Modeling listener variability in prosody perception
using transcription and imitation as indirect measures of linguistic processing
Jennifer Cole, University of Illinois

The speech channel conveys not only the lexical content of an utterance, from which its syntactic and semantic forms can be constructed, but also pragmatic and discourse information that is encoded through prosody. Moreover, the same acoustic dimensions that signal linguistic meaning also communicate rich information about the speaker's identity, their emotional state, and even conditions of the physical environment. The interaction among these factors results in massive acoustic variability at every level of linguistic meaning—e.g., the acoustic realization of a word varies as a function of its prosodic context, the communication context, and speaker's identity. This talk focuses on variability in prosodic form (accenting and phrasing), and presents findings from my recent research collaborations investigating how variability due to linguistic context and speaker identity may influence listeners' perception of prosody. In our studies of prosody in spontaneous speech drawn from the Buckeye Corpus (Ohio State U) and the American English Maptask Corpus (MIT), we assess listeners' perceptual responses using two indirect measures of phonological and phonetic processing: real-time prosody transcription, and prosody imitation. Results from these studies show (i) listeners adapt to an individual speaker's encoding of prosodic form; (ii) correlations between prosody and factors of the linguistic context allow top-down processing to guide prosody perception; and (iii) listeners attend more to the phonological prosodic form than to the phonetic details of its encoding. I will discuss implications of this research for theories of the phonological encoding of prosody, addressing questions of the discreteness of prosodic features and emergent grammar.