

References

- Bybee, Joan. 2005. The impact of usage on representation: Grammar is usage and usage is grammar. *Presidential address at the Annual Meeting of the Linguistic Society of America*, 8 January 2005.
- Denison, David & Marianne Hundt. 2013. Defining relatives. *Journal of English Linguistics* 41(2), 135–67.
- Fischer, Olga. 1992. Syntax. In Norman Blake (ed.), *The Cambridge history of the English language*, vol. II: 1076–1476, 207–408. Cambridge: Cambridge University Press.
- Kortmann, Bernd & Kerstin Lunkenheimer. 2013. *The electronic atlas of varieties of English*. Leipzig: Max Planck Institute for Evolutionary Anthropology. <http://ewave-atlas.org>
- Mesthrie, Rajend. 2010. New Englishes and the native speaker debate. *Language Sciences* 32, 594–601.
- Prince, Ellen. 1995. On *kind*-sentences, resumptive pronouns, and relative clauses. In Gregory N. Guy, John G. Baugh, Deborah Schiffrin & Crawford Feagin (eds.), *Towards a social science of language: A festschrift for William Labov*, 223–35. Cambridge: Cambridge University Press.
- Radford, Andrew. 2018. *Colloquial English: Structure and variation*. Cambridge: Cambridge University Press.
- Rissanen, Matti. 1999. Syntax. In Roger Lass (ed.), *The Cambridge history of the English language*, vol. III: 1476–1776, 187–331. Cambridge: Cambridge University Press.
- Rizzi, Luigi. 2015. Cartography, criteria and labeling. In Ur Shlonsky (ed.), *Beyond functional sequence: The cartography of syntactic structures*, 314–38. Oxford: Oxford University Press.
- Traugott, Elizabeth C. 1992. Syntax. In Richard Hogg (ed.), *The Cambridge history of the English language*, vol. I: *The beginnings to 1066*, 168–289. Cambridge: Cambridge University Press.
- Yáñez-Bouza, Nuria. 2007. Preposition stranding and prescriptivism in English from 1500 to 1900: A corpus-based approach. PhD dissertation, The University of Manchester.

(Received 24 February 2020)

doi:[10.1017/S1360674320000179](https://doi.org/10.1017/S1360674320000179)

Alexandra U. Esimaje, Ulrike Gut and Bassey E. Antia (eds.), *Corpus linguistics and African Englishes*. Amsterdam: John Benjamins, 2019. Pp. ix + 403. ISBN 9789027202192.

Reviewed by David Jowitt, University of Jos

This book, volume 88 in the John Benjamins Studies in Corpus Linguistics series, brings together a number of studies concerned with corpus linguistics – henceforth ‘CL’ – and its relevance to the study of African varieties of English. It is the first book of the kind to be published, a fact significant in view of certain features that have marked linguistic research in the contemporary era. On the one hand, there is the now long-standing scholarly interest in worldwide varieties of English, together with its eventual reinforcement by another development, the growing use of electronic corpora for the study of language. On the other hand, the number of published corpus-based studies of African varieties of

English has been comparatively limited, and the editors of this book point out that in CL research these are 'still severely underrepresented'. Alexandra Esimaje in her main chapter suggests that the reason is a reluctance in Africa to use computer technology for linguistic research.

The editors also assert, however, that the picture is changing; and being anxious to encourage the change, one of their overall tasks is to provide clear explanations of the characteristics and aims of CL for the benefit of African researchers. They have surely succeeded, especially as everything about the book, the John Benjamins production as well as the writing, is of a high standard. It must be pointed out, however, that one of the reasons for the earlier underrepresentation is that lecturers in African universities battle with problems of daily survival, and cannot afford to buy books of this sort. Publishers in the Western world should think of ways of making their books more accessible to African readers, since such readers constitute the principal obvious market.

The book comprises an Introduction and fifteen chapters, and it is well structured in relation to its overall objective. The seven chapters of part I seek to draw attention to general issues of corpus compilation and applicability. The five chapters of part II focus on several particular African English varieties, showing how the study of a variety is enriched by the use of a corpus pertaining to that variety. The three chapters of part III deal with applications of corpora, and it might be considered an artistic flourish that the final chapter suggests their applicability in a field well beyond variational linguistics or education, namely political discourse.

