

Prosodic phrasing is integrated in segmental speech perception: Evidence from the Korean Accentual Phrase

Jeremy Steffman¹ Sahyang Kim² Taehong Cho³ Sun-Ah Jun⁴

¹Northwestern University ²Hongik University ³Hanyang University ⁴UCLA

January 9
LSA 2021

contact: jeremy.steffman@northwestern.edu

Introduction

Prosodic phrasing fine-tunes how speakers articulate segmental contrasts¹

- including cues in the temporal domain: VOT, vowel duration, etc.²

This talk is about how these influences translate into the **perception** of temporal cues

- this is currently an open question³

¹e.g., Keating 2006

²e.g., Cho 2015; Cole et al. 2007; Nakai et al. 2009; Shepherd 2008

³Mitterer et al. 2016; Steffman 2019

Introduction

Consider domain-initial lengthening of voice onset time (VOT):

- in e.g., English, VOT is longer at higher-level prosodic domains¹
- Q: do listeners integrate phrasing with their perception of VOT?

Kim and Cho (2013)

Does being phrase-initial versus -medial mediate VOT categorization?

IP-initial VOT

(Let's hear)IP (/p-ba/ now)IP

IP-medial VOT

(Let's hear /p-ba/ now)IP

Result: longer VOT required for a voiceless /p/ response when IP-initial (fewer IP-initial /p/ responses)

- listeners expect longer VOT following an IP boundary
- i.e. integrate phrasing with perception of VOT

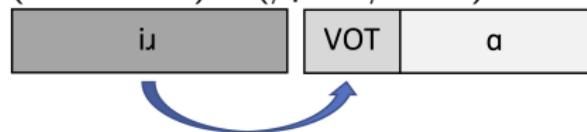
¹Cho and Keating 2009

Introduction

However, durational context might also explain this effect

IP-initial VOT

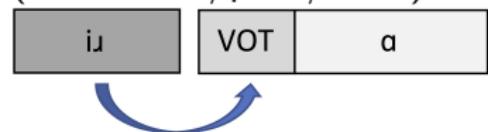
(Let's hear)IP (/p-ba/ now)IP



- longer rhyme preceding VOT
- short VOT percept, fewer /p/

IP-medial VOT

(Let's hear /p-ba/ now)IP



- shorter rhyme preceding VOT
- long VOT percept, more /p/

Thus two possible accounts¹

- ① listeners integrate prosodic phrasing with their expectations of how a cue should be realized (due to domain-initial strengthening)
- ② durational context effects (= general auditory processing²)

¹Mitterer et al. 2016; Steffman 2019

²Newman and Sawusch 2009; Wade and Holt 2005

This talk

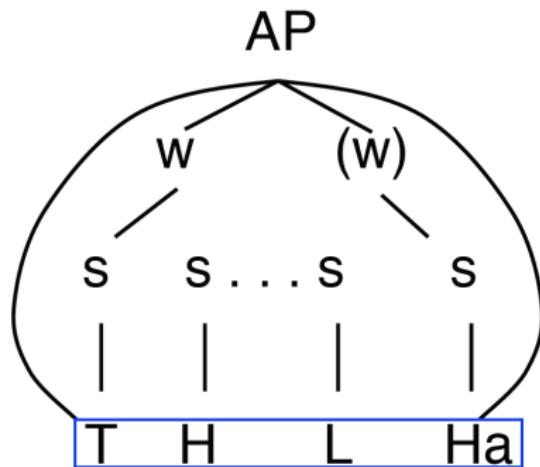
We test if a small phrasal domain in Seoul Korean influences perception of temporal cues

- phrasing cued only by pitch - eliminating durational context effects

Our goal:

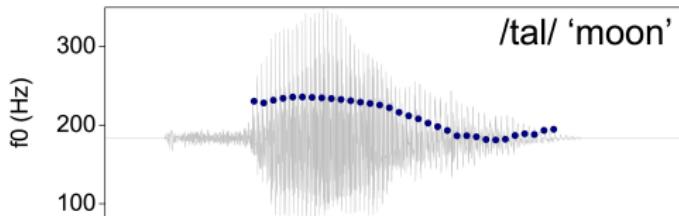
test if phrasing influences listeners' perception of temporal contrasts in a way that is definitively independent of durational context

The Korean Accentual Phrase (AP)¹

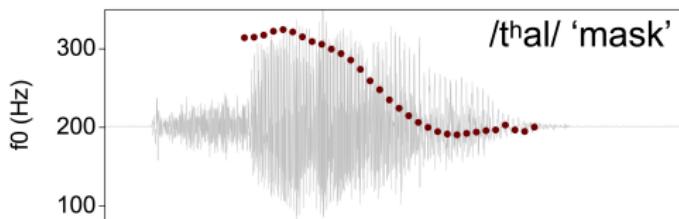


- initial tone dependent on segmental laryngeal features
- final H tone cues AP boundary

