

# Frequency-predicted shifts don't require word-specific phonetic details

Chelsea Sanker

chelsea.sanker@yale.edu

6 December 2021

Fifth Edinburgh Symposium on Historical Phonology

# Phonetic details in the phonological representation

- Some changes seem to depend on lexical frequency, potentially reflecting word-specific exemplar memories and lexical diffusion (Pierrehumbert 2002, Bybee 2012)
- But frequency conditioning can also be explained by higher frequency facilitating activation, causing more reduction (Kahn & Arnold 2012, Jurafsky et al. 2002)
- These accounts make different predictions about word-specific behavior

# Testing word-specific representations: Convergence

Convergence provides a test of whether phonetic details can differ in the representations of different words

Some studies find more convergence in lower frequency words, perhaps suggesting word-specific exemplar memories (Goldinger 1998; Babel 2010; Nielsen 2011; Dias & Rosenblum 2016), but:

- Other studies have failed to replicate the effect (Pardo et al. 2013; Pardo et al. 2017)
- Different behaviors based on lexical frequency are an indirect test of word-specific convergence
- In a direct test of exposure to different words shifted in different directions, Sanker (2021) finds no evidence for word-specific convergence

# This study

- Previous work has shown that VOT is shorter in higher-frequency words (e.g. Yao 2009; Chodroff & Wilson 2017)
- We can test whether frequency itself is impacting expected VOT or if listeners are remembering the particular VOT that they have heard for each word
- By manipulating both (1) frequency, based on recent repeated exposure and (2) the VOT in exposure items

# This study

I present results from a word identification task:

- How does exposure to particular lexical items influence listeners' VOT category boundary?
  - The category boundary is lowered after repeated exposure to the word, even if exposure items had lengthened VOT
- What do the effects suggest about frequency-related sound change?
  - Suggests that frequency-based differences in VOT are caused not by word-specific exemplars but by frequency influencing speed of access and corresponding reduction

# Participants

Participants were:

- 96 native speakers of American English (mean age 27.8; 36 male, 58 female, 2 nonbinary)
- No reported speech or hearing disorders
- Participated online through Qualtrics, paid through Prolific

# Task design

- First an **exposure phase**: Categorizing vowel duration for certain words (the categorization task was just to ensure that participants listened closely to the exposure items)
- Then a **testing phase**: Listening to words in isolation and identifying each as either starting with an aspirated or unaspirated stop (e.g. *town* vs *down*)

# Exposure Phase

In the exposure phase:

- Participants heard a word and categorized the vowel in it as long or short in duration
- Three exposure conditions for each word (balanced across participants), all starting with aspirated stops:
  - ① lengthened VOT (mean 137 ms), 4 words
  - ② shortened VOT (mean 51 ms), 4 words
  - ③ no exposure (control), 16 words
- Each exposure word appeared 9 times
- For each participant, the manipulation was consistent for all instances of a word, e.g. *town*, *pan*, *pet*, *could* with long VOT, and *tomb*, *peep*, *cob*, *cull* with short VOT



# Test Phase

In the test phase:

- Participants heard a word and identified it as one of two words differing only in aspiration of the initial stop (e.g. *town*, *down* – both items roughly matching in lexical frequency)
- 24 items manipulated along a 3-step VOT continuum (72 total stimuli)
  - The small number of steps was to reduce the possibility that additional exposure might obscure effects of the exposure manipulation
- 8 of the voiceless aspirated words had appeared during exposure (4 lengthened VOT, 4 shortened VOT), and 16 had not appeared
- Half high-frequency lexical items, half low-frequency

# Hypothesis: Word-specific convergence

Hypothesis A: Listeners' expectations will align with the VOT heard for each word during exposure – longer VOT for words heard with lengthened VOT and shorter VOT for words heard with shortened VOT

# Hypothesis: Shortening based on any recent exposure

Hypothesis B: Recent repeated exposure to a word will make listeners retrieve it more quickly, resulting in shorter expected duration (including VOT), regardless of whether the exposure tokens had shortened or lengthened VOT

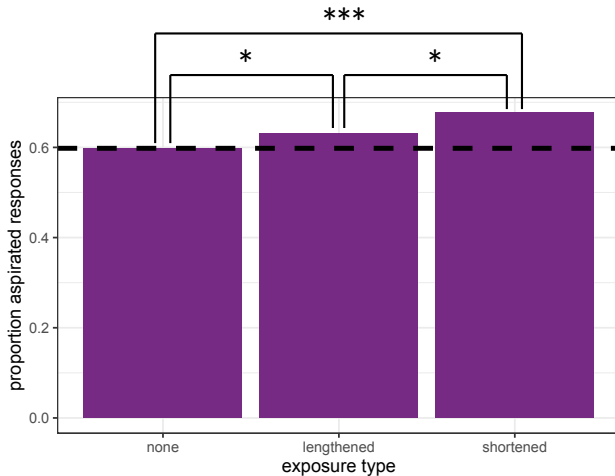
# Logistic mixed effects model for selecting the voiceless aspirated onset

	Estimate	Std. Error	z value	p value
(Intercept)	0.514	0.161	3.19	0.00145
ExposureType Lengthened	0.186	0.0764	2.43	0.0151
ExposureType Shortened	0.425	0.0779	5.46	< 0.0001

*Reference Levels: ExposureType = None*

Listeners accepted the aspirated stop as having a shorter VOT when they had recently been exposed to that word, both with shortened VOT exposure and lengthened VOT exposure

