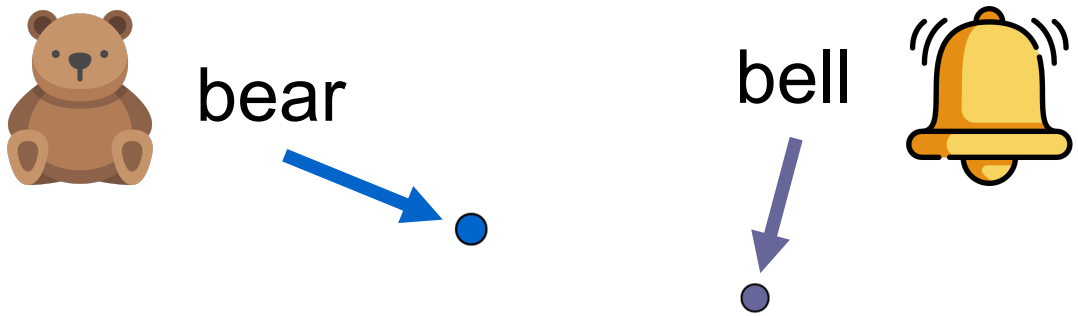
CAMP[7]

Effects of Contrastive Hyper-articulation and Word Predictability on VOT Production

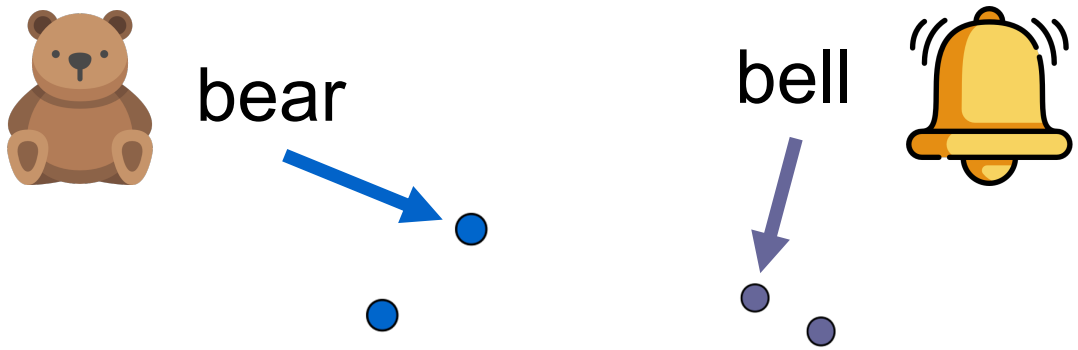
Will Chih-Chao Chang & Connor Mayer

Phonetic Variation in Speech Production



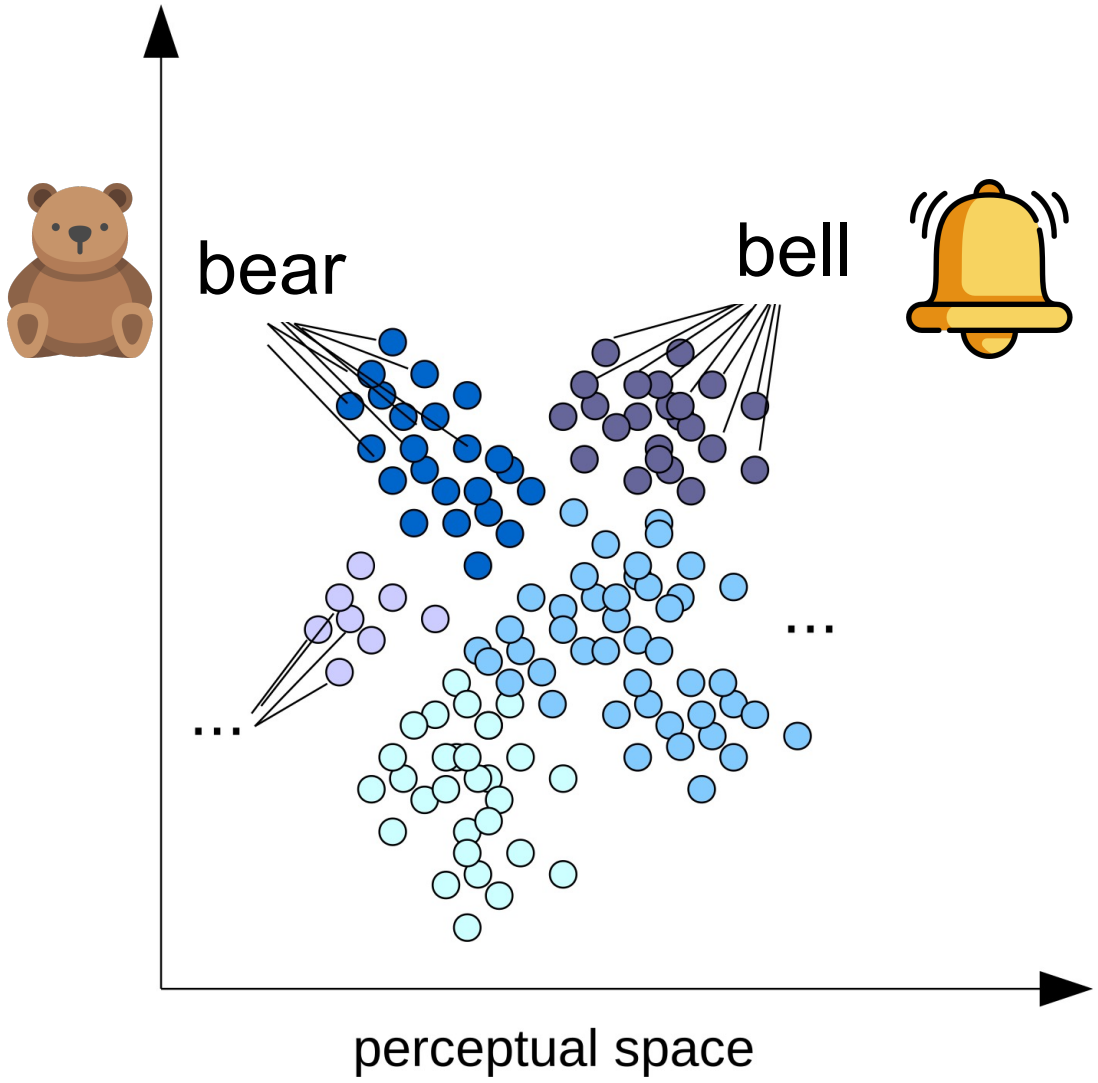
(Schweitzer, 2019)

