LEVELS OF THE PROSODIC HIERARCHY IN ENGLISH

Molly Shilman

UCLA Department of Linguistics
mshilman@humanities.ucla.edu

ABSTRACT

A relationship between prosody on the one hand and syntax and semantics on the other is generally acknowledged. Prosody is also influenced by other factors, both linguistic and extra-linguistic. Previous studies have found support for either a mapping between syntactic and prosodic structure or the use of a listener-directed strategy to avoid ambiguity. The current study supports the view that syntactic structure is cued by prosody even in the absence of ambiguity. It also reports evidence of a prosodic phrase in English, called an Accentual Phrase, that is larger than a word and smaller than an intermediate phrase.

Keywords: Prosody, English, syntax-prosody mapping, prosodic hierarchy, Accentual Phrase

1. INTRODUCTION

Starting with work by Selkirk (e.g. [9]), linguists' study of prosody has made extensive use of tree structures. Independent of the particulars of formalism, some relationship between prosodic tree structures on the one hand and syntactic and semantic tree structures on the other is widely acknowledged. For example, some syntactic constructions, such as lists, relative clauses, and parentheticals, have characteristic intonation contours. Apart from such cases, our understanding of how syntactic structure affects prosody is more limited: we know speakers are capable of using prosody for disambiguation (see, e.g., [9], [12], [14]), but we know considerably less about the syntax-prosody relation when disambiguation is not required. Some researchers have found that prosody reflects syntactic structure even outside of disambiguation contexts, as in [9], but conflicting results have also been found, as in [12]. In both of these studies, the syntactic structure found to influence prosody is the site of PP attachment.

In general, studies that show disambiguation by prosody for a larger range of constructions, such as [14], indicate that the relative degree of syntactic juncture may be mirrored by the strength of a corresponding prosodic boundary. This seems intuitive and is roughly consistent with proposals about the syntax-prosody interface such as those in [4], [6], and [12]. However, taken as a generalization, this would lead us to expect more degrees of prosodic juncture than the relatively sparse prosodic constituent inventory employed in the Intonational Phonology model of English proposed in [3] and in the Mainstream American English Tone and Break Index transcription system (MAE_ToBI) of [1], which use only the word, intermediate phrase (ip), and intonational phrase (IP). The current study looks for prosodic differences between three syntactic constructions: one that induces an ip boundary and two others—of different strengths in the syntax—that tend to correspond to only word-level boundaries in MAE_ToBI labeling.

The use of unambiguous sentences and control of the position of pitch accents allow for any differences in prosodic phrasing to be attributed to properties of the syntax-phonology interface rather than to a potentially extra-linguistic strategy for avoiding ambiguity. This paper reports the results of the first experiment in an ongoing series; the results reported here are consistent with at least one additional level of prosodic structure between the word and the intermediate phrase. Following [3] and [7], this prosodic unit will be called an Accentual Phrase (AP). The nature and purpose of follow-up work in progress is described at the end of the discussion section.

2. THE EXPERIMENT

A production experiment designed to test for phonetic evidence of the existence of a prosodic boundary smaller than an intermediate phrase and larger than a word in English was conducted. In order to induce the production of different prosodic boundary strengths, the syntactic structure of the target sequence was varied as explained below. The phonetic measure used was final lengthening, the tendency of phrase-final material to have a longer duration than the same material has in non-final position, as reported in [2] and [14].

2.1. Experimental design

Script items in this experiment contained two-word sequences in which the two words were predicted to be separated by three different levels of boundary: word, Accentual Phrase (AP), and ip. (See section 3 for discussion of the possible relationship between AP and pitch accent.)

Three different syntactic constructions were selected to induce the three different levels of boundary. To determine which syntactic structures were likely to cause production of a word boundary versus an AP boundary, several proposals about the mapping of syntax to prosody were consulted; see [4], [6], [12]. All of these proposals predict that an Adjective-Noun (Adj-N) sequence should be capable of belonging to a single AP and a Subject-Verb (NP-VP) sequence should not be. To induce production of an ip boundary, the juncture between a restrictive relative clause and a verb (RC-VP) was used; this is not based on the prediction of any specific theory, but rather on observation.

