Multiple-Talker Speech Processing:

Cognitive Costs in Audio-Only and Audio-Visual Contexts

Presenter: Chi-Hyun Kim, University of Chicago

Talker differences result in variability in the relationship between acoustic patterns and phonetic categories. Despite this difficulty, listeners are adept at comprehending speech in multiple-talker contexts, albeit at a cognitive cost. So far, this cost has been demonstrated only in audio-only speech. Other work in single-talker contexts have shown, however, that when listeners are able to see the talker's speaking face, speech recognition is improved under adverse listening conditions. Does seeing a talker's face reduce the cost of recognition in multiple-talker contexts? We used a speeded word-monitoring task in which listeners make quick judgments about target-word identity in single- and multiple-talker contexts. Results show better recognition performance in single- talker conditions compared to multiple-talker conditions for both audio-only and audio-visual speech. Resolving talker variability does not seem to be made easier by face information. In light of this finding, we discuss the critical roles played by listener expectations and by information from the auditory stream.