Published in final edited form as:
Lingua. 2009 July 1; 119(7): 1066–1082. doi:10.1016/j.lingua.2008.11.003.

How prosody constrains comprehension: A limited effect of prosodic packaging

Katy Carlson[a], **Lyn Frazier**[b], and **Charles Clifton Jr.**[c]

[a]419 Combs, Morehead State University, Morehead, KY 40371 USA k.carlson@morehead-st.edu, +1-606-783-2782 (tel), +1-606-783-9112 (fax)

[b]Department of Linguistics, University of Massachusetts, Amherst, MA 01003 USA lyn@linguist.umass.edu

[c]Department of Psychology, University of Massachusetts, Amherst, MA 01003 USA cec@psych.umass.edu

Abstract

Prosody has a large impact on language processing. We contrast two views of how prosody and intonation might exert their effects. On a 'prosodic packaging' approach, prosodic boundaries structure the linguistic input into perceptual and memory units, with the consequence that material in earlier packages is less accessible for linguistic processing than material in the current package. This approach claims that such lessened accessibility holds true for the comprehension of all constructions, regardless of the particular kind of linguistic dependency that needs to be established using the earlier constituent. A 'specialized role' approach, by contrast, attributes to prosodic boundaries a role in making grouping decisions when building hierarchical structure, but attributes to pitch accents the major role in determining the accessibility of a constituent. The results of four listening studies with replacive sentences (Diane thought Patrick was entertaining, not Louise) support the predictions of the specialized role hypothesis over the prosodic packaging approach.

1. Introduction

The way that a sentence is spoken matters. The presence or absence of prosodic boundaries at particular locations in a sentence plays an important role in sentence comprehension (Carroll & Slowiaczek, 1987; Kjelgaard & Speer, 1999; Lehiste, 1973; Price, Ostendorf, Shattuck-Hufnagel & Fong, 1991; Pynte & Prieur, 1996; Schafer, Speer, Warren, & White, 2000; Steinhauer, Alter, & Friederici, 1999; Warren, Grabe, & Nolan, 1995). The basic finding is that a phrase is more often taken to modify, or otherwise be in construction with, an earlier phrase if the two are not separated by a prosodic boundary than if they are so separated. Further, the presence and location of pitch accents has been shown to affect sentence comprehension (Carlson, 2001, 2002; Frazier & Clifton, 1998; Schafer, Carlson, Clifton, & Frazier, 2000; Schafer, Carter, Clifton, & Frazier, 1996; Weber, Grice & Crocker, 2006; Welby, 2003). The pitch accent can signal that a phrase is new or contrastive or otherwise focused, or provide information about the syntactic role of a word or phrase. The present paper explores the effects of prosodic features on how accessible phrases are for one aspect of interpretation, the choice of a contrasting antecedent in a replacive construction. It is part of an effort to understand the mechanisms behind the effects of prosodic boundaries

Correspondence to: Katy Carlson.

and pitch accents and to explore the linguistic environments in which each has an effect on comprehension.

On one view, prosodic structure, especially prosodic phrasing, may impose an organization on the linguistic input. For example, it might demarcate the 'phrasal packages' formed in a parser, along the lines of the Sausage Machine (Frazier and Fodor, 1978). All processing of auditory input might then be influenced by the accessibility relations implicit in prosodic packages. Material in an earlier prosodic package should, on this view, be less accessible than material in a later prosodic package, other things being equal. Very early work in psycholinguistics established that material in an earlier syntactic and prosodic phrase was less accessible for memory-based judgments than material in the current phrase (Jarvella & Herman, 1972; Caplan, 1972) and other work (e.g., Scharpff & van Heuven, 1988) shows effects on intelligibility of extending the duration of a prosodic boundary. Further, if prosodic packages are simply the form in which the linguistic input is held in memory for linguistic interpretation (cf. Speer, Kjelgaard, & Dobroth 1996), then the lesser accessibility of material in earlier prosodic packages than in more recent ones should hold for all types of grammatical dependencies, phrase structures, binding, coreference, ellipsis, etc. Even though accessibility for memory judgments and accessibility for linguistic interpretation may or may not be intimately related, the fact that some evidence suggests that prosodic (and syntactic) boundaries affect both provides some motivation for considering the prosodic packaging hypothesis.

A partially contrasting view might consider the separate influences of two types of suprasegmental properties of the speech signal: those related to prosodic grouping versus prosodic prominence. Prosodic phrasing, which is indicated in English by temporal changes (slowing, pausing) as well as tonal ones (boundary tones), is thought to constrain the syntactic grouping or phrase structure of the sentence (Lehiste, 1973; Price et al., 1991) and hierarchical structure (Hirotani, 2004). Prosodic prominence, conveyed by pitch, duration, and intensity, influences other areas of processing, such as information structure (Schwarzschild, 1999), pronoun resolution (Balogh, 2004; Hirschberg & Ward, 1991), reference resolution (Dahan, Tanenhaus, & Chambers, 2002; Eberhard, Sedivy, Spivey-Knowlton, & Tanenhaus, 1995), ellipsis (Carlson, 2001, 2002; Frazier & Clifton, 1998) and lexical status (e.g., whether who is an interrogative or a relative pronoun: Schafer, Carlson, et al., 2000).

These two views make fundamentally distinct claims about the role of prosodic boundaries. Consider the case of replacive constructions, such as Roger thought Alice was reliable, not Andrew, to be investigated here. The number of prosodic boundaries intervening between a remnant of ellipsis (not Andrew) and its potential antecedents/correlates (Roger or Alice) should matter according to the prosodic packaging view. If one boundary makes the potential antecedent less accessible, additional boundaries should make it even less accessible. However, according to the 'specialized role' view, on which prosodic grouping affects hierarchical structure while prominence affects focus and antecedent choice, the number of boundaries should not matter. Prominence should matter. Below we present three experiments designed to test these predictions.

In recent work, we have argued that the interpretation of an optional boundary before an ambiguously-attached constituent is influenced by the presence and size of earlier prosodic boundaries. A boundary of a given phonological size affects comprehension if it is larger[1] than certain other, earlier, potential boundaries in an utterance within a syntactically relevant

[1]Larger in phonological terms: an Intonational Phrase Boundary (IPh) is larger than an intermediate phrase boundary (ip) in the descriptive system we use (cf. Beckman & Ayers, 1993; Pierrehumbert and Beckman, 1988; Shattuck-Hufnagel & Turk, 1996).

domain. For example, Carlson, Clifton and Frazier (2001) showed that the impact of a boundary before the PP after the party depended on whether there was an earlier boundary after learned that was larger than the boundary before the PP.

(1) John learned that Bill telephoned after the party.

If an early boundary immediately after learned was larger, then low attachment of the after-phrase was favored (so that it modified telephoned, not learned). If the boundary immediately preceding the PP was larger, then more high attachments were observed. Clifton, Carlson and Frazier (2002) extended these findings to a variety of other syntactic ambiguities involving conjunction (old men and women with large houses), larger or smaller possessive phrases (John and Sheila's in-laws), low versus high attachment of a head noun (the daughter of the Pharaoh's son) and ambiguous relative clause attachment (the daughter of the colonel who was on the balcony). The observed effect of adding the early IPh boundary was always significant and quite substantial, but it did vary among the different constructions. For instance, adding the boundary decreased the frequency of high attachment choices by an average of 14% (from 31% to 17%) in Experiment 3 of Clifton et al. (2002), and between 14% and 33% in the three different constructions of Experiment 2 of that paper.

