A method for learning features from observed phonological classes

Connor Mayer, UCLA   Robert Daland, Independent
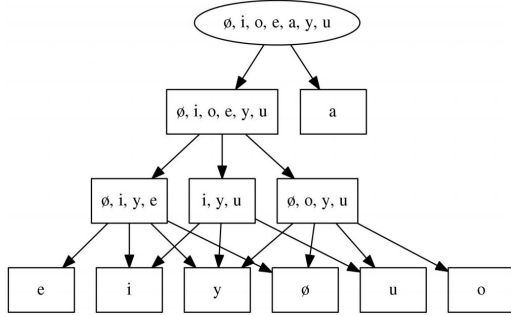

Fig. 1: A vowel class system

Classic texts (e.g., Chomsky & Halle, 1968) have assumed phonological features are universal, grounded in properties of the human vocal tract and perceptual system. More recent work has proposed theories of *emergent features* (e.g., Mielke, 2008; Archangeli & Pulleyblank, 2015, 2018). Under such theories, learners first identify classes of sounds in the input that share properties along one or more dimensions (e.g., articulatory, acoustic, distributional, or some combination of these) and then derive a symbolic feature system from these classes. Because features are not tied to a particular phonetic property, and may vary across languages, these theories provide accounts for phonetically disparate classes, variable segment patterning across languages, and other problematic cases for universal features. The goal of this project is to model feature learning as an emergent system. We describe a collection of algorithms that take as input a set of phonological classes and output a feature system that captures those classes. First, the input classes are arranged hierarchically based on parent/child relationships between them. An example for a simple vowel system is shown in Fig. 1. Second, the intersectional closure of the class system is calculated (Fig. 2), which includes all classes that must be generated by any featurization that characterizes the original class system (e.g., if there are [+back] and [+high] classes, there must
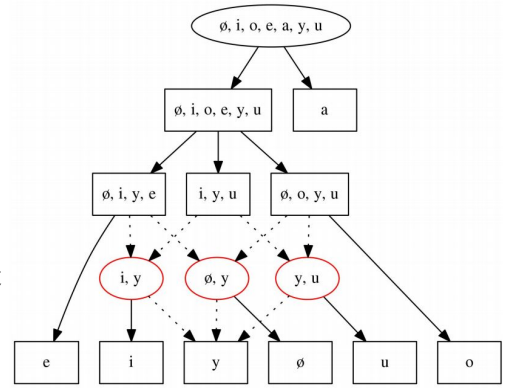

Fig. 2: The intersectional closure

be a [+back, +high] class). Finally, new feature/value pairs are generated for just those cases where they are necessary to distinguish a class from its parent. Because classes with more than one parent can be uniquely identified by the union of their parents' features, this corresponds exactly to classes that have only a single parent in the intersectional closure. The algorithms differ in whether they derive privative or ternary (underspecified), or binary (fully specified) feature systems, and whether they infer the existence of classes not in the input. An example of privative specification is shown in Fig. 3 (feature names added for readability). This work presents a deterministic method for deriving feature systems from learned classes, and generates testable empirical predictions about the types of feature systems learned by humans.
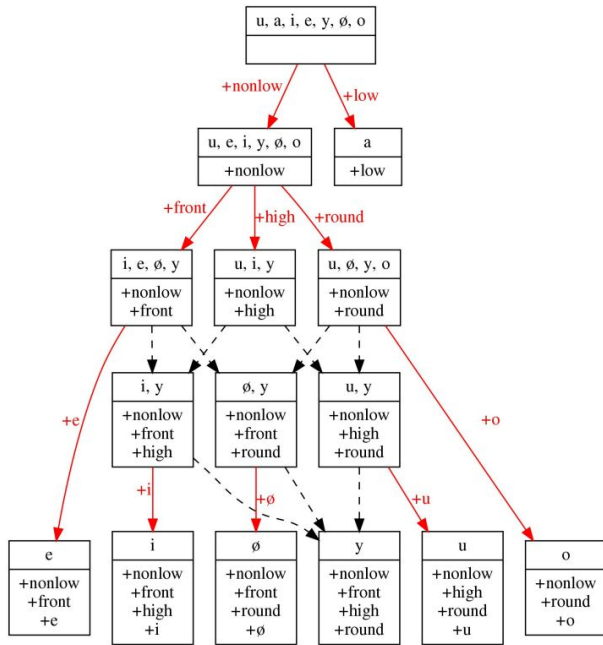

Fig. 3: Privative featurization

References

Archangeli, D., & Pulleyblank, D. (2015). Phonology without universal grammar. Frontiers in Psychology, 6, 1229.

Archangeli, D., & Pulleyblank, D. (2018). Phonology as an emergent system. In S. Hannahs & A. Bosch (Eds.), The Routledge Handbook of Phonological Theory (p. 476-503). London: Routledge.

Chomsky, N., & Halle, M. (1968). The Sound Pattern of English. New York: Harper & Row.

Mielke, J. (2008). The Emergence of Distinctive Features. Oxford: Oxford University Press.