Prosodic boundary strength and the location of pre-nuclear phrasal prominence

Soundess Azzabou-Kacem, Alice Turk & Ellen Gurman Bard

Linguistics and English Language, The University of Edinburgh, Edinburgh, UK
S.Azzabou-kacem@sms.ed.ac.uk; turk@ling.ed.ac.uk; e.bard@ed.ac.uk

ABSTRACT

The assignment of pre-nuclear phrasal prominence in sequences like *THIRteen **MEN** (vs. thir**TEEN**)* is not fully understood. We propose that the location of pre-nuclear prominence in *thirteen* is sensitive to the strength of the prosodic boundary separating the clashing words *thirteen* and *men*: a weaker prosodic boundary should invite early prominence on *THIRteen*. Eleven speakers read 14 potentially clashing word pairs in syntactic contexts likely to elicit a stronger boundary (NP+ VP: *How does the **canteen cook** these days?)* or a weaker boundary (Pre-modifier+NOUN: *Who is the **canteen cook** these days*?). Syntactic condition had significant effects in the predicted direction on the strength of the critical prosodic boundary. Results suggest that boundary strength modifies the relative prominence of the two syllables in the first word (e.g. *canteen)*, but analyses of F0 and duration suggest that the effect is primarily, or perhaps even exclusively, localised on the 2nd syllable, e.g. *-teen*.

Keywords: Pre-nuclear phrasal stress, Prosodic boundary strength, the Rhythm Rule, Stress clash.

1. BACKGROUND

Almost 40 years after the publication of Liberman and Prince's landmark paper on stress and linguistic rhythm [13] the location of pre-nuclear phrasal prominence in sequences like *THIRteen **MEN*** (*vs. thir**TEEN***), is not fully understood. The main proposal, *e.g.* [12, 13] within the framework of metrical phonology, was that early post-lexical prominence on the initial syllable *thir*- of *THIRteen* is the output of the rhythm rule, an **optional** rule where the late stress in *thirTEEN* 'shifts' to the initial syllable to produce a more eurythmic pattern [10], following the principle of rhythmic alternation [14] *interalia*. Most studies within the metrical framework identify foot-level stress clash as the trigger of early post-lexical prominence or "stress shift" in their accounts. Yet early prominence is known to occur without apparent stress clash, [5]; [2]; [6, 7]; [17].

Instead, Gussenhoven [8, 9] and Shattuck-Hufnagel [15-18] view early post-lexical prominence as one of the possible outcomes of the distribution of pitch accents across the intonational phrase domain, a domain delimited by two phrase boundaries, with a pair of pitch accents near the phrase edges [4]. These models link early prominence to structural factors, namely to the position of the word carrying pre-nuclear stress (*e.g. thirteen*) in the phrase. Gussenhoven argues that early prominence on the first (hereafter 'target') word (is the result of accent deletion on its second syllable, *e.g. -dee* in *Dundee#* as this syllable loses its right-edge peripheral position (hereafter '#') when it is embedded in a longer phrase (*Dundee marmalade#); cf.* Horne [11] for supporting research. Shattuck-Hufnagel and colleagues explain early post-lexical prominence, *e.g.* in *#MAssaCHUsetts **GO**vernor* as pitch accent placement on the first available accentable syllable after a left-edge initial prosodic boundary. Thus, early prominence placement plays an edge-marking demarcation role for the onset of the phonological phrase. See also Selkirk [14] for an alternative strategy for stress-clash resolution.

Recent research on prosodic constituency attests a hierarchy of boundaries of gradient strength, [19]; [20]. Following Gussenhoven [8, 9] and Shattuck-Hufnagel et al. [17] we suggest that pre-nuclear prominence could serve as edge demarcation. Additionally, we suggest that early prominence is sensitive not only to the location of prosodic boundaries within and around the critical phrase but also sensitive to their relative strength. Following

Astésano et al. [1], who demonstrated that the occurrence of optional Initial Accent in French depended on prosodic boundary strength, we hypothesise that the likelihood of early pre-nuclear prominence may be influenced by boundary strength, as well as the occurrence of so-called stress clash.

