

Vowel Harmony as a Distributional Learning Problem

Spencer Caplan

University of Pennsylvania, Philadelphia, Pennsylvania, United States

Jordan Kodner

University of Pennsylvania, Philadelphia, Pennsylvania, United States

Abstract: Vowel harmony is a class of phonotactic restrictions in which vowels in a language are divided into two or more subclasses, and words must contain only vowels from only one such subclass regardless of intervening consonants. Languages worldwide (Turkish, Finnish, Mongolian, Warlpiri, but not English) exhibit vowel harmony. The opacity of such potentially long distance alternations poses a challenge for the learner. Nevertheless, infants are sensitive to vowel harmony alternations at as young as seven months. We present a computational model for vowel harmony acquisition. By normalizing transitional probabilities over the vowel tier, and making minimal assumptions about the phonology, we successfully determine which test languages have harmony processes and correctly categorize their vowels into harmonizing classes. Using universal typological patterns to inform the search space, we find that phenomena which appear opaque can be captured by simple distributional learning.