Phonetic Variation in Speech Production



(Schweitzer, 2019)

Phonetic Variation in Speech Production



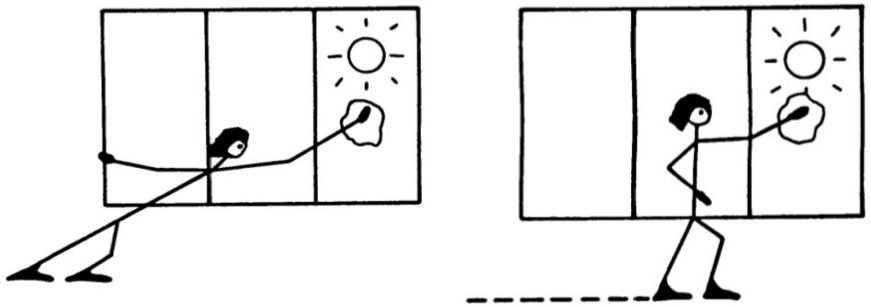
(Schweitzer, 2019)

Hyper- and Hypo-articulation (H&H) Theory

Hypo-articulation ⟷ Hyper-articulation

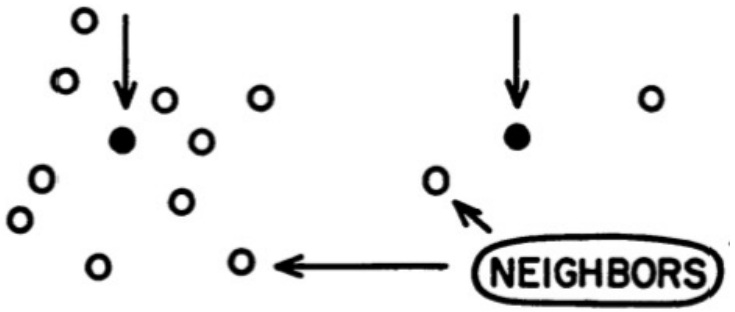
System-oriented control

To reduce behavior cost for speakers



Output-oriented control

To increase signal contrast for listeners



(Lindblom, 1990)

Lexical Competition and Contrastive Hyper-articulation

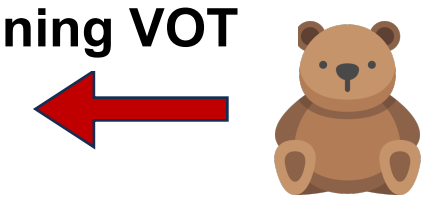
Hypo-articulation

Hyper-articulation

More extreme VOT for words with minimal pair competitors in a read-aloud task, and even more so in a conversational context (Baese-Berk & Goldrick, 2009)

Contrastive hyper-articulation is implemented to increase the perceptual distance between target words and phonetically-specific minimal pair competitors (Nelson & Wedel, 2017)