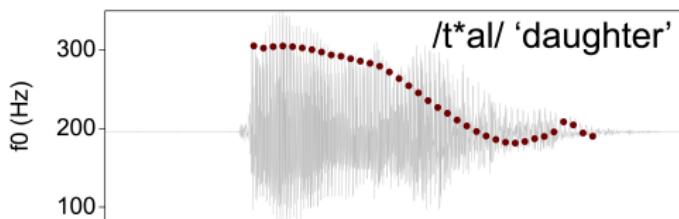
Korean laryngeal contrasts



Lenis/lax:
long lag VOT
L tone AP-initially
(not in this talk)



Aspirated:
long lag VOT
H tone AP-initially



Fortis/tense:
short lag VOT
H tone AP-initially
lengthened post-cons. vowel

Temporal cues

ASPIRATED



FORTIS

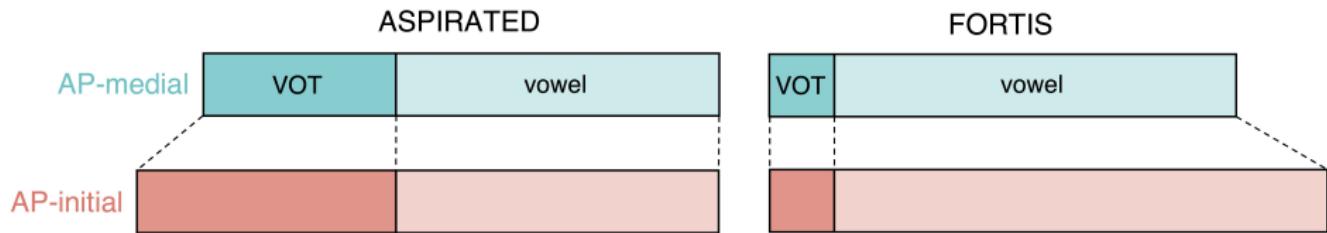


- ① longer VOT cues aspirated¹
- ② longer vowel duration cues fortis²

¹Kim et al. 2012

²Choi 2011

Prosodic modulation of cues: AP-initial strengthening



AP-initial stops show contrast enhancement

- aspirated: lengthening of AP-initial VOT¹
- fortis: lengthening of post-stop vowel in AP-initial CV²

¹Cho and Jun 2000; Keating et al. 2004

²Cho and Keating 2001

Experiment 1: VOT

Two alternative forced choice task

- VOT continuum from velar /k*i/ (30 ms) to /k^hi/ (80 ms)
- Target sound cued as either AP-initial or as AP-medial by changes in contextual pitch only
- Carrier phrase: “That’s what we inserted/grew”
(그거 우리가 끼웠어/키웠어)

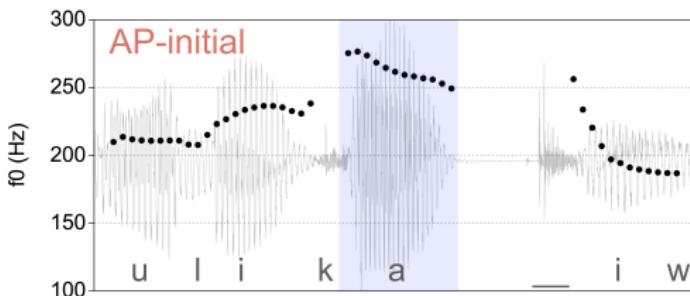
AP-initial

(kwkΛ)_{AP} (ulika)_{AP} (?iwa^s*Λ)_{AP}

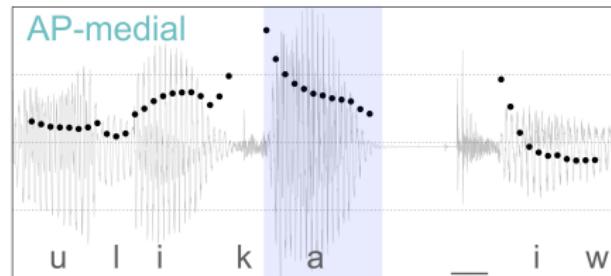
AP-medial

(kwkΛ)_{AP} (ulika ?iwa^s*Λ)_{AP}

Stimuli



(kukʌ) (ulika) (?iwa^s*ʌ)

