

# 1. English around the Globe: Local and Global Perspectives on Social and Regional Variation

Paula Rautionaho, ORCID 0000-0002-5239-8407

Hanna Parviainen, ORCID 0000-0002-0723-5692

Mark Kaunisto, ORCID 0000-0002-0084-1067

Arja Nurmi, ORCID 0000-0002-6446-9356

## Abstract

The World Englishes research paradigm focuses on processes of globalization and the linguistic change and variation resulting from the new contexts of use and the multitude of different backgrounds its speakers today have. Importantly, taking different levels of variation - intra-linguistic, regional, and social variation, as well as stylistic variation - into account offers us an opportunity to a more comprehensive understanding of both stable variation and the linguistic change springing from it. While each of the types of variation merits research on its own, in order to better understand their nature, it is beneficial to bring these different strands of linguistics together to match and compare findings and techniques.

Grammatical variation in English is a multi-faceted phenomenon, tied to space and time and yet constantly on the move. Grammatical constructions develop according to need and usage, but also as a result of short-lived fancies. From its insular origins, the English language has evolved into a complex set of different kinds of Englishes spoken and written around the world, be they native or non-native varieties, spoken by learners or facilitating global communication as a lingua franca. This global proliferation of the language has led to an increasing rate of change and variation, as English finds new contexts of use and comes into contact with speakers from a multitude of different language backgrounds using the language for a variety of purposes and tasks. The results of these globalization processes are the object of study in the World Englishes paradigm; the study of World Englishes involves the identification of different varieties of English, of converging and diverging features between them, and the analysis of how the varieties have been and are being shaped by their sociolinguistic histories and their multicultural backgrounds. The plural form

“Englishes”, rather than the singular “English”, stresses the fact that English no longer has one major variety that all speakers of the language consider the standard, a source of authority and prestige (Mesthrie and Bhatt 2008, 3). Following McKay (2010, 91), among others, the World Englishes paradigm is regarded as placing “all varieties of English on par with each other without any one being a reference point”. In the present volume, the term World Englishes is understood as referring to the full range of Englishes, be they native, non-native, or learner varieties.

World Englishes can be approached from a local or a global perspective. As Filppula, Klemola, and Sharma (2017, 3) point out, there is now sufficient corpus evidence to explore World Englishes from not only comparative but also theoretical perspectives. The local perspective focuses on one variety at a time, giving a detailed account of the intra-varietal variation or change studied, whereas the global perspective tackles a number of varieties at the same time, aiming at a detailed comparison of variation or change as it occurs in those varieties. While the majority of studies taking a local perspective focus on the major varieties that have native speakers, such as American English, they can just as well focus on a non-native variety, such as Hong Kong English. The global perspective, on the other hand, often juxtaposes native varieties with non-native or learner varieties, resulting in a birds-eye-view of the variation within the varieties included. At the same time, traditional ways of comparing native and non-native varieties have some persistent blind spots. Just the focus on discovering differences rather than also identifying similarities can obscure a great deal of the big picture (Hansen 2018, 49). There is also often an assumption that not only are changes spreading from native varieties to non-native ones, but that the associated extralinguistic forces behind these changes in different societies will be the same (see eg., Hansen 2018, 47). A great deal of work needs to be carried out to find the trends that are truly global in Englishes and to distinguish those from regional or local ones.

Another set of perspectives that can be taken in studies of language variation and change is social vs. regional variation. Studies focusing on social variation are tuned in on the effects that social parameters, such as a speaker’s age, gender or social status, have on language use, while studies focusing on regional variation are interested in its geographical or socio-political causes. Regional variation can be approached from both local and global perspectives, depending on the geographical area in focus. Social variation, on the other hand, is so far more or less confined to the local level – the study of social variation requires, first of all, metadata on the speakers, which in many cases is

not yet available for less central varieties. Secondly, the metadata should be comparable, which is not a straightforward matter when it comes to comparing varieties spoken in different societies which may be based on highly diverging social class systems, for instance. The cross-fertilization of social variation on a global scale is gaining momentum, however, as more and more detailed metadata is being collected for World Englishes on a larger scale.

Bringing the different understandings of linguistic variation into one discussion is important. Taking intra-linguistic variation, regional variation, and social variation, as well as stylistic variation into account allows us to gain a more comprehensive understanding of both stable variation and the linguistic change that is rooted in it. While there is great merit in studying each of the types of variation on its own, in order to better understand their nature, it is also useful to find moments of bringing these different strands of linguistics together, to match and compare findings and techniques.

