Nonnative Speech Recognition Under High-Variability Listening Conditions
Terrin Tamati, Indiana University

In everyday life, listeners encounter enormous amounts of natural variability in the speech signal, such as region of origin, gender, and socioeconomic status. Although native listeners can use knowledge of sociolinguistic variability in communication, it can also impede speech recognition, for example, when a talker is from an unfamiliar dialect region or when there noise or competing speech in the background. These listening environments may also be challenging to second language (L2) learners, who have less experience with the phonological system and less knowledge of sociolinguistic variation of the L2 and, as such, may have difficulty adjusting to this type of variability for successful speech recognition. This difficulty may also contribute to individual differences in L2 communicative abilities.

Twenty-five native Korean learners of American English, who had only resided in Bloomington, Indiana for 1-27 months, completed two sentence recognition tasks with sentences from two sentence recognition tests: PRESTO (Felty, 2008), which contains sentences from male and female talkers from different U.S. dialect regions, and HINT (Nilsson, Soli, & Sullivan, 1994), which contains sentences from one male talker of a General American English variety. Sentences were presented in six-talker babble at 4 signal-to-noise ratios (SNRs). On each trial, listeners heard a single sentence and responded by typing the words they recognized. Responses were scored by keywords correct. Listeners also completed a forced-choice regional dialect categorization task, talker and gender discrimination tasks, a word familiarity questionnaire, and a self-report questionnaire of executive function. Results showed that PRESTO sentences were challenging for the nonnative listeners, especially at more difficult SNRs. Accuracy on HINT sentences was significantly greater than PRESTO sentences (all p’s <= .023). Nonnative listeners were significantly worse than a native comparison group on PRESTO and HINT (all p’s < .001), but differences in performance were smaller for HINT (HINT, 21.7%; PRESTO, 32.8%). Nonnative listeners were also significantly less accurate at dialect categorization (p < .001) and less familiar with English words (p < .001), but did not differ on talker and gender discrimination. Individual performance on PRESTO was significantly correlated with gender discrimination accuracy and self-report measures of executive function, but not with dialect categorization, talker discrimination, or vocabulary measures.

In the current study, nonnative listeners had much difficulty understanding speech in high-variability, adverse listening conditions. While their performance was less accurate than native listeners on both sentences types, PRESTO sentences were especially difficult. Additional tests and self-report questionnaires suggest that the difficulty on PRESTO may be due to lack of knowledge of regional dialects and a smaller vocabulary size. In addition, individual differences may be related to executive function and processing of gender information. Taken together, results suggest that previous experience with sociolinguistic variability is crucial to understanding speech in adverse listening conditions and that individual differences in L2 speech perception may be related to neurocognitive and associated perceptual abilities. These findings contribute to our understanding of how L2 learners handle sociolinguistic variation, and help to identify areas where learners might benefit from focused training or explicit instruction.