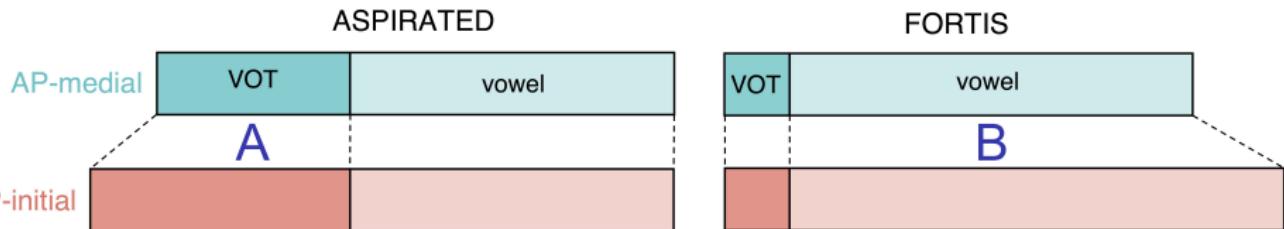


(kukʌ) (ulika) ?iwa^s*ʌ

Manipulated presence/absence of AP-final H on pre-target syllable

- pitch resynthesized (via PSOLA)
- all other aspects of carrier identical

Predictions



Prediction A

- Expectation of AP-initial lengthening of VOT
- AP-initial VOT perceived as relatively short (= fortis)
- **increased initial fortis resp.**

Prediction B

- Expectation of AP-initial lengthening of vowel
- AP-initial vowel perceived as relatively short (= asp.)
- **decreased initial fortis resp.**

Method

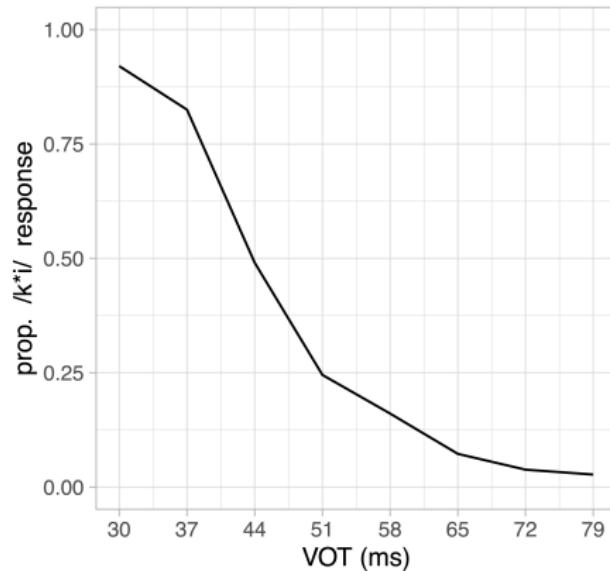
- 30 native speakers of Seoul Korean recruited
- remote participation online
- target categorized as ॥ (/k*i/) or ㅋ (/k^hi/) by key press
- 224 randomized trials (8 VOT steps × 2 AP contexts × 16 reps)

Statistical analysis: log-link Bayesian mixed-effects regression¹

- $\text{response} \sim \text{VOT} * \text{context} + (1 + \text{VOT} * \text{context} \mid \text{participant})$
- uninformative priors

¹Bürkner et al. 2017

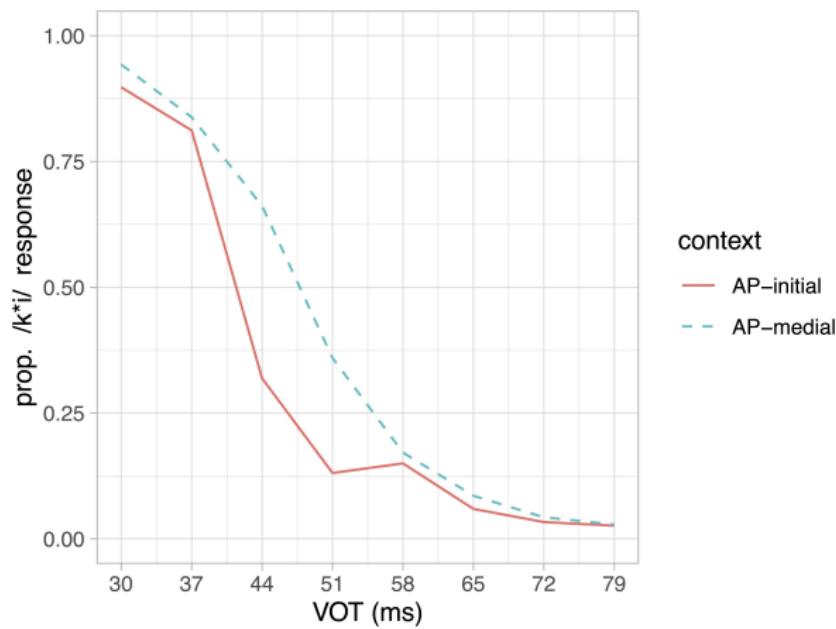
Results



Expected influence of VOT on categorization¹

¹ $\beta = -3.07$, 95%CI=[-3.59,-2.60]