As Sharma (2017, 252) points out, there is still a great deal of work to be done in building bridges between the study of World Englishes and sociolinguistic theory. There is a great deal of work still to be done in studying internal variety in these varieties (Hansen 2018, 47). Just as the history of English has provided a proving ground for stratificational sociolinguistics (see e.g., Nevalainen and Raumolin-Brunberg 2017; Hernández-Campoy and Conde-Silvestre 2012), the study of indigenized Englishes can allow us to test the general applicability and theoretical robustness of accepted sociolinguistic principles, such as uniform evaluation, the linguistic conformity of women, the apparent time hypothesis, or the role of identity, to mention a few (Sharma 2017, 233–234). Much of the mainstay of sociolinguistic work has risen in environments where the assumptions include monolingual speakers and speech communities, a Western urban culture, and non-mobile groups. These are all to some extent incompatible with the realities of most World Englishes.

As we approach World Englishes with the tools of sociolinguistics, we need to be careful to apply parameters in a way which is consistent with the society being studied. Any depictions of class are culture-specific and each culture needs its own description of social status. There is a concomitant need to evaluate how the social structure of a given society changes our assumptions of, for instance, everyone aiming at a shared standard, which can be problematic particularly in bi- and multilingual contexts for bi- and multilingual speakers (Sharma 2017, 238–239). The style aimed

for in more formal speech can also differ from the expected norm: not for instance aiming for Inner Circle prestigious norms but for local prestige norms (Sharma 2017, 240–241). Gender roles are similarly different in different societies, and particularly in contexts of migration can shift in the process of acclimatization, and the role of peers and family can differ distinctly from established sociolinguistic tenets when observed in non-Western communities (Sharma 2017, 240–242). The role of multilingual communities shows up also in the case of the apparent time hypothesis as it seems to relate to assumptions about stable monolingual communities (Sharma 2017, 243). Multiethnic communities can also contribute towards atypical communicative practices and the formation of new lects in the heart of supposedly native-speaking Inner Circle areas, such as Multicultural London English (Bullock et al. 2017, 222; Cheshire et al. 2008). All this brings to light the vital need to approach all World Englishes from the perspective of social variation.

The question of multilingual communities and the impact of substrate languages is important for both English as a Second Language and English as a Foreign Language varieties. For ESL varieties in particular, the questions of who actually speaks English in that society and what is the gamut of languages spoken as first language is highlighted particularly in cases like Indian English, where it can be suggested that there are many Indian Englishes rather than just one (Lange 2017, 32–33). While the *International Corpus of English* has provided useful first steps, a new generation of corpora will be needed to account for this type of variation. Incorporating such complexities into variationist corpus-based and sociolinguistically informed research is in many ways still to come.

In recent years research into variationist linguistics has greatly benefitted from the application of advanced statistical techniques, allowing us to understand the more complex patterns of variation in the increasingly complex data we approach. Particularly in the context of applying sociolinguistic parameters to the complex data of World Englishes, being able to see patterns through the help of modern methods is a benefit.

This thematic volume focuses on the evolution and present state of a number of grammatical features in Englishes spoken on the British Isles and beyond; the corpus linguistic studies included in the volume investigate grammatical variation in varieties such as British, Irish and American Englishes, Indian, Singaporean and Hong Kong Englishes, Nigerian, Ghanaian and South African Englishes, as well as Englishes spoken in Finland, Sweden, Japan, and China. This combination of

local and global perspectives allows us to investigate divergent and convergent trends in the English language complex, and the extent to which the catalysts for change are shared in time and space. The features investigated include modal auxiliaries, complementation patterns, prepositions, and forms of BE, among others, and the themes addressed range from social patterns to social networks, from contact linguistics to geographical proximity, and from individual grammatical features to alternations.

The nine chapters in this volume track patterns of change in the past, the present and the future. Diachronic patterns in the past go back to Late Middle English (Nevalainen) but focus also on recent trends in the past decades in native varieties (Rautionaho and Kaunisto), as well as in contact varieties (Filppula, Fuchs). Present-day variation is charted in varieties ranging from Africa (Romasanta) to Europe (Paulasto and Meriläinen), with each of the chapters taking a pan-world approach to grammatical variation. Future directions are mapped with more methodological approaches into new kinds of data (Ronan, Laitinen, and Fatemi) and new avenues of corpus linguistics (Hoffmann, Arndt-Lappe, and Uhrig). Together the chapters provide a fresh and detailed view of variation in the local and the global sphere, with the help of corpus linguistics and varied statistical methods. The chapters are introduced below.