A version of the prosodic packaging view consistent with our earlier data is that a prosodic boundary, especially one that is phonologically larger than earlier relevant boundaries, serves to make material that precedes it less accessible to the processor when analyzing material that follows the boundary (see Schafer, 1997; Schafer & Speer, 1998, for interesting ideas and data concerning the availability of material inside or outside of the current prosodic phrase). We extend this hypothesis to propose that constituents that are two prosodic phrases back should be less accessible for linguistic interpretation than constituents that are one prosodic phrase back. Experiments 1 (and its replication), 2a, and 2b test this hypothesis. Experiment 3 tests the specialized role hypothesis, according to which prominence information, but not prosodic boundaries, should influence the choice of phrases as antecedents/correlates. In particular, phrases that receive a pitch accent are more accessible for interpretation than phrases that do not receive a pitch accent. The pitch accent may convey focus as well as prominence, further increasing the accessibility of the accented material.

The experiments reported below use ellipsis sentences called replacives. We follow the analysis of the replacive construction that is given by Reinhart (1991), who argues that the replacive phrase ((not) Jessie, in (2)) is simply adjoined to a clause. We assume that the replacive is always adjoined to the root clause of a sentence.[2] The interpretation of the sentence is generated by raising the correlate (Ian), which is the element in the full clause which contrasts with the replacive phrase, at Logical Form (LF). The resulting open proposition "x left" is then applied to the correlate and, in its negated form, to the replacive (Jessie). On this account an ambiguous replacive like (3) is not a constituent structure ambiguity. All that differs between the two interpretations is the choice of a correlate (the phrase to be raised), either the matrix subject Robert or the embedded subject Ian. This

[2]Sentences like (i) below should be ungrammatical if the replacive needed to be attached to the minimal clause containing its correlate. In (i) the adverbial phrase next week at the executive meeting must attach to the higher clause and thus the replacive too may only attach to the higher clause (without crossed branches, which are generally prohibited). Nevertheless, it may take as its correlate the subject of the embedded clause, Mary.

i. [Imagine that the secretary, Mary, quit last week. Someone who knows this says:] John will announce that MARY quit next week at the executive meeting, not SUE.

property of replacives allows us to ask what effects different prosodic factors have on the resolution of ambiguities, not of phrase structure, but rather concerning the choice of antecedent/correlate.

(2) Ian left, not Jessie.

(3) Robert claimed Ian left, not Jessie.

2. Experiment 1

To test the prosodic packaging hypothesis with respect to the influence of prosodic boundaries, an auditory questionnaire was conducted using ambiguous replacive sentences. The matrix and the embedded subject noun phrases were proper names, so each was appropriate as the correlate of the replacive proper name. Participants were asked to answer a question following the replacive sentence. The question disambiguated whether the ambiguous replacive contrasted with the matrix or embedded subject (i.e., had a matrix vs. embedded interpretation, respectively). With these materials, the prosodic packaging hypothesis would predict less frequent choices of the matrix correlate when two prosodic boundaries separated it from the replacive than when only one prosodic boundary intervened. According to the specialized role hypothesis, no effect of prosodic boundary presence or strength is expected.

2.1. Methods

2.1.1. Materials—Twenty-four ambiguous sentences were constructed with proper names as subjects of the matrix and embedded clauses, followed by a replacive ("not" + proper name), as illustrated in (4). All items appear in Appendix A.

(4) a. Diane thought Patrick was entertaining,// not Louise.

b. Diane thought // Patrick was entertaining,// not Louise.

(Note: // = IPh boundary)

Each was recorded in two prosodic versions, with an IPh boundary before the replacive in both conditions and an additional IPh boundary after the matrix verb in condition (b). The names were all pronounced with pitch accents. Acoustic measurements of the properties of these stimuli appear in Tables 1–2. These measurements show that the IPh boundaries were produced as L−H% continuation rises. In condition (a), the F0 of the matrix verbs (i.e., *thought*) remained relatively high, instead of dipping down for a boundary as in condition (b). The matrix verbs in (a) were also much shorter in duration than the same verbs in condition (b), which did show a continuation rise. See Figures 1–2 for visual representations of the intonation of an example sentence.

The items were assigned to two counterbalanced lists, such that each list contained 12 sentences in each prosodic form. These items were combined with 92 other items, including sentences from unrelated experiments and filler items.

2.1.2. Participants and procedures—Seventy-two University of Massachusetts undergraduates participated in individual sessions, listening to sentences over loudspeakers in a sound-proof chamber. After receiving instructions and a 7-item practice list, a computer

played the 116 digitized sentences in the experiment in a different random order to each participant. After each sentence was played, a two-choice question appeared on the computer screen. The question disambiguated between matrix and embedded interpretations (e.g., (4) was followed by What about Louise? She didn't think something about Patrick or Diane thought she wasn't entertaining), and the participant was to pull a response trigger under the answer that best fit his or her own interpretation of the sentence. In the primary experiment, participants were instructed to indicate their own initial intuitive understanding of the sentence just heard, but in a secondary experiment (mentioned below) they were instead instructed to choose the answer they thought the speaker intended. The chosen answers to these questions were recorded by the computer.

2.2. Results

The results are very simple: The percentage of matrix subject choices was 9.6% (95% confidence interval (CI) 8.25 to 10.95[3]) when only one IPh boundary intervened between the matrix subject and the replacive, and 7.7% (95% CI 6.35 to 9.05) when two boundaries intervened. A replication of Experiment 1 on 60 other participants, with the only change being that the participants were instructed to indicate the meaning that they thought the speaker was trying to convey rather than simply indicating their personal interpretation of the utterance, yielded percentages of 9.5% (95% CI 8.17 to 10.83) and 8.9% (95% CI 7.57 to 10.23). In neither case did the difference between the values approach significance ($t(71) = 1.41$, $p = .16$ for the original study; $t(59) = 0.399$ for the replication). The difference is numerically small: 1.9% with a 95% repeated-measures CI of −0.8% to 4.6% in Experiment 1, and 0.6 % with a 95% CI of −2.1% to 3.2% in the replication, and small compared to the effect size of an early boundary in the constituent structure ambiguity results discussed in the Introduction, which ranged from 14 to 33%.

There are, however, concerns with making this comparison. The baseline percentage of matrix subject choices in the present experiment was under 10%, whereas the baseline percentage of matrix attachment choices in the previous research ranged from 81 to 29%. It is possible that Experiment 1 was subject to a floor effect, where it would be very hard to push the percentage of matrix subject choices lower. We address this possibility in Experiments 2a and 2b. A second concern is that, because of the nature of the binomial distribution (e.g., highest variance around 50%), absolute percentage differences cannot legitimately be compared across different regions of probability space. One way to address this concern is to compute odds ratios.[4] The odds of observing a given effect (e.g., interpreting a replacive as having a matrix subject antecedent) in one experimental condition is simply the probability of observing that effect divided by the probability of not observing that effect (e.g., if in one condition, 10% of the subjects give a matrix subject interpretation, the odds of doing so are .10/.90, or .11). An odds ratio is the ratio of the odds in one experimental condition to the odds in a second condition. If there is no difference between the conditions, the odds ratio is expected to be 1.0. Assuming that the numerator of this ratio is the numerically larger odds, the odds ratio can be interpreted as a measure of how much larger the odds are in the first condition than in the second condition. In Experiment 1, the odds ratio (interpreted as reflecting the extent to which adding the early prosodic boundary decreased the percentage of matrix subject interpretations) was 1.26 (95% CI 0.90 to 1.77), and in its replication, 1.07 (95% CI .75 to 1.53). These values can be compared to the odds ratio that adding an early boundary decreased the percentage of matrix attachment interpretations in the experiments cited from Clifton et al. (2002). The odds ratio for

[3]Confidence intervals on individual means throughout the paper are based on the between-subjects variability for each data cell, adjusted for the mean value for each subject, as recommended by Bakeman and McArthur, 1996. However, confidence intervals on the differences between means are computed using a single difference score for each subject.

[4]We thank an anonymous reviewer for calling our attention to this possibility.

Experiment 3 of Clifton et al. was 2.20 (95% CI 1.53 to 3.13) and pooled across the conditions of Clifton et al.'s Experiment 2, 2.47 (95% CI 1.95 to 3.13).