Because different segments have inherently different durations, it was necessary to hold the segmental string in the target region (that is, around the boundary) constant. In order to allow this, category-changing homophones were used. The first word in the target sequence (WORD1) was a noun-adjective homophone with three syllables and antepenultimate stress ending in a voiced continuant (*fugitive, natural, radical*) and the second word (WORD2) was a noun-verb homophone with initial stress starting with a voiceless stop (*pilot, target, partner*). This allowed for a clean division between the two words and kept the stress pattern constant across different groups of items.

In order to avoid a potential confound from the effect of lengthening under pitch accent found in [5], the position of pitch accent on the target words was also controlled. To elicit a consistent pitch accent pattern across sentence types and speakers, script items were two sentences long, with the first sentence (context sentence) providing background information for the second sentence (target sentence). Speakers read both the context sentence and the target sentence aloud. The context sentence encouraged the production of pitch accent on WORD1 and no accent on WORD2, as in the pair (1) and (2), which made up a single script item in the experiment. The use of *partners* in (1) discourages accenting of this word in the target sequence *radical partner* in (2), while the contrast between *traditional* and *radical* encourages placement of pitch accent on *radical* in (2).

(1) Before the last meeting, only the more traditional of the partners was in favor of the firm buying the new subsidiary.

(2) After the last meeting, the radical partner supported the responsible purchase, as well.

In addition to the factors already discussed, the total length of each target sentence in syllables and the starting position of the two-word target string were kept constant. Table 1 shows the target sentences from one of the three sets of script items that were used in the experiment.

Table 1: One of the three sets of script items, two-word target sequence underlined

Break	Example
Word	After the last meeting, the radical partner supported the responsible purchase, as well.
AP	After the last meeting, the radical partnered with the lead conservative on some key issues.
ip	Someone who was known as a radical partnered with the lead conservative on some key issues.

2.2. Subjects and procedure

Seven native speakers of American English, all in their 20s, participated in the experiment. Although most subjects had studied linguistics, none had studied prosody. All subjects were unpaid volunteers.

Subjects were digitally recorded in the sound booth in the UCLA Phonetics Lab (22.1 kHz, 16 bits). Subjects were given booklets with one item per page and asked to read each page aloud as naturally as possible. Real target items alternated with fillers of comparable length. Items were given in a pseudo-random order. The order was not held constant across subjects. Each subject produced three repetitions of each target sentence.

Despite the presence of the context sentence, not all utterances were produced with the desired accent pattern. Therefore, tokens were screened by accent pattern (following MAE_ToBI conventions as in [1]) and those that were not produced with the desired accent pattern were eliminated; for all of the subjects whose results are included here, there was at least one token of each target sentence produced with the desired accent pattern. (The RC-VP condition proved to be an exception, as subjects were very resistant to de-accenting WORD2 in these sentences.) Measurements from

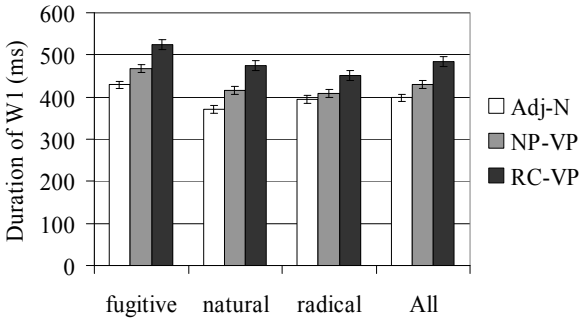
the remaining tokens were averaged, resulting in a single value of each measure for each target item from each speaker.

The measurements taken from these recordings were of the length of the entire WORD1 and the length of the end of WORD1, called END. The END of WORD1 began at the last reliable segmentation point in the word; thus, the end of *fugitive* was [ɪv], the end of *natural* [ɹɨ], and the end of *radical* [ɨ].

2.3. Results

Overall, the results show a pattern of increasing final lengthening as predicted: both of the duration measurements are smallest for the Adj-N items, in the middle for the NP-VP items, and largest for the RC-VP items. The measurements for duration of WORD1 are shown in Figure 1.

