Ollie Sayeed, UPenn, and Andrea Ceolin, UPenn

"Markedness" is an epiphenomenon of random phonetically grounded sound change

Introduction. The concept of "markedness" has been influential in phonology for almost a century, but recent movements in the field have argued that the concept should be eliminated from phonological theories. We argue, using a simple mathematical model based on Evolutionary Phonology (EP; Blevins 2004), that markedness is an epiphenomenon of random, phonetically grounded sound change.

Background. Theoretical phonology has found it useful to describe some segments as more "marked" than others, referring to a cluster of language-internal and -external properties (Jakobson 1941, Haspelmath 2006). Marked segments have low frequency within a language; marked segments have low frequency across languages; marked segments have a more restricted distribution within a language; marked segments are acquired later in infancy; marked segments are more likely to be impaired in cases of aphasia.

Starting with Trubetzkoy (1939), there has been a tradition of encoding "markedness" in the grammar, from the universal markedness rules of SPE to the markedness constraints of OT and its descendants. A recently popular alternative philosophy, EP, has argued that the concept of representational markedness is unnecessary (Blevins 2004, Hale and Reiss 2008, Samuels 2017). Instead, EP treats markedness as emergent from diachrony.

Model: random splits and mergers. We propose a simple abstract model of sound change as a discrete-time stochastic process of random splitting and merging of phonemic categories. Following EP, "marked" segments are those that have a low probability of being created by sound change and/or a high probability of being lost through sound change. Formalizing these two concepts, we define 'splitwise marked' segments to have a low probability of being created by a split, and 'mergerwise marked' segments to have a high probability of being lost by a merger. We simulate a model that randomly applies a split or a merger to an artificial phoneme inventory at each time step, with a diachronic bias against marked segments.

Predictions: phoneme frequency distributions. In the split-and-merger model, sound change belongs to a class of "random fragmentation and aggregation" processes (Banavar et al. 2004), whose fixed points are power-law frequency distributions over the elements being split and merged. Phoneme type and token frequencies in natural languages (Martin 2007) follow a Yule-Simon distribution (Simon 1955, Tambovtsev and Martindale 2007), a power-law distribution of which a special case is the more famous Zipf distribution. Simulations of our model show long-tailed distributions from an initial flat distribution, suggesting that sound change is one factor enforcing the Zipf distribution over time.

Predictions: within-language and across-language frequency. Empirically, across-language frequency correlates well with within-language frequency (shown by Gordon 2016 using a survey of 32 languages from WALS). The split-and-merger model derives this link from random sound change: it follows analytically that the expected frequency of an unmarked segment is higher than that of a marked segments within runs and across runs, and this is backed up qualitatively by simulations. We conclude that both i) splitwise marked and ii) mergerwise marked segments are shown to have both a) low within-language and b) low across-language frequencies as a consequence of phonetically grounded sound change.

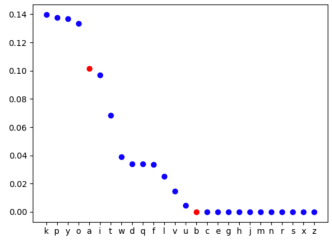


Fig. 1

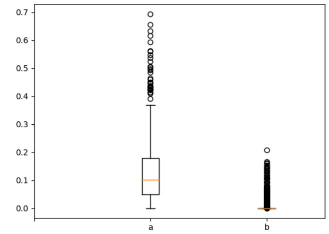
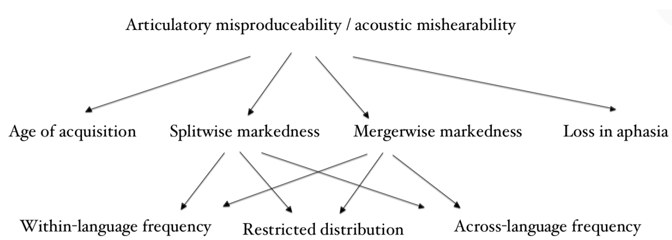


Fig. 2

We ran a simulation of the split-merger process for 100 iterations with 20 phonemes labelled a-t. In this example, mergers are biased towards /a/ and against /b/. Fig. 1 shows a typical run, where /a/ has become high frequency and /b/ has fallen out of the language. Fig. 2 shows the average frequencies of /a/ and /b/ across 1000 parallel runs given equal starting frequencies of 0.05; /a/ has median frequency of 0.1, while /b/ disappears in most runs.

Predictions: synchronic distributions. The model can be naturally adapted to apply mergers and *conditioned* splits with a randomly chosen environment can be applied to a miniature lexicon of strings. The model predicts that marked phonemes show a more restricted synchronic distribution after several runs, in line with traditional assumptions; in this case, the results are dependent on further parameters, like the initial distribution of the phonemes and the rate of lexical replacement.

Other correlates of markedness. Romani et al. (2017) find that age of acquisition and probability of impairment in aphasia show correlation with cross-linguistic frequency independently of within-language frequency. Both of these psycholinguistic correlations can be seen as effects of articulatory or perceptual difficulty, of the type that also underlies biases in sound change. The causal relationships between correlates of markedness start with these phonetic factors and ultimately trickle downstream into synchronic tendencies.



Conclusion. Both the power-law frequency distribution of phonemes in a language and the cluster of properties associated with "markedness" can be thought of as epiphenomena of phonetically grounded sound change. A random split-and-merger model predicts the attested language-internal and typological correlations.

References. Banavar, J. R. et al. (2004). Scale-free behavior and universality in random fragmentation and aggregation. **Blevins, J.** (2004). Evolutionary phonology: The emergence of sound patterns. **Gordon, M. K.** (2016). Phonological typology. **Jakobson, R.** (1941). Child language, aphasia and phonological universals. **Hale, M., & Reiss, C.** (2008). The phonological enterprise. **Haspelmath, M.** (2006). Against markedness (and what to replace it with). **Martin, A. T.** (2007). The evolving lexicon. **Romani, C. et al.** (2017). Comparing phoneme frequency, age of acquisition, and loss in aphasia: Implications for phonological universals. **Samuels, B.** (ed.) (2017). Beyond markedness in formal phonology. **Simon, H. A.** (1955). On a class of skew distribution functions. **Tambovtsev, Y., & Martindale, C.** (2007). Phoneme frequencies follow a Yule distribution. **Trubetzkoy, N.** (1939) Principles of Phonology.