Neurophysiological encoding of VOT and F0 in voicing perception

Chao Han[1], Katie Ma[1], Yoonjung Kang[1,2], Philip J. Monahan[1,2,3]

1. Department of Language Studies, University of Toronto Scarborough
2. Department of Linguistics, University of Toronto
3. Department of Psychology, University of Toronto Scarborough

CAP LAB COMPUTATION AND PSYCHOLINGUISTICS LABORATORY

UNIVERSITY OF TORONTO

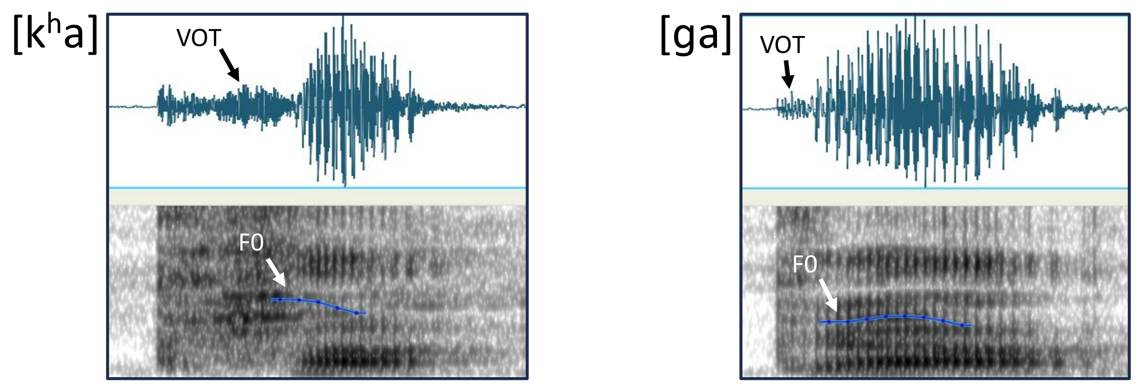
QR code

PREVIEW

- Both VOT and F0 support English consonant voicing categorization.
- Individuals differ in their VOT and F0 cue weights.
- MMN does not track individual VOT and F0 cue weights.

BACKGROUND

- English stop voicing contrasts primarily cued by voice onset time (VOT), secondarily by fundamental frequency (F0) [1].



- Individual variation in cue weighting [2].
- Larger mismatch negativity (MMN) response linked to more robust phonological contrast encoding [3].

Research Questions:

1. Do secondary cues pre-attentively engage phonological encoding?
2. Do MMN responses reflect participant-level variation in cue-weighting?

