

Naïve Categorization of American English Vowels

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 (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 VOT boundary between voiced and voiceless stop percepts or the F1 cross-over point between the KIT and DRESS vowels (using Wells's, 1982, key words). 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, perception-based phenomena cannot be fully interpreted.

The results of this study, then, provide a first step towards understanding how naïve listeners experience and group vowel sounds. Instead of providing category or token labels and having listeners perform a difference-judgement task, a modified pilesort task was used to allow listeners to construct vowel categories/groups as they perceived them. 35 undergraduates at the University of Texas heard isolated natural-speech tokens of 14 vowels of American English (the vowels of FLEECE, KIT, DRESS, TRAP, LOT, THOUGHT, STRUT, FOOT, GOAT, GOOSE, FACE, PRICE, CHOICE, MOUTH). Listeners were asked to group vowel tokens into 2, 3, 4, 5, 7, and 9 categories. All listeners were previously familiar with the speaker who provided the token data. No listeners had previous (formal) experience with phonetics or the categorization of vowel sounds. Results are interpreted via a similarity matrix; multidimensional scaling is used to visualize these findings.

Although listeners report a range of idiosyncratic grouping choices (e.g., Listener A reports grouping tokens according to assumed spelling rules while Listener B's grouping corresponds to a tense/lax vowel distinction), overall patterns of vowel categorization do emerge, generating a kind of folk-taxonomy of American English vowels. Surprisingly, though there is some evidence for the psychological reality of traditional phonological dimensions of vowels, there is stronger evidence in support of Labov's (1994) "peripheral" dichotomy of vowels. Specifically, the "non-peripheral" vowels of the KIT, DRESS, TRAP, LOT, STRUT, and FOOT sets are highly similar in naïve listener groupings. Additional groupings include high-back round vowels of the GOOSE and GOAT sets and front-gliding diphthongs with high front monophthongs (grouping the PRICE, FACE, and FLEECE sets together).

Since the listeners performing the categorization task 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 from these results, then, may lend objective support to the post-hoc construction of the feature [+/- peripheral] in Labov (1994) and related work. The ways in which these naïve groupings deviate from those of expert phoneticians can not only shed light on speech perception and language change phenomena but also provide a principled benchmark from which future work on vowel perception, categorization, and change can proceed.

(498 words)