

Available online at [www.sciencedirect.com](http://www.sciencedirect.com)

Lingua

Lingua □ (□□□□) □-□

[www.elsevier.com/locate/lingua](http://www.elsevier.com/locate/lingua)

# Don't break, or do: prosodic boundary preferences

Lyn Frazier\*, Charles Clifton, Jr., Katy Carlson

*Department of Linguistics, University of Massachusetts, Amherst, MA 01003, USA*

Received 16 May 2002; received in revised form 20 February 2003

## Abstract

Four naturalness judgment experiments were conducted to test different hypotheses about prosodic phrasing. The hypothesis that syntactic constituents should not be broken into distinct prosodic phrases [as in Truckenbrodt's Wrap constraint (Truckenbrodt, H., 1995. *Phonological Phrases: Their Relation to Syntax, Focus, and Prominence*. Unpublished PhD Dissertation, MIT.)] was less predictive of the results of Experiments 1 and 2 than the hypothesis that constituents may be freely divided into prosodic phrases, as long as the resulting phrases are semantically coherent [Selkirk, E., 1984. *Phonology and Syntax: The Relation Between Sound and Structure*. MIT Press, Cambridge, MA.]. The results of two further experiments confirmed Watson and Gibson's (Watson, D. G., Gibson, E., 2001. *Linguistic structure and intonational phrasing*. Paper presented at the Fourteenth Annual CUNY Conference on Human Sentence Processing, Philadelphia, 15–17 March 2001.) claim that prosodic breaks are natural before long upcoming constituents, but did not support their hypothesis that the distance between a new item and its integration site is what motivates the presence of a prosodic phrase boundary. The results are interpreted as further evidence that the use of high level breaks in language comprehension is not governed by an invariant local mapping from syntax or processing considerations to prosody/intonation, but is related to the overall pattern of intonational choices made.

© 2003 Published by Elsevier Science B.V.

**Keywords:** Prosody; Prosodic boundary preferences; Psycholinguistics

## 1. Introduction

While speakers of English can be quite effective in prosodically phrasing their utterances to get the intended message across (see Schafer et al., 2000, for a recent

\* Corresponding author. Tel.: +1-413-5450996; fax: +1-413-5452175.

E-mail address: [lyn@linguist.umass.edu](mailto:lyn@linguist.umass.edu) (L. Frazier).

convincing demonstration), they also have a rather wide range of options about how they can divide a spoken sentence into prosodic constituents. At lower levels of the prosodic hierarchy, the conditions on forming a prosodic constituent are quite strict. For example, there is little variability with respect to the formation of a phonological word. However, above the phonological word, there appears to be a lot of optionality concerning where and if a prosodic break should occur. Linguistic analyses (Selkirk, 1984, 1999; Gussenhoven, 1999) make it clear that the syntactic structure of a sentence affects but does not dictate its prosodic phrasing, and in relatively natural production tasks like that used by Schafer et al. (2000), speakers vary in the prosodic or intonational phrasings they use.

The present research examines several different approaches to the question of what determines the appropriateness of a given choice of intonational phrasing. The first two experiments pit a syntactically-based constraint on intonational phrasing against a semantically-based constraint. The second two experiments examine factors that are motivated by processing considerations, namely the length of a constituent and the distance over which interpretative integration must take place. The outcome of the experiments suggests that multiple factors influence the naturalness of a given intonational phrasing, but also indicates that some factors are of lesser importance than others.

Experiments 1 and 2 contrast two broadly different linguistic approaches to the naturalness of prosodic breaks. One approach employs a basic constraint against breaking up syntactic phrases, inspired by Wrap. Wrap was originally proposed by Truckenbrodt (1995) as one of a small number of violable universal constraints in an optimality theory (OT) grammar. In simple terms, Wrap says that prosodic phrasings that do not break up a syntactic phrase are favored over prosodic phrasings that do, unless breaking up a syntactic phrase is required for some reason, e.g. to satisfy some higher-ranked constraint. As phrased by Truckenbrodt, Wrap does not apply to adjoined nodes or sentence nodes, though it is not clear conceptually why it should not. Our concern in the present paper is not the particular details of any prohibition against breaking up syntactic phrases, but the question of whether we can find evidence for a prohibition against breaking up syntactic phrases below the level of the root sentence node. We will call this prohibition generalized Wrap.

A semantically-based linguistic approach might claim instead that syntactic phrases may be broken into parts freely providing that the resulting parts are semantically sensible. In other words, given the syntactic constituents  $x\ y\ z$ , if  $x$  and  $y$  belong together semantically but  $y$  and  $z$  do not, one may separately phrase each of  $x$ ,  $y$ , and  $z$ , but one may not place  $x$  alone and group  $y$  and  $z$  together. If only  $x$  and  $z$  belong together semantically, one may place  $x$ ,  $y$ , and  $z$  together or each alone but one may not group  $y$  and  $z$  together without  $x$  (or  $x$  and  $y$  together, without  $z$ ). Essentially, this constraint punishes any prosodic phrasing which groups together semantically unrelated elements. We will call this the semantic coherence constraint (SCC).

The SCC was inspired by an earlier, more specific version of the semantic approach proposed by Selkirk (1984). Her Sense Unit Condition, given in (1), places a semantic constraint on the prosodic constituents resulting from breaking up a syntactic unit.

## (1) The Sense Unit Condition on Intonational Phrasing:

The immediate constituents of an intonational phrase must together form a sense unit.

Two constituents,  $C_i$ ,  $C_j$  form a sense unit if (a) or (b) is true of the semantic interpretation of the sentence:

- a:  $C_i$  modifies  $C_j$  (a head)
- b:  $C_i$  is an argument of  $C_j$  (a head).

Either the Sense Unit Condition, or the more general semantic coherence constraint discussed above, will limit the ways in which a sentence may be broken into intonational parts. Although (1) refers to an intonational phrase, it is clear from the examples discussed by Selkirk that smaller prosodic boundaries must also be regulated. Assuming Pierrehumbert and Beckman's prosodic system (Beckman and Ayers, 1993; Pierrehumbert, 1980; Pierrehumbert and Beckman, 1988), the condition must apply to intermediate phrases (ip) as well as full intonational phrases (IPh).

Neither of these semantic constraints is fully satisfactory in its present form: the SCC needs more explicitness in its description of semantic belonging, while the Sense Unit Condition's specific predictions are occasionally too strong.<sup>1</sup> Additionally, generalized Wrap and the semantic coherence constraint (SCC) based on Selkirk's (1984) Sense Unit Condition do not exhaust the space of relevant constraints. For example, constraints favoring units of particular (relative) sizes will likely also be needed to fully capture the placement of phonological boundaries. Nevertheless, the two proposed constraints make clear and in some cases conflicting predictions on their own, and we believe that empirical testing of these hypotheses will increase understanding of the prosody-syntax mapping.

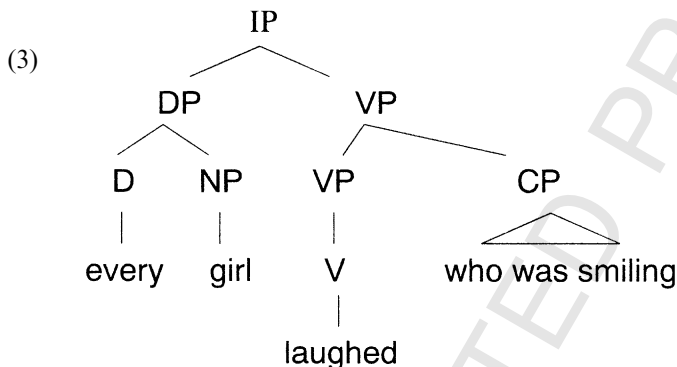
Experiments 1 and 2 contrast our constraints' predictions in four sentence types, illustrated below in (2), (4), (5) and (6), with predictions summarized in the columns to the right of the examples.<sup>2</sup> Consider (2), an extraposed relative clause sentence (Type I).

