Exploring Systematic Phonological Cues in Language: A Comparative Study Across 60 Languages from 13 Families

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 differs from animal communication in many respects, most prominently by having the capacity for great flexibility and arbitrariness in its expression which has evolved in the hominin lineage since speciation from the last common ancestor (Watson et al. 2022; C. F. Hockett and C. D. Hockett 1960). However, at some point in evolutionary history, non-arbitrary constraints have evolved to render language more efficient and easier to process to adapt to the needs of communication over generations by accumulating learning preferences (Kirby, Griffiths, & Smith, 2014; Motamedi et al., 2022).

This differentiation is explored in our study which focuses on the prevalence of systematicity across languages. The research questions are whether systematicity is distinguishing between word classes (open and closed) and if it is language-specific as evidenced by previous studies (Dingemanse et al., 2015). Corpus studies have revealed that some languages showcase systematic constraints, such as subtle systematic phonological cues to differentiate between word classes and phonological categories (Kelly 1992; Monaghan, Chater, et al. 2005). These cues provide cognitive advantages resulting in ease of processing, improved language comprehension and acquisition of languages (Raviv, Heer Kloots, et al. 2021; Fitneva et al. 2009; Monaghan, Christiansen, and Fitneva 2011). Understanding the diversity of systematicity is crucial in uncovering its cognitive advantages, such as enhanced memory processing, learnability, and acquisition (Raviv et al., 2021; Monaghan et al., 2012; Fitneva et al., 2009), as well as its significant role in the emergence and evolution of expansive lexical and grammatical inventories.

Recent studies using novel computational and statistical methods have underscored the increasing relevance of systematicity (Raviv & Arnon, 2018;

Pimentel et al., 2019; Nölle et al., 2018). However, many prior investigations were limited to a narrow sample predominantly biased towards modern Western European languages or analyzed only a limited number of words, constraining the generalizability of findings to other linguistic contexts. To address this gap, we conducted an extensive analysis encompassing grammatical data from 40 modern and ancient Indo-European languages, alongside 20 languages belonging to 12 distinct language families. The data was compiled from language-specific corpora, grammars as well as comparative language data bases. Specifically, we scrutinized phonological cues pertaining in the initial phoneme, thereby capturing the initial word recognition advantages conferred by systematicity (Trott et al., 2019; Tamariz, 2008).

With a Bayesian logistic regression model, we investigated the relationship between phonological cues and systematicity. A strong amount of systematicity is defined as the data points aggregating in the upper quantile of either open or closed class. The posterior probability results show how much evidence there is for a cue within a language occurring either above or below zero. Phonological cues with posterior probability values close to 1.0 are well supported (Greenhill, Gray, et al. 2009). Fig. 1 demonstrates the ubiquity of systematic patterns across all languages, on the clade level and across phonological categories.

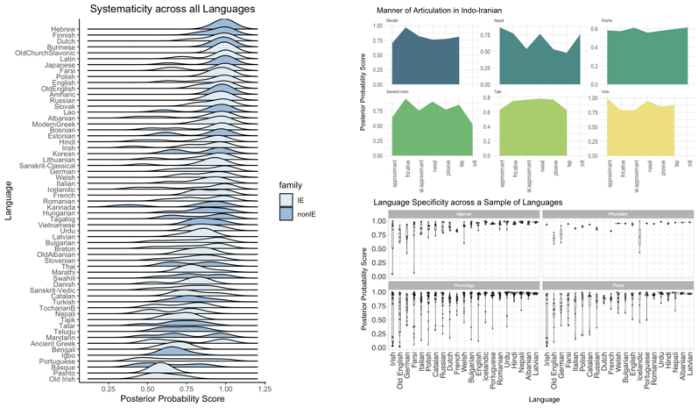

Figure 1. Systematicity patterns across all languages within the Indo-Iranian language clade and across all phonological categories (Place and Manner of Articulation, Phonology and Phonation).

This recurrent pattern was observed in other clades and phonological categories, such as place of articulation, phonation, and individual phonemic units. Consequently, it can be assumed that systematicity is present across all observed languages which seems to be a cross-linguistic pattern of distinct phoneme distribution in initial word segments between open and closed word classes. This effect, however, was not observed within the open word class, contrary to prior research. These discoveries underscore the pervasiveness and diversity of non-arbitrariness in a variety of global languages.

Acknowledgments

This work was supported by the NCCR Evolving Language. Swiss National Science Foundation Agreement Nr. 51N40_180888.

References

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.

Gray, R. D., Drummond, A. J., & Greenhill, S. J. (2009). Language phylogenies reveal expansion pulses and pauses in Pacific settlement. *science*, *323*(5913), 479-483.

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.

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., & Fitneva, S. A. (2011). The arbitrariness of the sign: learning advantages from the structure of the vocabulary. *Journal of Experimental Psychology: General*, *140*(3), 325.

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.

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.

Trott, S., Semenuks, A., & Bergen, B. (2019). Sub-morphemic form-meaning systematicity: the impact of onset phones on word concreteness. In *CogSci* (p. 3373).

Watson, S. K., Filippi, P., Gasparri, L., Falk, N., Tamer, N., Widmer, P., ... & Glock, H. J. (2022). Optionality in animal communication: a novel framework for examining the evolution of arbitrariness. *Biological Reviews*, *97*(6), 2057-2075.