

Accent-free prosodic phrases? Accents and phrasing in the post-nuclear domain

Elisabeth Norcliffe, T. Florian Jaeger

Department of Linguistics
Stanford University, USA

enorcliffe@stanford.edu

tiflo@stanford.edu

Abstract

Some theories of English intonation, e.g. [1], [2], posit a crucial connection between pitch accent placement and the identification of prosodic structure, stating that every prosodic phrase must contain at least one pitch accent. We provide new evidence from the empirical study of prosodic phrasing in the post-nuclear domain to show that this view cannot be maintained. Our experimental study, based on the prosodic disambiguation of verb-particle vs. verb-preposition readings in the post-nuclear domain, revealed significant length based prosodic differences between the two readings. This argues that the post-nuclear domain is divided into more than one prosodic phrase, the second of which doesn't bear any pitch accent.

1. Introduction

In Autosegmental and Metrical theoretical frameworks [1], [2], the prosody of an utterance is treated as a hierarchical structure organized into phonologically determined constituents and heads. Within this general framework, many prevalent theories of English intonation include some form of the following pitch accent requirement: prosodic units above the level of the prosodic word contain at least one pitch accent [1], [2], [3]. We shall henceforth refer to these theories as P(itch)A(ccent)R(equirement) theories. In the present study we provide new empirical evidence from the post nuclear domain which challenges PAR.

The remainder of this paper is organized as follows: §2 presents an overview of PAR theories of English intonation and then turns specifically to the post nuclear domain in order to illustrate the problems that arise if we try to extend the claims of PAR theories to this domain. §3 presents our experiment, which tests the predictions of PAR theories. On the basis of our results we conclude that current PAR theories are unable to account for the observed data. In §4 we close by discussing several alternatives.

2. The Pitch Accent Requirement

In prevalent theories of prosodic phrasing, the composition of English prosodic structure (phrasing) is intrinsically tied to pitch accent placement. Pitch accent may be defined as a tone that is contrastively localized to a syllable [2]. That is, a pitch accent is associated with some designated syllable, and there may be a minimal contrast between utterances in terms of which syllable this is. In English, the pitch accent associates to a stress-prominent syllable in a word. Its placement in English often has to do with information focus, that is, it encodes prag-

matic meaning connected with the discourse context and its relationship to the semantic value associated with the focused word or phrase [2], [4]. Other factors influencing pitch accent placement involve givenness of information, or default placement in neutral utterances.

According to PAR theories, it is a requirement of English that “there must be at least one pitch accent somewhere in every (prosodic) phrase” [2]. Two levels of prosodic phrasing are identified above that of the prosodic word c.f. [1], [2], [3]. The lower level is termed in [1] and [2] the ‘intermediate phrase’, which is marked by the distribution of tones (H-, L-) defining its right edge. These tones are not restricted purely to the phrase boundary, but rather spread across the space following the obligatory or last pitch accent. The highest level prosodic constituent for English is the ‘intonational phrase’. It consists of at least one intermediate phrase, and is defined according to the distribution of H% and L% boundary tones, at this level strictly localized to the phrase boundary.

Thus, it is the placement of pitch accent that structurally defines for PAR theories the two qualitative levels of prosodic structure identified as existing above the minimal level of the prosodic word. Because of the constraint that every intermediate phrase must have at least one pitch accent, the pitch accent is therefore treated as crucial in the definition of English prosodic structure.

2.1. Predictions of PAR theories: the post-nuclear domain

The term ‘nuclear accent’ refers to the strongest, and (usually) the last pitch accent in a prosodic phrase, as indicated by the capitalized BOOK in (1).

- (1) A1: What did John do?
A2: John gave me a BOOK.

Nuclear accent can mark focus or corrective contrasts. When a nuclear accent falls on a word early in a phrase because focus or contrast is associated with it, any following words are de-accented. This stretch of de-accented material is called the post-nuclear domain. Consider (2), where the focus falling on ‘John’ leaves all subsequent material in the phrase de-accented (indicated by underlining).

- (2) A1: Did Andrew give you the book?
A2: No, JOHN gave me the book.

