Naïve Categorization of American English Vowels

Douglas S. Bigham, The University of Texas at Austin, douglas.s.bigham@gmail.com

This work looks at the categorization of vowel sounds by speakers of American English. Many studies have shown evidence for the categorical perception of speech sounds (see, e.g., Liberman et al., 1957; Beddor & Strange, 1982; Beddor, 1992; Lotto et al., 1998; Diehl et al., 2004). Typically, perception studies investigate the categorical boundaries between percepts with closely related acoustic structures, such as the F1 cross-over point between /ɪ/ and /ɛ/ percepts or the VOT boundary between voiced and voiceless stop percepts. Although listener perception generally coincides with the pre-established categories of expert phoneticians, no major studies currently exist that begin without these kinds of *a priori* category assumptions. Without knowing the cognitive reality of these categories as naïve listeners experience them, categorical perception phenomena cannot be fully interpreted.

This study works toward filling this gap in our understanding of the categorical perception of vowel sounds by presenting results based on the naïve-categorization views of listeners. Instead of providing category or token labels and having listeners perform a difference-judgement task, listeners constructed vowel categories as they perceived them. ~35 undergraduates at the University of Texas heard isolated tokens of 11-14 standard vowels of American English (/i, ɪ, e, ɛ, æ, ɑ, ɔ, o, ʊ, u, ʌ/; also, /aɪ, ɔɪ, aʊ/ for some listeners). Listeners were asked to group vowel tokens into 2, 3, 4, and 5 categories. All listeners were previously familiar with the speaker who provided the token data. No speakers had previous experience with phonetics or the phonetic categories of vowel sounds. Results are interpreted via a correlation matrix of vowel x category groupings.

Although listeners show a range of idiosyncratic grouping variables (e.g., Listener A groups tokens according to expected phonological categories of front/back, high/low, tense/lax while Listener B groups tokens according to assumed spelling rules), overall patterns of vowel categorization do emerge. Surprisingly, though there is some evidence for the psychological reality of traditional phonological categories of vowels, there is stronger evidence for a circuit of low and lax vowels, similar to Labov's (1994) "peripheral" dichotomy of vowels. That is, the vowels /ɛ, æ, ɑ, ʌ, (ɪ, ɔ)/ are highly correlated in naïve listener groupings. Additional highly correlated groupings include back-round vowels: /u, o, (ʊ)/; and front-gliding diphthongs with high front monophthongs: /aɪ, ɔɪ, e, i/.

Since the listener-subjects performing the categorization are drawn from a wide variety of dialect backgrounds, it is unlikely that they are simply responding to an underlying awareness of vowel system changes. The evidence for these patterns, then, may lend objective support to the posthoc construction of the feature [+/- peripheral] in Labov (1994) and related work or help explain the connectedness of vowels for historical work. The ways in which these naïve groupings deviate from the categories of expert phoneticians can not only shed light on language change phenomena but also provide a principled benchmark from which future work on vowel perception, categorization, and change can proceed.