Effects of vowel duration and lexical frequency in perceptual word identification

Chelsea Sanker sanker@stanford.edu
California Meeting on Psycholinguistics, 2024

Introduction

- Higher frequency words are produced with shorter duration (e.g. Jurafsky et al 2002)
- Listeners more often identify ambiguous stimuli as higher frequency words (Luce & Pisoni 1998)
- Potential accounts include Exemplar Theory (e.g. Connine et al 2008) and the Neighborhood Activation Model (e.g. Luce & Pisoni 1998)

This study:

- Examines what underlies frequency effects in perceptual decisions, based on interactions with vowel duration and background noise

Methodology

Participants: 9 native speakers of English

Task: Word identification in noise

- identifying each stimulus as matching one of two written English words
- response options differed phonologically just in the initial consonant

Stimuli: 208 items (26 response option pairs)

- *lexical frequency* (high, low)
- *vowel duration* (20% shortened, 20% lengthened)
- *noise* (quiet noise, louder noise)
- seven consonant decision types: b-d, b-g, b-p, d-t, k-p, k-t, p-t
- for each consonant decision type, there was one pair for which each consonant was in the higher frequency word (e.g. pack-tack, peach-teach)

Regression model

	Estimate	SE	z value	p value
(Intercept)	2.6848	0.3925	6.840	< 0.001
VowelDuration Short	0.5725	0.2141	2.674	0.00749
NoiseType Quieter	0.2151	0.2130	1.010	0.31245
StimulusFrequency Lower	-5.4678	0.5311	-10.295	< 0.001
VowelDuration Short * NoiseType Quieter	-0.5247	0.3016	-1.740	0.08189

Table 1: Logistic mixed effects model for responses of the higher frequency word (e.g. pack vs tack, teach vs peach). Reference levels: VowelDuration = Short, NoiseType = Louder, StimulusFrequency = Higher. Random intercepts for participant and stimulus word.

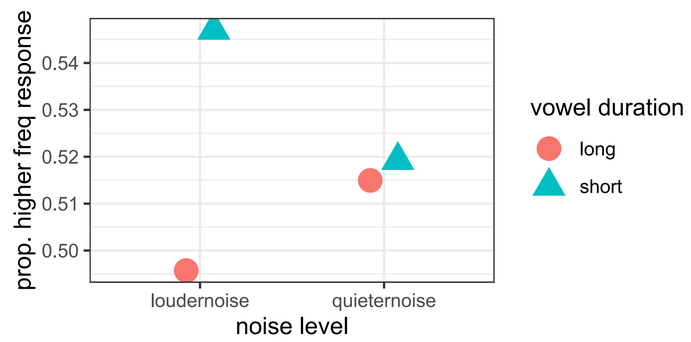


Figure 1: Proportion of the higher frequency response by noise type and vowel duration.

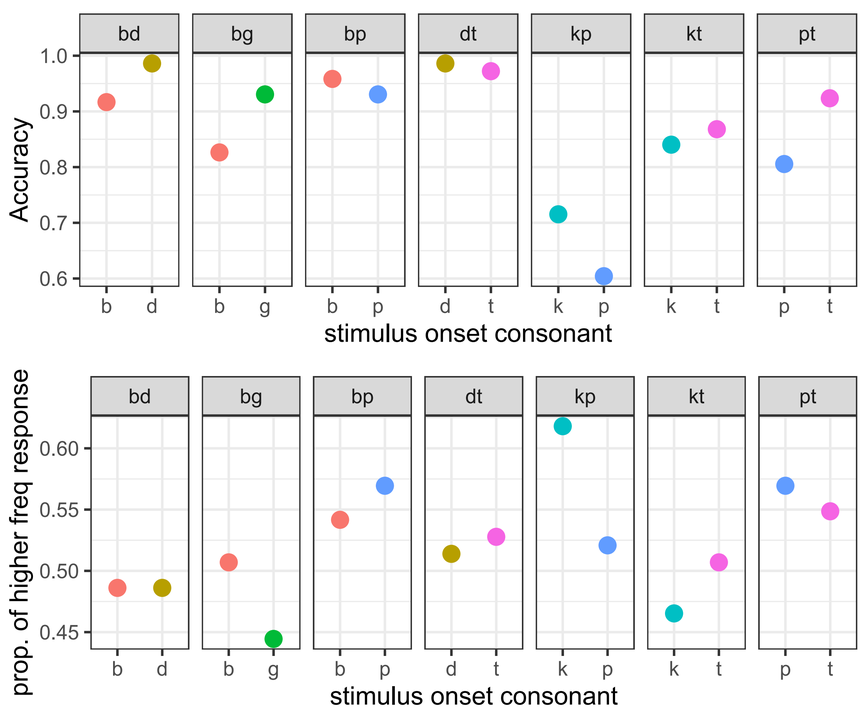
- Listeners are more likely to select the higher frequency response when the stimulus has a short vowel (53.3% of responses) rather than a long vowel (50.5%)
- This effect of vowel duration was only apparent with louder masking noise
- Louder noise doesn't have a clear overall effect on the frequency of the selected response (cf. e.g. Sommers et al 1997)

