i.e. that "inherent in both corpus linguistics and African Englishes is a view of language as a closed and enumerable set of features that are consigned to specific spaces or geographical borders", p. 4), but remain relatively vague in how this volume addresses or overcomes this problem – and, in fact, fewer than half the contributions even problematize similar notions. The volume is already an ambitious collection and this further extension in scope would not have been necessary.

Nonetheless, *Corpus linguistics and African Englishes* constitutes an overall accessible, comprehensive, and diverse collection of corpus linguistic approaches in the context of African Englishes. Although the volume's diversity and breadth in scope pose a challenge to its overall coherence and framing and, eventually, leave a somewhat fragmentary impression on the reader, it manages well to a) shed light on a range of methodologies for corpus compilation and analyses in different African contexts (especially those that have not been well-researched, such as Namibia and Uganda), and b) give an overview of recent structural, sociolinguistic, and educational perspectives on several varieties of English spoken on the African continent. While the volume is primarily targeted at readers with little previous knowledge of either corpus linguistics or African Englishes, more advanced readers should also easily be able to draw inspiration from the specialized corpora introduced and new insights from the analyses presented. Thus, it represents a valuable resource both for researchers and instructors interested in corpus linguistics and/or under-researched varieties of English in Sub-Saharan Africa.