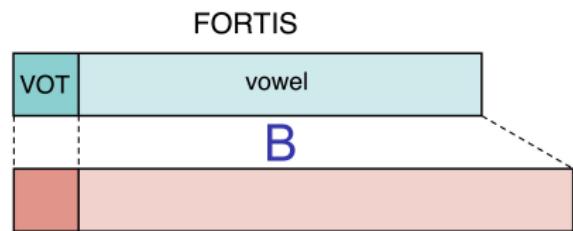
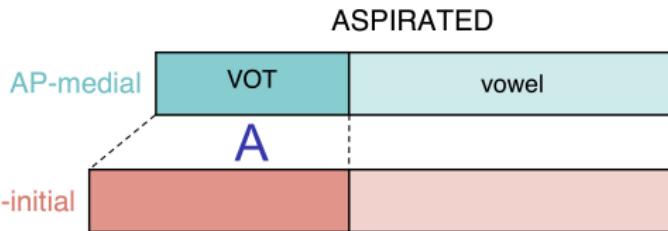
Results



Credible influence of context: decreased fortis responses AP-initially¹

¹ $\beta = -1.13$, 95%CI=[-1.53,-0.79]

Predictions



XPrediction A

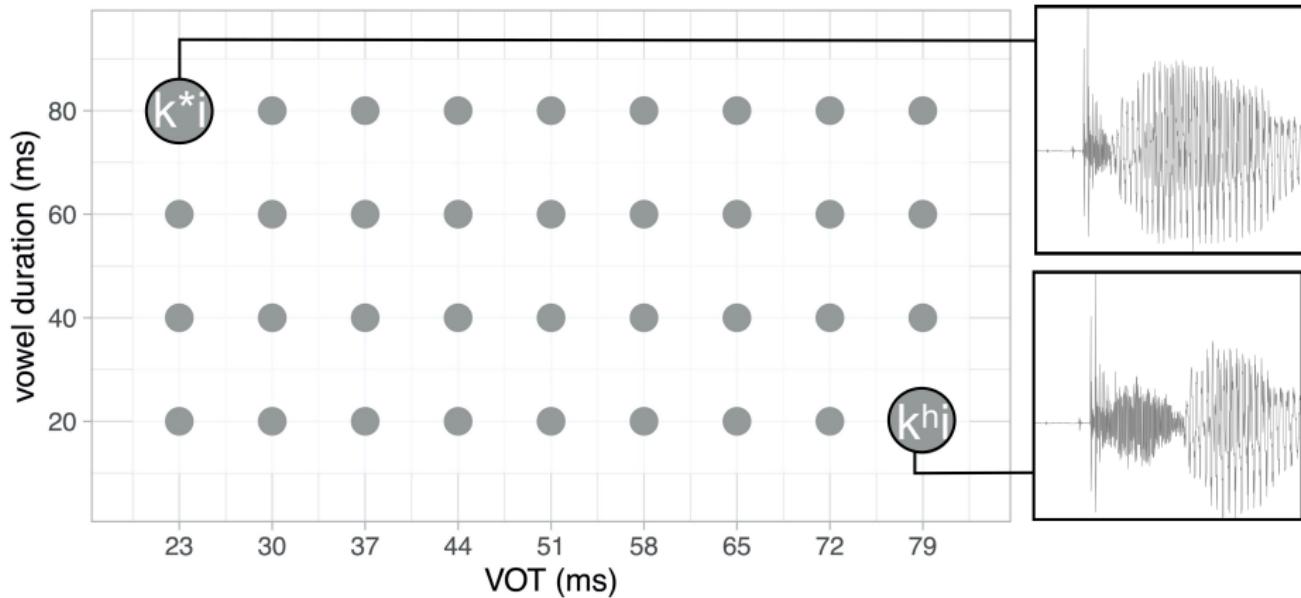
- Expectation of AP-initial lengthening of VOT
- AP-initial VOT perceived as relatively short (= fortis)
- **increased initial fortis resp.**

✓Prediction B

- Expectation of AP-initial lengthening of vowel
- AP-initial vowel perceived as relatively short (= asp.)
- **decreased initial fortis resp.**

Experiment 2: vowel duration and VOT

Same context manipulation - added orthogonal variation in vowel duration:



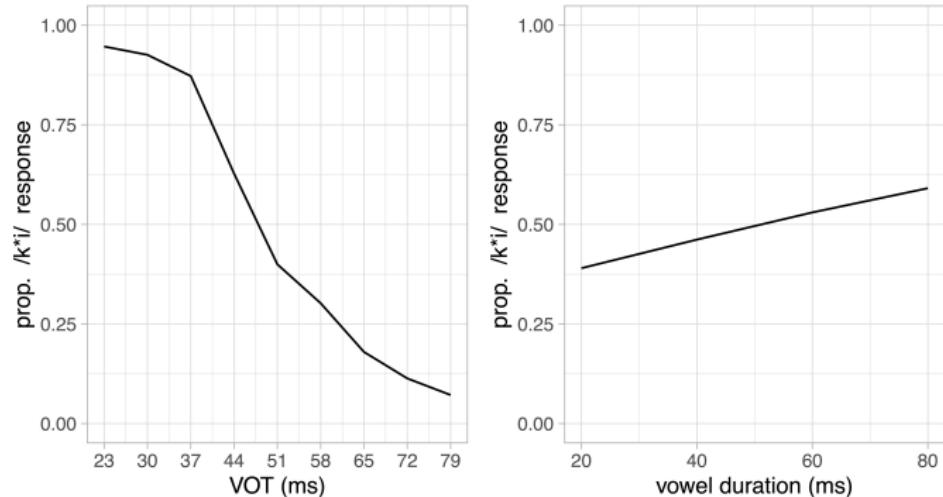
Methods

- 32 participants
- 216 randomized trials
(9 VOT steps \times 4 vowel dur. steps \times 2 AP contexts \times 3 reps)

Statistical analysis: same as Experiment 1

- now including vowel duration

Results: continuum



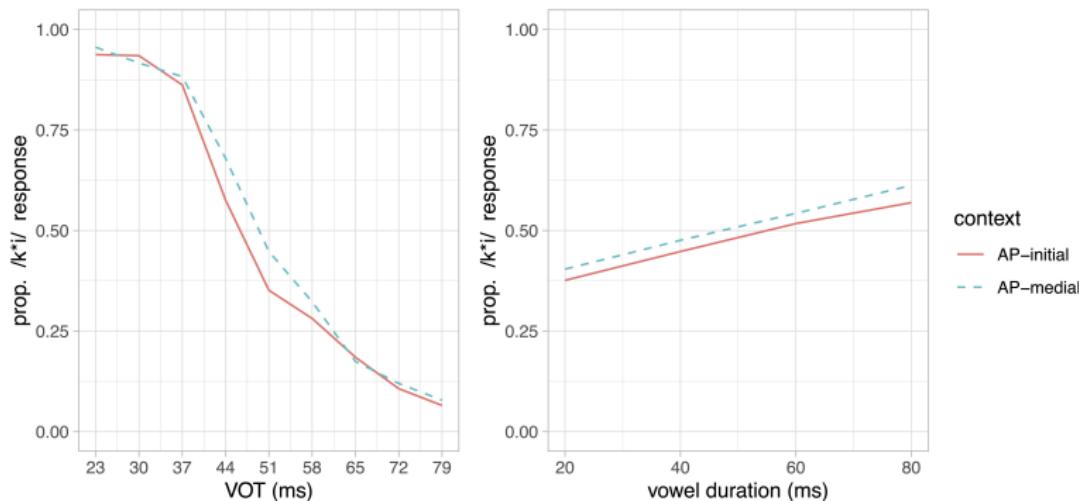
Credible influence of both cues¹

- vowel duration used to distinguish the contrast, though to a lesser extent than VOT

¹VOT: $\beta = -2.65$, 95%CI=[-3.05,-2.26]

vowel duration: $\beta = 0.71$, 95%CI=[0.58,0.84]

Results: context



Credible effect of context: decreased fortis responses in initial position¹

- consistent with Prediction B: expectation of AP-initial vowel lengthening

¹ $\beta = -0.30$, 95%CI=[-0.52,-0.09]

Summary of findings

Pitch-based AP phrasing shift listeners perception of a contrast signaled by temporal cues

- with an apparent restriction in our case to the patterning of vowel duration

Take home message

Prosodic phrasing mediates perception of temporal cues independent of temporal context

¹Brown et al. 2015; Cho et al. 2007

Some further directions

- Explore how these effects interact with other processing related to the AP, e.g. phonological inferencing and word segmentation¹
- Test how these effects are reflected in online processing²
- Test other prosodically-modulated temporal contrasts in Korean and elsewhere³

¹Kim and Cho 2009; Kim et al. 2018

²cf. Mitterer et al. 2019

³cf. Steffman and Katsuda 2020

Thank you!

We thank Hyunjung Joo, Jungyun Seo, and Seungwoo Baek for their help with data acquisition, and Jinyoung Jo for help with recording example materials and piloting.