To test our proposal, we varied boundary strength for the same lexical and metrical structures. We manipulated the boundary between the two clashing words, like *antique sink* or *canteen cook* by placing them in contexts where they would be parsed as either Pre-modifier + NOUN (*Who is the canteen cook these days?*) or as NP + VP (*How do the canteen cook these days?*).

If our prediction is correct, this manipulation should affect both boundary strength and the location of pre-nuclear prominence. *i.e.* the Pre-Modifier + NOUN condition, should give rise to a weaker internal boundary and a greater tendency to accent the initial accentable syllable in the phrase than does the NP + VP condition where the stronger internal boundary encourages the assignment of pre-nuclear phrasal stress to the second accentable syllable (*e.g. –teen* in *canteen cook*).

In this paper we report on the temporal and F0 results for the two syllables of the pre-nuclear target, *e.g. can-* and *-teen* in *canteen*.

2. MATERIAL

2.1. Clash sequences

We designed a clash sequence (*e.g. canteen COOK)* with a minimal number of syllables: a disyllabic Word 1 (*e.g. canteen*) followed by a monosyllabic Word 2 (*e.g cook*), where Word 2 is designed to carry the main phrasal stress and Word 1 the pre-nuclear phrasal stress. We constructed 14 homophonous pairs where Word 2 could be parsed either as noun or as verb (*e.g. antique sink*, *Chinese watch*, *Bombay means, shampoo foam, etc…*). These constitute the critical phrases in the experiment.

2.2. Experimental design

2.2.1. Targets

The critical phrases for both structures in a pair were embedded in carrying sentences matched for number of syllables, and were preceded by sentences designed to elicit pre-nuclear prominence on Word 1, and nuclear prominence on Word 2, *e.g.*:

Critical sentence-trio (Pre-modifier + NOUN)
Context: *Zoe got up to some mischief at nursery once again!*
Leading sentence: *She ate the salted dough again.*
Target sentence: *She ate the shampoo foam again.*

Critical sentence-trio (NP+VP)
Context: *Zoe had a wonderful time at nursery this week.*
Leading sentence: *She made the water drip again.*
Target sentence: *She made the shampoo foam again.*

2.2.2. Fillers

Sixteen fillers used the same sentence-trio structure as the targets. Filler-Word-1s were tri-syllabic with stress on the second syllable, *e.g. potato* in *potato omelette.*

Target and filler trios alternated. The items were pseudo-randomised across 2 parts of the experiment with each item appearing in the same order in one part as its homophonous counterpart in the other.

2.3. Participants

Data were analysed from 11 participants (1 male and 10 female), who were students or recent alumni of the University of Edinburgh.

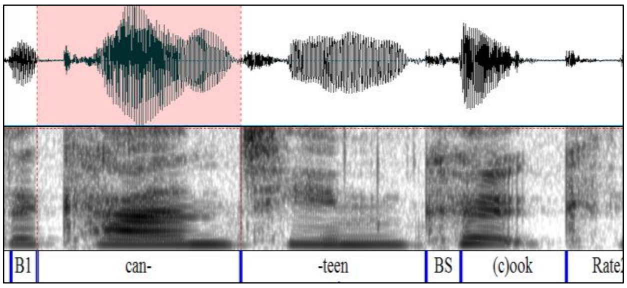
2.4. Methodology

Participants were asked first to read each sentence trio silently and understand the relationships among sentences; then, to produce the sentences so as to reflect those relationships.

Sound recordings were made using a high-quality microphone and digitised at a sampling rate of 48 KHz, at a bit depth of 16. These were edited using *Audacity* and manually segmented using *Praat* [3]. Syllabification normally followed the maximal onset principle. When the segmentation of syllable offsets was likely to be unreliable, as *e.g.* when the syllable boundary occurred between *e.g.* /n/ and /t/ in *canteen*, syllable 1 was taken to end at the release of the /t/ (*cf*. highlighted interval in figure 1).

