The source and the signal: An integrated framework for talker identification and speech processing

Tyler Perrachione

Department of Speech, Language, and Hearing Sciences, Boston University

tkp@bu.edu

There are bidirectional dependencies between talker identification (knowing who is speaking) and speech processing (recognizing what is being said). While classically studied separately, decades of research in psycholinguistics and cognitive psychology now convincingly show that human listeners process these two types of information simultaneously and integrally. Such integral processing of voice and speech is often mutually advantageous to both understanding what was said and recognizing who said it. However, accommodating the cognitive demands of a system that evolved to decode these two signals simultaneously is also sometimes detrimental to fast and accurate talker identification or speech perception. Investigating speech processing and talker identification through an integrated framework provides more parsimonious answers to key questions in both of these areas: Why is listening to several talkers, even one at a time, more effortful than listening to a single talker? How do listeners learn to identify voices when most vocal interactions prioritize speech comprehension? And why are listeners so much better at identifying talkers in their native language than in an unfamiliar foreign language? Ultimately, theoretical advances in both speech perception and talker identification should inform one another: Speech must be recognized in the context of phonetic variability across talkers. And talker identification is enhanced when listeners' linguistic knowledge lets them include talkers' phonetic idiosyncrasies in their representation of talkers' identities