This work was supported by the Ministry of Education of the Republic of Korea and the National Research Foundation of Korea (NRF-2018S1A5A2A03036736).

References |

- Brown, M., Salverda, A. P., Dilley, L., and Tanenhaus, M. K. (2015). Metrical expectations from preceding prosody influence perception of lexical stress. *Journal of Experimental Psychology: Human Perception and Performance*, 41(2):306.
- Bürkner, P.-C. et al. (2017). brms: An R package for Bayesian multilevel models using Stan. *Journal of statistical software*, 80(1):1–28.
- Cho, T. (2015). Language Effects on Timing at the Segmental and Suprasegmental Levels. In Redford, M. A., editor, *The Handbook of Speech Production*, pages 505–529. John Wiley & Sons, Inc.
- Cho, T. and Jun, S.-A. (2000). Domain-initial strengthening as enhancement of laryngeal features: Aerodynamic evidence from Korean. *UCLA working papers in phonetics*, pages 57–70.
- Cho, T. and Keating, P. (2001). Articulatory and acoustic studies on domain-initial strengthening in Korean. *Journal of phonetics*, 29(2):155–190.
- Cho, T. and Keating, P. (2009). Effects of initial position versus prominence in english. *Journal of Phonetics*, 37(4):466–485.
- Cho, T., McQueen, J. M., and Cox, E. A. (2007). Prosodically driven phonetic detail in speech processing: The case of domain-initial strengthening in English. *Journal of Phonetics*, 35(2):210–243.
- Choi, H. (2011). Vowel duration as a perceptual cue for preceding stop laryngeal contrast in korean. In *ICPhS*, volume 17, pages 17–21.
- Cole, J., Kim, H., Choi, H., and Hasegawa-Johnson, M. (2007). Prosodic effects on acoustic cues to stop voicing and place of articulation: Evidence from radio news speech. *Journal of Phonetics*, 35(2):180–209.
- Jun, S.-A. (1996). *The phonetics and phonology of Korean prosody: Intonational phonology and prosodic structure*. Taylor & Francis.
- Jun, S.-A. (1998). The accentual phrase in the Korean prosodic hierarchy. *Phonology*, pages 189–226.
- Jun, S.-A. (2000). K-tobi (korean tobi) labelling conventions. *Speech Sciences*, 7(1):143–170.
- Jun, S.-A. and Cha, J. (2015). High-toned [il] in korean: Phonetics, intonational phonology, and sound change. *Journal of Phonetics*, 51:93–108.
- Keating, P. (2006). Phonetic encoding of prosodic structure. *Speech production: Models, phonetic processes, and techniques*, pages 167–186.
- Keating, P., Cho, T., Fougeron, C., and Hsu, C.-S. (2004). Domain-initial articulatory strengthening in four languages. *Phonetic interpretation: Papers in laboratory phonology VI*, pages 143–161.
- Kim, S. and Cho, T. (2009). The use of phrase-level prosodic information in lexical segmentation: Evidence from word-spotting experiments in Korean. *The Journal of the Acoustical Society of America*, 125(5):3373–3386.