METHODS

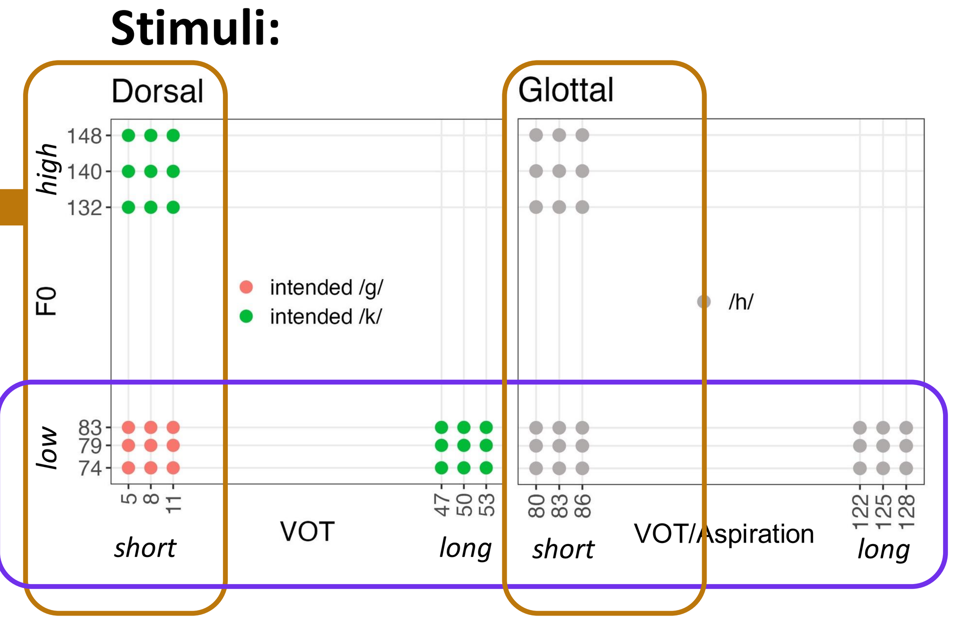
Participants

- English speakers

F0 group (n = 18)

VOT group (n = 20)

Stimuli:

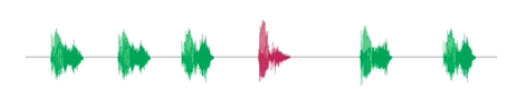


Procedure

Task 1: behavioral
Phoneme identification task

Choose the sound you hear

Task 2: EEG
Many-to-one oddball paradigm



2 groups × 2 places of articulation × 2 deviants

	F0 group (varying F0, fixed VOT)	**VOT group** (varying F0, fixed F0)
Dorsal /ka/ vs. /ga/	high–high–high–low–high	long–long–long–short–long
	low–low–low–high–low	short–short–short–long–short
Glottal /ha/	high–high–high–low–high	long–long–long–short–long
	low–low–low–high–low	short–short–short–long–short

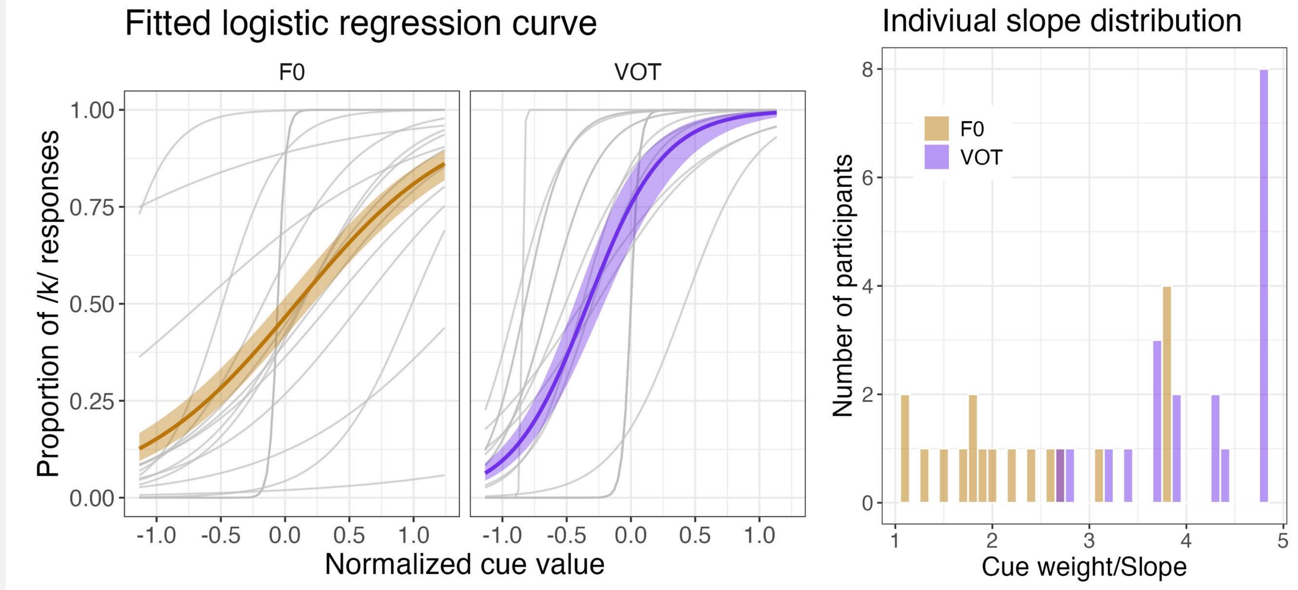
RESULTS: Behavioral cue weights

Mixed-effects logistic regression for each group.