The effect of adding an early boundary on the frequency of choosing the matrix subject thus appeared to be substantially smaller than the effect size of an early boundary in the constituent structure ambiguity results discussed in the Introduction. While we cannot confidently assert that there is no effect of an early boundary in the replacive constructions under study, we do propose that the most satisfactory theoretical proposal would claim that any role for number of boundaries in the replacive sentences of Experiment 1 is, at best, far smaller than the role it plays in interpreting constituent structure ambiguities.

2.3. Discussion

The results of the experiment failed to support the predictions of the prosodic packaging hypothesis with respect to prosodic boundaries. This suggests that the accessibility of linguistic material such as a correlate and remnant is not substantially affected by the presence of intervening prosodic phrase boundaries. This result is informative in light of our (and other's) previous demonstrations that the presence and relative sizes of prosodic boundaries after attachment sites affect the interpretation of sentences with phrase structure ambiguities. Instead, given our assumption that the ambiguity involved in replacives is not a phrase structure ambiguity, the results support the specialized role hypothesis.

However, perhaps examining final interpretations of ambiguous sentences is not the best way to test the prosodic packaging hypothesis. Also, in Experiment 1, we may have run into a floor effect due to the strength of the embedded subject bias. A more sensitive test might be to examine the response times for comprehending strongly biased or unambiguous sentences, as opposed to the sentences used in Experiment 1, whose content permitted either the matrix or the embedded subject to be a plausible correlate of the replacive. With strongly biased or unambiguous test sentences, however, we expect increased accessibility of a constituent to have the effect of speeding comprehension of a sentence in various situations: when that constituent serves as the attachment site for an ambiguously-attached constituent (such as the prepositional phrase in (1)); when it is copied into an elliptical constituent (as in sentences like Sarah left her boyfriend last May. Tina did too; cf. Frazier & Clifton 2001); or, according to the prosodic packaging hypothesis, when it serves as the correlate of a replacive construction.

3. Experiment 2A

To test the prosodic packaging hypothesis, a timed auditory comprehension ("got-it") study was conducted. Thirty-two replacive sentences were constructed, with two lexically different versions of each. One version, illustrated in (5a–d), ended with a replacive strongly biased toward a matrix subject reading in that the replacive (not Lenora) was most parallel, or similar in form, to the matrix subject (Renee). The second version, illustrated in (6), was identical to (5) up to the replacive (not the board), which was biased to be most similar to the embedded subject (the committee). Parallelism of just this sort has been shown to affect the resolution of other ellipsis sentences, like gapping and comparatives (Carlson, 2001, 2002).

Each lexical version of each sentence was then recorded in four prosodic conditions. The replacive constituent (not Lenora or not the board) was immediately preceded by either an intermediate prosodic boundary (ip) or by an intonational phrase boundary (IPh). The matrix verb (realized in the examples) was orthogonally followed by an intonational phrase boundary (IPh) or no boundary at all.

(5) a. Renee realized the committee debated the bill / not Lenora.

b. Renee realized // the committee debated the bill / not Lenora.

c. Renee realized the committee debated the bill // not Lenora.

d. Renee realized // the committee debated the bill // not Lenora.

(6) a. Renee realized the committee debated the bill / not the board.

b. Renee realized // the committee debated the bill / not the board.

c. Renee realized the committee debated the bill // not the board.

d. Renee realized // the committee debated the bill // not the board

(Note: / = ip boundary; // = IPh boundary)

Our earlier work on the effects of boundary size on the preferred interpretation of constituent structure ambiguities suggests that conditions whose later boundary is larger than any relevant earlier boundary (i.e., the a-forms and the c-forms) should behave alike. In particular, the late boundary should disfavor interpreting the ambiguous phrase with material immediately before the prosodic break, and bias toward the interpretation where the ambiguous phrase is grouped with earlier material, such as the verb realized (a matrix interpretation). In the two-boundary conditions with the larger of the two boundaries appearing early (b-forms), the favored interpretation should group the ambiguous phrase together with material lower in the tree. Finally, equal late and earlier boundaries, as in the d-forms, should result in more balanced high vs. low interpretations. The lexical content of the sentences in Experiment 2 was designed to bias them toward a matrix or an embedded subject interpretation (in contrast to the lexically-ambiguous sentences of Experiment 1). We expected that conditions where this lexical bias matched the bias provided by the pattern of prosodic boundaries would be quicker and easier (i.e., (5a,c) easier than (6a,c) and (5b); (6b) easier than (5b); (5d) and (6d) intermediate).

The prosodic packaging hypothesis leads us to further expect that, relative to the a-forms, the addition of the second prosodic boundary in the b-forms should slow down the processing of the matrix subject conditions (5) more than the embedded subject conditions (6). In other words, an interaction would be expected: the difference in processing time between (5b) and (5a) should be larger, due to the difficulty of accessing a correlate two prosodic phrases back in (5b), than the difference between (6b) and (6a), where the additional prosodic break does not intervene between the replacive and the most parallel (embedded) subject. A similar contrast holds for the (c) and (d) examples. If increased prosodic accessibility also increases the likelihood of a constituent being chosen as the correlate of a replacive, then the number of matrix subject responses in (5b) should be smaller than the number of matrix subject responses in (5a) (although it is possible that the choice of correlate will be largely determined by the lexical form and content of the replacive, i.e., the matrix subject will generally be chosen for sentences like (5) and the embedded subject for sentences like (6), regardless of prosody).

3.1. Methods

3.1.1. Materials—Thirty-two sentences were constructed in two lexical versions each, illustrated in (5) and (6) for the matrix and the embedded subject cases, respectively. In the matrix subject conditions, the replacive noun phrase was similar to the matrix subject, while in the embedded subject conditions, it was most similar to the embedded subject. All sentences appear in Appendix B. The experiment was designed as the factorial combination of matrix vs. embedded subject resolution, ip vs. IPh as the late boundary (before the replacive constituent), and no boundary vs. an IPh boundary as the early boundary. Acoustic measurements of the conditions are in Tables 1–2. On the matrix verb (realized), the conditions with no boundary (a,c) show relatively high F0s in contrast to the deep fall-rise of the L−H% IPh boundaries (b,d), and durational measurements also show greatly increased length in the boundary conditions. Before the replacive, there was always a prosodic boundary, but the F0 averages and durations show clear pitch and length differences between the L−H% IPh boundary conditions (c–d) and the simple fall of the L- ip boundary conditions (a–b). The F0 peak measurements on the matrix and embedded subjects show H* accents on each.

The resulting 32 × 8 utterances were digitized (16 bits, 22 kHz) and assigned to eight separate counterbalanced presentation lists. Each list included 32 other experimental sentences (investigating verb phrase ellipsis and adverb scope) plus 26 varied filler sentences.

3.1.2. Participants and procedures—Forty-eight University of Massachusetts students, all native speakers of English, were tested individually in a sound-deadened chamber. They were instructed that they would hear some sentences played by a computer over speakers at a comfortable listening level, and they were to attempt to understand each sentence. If they understood a sentence to their satisfaction, they were to pull the right-hand trigger on a response console as quickly as they could, once the sentence was over (a ‘got-it’ response). If, however, they found the sentence difficult to understand or confusing (described to them as they did not 'get it'), they were to pull the left-hand trigger. Because of an error in the experimental script, one item (item 25) in the Appendix was presented incorrectly in one condition. This item was eliminated from the analyses.

A pull of the right trigger, indicating understanding, was followed by the visual presentation of a question and two alternative answers, displayed to the right vs. the left underneath the question. The participant was to answer the question by pulling the console trigger on the corresponding side. In items 1–16, the question asked only about the predicate of the main or the subordinate clause (e.g., Who debated the bill?), and provided as possible answers the actual subject of this predicate (the committee; correct) and the remnant of the replacive (Lenora; incorrect). In items 17–32, the question targeted the interpretation of the replacive itself, and was formulated in the negative (e.g., What wasn't true about Lenora?), with the matrix predicate (in this example, realized something) and the subordinate clause predicate (debated the bill) provided as alternatives.