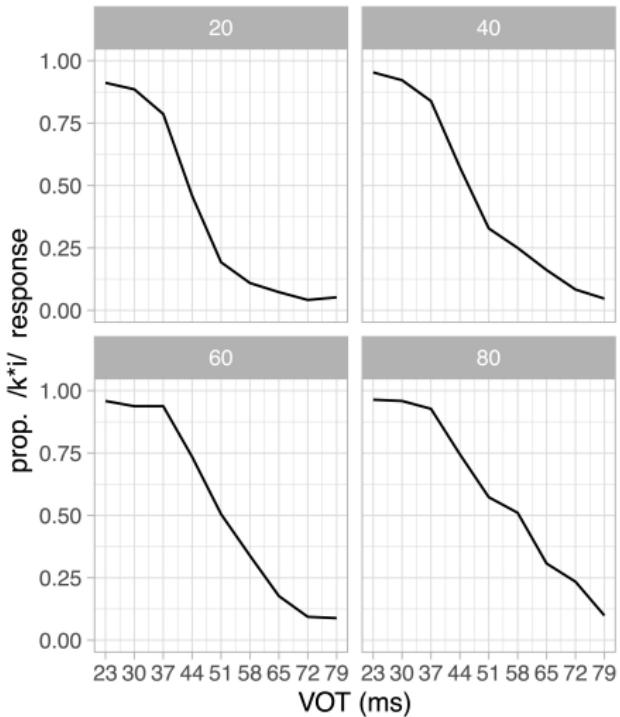
References II

- Kim, S. and Cho, T. (2013). Prosodic boundary information modulates phonetic categorization. *The Journal of the Acoustical Society of America*, 134(1):EL19–EL25.
- Kim, S., Cho, T., and McQueen, J. M. (2012). Phonetic richness can outweigh prosodically-driven phonological knowledge when learning words in an artificial language. *Journal of Phonetics*, 40(3):443–452.
- Kim, S., Mitterer, H., and Cho, T. (2018). A time course of prosodic modulation in phonological inferencing: The case of Korean post-obstruent tensing. *PloS one*, 13(8).
- Mitterer, H., Cho, T., and Kim, S. (2016). How does prosody influence speech categorization? *Journal of Phonetics*, 54:68–79.
- Mitterer, H., Kim, S., and Cho, T. (2019). The glottal stop between segmental and suprasegmental processing: The case of maltese. *Journal of Memory and Language*, 108:104034.
- Nakai, S., Kunnari, S., Turk, A., Suomi, K., and Ylitalo, R. (2009). Utterance-final lengthening and quantity in northern finnish. *Journal of phonetics*, 37(1):29–45.
- Newman, R. S. and Sawusch, J. R. (2009). Perceptual normalization for speaking rate iii: Effects of the rate of one voice on perception of another. *Journal of phonetics*, 37(1):46–65.
- Shepherd, M. A. (2008). The scope and effects of preboundary prosodic lengthening in japanese. *USC Working Papers in Linguistics*, 4:1–14.
- Steffman, J. (2019). Intonational structure mediates speech rate normalization in the perception of segmental categories. *Journal of Phonetics*, 74:114–129.
- Steffman, J. and Katsuda, H. (2020). Intonational structure influences perception of contrastive vowel length: The case of phrase-final lengthening in tokyo japanese. *Language and Speech*, page 0023830920971842.
- Wade, T. and Holt, L. L. (2005). Perceptual effects of preceding nonspeech rate on temporal properties of speech categories. *Perception & psychophysics*, 67(6):939–950.