Although we are unaware of any specific PAR theories that explicitly address the possibility of prosodic phrasing in the post-nuclear domain (see [5] for related issues), it follows from the PAR that there should be no post-nuclear prosodic breaks. This

would entail a prosodic phrase without a pitch accent. This prediction did not conform to our intuitions, and thus provided the motivation for the present production experiment.

3. The experiment

The purpose of our experiment was to gather production data to test whether, contra PAR predictions, prosodic phrasing can occur in the post-nuclear domain. We took syntactically ambiguous stimuli known to be disambiguated by prosodic phrasing via pre-boundary lengthening (a.k.a. phrase final lengthening). Unlike earlier studies, however, we elicited stimuli in such a way that the locus of ambiguity fell in the post-nuclear domain, thereby creating a potential conflict between the need to create prosodic phrases to disambiguate on the one hand, and the PAR, on the other.

3.1. Disambiguating syntax via prosodic phrasing

Various prosodic effects over and above that of pitch accent placement have been identified as correlates of prosodic phrasing, one of the most reliable being pre-boundary lengthening. Pre-boundary lengthening corresponds to the duration of the final (stressed) rime of a prosodic phrase being longer than the duration of the same rime when it does not occur phrase-finally.

Various studies have shown via durational effects that there is a correlation between prosodic phrasing and syntactic phrase structure. [6], for example, investigated duration at syntactic boundaries, finding significant lengthening at different types of syntactic junctures. [7], in a series of disambiguation experiments, investigated the ability of listeners to disambiguate surface structurally identical sentences for different types of syntactic structures on the basis of prosodic cues, principally duration. They found, for example, that left vs. right attachments (as in (3)), were reliably disambiguated by means of prosodic phrasing.

- (3) (a) In spring there was always more work to do on the farm. May was the hardest month. *They rose early in May.*
(b) Bears sleep all winter long, usually coming out of hibernation in late April, but this year they were a little slow. *They rose early in May.*

For the purposes of our experiment nothing hinged on whether such syntax-prosody correlations involve a direct mapping between these two domains (for discussion, see e.g. [8]), and we do not touch on this question here. Our aim was rather to utilize the fact that such a correlation exists, in order to elicit prosodic cues to prosodic phrasing. We did this by adopting a structural ambiguity, that of preposition vs. particle readings. Examples (4) and (5) illustrate this ambiguity type, based on similar examples used in [7]. In both examples, speaker A1 provides disambiguating discourse. A2 is identical for both examples in terms of the surface string. Note, however, that the readings (as brought about by the preceding disambiguating sentences in A1 in each), and concomitantly the underlying syntactic phrasings, are distinct. In (4), A2 is the Verb-Preposition reading, in which the verb 'win' combines with a PP (headed by 'o! ver').

- (4) A1: Heartless violence led to a bloody victory.
A2: The Vikings won] over their enemies.
(5) A1: Gentle persuasion led to a friendly settlement.
A2: The Vikings won over] their enemies

[7] showed that pre-boundary lengthening occurs as a prosodic cue disambiguating the two readings. That is, in (4), the verb 'won' is in pre-boundary position and correspondingly is subject to pre-boundary lengthening. In (5), by contrast, the particle 'over' is in pre-boundary position, and so for this reading it is now this element which is subject to pre-boundary lengthening.

Unlike earlier disambiguation studies, we elicited targets so that a nuclear accent fell before the locus of ambiguity, as in (6).

- (6) A1: [disambiguating discourse]
B1: So the Romans won over their enemies?
A2: No, the VIKINGS won over their enemies.

The contrastive focus on the word VIKINGS brings the nuclear accent onto this word, such that all material following the focused word, including the locus of ambiguity, is de-accented. By this method, the claim of PAR theories that prosodic phrasing is determined by the presence of a pitch accent could be directly tested.

3.2. Procedure and subjects

For each target sentence, the subject was asked to read a disambiguating context which contained the target sentence without contrastive focus (A1 in 6). A confederate then asked a question related to the agent of the sentence (B1), which required a corrective answer from the subject (A2). This corrective answer elicited focus on the agent, in order to obtain a de-accented post nuclear stretch on the remainder of the sentence.

The speech material was read from counterbalanced lists by four graduate students at Stanford university, one male and three females. The subjects were asked to familiarize themselves with each passage before reading it. The recordings were undertaken in a sound attenuated room and made onto DAT. They were then downsampled to 11025 kHz, and transferred to a Sun workstation for acoustic analysis using XWAV.

