

# Corpus-based descriptions of certain aspects of the lexical and phonological characteristics of eight varieties of World Englishes

Leah Gilner

Aichi University

leahgilner@gmail.com

## Abstract

This presentation will provide an interim report on an ongoing investigation into corpus-based approaches to the description of world Englishes and English as a lingua franca phonology.

## Keywords

global Englishes, phonology, functional load

## Introduction

The description of the English language has become more complex and nuanced with the acknowledgment of the variation, variability, and variety that accompanies the expansion of the users and uses of the language. Accounting for the full range of its present-day forms and functions is evidently an ambitious goal given the estimated 2 billion users around the world (Crystal, 2003). Results from the investigation described in this presentation will support this endeavor by providing corpus-based lexical distribution data and usage-based descriptions of certain aspects of phonological systems that contribute to the global English chorus of voices.

## 1 Significance

The relevance of this investigation resides in the fact that linguistic description of the English language has historically been based on a limited, and limiting, perception of the object of inquiry framed within conventions and behaviors claimed to be exhibited, and solely exhibited, by typically monolingual, mother tongue speakers. The rapid and extensive diversification of speakers who use English to communicate with one another has effectively overwhelmed established frameworks. This investigation approaches the task of descriptions by targeting the lexical and phonological domains and seeks

to augment currently available descriptions.

## 2 Rationale

The varieties under investigation represent English used by local speech communities in the following regions of the world (in alphabetical order): Canada, East Africa, Hong Kong, India, Ireland, Jamaica, the Philippines, and Singapore. Each of these locales is particular in its historical, social, political, and economic profiles. The role of the English language within each milieu is similarly particular, acquiring additional differentiation as each society follows its own evolutionary trajectory. Furthermore, various languages cohabit in most of these environments and, naturally, multiple registers coexist in all of them. In this manner, the linguistic habitus of the members of each community of speakers reflects an intricate and idiosyncratic interplay of a wide array of linguistic, social, interactive, and cognitive factors (Blommaert & Backus, 2012; Blythe & Croft, 2009; Canagarajah, 2018; Mufwene & Vigouroux, 2012; Van Rooy, 2010; Wei, 2018). Descriptive endeavors like the described in this presentation should promote speculation and theorization about contemporary language use.

## 3 Methodology

The analytical approach can be described as usage-based because it targets not only structure but "two other important aspects of language phenomenon – the material content or substance of language, and language use" (Bybee, 2001, p. 2). The realization of this methodology starts by identifying the words that dominate discourse, effectively isolating the lexical mass of the communicative events contained in the corpora under investigation. It culminates in the

examination of the phonemic characteristics of those words, using functional load (FL) to quantify the systemic relationships that emerge from the lexicon.

The FL framework quantifies the relative amount of work that elements in a linguistic class do in relation to each other by marrying structural descriptions with frequency information. In this way, FL measures add granularity to structural descriptions because they further qualify linguistic constructions in terms of usage and thus provide an indication of their communicative weight. This approach can be seen as a way to "bring linguistics research closer to living reality" by unifying the Saussurean notions of parole and langue (Mathesius, 1929, p. 138).

Operationalizing the construct of FL necessitates two types of datasets: 1) corpus-based lexical distribution metrics and 2) phonological representations of lexical datasets. At present, the datasets available for the English language are largely confined to British and American varieties. The current investigation will produce both types of datasets for eight other varieties of English thereby accounting for a substantial portion of the potential pool of ELF interactants, approximately half of the estimated 2 billion (Crystal, 2003). Furthermore, the analytical framework will establish a replicable precedent that will facilitate subsequent analyses of phonological systems, including that of individuals.

#### 4 Potential applications

Results like the ones provided by this investigation are becoming increasingly relevant as functionalist, usage-based perspectives of language, interaction, and cognition become more widely accepted among scholars (Hruschka et al., 2009; Mauranen, 2018; The Five Graces Group et al., 2009). Moving from speculation and theorization to assertion and explanation requires observation and empirical validation. The outcome of the analytical approach adopted here will supplement structural descriptions of the phonemic inventories of eight varieties of English with FL quantifications. Results will potentially shed light on the internal functional organization of each phonological system and further understanding regarding the role of phonology in mental representations and speech processing.

#### 5 References

- Blommaert, J., & Backus, A. (2012). Superdiverse repertoires and the individual. In I. de Saint-Jacques & J.-J. Weber (Eds.), *Multimodality and Multilingualism: Current Challenges for Educational Studies* (pp. 11–32). Rotterdam: Sense Publishers.
- Blythe, R. A., & Croft, W. A. (2009). The Speech Community in Evolutionary Language Dynamics. *Language Learning*, 59, 47–63.
- Bybee, J. (2001). *Phonology and language use*. Cambridge, [England]; New York: Cambridge University Press.
- Canagarajah, S. (2018). Translingual Practice as Spatial Repertoires: Expanding the Paradigm beyond Structuralist Orientations. *Applied Linguistics*, 39(1), 31–54.
- Crystal, D. (2003). *English as a global language*. Cambridge, UK; New York: Cambridge University Press.
- Hruschka, D. J., Christiansen, M. H., Blythe, R. A., Croft, W., Heggarty, P., Mufwene, S., ... Poplack, S. (2009). Building social cognitive models of language change. *TICS Trends in Cognitive Sciences*, 13(11), 464–469.
- Mathesius, V. (1929). Functional linguistics. In J. Vachek & L. Dušová (Eds.), *Praguiana* (pp. 121–142). Academia, Publ. House of the Czechoslovak Academy of Sciences, Prague: John Benjamins.
- Mauranen, A. (2018). Second Language Acquisition, world Englishes, and English as a Lingua Franca (ELF). *World Englishes*, 37(1), 106–119.
- Mufwene, S., & Vigouroux, C. (2012). Individuals, populations, and timespace: Perspectives on the ecology of language. *Linguistique Cahiers de Linguistique*, 38(2), 111–137.
- The Five Graces Group, Beckner, C., Blythe, R., Bybee, J., Christiansen, M. H., Croft, W., ... Schoenemann, T. (2009). Language Is a Complex Adaptive System: Position Paper. *Language Learning*, 59, 1–26.
- Van Rooy, B. (2010). Social and linguistic perspectives on variability in world Englishes. *World Englishes*, 29(1), 3–20.
- Wei, L. (2018). Translanguaging as a Practical Theory of Language. *Applied Linguistics*, 39(1), 9–30.