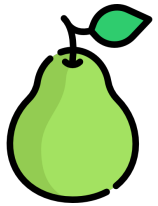
Shortening VOT



bear

bell

Lengthening VOT



pear

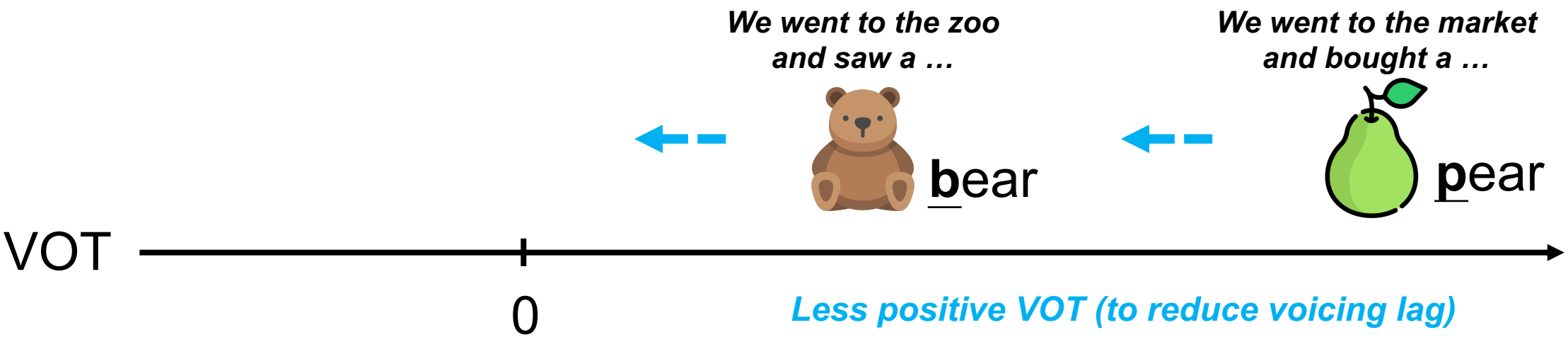
*p*ell

Predictability-based Phonetic Reduction

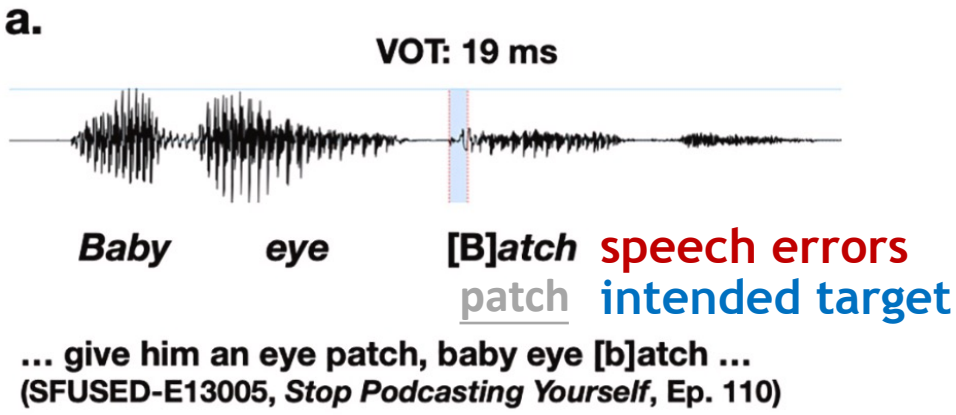
Hypo-articulation — Hyper-articulation

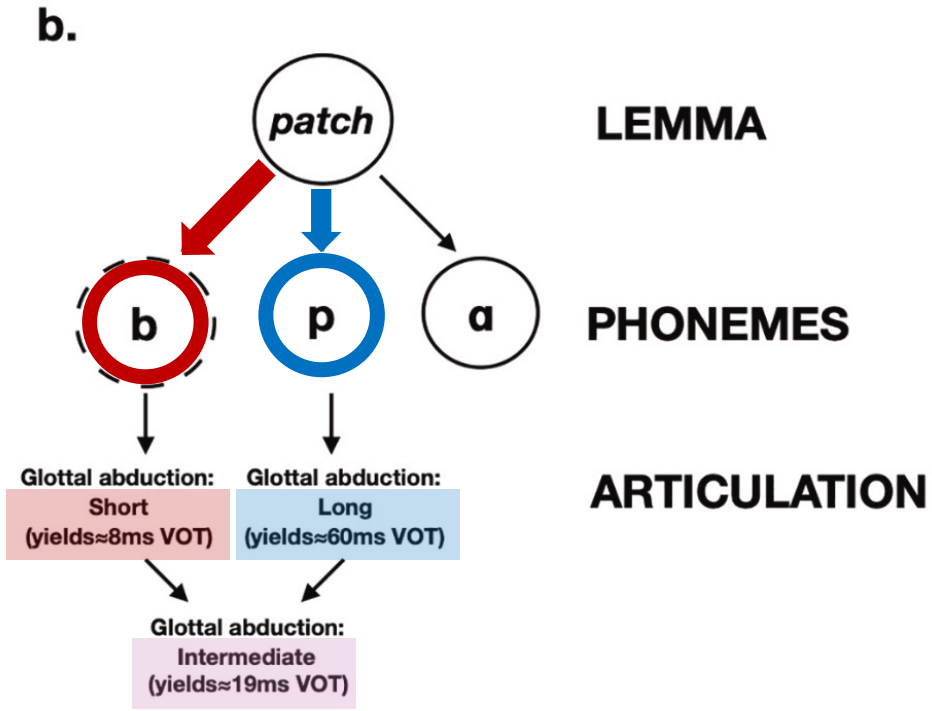
Reduced word forms and less strongly articulated segments when linguistic units are contextually more predictable (Probabilistic Reduction Hypothesis, Jurafsky et al., 2001; Baker & Bradlow, 2009; Hall et al., 2018)

For stop consonants, short-lag VOT requires less complex neuromuscular control compared to pre-voicing or long-lag VOT (Buckingham, 1998)



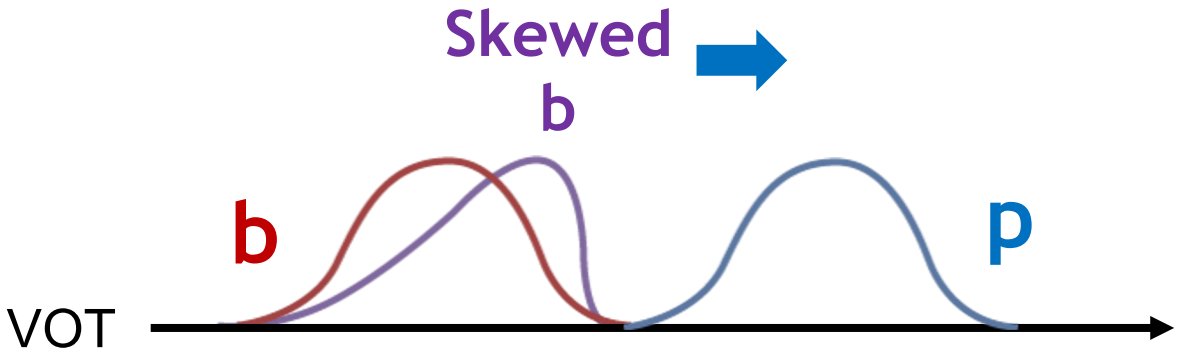
Production-internal Interaction Account: Cascading activation





Both experimentally-induced and spontaneous **speech errors leave acoustic "traces" of the intended target**

→ the cascade of partially activated phonological representations of the target consonant into articulatory processes



(Goldrick & Blumstein, 2006; Alderete et al., 2021)

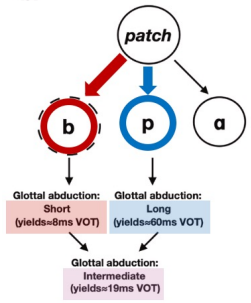
Interim Summary

H&H Theory

Phonetic realizations are adjusted based on communicative goals to enhance signal contrast for listeners (hyper-articulation) and to reduce articulatory efforts for speakers (hypo-articulation).

Predictability-based Phonetic Reduction

Reduced word forms and less strongly articulated segments when linguistic units are contextually more predictable.

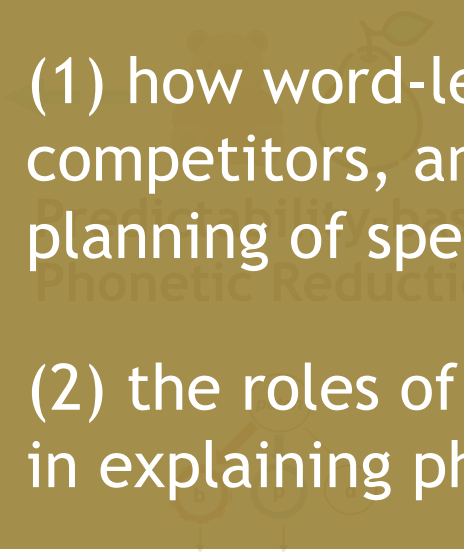


Production-Internal Cascading Activation

Cascading activation between lexical, phonological and phonetic representations can influence the phonetic properties of speech sounds, which makes the pronunciation of target words to be skewed toward non-target (but somewhat activated) words.

