

Affect Structures Variation in Vowel Quality: The Influence of Smiling on the Front Lax Vowels in California

Robert J. Podesva

While sociolinguists have long recognized that variation patterns are structured by social category membership, variationists have only recently turned toward more “interactionally specific” (Bucholtz and Hall 2005) dimensions of identity like affect (Johnson 2006, Eckert 2010, Podesva et al. 2015). Affective approaches to variation have demonstrated that vowels produced when speakers express positive affect are characterized by higher formant frequencies, particularly F2. Most of this work focuses on back vowels, which are fronting (higher F2) in many varieties of North American English. It is unclear whether fronting should be directly attributed to the expression of positive affect, or whether it is an automatic consequence of smiling, which frequently accompanies emotions with positive valence (lip spreading shortens the vocal tract, thus elevating F2). This paper uses audiovisual recordings of speakers from California to examine the front lax vowels, which are lowering (higher F1) across the state (Van Hofwegen et al. 2016), to test whether expressions of positive affect also influence F1 (which is not directly influenced by the physical act of smiling). I show that speakers produce lower (higher F1) front lax vowels when they smile, revealing a case where higher formant frequencies cannot be considered a straightforward acoustic consequence of lip configuration.

Data consist of unscripted dyadic interactions (30 minutes) between speakers recorded in a large sound-attenuated room staged like a living room. Recordings were transcribed and force-aligned, and a number of acoustic measurements were collected every 10 ms. Median F1 and F2 were taken for every vowel interval (KIT, DRESS, LOT), and normalized following Lobanov (1971). Smiling intervals were identified by running each video frame through a Haar cascade classifier trained on open-source data annotated for incidence of smiling. Mixed-effects regression models were constructed for F1 and F2, including linguistic (phrase position, segment duration, phonological environment) and social predictors (\pm smiling, sex, ethnicity), and random effects of speaker and word.

Results indicate that all three vowels are fronter (higher F2) in phrases when speakers are smiling ($p < 0.05$). This finding can be viewed as an acoustic consequence of smiling and is consistent with previous findings for back vowels in California. Regarding F1, all three vowels are lower (higher F1) in phrases when speakers are smiling ($p < 0.01$). Thus, vowels are more advanced (lowered) in the direction of the regional vowel shift when speakers are smiling, a pattern which cannot be understood as a reflex of smiling. Interestingly, speaker sex predicted neither F1 nor F2, suggesting that affective factors like smiling have at least as strong an effect as more well-established social factors.

In conclusion, affect structures vowel quality variation. Lip configuration does not wholly explain the observed patterns, suggesting that other articulators (tongue position, larynx height) are also recruited to express positive affect while smiling. This study raises questions about the extent to which regional accents are ideologically tied to particular affective valences, and calls for additional work on non-Western varieties. The approach

taken here, which harnesses computer vision techniques to objectively quantify one dimension of affect, can facilitate such work.