# Proportion of aspirated responses



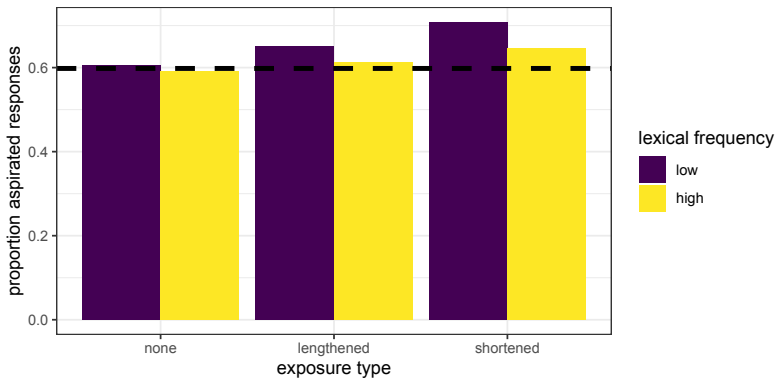
# Effects of Lexical Frequency

There was no significant interaction between lexical frequency and the effect of exposure.

	Estimate	Std. Error	z value	p value
(Intercept)	0.541	0.198	2.74	0.00621
ExposureType Lengthened	0.299	0.116	2.58	0.00986
ExposureType Shortened	0.546	0.114	4.79	< 0.0001
FrequencyCategory High	-0.053	0.229	-0.231	0.817
ExpType Lengthened : FreqCat High	-0.219	0.166	-1.32	0.188
ExpType Shortened : FreqCat High	-0.247	0.169	-1.46	0.144

*Reference Levels: ExposureType = None, FrequencyCategory = low*

# Proportion of aspirated responses, incl. frequency



# Caveat

Caveat: Since this is perception data, we can't exclude the possibility of the results reflecting a response preference



## Word specific detail?

- The category boundary is lowered in both VOT manipulation conditions, indicating that the shift is due to increased local frequency
- Differences in VOT are based on this local frequency; salient words are accessed more easily and thus have shorter expected duration
- Suggests that words do not have distinct VOT targets, but lexical frequency impacts the realization of VOT and expectations in perceptual access

# Shortening vs Lengthening

- Though both conditions result in lower VOT boundaries, the effect is larger with exposure to shortened VOT
- May suggest that listeners do have word-specific VOT targets, in addition to frequency itself affecting VOT
- Might have alternative explanations, e.g. a repeated word with short VOT has consistent cues for local frequency while they are in conflict if the repeated item has long VOT

# Implications for sound change

Apparent frequency-sensitive sound changes can be explained without requiring each word to have independent phonetic details:

- Reduction-driven shifts may first be apparent in high-frequency words, but the prototype for the entire category is shifted
- Uniform underlying phonetic targets across words, with differences in form caused by processes in lexical access

# References

- Babel, M. 2010. Dialect divergence and convergence in New Zealand English. *Language in Society*, 39, 437–456.
- Bybee, J. 2012. Patterns of lexical diffusion and articulatory motivations for sound change. In M.-J. Solé & D. Recasens (Eds.), *The initiation of sound change: Perception, production, and social factors* (pp. 211–234). Amsterdam: John Benjamins.
- Chodroff, E., & Wilson, C. 2017. Structure in talker-specific phonetic realization: Covariation of stop consonant VOT in American English. *Journal of Phonetics*, 61, 30–47.
- Dias, J. W. & Rosenblum, L. D. 2016. Visibility of speech articulation enhances auditory phonetic convergence. *Attention, Perception, & Psychophysics*, 78, 317–333.
- Goldinger, S. D. 1998. Echoes of echoes? An episodic theory of lexical access. *Psychological Review*, 105(2), 251–279.
- Jurafsky, D., Bell, A., & Girand, C. 2002. The role of the lemma in form variation. In C. Gussenhoven, & N. Warner (Eds.), *Laboratory Phonology VII* (pp. 3–34). Berlin: Mouton de Gruyter.
- Kahn, J.M., & Arnold, J.E. 2012. A processing-centered look at the contribution of givenness to durational reduction. *Journal of Memory and Language*, 67, 311–325.
- Nielsen, K. 2011. Specificity and abstractness of VOT imitation. *Journal of Phonetics*, 39(2), 132–142.
- Pardo, J.S., Jordan, K., Mallari, R., Scanlon, C., & Lewandowski, E. 2013. Phonetic convergence in shadowed speech: The relation between acoustic and perceptual measures. *Journal of Memory and Language*, 69, 183–195.
- Pardo, J.S., Urmanche, A., Wilman, S. & Wiener, J. 2017. Phonetic convergence across multiple measures and model talkers. *Attention, Perception, & Psychophysics*, 79, 637–659.
- Pierrehumbert, J. 2002. Word-specific phonetics. In C. Gussenhoven & N. Warner (Eds.), *Laboratory Phonology VII* (pp. 101–139). Berlin: Mouton de Gruyter.
- Sanker, C. 2021. Convergence doesn't show lexically-specific phonetic detail. *Supplementary proceedings of the 2020 Annual Meeting on Phonology*.
- Yao, Y. 2009. Understanding VOT variation in spontaneous speech. *UC Berkeley Phonology Lab Annual Report*, 5, 29–43.