The production and perception of a low back vowel merger

Katie Drager (*University of Hawaii at Manoa*), Rebecca Clifford (*University of Hawaii at Manoa*) and Jennifer Hay (*University of Canterbury*)

Exemplar Theory, a theory that is gaining popularity among sociolinguistics, posits that phonetically-rich representations of words are stored by listeners and that speakers use these word-level representations when producing speech (Johnson 1997, Pierrehumbert 2001). Additionally, hybrid models (Pierrehumbert 2006) assume a layer of phonemic abstraction. If episodic information is available at both the word and phoneme levels, then these may be accessed to different degrees for different types of tasks. One realm where each level of representation should be differently involved is in the production of real versus non-words because for words that have never been encountered, there is no existing word-level representation. Previous work has shown a production difference between real and non-words in a conditioned merger (Thomas and Hay 2005), but this has not yet been examined in a non-conditioned merger. The current study investigates the production and perception of real and non-words containing the vowels /ɑ/ and /ɔ/ for individuals who merge these vowels to varying degrees.

While the low back vowel merger has been studied, this study provides new insight by comparing the production and perception of real words to nonsense words. Participants from the western United States and Hawaii took part in a production task where they read, in isolation, real and nonsense words (e.g. cot and caught; chot and chawt) containing the target vowels. Participants then took part in a binary-forced choice identification task, where they listened to tokens produced by a speaker who maintains a distinction between the vowels. While all of the participants merged /ɑ/ and /ɔ/ in their production of at least some word pairs, the speakers merged to varying degrees. Based on how merged they were in production, participants were grouped into two categories: those who merged all tokens in all phonological contexts (fully merged), and those who merged some tokens in some contexts (nearly merged). Mixed effects models fit to the data indicate that 'nearly merged' participants were more likely to produce a distinction in nonsense words than in real words ($p<0.01$). In contrast, these same participants were more accurate at identifying real words than nonsense words during the perception task ($p<0.05$), with an average score of 73% for real words and 55% for nonsense words. Fully merged participants on the other hand were more accurate at identifying nonsense words than real words ($p<0.05$), with an average score of 53% correct for real words and 68% for nonsense words.

Taken together, we argue that the results provide evidence that models of production and perception must have both a phonetically-detailed word-based level of representation and an abstract phoneme level of representation. The word-based episodes dominate the production and perception of lexical items but only if those lexical forms exist. The phonemic representations drive the production and perception of non-words. The observed effect of this changing balance between lexical items and phonemic categories is speaker dependent; it is affected by the degree to which phonological categories are kept distinct for the individual.