Each participant first heard a practice list of 7 sentences and then heard the list of 90 experimental sentences, in an individually-randomized order. Since eight different counterbalancing lists were used, each participant heard four sentences in each of the eight conditions illustrated by (5) and (6); across all participants, each item was heard six times in each version. All responses and reaction times were recorded by the computer that controlled the experiment. "Got-it" response times were measured from the end of the sentence.

3.2. Results and Discussion

The primary data are the reaction times and frequencies of got-it (understood) responses. First, though, we will report data from the questions. Question-answering performance on the first 16 items (where the questions were about unambiguous aspects of the sentences) was nearly 100% correct. Questions for the last 16 items (less the eliminated item 25), which asked about what the remnant of the replacive didn't do, were scored in terms of whether the participant chose the matrix vs. the embedded predicate. Since no question was presented when a subject indicated that they did not 'get' a sentence, there were substantial amounts of missing data for some conditions when the data were tabulated by subjects. However, there were no missing data in the by-items tabulation, so the by-items analysis can be presented. The percentages of choices of embedded predicates appear in Table 3. Analyses of variance of these percentages indicate a lower frequency of choice of the embedded predicate for sentences whose replacive remnants were biased toward having a matrix correlate than for sentences biased toward an embedded correlate ($F2(1, 14) = 19.17$, $p < .001$). No other effect or interaction approached significance (all $p > .25$). So while our attempt to lexically bias interpretation toward a subordinate correlate was successful, our attempt to bias toward a matrix correlate only increased the percentage of matrix choices from approximately 20% to 33%. Still, this is a far higher percentage than the less-than-10% choices of a matrix subject antecedent in Experiment 1. We submit that the manipulation induced a substantial difference in the interpretation of the matrix and embedded sentences, large enough to warrant examining the reading of the sentences (but small enough to justify a stronger bias manipulation, in Experiment 2B).

The got-it reaction time and proportion "got-it" results are also presented in Table 3, after eliminating responses that were over 8000 ms (a single trial) or under 100 ms [5]. The matrix-biased (not Lenora) examples (5) were responded to more slowly and accepted less frequently than the embedded subject-biased (not the board) examples (6). Both effects reached significance in analyses of variance permitting generalization to participants and to items (RT: $F1(1,47) = 12.0$, $p < .001$; $F2(1,30) = 13.89$, $p < .001$; proportion "got-it": $F1(1,47) = 22.19$, $p < .001$; $F2(1,30) = 112.60$, $p < .001$). The only other effect that reached significance appeared in the proportion "got-it" analysis: Items with a late ip were accepted less frequently than items with a late IPh boundary, 80.6 vs. 84.6% ($F1(1,47) = 8.785$, $p = .01$, $F2(1,30) = 8.644$, $p = 0.01$), presumably reflecting a difference in prosodic naturalness.

Examination of Table 3 could lead one to think that reaction times are particularly slow in the matrix conditions that had an early IPh boundary (1020 and 1056 ms), which could suggest that the presence of a prosodic boundary after the matrix verb (or, equivalently, the presence of a second boundary between this verb and the replacive) made the targeted antecedent of the replacive less available. However, the interaction (matrix vs. embedded resolution × presence vs. absence of the early boundary) that would lead one to trust this apparent effect was not significant ($F1(1,47) = 1.392$, $p = 0.24$) and there was no sign in the question-answer percentages that the presence of the boundary influenced choices. Further, the apparent effect of the early boundary was limited to a relatively small number of very long responses. Reducing the "got-it" RT cutoff from 8000 ms to 3000 ms eliminated 2% of the data, and reduced the size of the effect of presence of the early IPh boundary for matrix subject items from 74 ms to less than 2 ms. While we cannot securely conclude that the presence of an early boundary in a sentence that is biased toward a matrix interpretation has no effect on comprehension time, the evidence that it does have such an effect is scant.

[5]Reaction times in all experiments reported here are pooled over both "got-it" and "didn't understand" responses. If the two responses were separated out, the extensive amount of missing data – up to 94.2% in the worst case (see Table 3) – would prevent doing statistical analyses.

Apart from this very weak hint of support for an effect of an additional prosodic boundary on reaction times, the data failed to confirm the predictions of the prosodic packaging hypothesis. An additional prosodic boundary had at most a small effect on the availability of phrases as judged by the question answers and the proportion of 'got-it' responses. As can be gleaned from Table 3, the percentages of embedded subject interpretations in the conditions with and without an early boundary were 77 and 74% respectively, a nonsignificant difference of 2.5% (95% CI = −4.3% to 9.2%). Similarly, the proportions of 'got-it' responses for the matrix conditions that did and did not have an early prosodic boundary (which one might have expected to differ if the early boundary mattered) were 71 and 73% respectively, a difference of 2.6% (95% CI −0.2% to 6.4%). The size of the 'got-it' effect can be expressed as an odds ratio, for comparability with the Experiment 1 data; its value was 1.16 (95% CI .84 to 1.59). The data also did not show interesting differences based on the pattern of the size of prosodic boundaries (IPh vs. ip), probably because the preference for the embedded subject conditions drowned out any more subtle effects.

There are several possible explanations for the clear advantage of the embedded subject condition. One, following Carlson (2002), would attribute the advantage for the embedded subject antecedent to the observation that another, possibly preferred way to communicate the matrix subject interpretation would be to use verb phrase ellipsis, i.e., (But) Lenora didn't (see discussion in Carlson (2002)). Additionally, the relative recency of the correlate in the embedded subject condition might be responsible for its higher acceptance and faster responses. The current experiment and those below were not designed to distinguish among these hypotheses.

The materials in Experiment 2A were biased toward either the matrix subject or the embedded subject correlate, but they were not unambiguous. Experiment 2B strengthened the disambiguation in the sentences by manipulating the animacy of the subject noun phrases. In principle, the effect of prosodic boundaries might be stronger with fully disambiguated materials.

4. Experiment 2B

To further test the prosodic packaging hypothesis, a new set of 24 replacive sentences were constructed. These materials differed in two ways from those in Experiment 2A. First, they were essentially disambiguated to matrix vs. embedded interpretations of the replacive clause, while those in Experiment 2A were merely strongly biased. Second, Experiment 2B did not investigate the relative sizes of intonational phrase boundaries, but simply compared the effects of having one vs. two boundaries intervene between the matrix subject and the replacive clause. Experiment 2B thus permitted a second look at the suggestive but nonsignificant effect of one vs. two boundaries on the accessibility of the matrix subject as the correlate of the replacive.

4.1. Methods

4.1.1. Materials—The 24 experimental sentences had the same structure as those in Experiment 2A, except that the new sentences were fully disambiguated: the replacive constituent was inanimate for the embedded condition and human for the matrix condition, while the matrix and embedded predicates were anomalous or nearly so for inanimate and animate noun phrases, respectively. This is illustrated in (7) and (8), which present matrix subject and embedded subject sentences, respectively.

(7) a. Diane thought that the movie was well-edited // not Louisa.

b. Diane thought // that the movie was well-edited // not Louisa.

(8) a. Diane thought that the movie was well-edited // not the soundtrack.

b. Diane thought // that the movie was well-edited // not the soundtrack.

All sentences appear in Appendix C. A question relevant to the contrast between matrix and embedded subjects was made up for each sentence (e.g., What about Louisa/the soundtrack? Didn't think something or Was not well-edited).

