Integrating prosodic context in speech perception

Jeremy Steffman

Northwestern University
jeremy.steffman@northwestern.edu

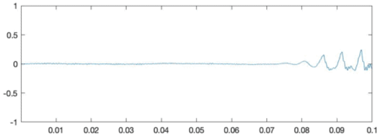
March 29, 2021

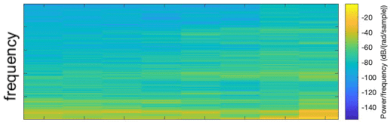
Introduction

Two parts of understanding spoken language

1. perceiving segmental contrasts → lexical items
2. perceiving prosodic features: grouping, prominence relations, etc.

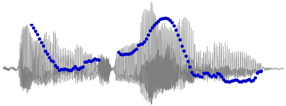
"her name is Marilyn"





(hɚneɪm ɪz meɹɪlɪn)$_{\text{intonational phrase}}$

L+H* L-L%



segmental processing

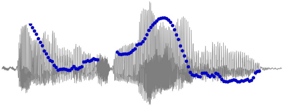
hɚneɪm ɪz meɹɪlɪn

prosodic processing

$(..........)_{IP}$

L+H*L-L%

Prosody and segment usually studied as independent in processing



↙		↘
segmental processing	⟺ ?	prosodic processing
↓		↓
hɚneɪm ɪz meɹɪlɪn		(...........)$_{IP}$ L+H*L-L%

We lack a good understanding of how they interact[1]

[1] Mitterer et al., 2016; Steffman, 2019

This talk

is about one type of interaction: the influence of prosodic context on the perception of segmental contrasts

Why should prosodic context matter?

Acoustic cues vary based on prosody and segment (cf. the notion of "suprasegmentals")

[1] Cho, 2015; Turk and Shattuck-Hufnagel, 2007
[2] Chen, 1970
[3] Lehiste, 1970
[4] Shepherd, 2008; Nakai et al., 2009
[5] Keating et al., 2004

Why should prosodic context matter?

Acoustic cues vary based on prosody and segment (cf. the notion of "suprasegmentals")

Vowel duration is longer...

- in prominent syllables[1]
- before prosodic boundaries[1]

---

[1] Cho, 2015; Turk and Shattuck-Hufnagel, 2007
[2] Chen, 1970
[3] Lehiste, 1970
[4] Shepherd, 2008; Nakai et al., 2009
[5] Keating et al., 2004

Why should prosodic context matter?

Acoustic cues vary based on prosody and segment (cf. the notion of "suprasegmentals")

Vowel duration is longer...

- in prominent syllables[1]
- before prosodic boundaries[1]
- before voiced obstruents[2]
- based on vowel height[3]
- in long vowels in languages where vowel length is contrastive[4]

[1] Cho, 2015; Turk and Shattuck-Hufnagel, 2007
[2] Chen, 1970
[3] Lehiste, 1970
[4] Shepherd, 2008; Nakai et al., 2009
[5] Keating et al., 2004

Why should prosodic context matter?

Acoustic cues vary based on prosody and segment (cf. the notion of "suprasegmentals")

Vowel duration is longer...

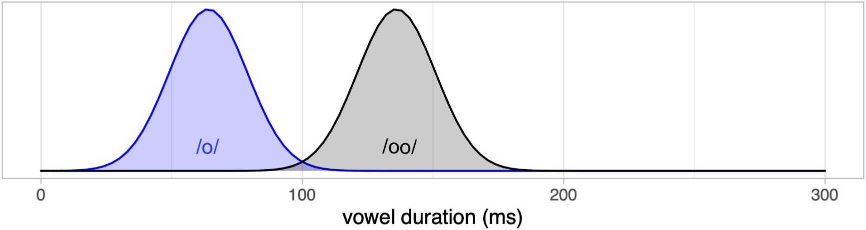
- in prominent syllables[1]
- before prosodic boundaries[1]
- before voiced obstruents[2]
- based on vowel height[3]
- in long vowels in languages where vowel length is contrastive[4]

This is the case for many cues, another well known case: VOT[5]

[1] Cho, 2015; Turk and Shattuck-Hufnagel, 2007
[2] Chen, 1970
[3] Lehiste, 1970
[4] Shepherd, 2008; Nakai et al., 2009
[5] Keating et al., 2004

Why should prosodic context matter?

Duration is contrastive in Japanese[1]

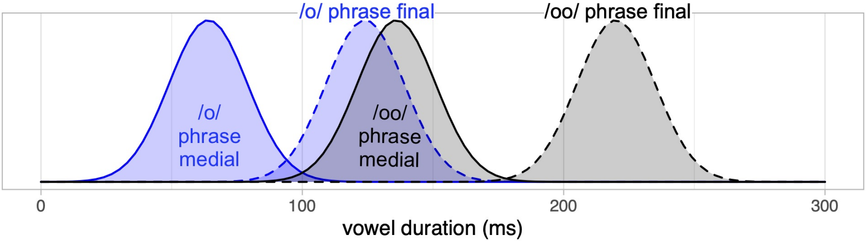


e.g. /dookjo/ "housemate" - /dookjoo/ "townmate"

[1]From Shepherd, 2008

Why should prosodic context matter?

And varies based on phrasing: longer phrase-finally[1]



Dual patterning based on both prosodic and segmental properties

[1] From Shepherd, 2008

Road map

Hypothesis: Determining segmental contrasts in speech involves **interaction** between segmental/prosodic processing

Question: What does does this interaction look like?

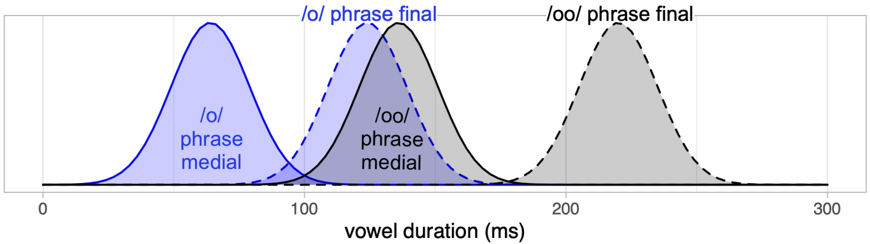
Section I
Do listeners care about prosodic boundaries in their perception of contrastive temporal cues?

Section II
Do listeners care about prosodic *prominence* their perception of spectral cues? (vowel formants)
And how is this information processed online?

I. Prosodic boundaries and contrastive vowel length in Japanese

Experiment 1: Contrastive vowel duration in Japanese

Recall this[1]



Do listeners care?

[1]From Shepherd, 2008

Experiment 1: Contrastive vowel duration in Japanese[1]

Will phrasing mediate listeners' expectation of how long a vowel should be?

Method: forced choice categorization of vowel duration continuum:

- Experiment 1a ($n = 26$): ʃi'ʃoo "master" - ʃi'ʃo "librarian"
- Experiment 1b ($n = 26$): dookjoo "housemate" - dookjo "townmate"

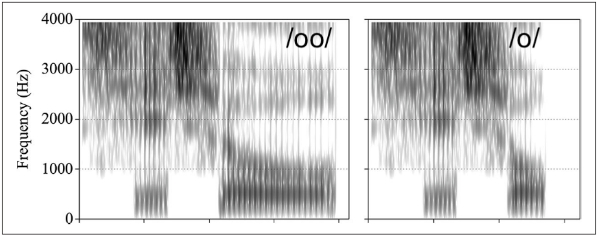
[1] Joint work with Hironori Katsuda (Steffman and Katsuda, 2020)

Experiment 1: Contrastive vowel duration in Japanese[1]

Will phrasing mediate listeners' expectation of how long a vowel should be?

Method: forced choice categorization of vowel duration continuum:

- Experiment 1a (n= 26): ʃi'ʃoo "master" - ʃi'ʃo "librarian"
- Experiment 1b (n= 26): dookjoo "housemate" - dookjo "townmate"



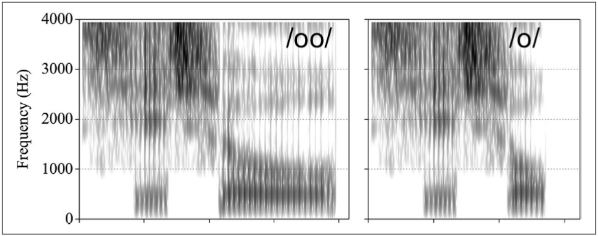
[1] Joint work with Hironori Katsuda (Steffman and Katsuda, 2020)

Experiment 1: Contrastive vowel duration in Japanese[1]

Will phrasing mediate listeners' expectation of how long a vowel should be?

