

The listener as a gate keeper between phonetic variation and phonological innovation

Numerous parallels between historical sound change and phonetic contextual variations in natural speech have motivated considerable research in phonetics, phonology, sociolinguistics, and historical linguistics on issues of how these phonetic variations become phonological. A particularly heated debate in the last three decades has been on whether a main cause of sound change is in the listener, as argued for by Ohala (1981, 1989, 1993) and other researchers (e.g., Beddor 2010; Blevins 1994; Harrington et al. 2008; Yu, 2011), or in the speaker, as argued for by Lindblom and his colleagues (1995) and other researchers (Boersma 1997; Hura, et al. 1992; Steriade 2001; Garrett & Johnson 2011). The purpose of the present study is to shed some light on a division of labor between speech production and speech perception for transmission of speech sounds between a speaker and a listener as well as to study the role of cognitive factors that seem to bridge production and perception.

To ascertain the relevant factors in both speech production and speech perception that might be preconditions for coarticulation-based sound changes, I investigated how speakers of American English produce, perceive, and repeat the high back vowel /u/ in fronting and non-fronting contexts. The production study examined the extent of contextual variation of /u/ in /dud/, /bud/, and /hud/ syllables uttered at various speaking rates. I found that: (1) the relative acoustic difference between the fronted /u/ and the non-fronted /u/ persisted across the elicited ranges of vowel duration; and (2) the degree of acoustic variability was less for the fronted /u/ than for the non-fronted /u/. These results indicate that speakers have a distinct and more narrowly specified articulatory target for the fronted /u/ in the alveolar context, separately from the /u/ in non-fronting contexts. A perception study examined each subject's /i/-/u/ category boundary in a /dit/-/dut/ and a /bip/-/bup/ syllable continuum separately. That study found evidence for (1) compensation for coarticulation (i.e., phonemic category boundary shift as a function of consonantal environment), (2) systematic individual variation in perceptual category judgments, and (3) similarity between the distributional properties of /u/ observed in the production experiment and the range of perceptual responses. Taken together, these results suggest that one source of systematic variation in speech perception is individual differences in listeners' phonological grammars (the perceptual category boundaries) used in speech perception, and that this grammar emerges in response to the structure of ambient language data to which the listeners have daily exposure. Finally, a vowel repetition study examined the acoustic quality of vowels that listeners repeated after hearing each of the CVC syllables used in the perception study. I found evidence that listeners' judgments about the auditory quality of heard vowel stimuli guide how they repeat the vowels, suggesting that perceptual interpretation determines mental representation of spoken inputs.

Based on these experiments, I contend that pronunciation variation emerges through a system of multiple causal loops between and among mechanisms of speech perception and speech production, speaker knowledge of the normative pronunciations in one's speech community, and the distributional structure of ambient language data. These properties of spoken language govern the output of communicative interactions among members in a speech community, and one such output is the knowledge of sub-phonemic variations as distinct pronunciation categories that speakers have for their pronunciation repertoire. Additionally, I argue that any speech community is in a constant state of readiness to adopt an innovative pronunciation as a new community norm, because members have rich pronunciation repertoire even when there is no observable community-level sound change.