Phonological Variation in Auditory Word Recognition

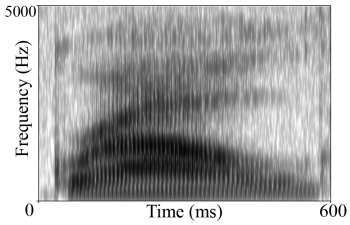
Introduction: Auditory word recognition studies generally focus on how listeners recognize words in homogeneous linguistic varieties; less is known about how lexical access functions in the presence of phonetic and phonological variability (but see Sumner and Samuel 2009). This study investigates how variation in the pronunciation and representation of sesquisyllabic words (trimoraic words comprising a diphthong followed by [l] or [ɹ], e.g. *owl*, *dire*) influences auditory word recognition in American English. Lavoie and Cohn (1999) and Tilsen et al. (2014) document significant variability in American English speakers' judgments about the weight of trimoraic words, with some judging them to comprise one syllable and others judging them to comprise more than one syllable. Tilsen et al. (2014) further show that the pronunciation of trimoraic words differs depending on the speaker's judgment of their phonological weight. Sesquisyllabic words thus provide a good test case for examining how phonological variability influences lexical access.

Method: Participants completed an auditory lexical decision task followed by a syllable judgment task. In the lexical decision task, participants judged the lexicality of 90 monomorphemic real words of English and 90 phonotactically legal non-words. The real and non-words comprised one-third monosyllabic words (σ, e.g. *wish*), one-third disyllabic words ($\sigma\sigma$, e.g. *coffee*), and one-third sesquisyllabic words ($\mu\mu\mu$, e.g. *growl*). Sesquisyllabic words exemplified the following rime-coda combinations: [aɪl, aʊl, oɪl, aɪɹ, aʊɹ].

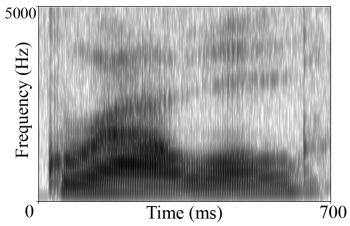
Materials were recorded by two monolingual English speakers. Speaker 1 was a male from Pittsburgh who judges and produces sesquisyllabic words as monosyllables. Speaker 2 was a female from Montana who judges and produces sesquisyllabic words as disyllables. Representative spectrograms of the word *growl* produced by Speaker 1 (top) and Speaker 2 (bottom) are shown below. Speaker 2 exhibits notable F2 movement while Speaker 1 monopthongizes diphthongs to some extent (typical of Pittsburgh English, see for example Labov et al. 2006). Two experimental lists were constructed with Speaker counterbalanced such that every participant heard every word type (σ, $\mu\mu\mu$, & $\sigma\sigma$) produced by both speakers, but responded to each word only once.

In the syllable judgment task, participants viewed written forms of the real words from the lexical decision task and judged whether each word *as they pronounced it* comprised one syllable or more than one. Participant's **structure preference** was calculated as the proportion of sesquisyllables that they judged as having more than one syllable.

Speaker 1: *growl* = σ



Speaker 2: *growl* = $\sigma\sigma$



Hypothesis: Dissimilarity between listener phonology and speaker pronunciation will inhibit lexical access: listeners who judge sesquisyllabic words to be heavier than one syllable will respond more slowly to Speaker 1's monosyllabic tokens compared to Speaker 2's disyllabic tokens, and vice versa.

Preliminary Results: Data collection is underway; so far 12 native monolingual English speakers have participated (average age = 21.31). We aim to have data from an additional 18-20 participants by October, 2019.

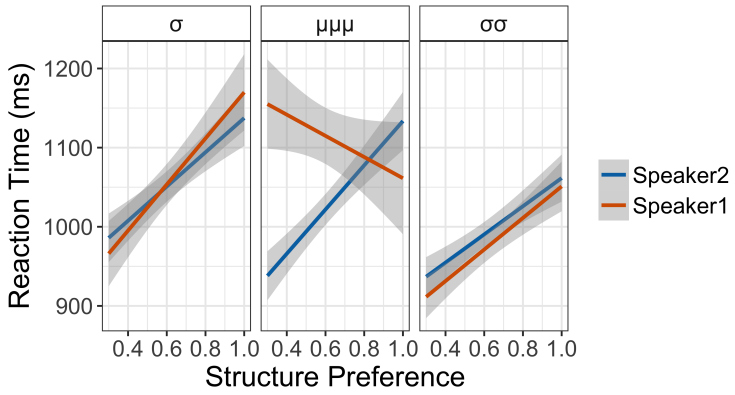


Figure 1: Reaction time for accurate responses to real word targets.

Figure 1 shows RT for accurate responses to real word targets as a function of structure (σ, $\mu\mu\mu$, or $\sigma\sigma$), speaker, and participants' structure preference (1.0 = judges sesquisyllables to be heavier than one syllable). While statistical analysis would be premature, **two patterns stand out:** listeners who judge sesquisyllables to be heavier than one syllable are slower overall and participants with strongly monosyllabic judgments are slower to respond to Speaker 1's monosyllabic tokens of trimoraic words than to Speaker 2's disyllabic tokens.

Discussion: The data do not currently support our hypothesis that similarity between listener and speaker phonology predicts ease of lexical access. Instead, they may suggest multiple *modes* of lexical access. Participants who were slower overall were unaffected by the dialectal variants in the experiment while participants who were faster overall were slowed by Speaker 1's 'non-standard' monosyllabic, monophthongized tokens of sesquisyllables. The first group may employ a less conservative mode of lexical access in which they tolerate greater dissimilarity between target and lexical candidates, resulting in slower winnowing of the candidate set but less sensitivity to variation. The second group may employ a more conservative mode of lexical access in which they require greater similarity between target and lexical candidate, winnowing the candidate set more quickly but resulting in back-tracking just in case the target differs too much from their representation of the intended item. Why these 'modes' correlate with phonological judgments about the weight of sesquisyllables is not immediately clear; both may be correlated with a third factor, such as experience with heterogeneous dialects (Sumner and Samuel 2009). The current study raises several questions about the impact of phonological variation on lexical access, and more work is needed.

References

Labov, W., Ash, S., and Boberg, C. (2006). *Atlas of North American English: Phonology and Phonetics*. Mouton de Gruyter.

Lavoie, L. and Cohn, A. (1999). Sesquisyllables of English: The structure of vowel-liquid syllables. In *Proceedings of the XIVth International Congress of Phonetic Sciences*.

Sumner, M. and Samuel, A. G. (2009). The effect of experience on the perception and representation of dialect variants. *Journal of Memory and Language*, 60(4):487–501.

Tilsen, S., Cohn, A., and Ricciardi, E. (2014). Syllable count judgments and durations of liquid rimes in English. *Cornell Working Papers in Phonetics and Phonology*.