Method: forced choice categorization of vowel duration continuum:

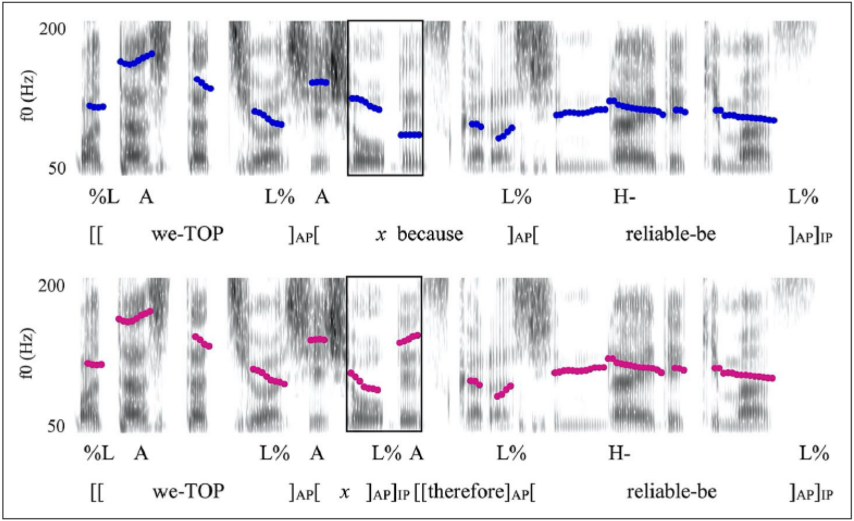
- Experiment 1a (n= 26): ʃi'ʃoo "master" - ʃi'ʃo "librarian"
- Experiment 1b (n= 26): dookjoo "housemate" - dookjo "townmate"



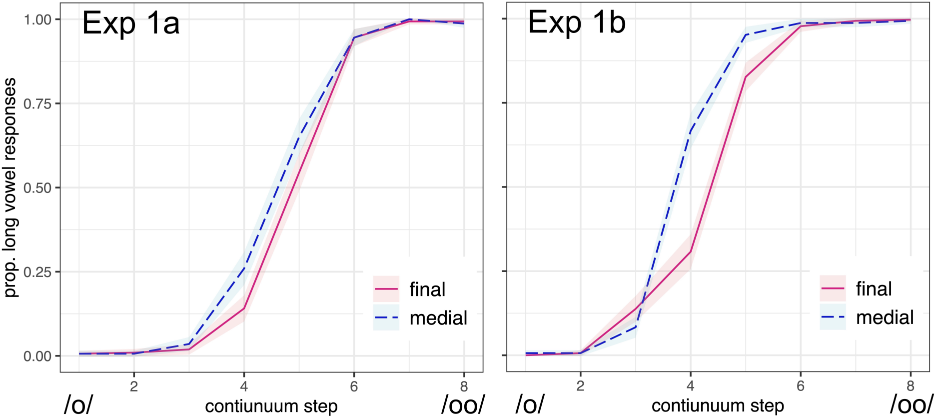
- cued as phrase final or phrase medial by intonational tunes in a phrase
  we are [TARGET] therefore we are reliable
- only pitch varies across phrase final/ phrase medial conditions

[1] Joint work with Hironori Katsuda (Steffman and Katsuda, 2020)

Experiment 1a stimuli



Experiment 1: Results[1]



[1] Exp 1a phrasing β= -0.91; 95%CI = [-1.54,-0.31]
Exp 1b phrasing β= -1.54; 95%CI = [-2.14,-1.06]

What we've seen so far...

Listeners modulate their perception of temporal contrasts as a function of prosodic phrasing

Up next...

- what about non-temporal cues?
- what about prosodic prominence?

II. Contextual prominence in vowel perception

The test case: Prominence strengthening in American English vowels

Vowel articulations/acoustics are modulated by prominence[1]
(= *prominence strengthening*)

Two effects:

1. Sonority expansion: expansion of the oral cavity
   - where "sonorous" = more open[2]
   - makes a louder vowel

[1] e.g., Beckman et al., 1992; Cho, 2005; de Jong et al., 1993; Erickson, 2002
[2] Silverman and Pierrehumbert, 1990
[3] Dilley et al., 1996; Garellek, 2013, 2014

The test case: Prominence strengthening in American English vowels

Vowel articulations/acoustics are modulated by prominence[1]
(= *prominence strengthening*)

Two effects:

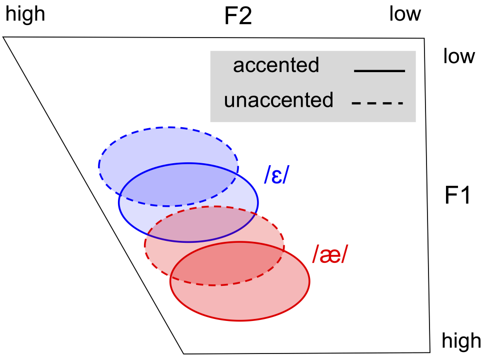
1. Sonority expansion: expansion of the oral cavity
   - where "sonorous" = more open[2]
   - makes a louder vowel
2. Voice quality strengthening: glottalization
   - glottalization and production of [ʔ] precede prominent vowel-initial words [3]
   - though not always

[1] e.g., Beckman et al., 1992; Cho, 2005; de Jong et al., 1993; Erickson, 2002
[2] Silverman and Pierrehumbert, 1990
[3] Dilley et al., 1996; Garellek, 2013, 2014

Sonority expansion effects on vowel formants

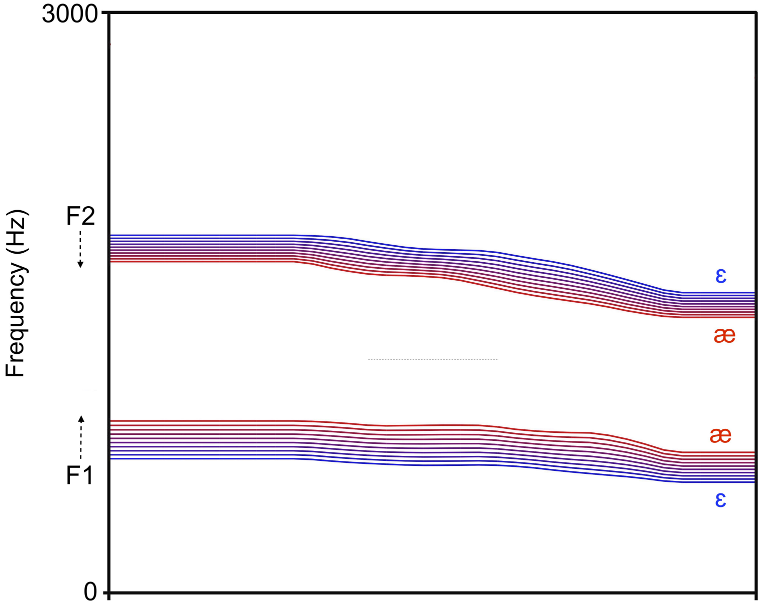
Sonority expansion entails:

- increased opening, jaw lowering: raised F1[1]
- for some vowels, retraction the tongue: lowered F2[2]



[1] Erickson, 2002; Van Summers, 1987

[2] Cho, 2005; Erickson, 2002


3000
F2
Frequency (Hz)
ɛ
æ
æ
F1
ɛ
0

Experiment 2 and 3

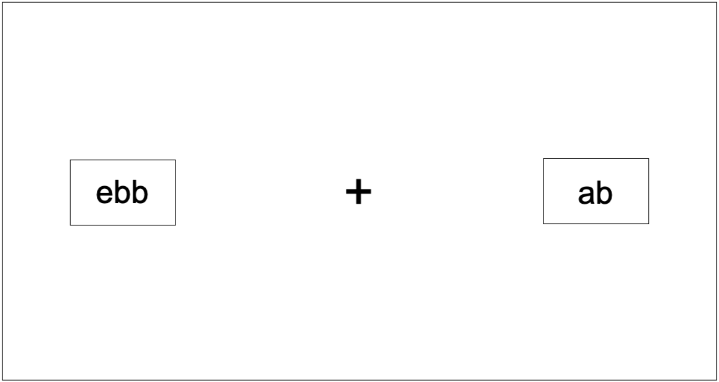
Two visual word eyetracking studies (n = 36 in each)