The four participants read six target sentences for each condition (the two possible readings), producing 48 target sentences. For each target sentence, vowel duration measurements were taken for the rime of the final syllable of the verb, and for the rime of the preposition/particle. A third measure was derived from the duration of the verb rime and the particle/preposition rime by dividing the verb rime by particle/preposition rime. This was done to test for relative phrase boundary strength since it has been suggested that the strength of a phrase boundary is not marked absolutely, but depends on the strength of the surrounding phrase boundaries, e.g. [9], [10]. That is, phrase final lengthening of a rime X in reading A may not show up as an absolute increase in duration of X compared to X' in reading B but rather in the *relative* length of X given the length of surrounding rimes in A. Using the ratio of the verb and the particle/preposition rime lengths constitutes such a measure of the relative length of each of the rimes.

For each of the three measures, we conducted a repeated measures analysis of variance with SYNTAX/READING (Verb-Particle vs. Verb-Preposition) as the only factor. Both subject (F1) and item (F2) analyses were performed.

3.3. Predictions

If prosodic phrasing takes place in the post-nuclear domain, then we predicted that pre-boundary lengthening should cause significant lengthening of the verb rime in the Verb-Preposition reading, and of the particle rime in the Verb-Particle reading.

If speakers do not consistently mark phrase boundaries by absolute lengthening, but rather use relative pre-boundary lengthening, the ratio of the verb rime's duration and the particle/preposition rime's duration should be significantly higher in the Verb-Preposition reading (since the verb should be relatively long and the preposition relatively short). In contrast, PAR theories would predict there to be no pre-boundary lengthening unless there is a second pitch accent following the contrastive accent. We return to this issue in §4.

3.4. Results

We found the following differences between the two readings:

- The verb rime was longer in the Verb-Preposition reading, with a mean difference of 18ms ($F(1,3)=3.0$, $p=0.18$; $F(1,5)=4.1$, $p=0.1$).
- The verb particle was significantly longer in the Verb-Particle reading, with a mean difference of 28ms ($F(1,3)=15.5$, $p=0.03$; $F(1,5)=6.5$, $p=0.05$).
- The ratio of the two rime lengths differed significantly between the two readings ($F(1,3)=22.7$, $p=0.02$; $F(1,5)=13.5$, $p=.01$)

3.4.1. Secondary accents in the post-nuclear stretch

Taking an overly conservative approach, we also checked each target sentence for any possible presence of a pitch accent. Despite the presence of the nuclear accent in each, which should, according to PAR theories, have ruled out the presence of any further pitch accent in the remainder of the phrase (see [5] on post-nuclear prominences), some of our data did contain secondary pitch accents in the post nuclear stretch (around 27% of the sample). We assume this to be partially due to reader disfluencies, though it may be correlated with what has been observed for post nuclear stretches elsewhere, in, for example, European Portuguese and Maltese ([11], [12] respectively), namely the presence of a post-focal accent, characterized by a compressed pitch range. We cannot address these post-nuclear phenomena here, save to observe that for the immediate purposes of our study, our results may have been influenced by the presence of secondary pitch accents. We therefore removed all accented examples from our sample, and also any non-accented examples whose pair was accented. This reduced the sample size by 50%.

3.4.2. Results for the reduced sample size

With the accented material removed, the same trends were nevertheless observed. (Significance predictably decreased, due to the reduced sample size ($N=24$):

- The ratio of the verb rime and the verb particle lengths still strongly differed between the two readings, with a mean difference of 58ms ($F(1,2)=6.1$, $p=0.13$; $F(1,4)=9.3$, $p=.04$)

As in the full data set, this difference was mainly due to the absolute lengthening of the particle in the Verb-Particle reading.

3.4.3. Summary of results

The results lend support for the presence of prosodic phrasing in the post-nuclear domain. While lengthening of the verb rime in the Verb-preposition reading only *approached* significance, lengthening of the particle in the Verb-Particle reading

was much stronger. Moreover, although not conclusive, the significant effect on the ratio of the two verb rime is compatible with the hypothesis that lengthening is a relative phenomenon.