Geographically speaking, the coverage of the book is wide, although for obvious reasons only those countries already represented by a corpus are covered. In some cases the corpus is a 'national' component of the *International Corpus of English* (ICE); but several chapters make use of non-ICE corpora. Nigeria and Cameroon are both well represented in the book. Kenyan English is not, even though Kenya supplied a great number of the texts making up ICE-East Africa, which was compiled in the early days of the ICE project. Namibia, where a possibly distinct variety is only beginning to take shape, is represented by two chapters. Perhaps eventually another book of this kind will give space to the Englishes of other African countries, including Ethiopia and Rwanda, where English is the official language and also the language of instruction at higher levels, and francophone and lusophone countries, at present unexplored territory.

Each of the first two chapters of part I can serve as a lucid and comprehensive guide to CL. Each thus achieves the overall Africa-oriented objective, while proving potentially useful to anyone who wishes to know more about the subject. In their wide-ranging chapter 'What is corpus linguistics?', Alexandra Esimaje and Ulrike Gut discuss first the concept of corpus linguistics, then key concepts in corpus linguistics (an interesting difference), and then types and applications of corpora. The first section includes a few remarks on the question of the relationship of CL to linguistic theory, and the authors suggest that CL tends to support the idea that the essence of the structure of language lies in the patterning of lexis. In the second main section we find an extensive discussion of the concepts of the design, content, size, balance, representativeness,

output and annotation of corpora; while the third section describes an array of corpus types – ‘general’, ‘specialized’, ‘comparable and parallel’, ‘learner’ and ‘diachronic’, and also gives an outline of available software tools.

The second of these two chapters, ‘Corpus-based research on English in Africa’, by Robert Fuchs, Bertus van Rooy and Ulrike Gut, is more explicitly addressed to African needs. It highlights the usefulness of the ICE and also the *Global Web-Based English Corpus* (GloWbE), with its 1.9 billion words drawn from twenty varieties of English around the world including some African ones, but it also mentions several other corpora, including a corpus of Zimbabwean English (a variety so far understudied) that was compiled by the late Marungudzi (2017). A list of corpora shown on p. 45 gives important indications as to whether or not there is a fee for access.

The latter part of the chapter presents three ‘African’ case studies, cunningly related to lexis, syntax and phonology respectively. The first uses the GloWbE to study the frequency in several varieties of *brother* and words derived from it (such as *bro*), and this was a felicitous choice, given that, as the authors say, ‘this word is . . . claimed to be used with a wider semantic range than just the meaning of a male sibling in sub-Saharan Africa’ (p. 52). Nigeria turns out to have the highest frequency for nearly all the variants. The second study investigates the frequency in African varieties of the present perfect tense, and finds that it is used less frequently in most such varieties than in British English. The third study investigates /h/-deletion in Nigerian English, and suggests that this occurs across all ethnic groups in Nigeria, although the figures actually indicate that it is more frequent among mother-tongue Yoruba speakers.

Subsequent chapters in part I have a common principal theme: the description of more unusual types of corpus. In ‘The purpose, design and use of the Corpus of Nigerian and Cameroonian English Learner Language (Conacell)’, Esimaje describes the 443,000-word corpus that she compiled, with texts drawn from essays and examination scripts written by university and secondary-school students in the two countries indicated. She emphasizes that the use of a learner corpus is attractive for those involved in language education – and one may add that it should be all the more attractive given that in Africa (and elsewhere) English is learned mainly through the formal education system. Such a corpus is naturally used to investigate the incidence of errors, and Esimaje’s two case studies investigate those of spelling and tense use respectively. The first reveals that, apart from errors of capitalization, spelling errors are infrequent; the second reveals that the vast majority of tense errors occurred because the present or unmarked form of the verb was used where the past was required. This finding is worth setting beside that presented by van Rooy in his chapter (see below).

The corpus of Namibian English that Helen Steigertahl describes in the next chapter (‘Introducing a corpus of English(es) spoken in post-independence Namibia: Insights into corpus design and quantitative analyses’) is novel in being composed of sociolinguistic interviews à la Labov. The evidence provided by the 77 educated interviewees for a distinctively Namibian variety of English is perhaps rather slight: thus, although they showed a tendency to use *will* rather than *would* with a hypothetical future meaning, and deontic *must* in preference to *shall/should*, English

modals are in general highly versatile in their meanings. In her introductory section Steigertahl points out that English was proclaimed Namibia's sole official language when the country at last secured its independence from South Africa in 1990, but she might also have said that before then English had become well established in the country, undoubtedly because of the influence of English-speaking South Africans during the 70-year-long period of South African rule.