Each lexical sentence condition was recorded in two prosodic versions. Both versions were pronounced with an intonational phrase boundary (IPh) before the replacive. In the a-forms, this was the only prosodic boundary (above the phonological word level). In the b-forms, there was also an IPh boundary after the matrix verb. The acoustic measurements in Tables 1–2 show the L−H% continuation rises which were found at each prosodic boundary location, as well as the increased duration of the first verb in the (b) conditions where it preceded a boundary. The resulting sentences were recorded and digitized as in Experiment 2A, and assigned to four counterbalanced lists so that each list contained six instances of each of the four conditions (matrix vs. embedded × one vs. two phrase boundaries) illustrated in (7) and (8). These sentences were combined with 92 other sentences, some from unrelated experiments and some fillers, to make up the full presentation lists.

4.1.2. Participants and procedures—Forty-eight undergraduates at the University of Massachusetts were tested using the same procedures as described in Experiment 2A. Twelve participants received each of the four lists described above, in individually-randomized orders.

4.2. Results and Discussion

Questions were asked following a "got-it" response to a sentence. Question-answering accuracy averaged 83% for matrix subject items and 88% for embedded subject items, with only minor variation between intonational conditions.

As in Experiment 2A, the data of interest were the percentages and reaction times of "got-it" responses. These appear in Table 4. The differences among reaction times were nonsignificant in analyses of variance treating correlate and prosodic pattern as factors (all $Fs < 1$). "Got-it" responses were more frequent for embedded subject than for matrix subject items ($F1(1,47) = 27.83$, $F2(1, 23) = 155.01$, $p < .001$). All effects involving prosodic patterns had $Fs < 1$.

There was thus little indication that the pattern of prosodic boundaries affected the accessibility of the correlate of a replacive. The difference in percentages of ‘got-it’ responses for matrix sentences with and without an early prosodic boundary, which could have been expected to reflect an effect of the early boundary, was only 3.8% (95% CI = −0.1% to 8.5%) (and the corresponding odds ratio was 1.19, with a 95% CI of .84 to 1.68). The hint of an effect of number of boundaries that appeared in the RT data in Experiment 2A disappeared in Experiment 2B. The results of Experiment 2 as a whole failed to support the predictions of the prosodic packaging hypothesis. In sum, none of the experiments provide evidence that the interpretation of replacives (whether ambiguous, strongly biased, or unambiguous) is more difficult when several prosodic boundaries or stronger prosodic boundaries intervene between the replacive and its correlate. This suggests, then, that the specialized role hypothesis may be more successful in predicting the interpretation of

ellipsis sentences. To see whether the interpretation of replacives *is* influenced by prominence information, Experiment 3 manipulated accents rather than prosodic boundaries.

5. Experiment 3

Experiment 3 investigated whether the interpretation of ambiguous replacives can be influenced by the presence of a prominent pitch accent on one or another potential correlate, as the specialized role hypothesis predicts. If more prominence, as conferred by the pitch accent, results in greater accessibility in a discourse representation, then accented phrases should more often be chosen as the correlate of a replacive than unaccented phrases. Also, there is reason to believe that ellipsis sentences should respond to prosodic prominence based on previous prosodic studies (e.g., Carlson, 2001; Frazier & Clifton, 1998) as well as syntactic and thematic theory (e.g., Merchant, 2001).

5.1. Methods

5.1.1. Materials—The 24 sentences used were the same as those used in Experiment 1, and were paired with very similar questions that distinguished between matrix and embedded subject correlates of the replacive. The sentences were recorded in three prosodically different conditions, as indicated in (9).

(9) a. ROGER insisted that Alice was reliable // not ANDREW

b. Roger insisted that ALICE was reliable // not ANDREW

c. ROGER insisted that ALICE was reliable // not ANDREW.

All sentences had an IPh boundary, a continuation rise, before the replacive phrase, but there were no prosodic boundaries preceding that point. Version (9a) was recorded with prominent L+H* pitch accents on the matrix subject and the replacive (Andrew); (9b) was recorded with L+H* accents on the embedded subject and the replacive; and (9c) was recorded with L+H* accents on each noun phrase. Acoustic measurements of the F0s are shown in Table 1.[6]

There is an intonational asymmetry between the prosodic contours in (9a–b), in that the embedded subject (Alice) in (9a) can be completely de-accented, while it is infelicitous to produce a sentence like (9b) with no accent at all on the matrix subject (Roger). This is probably due to the phonological demand for an early accent within a prosodic phrase (Shattuck-Hufnagel, Ostendorf, & Ross, 1994; Ladd, 1996). It is true, therefore, that the matrix subjects in sentences in condition (b) had a H* accent. Such accents were not nearly as prominent as the L+H* accents placed on other arguments. The mean F0 for the H* accents was 212 Hz, with a standard deviation of 11 Hz; the mean F0s for L+H* accented subjects were well over 300 Hz with similar standard deviations. There was no overlap between the F0 peaks which H* and L+H* accents reached. If this matrix subject accenting has an effect despite the F0 differences, though, we would expect that conditions (b) and (c) would produce similar response percentages.

The experimental sentences were assigned to three counterbalanced lists, each with eight items in each of the three prosodic conditions. They were combined with 74 other items of varying constructions.

[6]One sentence–number 3 in Appendix A–was not played correctly in the both accent (c) condition, and was therefore dropped from the analyses.

5.1.2. Participants and procedures—Forty-eight University of Massachusetts undergraduates were tested in individual sessions. The procedures used were the same as those described in the Methods section of Experiment 1.

5.2. Results and Discussion

The mean percentage of responses choosing the matrix subject as the correlate of the replacive and the mean question-answering times for all responses appear in Table 5.

The data were analyzed in an analysis of variance that included treatments as a within-subjects effect and counterbalancing groups as a between-subjects factor (see Pollatsek & Well, 1995, for justification). There were no significant differences among the mean question-answering times ($F1(2,92) < 1.0$; $F2(2,42) < 1.0$). However, the position of the pitch accent did significantly affect the choice of the correlate of the replacive ($F1(2,90) = 33.245$, $p < .001$; $F2(2, 42) = 16.99$, $p < .001$). Separate t-tests (Bonferroni-corrected) indicated that the percentage in the matrix subject accented condition (9a) was greater than the percentages in each of the other two conditions (9b, 9c) (in both cases, $t(47) > 6.0$, $p < .001$), which did not differ from each other ($t(47) = 0.14$). The odds ratio of this effect size was 3.53 (95% CI 2.61 to 4.92). Prosodic prominence did affect the interpretation of replacives. The fact that (9a) was greater than (9c) shows that having a prominent accent on the matrix subject alone increased the number of matrix subject interpretations compared to having an accent on both possible antecedents. The fact that (9a) was greater than (9b) shows that having a prominent accent on the matrix subject alone increases the number of matrix subject interpretations compared to having an accent only on the already favored antecedent. The fact that the embedded subject accent (9b) and the equal-accent (9c) pronunciations resulted in only a very small difference (0.3%, 95% CI = −3.4% to 3.9%; odds ratio of 1.03, 95% CI .65 to 1.63) in numbers of embedded interpretations suggests that an accent favoring an already favored antecedent may have little effect, or that the presence of H* accents on the matrix subjects in (9b) may have had a similar effect as the L+H* accents in (9c).

6. Conclusions

The prosodic packaging hypothesis proposed that the accessibility of all linguistic material should be affected by prosodic boundaries. The first four experiments of this project (Experiment 1, its replication, Experiment 2a, and Experiment 2b), however, failed to find any significant effects of prosodic boundaries on the interpretation of replacive sentences, and while they could not in principle show that there is no effect, they indicate that any effect is small in comparison to previously-reported effects of an early prosodic boundary in other grammatical constructions. These results should not be taken as evidence that linguistic material one vs. two prosodic phrases back is equally accessible to the parser for all purposes (or equally accessible to other cognitive functions). Instead, we believe that these data reveal a limitation on the environments where prosodic boundaries affect interpretation. The replacive sentences used in this project differ from many other ambiguous sentences in that they are not phrase-structurally ambiguous.