<sup>1</sup> To accommodate examples like (i), Selkirk allows just the head of a phrase and the head of its argument to form a sense unit. (i) (This is the cat) (that ate the rat.) In cases like (i) this will suffice. But with prenominal modifiers it would allow (*pretty*)(*girls left*), which does not seem fully natural. Further, Hirotani (2000) has shown that semantically unrelated moved constituents may at times prefer to occur in the same prosodic phrase in complex Japanese sentences, if the constituents are moved from the same clause. This prosodic phrasing would allow them to be held in memory as one unit, Hirotani proposes, until the perceiver receives the clause from which they have been extracted. In this case, prosodic phrasing may aid interpretation and memory, but not in ways explicitly captured by the Sense Unit condition or the semantic coherence condition.

<sup>2</sup> Caution: the columns after the examples may look similar to the tableaux of OT grammars, but that is not their status. The "Gwrap" column simply summarizes the predictions of the generalized Wrap principle for the corresponding example, e.g., (2a) violates Wrap, hence an asterisk following (2a), but (2b) does not violate it and therefore receives a "✓." (Note that the \* indicates that the principle is violated one or more times. Unlike OT, more violations do not result in more \*'s.) The predictions of the semantic constraint are summarized to the right. We treat the two principles as if they represent distinct hypotheses, e.g. distinct grammars, not different principles in the same grammar (although in fact they could co-exist in an OT grammar).

			GWrap	SCC
(2)	a.	(Every girl laughed)(who was smiling)	*	✓
	b.	(Every girl)(laughed who was smiling)	✓	*

The parentheses in (2) and the other examples in this paper stand for prosodic boundaries above the level of phonological words. In our materials, these prosodic boundaries are always intonational phrase boundaries, which end with a phrase tone (H-, L-) followed by a boundary tone (H%, L%; Pierrehumbert, 1980; Pierrehumbert and Beckman, 1988). Assuming that extraposed relatives are adjoined to VP,<sup>3</sup> the structure of (2) is given in (3):



The phrasing in (2a) breaks up the VP, violating generalized Wrap. Thus GWrap predicts that (2a) should be unacceptable but (2b) should be acceptable. The semantic coherence constraint predicts the opposite. In (2a) *every girl* and *laughed* can be related semantically, though questions could arise concerning the presence (and interpretation) of a trace left by the extraposed relative. But in (2b) the unit *laughed who was smiling* clearly violates the SCC, because *who was smiling* semantically belongs with *every girl* and not directly with *laughed*. In other words, this violates the constraint that in *x y z* if *x* and *z* belong together and *y* and *z* do not, *y* and *z* cannot form a unit on their own.

In (2) one might claim that a preference for constituents that are approximately equal in size (cf. Gee and Grosjean, 1983; Fodor, 1998), which we will call Balance, explains the intuitive preference for (2a), instead of the SCC. But consider (4), a separated particle sentence (Type II):

			GWrap	SCC	Bal
(4)	a.	(Jason)(called all the participants over)	✓	✓	*
	b.	(Jason called)(all the participants over)	*	*	✓

<sup>3</sup> Examples with VP ellipsis, like (i), suggest that an extraposed relative clause may adjoin to either VP or IP, but that VP is preferred. (i) A girl left who was laughing and so did a boy. Perceivers prefer to interpret the second conjunct of (i) as *and a boy left who was laughing*, with the relative inside the copied VP (instead of as *a boy left*, in which only the verb is in the VP and thus copied).

In (4), Balance would favor (4b). In contrast, generalized Wrap predicts a preference for (4a), because the phrasing in (4b) breaks up the VP. The semantic coherence constraint also favors (4a), because the phrase *all the participants over* in (4b) groups together the object of the verb with the verb's particle. This is a case of *y* and *z* being grouped together when *x* and *z* (or *x* and *y*) are the elements that belong together. Consequently if both (2a) and (4a) are judged more acceptable than their counterparts, only the SCC could explain the overall pattern of data.

Next consider (5), a modified subject sentence (Type III).

		GWrap	SCC
(5)	a. (A lady)(from the accounting office)(visited.)	*	✓
	b. (A lady)(from the accounting office visited.)	*	*

In (5a) and (5b), generalized Wrap is violated because the subject DP/NP has been broken up. So (5a) and (5b) are both predicted to be disfavored. By contrast, the semantic coherence constraint predicts that it is acceptable to break up the sentence into the parts in (5a). However, (5b) is predicted to be unacceptable because the PP should not appear in a unit that contains the verb but not the head noun, as it belongs with the head of the subject NP. Thus this is a case of *x y z* when only *x* and *y* belong together semantically, but *y* and *z* have been grouped together without *x*. (Similarly, the Sense Unit Condition would disfavor (5b) because a head (the verb) is grouped together with a PP that is neither its argument nor an adjunct modifying it.)

Turning to (6), a sentence complement sentence (Type IV), it is clear that the VP is broken up in both (6a) and (6b). Thus Generalized Wrap does not predict differences between them but might predict both to be disfavored.

		Gwrap	SCC
(6)	a. (John promised Mary)(that he would go)	*	✓
	b. (John promised)(Mary that he would go)	*	*

However, the semantic coherence constraint does make a distinction between them. Specifically (6b) should be unacceptable because *Mary*, which is an argument of the verb *promise*, should either occur on its own or in a unit containing the head verb *promise* rather than being phrased together with the verb's CP complement.

## 2. Experiment 1

Ten sentences of each of the above four types (extraposed relative clause, separated particle, modified subject, and sentence complement sentences) were constructed. Each was recorded twice: once with the a-prosody and once with the b-prosody (see below, under Materials). Both Experiments 1 and 2 were designed to

measure the prosodic naturalness or acceptability of the utterances. In Experiment 1, one version of each sentence was played to participants who were asked to rate the acceptability of the prosody, described as the “melody and rhythm,” of the sentence. In Experiment 2, the same items were tested but the critical sentences were presented as pairs. Participants were asked to make a relative judgment about which member of the pair sounded better and then indicate whether the less good member was or was not acceptable. Experiment 2 was conducted both as a replication and to compare the absolute judgments in Experiment 1 with presumably more natural (though more complicated) relative judgments.

## 2.1. Method

### 2.1.1. Materials

The 10 sentences of each of the four types described above appear in [Appendix A](#). Each was recorded in a sound-deadened chamber by a trained phonologist, who placed an intonational phrase boundary (IPh) at the points indicated by the parentheses in examples (2), (4), (5), and (6), and by the / marks in [Appendix A](#). Each sentence was digitized (16 bit, 22 kHz). Pitchtracks of typical utterances of sentences like those in example (2) appear in [Fig. 1](#). The intonation contours and timing they exhibit at the claimed IPh boundaries were generally typical of utterances of all four types of sentences. The prosodic analysis was verified by measuring several acoustic parameters of each word immediately preceding a boundary and the corresponding word when it did not precede an intended boundary [e.g., *girl* and *laughed* in (2)]. Specifically, we measured the duration of the word, its minimum F0 (fundamental frequency) value, and the F0 value at the end of the word. An IPh boundary in these materials is typically characterized by durational lengthening of a preceding word and by a fall then a rise in F0. [Table 1](#) provides the means of the measured parameters and indicates that the sentences generally had just this pattern where an IPh boundary was intended (compared to shorter durations and no fall-rise when no boundary was intended).

### 2.1.2. Participants and procedure

The resulting 40 sentences (with two versions of each) were separated into two counterbalanced lists, each with five instances of each version of each sentence type. They were supplemented with 44 additional unrelated sentences and two-sentence discourses (28 of which will be described in Experiments 3 and 4).