Interim Summary

Phonetic realizations are adjusted based on communicative goals to enhance signal contrast for listeners (hyper-articulation) and reduce articulatory effort for speakers (hypo-articulation).



It remains unclear:

(1) how word-level lexical factors, such as the existence of minimal pair competitors, and probabilistic factors, like word predictability, interact in the planning of speech production

(2) the roles of communicative goals and production-internal cascading activation in explaining phonetic variation

Reduced word forms and less strongly articulated segments when linguistic units are contextually more predictable

Cascading activation between lexical, phonological and phonetic representations can influence the phonetic properties of speech sounds, which makes the pronunciation of target words to be skewed toward non-target (but somewhat activated) words.

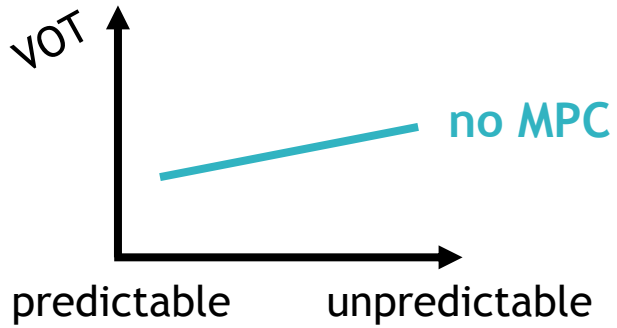
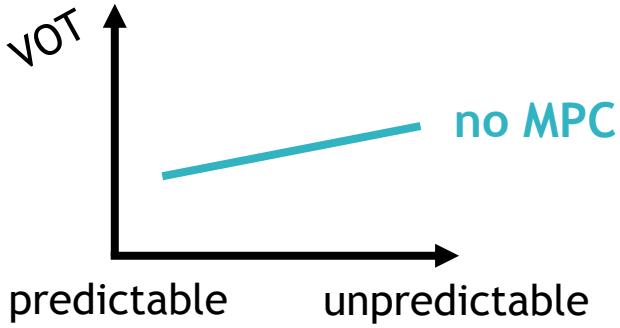
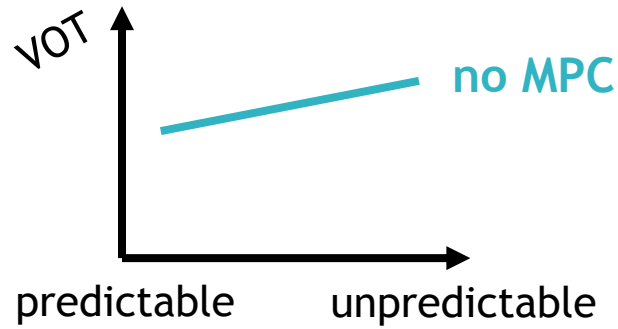
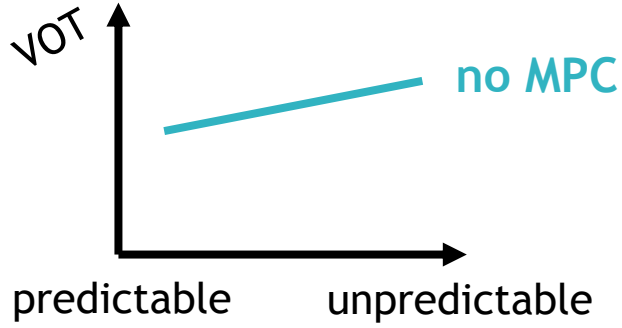
Research Questions and Predictions

Does word predictability affect the phonetic realization of stop consonants differently for words with or without minimal pair competitor (*MPC*)? → VOT in English voiced and voiceless stops

Voiced stops (no pre-voicing; positive VOTs)		
Voiceless stops		

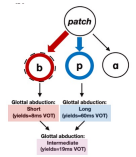
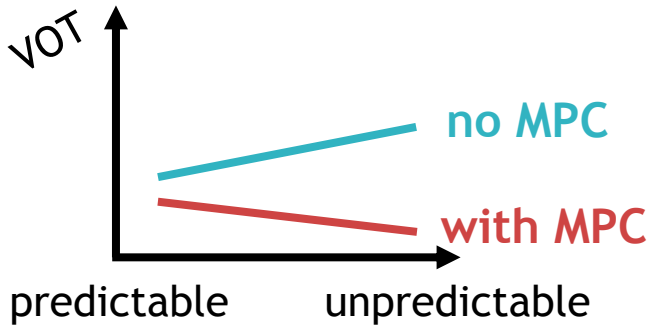
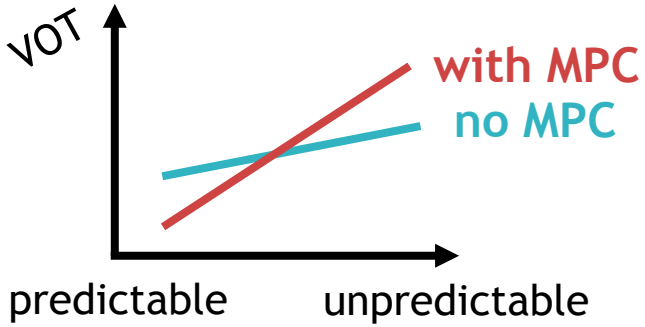
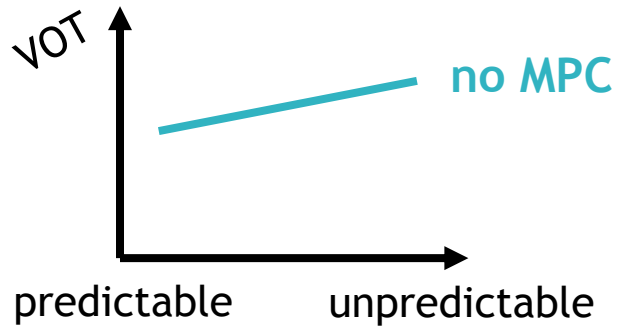
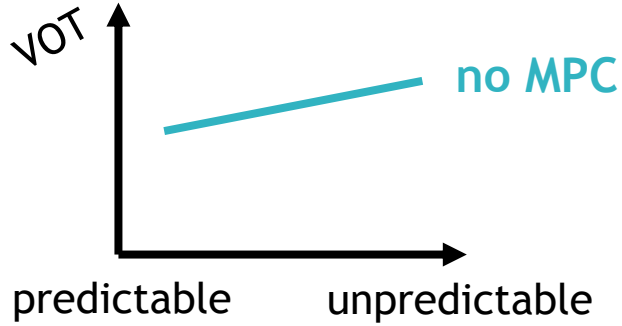
Research Questions and Predictions