- 6-step continuum categorized as "ebb" /ɛ/ or "ab" /æ/
  - central region of a piloted 10 step continuum - more ambiguous sounds
- Two contextual conditions: prominent/non-prominent

Experiment 2 and 3

Two visual word eyetracking studies (n = 36 in each)

- 6-step continuum categorized as "ebb" /ɛ/ or "ab" /æ/
  - central region of a piloted 10 step continuum - more ambiguous sounds
- Two contextual conditions: prominent/non-prominent



Experiment 2 conditions: "Phrasal prominence"

Two prominence conditions cuing a contrast in accentuation

1. I'll say *X* now *nuclear pitch accent - NPA* (prominent)
   H* H* L-L%
2. I'll SAY *x* now *post-focus* (non-prominent)
   L+H* L-L%

Experiment 2 conditions: "Phrasal prominence"

Two prominence conditions cuing a contrast in accentuation

1. I'll say *X* now — *nuclear pitch accent - NPA* (prominent)
   H* H* L-L%
2. I'll SAY *x* now — *post-focus* (non-prominent)
   L+H* L-L%

the target is acoustically the same across conditions - prominence cued by context

Experiment 2 conditions: "Phrasal prominence"

Two prominence conditions cuing a contrast in accentuation

1. I'll say *X* now — *nuclear pitch accent - NPA* (prominent)
   H* H* L-L%
2. I'll SAY *x* now — *post-focus* (non-prominent)
   L+H* L-L%

the target is acoustically the same across conditions - prominence cued by context

Prediction

Higher F1/ lower F2 expected in the NPA condition, more tokens mapped to /ɛ/ therein: increased "ebb" responses in NPA.

Experiment 3 conditions: [ʔ] as a cue to prominence

Glottalization often co-occurs with phrasal prominence in vowel-initial words

Prominence manipulation: presence/absence of pre-target [ʔ]

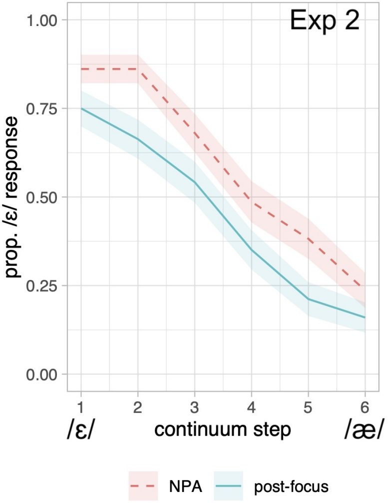
- "say the [target] now"

1. [seɪ ðə ʔ ɛb naʊ] *glottal stop*
   H* H* L-L%
2. [seɪ ðə ɛb naʊ] *no glottal stop*
   H* H* L-L%

Prediction:

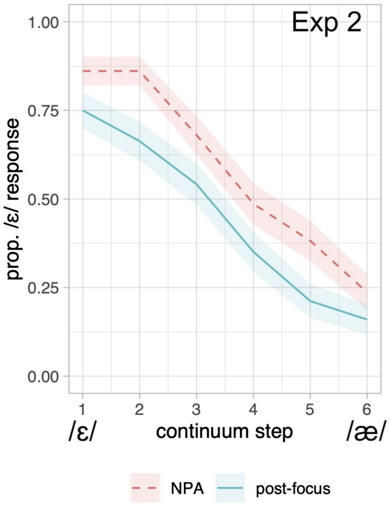
if [ʔ] cues prominence, increased "ebb" responses in the glottal stop condition - same as phrasal prominence

Categorization responses[1]



[1] Exp 2 prominence β= 0.91; 95%CI = [0.22,1.59]
Exp 3 prominence β= 2.66; 95%CI = [2.08,3.28]

Categorization responses[1]



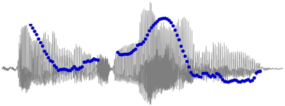
[1] Exp 2 prominence $\beta = 0.91$; 95%CI = [0.22,1.59]
Exp 3 prominence $\beta = 2.66$; 95%CI = [2.08,3.28]

What we've seen so far...

- Prosodic prominence matters in vowel perception
- Both phrasal context (Exp 2) and local context (Exp 3) generate adjustments in categorization

Up next:

- How are listeners integrating these cues as speech unfolds?



↙ ↘

segmental processing $\Longleftrightarrow$??? prosodic processing

Two ways this might work

1. Segments first: 
    - make your best guess about segmental material as soon as you can
    - wait to integrate this information with a prosodic representation, once you know more about utterance prosody

2. All at once: 
    - Why wait? Integrate preceding contextual cues with information about a segment as soon as you can

The segments first model

Prosodic analysis[1]

- simultaneous segmental and prosodic parses of the signal

1. Segmental analysis activates lexical hypotheses
2. Parsed prosody integrated in lexical competition (e.g., alignment of phrasal/word boundaries)

Prediction:

Prosodic context shows a delayed influence in segmental processing

Recent empirical support from prosodic *boundaries*[2]

[1] Cho et al., 2007

[2] Kim et al., 2018a; Mitterer et al., 2019

All at once

On the other hand...

- listeners rapidly integrate temporal and spectral context in perception in general[1]
- and some other contextual prosodic influences occur rapidly[2]

Prediction:

Prosodic context, which precedes segmental material, should show an immediate influence on listeners' interpretation of that material

[1] e.g., Reinisch and Sjerps, 2013
[2] e.g., Brown et al., 2011

"The segmental analysis activates all possible lexical hypotheses, and its activation is further modulated by the prosodic analysis at a relatively late stage in spoken-word recognition." (Mitterer et al., 2019)
(prosodic boundaries)

"listeners integrate [...] cues to prosodic structure in the earliest moments of spoken-word recognition." (Brown et al., 2011)
(rhythmic alternations)

Why prominence is interesting

Prominence in American English can be described in phonological/categorical terms

- nuclear accented > accented > unaccented
- pitch accent categories
  - L+H* > H* > L*

[1] Bishop et al., 2020; Grice et al., 2017

Why prominence is interesting

Prominence in American English can be described in phonological/categorical terms

- nuclear accented > accented > unaccented
- pitch accent categories
  - L+H* > H* > L*

But...

- various phonetic parameters impact prominence perception, including within pitch accent categories[1]
- "a linguistic entity is prosodically prominent when it stands out relative to an entity or a set of entities in its environment." (Terken and Hermes, 2000)

[1] Bishop et al., 2020; Grice et al., 2017

Experiments in this section

In both Experiment 2 and 3:

1. listeners hear variation in formants and contextual prominence
2. prominence cues only *precede* the target

Experiments in this section

In both Experiment 2 and 3:

1. listeners hear variation in formants and contextual prominence
2. prominence cues only *precede* the target

Experiments 2/3 differ in

1. the temporal scope of differences in preceding context
2. the phonological status of prominence differences across conditions?

Experiments in this section

In both Experiment 2 and 3:

1. listeners hear variation in formants and contextual prominence
2. prominence cues only *precede* the target

Experiments 2/3 differ in

1. the temporal scope of differences in preceding context
2. the phonological status of prominence differences across conditions?

