English (mor)phonotactics in first language acquisition: A corpus study

Paulina Zydorowicz
Adam Mickiewicz University

Phonotactics investigates permissible sound combinations in a language. English is a phonotactically rich language, with a subset of clusters violating available scales predicting the goodness of a cluster, e.g. *Sonority Sequencing Generalisation* (SSG; Selkirk 1984) or the *Net Auditory Distance* principle (NAD; Dziubalska-Kołaczyk 2009).

Traditionally, the goodness of clusters has been determined by SSG (Selkirk 1984). The principle entails measurements of distances between consonants expressed by the manner of articulation, which reflects the degree of aperture of the vocal tract. The SSG states that sonority of adjacent segments should decrease from the nucleus outward, and clusters which obey this pattern are deemed unmarked. An alternative approach to cluster evaluation has been proposed by Dziubalska-Kołaczyk (2009). The *Net Auditory Distance* principle formulates universal preferences for optimal clustering, depending on the length of a cluster and its word position. Cluster quality (preferred or dispreferred) is based on *three* criteria of consonant description: manner and place of articulation as well as the distinction between an obstruent and a sonorant in a sequence. The condition for a preferred word-initial or -final CC cluster states that the distance between the two neighbouring consonants (C_1C_2) must be greater than (or at least equal to) the distance between the vowel and the neighbouring consonant. Clusters which violate the NAD condition(s) are considered dispreferred. The second dimension of the study is the inclusion of morphology as a potential factor modulating cluster production rates (*past* vs *passed*) (cf. Dressler and Dziubalska-Kołaczyk 2006).

The goal of this contribution is to investigate (mor)phonotactic acquisition by English children. The first hypothesis predicts that preferred clusters will obtain lower reduction rates than dispreferred clusters. Secondly, it is predicted that clusters spanning a morphological boundary will be retained in production better than lexical clusters as the former encode important morphological information. The hypotheses will be tested empirically in a corpus-based study (CHILDES; MacWhinney, 2000; Rose & MacWhinney, 2014; Providence corpus by Demuth et al. 2006). A mixed effects logistic regression model will be fitted to the dataset with the lme4 package (Bates et al. 2015) in R (R Core Team 2018) to examine factors affecting cluster rendition (intact or modified) including markedness (NAD), word position, cluster size, the presence of a morphological boundary, logarithmic frequency, word length, and age. It is hoped that the study will enrich the crosslinguistic data pool regarding phonotactic acquisition.

References

Bates, Douglas, Martin Mächler, Ben Bolker and Steve Walker (2015), Fitting Linear Mixed-Effects Models Using lme4., *Journal of Statistical Software* 67(1), 1-48.

Demuth, Katherine, Jennifer Culbertson, and Jennifer Alter. (2006). Word-minimality, epenthesis, and coda licensing in the acquisition of English, *Language & Speech* 49, 137-174.

Dressler, Wolfgang U. and Katarzyna Dziubalska-Kołaczyk. (2006), Proposing Morphonotactics, *Rivista di Linguistica* 18 (2), 249-266

Dziubalska-Kołaczyk, Katarzyna (2009), NP extension: B&B phonotactics, *Poznań Studies in Contemporary Linguistics* 45 (1), 55-71.

Dziubalska-Kołaczyk, Katarzyna (2014), Explaining phonotactics using NAD, *Language Sciences* 46, 6-17.

MacWhinney, Brian (2000), *The CHILDES Project: Tools for analyzing talk*, Third Edition. Mahwah, NJ: Lawrence Erlbaum Associates.

R Core Team (2018), R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing. (https://www.R-project.org)

Rose, Yvan and Brian MacWhinney (2014), The PhonBank Project: Data and software-assisted methods for the study of phonology and phonological development, in J. Durand, U. Gut and G. Kristoffersen (eds), (2014), *The Oxford handbook of corpus phonology*, Oxford, UK: Oxford University Press, 380-401.

Selkirk, Elisabeth O. (1984), On the major class features and syllable theory, in M. Aronoff & R. T. Oehrle (eds), (1984), *Language sound structure*, Cambridge, MA: MIT Press, 107-136.