Does word predictability affect the phonetic realization of stop consonants differently for words with or without minimal pair competitor (*MPC*)? → VOT in English voiced and voiceless stops

	Predictability-based Phonetic Reduction ←	
Voiced stops (no pre-voicing; positive VOTs)		
Voiceless stops		

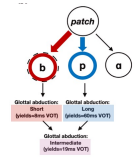
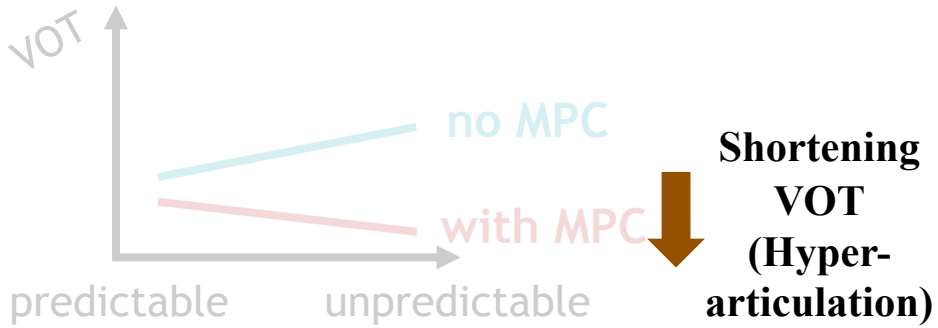
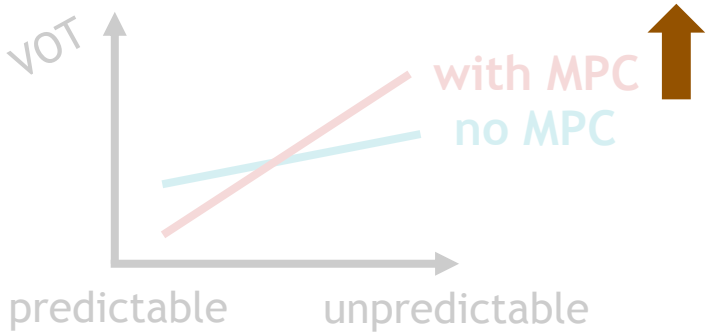
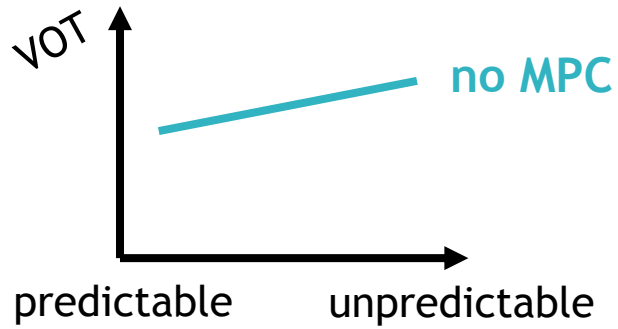
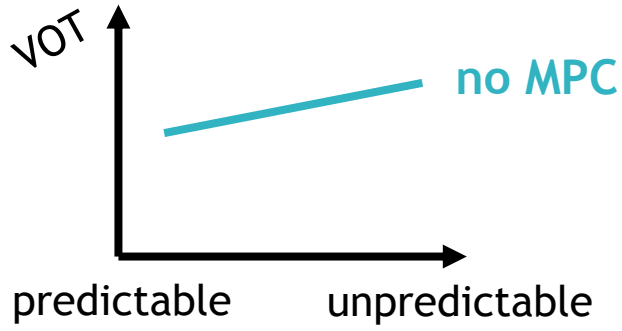
Research Questions and Predictions

Does word predictability affect the phonetic realization of stop consonants differently for words with or without minimal pair competitor (*MPC*)? → VOT in English voiced and voiceless stops

	Predictability-based Phonetic Reduction	
	Hypo-articulation ⟷ Hyper-articulation **H&H Theory**	 **Production-Internal Cascading Activation**
Voiced stops (no pre-voicing; positive VOTs)		
Voiceless stops		

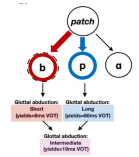
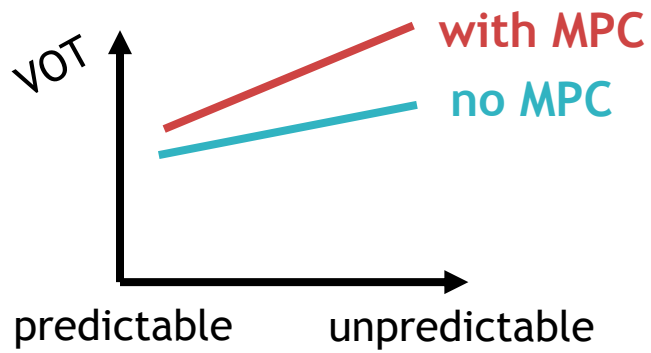
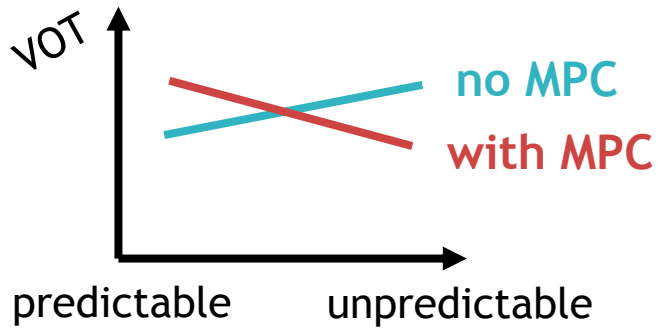
Research Questions and Predictions

Does word predictability affect the phonetic realization of stop consonants differently for words with or without minimal pair competitor (*MPC*)? → VOT in English voiced and voiceless stops