We do not think that some structures or dependency types just arbitrarily happen not to be affected by the number and relative size of prosodic boundaries, and others do. Rather we suspect that there is a principled distinction. Sentences with phrase structure ambiguities have often been substantially influenced by the presence and relative size of prosodic boundaries (see, e.g., Carlson et al., 2001; Clifton et al., 2002; Lehiste, 1973; Price et al., 1991; Schafer, 1997; Schafer et al., 2000; Speer et al., 1996; etc.). Hirotani (2004) showed that scopal items which must be c-commanded by their binder also are limited by prosodic boundaries: the binder and the bindee preferably occur in the same prosodic phrase (see

Hirotani for details). We suggest that the prosodic accessibility of attachment sites in structurally ambiguous sentences is indeed reduced by multiple intervening prosodic boundaries. But this is in sharp contrast to the antecedents of co-referential pronouns, ellipsis remnants, and replacives. Prosodic boundaries influence decisions about hierarchical structure but do not necessarily influence the accessibility of constituents preceding those boundaries for all grammatical dependencies, as formulated in the "specialized role hypothesis" discussed earlier.

In Experiment 3, the position of pitch accents did significantly affect the interpretation of replacive sentences, as expected if prosodic prominence influences ellipsis resolution. In particular, prominently accenting the matrix subject increased the likelihood of this otherwise dispreferred argument being chosen as the correlate of the replacive phrase. This suggests that the prominence or focus conveyed by pitch accents can affect the prosodic accessibility of a phrase, and that this type of accessibility is relevant to antecedent choice and discourse processing. See Foraker and McElree (2007) for a similar point with respect to pronouns focused by clefting.

The work of Hirotani (2004) dovetails neatly with this set of results. She proposed the Scope-Prosody Correspondence principle to explain the effects of prosodic boundaries on the interpretation of scope-ambiguous wh-questions in Japanese. This principle highlights the preference for the grammatical dependency between the wh-word and its trace position to be satisfied within one prosodic phrase, similar in spirit to the prosodic packaging hypothesis. But when she turned to pronominal ambiguities, like pronoun reference, prosodic phrasing had no effect.

Lehiste (1973) conducted a pioneering study of the role of prosody in a wide variety of different types of ambiguities. She found no effect of prosody on lexical and thematic ambiguities, but often strong effects on constituent structure ambiguities. She attributed the disambiguation of constituent structure ambiguities primarily to temporal or 'junctural' cues. Price et al. (1991) studied a number of different constituent structure ambiguities and found effects of prosodic boundaries on their preferred interpretation. They concluded that temporal and juncture cues were the primary supersegmental cues for disambiguating sentences, with pitch cues playing ONLY a supporting role. Both studies may be seen as a forerunner of the specialized role hypothesis, provided their claims are limited to hierarchical structure decisions (those involving phrase structure or c-command).

It is possible that the specialized role hypothesis pursued here is ultimately related to differences in the acoustic dimensions supporting grouping information, which are primarily temporal relations, and the acoustic dimensions supporting prominence information, which are primarily pitch information. Whether separate neural mechanisms underlie pitch perception and the sensory processing of temporal relations is a question of considerable interest (e.g., Gandour et al, 2002; Schirmer et al., 2001). If there are two somewhat separate channels of information at early/lower levels of processing, presumably they must be brought together into a prosodic representation. Precisely where these two types of information come together and to what extent they interact is a fascinating question, though one which goes beyond the scope of the present paper.

To conclude, prosodic boundaries have been shown to play a central role in sentence comprehension. But the present studies suggest their role is largely confined to influencing decisions about hierarchical structure, rather than also influencing the accessibility of constituents in memory for purposes of identifying antecedents of pronouns, elided constituents, and replacives. This does not entail that there could be no effect whatsoever of a prosodic boundary on the prominence of a constituent. For example, because focused

constituents are often followed by a prosodic boundary (and/or de-prosodification of following material; Schafer & Jun, 2005), a non-reduced but not prominently accented constituent followed by a prosodic boundary might be taken to be focused, and thereby accorded prominent status even though the constituent was not spoken with a high pitch peak. However, note that this sort of indirect effect of prosodic boundaries on prominence is a far cry from the claims of the prosodic packaging hypothesis, according to which the presence of a boundary should have automatically decreased the accessibility of all earlier constituents.

Acknowledgments

This research was supported in part by Kentucky NSF EPSCoR REG EPS-0132295 and Grant NSF-BCS 0090674, and by Grant Number HD18708 from the National Institute for Child Health and Human Development. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institute for Child Health and Human Development or the National Institutes of Health.

Appendix A

Materials used in Experiments 1 and 3

1. Roger insisted that Alice was reliable, not Andrew.
2. Diane thought Patrick was entertaining, not Louise.
3. Jeanette felt that Donald was depressing, not Melissa.
4. Rebecca argued that Josh was impolite, not Maria.
5. Jason found out Barbara was in New York, not Edward.
6. Dan knew that Polly would be impatient, not Matthew.
7. Simon promised that Lisa would visit, not Robbie.
8. Dominic guessed that Paula would be difficult, not Timothy.
9. Sharon said David would be agreeable, not Claudia.
10. Michelle noticed that Gary was sick, not Nicole.
11. Arthur suggested that Felicia would be helpful, not William.
12. Rafael claimed Anne would cook, not Dennis.
13. Kimberly thought Steven was ambitious, not Veronica.
14. Tony said that Anna was vivacious, not David.
15. Lillian realized that Walter was rude, not Rachel.
16. Jeremiah found out Lucinda was underage, not Alexander.
17. Abigail insisted that Roger was brilliant, not Alicia.
18. Peter guessed that Eva was Turkish, not Larry.
19. Carmen wrote that Antonio was friendly, not Elena.
20. Tricia hinted that Ronald was upset, not Brittany.
21. Wendy reported that Charles was outrageous, not Stacy.
22. Natalie reported that Tom was fired, not Jane.
23. Justin suspected Alicia was asthmatic, not Victor.

24. Julie implied that Gordon was wrong, not Maryann.

Appendix B

Materials used in Experiment 2A

1. Renee realized the committee debated the bill, not Lenora / the board.
2. Dionne thought the dog was well-trained, not Louann / the cat.
3. Jeanette felt that the rain was depressing, not Giselle / the hail.
4. Delilah argued that the paint was toxic, not Maria / the primer.
5. Jason found out that the water contained arsenic, not Juwan / the wood.
6. Dan knew that oil would be expensive, not Matt / gas.
7. Thomas promised the airfare would be cheap, not Robbie / lodging.
8. Dominic guessed that calculus would be difficult, not Timothy / chemistry.
9. Louisa said the readings would be easy, not Marlena / the paper.
10. Michelle noticed that the ceiling was painted, not Nicole / the wall.
11. Matthew suggested the aquarium would be the most fun, not William / the film.
12. Simon promised the appetizers would be delicious, not Dennis / dinner.
13. Penelope argued that the local news was important, not Veronica / the national news.
14. Ramona knew that the quartet was accomplished, not Dolores / the chorus.
15. Lydia realized that the accounting was illegal, not Laverne / the merger.
16. Jeremiah found out that the profits were inflated, not Alexander / the expenses.
17. Abigail insisted that the melody was repetitive, not Melanie / the rhythm.
18. Peter guessed that tulips are Turkish, not Walter / dahlias.
19. Maria wrote that the sculptures were fantastic, not Elena / the paintings.
20. Tricia hinted that the novel was racy, not Tallulah / the movie.
21. Pam reported that CDs were outlawed in Turkestan, not Kate / kites.
22. Pearl reported that broccoli was healthful, not Lynne / figs.
23. Jean-Luc suspected that the ceramics were just decorative, not Pierre / the rugs.
24. Emmylou realized that MapQuest was inaccurate, not Marianne / the state maps.
25. Andrea implied that horned Dorset sheep were rare, not Marina / Merinos.
26. Roger insisted that VWs are reliable, not Andrew / Hondas.
27. Tony said that the tips were lucrative, not Gary / the wages.
28. Dinah found out that the rumor was incorrect, not Desireé / the resumé.
29. Maxwell noticed that the maple was diseased, not Monroe / the oak.
30. Tim insisted that psychology is interesting, not Lance / math.
31. Pamela maintained that science fiction was edifying, not Natalie / fantasy.