Prediction in both: prominent contexts should lead to increased looks to "ebb" - same as categorization

Predictions: timecourse

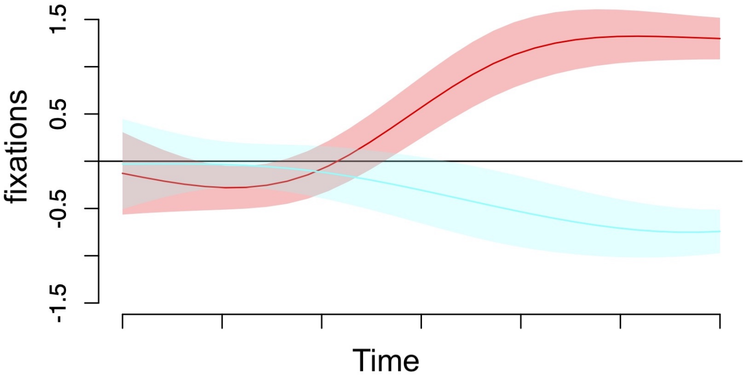
Segments first

- P1 Given (1) formants, and (2) contextual prominence, formants will impact processing *before* prominence asynchronous use of cues
- P2 Perception of formants early in processing *won't vary* across conditions

All at once

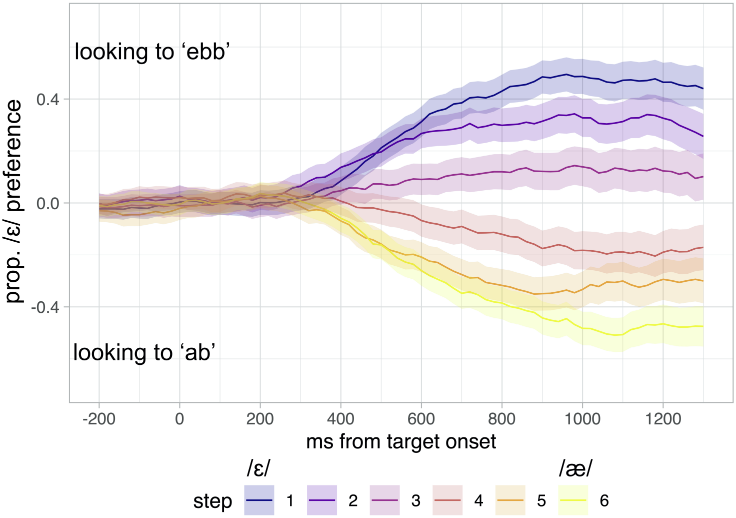
- P1 Given (1) formants, and (2) contextual prominence, both will impact processing immediately synchronous use of cues
- P2 Perception of formants early in processing *will vary* across conditions

GAMM modeling

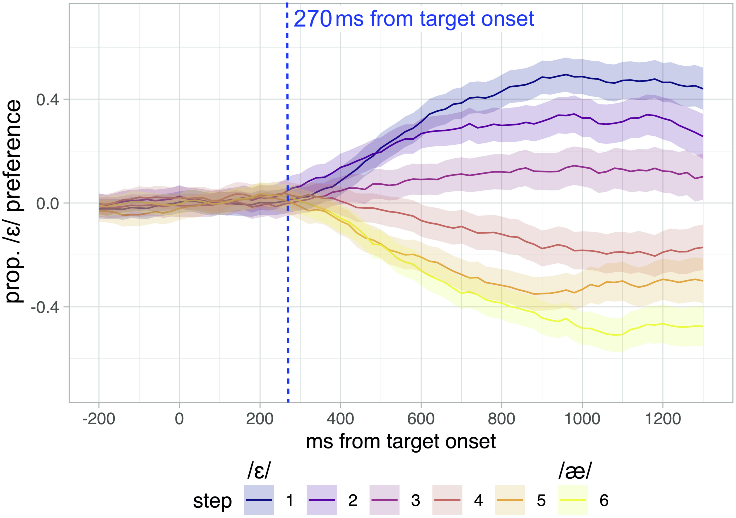


When in time do smooths diverge, and how do they change over time?

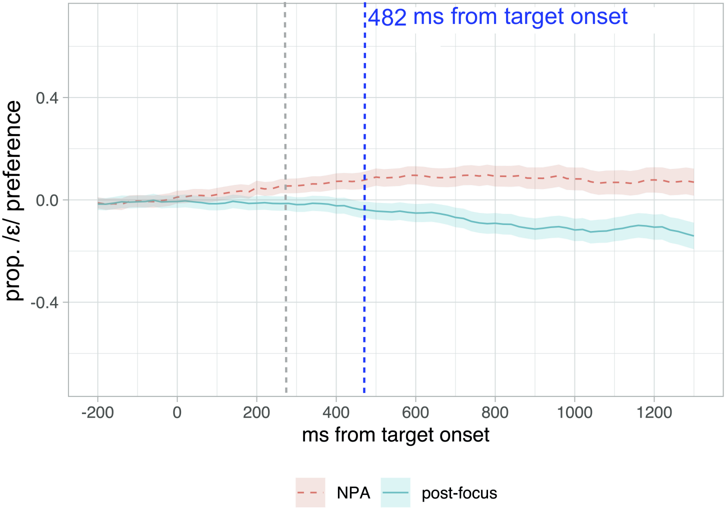
Experiment 2 results: Formants



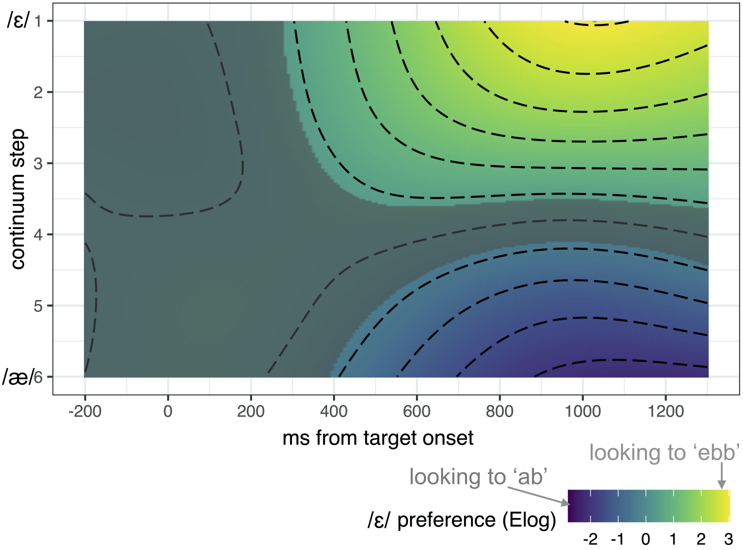
Experiment 2 results: Formants

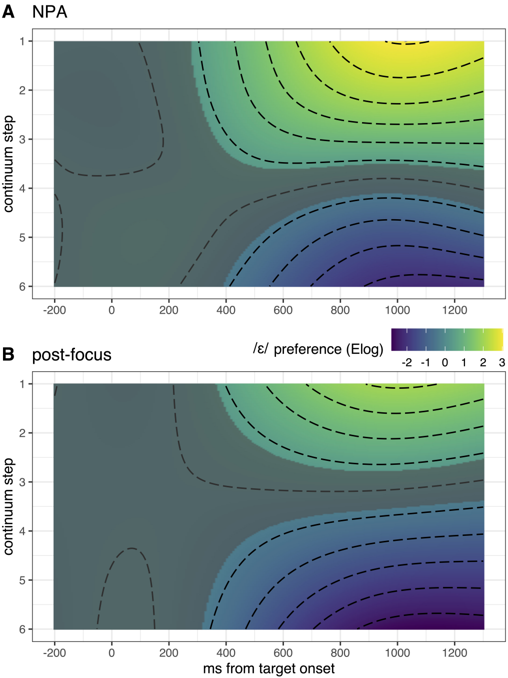


Experiment 2 results: Prominence



Experiment 2 surface plots



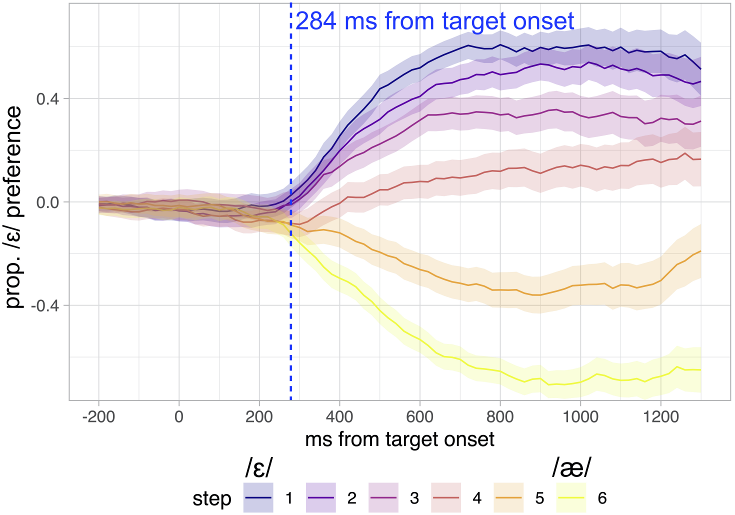

A
NPA
B
post-focus
/ɛ/ preference (Elog)
-2
-1
0
1
2
3
continuum step
1
2
3
4
5
6
-200
0
200
400
600
800
1000
1200
ms from target onset

What we've seen so far...