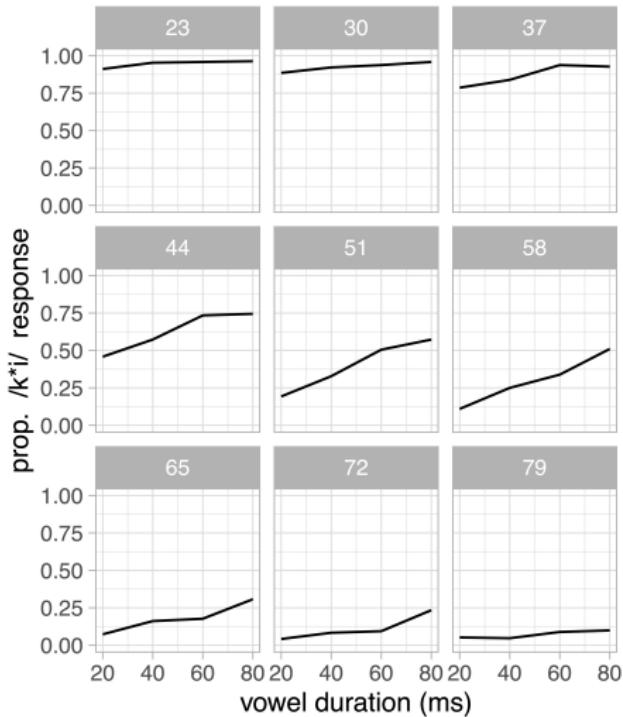
Appendix slides

Results: continuum 2

VOT cat. at vowel dur. steps

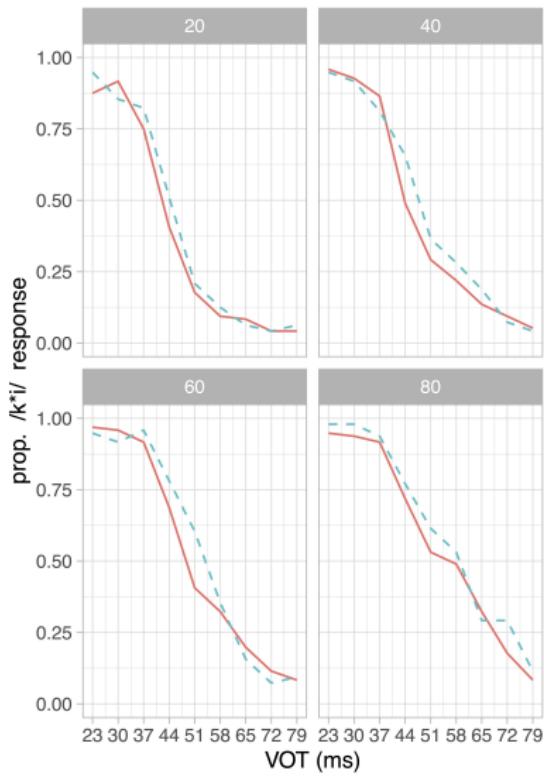


vowel dur. cat. at VOT steps

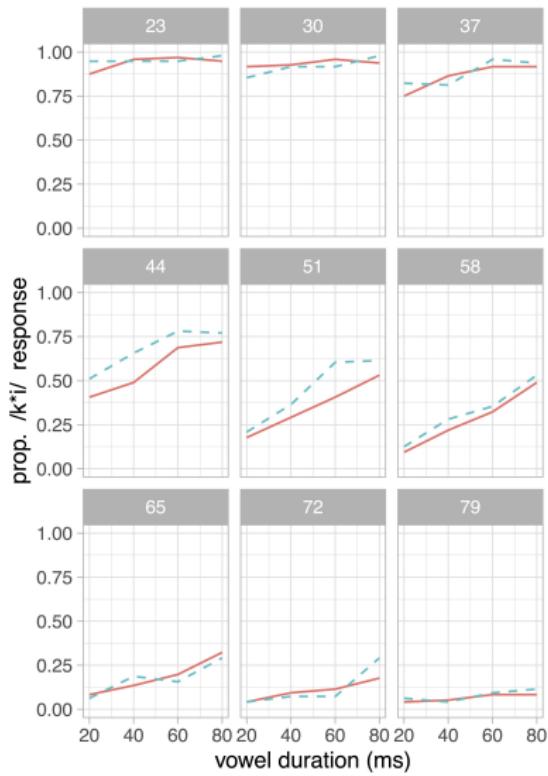


Exp 2 context effects split by continuum

VOT cat. at vowel dur. steps

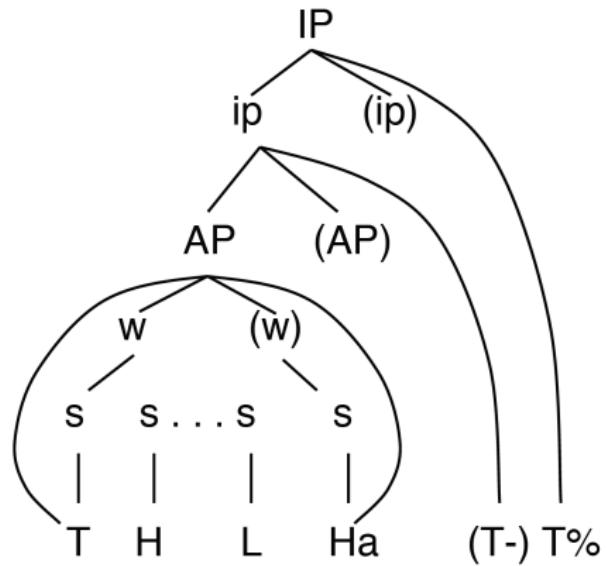


vowel dur. cat. at VOT steps



context
— AP-initial
- - - AP-medial

Korean prosody and phrasing¹



¹From Jun and Cha (2015)

Speculation RE the importance of vowel duration and context

Why the apparent prevalence of phrasing effects for vowel duration only?

- post-aspirated vowels also lengthen slightly AP-initially, though to a much smaller extent than post-fortis vowels¹
- in contrast fortis stop VOT *shortens* slightly AP-initially, though to a much smaller extent than aspirated stop VOT lengthens²
 - ▶ when a stop is ambiguous in terms of VOT, should VOT lengthen or shorten as a function of prosody? listener may not know what to expect
 - ▶ cf. vowel duration, which at the very least does not show a conflicting pattern of modulation

¹Cho and Keating 2001

²Cho and Jun 2000

Exp. 1 model output

	Estimate	Est. Error	L-95% CI	U-95%CI	cred.?
intercept	-1.79	0.22	-2.23	-1.35	✓
context	-1.13	0.19	-1.53	-0.79	✓
VOT	-3.07	0.25	-3.59	-2.60	✓
VOT:context	-0.12	0.20	-0.54	0.24	

Exp. 2 model output

	Estimate	Est. Error	L-95% CI	U-95%CI	cred.?
intercept	-0.14	0.14	-0.42	0.14	
VOT	-2.65	0.20	-3.05	-2.26	✓
vdur	0.71	0.07	0.58	0.84	✓
context	-0.30	0.11	-0.52	-0.09	✓
VOT:vdur	0.24	0.08	0.09	0.40	✓
VOT:context	-0.01	0.11	-0.24	0.21	
vdur:context	-0.01	0.08	-0.17	0.15	
VOT:vdur:context	0.04	0.11	-0.17	0.28	