Thirty-two undergraduate students at the University of Massachusetts participated in individual half-hour sessions for course credit. Each was tested in a sound-deadened chamber. A session began with instructions that the participant was to rate each sentence or discourse for its rhythm and melody, deciding whether it was pronounced with a natural, normal, acceptable rhythm and melody. They were instructed to use a five point scale, where 5 is the best and 1 is the worst, and were told not to rate how much sense the sentence makes apart from its rhythm and melody. The subject announced the rating to the experimenter who recorded it into the computer. A six-sentence practice list was first presented by playing the digitized

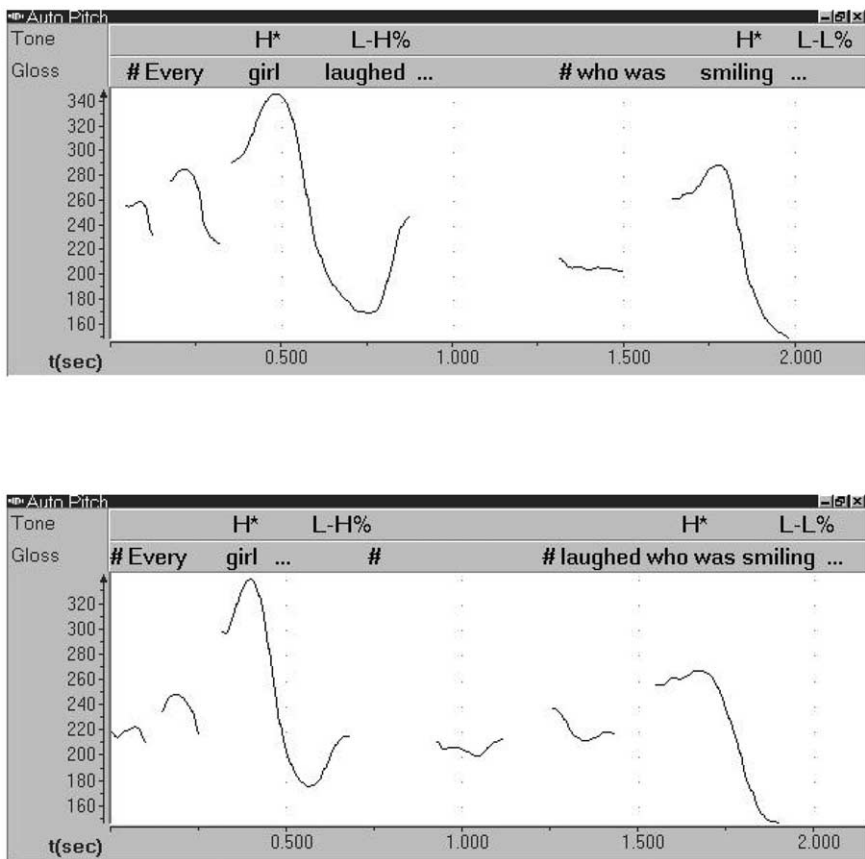


Fig. 1. Pitchtracks of typical utterances of the extraposed relative clause sentences [as in (2)] used in Experiments 1 and 2.

files, and then the experimenter answered any questions the participant had. Following that, the participant received an individually-randomized presentation of one of the two counterbalanced lists (16 participants to each list).

## 2.2. Results

The mean ratings are presented in Table 2. The version of the extraposed relative clause sentences (Type I) favored by the Semantic Coherence Constraint was rated significantly higher than the version favored by generalized Wrap [(2a) vs (2b),  $F(1,31)=20.03$ ,  $F(1,9)=20.27$ ,  $P<0.001$ ] ( $F1$  designates the results of an analysis permitting generalization to participants, and  $F2$  the results of an analysis permitting generalization to items). The version of the separated particle sentences (Type II) favored by generalized Wrap and the SCC was rated higher than the version favored by balance [(4a) vs (4b),  $F(1, 31)=52.37$ ;  $F(1, 9)=95.63$ ,  $P<0.001$ ]. The

Table 1  
Acoustic measurements on critical words (with example words), Experiment 1; mean durations (ms) and F0 values (Hz). **Boldfaced values** indicate a word with an assumed L–H% intonational boundary

Sentence group and version		Word and measure					
		Word 1			Word 2		
		Duration	F0 min	F0 end	Duration	F0 min	F0 end
I. Girl laughed							
A.		432	193	197	<b>630</b>	<b>178</b>	<b>232</b>
B.		<b>645</b>	<b>169</b>	<b>232</b>	391	193	197
II. Jason called							
A.		<b>561</b>	<b>168</b>	<b>229</b>	265	184	187
B.		372	199	200	<b>509</b>	<b>167</b>	<b>226</b>
III. Lady office							
A.		<b>584</b>	<b>175</b>	<b>231</b>	<b>512</b>	<b>170</b>	<b>231</b>
B.		<b>572</b>	<b>172</b>	<b>226</b>	305	204	208
IV. Promised Mary							
A.		358	180	180	<b>448</b>	<b>175</b>	<b>243</b>
B.		<b>572</b>	<b>171</b>	<b>247</b>	246	184	190

Table 2  
Mean sentence ratings (Experiment 1) and mean percentage preference (Experiment 2)

Sentence type with example			Expt 1	Expt 2
I.	a	(Every girl laughed)(who was smiling)	3.30	70%
	b.	(Every girl) (laughed who was smiling)	2.87	
II.	a.	(Jason)(called the participants over)	3.93	80%
	b.	(Jason called)(all the participants over)	3.05	
III.	a.	(A lady)(from the accounting office)(visited)	3.70	66%
	b.	(A lady)(from the accounting office visited)	3.70	
IV.	a.	(Josh promised Mary)(that he would go)	3.94	80%
	b.	(Josh promised)(Mary that he would go)	2.99	

two versions of the modified subject sentences (Type III) did not differ in prosodic acceptability [(5a) vs. (5b),  $F < 1$ ]. GWrap made no predictions about the relative acceptability of the sentence complement sentences (Type IV), but the version favored by the SCC was rated higher than the other [(6a) vs (6b),  $F(1, 31) = 101.48$ ,  $F(2, 9) = 49.69$ ,  $P < 0.001$ ].

2.3. Discussion

The ratings supported the predictions of the semantic coherence constraint in three cases of four, all except the Type III modified subject sentences. For the extraposed relative clause sentences and sentence complement sentences, a prosodic rendition which broke up a syntactic constituent was favored or highly acceptable, as long as the phrases thus created were semantically coherent, contrary to general-



ized Wrap. The separated particle sentences (Type II) provided no support for Balance. It is possible that participants found each version of the Type III sentences fairly acceptable (for reasons to be discussed later) and that the rating task was not refined enough to pick up subtle differences in acceptability. Experiment 2 was conducted using the same materials but a different, two-choice, technique for determining acceptability.

### 3. Experiment 2

#### 3.1. Method

Forty-eight undergraduates were tested in the same environment used in Experiment 1, and the materials used (including the filler and practice items) were the same as those used in Experiment 1. However, on each trial, a participant heard the two versions of each sentence in succession and indicated by pulling one of two response triggers whether the first or the second was the better one in terms of its rhythm and melody. Which version occurred first was randomly determined on each trial. Following the participant's response, a message appeared on a video screen asking the participant to indicate by pulling one of two triggers whether the unpreferred version was acceptable or not. The participant could re-hear the unpreferred version as often as needed by pressing a thumb button. The primary measure was the percentage of times each version was initially preferred. A secondary measure was the percentage of times each version was classified "acceptable" when it had been initially the unpreferred version.

#### 3.2. Results

Table 2 presents the percentages of time the version favored by the SCC was chosen as the more preferred one. In each case, the percentage was significantly greater than 50%, with F1 and F2 ranging from 43.26 to 212.83,  $P < 0.001$ , for sentences of Types I, II, and IV. While the two prosodic versions of the Type III sentences had not differed from each other in the rating data of Experiment 1, the forced choice response of Experiment 2 proved more sensitive, with the version favored by the SCC (6a) chosen as preferred 66% of the time, significantly greater than 50% [ $F(1,47) = 21.80$ ,  $P < 0.001$ ;  $F(1, 9) = 38.61$ ,  $P < 0.001$ ].