	Predictability-based Phonetic Reduction	
	Hypo-articulation ⟷ Hyper-articulation **H&H Theory**	 **Production-Internal Cascading Activation**
Voiced stops (no pre-voicing; positive VOTs)	 **Shortening VOT (Hyper-articulation)**	 **Longer VOT (skewed toward voiceless competitors)**
Voiceless stops		

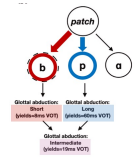
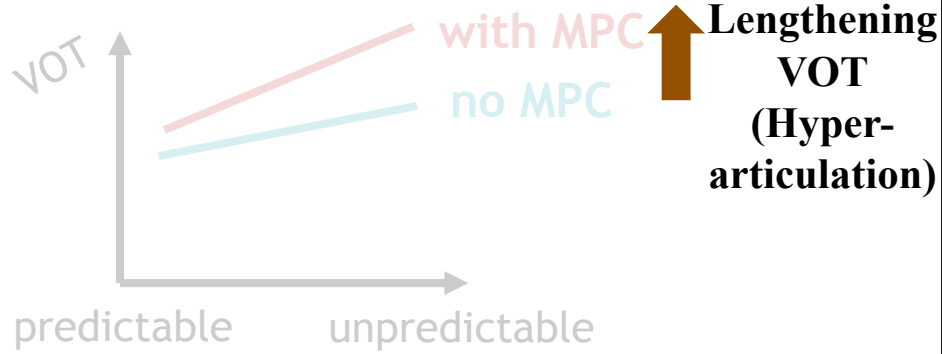
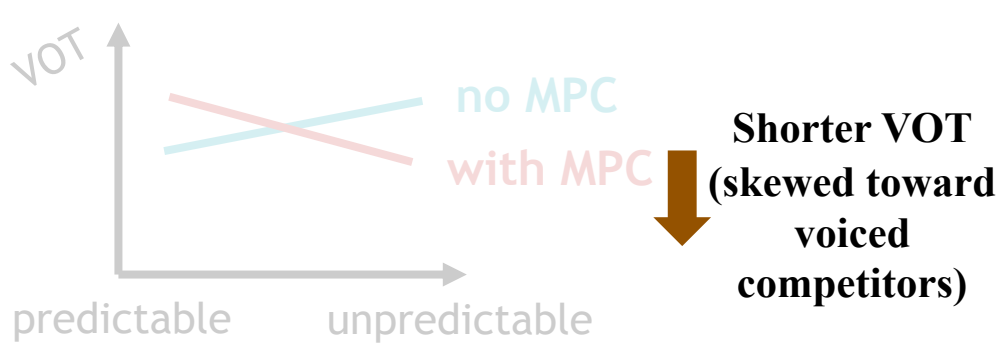
Research Questions and Predictions

Does word predictability affect the phonetic realization of stop consonants differently for words with or without minimal pair competitor (*MPC*)? → VOT in English voiced and voiceless stops

	Predictability-based Phonetic Reduction	
	Hypo-articulation ⟷ Hyper-articulation **H&H Theory**	 **Production-Internal Cascading Activation**
Voiced stops (no pre-voicing; positive VOTs)		
Voiceless stops		

Research Questions and Predictions

Does word predictability affect the phonetic realization of stop consonants differently for words with or without minimal pair competitor (*MPC*)? → VOT in English voiced and voiceless stops

	Predictability-based Phonetic Reduction	
	Hypo-articulation ⟷ Hyper-articulation **H&H Theory**	 **Production-Internal Cascading Activation**
Voiced stops (no pre-voicing; positive VOTs)		
Voiceless stops		

Method

Corpus data: Mixer 6 corpus (Brandschain et al. 2010, Brandschain et al. 2013; Chodroff et al. 2016); 45-minute reading speech, 179 AE speakers (102 female)

Dataset: Xie et al. (2023)'s analysis originally from Chodroff & Wilson (2018) on word-initial stops; full sentences reconstructed from the Mixer 6 corpus.

- include only lexical words, not function words
- for each stop consonant, tokens with VOT values more than 2.5 standard deviation from the specific speaker's mean were excluded (Chodroff & Wilson, 2018)
  → 12699 voiced stop tokens and 18423 voiceless stop tokens in total
- positive VOTs were automatically extracted by AutoVOT (Keshet et al., 2014)

Method

Fixed factors:

(1) Existence of Minimal Pair Competitor (MPC):

- Carnegie-Mellon University Pronouncing Dictionary that differed only in the voicing of the initial segment, and that had a CELEX frequency > 1
- 1 = words with MPC; 0 = words without MPC

(2) **Target Surprisal (TS)**

- the negative log (base 2) probability of each target word conditioned on the preceding context in each sentence, calculated by GPT-2 language model (Radford et al., 2019) using `surprisal` package (Sathe, 2023)

Covariates:

Place of articulation, Number of syllables, Position in utterance, Speaking rate, Lemma frequency, Voicing of preceding segment, Height of following vowel

All continuous variables were centered, and then linearly scaled to a range between -1 and 1 to facilitate model convergence (Nelson & Wedel, 2017)

Results – Linear Mixed-Effect Models

Models were fitted for voiced and voiceless stops separately:

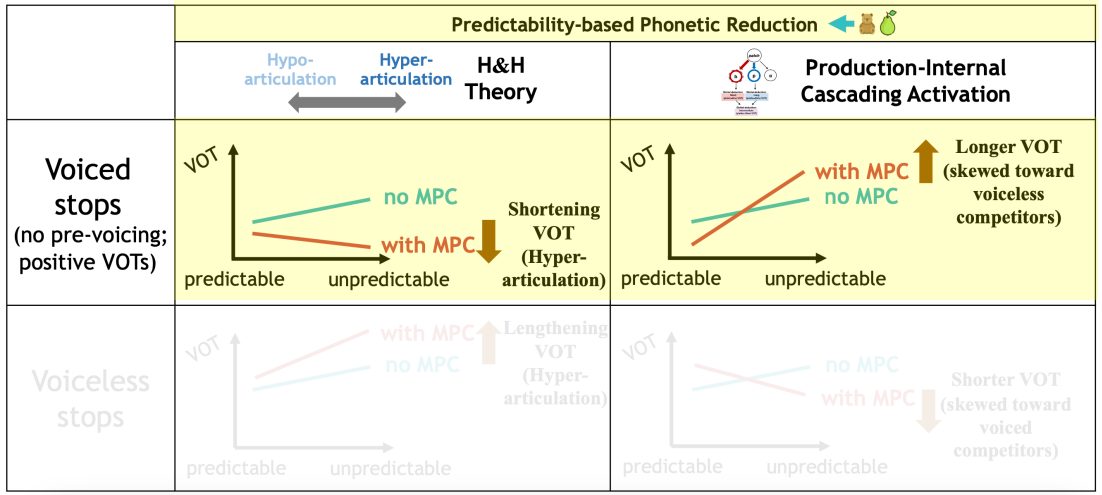
Model for voiced stops:

