Speaker- vs. Listener-based Explanations for Dispersion Effects in Sound Change

Tommy Denby
Northwestern University
tdenby1@gmail.com

Grant McGuire
UC Santa Cruz
grantlmcguire@gmail.com

Jaye Padgett
UC Santa Cruz
jayepadgett@gmail.com

Phonetic dispersion—defined as the maintenance of sufficient pyschoacoustic distance between phonemic categories such that they remain distinct—has been proposed as the driving force behind universal trends within phoneme inventories and a number of closely related sound-change phenomena, such as vowel chain shifts and compensatory sound change. While the need to appeal to dispersion in sound change seems clear, the precise mechanism by which it comes about has remained elusive. The results of our experiments, testing both speaker-based and listener-based mechanisms, are reviewed here.

Speaker-based accounts (e.g. Liljencrants and Lindblom, 1972; Lindblom, 1986) posit that speakers are sensitive to the communicative needs of listeners, and adjust their production based on these needs. There is strong evidence to support a broad version of this claim, given that speakers hyperarticulate vowel productions when prompted for "clear" speech in a laboratory setting (e.g. Moon and Lindblom 1994; Uchanski, 2005) and can tailor their speech for various settings. Yet these types of global effects do not account for the hyperarticulation of individual phonemes to preserve a particular contrast, and thus do not explain dispersion effects such as chain shifts. Further, they can actually be explained by appeal to a listener-based approach (McGuire et al. 2010).

The listener-based hypothesis, (Labov, 1994: 587; Wedel, 2006; Denby, 2013), posits that phonetically unambiguous productions will influence future productions of the listener more than ambiguous productions. The mechanism that drives this is a filter: not all ambiguous productions are stored to phonetic memory (or are stored but decay faster than their unambiguous counterparts). In turn, these unstored productions do not update the phonemic categories of the listener, and are not reflected in that listener's future productions. The filter acts as a buffer between phonemic categories, reducing the number of productions stored between phonemic categories relative to those at the center of the distribution.

The first set of experiments sought to test the speaker-based account by examining subjects' online control over dispersion of individual vowels. Subjects were visually prompted to pronounce monosyllabic words containing one of three adjacent vowel categories, e.g., [ɪ,ɛ,æ]. Subjects were told they were testing speech recognition software. For some trials the presentation software (E-prime) appeared to incorrectly recognize the production and prompted further (hopefully hyperarticulated) productions, with up to three additional repetitions possible. The productions of interest to us involve the middle in the vowel triplet, e.g., [ɛ]: in one condition, subjects were led to think the program misheard [ɪ] (e.g. reporting recognizing "pick" instead of "peck"); in another the program "misheard" [æ]. The hypothesis was that productions of [ɛ] in the former conditions would be lower than those in the latter. The results of these experiments, under various conditions, have been null, giving no support so far for the speaker-based approach to dispersion.

In an experiment designed to test the listener-based approach, subjects heard ambiguous target words and unambiguous control words in noise and were asked to identify them, following Goldinger (1996).

Subjects were exposed to a word list containing minimal pairs, each of which was monosyllabic and stop-initial, differing only in first-segment voicing (e.g. *tip*/*dip*). Half of these pairs were normal, unambiguous productions, while the stop-initial VOT of the other half were manipulated to be somewhat ambiguous by replacing periods of the onset of the vowel of the voiced member with aspiration from the voiceless member. An ambiguous stimulus skewed towards a voiced initial stop, and one skewed towards voiceless, were created for each minimal pair. After each stimulus, subjects typed the word they heard using a keyboard.

If subjects store these words, their accuracy should improve with every exposure. If however, they do not store ambiguous productions, their accuracy should not improve after repeated exposures, or improve less than it does for unambiguous productions. Using *d'* scores, a repeated-measures ANOVA (subjects treated as random effects, with experimental block and stimulus ambiguity as factors) reported that differences in improvement between ambiguous and unambiguous stimuli are significant (Subject*Ambiguity, $F(1,24)$, $p < .0005$; Subject*Block*Ambiguity, $F(3,72)$, $p < .05$). Results also indicated, however, that the manipulation of the stimuli was flawed, as subjects had a large bias towards hearing ambiguous stimuli, regardless of intended voicing, as voiced, and that this may represent a serious confound. Nevertheless, this provides strong initial evidence for a listener-based approach. Currently, a follow-up study is being implemented to further test and expand these results.

Denby, Thomas (2013). The filtering listener: dispersion in exemplar theory. MA thesis, UC Santa Cruz.

Goldinger, Stephen D. (1996). Words and voices: episodic traces in spoken word identification and recognition memory. *Journal of Experimental Psychology* 22.5, 1166-1183.

Labov, William (1994). Principles of linguistic change. Oxford: Blackwell.

Liljencrants, Johan & Björn Lindblom (1972). Numerical simulation of vowel quality systems: the role of perceptual contrast. *Language* 48.4, 839-862.

Lindblom, Björn (1986). Phonetic universals in vowel systems. In Ohala, J. J. & J. J. Jaeger (eds.) Experimental phonology. Orlando: Academic Press, 13-44.

McGuire, Grant, Jaye Padgett, Lillian V. Clark, Nathan P. Hinchey, Rachel McClellan, Kaitlyn Pavlina, Elan Samuel & Alex Wolfe (2010). On the bases of vowel dispersion: an experiment. Slides of talk presented at the annual meeting of the LSA, Baltimore.

Moon, Seung Jae & Björn Lindblom (1994). Interaction between duration, context, and speaking style in English stressed vowels. *Journal of the Acoustical Society of America* 96.1, 40-55.

Uchanski, Rosalie M. (2005). Clear speech. In Pisoni, D. B. & R. E. Remez (eds.) *The handbook of speech perception.* Malden, MA: Blackwell Publishing, 207-235.

Wedel, Andrew (2006). Exemplar models, evolution and language change. *The Linguistic Review* 23, 247-274.