Language Science Colloquium 2016

Job Candidate Talk
Time: 2:30-4:00 pm
Location: 1321 Social and Behavioral Sciences Gateway Building
Refreshments will be served.

Ewan Dunbar

WED, Feb 17th

Title: Quantitative investigations of speech sounds and sound systems

Abstract: What are the atoms of speech? Speech sounds are transcribed with phonetic symbols ("path": /p-æ-θ/), but it isn't these written symbols that go in our ear. Sounds have structured cognitive representations extracted ultimately from the auditory signal, and a central goal of speech science is to characterize their representational space. Moreover, according to emergent feature theory, the representational space for sounds is determined, language by language, by the phonological patterning of sounds in that language. I present new approaches to two pertinent questions. What bits of information about sounds ("features") can be extracted from speech? How do those features combine to yield the sound inventories of the world?

The first study takes three kinds of information that could each be the raw material for a representation of speech sounds: (acoustic) filterbank features extracted from a speech corpus, (articulatory) oral cavity distances from ultrasound, and (phonotactic) representations learned by training a neural network to predict phones from their context in a transcript of conversational English. I present a method for evaluating how well each space encodes voicing, manner, and place features in English consonants. The phonotactic space turns out not to code these features, while the combined acoustic and articulatory representations code them almost perfectly. The results are surprising for the emergent feature view, as they would imply that voicing, manner, and place are not basic representational dimensions for English consonants, contrary to other evidence.

The second study looks at the sets of sounds used in a large sample of languages (their sound inventories). I show that what sounds are in an inventory is strongly influenced by the way the sounds relate to one another, with respect to a space based on binary distinctive features. Specifically, languages tend to have more pairs of sounds differing only in one feature than expected by chance ([p] is to [b] as [s] is to [z], etc.: they differ only in voicing), and languages also tend to fill out their inventories in a balanced way, having, for example, about the same number of voiced as voiceless sounds. It is unexpected for the emergent feature view that geometric tendencies defined in terms of a language-independent set of feature dimensions should hold across the world's languages.

The first study also addresses the increasingly important methodological problem of giving cognitive interpretations to rich vector representations (brain imaging data, learned neural network embeddings), while the second deals with the problem that sound inventories' properties depend the way the inventory is analyzed into contrastive features. Large scale work in typology has been held back by the fact that properties of the language depend on the analysis, and linguists usually work by hand to choose between analyses. I show here that the result is robust to averaging over a large sample of possible analyses, an approach made possible by cheap computing power and one that I suggest is the way forward for studies of linguistic typology.