Nevalainen approaches a central question of variationist linguistics, discussing whether social or situational variation is more significant. She builds bridges between modern sociolinguistic studies starting from interview data and corpus linguists working with both written and spoken data. Due to the long diachrony available for comparisons, she is able to show through her metastudy that the phase of a change is relevant: for earlier phases social variation is more relevant while for the final phases register variation takes over.

Filppula examines the variation between the uses of the semi-auxiliaries *have to* and *have got to*, describing the emergence and differences in usage of the two structures in British and Irish English. He studies corpus data of both written and spoken language data representing standard and non-standard varieties, and compares the findings to those in Indian English and Hebridean English, detailing the different paths that British and Irish varieties have taken with regard to the use of the variant structures.

Rautionaho and Kaunisto focus on recent spoken British English, taking advantage of the two versions of the *British National Corpus* (1990s and 2014). They study the variation of *was* and *were* with pronoun subjects *we*, *you*, and *they*, showing that while both intra-linguistic and sociolinguistic parameters play their role, it is the sociolinguistic variation in terms of age and social class that better explains the patterns of variation observed.

Romasanta's chapter looks into the clausal complementation patterns of the matrix verb *regret*, which allows both finite *that*-clauses and non-finite *to*-infinitival clauses as its complements. Through multidimensional aggregational analyses and random forest analyses, her study investigates the extra- and intra-linguistic factors affecting the choices between the patterns in seven varieties of English. The postcolonial varieties examined show a preference for the use of finite patterns, suggesting that the principle of maximization of transparency in these varieties explains the distribution between the patterns.

Paulasto and Meriläinen examine the omission of standard English prepositions across different types of Englishes, with the aim of charting the contexts where the omission occurs and identifying commonalities and differences between omission patterns across native, second-language, and foreign-language varieties. With their corpus approach, Paulasto and Meriläinen are able to show that the range of contexts for preposition omission is much broader than documented in earlier studies relying on smaller sets of data, and that L1 influence, semantic redundancy as well as cognitive processes of non-native language use influence the omission patterns in various ways.

Fuchs uses a number of different corpora to investigate the use of the negative scalar conjunction *and that too* in Asian varieties of English. His results indicate that this feature, which has become increasingly rare in British English, has remained in continuous use in Indian English up to the present and thus, the feature is a case of colonial lag (also called "feature retention"). Furthermore, the results demonstrate existing usage, although to a more varying degree, also in other South Asian varieties and beyond, to English varieties spoken in South-East Asia.

Ronan presents a qualitative and quantitative study on the use of the expressive marker *too bad* across different genres and time using various corpora. The results of her study show that the

feature first appeared in US film and TV discourse during the latter half of the 20th century, after which it spread to both new informal genres and to varieties of English spoken around the world.

Laitinen and Fatemi present an algorithmic method for adding social information to Twitter data, which is crucially lacking in social metadata, with the aim of allowing the study of diffusion of language change. They argue that using participant-centred information, such as size and structure of social networks, as a proxy for social information improves the validity of social media data, and present a case study of how networks of varying strength in the UK and the Nordic region condition ongoing change in modal usage.

The chapter by Hoffmann, Arndt-Lappe and Uhrig explores the possible effects that the rhythmic characteristics of different L1 languages may have on the production of English bigrams by L2 speakers of English. Examining bigrams in the 20 subcomponents of the 1.9-billion-word GloWbE corpus, the authors find support to their hypothesis that speakers from syllable-timed linguistic backgrounds do not adhere to the same kind of rhythmic constraints in English as speakers whose first language is stress-timed.

The studies in the chapters showcase a wide spectrum of application of analytical methods, involving the use of different types of statistical analyses as well as databases of different sizes. In some studies, the authors have made use of previously available corpora, but some chapters also involve the use of self-compiled sets of data. The useability of different types of datasets when studying variation globally is in itself a challenging issue, and a good case in this regard is the study involving Twitter data in Chapter 9. A list of the corpora and databases used in this volume is as follows:

- *British National Corpus* (1994 version, the demographically sampled section; Chapter 4)
- *British National Corpus 2014* (Chapter 4)
- *Corpus of Contemporary American English* (Chapter 8)
- *Corpus of Global Web-based English* (Chapters 5, 7, 8, and 10)
- *Corpus of Historical American English* (Chapter 8)
- *Corpus of Early English Correspondence* (Chapter 2)
- “The Cregeen Texts” (the island of Tìree, the Inner Hebrides; Chapter 3)