Figure 1: Duration of WORD1 by lexical item and overall, bars are standard error

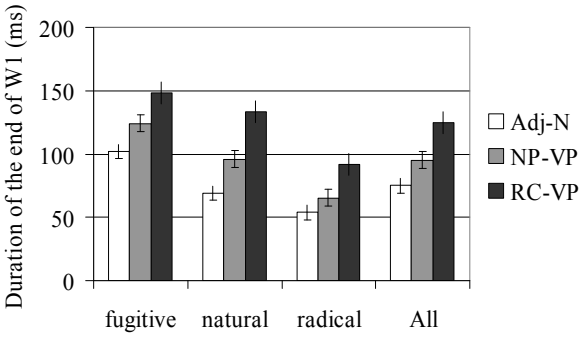


A two-factor Repeated Measures ANOVA (syntactic construction by word, $\alpha = 0.05$) with Huynh-Feldt correction was run on these results; both effects were significant, as shown in Table 2**.** The effect of word reflects the inherently different durations of the lexical items used, and the effect of construction indicates that a difference in lengthening is tied to syntactic construction. The interaction of construction and word reflects the different behavior of *radical* as compared to the other words.

Paired t-tests ($\alpha = 0.025$) show that a true three-way distinction between each pair of constructions exists. For the Adj-N vs. NP-VP comparison, $t(20) = 4.8$, $p < 0.01$; for the NP-VP vs. RC-VP comparison, $t(20) = 9.3$, $p < 0.01$.

As mentioned above, the measure END shows the same pattern, as shown in Figure 2.

Figure 2: Duration of END by lexical item and overall, bars are standard error



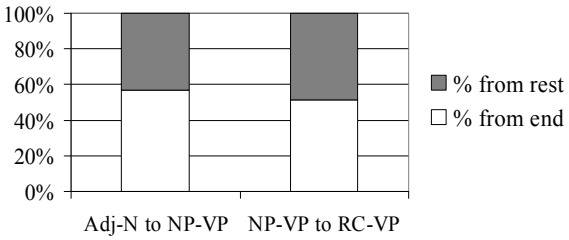
The same statistical tests were conducted on these results as on the WORD1 measurements, and the results were similar: the RM-ANOVA (in Table 2) shows main effects of both construction and word, although the interaction is not significant (i.e. all words behave similarly). The paired t-tests again support a three-way distinction in length. For Adj-N vs. NP-VP, $t(20) = 4.8$, $p < 0.01$; for NP-VP vs. RC-VP, $t(20) = 6.9$, $p < 0.01$.

Table 2: RM-ANOVA results for both measures

	WORD1	**END**
Construx.	F(2, 12)=61.0, p<0.01	F(1.4, 8.6)=31.6, p<0.01
Word	F(1.6, 9.8)=23.8, p<0.01	F(2, 12)=40.2, p<0.01
Interaction	F(3.5, 21.1)=3.6, p=0.02	F(3.6, 21.4)=2.4, p>0.09

To understand whether lengthening of WORD1 is truly final (i.e. concentrated at the end of the word), the percentage of this lengthening coming from END was calculated. As shown in Figure 3, more than half of the lengthening of WORD1 comes from END—lengthening is localized at the end of the word rather than evenly distributed or concentrated on the main-stress syllable. This is consistent with final lengthening as in [14] rather than with another type of lengthening, such as the lengthening under accent reported in [5].

Figure 3: Percentage of WORD1 lengthening contributed by END



3. DISCUSSION

The results are consistent with the presence of an AP in English: the degree of final lengthening at the

end of the syntactic subject falls between the lengthening at the right word edge and that at the right ip edge. This suggests the presence of a boundary greater than a word but smaller than an ip.

To further support the proposal that the degree of break between WORD1 and WORD2 in the NP-VP condition is a distinct level of boundary, an additional pair of tests was conducted on the measurements discussed above. These tests verified that the lengthening found in the NP-VP condition is truly intermediate in degree between the word and ip boundaries rather than representing a mixture of tokens produced with word-boundary lengthening and tokens produced with ip-boundary lengthening. If the NP-VP items were produced with a mixture of different boundary types, the length of WORD1 (and END) in this condition should cover the full range of values occupied by measurements from both of the other conditions, and the standard deviation (SD) of the NP-VP tokens should reflect this broad distribution. Monte Carlo tests were run on both sets of measurements to test this hypothesis.

