GMU Linguistics Colloquium Series

Word representation in speech recognition

Matthew C. Kelley
Postdoctoral Researcher
University of Washington

A core question in phonetics, phonology, and computational linguistics is how words should be represented, both for data analysis and cognitive modeling. In computational and human speech recognition, words are commonly represented as strings of speech segments. Segment strings are convenient and compact representations for words, but there are also concomitant drawbacks to only and exclusively using this approach. In this talk, I discuss this choice of representation and how it has shaped parts of both automatic speech recognition and spoken word recognition. Neural networks used in speech recognition are usually trained to separate acoustically similar speech segments as much as possible instead of exploiting the shared acoustic structure between similar segments. The first study I present examines how performance changes if the neural network is allowed to classify sounds as multiple similar segments at once. This change is implemented in a forced aligner that automatically places boundaries between speech segments in recordings. Additionally, in spoken word recognition, a word is said to induce greater amounts of lexical competition if it has a high degree of "sound similarity" with other words. Yet, sound similarity is calculated based on textually comparing strings of speech segments. The second series of studies I present attempts to bridge theory and practice in speech perception by assessing sound similarity with acoustic comparisons that reflect human cognition. The comparison technique itself originates from automatic speech recognition and can be performed across whole words. Together, these studies suggest that it may be advantageous to use nonsegmental representations for words in some scenarios. In addition, they highlight how automatic speech recognition might benefit from (re)considering how speech is mapped onto discrete objects.

HORIZON HALL Large conference room 4225
Zoom link: https://gmu.zoom.us/j/98032980617

WEDNESDAY,
MARCH 29, 2023
3:00 PM

GEORGE MASON UNIVERSITY