In 'The historical corpus of English in Ghana (HiCE Ghana): Motivation, compilation, opportunities', Thorsten Brato presents another type of corpus, a 'historical' corpus, which he has developed with texts stored in the university library at Frankfurt, and uses to deepen our understanding of Ghanaian English. He points out that, though 'historical', his is a synchronic and not a diachronic corpus, because its texts were written by Ghanaians during a short period, from 1966 to 1975. It can nevertheless be used for a diachronic purpose, that of comparing Ghanaian English of some fifty years ago to that of today, represented here by ICE-Ghana.

In their chapter, 'Addressing a coverage gap in African Englishes: The tagged corpus of Cameroon Pidgin English', Gabriel Ozón, Sarah Fitzgerald and Melanie Green describe yet another type of corpus, one devoted to Cameroonian Pidgin English, CPE. This is a lingua franca in the anglophone western part of Cameroon but it is also commonly used in cities throughout the country. A problem for the compilation of a corpus of Pidgin texts is the lack of a generally accepted orthography. The authors adopt devices of their own, such as distinguishing the homophones pronounced /fo/ as *foe* (a preposition or particle equivalent to standard English 'for') and *fo* ('four'), respectively. Corpus searches show that certain high-frequency verbs feature in Cameroonian Pidgin, such as *goe*, *meik* and *get*, and some of these participate in serial verb constructions (SVCs), such as *jos ron goe* ('just run away'). Moreover, the tagging of the corpus makes it possible to distinguish and quantify instances of a verb like *goe* functioning as a serial verb or alternatively as a lexical verb, the latter being more frequent.

A corpus of Cameroonian English, as distinct from CPE, was compiled by Daniel Nkemleke and is described in the next chapter, 'Practical corpus linguistics: Designing and exploiting a written corpus for research with special reference to Cameroon English.' Nkemleke uses it chiefly to investigate the frequency and the meanings of the modal verbs, especially *must* and *should*, and finds that these are less frequent in Cameroonian than in British English, with more than half the instances of *must* occurring in just one text type. Distinguishing 'root' and 'epistemic' meanings of modals, he also finds that the first is more common than the second, as in British English, only more so. One observation worth making here is that while the coming of CL has put the collection of linguistic data on a firmly objective basis, subjectivity still plays a role. Thus, ultimately, the corpus itself, even though it provides contexts of use of a particular modal, cannot pronounce on the meaning of the modal in a particular context, to decide whether it is 'root' or 'epistemic'; this is a task for an alert researcher – as it has always been.

Van Rooy in the first of the part II chapters ('Evaluating explanations for past-time reference with unmarked verb forms in African Englishes') investigates the frequency

of and possible explanations for the use of unmarked verb forms with past-time reference in Nigerian English and in Black South African English (BSAE). He does so using ICE-Nigeria for the former and de Klerk's (2006) corpus of Xhosa-English to represent the latter. The non-use of past-tense forms in non-native varieties of English is a well-known phenomenon. Explanations that have been offered for its occurrence in BSAE are that it represents the 'conversational historic present'; the use of time adverbials make past-tense marking unnecessary; phonological reduction; and level of education. Van Rooy finds that these explanations are not borne out by the corpus data. Thus in both varieties only a few unmarked verb forms occur with accompanying adverbials, and the presence of a time adverbial is actually 'associated with a higher likelihood of overt past tense inflection of the verb'. Moreover, the frequency of the non-use of such marking is rather low (20–22%). It should therefore not be regarded as a 'typical' (and so perhaps ultimately acceptable) feature of the varieties concerned.

Ulrike Gut and Foluke Olayinka Unuabonah's chapter on the use of stance markers, 'The use of stance markers in West African Englishes', is also comparative, since it uses the three relevant ICE components to study and compare the occurrence of such markers in Nigerian English, Ghanaian English and British English. Stance markers fall into different groups, expressing 'certainty' (such as *obviously*), or 'doubt/possibility' (e.g. *I think*), 'attitude/perspective' (e.g. *fortunately*) and emphasis (*really*). It is not easy to see how *usually* features in the list at all, but it is shown to have a low rate of occurrence anyway, and in Nigerian English it is synonymous with the Nigerianism and probably much more common *most times* (for which, see Jowitt 2019). Generally the frequency of stance markers is shown to be much higher in British English than in the two West African varieties. At the same time the authors conclude that their study yields mixed results with regard to the question whether there exists one pan-West-African variety.