32. Paolo suggested that the salary was adequate, not Antonia / the benefits.

Appendix C

Materials used in Experiment 2B

1. Roger insisted that VWs are reliable, not Andrew / Hondas.

2. Diane thought that the movie was well-edited, not Louisa / the soundtrack.

3. Jeanette felt that the car was skidding, not Melissa / the trailer.

4. Rebecca argued that the paint was toxic, not Maria / the primer.

5. Jason found out the well water contained arsenic, not Edward / the pipes.

6. Dan knew that oil prices would rise, not Matthew / gas prices.

7. Simon promised that the airfare would be cheap, not Robbie / lodging.

8. Dominic guessed that calculus would be a required course, not Timothy / chemistry.

9. Sharon said many readings would be assigned, not Claudia / many papers.

10. Michelle noticed that the ceiling was painted, not Nicole / the wall.

11. Arthur suggested that the aquarium would be crowded, not William / the zoo.

12. Rafael claimed the appetizers would be delicious, not Dennis / dinner.

13. Kimberly thought the local news was presented well, not Veronica / the national news.

14. Tony said that the tips were lucrative, not David / the wages.

15. Lillian realized that the accounting was illegal, not Rachel / the merger.

16. Jeremiah found out the profits were inflated, not Alexander / the expenses.

17. Abigail insisted that the melody was classical, not Alicia / the rhythm.

18. Peter guessed that tulips are bulbs, not Larry / roses.

19. Carmen wrote that the sculptures were defaced, not Elena / the paintings.

20. Tricia hinted that the novel had a good plot, not Brittany / the movie.

21. Wendy reported that whole CDs were downloaded, not Stacy / the files.

22. Natalie reported that broccoli was healthful, not Jane / grapes.

23. Justin suspected the ceramics were handmade, not Victor / the rugs.

24. Julie implied that MapQuest was user-friendly, not Maryann / the atlas.

References

Bakeman R, McArthur D. Picturing repeated measures: Comments on Loftus, Morrison, and Others. Behavior Research Methods, Instruments, and Computers. 1996; 28:584–589.

Balogh, JE. Unpublished PhD Dissertation. San Diego: University of California; 2004. Pronouns, prosody, and the discourse anaphora weighting approach.

Beckman ME, Ayers GM. Guidelines for ToBI labelling, version 2.0. Ohio State University. 1993

Caplan D. Clause boundaries and recognition latencies for words in sentences. Perception and Psychophysics. 1972; 12:73–77.

Carlson K. The effects of parallelism and prosody on the processing of gapping structures. Language and Speech. 2001; 44:1–26. [PubMed: 11430186]

Carlson, K. Parallelism and prosody in the processing of ellipsis sentences. New York and London: Routledge; 2002.

Carlson K, Clifton C Jr, Frazier L. Prosodic boundaries in adjunct attachment. Journal of Memory and Language. 2001; 45:58–81.

Carroll, PJ.; Slowiaczek, ML. Modes and modules: Multiple pathways to the language processor. In: Garfield, JL., editor. Modularity in knowledge representation and natural-language understanding. Cambridge, MA: MIT Press; 1987. p. 221-247.

Clifton C Jr, Carlson K, Frazier L. Informative prosodic boundaries. Language and Speech. 2002; 45:87–114. [PubMed: 12613557]

Cohen, J. Statistical power analysis for the behavioral sciences. 2nd ed.. Hillsdale, NJ: Lawrence Earlbaum Associates; 1988.

Dahan D, Tanenhaus MK, Chambers CG. Accent and reference resolution in spoken-language comprehension. Journal of Memory and Language. 2002; 47:292–314.

Dunlop WP, Cortina JM, Vaslow JB, Burke MJ. Meta-analysis of experiments with matched groups or repeated measures designs. Psychological Methods. 1996; 1:170–177.

Eberhard K, Sedivy JC, Spivey-Knowlton MJ, Tanenhaus MK. Eye movements as a window into real-time spoken language processing in natural contexts. Journal of Psycholinguistic Research. 1995; 24:409–436. [PubMed: 8531168]

Foraker S, McElree B. The role of prominence in pronoun resolution: Active vs. passive representations. Journal of Memory and Language. 2007; 56:357–383.

Frazier L, Clifton C Jr. Comprehension of sluiced sentences. Language and Cognitive Processes. 1998; 13:499–520.

Frazier L, Clifton C Jr. Parsing coordinates and ellipsis: Copy α. Syntax. 2001; 4:1–22.

Frazier L, Fodor JD. The Sausage Machine: A new two stage parsing model. Cognition. 1978; 6:291–325.

Gandour J, Wong D, Lowe M, Dzemidzic M, Satthamnuwong N, Tong Y, Li X. A cross-linguistic fMRI study of spectral and temporal cues underlying phonological processing. Journal of Cognitive Neuroscience. 2002; 14(7):1076–1087. [PubMed: 12419130]

Hirschberg J, Ward G. Accent and bound anaphora. Cognitive Linguistics. 1991; 2:101–121.

Hirotani, M. Unpublished PhD Dissertation. Amherst, MA: University of Massachusetts; 2004. Prosody and LF interpretation: Processing Japanese wh-questions.

Jarvella RJ, Herman SJ. Clause structure of sentences and speech processing. Perception and Psychophysics. 1972; 11:381–384.

Kjelgaard MM, Speer SR. Prosodic facilitation and interference in the resolution of temporary syntactic closure ambiguity. Journal of Memory and Language. 1999; 40:153–194.

Ladd, DR. Intonational phonology. Cambridge: Cambridge University Press; 1996.

Lehiste I. Phonetic disambiguation of syntactic ambiguity. Glossa. 1973; 7:102–122.

Merchant, J. The syntax of silence: Sluicing, identity, and islands in ellipsis. Oxford: Oxford University Press; 2001.

Pierrehumbert, J.; Beckman, M. Japanese Tone Structure. Cambridge, MA: MIT Press; 1988.

Pierrehumbert, JB.; Hirschberg, J. The meaning of intonational contours in the interpretation of discourse. In: Cohen, PR.; Morgan, J.; Pollack, ME., editors. Intentions in communication. Cambridge, MA: MIT Press; 1990. p. 271-311.

Pollatsek A, Well AD. On the use of counterbalanced designs in cognitive research: A suggestion for a better and more powerful analysis. Journal of Experimental Psychology: Learning, Memory, and Cognition. 1995; 21:785–794.

Price PJ, Ostendorf M, Shattuck-Hufnagel S, Fong C. The use of prosody in syntactic disambiguation. Journal of the Acoustical Society of America. 1991; 90:2956–2970. [PubMed: 1787237]

Pynte J, Prieur B. Prosodic breaks and attachment decisions in sentence parsing. Language and Cognitive Processes. 1996; 11:165–191.

Reinhart, T. Elliptic conjunctions-non-quantificational LF. In: Kasher, A., editor. The Chomskyan turn. Cambridge, MA: MIT Press; 1991. p. 360-384.

Schafer, A. Unpublished PhD Dissertation. Amherst, MA: University of Massachusetts; 1997. Prosodic parsing: The role of prosody in sentence comprehension.

Schafer A, Carlson K, Clifton C Jr, Frazier L. Focus and the interpretation of pitch accent: Disambiguating embedded questions. Language and Speech. 2000; 43:75–105.