Table 3 presents the percentages of times that participants indicated that the version that they dispreferred on a given trial was nonetheless acceptable. For each type of item except for Type III modified subject sentences, participants were more frequently willing to accept the version favored by the Semantic Coherence Constraint than the version favored by Generalized Wrap or Balance. Varying numbers of participants (between 5 and 13) had to be eliminated from the by-subjects analyses because they never indicated that one version of a sentence type was unacceptable, which is reflected in the d.f. of the following tests: Type I:  $F(1, 40) = 3.78$ ,  $P < 0.06$ ;  $F(1, 9) = 6.34$ ,  $P < 0.04$ ; Type II:  $F(1, 41) = 6.32$ ,  $P < 0.02$ ;  $F(1, 9) = 32.30$ ,  $P$

Table 3  
Mean percentages with which initially unpreferred version was indicated to be acceptable

			Experiment 2
Sentence type with example			Percentage “acceptable” when initially disfavored
I.	a	(Every girl laughed)(who was smiling)	58
	b.	(Every girl) (laughed who was smiling)	46
II.	a.	(Jason)(called the participants over)	66
	b.	(Jason called)(all the participants over)	49
III.	a.	(A lady)(from the accounting office)(visited)	78
	b.	(A lady)(from the accounting office visited)	69
IV.	a.	(Josh promised Mary)(that he would go)	68
	b.	(Josh promised)(Mary that he would go)	47

<0.001; Type IV:  $F(1, 47) = 6.99$ ,  $P < 0.02$ ;  $F(1, 9) = 46.98$ ,  $P < 0.001$ . Both versions of the Type III modified subject sentences were generally rated as acceptable even when they were initially dispreferred, and Version a, favored by the SCC, was rated as acceptable significantly more than Version b only by items, but not by subjects:  $F(1, 42) = 2.31$ ,  $P > 0.13$ ;  $F(1, 9) = 14.65$ ,  $P < 0.01$ .

3.3. Discussion

In general, both the absolute and the relative judgments supported the predictions of the semantic coherence constraint. Breaking up syntactic constituents is fine as long as the resulting unit is semantically sensible. The data do not provide any support for generalized Wrap. In particular, (2a), (5a), and (6) show that breaking up syntactic constituents such as a VP or DP can be perfectly acceptable. The data also suggest that if balance plays a role in prosodic phrase, it must be a minor role, subordinate to the semantic conditions.

The one unexpected finding is the high absolute rating received by (5b): (A lady)(from the accounting office visited). The relative judgment data of Experiment 2 did show a clear (66%) preference for the a-form. However, the absolute rating data of Experiment 1 did not. The oddity may be due to acoustic characteristics of the materials: (5b) is so odd that the speaker may have often introduced a very subtle break before the verb [*visited* in (5b)]. However, the break was clearly not an IPh boundary in (5b), as can be seen in Table 1. There was little or no tonal movement on the word before the break (*office* in Table 1) and no evidence of lengthening of this word. Further, listening to the sentences gave little or no evidence of an ip boundary that might have been enough to “save” sentences like (5b) when a better alternative was not available. We conclude that while the difference observed in Experiment 2 does support the Semantic Coherence Constraint, other unidentified factors may also influence the acceptability of different prosodic phrasings.

The data from Experiments 1 and 2 support the spirit of Selkirk’s Sense Unit Condition, if not necessarily the specifics, and certainly support the more general

Semantic Coherence Constraint. We take these results as evidence that the theoretical notion of semantic coherence is important in the evaluation of prosodic phrasing and should be considered in future processing work. Now we turn to the effects of processing factors on prosody.

#### 4. Experiment 3

Experiments 3 and 4 examine the effect that two factors motivated by processing considerations may have on the naturalness of different prosodic phrasings. In particular, the experiments examine the approach that Watson and Gibson (2001) take to intonational phrasing. Watson and Gibson are primarily concerned with where a speaker places an intonational phrase break and suggest that just two factors are important. One is the distance between the syntactic head of the upcoming phonological phrase and the head's syntactic integration site. Longer integration distances create processing difficulty and this increases the probability of an intonational boundary between two phonological phrases.<sup>4</sup> The second factor is the eventual size (in phonological phrases) of the upcoming constituent. The larger the constituent the higher the probability of an intonational phrase boundary. The two factors are summed and the sum is claimed to predict the probability of an intonational phrase boundary.

In a production study, Watson and Gibson (2001) varied the integration distance of the main verb and the head of its subject NP by manipulating the length of an intervening relative clause. The longer the relative clause, the more likely the speaker was to place an intonational phrase boundary before the main verb. In a comprehension study, Watson and Gibson (2001) expected listeners to be sensitive to the same factors that had mattered in production. Watkins and Gibson placed an intonational boundary at point (a) in sentences like (7), before a long distance integration of *to the client* to *showed*, or at point (b), before a long relative clause.

- (7) The detective showed the blurry picture of the diamond (a) to the client  
(b) who was in the office that was on the fourteenth floor.

Listeners rated sentences with a boundary at either (a) or (b) as less complex than sentences without the boundary. Although Watson and Gibson did not test it, presumably having both breaks would also be predicted to aid comprehension.

To test Watson and Gibson's hypotheses, we investigated sentences like those in (8) where there is a long distance integration between the trace (*t*) and the interrogative *wh*-phrase.

<sup>4</sup> Later work has persuaded these authors to give up the claim that integration distance affects the appropriateness of a prosodic boundary (personal communication, D. G. Watson, March, 2002; Watson and Gibson, 2002). The results we report here are consistent with this change of position.

(8) a. (Which film ) (did Anthony say Timothy wanted Mary) (to see *t*?)

integration: 0 0 0 0 0 5

size: 6 5 4 3 2 1

b. (Which film did Anthony say) (Timothy wanted Mary to see *t*?)

integration: 0 0 0 0 0 5

size: 6 5 4 3 2 1

c. (Anthony said Timothy wanted Mary)( to see the film.)

integration: 0 0 0 0 0 0

size: 6 5 4 3 2 1

d. (Anthony said)(Timothy wanted Mary to see the film.)

integration: 0 0 0 0 0 0

size: 6 5 4 3 2 1

The numbers below specific words in (8) designate the integration and size indices. The integration index indicates how many phonological words the specific word must be integrated across to relate it to other material in the sentence.<sup>5</sup> Watson and Gibson's integration factor predicts that each phonological word has an integration cost of zero until *see*, since no integration across intervening phonological phrases is required until *to see* is encountered. At this point the trace in the position of the object of *see* must be integrated across five intervening items. The large integration index of "5" below *see* in (8a, b) increases the probability of a boundary immediately before *see*; such a boundary should be helpful to a listener, because it occurs before a difficult integration.

In contrast to the integration index, which is associated with the word after the boundary of interest, the size index appears under the word before the boundary of interest. It indicates how many phonological words the upcoming syntactic constituent contains. A boundary after *which film*, which appears in (8a), should help processing because the constituent *did Anthony say Timothy wanted Mary to see* is long, containing 6 phonological words. Therefore the two boundaries in (8a) are at just the right places, and (8a) should be preferred to (8b). By contrast, the boundary in (8b) is at a position where the integration index is zero, instead of 5 at the second break in (8a), and the size index is 4 instead of 6.

We also tested the declarative counterparts of (8a and b), namely, (8c and d). The integration index does not choose between the phrasings in (8c) and (8d), since no long-distance integration is necessary in either sentence. However, the size index will

<sup>5</sup> The term "phonological word" in this context designates each word except *did*, *to*, *the* and possessive pronouns, following Watson and Gibson.

favor the phrasing in (8d) because the break in (8d) coincides with the beginning of a larger constituent than the break in (8c)

Subject questions like those in (9) were also tested.