The late Alexander Kautzsch, in his chapter entitled 'Namibian English on the web: Lexical and morphosyntactic features in a Corpus of Namibian Online Newspapers (CNamON)', describes his *Corpus of Namibian Online Newspapers* (CNamON) and uses it to discuss whether or not Namibian newspapers provide evidence of the emergence of a distinctive English variety. The extensive amount of data presented serves no more than does the more limited evidence presented by Steigentahl in her chapter to answer the question in the positive. Thus, although Namibian English unlike South African English contains several German loanwords – a reflection of Namibia's first experience of colonial rule – most of these turn out to be personal names or place names. A few 'new' and perhaps distinctive expressions are *struggle kids* (Namibians who grew up in exile) and *born-frees* (those born in the post-independence era).

Brato's second chapter in the book, 'Lexical expansion in Ghanaian English from a diachronic perspective: A structural and semantic analysis', continues his study of Ghanaian English begun with the first, and he uses for data both HiCE (the corpus of older Ghanaian English described in the first chapter) and ICE-Ghana to investigate lexical expansion in this variety. Lexical expansion, he rightly points out, is at the heart of Schneider's (2007) Dynamic Model of the evolution of Postcolonial Englishes, and

it is marked by borrowings, word-formation processes and semantic shift. He identifies over 1,000 Ghanaianisms, and they, and borrowings in particular, are more numerous in HiCE than in ICE-Ghana, which shows that many of them have been long established. Sometimes the author's judgement as to what counts as a Ghanaianism is questionable: thus *aim at* + 'progressive' (actually the *-ing* form of another verb), with the meaning 'intend to + base form of another verb', is not distinctively Ghanaian and is standard in native-speaker varieties.

Bebwa Isingoma and Christiane Meierkord's 'Capturing the lexicon of Ugandan English' is welcome as the only chapter in the book exclusively concerned with a variety of East African English. It mentions a number of 'Ugandanisms' (not a word the authors use), under the usual categories of lexical borrowing, calques, morphological and phrasal innovations, and semantic change, the latter including *stroke* as a verb meaning 'take an additional subject of study'. *Good enough*, mentioned as a Ugandan innovation and meaning 'luckily', also occurs in Nigerian English. The chapter is also useful for stressing the importance of 'multi-method approaches' – the use of newspapers, for example – 'that add to pure corpus analyses', and may provide some useful balance to such analyses. Thus the borrowing *empaako*, meaning 'praise name', is over-represented in ICE-Uganda because many tokens of it occur in just one academic paper, and hardly occurs in the authors' own 'web-based' corpus of Ugandan English.

Adeyemi Iyabo in her chapter at the beginning of part III, 'A corpus-based analysis of conjunctive cohesion in English essays of Nigerian university learners', studies 'conjunctive cohesion' in the writing of Nigerian university students, with data supplied by a special *Nigerian Learner English Corpus* (NLEC). She uses it to search for what Halliday & Hasan (1976) call 'conjunctions', a class of words that includes conjunctions in the traditional sense but also the 'conjuncts' of the Quirk *et al.* (1985) taxonomy of adverbials; she therefore identifies 'additive', 'causal', 'adversative' and 'temporal' conjunctions. Nigerian usage is compared with native-speaker usage as found in the *Louvain Corpus of Native English Essays* (LOCNESS). The results suggest that Nigerian students use a greater number of additive and causal conjunctions than British and American students, and a much smaller number of adversative conjunctions, with *but* accounting for over 60 per cent of the adversative tokens in NLEC; the frequency of temporal conjunctions is about the same in the two corpora. One additive not mentioned, though it often features in Nigerian usage, is *more so* with the same meaning as *moreover*. The author urges that Nigerian students need to use more adversatives in order to enhance the logicity of their texts.

Using a corpus of Cameroonian MA students' theses, Josef J. Schmied in his part III chapter, 'African corpora for standards in African academic English: Case studies on prepositions', addresses standards in African academic English. He rightly says that this is an urgent question. (Iyabo for one has just confirmed it.) Schmied argues that if non-native standards are adopted, and if they are to exclude common learner errors, they need to meet certain criteria, including frequent usage, international usage (i.e. the usage of non-native environments) and functional appropriateness. The discussion

focuses on prepositions in verb + preposition constructions because, Schmied says, these are a particular candidate for acceptance of variability. One good example is *discuss about*, common in Indian as well as African usage; Schmied points out that it has been stigmatized by native speakers, yet it causes no communicative problems. Questions might be asked concerning *to*, which is here treated generally as a preposition and in the African corpora used by Schmied has the highest frequency after *of*. Yet, as he points out, this is chiefly because of its frequency in *to*-infinitive constructions – and in this environment it is doubtful whether it should be classified as a preposition at all. In generative grammar since the 1980s it has been treated (also controversially) as an auxiliary.