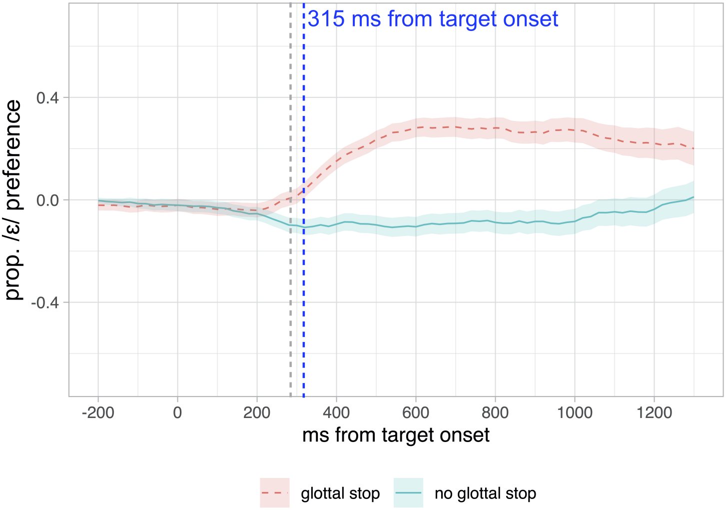
1. Phrasal prominence effects are overall delayed relative to the effect of formant cues
2. But there are subtle immediate impacts on formant perception

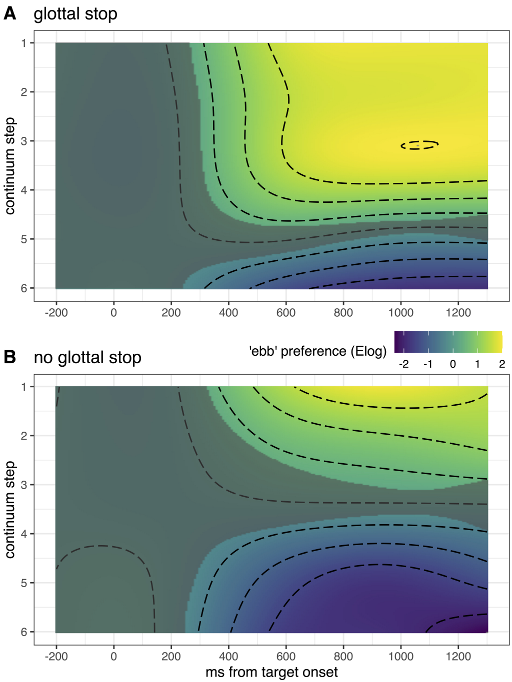
What about [ʔ]?

Experiment 3 results: Formants



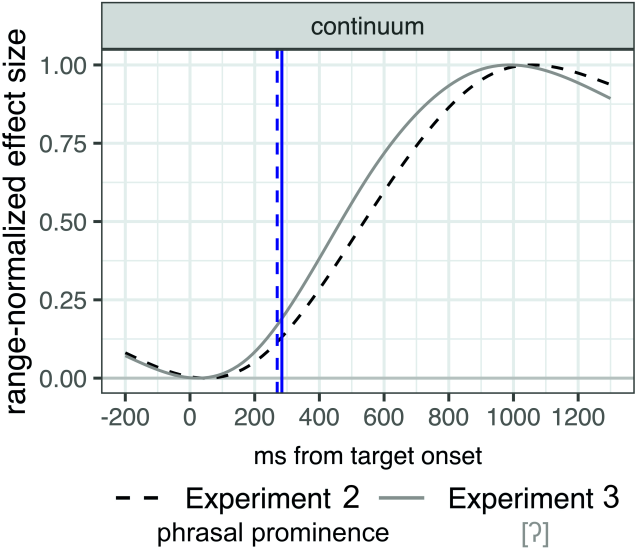
Experiment 3 results: Prominence



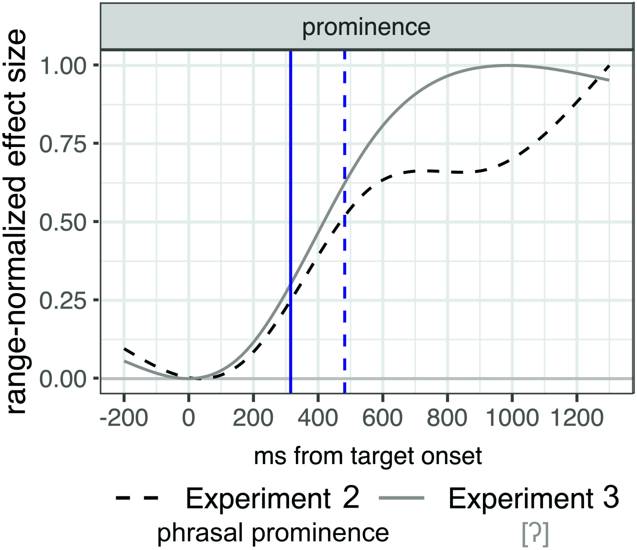


A glottal stop
continuum step
1
2
3
4
5
6
-200
0
200
400
600
800
1000
1200
B no glottal stop
'ebb' preference (Elog)
-2
-1
0
1
2
continuum step
1
2
3
4
5
6
-200
0
200
400
600
800
1000
1200
ms from target onset

Comparing Experiment 2 and 3: Continua



Comparing Experiment 2 and 3: Prominence



Processing prominence and segment

1. When [ʔ] precedes a vowel....

 don’t wait to use prominence information - process in tandem with formant cues

1. When prominence is conveyed by preceding tunes/duration...

  subtler early influences grow slowly reaching a maximum relatively later in processing

This is a notable departure from strictly delayed influence attested for prosodic boundaries[1]

[1] Kim et al., 2018a; Mitterer et al., 2019

Take home messages

1. Does prosodic context shape speech perception?

- Yes
- Both prosodic boundaries (Section I) and prosodic prominence (Section II) mediate perception of segmental contrasts

2. How is prominence processed in relation to segmental material?

- Prominence is integrated immediately with cues to a contrast
- When context is global/phrasal - later-stage reinforcement of prominence effects
- Listeners consider the phonetics and phonology of prominence in processing?

Going forwards

On the processing side:

- Does incoherence between prosodic context and cues hinder processing?
  - Are certain types of incoherence more inhibitory?

Going forwards

On the processing side:

- Does incoherence between prosodic context and cues hinder processing?
    - Are certain types of incoherence more inhibitory?
- How bidirectional are these influences?
    - How do "segmental" cues influence interpretation of prosodic features? (see Mitterer et al., 2020)

Going forwards

On the processing side:

- Does incoherence between prosodic context and cues hinder processing?
  - Are certain types of incoherence more inhibitory?
- How bidirectional are these influences?
  - How do "segmental" cues influence interpretation of prosodic features? (see Mitterer et al., 2020)
- How do these effects fit into existing models of spoken word recognition?

On the typological side:

- How do language-specific patterns of vowel reduction and strengthening translate to perception?
- How does prominence mediate processing in edge-prominence languages?

An abridged thank you

- To Sun-Ah Jun, Pat Keating, Megha Sundara and Taehong Cho
- To RAs for these projects:

  Danielle Bagnas, Qingxia Guo, Jae Weller and Bryan Gonzalez

- To you all!

References

References I

Abramson, A. S. (1976). Laryngeal timing in consonant distinctions. *Haskins Laboratory Status Report on Speech Research, SR-47*, pages 105–112.

Abramson, A. S. and Whalen, D. H. (2017). Voice Onset Time (VOT) at 50: Theoretical and practical issues in measuring voicing distinctions. *Journal of phonetics*, 63:75–86.

Beckman, M. E. (1996). The parsing of prosody. *Language and cognitive processes*, 11(1-2):17–68.

Beckman, M. E., Edwards, J., and Fletcher, J. (1992). *Prosodic structure and tempo in a sonority model of articulatory dynamics*, page 68–89. Papers in Laboratory Phonology. Cambridge University Press.

Bishop, J., Kuo, G., and Kim, B. (2020). Phonology, phonetics, and signal-extrinsic factors in the perception of prosodic prominence: Evidence from rapid prosody transcription. *Journal of Phonetics*, 82:100977.

