Prominence effects on pre-lexical processing
time-course evidence from vowel perception

UCLA Department of Linguistics
Jeremy Steffman jsteffman@ucla.edu
CUNY 2020

I. Introduction

Two important components of spoken language processing:

- determining segments, lexical contrasts → segmental processing
- determining phrasal grouping and prominence → prosodic processing

These two processes usually assumed to operate fairly independently [1]

- however, acoustic info which specifies segment/prosody overlap – prosody shapes segmental cues [2,3] – how should listeners contend with this?

Proposed non-independence of prosodic and segmental processing [3,4]

- prosodic structure as a mediating influence in segmental perception

"Prosody Analyzer" [4]: integration of prosody + segment via lexical competition:
(1) segmental information activates lexical hypothesis
(2) prosodic context modulates lexical activation/competition

Prediction: prosodic contextual effects should occur later in time – esp. following the uptake of segmental cues – some evidence for this from boundary effects [3,5]

Goals of the project

- Test if/how *phrasal prominence* influences perception of spectral cues (Exp. 1)
  - previous work focuses on phrasal boundaries, and durational cues
- Two accounts tested by observing the time-course of prominence effects:
  (a) Prosody Analyzer as a mechanism of integration – delayed time-course
  (b) Prominence as a standard acoustic pattern – no delay in integration [6]

II. Experiment 1 - methods

The test case: prominence driven "sonority expansion" – phrasally prominent vowel articulations show:

- more jaw movement, lowered/backed lingual articulations [7,8]
- this modulates the acoustic structure of vowels: **formants** (= resonance frequencies that characterize different vowels)
- raised first formant (F1) & lowered second formant (F2) [8,9]

F1 and F2 are crucial for making vowel category distinctions
- lexical contrasts: e.g. "head" /ɛ/ versus "had" /æ/.

- F1 & F2 are thus shaped by both **prosody** and **segment**

Task: 2AFC categorization task (n = 30)
Target: 10-step F1-F2 phonetic continuum from "ebb" to "ab" (Fig.1)

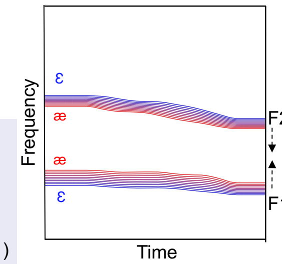

Fig. 1 : Formant tracks showing variation in F1 and F2 along the /ɛ/ - /æ/ continuum**.** The innermost red lines are for the vowel in "had" the outermost blue lines are for the vowel in "head". Note /æ/ has higher F1 and lower F2 than /ɛ/.

Prominence manipulation: two carrier phrase conditions

Nuclear pitch accent (NPA):	I'll say [target] now
target is prominent	H* H* L-L%
Post-Focus :	I'll **say** [target] now
target is non-prominent	L+H* L-L%

Predictions: in prominent contexts: attribution of F1 raising + F2 lowering to prominence marking (not segmental contrast) would lead to more /ɛ/ percepts – i.e. "this is prominent /ɛ/, not /æ/"

- In other words - /ɛ/ becomes acoustically more like /æ/ when prominent – will listeners account for this?

III. Experiment 1 - results

Main effect of prominence confirms predictions – increased "ebb" responses when the target is prominent

Evidence for phrasal prominence effects in vowel contrast perception

- aligns with previous prosodic boundary findings [3,5]

With this finding in hand, we can test how this information is processed online

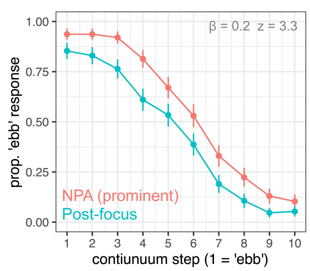

Fig. 2 : "ebb" responses as a function of prominence and continuum step

IV. Experiment 2 - method

2AFC visual world eyetracking task, following [10,11] (n = 32)

- listeners look to and click on orthographic target word representations
- same stimuli as Exp. 1, using the 6 most ambiguous continuum steps [12]
- time-course modeled as log-transformed **preference** for "ebb" [11] using GAMM
  - full parametric terms, smooths for *prominence, time* and *continuum - ti()* tensor product used to model continuum effects over time, random smooths

Time course predictions

- Following **(a)** Prosodic Analysis – formant cues activate lexical hypotheses – prosodic structure mediates lexical competition
  - Prediction: **asynchrony** in the use of formant cues and prominence
  - Spectral cues are used rapidly [11] – a benchmark for what counts as "fast"
- Following **(b)** prominence as a contextual acoustic pattern – **synchronous** use of formant and prominence cues, standard for context effects (e.g. speech rate, spectral context [11])
  - compatible with models where contextual information recodes cue values [13] – or cue integration [6] – no explicit prosodic analysis on the part of the comprehension system

V. Experiment 2 - results

(i) Click responses show prominence increases clicks on "ebb" (Fig. 3)

- replicates categorization responses in Exp. 1

(ii) Eye movement data: both manipulations show effects online (Fig. 4)

- graded preference for "ebb" along the continuum (panel A) – sanity check, showing listener sensitivity to fine-grained F1-F2 differences
- clear online influence of prominence manipulation (panel B) – though relatively small

(iii) Time course assessment (Fig. 5 at right)

- formant information is used relatively early in time –standard time course for intrinsic spectral cues [11] (panel C)
- The observed prominence effect is **asynchronous,** following formant cues by 300ms (panel D)

Take home messages

1. Listeners incorporate prosodic prominence in segmental processing
   - Prominence effects are delayed – consistent with recent proposals - later stage integration of prosody in lexical competition

Further questions

1. Does signal-extrinsic prominence (e.g. information structural cues from context) also show this effect? cf. [14]
2. What perception/processing model best accounts for these findings?