- Fixed effect: normalized cue value
- Random effects: Random intercepts and random by-participant slopes for normalized cue value

Regression coefficients extracted for each participants to index individual cue weights.

- VOT group: β = 4.3, $p < .001$
- F0 group: β = 2.6, $p < .001$
- A significant difference between VOT and F0 slopes: $t = 6.3$, $p < .001$

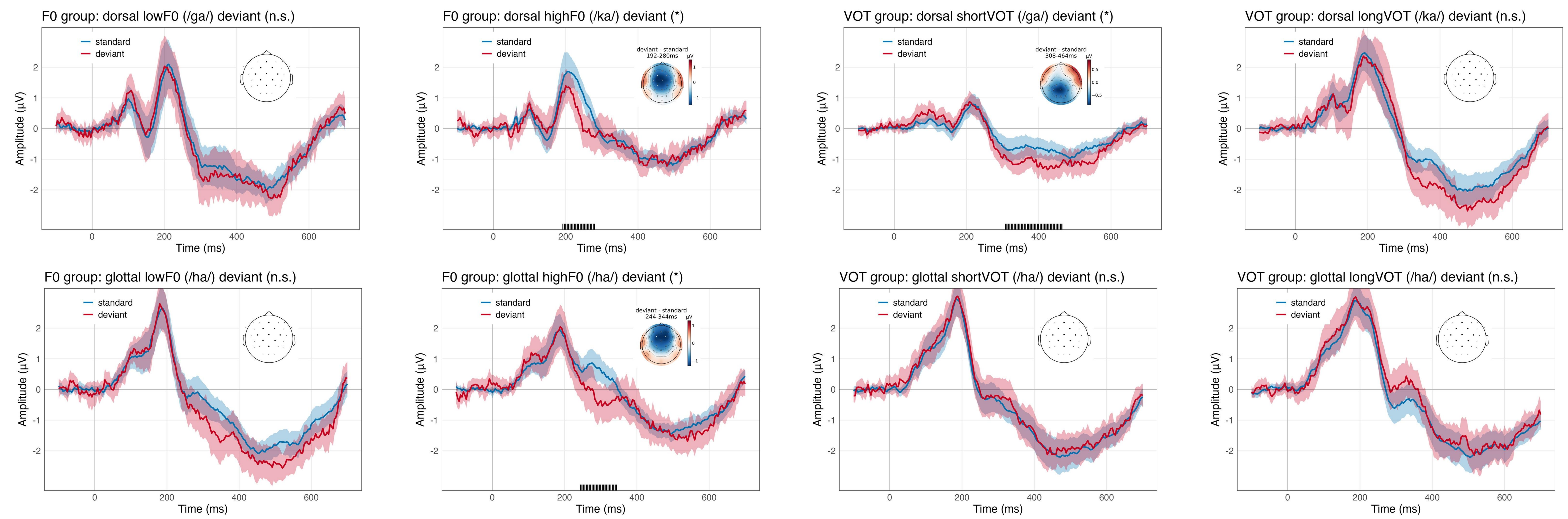


RESULTS: MMN

- Identity MMN: subtract ERPs to same stimulus, standard presentation minus deviant presentation across blocks.
- MMN amplitude: assessed by spatiotemporal cluster-based permutation tests (cluster formation threshold p = 0.05, two-tailed; cluster-level threshold: p = 0.05).

(* = significant cluster found; n.s. = no significant cluster)

	F0 group	VOT group
Dorsal /ka/ vs. /ga/	lowF0 (n.s.)	shortVOT (*)
	highF0 (*)	longVOT (n.s.)
Glottal /ha/	lowF0 (n.s.)	shortVOT (n.s.)
	highF0 (*)	longVOT (n.s.)