(9)	a.	(Which assistant did Marcia announce) (t hated his job?)			
integration:		0	0	2	0
size:	4	3	2	1	
	b.	(Which assistant) (did Marcia announce t hated his job?)			
integration:		0	0	2	0
size:	4	3	2	1	

The trace (subject of *hated*) in (9) must be integrated across two phonological words and thus a break before the verb (*hated*) should be helpful. For the size parameter, however, the longest upcoming constituent is the *C'* constituent (*did Marcia announce hated his job*). If both size and integration contribute to boundary preferences (as claimed by Watson and Gibson, 2001), then summing the indices under (9) makes it clear that (9b), where a boundary occurs at a position with a summed index of 4, should be similar to (9a), where the boundary occurs at a position which also has an index of 4.<sup>6</sup> However, if only size and not integration affect the appropriateness of a boundary (see footnote 4), then (9b) should be preferred to (9a).<sup>7</sup>

Although sentences (8) and (9) were not designed to provide strong tests of the principles examined in Experiments 1 and 2 (generalized Wrap, SCC, and balance), these principles do apply to the sentences. Sentences (8a) and (8c) can be argued to violate generalized Wrap, in that they both contain a prosodic boundary that places part of a sentence complement into one prosodic unit and the remainder into a different unit. Sentence (8b) seems to violate the SCC in that it separates the verb of the complement sentence from its (wh-moved) object without separating this moved phrase from material it is unrelated to. In other words, *which film did Anthony say* does not form a coherent unit. Sentence (8b), however, is most consistent with the balance principle. Sentence (8d) is consistent with both generalized Wrap and SCC. A similar analysis indicates that (9a) is consistent with generalized Wrap but not SCC, while the opposite holds for (9b).

<sup>6</sup> Six of the eight items in this subexperiment contained four phonological words in the *C'*. One item contained five phonological words in *C'* and one contained six. The predictions of Watson and Gibson's summed index do not differ for these items.

<sup>7</sup> Nagel et al. (1994) proposed that wh-traces, including subject traces, are marked prosodically by long durations and large pitch excursions on the word preceding the trace. However, subsequent research strongly suggests that the effects they observed are in fact due to the presence of intonational phrase boundaries that happened to occur at the position of the trace. In a production study, Straub et al. (2001) showed that traces give rise to long durations and large pitch excursions only when they co-occur with an intonational phrase boundary, in which case the boundary alone could explain the effects. When a trace occurred internal to an intonational phrase, neither long durations nor large pitch excursions were observed on the pre-trace word.

#### 4.1. Method

Twenty sentences like (8) were constructed and recorded in four versions each (question vs. declarative, each with two prosodic phrasings). Eight sentences like those in (9) were recorded, in two versions each. The same speaker who recorded the materials of Experiments 1 and 2 recorded the Experiment 3 (and Experiment 4) materials in the same setting. She introduced an IPh boundary at the points marked by parentheses in (8) and (9) and by ‘/’ in [Appendix B. Figs. 2 and 3](#) present pitchtracks of typical utterances of all six sentence types. The presence of fall-rise contours and lengthening/pausing at IPh boundaries can be clearly observed.

The experiment was run concurrently with Experiment 1 and the procedures used were identical except that Experiment 3 required four different counter-balanced lists to permit each version of each sentence like (8) to be tested equally often.

#### 4.2. Results

The results are presented in [Table 4](#). Turning first to the object questions (8a,b) and their declarative counterparts (8c,d), the declarative with a prosodic boundary after the matrix verb (8d) was rated much better than the other three forms (8a,b,c), resulting in a significant interaction between question/declarative and boundary position [ $F(1, 31) = 24.92$ ;  $F(1, 19) = 27.83$ ,  $P < 0.001$ ] as well as significant main effects of both factors. Although the Watson and Gibson size hypothesis is consistent with the high rating for (8d), their integration measure is not consistent with the comparable ratings for (8a) and (8b). (8a) was predicted to be better than (8b), with breaks at two positions with summed indices totaling 6 vs. a break at one position with a 4 index.

Turning to (9), both (9a) and (9b) received a relatively high rating (3.60 and 3.84 respectively) and they differed significantly from each other only in the analysis by participants [ $F(1, 31) = 4.54$ ,  $P < 0.05$ ;  $F(1, 9) = 1.80$ ,  $P > 0.20$ ]. The direction of the difference suggests that boundaries are favored before long constituents, supporting the size hypothesis, but this result cannot independently disconfirm the prediction of equal acceptability for (9a) and (b) which was given by the integration plus size theory.

### 5. Experiment 4

#### 5.1. Method

The final experiment used the relative naturalness judgment task of Experiment 2 to test the sentences examined in Experiment 3. The experiment was conducted concurrently with Experiment 2, and used the materials from Experiment 3, excluding the declarative versions of (8).

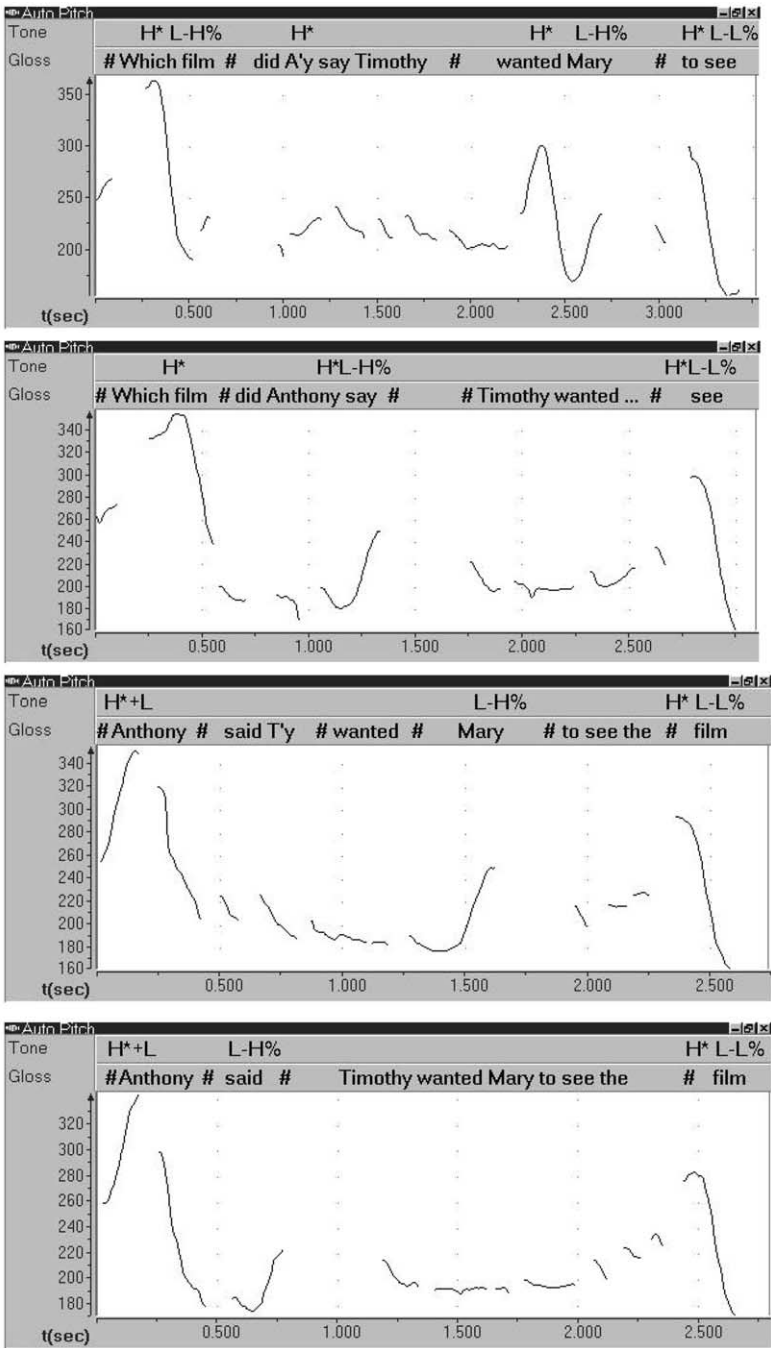


Fig. 2. Pitchtracks of typical utterances of Experiment 3 and 4 sentences as in (8).

