Daniel Duncan
Newcastle University

RUSE 2019

Competing representations in sociophonetic variation?

An implicit assumption in the analysis of sociophonetic variation is that speakers have one mental representation (a phoneme or set of allophones) and one phonological system (knowledge of which variant surfaces in which environments). This assumption may be seen, for example, in the typical use of linear regression analysis of acoustic data: the intercept reflects the mean production for a given baseline (the single representation), and language-internal and -external factors then have effects represented as a change from the baseline (the single phonological system). Recent empirical evidence problematizes the latter part of this assumption; Sneller (2018) suggests that speakers in Philadelphia show intraspeaker variation between two competing phonological systems with respect to the TRAP vowel (traditional short-a split vs. nasal system). If the assumption of a single phonological system does not appear to hold, we may ask whether the same is true of the assumption of a single representation. For example, could speakers have competing phonemic representations for a given vowel? This paper draws on empirical data from Greater St. Louis to suggest that this is indeed the case for some speakers.

I test this question by examining intraspeaker variation in production of the LOT vowel among 52 white middle-class females, born 1896-1992. Approximately 15 minutes of informal conversation with the women were transcribed and forced-aligned with FAVE (Rosenfelder et al. 2014), and vowel formants were extracted at 40% of the vowel duration using a Praat script. Vowels were Lobanov normalized and rescaled using the overall speaker mean and standard deviation, and then pre-obstruent tokens of F2 of the LOT vowel were extracted from the data (n=2337, 19-89 per speaker). F2 of LOT was selected for emphasis because the St. Louis area participated in the Northern Cities Shift, which involves fronting of the vowel (Duncan 2018, *inter alia*).

I model speaker representations with two hidden Markov models (Visser and Speekenbrink 2010): one which assumes two states and one which assumes one. These models iterate through the data to attempt to find the mean value for each of a specified number of states. When there is more than one, the model predicts transition probabilities that represent how likely the system is to switch between states. In the present data, we may interpret states as representations/targets: one state will have a higher F2 value as a mean than the other (representing fronter/backer vowels), and the transition probabilities thus model the speaker selecting between fronter and backer representations. For each speaker, I assess model fit using the Bayesian Information Criterion; the model with the lower BIC is selected. I interpret the model with two states fitting the data better as indicating that the data supports a view of competing representations, and that speakers have both a fronter and backer target for the vowel. I interpret the model with one state fitting the data better as indicating that intraspeaker variation is based on production targeting one representation, even as it may include fronter and backer tokens.

BIC was lower for the two-state model in 14/52 model pairs (26.9%). Given that the implicit assumption is that no two-state models would be selected, this is surprisingly high. The two states appear to represent quite distinct targets; the average 'front' state was 1638.8 Hz, while the average 'back' state was 1324.7 Hz. The average range of 314.1 Hz suggests that speakers for whom the two-state model was selected seem to in fact have two targets. The targets additionally appear to largely not be due to speech errors or lexical items lagging/leading change. Were this the case, we would expect to find that transition probabilities predict that one state is quite likely to be followed by the other, but not vice versa. This would indicate that one state is the speaker's representation of the LOT vowel, but that occasional outliers (whether speech error or lexical exception) are uttered before speakers almost certainly return to their primary target. While this was the case for two speakers, it was not the case for 12/14 speakers, who were quite likely to continue in one state upon having uttered a token in that state (p at least .686). These 12 speakers appear to truly vary between two representations.

The question, of course, is why these speakers and no others would do so. Having multiple representations does not appear to be conditioned by the social factors of age, education level, or location (urban vs. suburban) within Greater St. Louis. The speakers' ages additionally do not appear to cluster around a period of change (whether advance/retreat of fronting) in F2 of LOT. Nevertheless, I hypothesize that having multiple representations may play some role in sound change, and may perhaps arise due to as yet unknown social or cognitive factors. For now, we may say that individuals may have competing representations, and that whether an individual does is itself variable. I call for additional research into this problem, as replication of this result should lead us to more deeply consider the origin and role of competing representations.

Daniel Duncan
Newcastle University

References:

Duncan, Daniel. 2018. Language variation and change in the geographies of suburbs. Doctoral dissertation, New York University.

Rosenfelder, Ingrid, Josef Fruehwald, Keelan Evanini, Scott Seyfarth, Kyle Gorman, Hilary Prichard, and Jiahong Yuan. 2014. FAVE (Forced Alignment and Vowel Extraction) Program Suite v1.2.2 10.5281/zenodo.22281.

Sneller, Betsy. 2018. Mechanisms of phonological change. Doctoral dissertation, University of Pennsylvania.

Visser, I, and M. Speekenbrink. 2010. depmixS4: An R Package for Hidden Markov Models. *Journal of Statistical Software*, 36(7): 1-21. URL http://www.jstatsoft.org/v36/i07/.