RESULTS: Brain-behavior relations

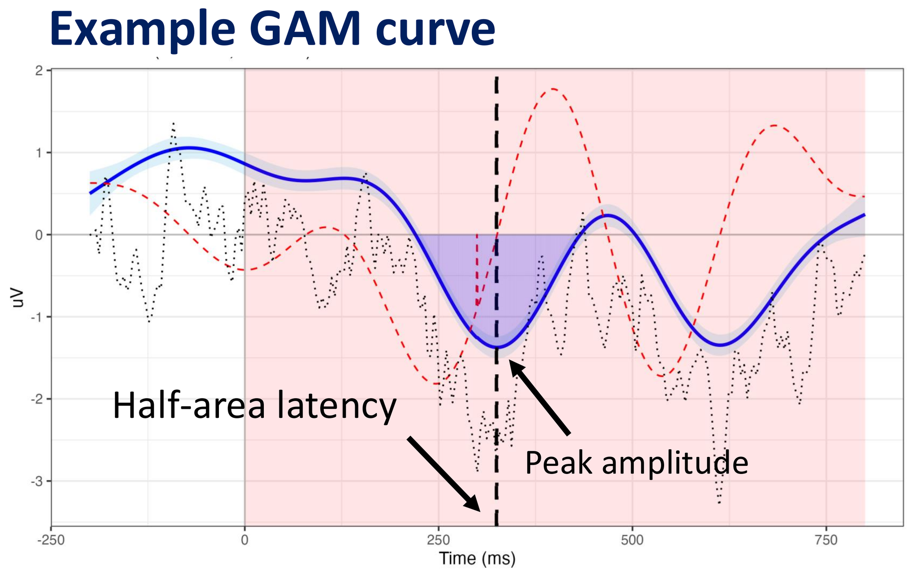
Individual brain responses for each deviant condition, modeled using generalized additive modeling (GAM) [4]. Two measures extracted:

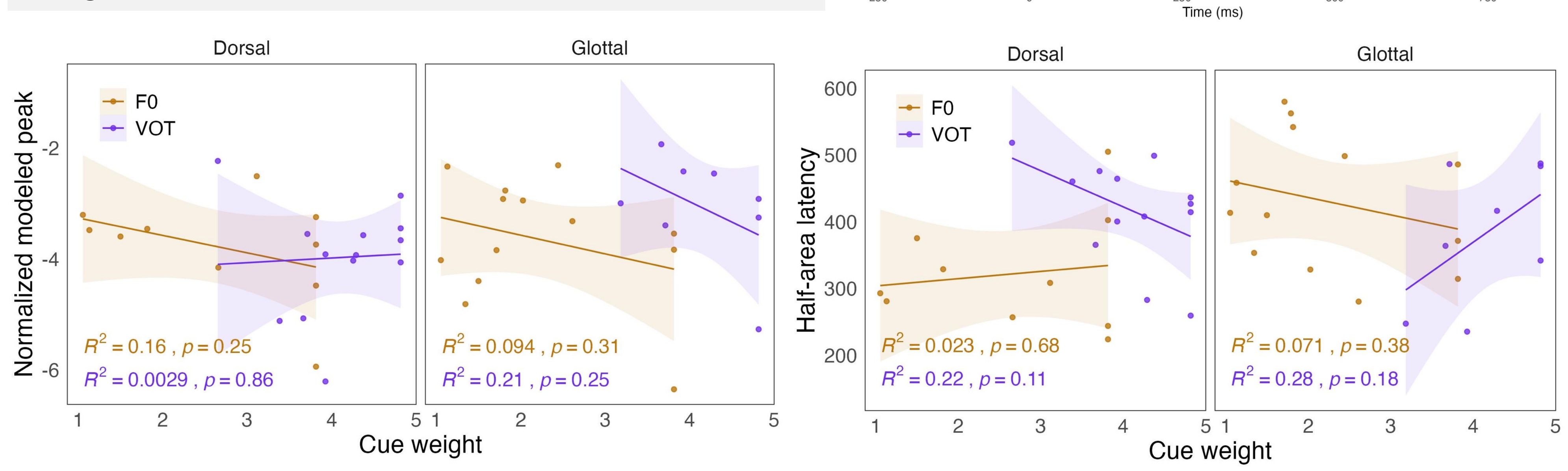
- Normalized modeled peak: MMN peak amplitude normalized by standard error: peak / (1.96 * SE)
- Half-area latency: Time point at which 50% of the area within the negative deflection is reached.