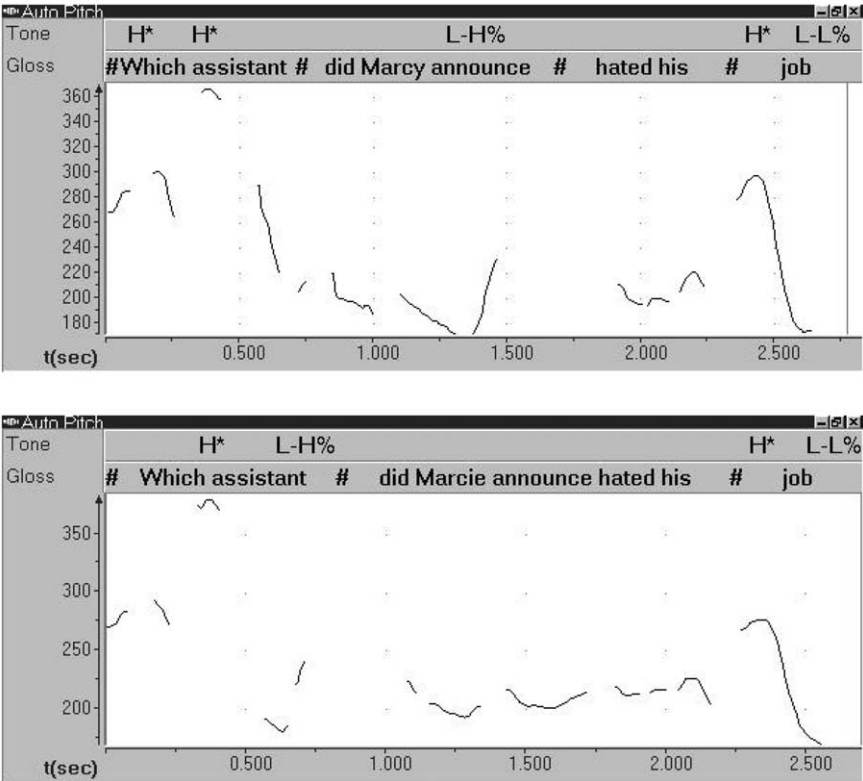


Fig. 3. Pitchtracks of typical utterances of Experiment 3 and 4 sentences as in (9).

5.2. Results

Table 4 presents the percentages of times that (8a) was chosen over (8b), and that (9a) was chosen over (9b). The former percentage was not significantly different from 50% [ $F(1, 47) = 2.11, P > 0.15$ ;  $F(1, 19) = 1.94, P > 0.15$ ]. The latter percentage, however, was significantly less than 50%, both by participants and items [ $F(1, 47) = 22.44$ ;  $F(1, 7) = 32.91, P < 0.001$ ]. This preference for the phrasing in

Table 4  
Mean sentence ratings (Experiment 3) and mean percentage preference (Experiment 4)

		Expt 3	Expt 4
(8)	a. (Which film)(did A say T wanted M)(to see ____?)	2.89	47%
	b. (Which film did A say)(T. wanted M to see ____?)	2.86	
	c. (A said T wanted Mary)(to see the film)	2.82	–
	d. (A said)(T wanted Mary to see the film)	3.58	
(9)	a. (Which assistant did Marcia announce)(____ hated his job?)	3.60	36%
	b. (Which assistant)(did Marcia announce ____ hated his job?)	3.84	



(9b) over that in (9a) reinforces the tendency observed in Experiment 3 for a boundary to be preferred before a long constituent.

The secondary response measure gathered in Experiment 4, the percentage of times the version that a participant indicated to be dispreferred was nonetheless rated as acceptable, did not differ between items (8a) and (8b) [49 vs 54%,  $F(1, 47) = 2.27$ ,  $P > 0.10$ ;  $F(1, 19) < 1$ ]. However, there was a suggestion that participants more frequently indicated accepting (9b) when they had initially preferred (9a) than the opposite, 79 vs. 66% of the time [ $F(1, 43) = 6.52$ ,  $P < 0.02$ ;  $F(1, 7) = 3.91$ ,  $P < 0.10$ ], providing some additional evidence that (9b) was favored over (9a).

### 5.3. Discussion

The clear superiority of (8d) over (8a–c) is congruent with the fact it is the only version that is consistent with both generalized Wrap and the semantic coherence constraint, but it also fits with Watson and Gibson's (2001) claim that length considerations favor breaks before long upcoming constituents. The tendency for (9b) to be preferred over (9a) is also consistent with the claim that boundaries are favored before long constituents. The results give no support to the proposal that prosodic boundaries are also favored before a point where integration is required over a long distance. They also suggest that Balance, which had favored (8b), is at most a weak principle.

## 6. General discussion

We have examined the role of several principles in characterizing the intonational phrasing of sentences: generalized Wrap (don't break up constituents); semantic coherence constraint (don't form semantically incoherent groups); balance (favor equi-sized units); integration (break before long distance integrations); and length of upcoming constituent (break before long upcoming constituents).

The data do not provide evidence supporting generalized Wrap or the Integration principle, and Balance must be a low ranked principle as indicated by the results for both (4b) and (8). The semantic coherence constraint and the importance of length of an upcoming constituent are clearly supported. However both principles need to be formulated more precisely. The semantic coherence constraint depends on a notion of what goes together semantically, without providing an explicit definition of the notion. The length principle also requires further attention. Intuitions suggest that the relative length of prosodic phrases may play a role in determining the effective weight of a phrase, and that focus may contribute to weight (Arnold et al., 2000). Further, the interaction of the SSC and length principles must be explored.

It is possible to take our data as arguing against the existence of any prosodic constraint (like Wrap) which prohibits breaking a syntactic constituent into prosodic phrases. Limiting attention to the higher level prosodic categories (intermediate phrases and intonational phrases), though, there is nothing intrinsically better about avoiding phrase breaks than having phrase breaks. In our experimental sentences,

pronunciations in which breaks occurred were rated as being very natural (e.g. 3–4 on a 5 point scale) even if a break occurred internal to a constituent. This holds even for sentences presented out of the blue, without narrow focus which might have motivated a (following) phrase break, and even for relatively short sentences that might in principle have occurred without a higher level boundary. What matters is not the simple presence or absence of a break, but whether the resulting prosodic phrases are semantically coherent. Sentences in which strings of constituents which do not semantically belong together are prosodically grouped together are rated as less natural than sentences with breaks creating semantically coherent groups.

The results of auditory comprehension studies (Carlson et al., 2001; Clifton et al., in press) have led us to argue that absolute break size (intermediate phrase versus intonational phrase) is not important to the listener's interpretation of a sentence. Instead, the listener is concerned with relative break size: a break is informative to the extent that it is, say, larger than some other break in the sentence. The present results fit well with the conclusions of this earlier work: regularities in prosodic phrasing appear to be concerned not with purely local properties, such as this particular placement of a boundary, or this particular kind of boundary, etc., but instead they are concerned with the patterns resulting from particular intonational choices. What's informative about optional higher level prosodic breaks is not whether they occur or what size the break is, but where else they do or do not occur (and thus what semantically related or unrelated elements are grouped together) and the relative size of those breaks. This level of abstract informativity, interpreting a speaker's intonational decisions relative to the speaker's other intonational decisions in the utterance, makes sense in light of the fact that length is important to the articulatory system (which is parasitic on our respiratory mechanisms) but not to the syntactic/semantic systems. In other words, invariant intonational translation of particular syntactic phrases would be difficult given that the phrase in question might be one syllable long or 25. Instead of an invariant intonational interpretation of a particular syntactic or semantic structure, what human languages seem to have is a flexible system allowing considerable optionality in the prosodic phrasing of a sentence (see Schafer et al., 2000). But flexibility does not necessarily tolerate irrationality or arbitrariness. The particular options chosen by the speaker must be coherent in light of that speaker's other choices and the speaker's intent.

