Native language differences in adaptation to altered auditory feedback
Sarah Bakst, John Houde, Susan Lin, and Keith Johnson
sbakst@wisc.edu, houde@phy.ucsf.edu, susanlin@berkeley.edu, keithjohnson@berkeley.edu

We listen to ourselves while we are speaking, comparing our acoustic output with our internal auditory speech targets. When this auditory feedback is altered and fed back to participants in real time, speakers change their well-practiced motor plans so that the altered version of their own speech better-matches the intended target (e.g. Houde & Jordan 1998). The experiment presented here adds ultrasound imaging to find out how speakers adapt to altered feedback. Both American English (AmE) and the Dravidian language Tamil (T) were tested to determine if language-specific importance of spectral cues in distinguishing phonemes affects adaptation. Tamil has five contrastive monophthongs {i, e, a, o, u}. Tamil also has a retroflex stops and liquids, which are primarily distinguished from their plain counterparts by F3. Thus the relative importance of F3 in distinguishing phonemes is greater in Tamil than in English, but F1 and F2 differentiate fewer vowels.

In each block, speakers gradually adapted to the perturbation during repetitions of a single word. In Block 1, F1 was raised (AmE ‘head’; T [pes] ‘speak!’). In Block 2, F2 was raised (AmE ‘hood’, T [po] ‘go!’). In Block 3, F3 was raised (AmE ‘heard’; T [a]r ‘cry!’). We hypothesized that Tamil speakers would adapt more than English speakers in the F3 condition and less in the F1 and F2 conditions.

To account for between-language vowel boundary differences and maximize visibility of articulatory differences in ultrasound, the perturbation differed between languages. In Tamil, the maximum perturbation was a 300 Hz shift upward in all blocks. In English, F1 and F2 were raised by 240 Hz, and F3 was raised by 450 Hz. Acoustic data was converted to Bark (perceptual scale) for analysis. The Tamil mean maximum perturbations were 2.5, 1.7, and 0.75 Bark, respectively. For English, they were 1.88, 1.07, and 1.3 Bark.

Both English (p < 0.01) and Tamil (p < 0.01) speakers compensated for a larger percentage of the F2 perturbation than the F1 perturbation. Tamil speakers barely adapted in the F1 condition (5%), compared with 15% for American English. However, Tamil speakers compensated for a larger percentage of the F2 perturbation (48%; mean acoustic change = 0.74 Bark) than English speakers did (31%; 0.37 Bark) (percent compensation not significantly different, but mean Bark difference was; p < 0.05). Perturbation differences may account for the differences in F1 behavior, but not for those in F2, where we would expect Tamil speakers to adapt less (Katseff et al. 2012). Due to difficulties with real-time tracking F3 during /r/, there were only reliable F3 data for 8 Tamil and 4 English speakers. Of these, Tamil speakers adapted more than English speakers. English speakers’ maximum change in production was 0.51 Bark, but Tamil speakers’ maximum change was 1.39 Bark (comparable minima). Ultrasound analysis is ongoing, and results will be presented to compare articulatory strategies between the two languages.

The results support the hypothesis that language-specific phonology and importance of spectral cues affect degree of adaptation: Tamil speakers adapted more strongly to the F2 perturbation than English speakers did; F2 may be a more important cue for dividing front and back vowels in Tamil than in English. Between-language differences in spectral cue salience may affect how speakers assess the accuracy of their productions.