For each test, 10,000 groups of values, balanced for the number of tokens actually contributed by each speaker, were randomly chosen from the Adj-N and RC-VP sets and the SD of each group was calculated. This provided a distribution of SDs against which the actual SD of the NP-VP tokens was compared. The probability (p) of obtaining the SD of the actual NP-VP tokens through mixed production of these tokens was calculated by dividing the number of groups from the simulation that had SD $\leq$ the SD of the real NP-VP tokens by the total number of groups (10,000). For both WORD1 and END, p was well below 0.05 (WORD1: $p = 0.01$, END: $p < 0.01$). This suggests that the degree of lengthening observed in the NP-VP tokens belongs to its own category and is not the result of variation in production of these tokens.

As alluded to in section 2, the additional level of phrasing proposed here (the AP) is in just the position in the prosodic hierarchy to be headed by the pre-nuclear pitch accent, which is a head-type prominence without a phrase to head in the MAE_ToBI labeling system. A follow-up experiment with the same target sentences but with pitch accent on both words in the target sentence (*fugitive pilot, natural target, radical partner*) is in progress. Comparing the results of these two experiments will allow us to see whether the pattern of lengthening shown here is dependent not only on the syntax, but also on the pitch accent pattern. This is the type of relation between the proposed AP and the pre-nuclear pitch accent that the patterns of other languages (see [7]) would lead us to expect in English, and positing such a relation makes a clear prediction: if the AP found here is headed by pitch accent, there should be such a boundary between WORD1 and WORD2 in both the Adj-N and NP-VP conditions when both words are pitch accented.

4. REFERENCES

[1] Beckman, M., Ayers-Elam, G. 1997. Guidelines for ToBI Labeling. Ms.

[2] Beckman, M., Edwards, J. 1990. Lengthening and shortening and the nature of prosodic constituency. *Proc. Lab Phon I* Ohio, 152-178.

[3] Beckman, M., Pierrehumbert, J. 1986. Intonational Structure in Japanese and English. *Phonology Yearbook* 3, 15-70.

[4] Büring, D., Gutiérrez-Bravo, R. 2001. Focus-related word order variation without the NSR: A prosody-based crosslinguistic analysis. *Proc. Syntax at Santa Cruz 3*, 41-58.

[5] Cambier-Langeveld, T., Turk, A. 1999. A cross-linguistic study of accentual lengthening: Dutch vs. English. *J. Phon.* 27, 171-206.

[6] Jacobs, J. 1991. Neutral stress and the position of heads. *Informationsstruktur und Grammatik*, Opladen: Westdeutscher Verlag, 220-244.

[7] Jun, S-A. (ed.) 2005. *Prosodic Typology: The Phonology of Intonation and Phrasing*. Oxford University Press, Oxford.

[8] Pierrehumbert, J. 1980. The phonology and phonetics of English intonation. Ph.D. diss., Department of Linguistics and Philosophy, MIT.

[9] Schafer, A., Speer, S., Warren P. 2005. Prosodic influences on the production and comprehension of syntactic ambiguity in a game-based conversation task. In: Tanenhaus, M., Trueswell, J. (eds.), *Approaches to Studying World Situated Language Use: Psycholinguistic, Linguistic and Computational Perspectives on Bridging the Product and Action Tradition*. Cambridge: MIT Press.

[10] Selkirk, E. 1978. On Prosodic Structure and its Relation to Syntactic Structure. Presented at the Sloan Workshop of the Mental Representation of Phonology, University of Massachusetts.

[11] Shattuck-Hufnagel, S., Turk, A. 1996. A prosody tutorial for investigators of auditory sentence processing. *J. Psycholinguistic Research* 25, 193-247.

[12] Snedecker, J, Trueswell, J. 2003. Using prosody to avoid ambiguity: Effects of speaker awareness and referential context. *J. Memory and Language* 48, 103-130.

[13] Truckenbrodt, H. 1999. On the relation between syntactic phrases and phonological phrases. *Linguistic Inquiry* 30, 219-255.

[14] Wightman, C., Shattuck-Hufnagel, S., Ostendorf, M., Price, P. 1992. Segmental durations in the vicinity of prosodic phrase boundaries. *J. Acoust. Soc. Am* 91, 1707-1717.