The current results together with already known facts about low level phonological phrasing (into phonological words or phonological phrases) suggest that a break is prohibited not by any general constraint against breaking up syntactic constituents, but simply by constraints on phonological words/phrases as well as constraints favoring semantically coherent groups. Otherwise a break may freely occur, and it will have interpretive consequences relative to the existence, size, and location of other breaks. Long phrases are of particular interest in this regard because on the one hand they may motivate the speaker to take the option of breaking in some location, but on the other, length per se may lessen the informativity of a break, since phonological rather than syntactic or semantic reasons may justify the existence of the break. We are currently testing this 'reduced informativity' hypothesis.

This approach to prosodic phrasing makes the prediction that the results of almost any prosodic boundary may be salvaged (unless it violates low level phonological constraints, for example by forcing a function word to be a phonological phrase by itself). Basically, a break should be bad only if it creates semantically incoherent groups, which may be salvaged by further breaks, or if the break results in a misleading pattern, e.g., a break at a small syntactic boundary but no break at a larger or higher syntactic boundary, which could also be salvaged by further breaks.

This is illustrated below for example (2b), repeated as (11).

(11) (Every girl) (laughed who was smiling)

If the second prosodic phrase is divided into (*laughed*) and (*who was smiling*), the sentence is predicted to be natural. Our intuitions support the prediction.

Let us consider some of the other dispreferred prosodic conditions, such as example (3b), repeated as (12):

(12) (Jason called) (all the participants over).

Presumably the offensive prosodic phrase in (12) is *all the participants over*. This in principle might be broken up. However, a short function word like *over* does not readily stand as a phonological phrase by itself. Consequently, this is one prosodic phrasing that may not easily be salvaged.

In contrast, though, see sentence (13), which repeats (5b).

(13) (Josh promised) (Mary that he would go.)

The break which creates the semantically incoherent phrase *Mary that he would go* should be salvageable if another break is introduced after *Mary*. This may be somewhat unnatural given the shortness of *Mary* but seems considerably better if we replace *Mary* with a long name: (*Josh promised*) (*Signora Carminati*) (*that he would go.*). Again intuitions seem to support the prediction. In short, even with our dispreferred sentences, it is not one particular break that is unnatural but the overall prosodic pattern it creates.

## Acknowledgements

This research was supported in part by NSF Grant BCS-0090674, "Prosody in Sentence Comprehension," to the University of Massachusetts. Katy Carlson is now at Northwestern University.

## Appendix A. Sentences used in Experiments 1 and 2

Note: A ‘/’ indicates an intonational phrase boundary. In each case, the version predicted in the text to be preferred appears first.

### I. Extraposed relative clause sentences

1. Every girl laughed / who was smiling; Every girl / laughed who was smiling.
2. No student complained / who was passing; No student / complained who was passing.
3. Each customer replied / who lived in the area; Each customer / replied who lived in the area.
4. Lots of patients returned / who didn't respond to the treatment; Lots of patients / returned who didn't respond to the treatment.
5. Many players retired / who had serious injuries; Many players / retired who had serious injuries.
6. Some scientists quit / who objected to the project; Some scientists / quit who objected to the project.
7. An assistant lied / who was afraid of the police; An assistant / lied who was afraid of the police.
8. Every child tripped / who was jumping rope; Every child / tripped who was jumping rope.
9. Many actors starved / who were very talented; Many actors / starved who were very talented.
10. No employee protested / who had children; No employee / protested who had children.

### II. Separated particle sentences

1. Jason / called all the participants over; Jason called / all the participants over.
2. Alice / asked every applicant back; Alice asked / every applicant back.
3. Martin / paid several lenders back; Martin paid / several lenders back.
4. Sally / invited the friendly neighbors over; Sally invited / the friendly neighbors over.
5. Molly / rowed the campers across; Molly rowed / the campers across.
6. Anthony / brought the lawn mower back; Anthony brought / the lawn mower back.
7. Lydia / passed the parsley-potatoes around; Lydia passed / the parsley-potatoes around.
8. All the kids / carried the grocery bags in; All the kids carried / the grocery bags in.
9. The angry actor / pushed the props over; The angry actor pushed / the props over.
10. The director / sent the notice around; The director sent / the notice around.

### III. Modified subject sentences

1. A lady / from the accounting office / visited; A lady / from the accounting office visited.
2. The electrician / from the physical plant / stopped by; The electrician / from the physical plant stopped by.
3. The contractor / for the housing project / was sued; The contractor / for the housing project was sued.
4. A waitress / at the coffee shop / ran away; A waitress / at the coffee shop ran away.
5. A tailor / at the boutique / quit; A tailor / at the boutique quit.
6. The nurse / at the clinic / helped Maria; The nurse / at the clinic helped Maria.
7. A lawyer / from the big law firm / was indicted; A lawyer / from the big law firm was indicted.
8. A garbage man / in San Francisco / was killed; A garbage man / in San Francisco was killed.
9. The mayor / of New York / should resign; The mayor / of New York should resign.
10. The swimmers / near the pier / shouted; The swimmers / near the pier shouted.

### IV. Sentence complement sentences

1. Josh promised Mary / that he would go; Josh promised / Mary that he would go.
2. Max assured Megan / that he'd show up; Max assured / Megan that he'd show up.
3. Anita persuaded Lou / that she could dance; Anita persuaded / Lou that she could dance.
4. Angela convinced Freddie / that she would help; Angela convinced / Freddie that she would help.
5. Martin promised Lisa / that he would change; Martin promised / Lisa that he would change.
6. Ian told Sue / that he resigned; Ian told / Sue that he resigned.
7. Marta warned Fritz / that she was sick; Marta warned / Fritz that she was sick.
8. Paula promised Ken / that she'd excel; Paula promised / Ken that she'd excel.
9. Niki persuaded Tom / that she worked hard; Niki persuaded / Tom that she worked hard.
10. Daniel assured Lucy / that he told the truth; Daniel assured / Lucy that he told the truth.

**Appendix B. Sentences used in Experiments 3 and 4***Sentences of type (8)*

1. Which film / did Anthony say Timothy wanted Mary / to see?  
Which film did Anthony say / Timothy wanted Mary to see?  
Anthony said Timothy wanted Mary / to see the film.  
Anthony said / Timothy wanted Mary to see the film.
2. Which house / did the realtor report that Misha intended / to buy?  
Which house did the realtor report / that Misha intended to buy?  
The realtor reported that Misha intended / to buy the house.  
The realtor reported / that Misha intended to buy the house.
3. Which bicycle / did Ma say Fred thought Billy / should ride?  
Which bicycle did Ma say / Fred thought Billy should ride?  
Ma said Fred thought Billy / should ride the bicycle.  
Ma said / Fred thought Billy should ride the bicycle.
4. Which book / did Lucy indicate that Tom said Maria / hated?  
Which book did Lucy indicate / that Tom said Maria hated?  
Lucy indicated that Tom said Maria / hated the book.  
Lucy indicated / that Tom said Maria hated the book.
5. Which experiment / did Tim hear Chuck say Anita / designed?  
Which experiment did Tim hear / Chuck say Anita designed?  
Tim heard Chuck said Anita / designed an experiment.  
Tim heard / Chuck said Anita designed an experiment.
6. Which trip / did the travel agent complain that she hated / to plan?  
Which trip did the travel agent complain / that she hated to plan?  
The travel agent complained that she hated / to plan the trip.  
The travel agent complained / that she hated to plan the trip.
7. Which rug / did the salesman pretend he wanted his wife / to keep?  
Which rug did the salesman pretend / he wanted his wife to keep?  
The salesman pretended he wanted his wife / to keep the rug.  
The salesman pretended / he wanted his wife to keep the rug.
8. Which coffee / did Barbara claim Bill said Tim / wouldn't drink?  
Which coffee did Barbara claim / Bill said Tim wouldn't drink?  
Barbara claimed Bill said Tim wouldn't / drink the coffee.  
Barbara claimed / Bill said Tim wouldn't drink the coffee.
9. Which report / did the chemist announce the editor indicated the journal / wouldn't print.  
Which report did the chemist announce / the editor indicated the journal wouldn't print.  
The chemist announced the editor indicated the journal / wouldn't print the report.  
The chemist announced / the editor indicated the journal wouldn't print the report.

