Prominence effects on the processing of spectral cues: testing sonority expansion in perception

Jeremy Steffman

It is well established that prosodic factors modulate phonetic detail in the realization of segments [1,2]. Recent studies have explored how listeners are sensitive to this prosodically conditioned phonetic detail in speech perception, focusing primarily on boundary phenomena (final lengthening, initial strengthening) and the perception of durational cues [3,4]. The present study extends this work by testing how prosodic *prominence* influences perception of *spectral* cues. In two experiments, perception of a vowel contrast under different types of prominence marking was tested.

One effect of prominence on formant structure is *sonority expansion*. Prominent vowels show expanded sonority, defined in this context as overall openness of the vocal tract [5]. Sonority expansion may be taken as a form of syntagmatic contrast enhancement for adjacent consonants or nearby vowels [2] (note the picture is more complicated for high vowels, where oral cavity expansion may conflict with specification of a paradigmatic [+high] feature [6]). Prominence-driven sonority expansion (for non-high vowels) results in increased amplitude of jaw movements [7], and lowered and backed lingual articulation [6], serving to expand the oral cavity and correlating with higher F1 and lower F2 values [8]. The present study therefore tests if listeners' perception of a vowel contrast (cued by changes in F1 and F2) is mediated by prominence. The contrast between American English [ɛ] and [æ] was chosen as a test case. [æ] generally has lower F2 and higher F1 than [ɛ] [9] and is therefore a more sonorous vowel in terms of F1/F2. In two experiments, listeners categorized a 10-step [ɛ] to [æ] continuum that varied in only F1 and F2 (created via LPC decomposition/resynthesis in Praat [10]). Given the dual nature of F1/F2 as cuing a phonemic contrast *and* varying along a sonority scale, it was predicted that prominence should mediate vowel perception: along the continuum, more sonorous ([æ]-like) F1/F2 values should be categorized as [ɛ] in prominent contexts where sonorous F1/F2 is expected, i.e. as compensation for sonority expansion. This predicts that prominent contexts will *increase [ɛ] responses.*

Prominence was cued by varying the phrasal prosodic context in which the target word appeared (Experiment 1), and by the presence of glottalization in a vowel-initial target word, which has been argued to have a prominence marking function in American English [11,12] (Experiment 2).

Both experiments were a 2AFC task, with the target word categorized as 'ebb' or 'ab'. In **Experiment 1**(n=30), prominence was manipulated by varying the phrasal prosody in the carrier sentence ("I'll say ebb/ab now"). The target word (identical across conditions) was placed in a broad focus context in which it bore nuclear prominence (the nuclear pitch accent (NPA) condition) or was post-focus, following narrowly focused "say" (the post-focus condition), therefore lacking prominence (see Fig. 1 below). In **Experiment 2** (n=30), prominence was cued by the presence/absence of a glottal stop preceding the target in a two word phrase "the ebb/ab". If a glottal stop signals prominence, following [11,12], a [ʔ]-initial target should therefore undergo expected sonority expansion, in similar fashion to a phrasally prominent target. Exp. 2 stimuli were created by manipulating F1/F2 of a target following the word "the" [ði], resulting in a [ðiɛb]-[ðiæb] continuum. A glottal stop from another production was cross-spliced preceding the target word to create an otherwise identical [ðiʔɛb]-[ðiʔæb] continuum (see Fig. 2).

Results (assessed with mixed-effects logistic regression, by-participant random intercepts, maximal random slopes) show a main effect of prominence in both Experiments. In Exp.1: the NPA condition significantly increased 'ebb' responses (β=0.42, z=3.26, Fig. 3). In Exp. 2: the presence of a glottal stop likewise increased 'ebb' responses (β=1.26, z=11.29, Fig. 4). Together, these results provide novel evidence that listeners incorporate prominence in their perception of vowel contrasts, exploiting sonority expansion patterns. Though Experiment 1 and 2 manipulated prominence in different ways, similar effects were observed, suggesting an abstract representation of prominence with various possible segmental and phrasal cues to which listeners are sensitive. Results will be further discussed in terms of their implications for models of speech perception, parallel representations of prosodic and segmental structures in speech processing [13], and perceptual mechanisms that may explain how prosodic and segmental cues interact, including cue integration, exemplar accounts, and abstract prosodic analysis.

Fig. 1

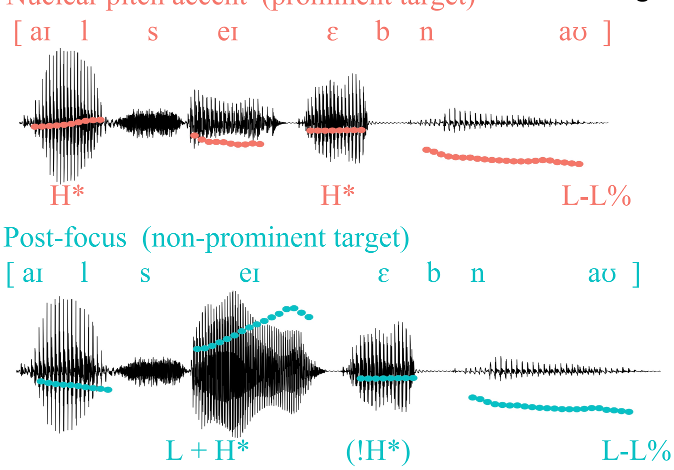

Fig. 2

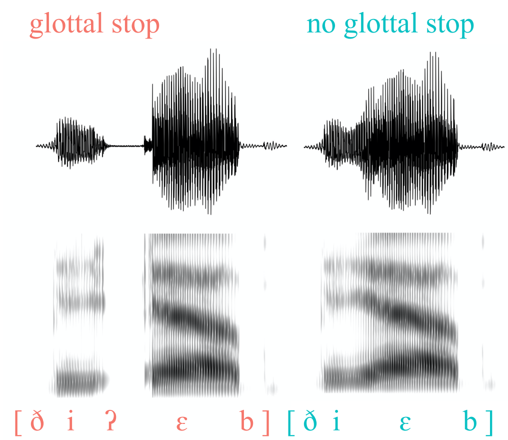

Fig. 1: Waveforms overlaid with pitch tracks showing the conditions in Exp. 1. The pitch range spans the maximum amplitude of the waveforms, and is 50-250 Hz. A segmental transcription is given above, and ToBI labels below the waveforms. The target is the same across conditions, such that perceived target prominence is purely contextual.