- *Google Books* corpus of British English (Chapter 7)
- *International Corpus of English* (spoken language sections from the Great Britain, Ireland, India, Singapore, and Nigeria components of the corpus; Chapters 3, 6, and 7)
- *International Corpus of Learner English* (subcorpora compiled in Finland, Sweden, Japan, and Hong Kong; Chapter 6)
- *Kolhapur Corpus* (Chapter 7)
- *Louvain Corpus of Native English Essays* (Chapter 6)
- Llandybie and North Wales corpora (Chapter 6)
- *The Movies Corpus* (Chapter 8)
- *Northern Irish Transcribed Corpus of Speech* (Chapter 3)
- Nineteenth-century Irish emigrants' letters (Chapter 3)
- *Survey of English Dialects* tape recordings (Chapters 2 and 6)
- transcriptions of self-recorded field recordings (east/south-west of Ireland; Chapter 3)
- Twitter (Chapter 9)
- "The Sabban Corpus" (Inner and Outer Hebrides; Chapter 3)

As a final note, it is worth pointing out that in this collection of studies on variation, the authors are quite conscious of the varied nature of the more precise linguistic phenomena examined as well as the methods used in the investigation of them. We find that this heterogeneity aptly reflects the overall state of the current work in the field. While volumes focussing on, for example, questions relating to individual methodological approaches or the use of datasets may in a more concerted fashion survey the strengths and weaknesses of such points of interest, the breadth of field, we believe, will only continue to increase. Yet the choice of analytical tools very much depends on the topic under study, which means that even more traditional methods of analysis should not be neglected. The variety in this respect is knowingly showcased and accepted.

### **Reference list**

Bullock, Barbara E., Lars Hinrichs, and Almeida Jacqueline Torribio. 2017. "World Englishes, Code-switching, and Convergence." In *The Oxford Handbook of World Englishes*, edited by Markku Filppula, Juhani Klemola, and Devyani Sharma, 211–231. Oxford: Oxford University Press.

Rautionaho, Paula, Hanna Parviainen, Mark Kaunisto & Arja Nurmi (2022). English around the Globe: Local and Global Perspectives on Social and Regional Variation. In Paula Rautionaho, Hanna Parviainen, Mark Kaunisto & Arja Nurmi (eds.) *Social and Regional Variation in World Englishes: Local and Global Perspectives*. New York & London: Routledge. 1–7.



Filppula, Markku, Juhani Klemola, and Devyani Sharma. 2017. "Introduction: World Englishes and Linguistic Theory." In *The Oxford Handbook of World Englishes*, edited by Markku Filppula, Juhani Klemola, and Devyani Sharma, 3–13. Oxford: Oxford University Press.

Hansen, Beke. 2018. *Corpus Linguistics and Sociolinguistics. A Study of Variation and Change in the Modal Systems of World Englishes*. Leiden: Brill Rodopi.

Hernández-Campoy, Juan Manuel and Juan Camilo Conde-Silvestre (eds.) 2012. *The Handbook of Historical Sociolinguistics*. Malden, MA: John Wiley & Sons.

Lange, Claudia. 2017. "Indian English or Indian Englishes? Accounting for Speakers' Multilingual Repertoires in Corpora of Postcolonial Englishes." In *Challenging the Myth of Monolingual Corpora*, edited by Arja Nurmi, Tanja Rütten, and Päivi Pahta, 16–38. Leiden: Brill Rodopi.

McKay, Sandra Lee. 2010. "English as an International Language." In *Sociolinguistics and Language Education*, edited by Nancy H. Hornberger and Sandra Lee McKay, 89–115. Clevedon: Multilingual Matters.

Mesthrie, Rajend and Rakesh M. Bhatt, eds. 2008. *World Englishes: The Study of New Linguistic Varieties*. Cambridge: Cambridge University Press.

Nevalainen, Terttu and Helena Raumolin-Brunberg. 2017. *Historical Sociolinguistics: Language Change in Tudor and Stuart English*. London: Routledge.

Sharma, Devyani. 2017. "World Englishes and Sociolinguistic Theory." In *The Oxford Handbook of World Englishes*, edited by Markku Filppula, Juhani Klemola, and Devyani Sharma, 232–251. Oxford: Oxford University Press.