Bassey E. Antia and Tamsyn Hendricks in the final chapter (‘Semiotic signature of transformation in a diachronic corpus of a South African political party’) are concerned to show that linguistic corpora can be used to investigate not only language variation but also broader social phenomena such as political discourse. Meanings in such discourse are moreover created not only by the central elements of language, but also by other areas of study such as onomastics and code-switching. The authors’ case study is a South African political party, the Democratic Alliance (DA), formerly known as the Democratic Party, and for it they used a corpus made up of campaign speeches, parliamentary debates and so on, and also party membership lists. The party has been in opposition to the dominant African National Congress (the ANC) and labelled by it as White- and male-dominated; but the authors show that it has become more Black-oriented and female-oriented and so more inclusive. Thus its spokespersons in their speeches now employ fewer Afrikaans words and more words from Black languages.

Collectively the chapters in this book contribute enormously to our knowledge and understanding of African varieties of English. They also demonstrate the great uses that can be made of corpora, especially when they are used together with other sources of data. If the book stimulates a higher level of investigation on the part of researchers in Africa, it will have achieved its main objective.

Reviewer’s address:

Department of English

University of Jos

P.M.B. 2084 Jos

Plateau State 930001

Nigeria

davidjowitt2003@yahoo.com

References

- de Klerk, Vivian. 2006. *Corpus linguistics and World Englishes: An analysis of Xhosa-English*. London: Continuum.
- Halliday, Michael & Ruqaiya Hasan. 1976. *Cohesion in English*. London: Longman.
- Jowitt, David. 2019. *Nigerian English*. Boston and Berlin: De Gruyter Mouton.

-
- Marungudzi, Thomas. 2017. Selected morphosyntactic features and variety status of Zimbabwean English: A corpus-based approach. PhD dissertation, North-Western University.
- Quirk, Randolph, Sidney Greenbaum, Geoffrey Leech & Jan Svartvik. 1985. *A comprehensive grammar of the English language*. London: Longman.
- Schneider, Edgar. 2007. *Postcolonial English: Varieties of English around the world*. Cambridge: Cambridge University Press.

(Received 31 March 2020)

doi:[10.1017/S1360674320000180](https://doi.org/10.1017/S1360674320000180)

Kristian A. Rusten, *Referential null subjects in Early English* (Oxford Studies in Diachronic and Historical Linguistics). Oxford: Oxford University Press, 2019. Pp. 272. ISBN: 9780198808237.

Reviewed by Nikolaos Lavidas, National and Kapodistrian University of Athens

Kristian Rusten's *Referential Null Subjects in Early English* provides significant evidence of the bidirectional relationship between historical linguistics (here, historical syntax) and diachronic corpora: it shows how the theory of historical syntax can be the basis for and the starting point of diachronic corpora investigations, but also how quantitative studies can trigger new directions for modern approaches to historical syntax. Rusten's study is a development of his University of Bergen PhD dissertation (2015) and succeeds perfectly in testing all major hypotheses on the status of null subjects in Old English (OE) and their development. Most importantly, Rusten manages to present a clear empirical analysis of null subjects in English, as well as of various relevant approaches according to several models, which serve as a stable basis for sophisticated quantitative research. The theoretical discussion and analysis of relevant examples and the main characteristics of null subjects combine coherently with the empirical surveys presented in the book. The well-prepared structure of the book, the well-presented research questions and methodology, the clear short summaries in various parts and the detailed discussion of all results and of their theoretical implications support a clear presentation of Rusten's main proposal that no dialect of OE had null subjects of the type that pro-drop languages have, but null subjects in OE and their development are related to a remnant of an earlier productive system and are licensed by argument ellipsis.

Rusten's study can be seen as a model study for corpus-based investigations that also rely on state-of-the-art statistical methods: Rusten provides an in-depth quantitative analysis of 217,239 overt and null referential pronominal subjects (referential subjects that 'are left unexpressed not due to conjunction reduction', p. 16) that are attested in 509 OE, Middle English (ME) and Early Modern English (eModE) texts. Rusten clearly states that he investigates the referential subjects of OE that are not compatible with the rules of Present-day English (PDE) and excludes expletive null subjects, cases