In the reduced sample, we still observed the same trend of pre-boundary lengthening in the post-nuclear domain, despite the absence of any type of pitch accent. We currently have no explanation why the verb lengthening is less significant than the particle/preposition lengthening. We plan to run more subjects to see whether this is an irregularity of our current data set or a stable difference.

These results support the notion that the post-nuclear domain is divided into more than one prosodic phrase, the second of which does not bear any pitch accent.

4. Discussion

This study both builds on the results of previous work which has pointed to a relationship between prosodic marking and syntactic phrase boundaries in non-focussed sentences and, significantly, provides some counter evidence to the claim of PAR theories that the identification of prosodic units above that of the prosodic word crucially relies on the presence of pitch accents.

In light of this second point, our findings therefore call for either a revision of PAR theories of English intonation, or otherwise recommend the development of alternative theories which do not rely solely on pitch accent in determining the constituency of English prosody. It is beyond the scope of this paper to adequately address the possibilities regarding the most fruitful manner of incorporating the findings of the present study into current theories of English prosodic structure; here we wish merely to suggest briefly some possible ways in which current theories might accommodate our results.

4.1. Possible revisions to existing PAR theories

As it stands, the prosodic level immediately above that of the prosodic word, the intermediate phrase, is defined by PAR theories with crucial reference to the presence of a pitch accent. There are no qualitatively marked prosodic constituent edges postulated between the intermediate phrase and the prosodic word, as [2] observes.

If we wish to allow for the possibility of prosodic phrasing despite the absence of any pitch accent, a possibility supported by the present study, then we suggest the following alternatives. One could either revise the definition of the intermediate phrase to allow for accent-less intermediate phrases, or, alternatively, the prosodic hierarchy could be enhanced by a new prosodic unit above the prosodic word and below the intermediate phrase. This second alternative would perhaps be en par with the accental phrase proposed for Japanese [1]. Japanese has lexically unaccented words, and, consequently, well formed utterances are possible without any pitch accents. [1] claim that this would be impossible in the English intonation system. If, as our study suggests, prosodic phrasing can occur in the absence of a pitch accent, then this may turn out to be a viable enhancement of English prosodic constituency structure.

4.2. Alternative theories

The observations made in this paper can be accounted for within theories of English intonation that loosen the PAR, e.g [13], [14], [10], [15]. As an illustration, we present an analysis in the most recent framework [10], [15], but, for the current purpose, earlier work on prosodic grids and recursive prosodic phrasing would make the same predictions (e.g. [13], [14]).

For Wagner [10], syntactic sisters are prosodically ‘matched’ whenever the functor precedes its argument. This ‘Sister Matching’ causes each sister to be mapped to a foot at the next higher grid line and all top line grid marks within each sister project up. Consider the construction of the verb phrase for the two different readings in (4) and (5). We assume that the particle first forms a constituent with the verb and then takes the NP complement, while the preposition first combines with its argument NP before the resulting PP combines with the verb. The predicted prosodic grids for the VP are given in (7) and (8) respectively, where foot structure/phrasing is indicated by ‘(|)’. The phonetic strength of a foot boundary increases with each grid line. Feet on higher grid lines have stronger boundaries than feet on lower grid lines. The resulting ‘Boundary Rank’, that is the relative strength of a boundary, is summarized below the grid.

(7) $\begin{array}{cccc|} | & x & | & x & x & | \\ | & x & | & x & | & x & | \\ 2 & & 2 & & 1 & & 2 \end{array}$ (Boundary Rank)
won over their enemies

(8) $\begin{array}{cccc|} | & x & x & | & x & | \\ | & x & | & x & | & x & | \\ 2 & & 1 & & 2 & & 2 \end{array}$ (Boundary Rank)
won over their enemies

Since the context in our experiment provides a contrasting alternative to the subject *the Vikings*, Wagner [15] predicts the VP to be subordinated (rather than coordinated) to the subject. In the case of subordination, only the (syntactically) projecting element receives a grid mark in the next higher level of phrasing. This means only the subject will project its grid marks up to the next grid line, and subject and VP together are mapped to a foot, resulting in the following prosodic grids for our examples:

