Consistency in difference: The relationship between articulatory variability and segment differentiation for individual speakers

Sarah Harper

The articulatory and acoustic properties of any one phonological segment can vary substantially from token to token (e.g., [1]-[3]). Previous research has demonstrated that the effect of factors conditioning variability on the realization of particular articulatory dimensions varies across segments in a language, with specific dimensions differing in their response to these factors as a function of their recruitment to achieve specific phonological tasks ([4], [5]). However, most work examining the relationship between variability in the realization of particular articulatory dimensions and phonological contrast has focused solely on how variation along a specific dimension is or is not constrained for certain segments in specific phonetic environments, making it highly context-specific in scope. The extent to which individual speakers differ in the overall variability they exhibit across contextual and stochastic factors is largely unknown, but of crucial importance for understanding how phonological contrasts are produced and, subsequently, how they may be perceived. This study tests the hypothesis that speakers who exhibit less variability along a particular articulatory dimension will use this dimension to differentiate contrasting pairs of segments to a greater extent than more variable speakers.

Kinematic data from 9,859 tokens of word-initial and -final /s/, /ʃ/, /l/, and /ɹ/ were analyzed from sentences read by 40 native speakers of American English in the Wisconsin XRMB Corpus [6]. At least 30 tokens per segment per speaker were used in the analysis. Velocity trajectories of pellets placed on the upper and lower lips and on the tongue tip, blade, body and dorsum were used to find the time of movement extremum for the articulatory gesture(s) used to form each segment. Constriction location, degree, and orientation were extracted for all gestures in each consonant at the time for movement extremum. The contribution of these articulatory dimensions to the production of the /s/~/ʃ/ and /l/~/ɹ/ contrasts was evaluated using logistic regression models fit to each speaker's data, with standardized coefficients from these models interpreted as indicating the extent to which each articulatory dimension differentiated the pair of contrasting segments. The coefficient of variation (CoV), a standardized measure of variability, was additionally calculated for each measurement in every segment for all speakers individually. The relationship between the amount of variability a speaker exhibits along a particular articulatory dimension and the extent to which this dimension distinguished the two contrasting segments was evaluated using both Pearson's product-moment correlation and t-tests.

The results of this experiment indicate that there is co-variation in the extent of the variability observed along particular articulatory dimensions for specific segments and the contribution of these dimensions to differentiating contrasting segment pairs, independent of context. The CoV values calculated for the various articulatory dimensions and the standardized regression coefficients for these dimensions within the logistic regression models differed substantially between individual speakers, indicating that speakers differed from one another both in the extent of the variability they exhibited along specific dimensions and in the extent to which these dimensions could be used to distinguish the segments examined. Specifically, individuals exhibiting higher-than-average regression coefficients for a specific articulatory dimension were found to have significantly lower CoV values along that dimension than individuals with lower-than-average regression coefficients (Fig. 2). This suggests that speakers who are less variable in their production of a particular articulatory dimension use this dimensions to a greater extent in differentiating contrasting segment pairs than more variable speakers.

Fig. 1: Schematic representations of constriction location, degree, and orientation calculations. All location calculations based on x-axis distance of pellet from teeth (gray line).
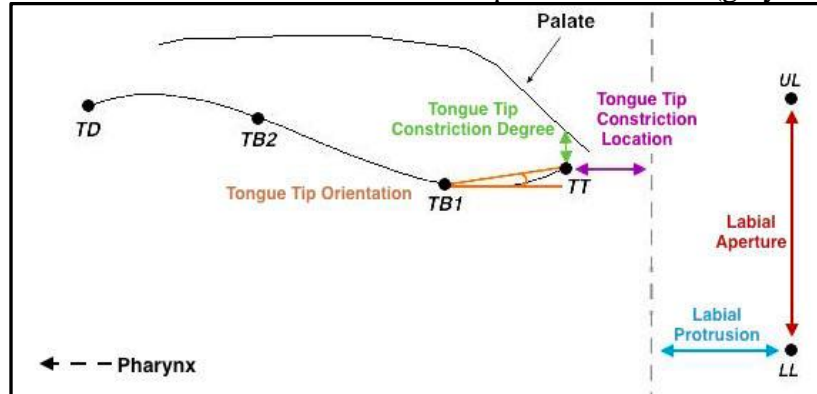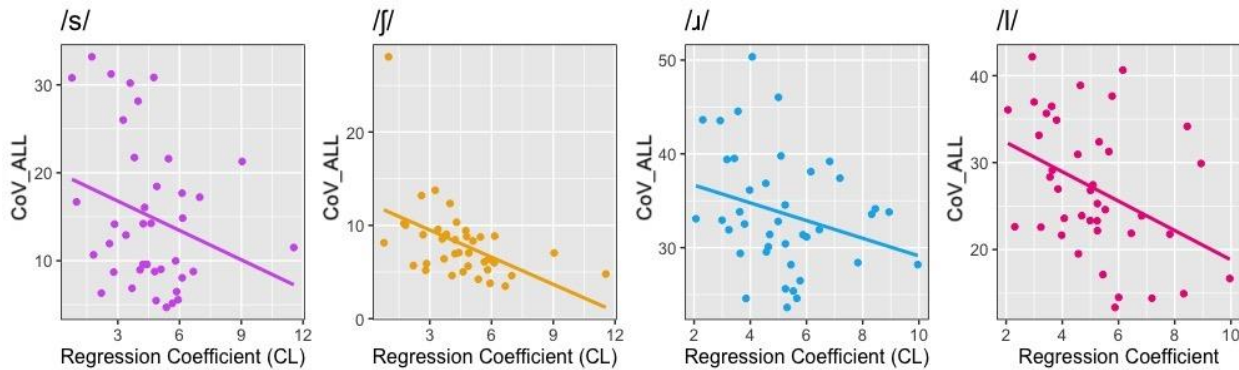

Fig. 2: Relationship between CoV values for the location of the anterior lingual constriction and regression coefficients across speakers for each segment.

References:
[1] Westbury, J., Hashi, M., and Lindstrom, J.M. 1998. Differences among speakers in lingual articulation for American English /ɹ/. Speech Communication, 26(3), 203-226.
[2] Blaydon, R.A.W, and Nolan, F. 1977. A video-fluorographic investigation of tip and blade alveolars in English. Journal of Phonetics, 5, 185-193.
[3] Noiray, A., Cathiard, M-A., Ménard, L., and Abry, C. 2011. Test of the movement expansion model: Anticipatory vowel lip protrusion and constriction in French and English speakers. Journal of the Acoustical Society of America, 129(1), 340-349.
[4] Keating, P.A. 1990. The window model of coarticulation: Articulatory evidence. In J. Kingston & M. Beckman (Eds.), Papers in Laboratory Phonology (Cambridge: Cambridge University Press), 451-470.
[5] Nieto-Castañon, A., Guenther, F., Perkell, J.S., and Curtin, H.D. 2005. A modeling investigation of articulatory variability and acoustic stability during American English /r/ production. Journal of the Acoustical Society of America, 117(5), 3196-3212.
[6] Westbury, J. 1994. X-ray microbeam speech production database user's handbook. J.R. Westbury: Madison, WI.