lmer(vot ~ minimal_pair*target_surprisal + poa + pos + syll + spk_rate + Freq_CobS_Lemmas + preceding_segment_voicing + following_vowel_height + (1 + target_surprisal | subj) + (1 | word), data=vcd_df)

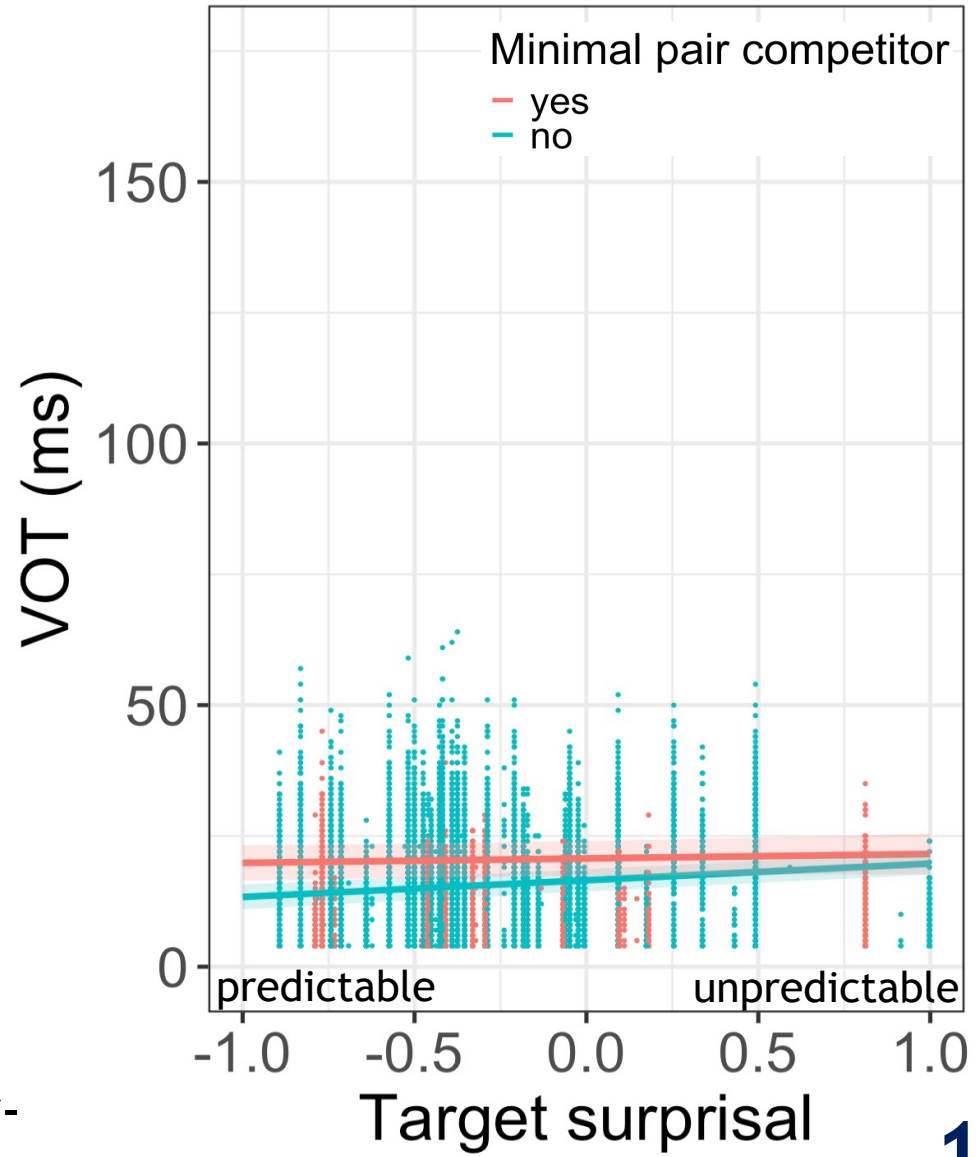
Model for voiceless stops:

lmer(vot ~ minimal_pair*target_surprisal + poa + pos + syll + spk_rate + Freq_CobS_Lemmas + preceding_segment_voicing + following_vowel_height + (1 + minimal_pair + target_surprisal | subj) + (1 | word), data=vcl_df)

