The role of syllable prominence in foreign accent perception

Elizabeth A. McCullough & Richard Wright

Previous investigations have found that listeners' perceptions of foreign accent are related to both segmental (Major, 1987; Munro, 1993) and prosodic (Anderson-Hsieh et al., 1992; Boula de Mareüil & Vieru-Dimulescu, 2006) aspects of the speech signal. In a recent study, McCullough (2013) reported that VOT, F1, and F2 correlated with listeners' ratings of accentedness in one- and two-syllable English productions, while vowel duration correlated with ratings in two-syllable productions only. The restriction of the duration effect to the two-syllable condition suggests that listeners may have been noticing differences in prominence between the two syllables (e.g., Baker et al., 2011) rather than using duration as an indicator of the tense-lax distinction. In the present study, relative syllable prominence was evaluated directly using a subset of the recordings and perceptual responses used by McCullough (2013).

In these materials, six talkers from each of 3 L1 backgrounds (American English, Korean, and Spanish) were recorded reading English stop-initial trochaic words in isolation. These productions were then played for 20 monolingual American English-speaking listeners, who saw the orthographic representation of the target word prior to hearing each production and then rated the perceived accentedness of each talker on a continuous scale. Segmental and prosodic dimensions of the speech were measured. Segmental dimensions included VOT and vowel quality (F1 and F2) in the first (stressed) syllable, which for each production were expressed as the absolute distance from the native talkers' mean. Prosodic dimensions included vowel duration, intensity, and f0, which were measured in both syllables of each production and expressed as ratios (first syllable/second syllable) to represent relative prominence.

Mixed-effects linear regression models with logit-transformed accentedness ratings as the dependent variable were created separately for each L1 background. Multiple models were created and compared by log-likelihood ratio testing. For each L1 group, the base model included only a random intercept for talker. The segmental model included this random intercept and fixed effects of VOT and vowel quality, while the prosodic model included this random intercept and fixed effects of duration, intensity, and f0. The full model included all fixed effects from the segmental and prosodic models as well as the random intercept for talker.

For L1 American English talkers, neither the segmental nor the prosodic model offered an improvement over the base model, although the full model did ($p < 0.05$), suggesting that the small amount of variation in ratings for native talkers arose from a complex assortment of properties. For L1 Korean and L1 Spanish talkers, the segmental model did not offer an improvement in fit, but the prosodic model did ($p < 0.05$). Additional models were created to identify which prosodic dimensions were responsible for this effect. For L1 Korean talkers, adding duration, intensity, or f0 to the base model improved the fit ($p < 0.05$). Based on AIC values, the best addition was duration; adding intensity or f0 to this model offered no additional improvement, and the model with duration was not different from the full model. For L1 Spanish talkers, adding duration or intensity to the base model was a significant improvement ($p < 0.05$). Again, duration was a slightly better addition; adding intensity to this model did not improve it further, and the model with duration did not differ from the full model. Thus, for both non-native groups, relative duration of the syllables was the best correlate for accentedness.

These results suggest that even for highly constrained stimuli, listeners evaluating foreign accentedness may attend primarily to prosodic rather than segmental information in the speech signal. Such effects would be expected to increase with phrase- or sentence-length productions.

References

Anderson-Hsieh, J., Johnson, R., & Koehler, K. (1992). The relationship between native speaker judgments of nonnative pronunciation and deviance in segmental, prosody, and syllable structure. *Language Learning*, *42*, 529-555.

Baker, R. E., Baese-Berk, M., Bonnasse-Gahot, L., Kim, M., Van Engen, K. J., & Bradlow, A. R. (2011). Word durations in non-native English. *Journal of Phonetics*, *39*(1), 1-17.

Boula de Mareüil, P. & Vieru-Dimulescu, B. (2006). The contribution of prosody to the perception of foreign accent. *Phonetica*, *63*, 247-267.

Major, R. C. (1987). English voiceless stop production by speakers of Brazilian Portuguese. *Journal of Phonetics*, *15*, 197-202.

McCullough, E. A. (2013). Acoustic correlates of perceived foreign accent in non-native English. Ph.D. dissertation, Ohio State University.

Munro, M. J. (1993). Productions of English vowels by native speakers of Arabic: Acoustic measurements and accentedness ratings. *Language and Speech*, *36*(1), 39-66.