Schafer A, Carter J, Clifton C Jr, Frazier L. Focus in relative clause construal. Language and Cognitive Processes. 1996; 11:135–163.

Schafer, AJ.; Jun, S-A. Effects of contrastive focus on prosodic cues to clause boundaries. Paper presented at the Eighteenth Annual CUNY Conference on Human Sentence Processing; Tucson, AZ. 2005.

Schafer, A.; Speer, SM. Prosodic phrasal effects on the resolution of within-category lexical ambiguity. Paper presented at the Eleventh Annual CUNY Conference on Human Sentence Processing; New Brunswick, NJ. 1998.

Schafer AJ, Speer SR, Warren P, White SD. Intonational disambiguation in sentence production and comprehension. Journal of Psycholinguistic Research. 2000; 29:169–182. [PubMed: 10709182]

Scharpff, PJ.; van Heuven, VJ. Effects of pause insertion on the intelligibility of low-quality speech. Proceedings of 7th FASE Symposium, SPEECH 88; Edinburgh, Scotland. 1988.

Schirmer A, Alter K, Kotz SA, Friederici AD. Lateralization of prosody during production: A lesion study. Brain and Language. 2001; 76:1–17. [PubMed: 11161351]

Schwarzschild R. Givenness, AvoidF and other constraints on the placement of accent. Natural Language Semantics. 1999; 7:141–177.

Selkirk, E. Phonology and syntax: The relation between sound and structure. Cambridge, MA: MIT Press; 1984.

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

Shattuck-Hufnagel S, Turk AE. A prosody tutorial for investigators of auditory sentence processing. Journal of Psycholinguistic Research. 1996; 25:193–247. [PubMed: 8667297]

Speer SR, Kjelgaard MM, Dobroth KM. The influence of prosodic structure on the resolution of temporary syntactic closure ambiguities. Journal of Psycholinguistic Research. 1996; 25:249–272. [PubMed: 8667298]

Steinhauer K, Alter K, Friederici AD. Brain potentials indicate immediate use of prosodic cues in natural speech. Nature Neuroscience. 1999; 2:191–196.

Swerts M, Geluykens R. Prosody as a marker of information flow in spoken discourse. Language and Speech. 1994; 37:21–43.

Warren P, Grabe E, Nolan F. Prosody, phonology and parsing in closure ambiguities. Language and Cognitive Processes. 1995; 10:457–486.

Weber A, Grice M, Crocker M. The role of prosody in the interpretation of structural ambiguities: A study of anticipatory eye movements. Cognition. 2006; 99(2):B63–B72. [PubMed: 16157327]

Welby P. Effects of pitch accent position, type, and status of focus projection. Language and Speech. 2003; 46:53–81. [PubMed: 14529111]

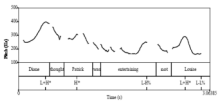

Figure 1.
Pitch track for Experiment 1, condition (a).

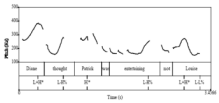

Figure 2.
Pitch track for Experiment 1, condition (b).

Table 1

Mean F0 Measurements, Experiments 1–3, in Hz, SDs in parentheses

	Matrix subject peak	L- or lowest F0 on verb 1, plus H%	Embedded subject peak	L- or lowest F0 at clause end, plus H%
Experiment 1				
a	374 (15)	269 (20)	271 (16)	159 (5), 230 (10)
b	361 (13)	158 (5), 234 (14)	272 (9)	156 (5), 230 (10)
Experiment 2A				
Matrix a	358 (22)	260 (20)	269 (15)	166 (7)
Matrix b	350 (22)	157 (6), 216 (14)	272 (12)	168 (8)
Matrix c	347 (22)	251 (19)	267 (10)	159 (7), 212 (9)
Matrix d	344 (21)	156 (7), 215 (13)	268 (9)	156 (6), 213 (9)
Embedded a	341 (22)	252 (17)	267 (12)	170 (11)
Embedded b	339 (21)	159 (6), 216 (14)	270 (12)	170 (10)
Embedded c	340 (22)	248 (18)	267 (13)	161 (8), 210 (9)
Embedded d	338 (23)	157 (7), 216 (9)	268 (11)	157 (5), 211 (10)
Experiment 2B				
Matrix a	375 (13)	256 (21)	254 (11)	158 (4), 218 (12)
Matrix b	361 (12)	157 (4), 223 (13)	262 (10)	154 (4), 218 (11)
Embedded a	366 (15)	251 (17)	254 (11)	154 (4), 219 (13)
Embedded b	358 (14)	154 (4), 219 (11)	258 (10)	151 (3), 216 (10)
Experiment 3				
a	354 (11)		188 (9)	161 (6), 230 (15)
b	213 (11)		323 (14)	156 (5), 232 (15)
c	344 (11)		254 (12)	153 (4), 211 (12)

Table 2

Mean Durations, Experiments 1–2, in ms, SDs in parentheses

	Matrix verb + pause	**Last word of embedded clause + pause**
Experiment 1		
a	366 (97)	
b	599 (105)	
Experiment 2A		
Matrix a	380 (101)	569 (93)
Matrix b	624 (97)	545 (88)
Matrix c	364 (94)	670 (71)
Matrix d	627 (96)	664 (79)
Embedded a	357 (96)	547 (95)
Embedded b	620 (97)	522 (95)
Embedded c	355 (96)	667 (87)
Embedded d	615 (93)	654 (78)
Experiment 2B		
Matrix a	383 (106)	
Matrix b	589 (90)	
Embedded a	360 (106)	
Embedded b	577 (111)	

Table 3

"Got-it" RTs, Percentage "Got-it" Responses, and Percentage Embedded answers, Experiment 2A (with 95% confidence intervals)

Correlate	Boundaries	RT	Got-it Percentage	% Embedded
Matrix	none, ip	960 (881–1039)	69.8 (63.7–75.9)	68.3 (59.5–77.2)
Matrix	IPh, ip	1020 (929–1112)	69.6 (63.2–76.1)	67.5 (57.5–77.4)
Matrix	none	966 (868–1066)	76.7 (71.0–82.5)	62.6 (52.5–72.6)
Matrix	IPh	1056 (956–1158)	71.7 (65.8–77.6)	65.3 (54.8–75.8)
Embedded	none	887 (798–976)	90.6 (84.8–96.4)	82.2 (74.0–90.5)
Embedded	IPh	826 (751–902)	92.5 (86.9–98.2)	85.3 (78.6–92.0)
Embedded	none	774 (681–868)	94.5 (89.7–99.2)	84.7 (74.5–94.9)
Embedded	IPh	839 (737–940)	95.7 (91.0–100)	90.0 (83.1–96.0)

Note: As explained in the text, the % Embedded question answers are means over items; the other means are calculated by subjects

Table 4

"Got-it" RTs and Percentage "Got-it" Responses, Experiment 2B (with 95% confidence intervals)

Correlate	Boundaries	RT	Percentage
Matrix	none, IPh	1229 (1142–1316)	68.0 (63.2–72.8)
Matrix	IPh, IPh	1289 (1174–1405)	64.2 (58.5–69.6)
Embedded	none, IPh	1228 (1110–1346)	90.9 (85.6–96.2)
Embedded	IPh, IPh	1227 (1116–1338)	89.9 (85.4–94.4)

Table 5

Percentages of Matrix Subject Choices as Correlate, and Question-answering Time (ms), Experiment 3 (with 95% Confidence Intervals)

Condition	Percentage Matrix Subject Choices	Mean Answer Time
Matrix subject accented	29.7 (25.9–33.5)	2603 (2520–2686)
Embedded subject accented	10.8 (8.3–13.3)	2504 (2414–2594)
Equal accents	10.5 (7.8–13.2)	2613 (2510–2716)