F0 was measured in intervals within the vowels of syllable 1 (*e.g. can-* in *canteen*) and syllable 2 (*-teen*), trimmed to exclude the first and last 15 ms of each interval. Critical phrases with low pitch accent on the initial syllable and those with a low up-stepping contour were excluded from the F0 analysis, as were their matching homophones in the other syntactic condition.

Figure 1: Waveform and spectrogram of the critical phrase exemplified by *canteen COOK*. B1 is the boundary before Word 1 and BS is Intra-phrase Boundary Strength. The highlighted interval is syllable 1 *-can*.



2.5. Variables

The main dependent variables were:

- Syllable 1 (s1) duration;
- Syllable 2 (s2) duration;
- Boundary Strength (BS) measure: onset closure or VOT of the initial consonant in Word 2 *e.g.* constriction duration of /s/ of *sink* in *antique sink*.
- Rate reference: the duration of a constant interval in the post-nuclear domain, *e.g.* the last word *again* in *She made the shampoo FOAM* ***again***.
- F01 and F02: mean fundamental frequency on syllables 1 and 2.
- F01stdev and F02stdev: standard deviation of fundamental frequency on syllable 1 and syllable 2, used to quantify the size of the F0 expansion over the vowel.
- F01Max, F02 Max and F03Max: respectively the mean of F0 peaks over the trimmed vowel intervals of syllable1, syllable2 and the phrase nucleus (e.g. *cook* in *canteen* ***COOK***) normalized in quarter tones. F03Max is used as an anchor point for s1 and s2 F0 variation.

3. ANALYSIS

For the purpose of valid comparisons across the syntactic conditions we analysed paired items *i.e.* one item per condition per speaker. Items were included in the analysis only if they had the phrase nucleus on Word 2, *e.g. cook* in *canteen* ***COOK***.

3.1. Syntactic condition affects boundary strength

To test for boundary strength differences across conditions (Pre-modifier + NOUN *vs.* NP + VP), we ran a multiple linear regression analysis on the normalised durations (*z* scores) of a boundary strength measure which was appropriate for each pair (*e.g.* VOT of /k/ at the onset of *cook*) on a subset of the data (the 11 out of 14 items, as for the 3 other items word 2 (*e.g. watch* in *Chinese watch)* started with an approximant). The predictors were Syntactic Condition, normalised rate (zRate) and Condition*zRate.

As table 1 shows, NP + VP sequences had slightly but significantly longer boundary strength intervals than Pre-modifier-NOUN sequences. ($F(3,172) = 3.46$; $p < .005$). Rate also had a significant effect ($\beta = 18$, $t = 2.34$, $p < .005$), but there was no significant interaction of Condition and Rate.

Table 1: Mean duration (in milliseconds) and standard deviations of Boundary strength interval across syntactic condition (Pre-modifier + NOUN vs. NP + VP).

Boundary strength interval	Pre-modifier + NOUN	NP+VP
Mean	86	92
Std. deviation	43	48

3.2. Syntactic condition affects Syllable 2 duration

A multiple linear regression on the raw durations of syllable 2, used: Condition, zRate, and zBS*zRate as predictors. The duration of the pre-nuclear clashing syllable (*cf.* table 2) is significantly longer, in the NP + VP condition than in the Pre-modifier + NOUN ($F(3,220) = 4.7$; $p < .05$). Rate had no significant main effect on the duration of syllable 2 across conditions, but there was an interaction of Rate and BS ($\beta = -.14$, $p < .005$), which shows that in rapid speech and weaker boundary strength, syllable 2 duration may vary less across conditions.

Table 2: Mean durations (in milliseconds) and standard deviations for syllable 2 (*e.g. –tique* in *antique*) across syntactic condition (Pre-modifier + NOUN *vs.* NP + VP)