Fig. 2: Waveforms and spectrograms (0-4kHz range) of the conditions for Exp. 2, using the most "ebb"-like continuum step. A segmental transcription is given below the spectrogram.

Fig. 3

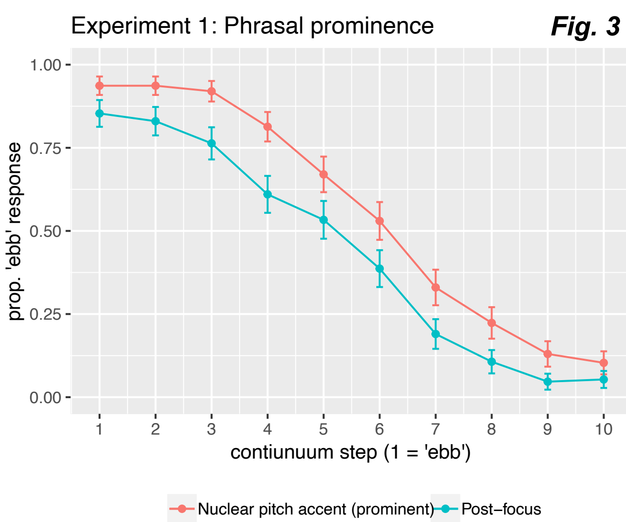

Fig. 4

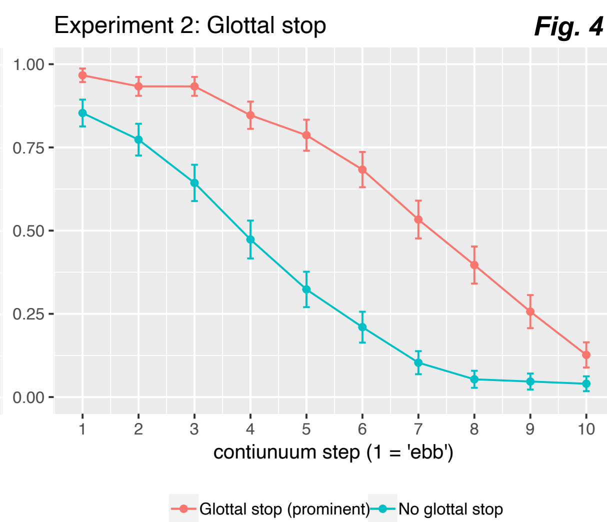

***Fig. 3** (at left) & **Fig. 4** (at right):* Categorization responses for Experiments 1 and 2, split by prominence condition (note in both cases the red line represents the more prominent condition). The *x* axis shows the continuum step, *y* axis shows the proportion of "ebb" responses in each condition. Error bars show 95% CI for each point.

References: **[1]** Cho, T. (2016). Prosodic Boundary Strengthening in the Phonetics–Prosody Interface. *Language and Linguistics Compass*, *10*(3), 120–141. **[2]** de Jong, K. J. (1995). The supraglottal articulation of prominence in English. *JASA*, *97*(1), 491–504. **[3]** Mitterer, H., Cho, T., & Kim, S. (2016). How does prosody influence speech categorization? *JPhon*, *54*, 68–79. **[4]** Steffman, J. (2019). Phrase-final lengthening modulates listeners' perception of vowel duration as a cue to coda stop voicing. *JASA*, *145*(6), EL560–EL566. **[5]** Silverman, K., & Pierrehumbert, J. (1990). The timing of prenuclear high accents in English. *Papers in laboratory phonology I*, 72-106. **[6]** Cho, T. (2005). Prosodic strengthening and featural enhancement: Evidence from acoustic and articulatory realizations of /ɑ,i/ in English. *JASA, 117*(6), 3867–3878. **[7]** Van Summers, W. (1987). Effects of stress and final-consonant voicing on vowel production. *JASA*, *82*(3), 847–863. **[8]** Lee, E. K., Cole, J., & Kim, H. (2006, May). Additive effects of phrase boundary on English accented vowels. In *Proceedings of the 3rd Speech Prosody Conference, Dresden, Germany*. **[9]** Clopper, C. G., Pisoni, D. B., & De Jong, K. (2005). Acoustic characteristics of the vowel systems of six regional varieties of American English. *JASA*, *118*(3), 1661-1676. **[10]** P. Boersma and D. Weenink, Praat: doing phonetics by computer [Computer program], 2019 **[11]** Dilley, L., Shattuck-Hufnagel, S., & Ostendorf, M. (1996). Glottalization of word-initial vowels as a function of prosodic structure. *JPhon*, *24*(4), 423–444. **[12]** Garellek, M. (2014). Voice quality strengthening and glottalization. *JPhon*, *45*, 106–113. **[13]** Mitterer, H., Kim, S., & Cho, T. (2019). The glottal stop between segmental and suprasegmental processing. *JML*, *108*, 104034.