LANGUAGES PREFER ROBUST PHONEMES

ANDREW WEDEL

Department of Linguistics
Tucson, AZ, USA
wedel@email.arizona.edu

BODO WINTER

Cognitive and Information Sciences Department, Univ. California
Merced, CA, USA
bwinter@ucmerced.edu

The signaling systems of all human languages are characterized by combinatorial structure: individually meaningless elements are combined to form meaningful units, words. A phoneme's overall contribution to conveying meaning is constrained by the degree to which it is perceptually distinct from other phonemes. If language evolution is influenced by pressures to transmit semantic information efficiently (e.g., Graff, 2012; Piantadosi et al., 2012), spoken languages should evolve to favor more perceptually distinctive phoneme types. Existing theoretical work suggests that those phonemes categories that are more perceptually distinct are transmitted more robustly across linguistic populations (Blevins, 2004; Winter, 2014).

In this paper, we present a test of this prediction: do phoneme inventories evolve to favor more robustly transmitted phonemes? To operationalize robustness, we use Graff (2012)'s symmetric phoneme confusability index, which is based on an English phoneme perception study from Miller and Nicely (1955). For each consonant phoneme within the English phoneme inventory, we ask whether its distinctiveness predicts how likely the phoneme will be found in other languages. And indeed, using the 1,672 languages represented in the PHOIBLE database (Moran et al., 2014), we show that the confusability of a phoneme is inversely correlated with its cross-linguistic frequency.

The evolution of language sound-systems as a whole has been proposed to be driven by changes in individual words, which then spread by diffusion through the lexicon (e.g., Wang, 1969, Wedel 2012). If phoneme inventories

evolve via transmission of phonemes through words, more robust phonemes should be on average found in more words (cf. Graff, 2012). We provide evidence in favor of this prediction by showing that scores on the Graff confusability index also significantly correlate with a set of within-language phoneme lexical frequencies, computed for Turkish, Mutsun, Mawukakan/Mahou, English, Dutch, German and Nama (7 languages from 5 different families): More robust phonemes occur in more words. Finally, the frequency with which consonant phonemes appear across the languages in PHOIBLE is also significantly correlated with within-language lexical frequencies. Specifically, the more languages in PHOIBLE that a phoneme appears in, the more words it tends to appear in within a given language. These results survive controls for language area and language family.

Taken together, our findings support the prediction that phoneme inventories evolve indirectly via a pathway involving transmission of individual words, and that the transmission of phonemes within words is constrained by their perceptual robustness. Our results firmly fit within a view that sees phoneme systems as evolving to meet the functional demands of language use (Wedel et al., 2013), and a view that sees language as more broadly evolved to support robust communication (Winter, 2014).

References

Blevins, J. (2004). *Evolutionary Phonology: The Emergence of Sound Patterns.* Cambridge: Cambridge University Press.

Graff, P. (2012). Communicative Efficiency in the Lexicon. PhD thesis, Massachusetts Institute of Technology.

Piantadosi, S. T., Tily, H., & Gibson, E. (2012). The communicative function of ambiguity in language. *Cognition*, 122: 280-291.

Miller, G. A., & Nicely, P. E. (1955). An analysis of perceptual confusions among some English consonants. *The Journal of the Acoustical Society of America*, 27: 338-352.

Moran, S., McCloy, D., & Wright, R. (eds.) (2014). *PHOIBLE Online*. Leipzig: Max Planck Institute for Evolutionary Anthropology. (Available online at http://phoible.org, Accessed on 2015-08-08.)

Wedel, A. (2012). Lexical contrast maintenance and the development of sublexical contrast systems. *Language and Cognition*, 4: 319-355.

Wedel, A., Kaplan, A., & Jackson, S. (2013). High functional load inhibits phonological contrast loss: A corpus study. *Cognition*, 128: 179-186.

Winter, B. (2014). Spoken language achieves robustness and evolvability by exploiting degeneracy and neutrality. *BioEssays*, 36: 960-967.