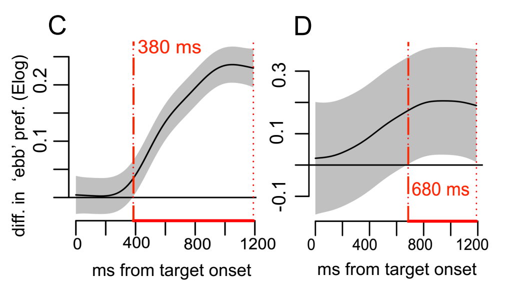

Fig. 5 : Smooth differences as a function of continuum step (C) and prominence (D) – values are log-transformed preference measures Divergence from zero represents a significant effect in time.

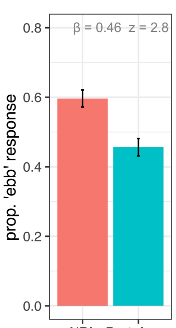

Fig. 3 : "ebb" responses as a function of prominence (pooled by continuum step).

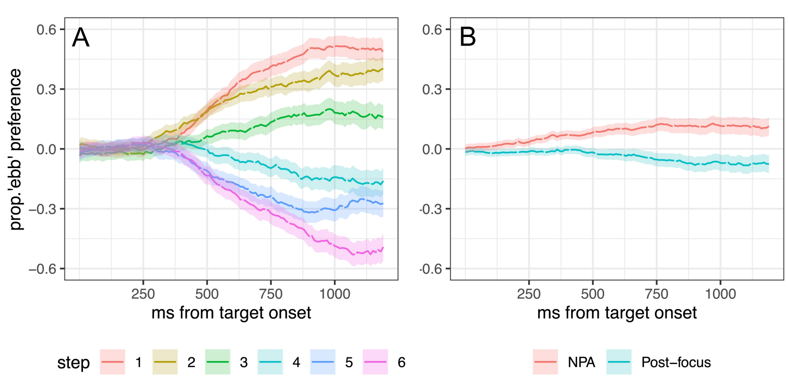

Fig. 4 : "ebb" preference in looks as a function of continuum step (A) – and prominence (B) – plotted preference measure calculated as prop. looks to "ebb" minus looks to "ab".

Acknowledgments: Thanks to Adam Royer for recording for stimuli, to Danielle Bagnas, Qingxia Guo, Juliana Casparian and James Weller for data collection, & to the UCLA Psycholinguistics seminar, Sun-Ah Jun, Pat Keating, Megha Sundara and Taehong Cho for feedback.

Selected refs: [1] Cutler, A., Dahan, D., & Van Donselaar, W. (1997). Prosody in the comprehension of spoken language: A literature review. *Language and speech*, *40*(2), 141-201. **[2]** Keating, P. (2006). Phonetic Encoding of Prosodic Structure. In J. Harrington & M. Tabain (Eds.), *Speech production: Models, phonetic processes, and techniques* (pp. 167–186). Macquarie Monographs in Cognitive Science, Psychology Press. **[3]** Cho, T., McQueen, J. M., & Cox, E. A. (2007). Prosodically driven phonetic detail in speech processing: The case of domain-initial strengthening in English. *JPhon*, *35*(2), 210–243. **[4]** Kim, S., Mitterer, H., & Cho, T. (2018). A time course of prosodic modulation in phonological inferencing: The case of Korean post-obstruent tensing. *PLOS ONE*, *13*(8). **[5]** Mitterer, H., Kim, S., & Cho, T. (2019). The glottal stop between segmental and suprasegmental processing: The case of Maltese. *JML*, *108*, 104034. **[6]** Toscano, J. C., & McMurray, B. (2012). Cue-integration and context effects in speech: Evidence against speaking-rate normalization. *Attention, Perception, & Psychophysics*, *74*(6), 1284–1301. **[7]** Van Summers, W. (1987). Effects of stress and final-consonant voicing on vowel production: Articulatory and acoustic analyses. *JASA*, *82*(3), 847–863. **[8]** 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. **[9]** Mo, Y., Cole, J., & Hasegawa-Johnson, M. (2009). Prosodic effects on vowel production: Evidence from formant structure. *INTERSPEECH*, 2535–2538. **[10]** Kingston, J., Levy, J., Rysling, A., & Staub, A. (2016). Eye movement evidence for an immediate Ganong effect. *Journal of Experimental Psychology. Human Perception and Performance*, *42*(12), 1969–1988. **[11]** Reinisch, E., & Sjerps, M. J. (2013). The uptake of spectral and temporal cues in vowel perception is rapidly influenced by context. *JPhon*, *41*(2), 101–116. **[12]** Mitterer, H., & Reinisch, E. (2013). No delays in application of perceptual learning in speech recognition: Evidence from eye tracking. *Journal of Memory and Language*, *69*(4), 527–545. **[13]** McMurray, B., & Jongman, A. (2011). What information is necessary for speech categorization? Harnessing variability in the speech signal by integrating cues computed relative to expectations. *Psychological Review*, *118*(2), 219–246. **[14]** Bishop, J. (2012). Information structural expectations in the perception of prosodic prominence. *Prosody and meaning*, *25*, 239.