Brown, M., Salverda, A. P., Dilley, L. C., and Tanenhaus, M. K. (2011). Expectations from preceding prosody influence segmentation in online sentence processing. *Psychonomic bulletin & review*, 18(6):1189–1196.

Chen, M. (1970). Vowel length variation as a function of the voicing of the consonant environment. *Phonetica*, 22(3):129–159.

Cho, T. (2005). Prosodic strengthening and featural enhancement: Evidence from acoustic and articulatory realizations of/ɑ, i/in english. *The Journal of the Acoustical Society of America*, 117(6):3867–3878.

References II

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 Keating, P. (2001). Articulatory and acoustic studies on domain-initial strengthening in Korean. *Journal of phonetics*, 29(2):155–190.

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.

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.

de Jong, K., Beckman, M. E., and Edwards, J. (1993). The interplay between prosodic structure and coarticulation. *Language and speech*, 36(2-3):197–212.

Dilley, L., Shattuck-Hufnagel, S., and Ostendorf, M. (1996). Glottalization of word-initial vowels as a function of prosodic structure. *Journal of phonetics*, 24(4):423–444.

Erickson, D. (2002). Articulation of extreme formant patterns for emphasized vowels. *Phonetica*, 59(2-3):134–149.

Garellek, M. (2013). *Production and perception of glottal stops*. PhD thesis, UCLA.

Garellek, M. (2014). Voice quality strengthening and glottalization. *Journal of Phonetics*, 45:106–113.

Grice, M., Ritter, S., Niemann, H., and Roettger, T. B. (2017). Integrating the discreteness and continuity of intonational categories. *Journal of Phonetics*, 64:90–107.

References III

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., Kim, J., and Cho, T. (2018a). Prosodic-structural modulation of stop voicing contrast along the vot continuum in trochaic and iambic words in american english. *Journal of Phonetics*, 71:65–80.

Kim, S., Mitterer, H., and Cho, T. (2018b). A time course of prosodic modulation in phonological inferencing: The case of Korean post-obstruent tensing. *PloS one*, 13(8).

Lehiste, I. (1970). *Suprasegmentals.* Massachusetts Institute of Technology Press.

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.

Mitterer, H., Kim, S., and Cho, T. (2020). The role of segmental information in syntactic processing through the syntax–prosody interface. *Language and Speech*, page 0023830920974401.

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.

Reinisch, E. and Sjerps, M. J. (2013). The uptake of spectral and temporal cues in vowel perception is rapidly influenced by context. *Journal of Phonetics*, 41(2):101–116.

References IV

Shepherd, M. A. (2008). The scope and effects of preboundary prosodic lengthening in japanese. *USC Working Papers in Linguistics*, 4:1–14.

Silverman, K. and Pierrehumbert, J. (1990). The timing of prenuclear high accents in English. In Beckman, M. E. and Kingston, J., editors, *Papers in Laboratory Phonology*, Papers in Laboratory Phonology, pages 72–106.

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. PMID: 33251945.

Terken, J. and Hermes, D. (2000). The perception of prosodic prominence. In *Prosody: Theory and experiment*, pages 89–127. Springer.

Turk, A. E. and Shattuck-Hufnagel, S. (2007). Multiple targets of phrase-final lengthening in american english words. *Journal of Phonetics*, 35(4):445–472.

Van Summers, W. (1987). Effects of stress and final-consonant voicing on vowel production: Articulatory and acoustic analyses. *The Journal of the Acoustical Society of America*, 82(3):847–863.

Appendix

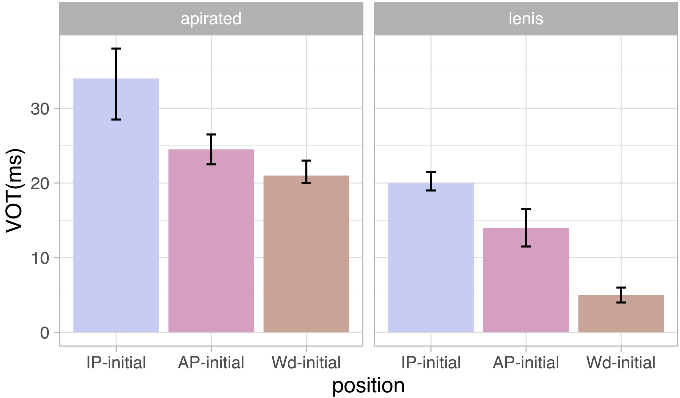
"Segmental" cues vary systematically based on prosody

E.g., voice onset time (VOT) is longer...

- in aspirated vs. unaspirated stops[1]
- in prominent syllables[2]
- at the beginning of prosodic domains[3]

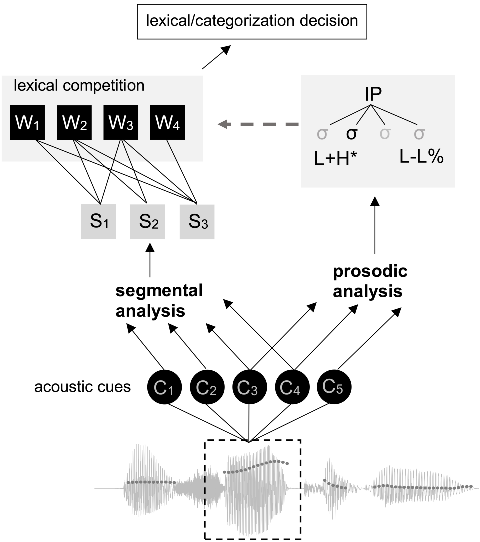
[1] Abramson, 1976; Abramson and Whalen, 2017
[2] Cole et al., 2007; Cho, 2015
[3] Keating et al., 2004

Dual patterning: Korean VOT

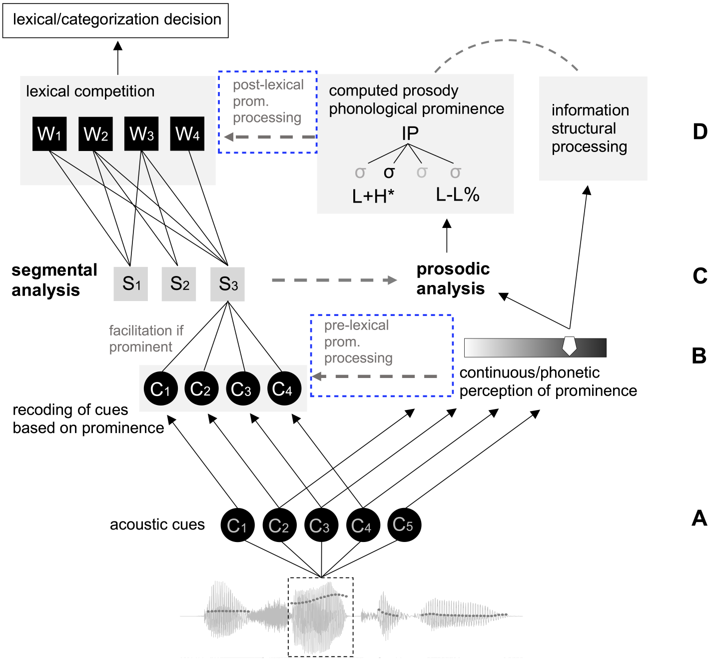


VOT from Korean /t^h/ and /t/, adopted from Cho and Keating (2001)

Schematic of the prosodic analysis model[1]

