Some candidates for evaluating fluency, accentedness, comprehensibility, and pleasantness of non-native speech

Dong Jin Kim, Jieun Lee, & Hanyong Park
Department of Linguistics, University of Wisconsin-Milwaukee
kim252@uwm.edu | jieunlee@uwm.edu | park27@uwm.edu

Our previous study has revealed that listeners’ evaluations on certain aspects of non-native speech are often associated with one another in varying degrees. For example, regarding fluency, accentedness, comprehensibility, and pleasantness of non-native utterances, native listeners’ ratings were all correlated to one another but the ratings between pleasantness and fluency were stronger between other dimensions. As a follow-up, we investigated which features of speech are the source of associations among perceptual dimensions in the previous study. In the previous study, we asked 24 native speakers of English to rate spontaneous speech samples, each about 10-seconds long and by 21 adult Korean learners of English, on a 9-point scale, for the aforementioned four dimensions. In the current study, we made a variety of measurements of speech samples, such as speech rate, number of pauses, pause duration, repair fluency, and so forth. Then, we examined the correlations between the measurements and the perceptual ratings. We assumed that a higher correlation between a certain measurement and ratings in a particular dimension indicate the salience of that measurement in that dimension. Then, we hypothesized that if a particular measurement is highly correlated to the ratings of multiple dimensions, the ratings of these multiple dimensions would be correlated with one another.

Preliminary results confirmed our hypothesis partially. We observed that the speech rate measurement was positively correlated only with the fluency and pleasantness ratings. This suggests that speech rate is a good candidate for evaluating fluency and pleasantness of non-native speech, explaining the strong correlation of ratings between these two dimensions. However, none of our measurements were correlated to all ratings for the four dimensions. Nonetheless, we observed that the repair fluency measurement (consisting of repetitions, hesitations, replacements, and false starts) was negatively correlated with all ratings, except for accentedness, suggesting this composite measure contributes to the association among fluency, comprehensibility, and pleasantness.

Though we have not found a single measurement correlated with the ratings of all four dimensions, the list of our measurements is not at all exhaustive. To test our hypothesis more fully, we have been examining more measurements as possible candidates for perceptual evaluation.