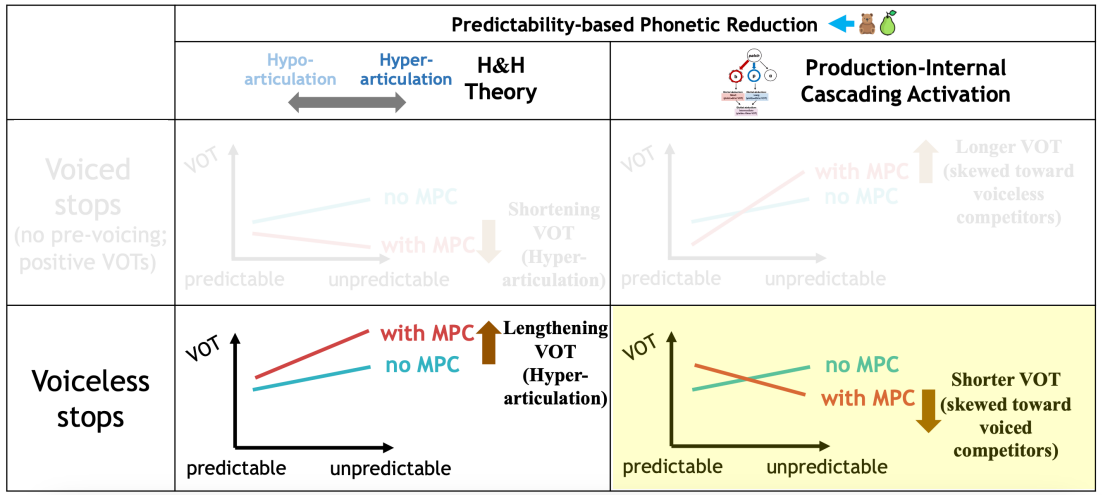
Results - Voiced stops



- *Longer* VOTs when words have MPC (β_{MPC}=4.171, p<.01)
  → Production-internal cascading activation skewing VOT values toward voiceless counterparts
- *Longer* VOTs when words without MPC are less predictable (β_{TS}=3.184, p<.001)
- However, VOTs are longer in words with MPC when they are less predictable, but to a lesser extent, given the negative interaction effect (β_{MPC*TS}=-2.326)
  → Predictability effect is mediated by contrastive hyper-articulation to main contrasts

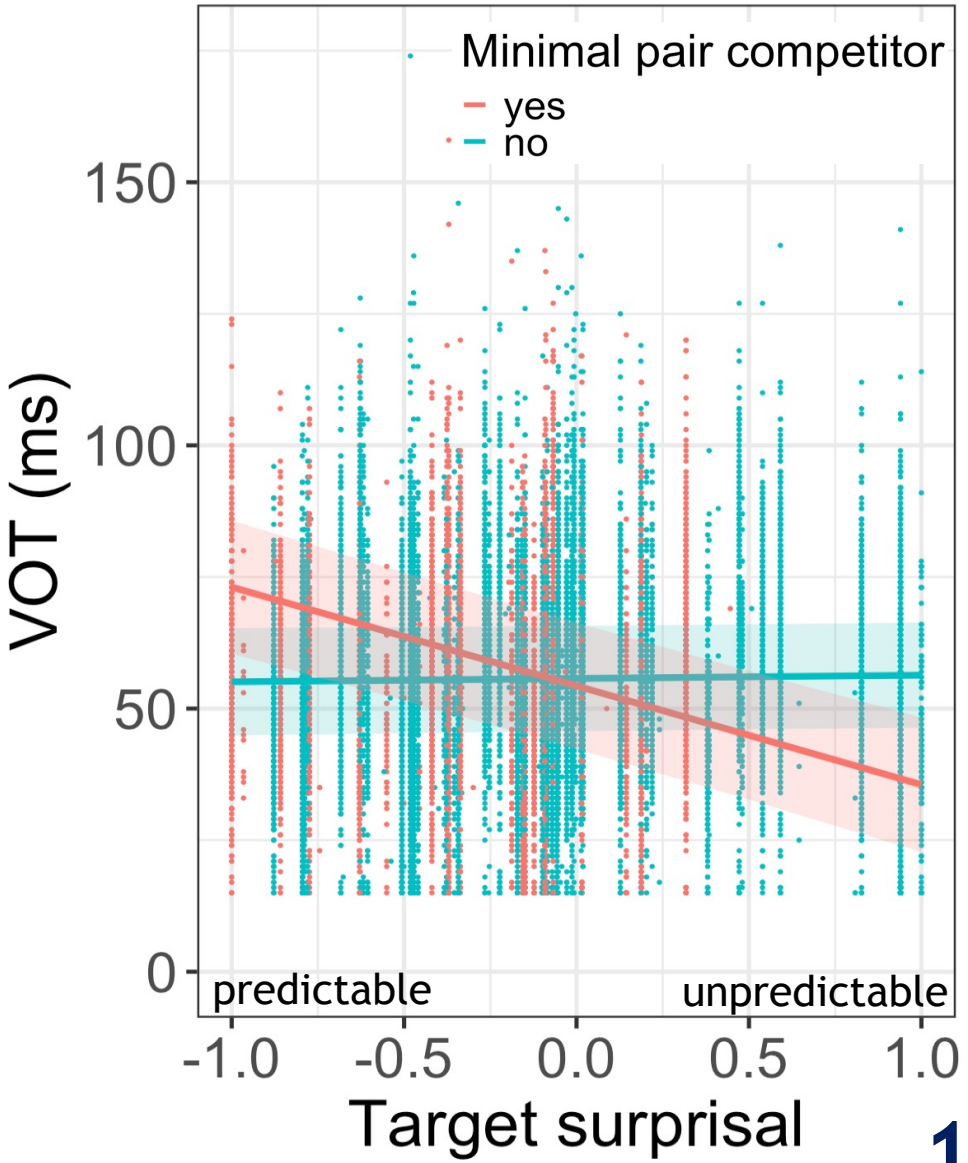


Results - Voiceless stops



- No main effects of MPC (β_{MPC}=-1.406, p=.64) and target surprisal (β_{TS}=0.639, p=.39)
- *Shorter* VOTs when words with MPC become less predictable (β_{MPC*TS}=-19.408, p<.001)

  → Relatively weaker activation of target words makes the cascading activation of competitor words more influential, leading to more skewed VOT values



Conclusion

Phonetic variations in speech production are conditioned by contrastive hyper-articulation, predictability-based phonetic reduction, and production-internal cascading activations.

For **voiced stops**, all three mechanisms contribute to the phonetic realization of VOTs, with predictability effect being mediated by contrastive hyper-articulation especially in less predictable contexts.

For **voiceless stops**, however, production-internal cascading activations primarily determine the phonetic realization of VOTs.

Conclusion

In the present study, production-internal cascading activation, rather than contrastive hyper-articulation, seems to better account for competition-driven phonetic variation.
→ Possibly due to lower communication demands in the reading task compared to a conversational context (cf. Baese-Berk & Goldrick, 2009; Nelson & Wedel, 2017).

Further investigation is needed to understand the complex interactions between lexical and probabilistic factors, as well as the potentially distinct phonetic implementations of voiced vs. voiceless stops.

CAMP[7]

Will Chih-Chao Chang
chihchc9@uci.edu

Connor Mayer
cjmayer@uci.edu

Thank you very much!