THE PREVALENCE OF SYSTEMATICITY IN INDO-EUROPEAN LANGUAGES

Nicole Tamer[*1,2], Paul Widmer[1,2]

[*]Corresponding Author: Nicole.tamer@uzh.ch

[1]Department of Comparative Language Science, University of Zurich, Zurich, Switzerland

[2]Center for the Interdisciplinary Study of Language Evolution, University of Zurich, Zurich, Switzerland

Human language has long been thought to be largely arbitrary, meaning that no intrinsic or logical connection exists between form and meaning (Hockett, 1960; de Saussure, 1983). Recent studies have revealed new insights by applying novel statistical methods, demonstrating the prevalence, and increasing relevance of non-arbitrary forms. These forms are commonly divided into iconicity, the resemblance-based mapping of form and meaning, (Winter & Perlman, 2021; Nielsen & Dingemanse, 2021) and systematicity, phonological cues predictive of grammatical categories (Raviv & Arnon, 2018; Pimentel et al., 2019; Nölle et al., 2018). An example of systematicity is the phonological distinction between open and closed word classes, the former commonly accept new words, whereas the latter does not, and within the open word class (Monaghan et al., 2007).

Because systematicity focuses on the relationships of large numbers of words to a small number of abstract categories, rather than the more frequent research on the relationships of single words to simple referential meanings, it has received less attention in the context of language evolution (Dingemanse et al., 2015). Languages tend to only exhibit subtle aspects of iconicity and these tend to decline over time in favour of more systematic or arbitrary forms (Little et al. 2017). Understanding the prevalence and mechanisms of systematicity is crucial to reveal more about its roots (cognitive advantages such as ease of processing, learnability, and acquisition advantages: Raviv et al., 2021; Monaghan et al., 2012; Fitneva et al., 2009) and role in the emergence and evolution of large grammatical and lexical inventories.

Corpus studies have revealed that some languages use subtle phonological and prosodic cues (e.g. stress, duration, voicing or phonotactics: Monaghan, et al., 2007) to differentiate between word classes and categories (Kelly 1992; Monaghan et al., 2005). However, previous studies included only a small sample

heavily biased towards modern Western European languages. These factors constrain the ability to generalise these findings to other languages which limits broader inference about the role of systematicity in language evolution. To address this research gap and carry out an in-depth analysis of the prevalence of systematicity, we collected grammatical information of 30 modern and ancient languages extracted from language-specific grammars, corpora, and dictionaries. Specifically, we examined phonological cues towards grammatical categories of the first phoneme to capture the initial word recognition advantages provided by systematicity (Tamariz, 2008; Dingemanse et al. 2015).

A simple Shannon entropy was calculated within each grammatical category. The significance was tested by bootstrapping the original distribution of tokens to create new populations of entropy measures which provided information about expected entropy and the uncertainty of the measure.

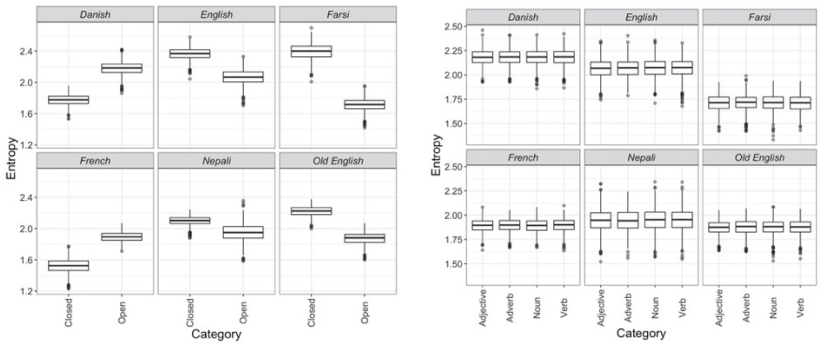

Figure 1. Entropy of the first word segment of closed and open word classes (left) with the highest differences and between grammatical categories of the open word class in a sample of languages.