Effects of the particular consonants and vowels

Misperceptions are sometimes directional

This can be particularly important if the lexical frequency of words containing particular phonemes also varies

Figure 2:
Top: Accuracy of identifications by consonant decision and stimulus consonant.
Bottom: Proportion of the higher frequency response by consonant decision and stimulus consonant.



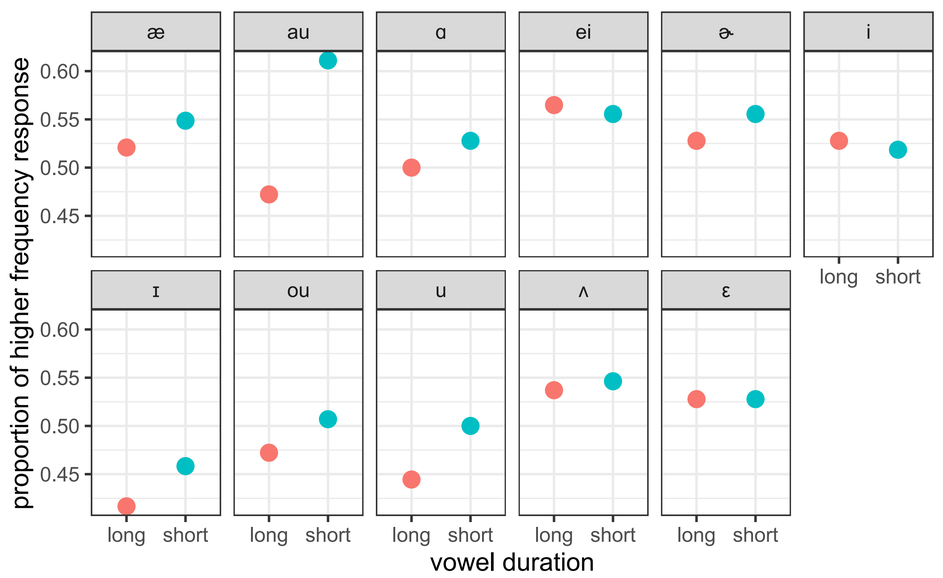


Figure 3: Proportion of the higher frequency response by vowel quality and vowel duration.

Conclusions

- **Exemplar account**: Words are associated with phonetically detailed instances of their form, so if a word is often produced with a short vowel (due to lexical frequency), listeners would be more likely to identify stimuli with short vowels as being that word
Short vowels produce a higher likelihood of high frequency response, but for long vowels, the odds are even – an exemplar explanation would predict higher odds of a low frequency response for a stimulus with a long vowel
- **Neighborhood Activation account**: Higher resting activation in higher frequency words allows poorer acoustic matches to raise activation levels to threshold during lexical access
The effect of vowel duration could be explained by longer duration producing greater activation of words containing that vowel (with masking noise, all stimuli are weaker acoustic matches) – a greater contribution from the stimulus will reduce the effect of resting activation level

References

Connine, C., Ranbom, I., & Patterson, D. 2008. Processing variant forms in spoken word recognition: The role of variant frequency. Perception & Psychophysics 70(3). 403-411.
Jurafsky, D., Bell, A., & Girand, C. 2002. The role of the lemma in form variation. In C. Gussenhoven & N. Warner (eds.), Laboratory Phonology VII, 3-34. Berlin: Mouton de Gruyter.
Luce, P., & Pisoni, D. 1998. Recognizing spoken words: The neighborhood activation model. Ear and Hearing 19(1). 1-36.
Sommers, M., Kirk, K., & Pisoni, D. 1997. Some considerations in evaluating spoken word recognition by normal-hearing, noise-masked normal-hearing, and cochlear implant listeners. I: The effects of response format. Ear and Hearing 18(2). 89-99.