Correlate individual responses with individual cue weights.

No significant correlations found across conditions.

Example GAM curve





DISCUSSION & REFERENCES

- Behaviorally, F0 supports English voicing contrasts.
- VOT MMN found for dorsal shortVOT but not for longVOT
  - Consistent with phonological accounts [5–7].
- F0 MMN comparable for dorsal and glottal conditions:
  - No evidence that F0 is engaged pre-attentively in phonological contrast encoding.
- No significant brain-behavior correlations.
  - Neural response encoding individual cue weights may degrade by MMN (150–300 ms).
  - Consistent with degradation of fine-phonetic details along the subcortical-cortical auditory pathway [8-10].

1. Kapnoula, E. C., Winn, M. B., Kong, E. J., Edwards, J. & McMurray, B. Evaluating the sources and functions of gradiency in phoneme categorization: An individual differences approach. *J. Exp. Psychol. Hum. Percept. Perform.* **43**, 1594–1611 (2017).
2. Kong, E. J. & Edwards, J. Individual differences in categorical perception of speech: Cue weighting and executive function. *J. Phon.* **59**, 40–57 (2016).
3. Díaz, B., Baus, C., Escera, C., Costa, A. & Sebastián-Gallés, N. Brain potentials to native phoneme discrimination reveal the origin of individual differences in learning the sounds of a second language. *Proc. Natl. Acad. Sci.* **105**, 16083–16088 (2008).
4. Meulman, N., Sprenger, S. A., Schmid, M. S. & Wieling, M. GAM-based individual difference measures for L2 ERP studies. *Res. Methods Appl. Linguist.* **2**, 100079 (2023).
5. Lahiri, A. & Reetz, H. Underspecified recognition. in *Laboratory Phonology 7* (eds Gussenhoven, C. & Warner, N.) 637–676 (Mouton de Gruyter, 2002). doi:10.1515/9783110197105.2.637.
6. Hestvik, A. & Durvasula, K. Neurobiological evidence for voicing underspecification in English. *Brain Lang.* **152**, 28–43 (2016).
7. Fu, Z. & Monahan, P. J. Extracting Phonetic Features From Natural Classes: A Mismatch Negativity Study of Mandarin Chinese Retroflex Consonants. *Front. Hum. Neurosci.* **15**, 1–15 (2021).
8. Tamura, S. & Sung, Y. Brainstem and early cortical auditory activities associated with language differences in acoustic cue weighting for voicing perception. *Neurosci. Lett.* **735**, 135154 (2020).
9. Toscano, J. C., McMurray, B., Dennhardt, J. & Luck, S. J. Continuous perception and graded categorization: Electrophysiological evidence for a linear relationship between the acoustic signal and perceptual encoding of speech. *Psychol. Sci.* **21**, 1532–1540 (2010).
10. Ou, J., Xiang, M. & Yu, A. C. L. Individual variability in subcortical neural encoding shapes phonetic cue weighting. *Sci. Rep.* **13**, 9991 (2023).

17th *Annual Meeting Society for the Neurobiology of Language*, September 12–14, 2025. Gallaudet University, Washington DC

We thank Jack Mahlmann, Carolyn Kortje, Shane de Dios, and Research Assistants in the CAP Lab for their support. This work was funded by SSHRC and NSERC of Canada.