10. Which tea leaves / did the palmist say she was sure the gypsy / wouldn't read?  
Which tea leaves did the palmist say / she was sure the gypsy wouldn't read?

The palmist said she was sure the gypsy / wouldn't read the tea leaves.

The palmist said / she was sure the gypsy wouldn't read the tea leaves.

11. Which assignment / did the students announce they didn't have enough time /  
to complete?

Which assignment did the students announce / they didn't have enough time to  
complete?

The students announced they didn't have enough time / to complete the assign-  
ment.

The students announced / they didn't have enough time to complete the assign-  
ment.

12. Which billboard / did the worker think the manager was wrong / to approve?  
Which billboard did the worker think / the manager was wrong to approve?

The worker thought the manager was wrong / to approve the billboard.

The worker thought / the manager was wrong to approve the billboard.

13. Which cat / did Elizabeth know Josh saw Meghan / feed?

Which cat did Elizabeth know / Josh saw Meghan feed?

Elizabeth knew Josh saw Meghan / feed the cat.

Elizabeth knew / Josh saw Meghan feed the cat.

14. Which rumor / did Max claim the reporter found out / Sam started?

Which rumor did Max claim / the reporter found out Sam started?

Max claimed the reporter found out Sam / started the rumor.

Max claimed / the reporter found out Sam started the rumor.

15. Which swing set / did Ana say the neighbor claimed Fred / fixed?

Which swing set did Ana say / the neighbor claimed Fred fixed?

Ana said the neighbor claimed Fred / fixed the swing set.

Ana said / the neighbor claimed Fred fixed the swing set.

16. Which quartet / did the newspaper report the University refused / to hire?

Which quartet did the newspaper report / the University refused to hire?

The newspaper reported the University refused / to hire the quartet.

The newspaper reported / the University refused to hire the quartet.

17. Which scandal / did Kathy announce the supervisor told her / to believe?

Which scandal did Kathy announce / the supervisor told her to believe?

Kathy announced the supervisor told her / to believe the scandal.

Kathy announced / the supervisor told her to believe the scandal.

18. Which medicine / did Lisa say her doctor claimed Mr. Wu / invented?

Which medicine did Lisa say / her doctor claimed Mr. Wu invented?

Lisa said her doctor claimed Mr.Wu / invented the medicine.

Lisa said / her doctor claimed Mr.Wu invented the medicine.

19. Which chair / did Lynne proudly announce her husband said Timmy /  
repaired?

Which chair did Lynne proudly announce / her husband said Timmy repaired?

Lynne proudly announced her husband said Timmy / repaired the chair.

Lynne proudly announced / her husband said Timmy repaired the chair.

20. Which concert / did the report indicate the promoters now refused / to back?  
 Which concert did the report indicate / the promoters now refused to back?  
 The report indicated that the promoters now refused / to back the concert.  
 The report indicated / that the promoters now refused to back the concert.

### *Sentences of type (9)*

1. Which assistant did Marcie announce / hated his job?  
 Which assistant / did Marcie announce hated his job?
2. Which report did Maxwell claim / was false?  
 Which report / did Maxwell claim was false?
3. Which wounded bird did Ellen think / would live?  
 Which wounded bird / did Ellen think would live?
4. Which recording did Rachel claim / was lost?  
 Which recording / did Rachel claim was lost?
5. Which pen did Karen think / lasted longest?  
 Which pen / did Karen think lasted longest?
6. Which neighbor did Mrs. Grimshaw indicate / was sick?  
 Which neighbor / did Mrs. Grimshaw indicate was sick?
7. Which oak tree did Thomas assert / was 200 years old?  
 Which oak tree / did Thomas assert was 200 years old?
8. Which guest did Carey suspect / hated cherries?  
 Which guest / did Carey suspect hated cherries?

### References

- Arnold, J.E., Wasow, T., Losongco, A., Ginstrom, R., 2000. Heaviness vs. newness: the effects of structural complexity and discourse status on constituent ordering. *Language* 76, 28–57.
- Beckman, M.E., Ayers, G.M., 1993. Guidelines for ToBI Labelling, Version 2.0. Ohio State University.
- Carlson, K., Clifton, Jr., C., Frazier, L., 2001. Prosodic boundaries in adjunct attachment. *Journal of Memory and Language* 45, 58–81.
- Clifton, Jr., C., Carlson, K., Frazier, L. Informative prosodic boundaries. *Language and Speech* (in press).
- Fodor, J.D., 1998. Learning to parse? *Journal of Psycholinguistic Research* 27, 285–319.
- Gee, J., Grosjean, F., 1983. Performance structure: a psycholinguistic and linguistic appraisal. *Cognitive Psychology* 15, 411–458.
- Hirotnani, M., 2000. Prosodic constraints on Japanese scrambling. Poster presented at the 13th Annual CUNY Conference on Human Sentence Processing, San Diego, 30 March–1 April, 2000.
- Nagel, H.N., Shapiro, L.P., Naway, R., 1994. Prosody and the processing of filler-gap sentences. *Journal of Psycholinguistic Research* 23, 473–485.
- Pierrehumbert, J. B., 1980. The Phonology and Phonetics of English Intonation. Unpublished PhD dissertation, MIT.
- Pierrehumbert, J., Beckman, M., 1988. *Japanese Tone Structure*. MIT Press, Cambridge, MA.
- Schafer, A.J., Speer, S.R., Warren, P., White, S.D., 2000. Intonational disambiguation in sentence production and comprehension. *Journal of Psycholinguistic Research* 29, 169–182.



- 1 Selkirk, E., 1984. *Phonology and Syntax: The Relation Between Sound and Structure*. MIT Press, Cam-  
2 bridge, MA.
- 3 Straub, K., Wilson, C., McCollum, C., Badecker, W., 2001. Prosodic structure and wh-Questions. *Journal*  
4 *of Psycholinguistic Research* 30, 379–394.
- 5 Truckenbrodt, H., 1995. *Phonological Phrases: Their Relation to Syntax, Focus, and Prominence*.  
6 Unpublished PhD Dissertation, MIT.
- 7 Watson, D. G., Gibson, E., 2001. Linguistic structure and intonational phrasing. Paper presented at the  
8 Fourteenth Annual CUNY Conference on Human Sentence Processing, Philadelphia, 15–17 March  
9 2001.
- 10 Watson, D. G., Gibson, E., 2002. When does prosody influence parsing? Paper presented at the Fifteenth  
11 Annual CUNY Conference on Human Sentence Processing, New York, 21–23 March 2002.
- 12
- 13
- 14
- 15
- 16
- 17
- 18
- 19
- 20
- 21
- 22
- 23
- 24
- 25
- 26
- 27
- 28
- 29
- 30
- 31
- 32
- 33
- 34
- 35
- 36
- 37
- 38
- 39
- 40
- 41
- 42
- 43
- 44
- 45