Syllable 2 duration	Pre-modifier + NOUN	NP + VP
Mean	221	244
Std. dev.	59	60

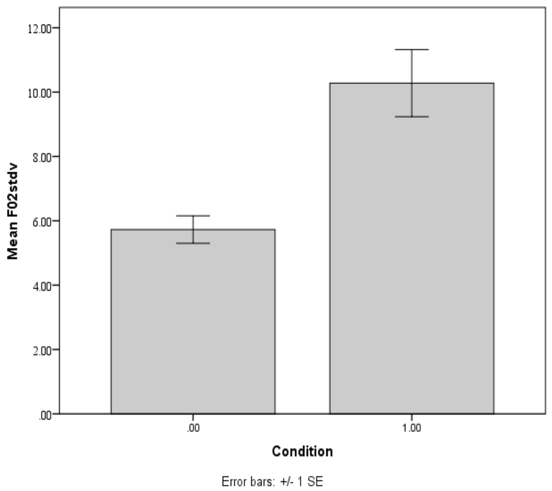
3.3. Syllable 2 F0 results

Table 3: Means (in quarter tones) of the F0 standard deviations of syllable 2 across syntactic conditions (Pre-modifier + NOUN *vs.* NP + VP)

F0stdev	Pre-modifier + NOUN	NP + VP
Mean	5.73	10.28
Std. error	.43	1.04

The mean F02 showed similar values across syntactic conditions (Pre-modifier + NOUN *vs.* NP + VP). But the size of the F0 expansion over the vowel of syllable 2 was significantly larger in the NP + VP condition, *e.g. She made the **shampoo foam** again* (*F*(1,220) = 18.53, *p* < *.0001), cf.* table 3 and figure 2.

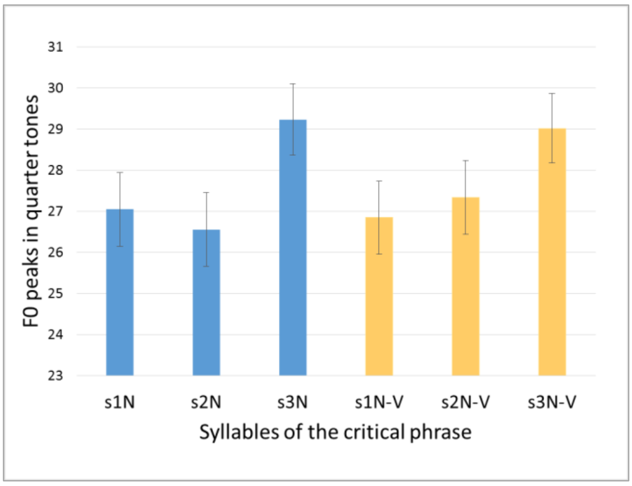
Figure 2: Variation of syllable F0 standard deviation values across conditions (in quarter tones. 0 = Pre-modifier + NOUN; 1 = NP + VP).



3.4. The pre-nuclear 'clashing' syllable is more prominent in the NP + VP condition

Because F0 effects will be affected by overall F0 level, we calculated the mean differences in F0 peaks relative to an anchor point the peak of the phrase nucleus on foam in e.g. *shampoo foam*. F03Max-F02Max was significantly smaller in the VP_NP condition (*M*=1.7 *SD* = 4) compared to the Pre-modifier + NOUN condition (*M*=2.66 *SD* = 3.66), *t*(1) = 1.87, *p* < .05), suggesting greater F0 prominence on Syllable 2 in the NP + VP condition (*cf.* figure 3).

Figure 3: F0 Peaks (and standard errors) for critical phrase syllables as a function of syntactic condition (N = Pre-modifier +NOUN; N-V = NP + VP; s1 and s2 comprise the target word; s3 is the nuclear syllable.



3.5. No clear evidence for modification of the absolute prominence of the initial syllable

Syllable 1, one of the likely sites for the location of pre-nuclear prominence, showed similar values for syllable F0 mean and F0stdev. Overall syllable duration was 10 milliseconds longer in the Pre-modifier + NOUN condition but this difference did not reach significance.

4. DISCUSSION

Since the boundary Strength interval was significantly longer for the NP + VP condition than for the Pre-modifier-NOUN condition, the present results show that varying the syntactic structure (Pre-modifier + NOUN *vs.* NP + VP) can elicit a successful manipulation of prosodic boundary strength. In addition, our predictions for the effect of syntactic locality on 1) the strength of the intra-phrase boundary and 2) on the prominence of the pre-nuclear clashing syllable (syllable 2) were met.