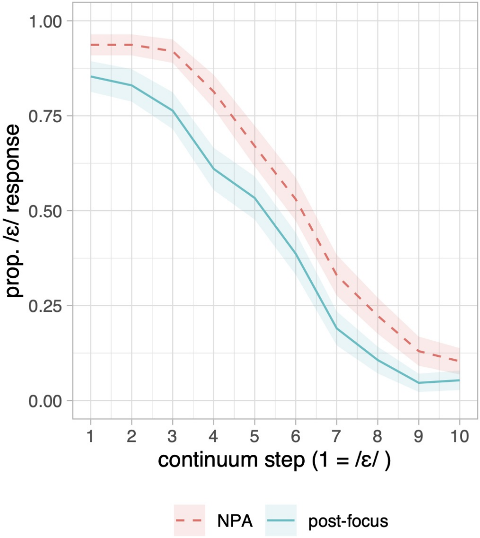


[1] Cho et al., 2007; Kim et al., 2018b; Mitterer et al., 2019



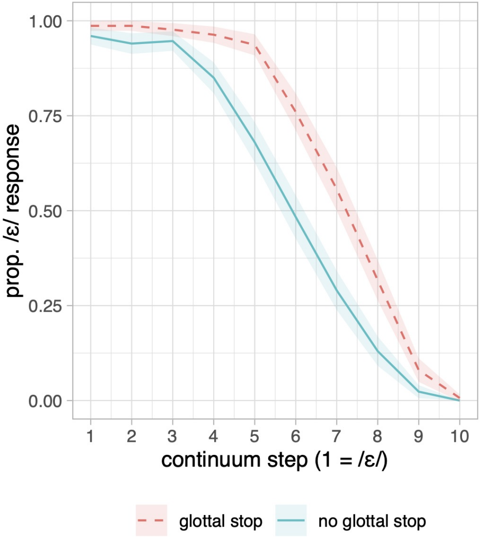
Multistage Assessment of Prominence in Processing (MAPP)

Offline pre-Experiment 2 results



main effect of prom.
$\beta = 0.83$; 95%CI = [0.27,1.39]

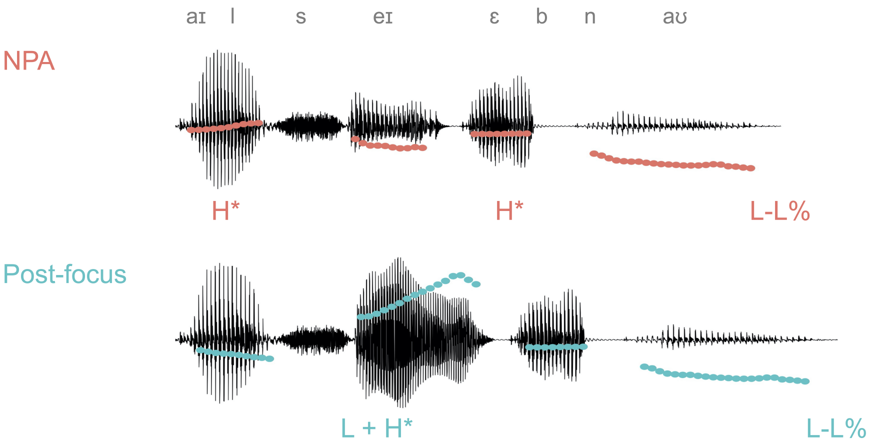
Offline pre-Experiment 3 results



main effect of prom.

$\beta = 1.75$; 95%CI = [1.31,2.22]

Experiment 2 stimuli



Experiment 3 stimuli

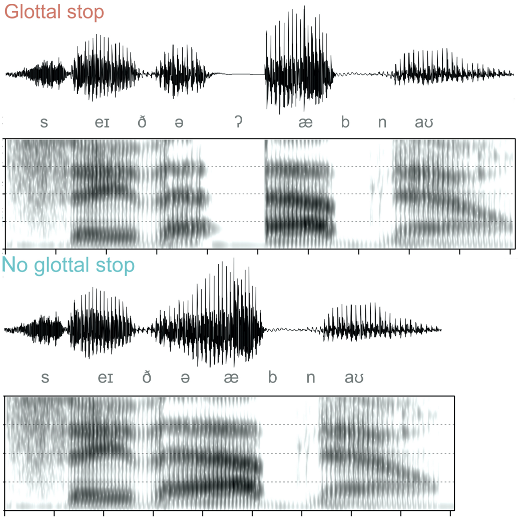
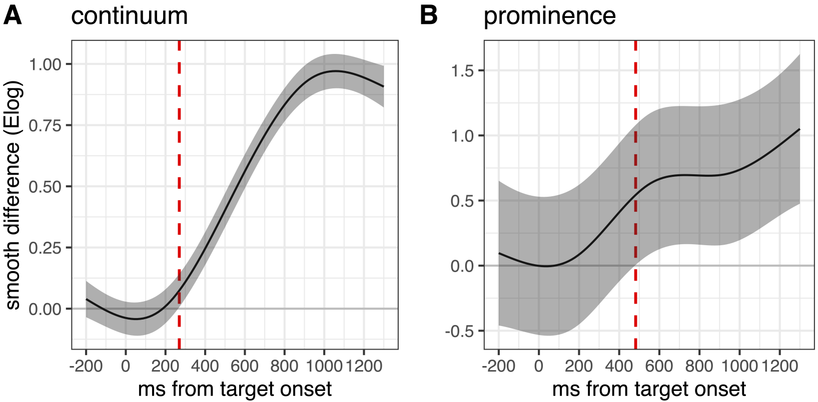
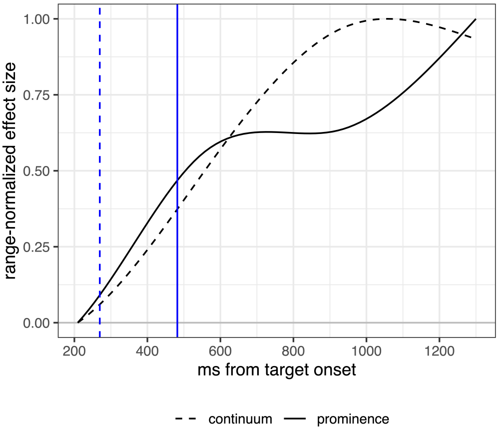


Figure: Waveforms and spectrograms of the Experiment 3 stimuli

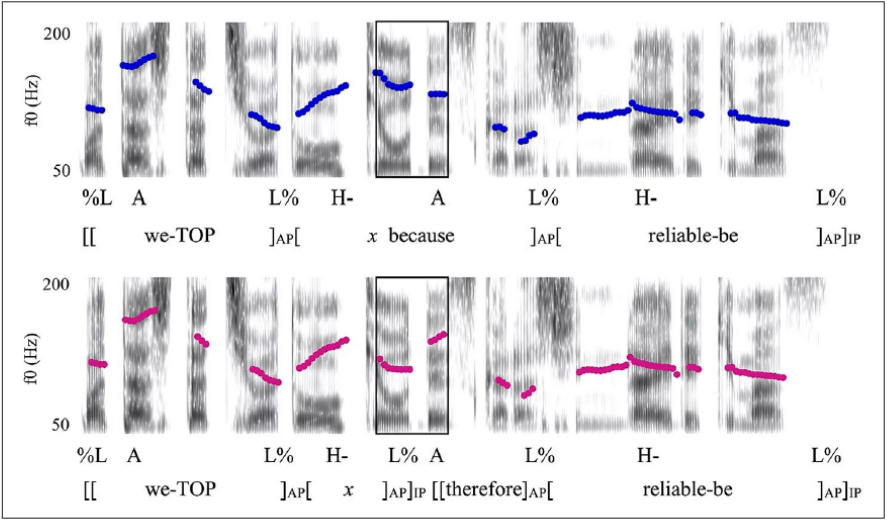
Exp 2 difference smooths



Range-normalized effects Experiment 2



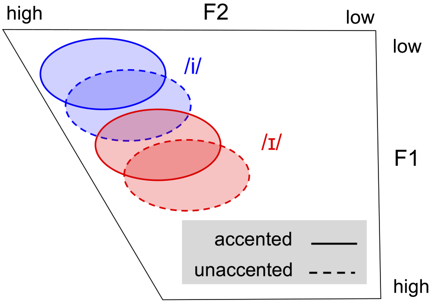
Experiment 1b stimuli



Bonus Experiment: High front vowels

High front vowels, contra sonority expansion, often show:

- Hyperarticulation: enhance features [+high, -back] - raise/front tongue (lower F1/higher F2)



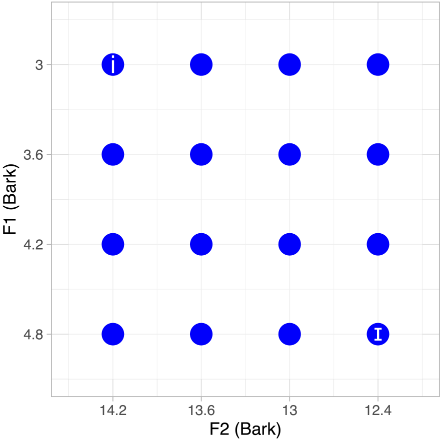
Tests perception of an /ɪ/ to /i/ continuum