We found significant differences in entropy measures of phoneme variability between open and closed word classes for every language across the Indo-European clades. No significant difference in variability was observed within the open word class itself when, for example, nouns were compared to verbs or adjectives with adverbs contrary to results of previous studies (Monaghan et al., 2007; Dingemanse et al., 2015).

We therefore conclude that a language-general pattern exists for differential variability in first word segment phoneme distribution between open and closed word classes. Intriguingly, this effect is not driven by variability between grammatical categories within the open word class contrary to previous findings. These findings demonstrate the critical and hitherto underestimated role that systematicity plays in the development and stability of certain grammatical macro-categories. Previous studies have not captured these aspects of systematicity, and these findings demonstrate how language behaves not only across the Indo-European language family but also within grammatical categories further highlighting the prevalence of non-arbitrariness in a sample of Indo-European languages.

References

De Saussure, F. (1983). *Course in General Linguistics*, Open Court.

Dingemanse, M., Blasi, D. E., Lupyan, G., Christiansen, M. H., & Monaghan, P. (2015). Arbitrariness, iconicity, and systematicity in language. *Trends in cognitive sciences*, *19*(10), 603-615.

Fitneva, S. A., Christiansen, M. H., & Monaghan, P. (2009). From sound to syntax: Phonological constraints on children's lexical categorization of new words. *Journal of child language*, *36*(5), 967-997.

Garrod, S., Fay, N., Lee, J., Oberlander, J., & MacLeod, T. (2007). Foundations of representation: where might graphical symbol systems come from?. *Cognitive science*, *31*(6), 961-987.

Hockett, C. F., & Hockett, C. D. (1960). The origin of speech. *Scientific American*, *203*(3), 88-97.

Kelly, P. (1992). Does the ear assist the eye in the long-term retention of lexis?. *IRAL: International Review of Applied Linguistics in Language Teaching*, *30*(2), 137.

Little, H., Rasilo, H., Van Der Ham, S., & Eryılmaz, K. (2017). Empirical approaches for investigating the origins of structure in speech. *Interaction Studies*, *18*(3), 330-351.

Monaghan, P., Mattock, K., & Walker, P. (2012). The role of sound symbolism in language learning. *Journal of experimental psychology: Learning, memory, and cognition*, *38*(5), 1152.

Monaghan, P., Christiansen, M. H., & Chater, N. (2007). The phonological-distributional coherence hypothesis: Cross-linguistic evidence in language acquisition. *Cognitive psychology*, *55*(4), 259-305.

Monaghan, P., Chater, N., & Christiansen, M. H. (2005). The differential role of phonological and distributional cues in grammatical categorisation. *Cognition*, *96*(2), 143-182.

Nielsen, A. K., & Dingemanse, M. (2021). Iconicity in word learning and beyond: A critical review. *Language and Speech*, *64*(1), 52-72.

Nölle, J., Staib, M., Fusaroli, R., & Tylén, K. (2018). The emergence of systematicity: How environmental and communicative factors shape a novel communication system. *Cognition*, *181*, 93-104.

Pimentel, T., McCarthy, A. D., Blasi, D. E., Roark, B., & Cotterell, R. (2019). Meaning to form: Measuring systematicity as information. *arXiv preprint arXiv:1906.05906*.

Raviv, L., de Heer Kloots, M., & Meyer, A. (2021). What makes a language easy to learn? A preregistered study on how systematic structure and community size affect language learnability. *Cognition*, *210*, 104620.

Raviv, L., & Arnon, I. (2018). Systematicity, but not compositionality: Examining the emergence of linguistic structure in children and adults using iterated learning. *Cognition*, *181*, 160-173.

Tamariz, M. (2008). Exploring systematicity between phonological and context-cooccurrence representations of the mental lexicon. *The Mental Lexicon*, *3*(2), 259-278.
Winter, B., & Perlman, M. (2021). Iconicity ratings really do measure iconicity, and they open a new window onto the nature of language. *Linguistics Vanguard*, *7*(1).