At the current stage of the analysis our results support Gussenhoven's [8] theory of pitch accent deletion over the first clashing syllable 2 (*e.g. –teen* in *canteen)* and [11]; [6]. Additional analyses of spectral tilt data will be required to confirm/disconfirm this conclusion.

5. REFERENCES

[1] Astésano, C., Gurman Bard, E., Turk, A. 2007. Structural Influences on Initial Accent Placement in French. *Language and Speech* 50 (3), 423 -446.

[2] Beckman, M., Edwards, J. 1994. Articulatory evidence for differentiating stress categories. In: Keating, P.A. (ed), *Papers in Laboratory Phonology III: Phonological Structure and Phonetic Form.* Cambridge, U.K.: Cambridge University Press, 1–33.

[3] Boersma, P., Weenink, D. 2015. *Praat: doing phonetics by computer* [Computer program]. Version 5.4.8, retrieved March 24, 2015 from http://www.praat.org.

[4] Bolinger, D. 1986. *Intonation and Its Parts: Melody in Spoken English*. Stanford: Stanford University Press.

[5] Cooper, W. E., Eady, S. J. 1986. Metrical phonology in speech production. *Journal of Memory and Language* 25, 369-384.

[6] Grabe, E., Warren, P. 1995. Stress Shift: Do speakers do it or listeners hear it? In: Connell, B. (ed), *Laboratory Phonology IV. Phonology and Phonetic Evidence*. Oxford: Oxford University Press, 95-110.

[7] Grabe, E., Warren, P., Nolan, F. 1994. Resolving category ambiguities - Evidence from stress shift. *Speech Communication* 15, 101-114.

[8] Gussenhoven, C. 1991. The English rhythm rule as an accent deletion rule. *Phonology* 8, 1-35.

[9] Gussenhoven, C. 1992. Sentence accents and argument structure. In: Roca, I. (ed), *Thematic Structure: Its Role in Grammar*. Berlin, New York: Foris Publications, 91-106.

[10] Hayes, B. 1984. The phonology of Rhythm in English. *Linguistic Inquiry* 15, 33-74.

[11] Horne, M. 1990. Empirical evidence for a deletion formualtion of the rhythm rule in English. *Linguistics* 28, 959-981.

[12] Liberman, M. 1975. *The intonational system of English. Distributed by Indiana University Linguistics Club*. MIT.

[13] Liberman, M., Prince, A. 1977. On stress and linguistic rhythm. *Linguistic Inquiry* 8, 249-336.

[14] Selkirk, E. 1984. *Phonology and syntax: the relation between sound and structure*. Cambridge, Mass: MIT Press.

[15] Shattuck-Hufnagel, S. 1991. Acoustic correlates of stress shift. *Proceedings International Congress of the Phonetic Sciences*. 266-269.

[16] Shattuck-Hufnagel, S. 2000. Phrase-level phonology in speech production planning: evidence for the role of prosodic structure. In: Horne, M. (ed), *Prosody: Theory and Experiment*: Kluwer Academic Publishers, 201-229.

[17] Shattuck-Hufnagel, S., Ostendorf, M., Ross, K. 1994. Stress shift and early pitch accent placement in lexical items in American English. *Journal of Phonetics* 22, 357-388.

[18] Shattuck-Hufnagel, S., Turk, A. 1996. A prosody tutorial for investigators of auditory sentence processing. *Journal of Psycholinguistic Research* 25(2), 193–247.

[19] Turk, A. 2010. Does prosodic constituency signal relative predictability? A Smooth Signal Redundancy hypothesis. *Journal of Laboratory Phonology* 1, 227-262.

[20] Wightman, C. W., Shattuck-Hufnagel, S., Ostendorf, M., Price, P. J. 1992. Segmental durations in the vicinity of prosodic phrase boundaries. *Journal of the Acoustical Society of America* 91(3), 1707-1717.

ACKNOWLEDGEMENTS

Many thanks to all participants and to Alisdair Tullo for help with *Praat* scripting. The current research has been supported by a University of Edinburgh Research Support Grant for participant payment.