(9) $\begin{array}{cccc|} | & x & & & & | \\ | & x & & | & x & | & x & | \\ | & x & & | & x & | & x & | \\ 3 & & & 2 & 2 & 1 & & 3 \end{array}$
The Vikings won over their enemies

(10) $\begin{array}{cccc|} | & x & & & & | \\ | & x & & | & x & | & x & | \\ | & x & & | & x & | & x & | \\ 3 & & & 2 & 1 & 2 & & 3 \end{array}$
The Vikings won over their enemies

Wagner [10] posits that exactly all top line grid marks are associated with an accent. Accents associated to grid marks that do not project all the way up are at least extremely reduced in pitch range, if not altogether absent (this can be seen as a relaxation of the PAR). Thus, Wagner [10], [15] predicts correctly that, for the phrasings in (9) and (10):

- The subject should be accented
- All following material should be de-accented (or at least drastically reduced in its pitch range)
- Since location and strength of phrase breaks is purely determined by the Boundary Rank prosodic phrasing in the post-nuclear domain is possible. More precisely, since pre-boundary lengthening correlates with the rank of the upcoming boundary:

- In the Verb-Particle reading, (9), the phrase break between the verb and the particle is *weaker* than the break between the particle and the direct object NP.

- In the Verb-Preposition reading, (10), the phrase break between the verb and the preposition is *stronger* than the break between the preposition and its NP.

As discussed, these predictions are met. More generally, the data presented provides support for alternatives to strict PAR-theories.

5. Acknowledgements

Thanks especially to Edward Flemming for the original intuition that it would be worthwhile running this experiment and for helpful feedback throughout. Thanks also to Michael Wagner for discussion, and to our anonymous reviewers.

6. References

- [1] Beckman, M. E. and Pierrehumbert, J. B., “Intonational Structure in Japanese and English”, *Phonology Yearbook*, 3:255–309, 1986.
- [2] Beckman, M.E., “The Parsing of Prosody”, *Language and Cognitive Processes*, 17-67, 1996.
- [3] Silverman, K., Beckman, M., Pitrelli, J., Ostendorf, M., Wightman, C., Price, P., Pierrehumbert, J., and Hirschberg, J. “ToBI: a standard for labelling English prosody”, *Proceedings of ICSLP 92(2)*:867-870, 1992.
- [4] Selkirk, E. “Sentence Prosody: Intonation, Stress, and Phrasing”, *The Handbook of Phonological Theory*, Basil Blackwell, Cambridge and Oxford, 1995.
- [5] Ladd, D. R., “Intonational phonology.” *Cambridge Studies in Linguistics 79*. Cambridge University Press, Cambridge, 1996.
- [6] Klatt, D. H. “Vowel lengthening is syntactically determined in a connected discourse”, *Journal of Phonetics*. 3, 1975.
- [7] Price, P. J., Ostendorf, S., Shattuck-Huffnagel, S., and Fong, “The use of prosody in syntactic disambiguation”, *J. Acoust. Soc. Amer.*, Vol. 9:2956–2970, 1991.
- [8] Nespor, Marina and Irene Vogel. *Prosodic phonology*. Foris, Dordrecht, 1986.
- [9] Carlson, K., Clifton, C. Jr., and Frazier, L. “Prosodic boundaries in adjunct attachment”. *Journal of Memory and Language*, 45:58–81, 2001.
- [10] Wagner, M., “Prosody as diagonalization of syntax. Evidence from complex predicates”. To appear in K. Moulton and M. Wolf (eds): *Proceeding of NELS 34*, SUNy Stony Brook, 2003. 2004.
- [11] Frota, Sónia. *Prosody and focus in European Portuguese*. Garland, New York and London. 2000
- [12] Vella, Alexandra. *Prosodic structure and intonation in Maltese: influence on Maltese English*. Ph.D. dissertation, Edinburgh University, 1995.
- [13] Gussenhoven, C. and Rietveld, A.C.M. “Intonation contours, prosodic structure, and preboundary lengthening”, *Journal of Phonetics* 20:283-303, 1992.
- [14] Liberman, M. Y., *The Intonational System of English*. Ph.D. thesis, MIT, 1975 (published by Garland Press, NY, 1979).
- [15] Wagner, M. PhD thesis (*Prosodic Phrasing*), forthcoming.