Prediction

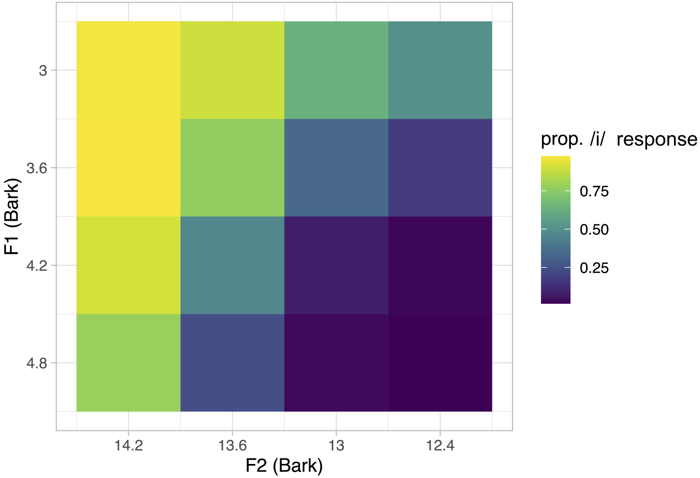
Lower F1/ Higher F2 expected in the NPA condition, fewer tokens mapped to /i/ therein:
decreased /i/ responses in NPA

Bonus Experiment

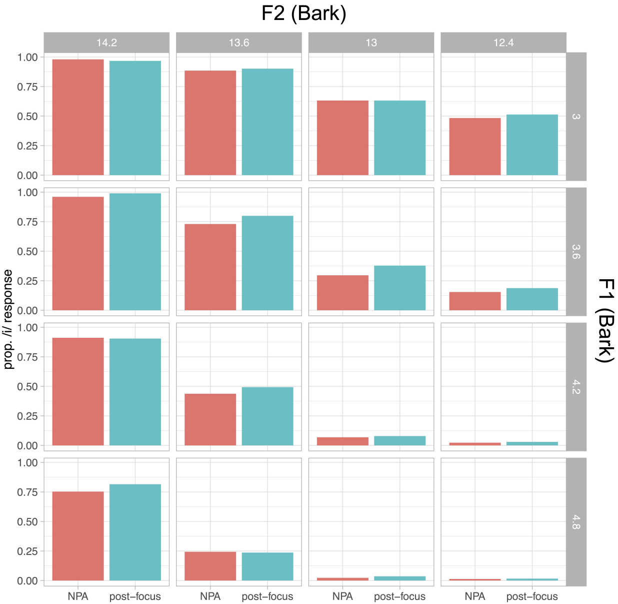
- "sit" to "seat" (/ɪ/ to /i/) continuum
- Same prominence manipulation as Experiment 1: NPA / post-focus



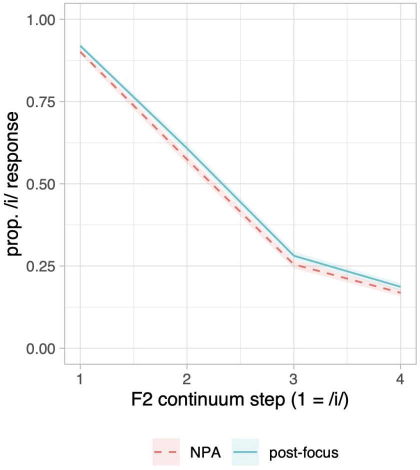
Bonus experiment results: Continuum



Bonus experiment results: Continuum/prominence



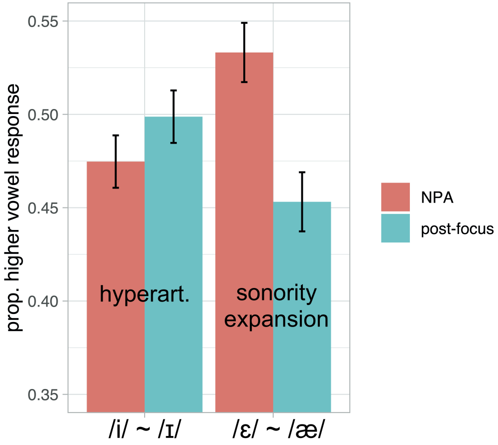
Bonus experiment results: Prominence



main effect of prom. β= -0.26; 95%CI = [-0.42,-0.10]

hyperarticulation expected under prominence

high vowels and non-high vowels



Evidence for contrast-specific expectations of prominence strengthening

Model outputs

Table: Experiment 3

	Estimate	Est. Error	L-95% CI	U-95%CI	credible?
intercept	0.05	0.17	-0.29	0.39	
prominence	0.83	0.28	0.27	1.39	✓
continuum	-2.57	0.28	-3.15	-2.03	✓
prom:cont	-0.24	0.13	-0.50	0.01	

Table: Experiment 5 click responses

	Estimate	Est. Error	L-95% CI	U-95%CI	credible?
intercept	0.09	0.16	-0.21	0.40	
prominence	0.91	0.35	0.22	1.59	✓
continuum	-1.56	0.20	-1.96	-1.17	✓
prom:cont	-0.13	0.12	-0.37	0.11	

Model outputs

Table: Experiment 4

	Estimate	Est. Error	L-95% CI	U-95%CI	credible?
intercept	1.18	0.16	0.87	1.50	✓
glottal stop	1.75	0.23	1.31	2.22	✓
continuum	-3.35	0.17	-3.71	-3.02	✓
glot:cont	-0.79	0.20	-1.20	-0.41	✓

Table: Experiment 6 click responses

	Estimate	Est. Error	L-95% CI	U-95%CI	credible?
intercept	1.11	0.17	0.78	1.45	✓
glottal stop	2.66	0.30	2.08	3.28	✓
continuum	-2.89	0.22	-3.33	-2.49	✓
glot:cont	-0.56	0.23	-1.03	-0.13	✓

Model outputs

Table: Bonus Experiment

	Estimate	Est. Error	L-95% CI	U-95%CI	credible?
intercept	-0.54	0.15	-0.84	-0.25	✓
prominence	-0.26	0.08	-0.42	-0.10	✓
F1	-1.80	0.15	-2.10	-1.52	✓
F2	2.63	0.18	2.28	2.99	✓
F1:F2	0.78	0.11	0.57	1.00	✓
F1:prominence	-0.01	0.10	-0.19	0.19	
F2:prominence	0.01	0.11	-0.20	0.22	
F1:F2:prominence	-0.01	0.10	-0.20	0.19	

Model outputs

Table: Experiment 6

	Estimate	Est. Error	L-95% CI	U-95%CI	credible?
intercept	-0.14	0.16	-0.44	0.17	
prominence	0.49	0.24	0.01	0.95	✓
continuum	-2.61	0.22	-3.05	-2.18	✓
prom:cont	-0.46	0.12	-0.71	-0.25	✓

Model outputs

Table: Model output for the GAMM used in Experiment 2

Parametric terms	Estimate	Est. Error	t-value	p-value
intercept	0.24	0.16	1.50	0.14
continuum	-1.63	0.09	-18.04	< 0.001
prominence	0.45	0.23	1.91	0.05
Smooth terms	**edf**	**ref df**	**F-value**	**p-value**
te(time, continuum; condition = NPA)	17.09	19.71	38.27	< 0.001
te(time, continuum; condition = post-focus)	8.99	9.52	66.32	< 0.001
s(time, participant; condition = NPA)	228.11	323.00	3.91	< 0.001
s(time, participant; condition = post-focus)	231.32	323.00	4.97	< 0.001

Model outputs

Table: Model output for the GAMM used in Experiment 4

Parametric terms	Estimate	Est. Error	t-value	p-value
intercept	0.37	0.08	4.39	< 0.001
continuum	-1.01	0.10	-9.86	< 0.001
glottal stop	0.52	0.12	4.45	< 0.001
Smooth terms	**edf**	**ref df**	**F-value**	**p-value**
te(time, continuum; condition = GS)	21.85	22.48	333.36	< 0.001
te(time, continuum; condition = no GS)	22.00	23.05	282.845	< 0.001
s(time, participant; condition = GS)	268.60	358.00	9.38	< 0.001
s(time, participant; condition = no GS)	284.13	358.00	10.24	< 0.001