INTONATIONAL PHRASING AND CONSTITUENCY IN LANGUAGE PRODUCTION AND COMPREHENSION[*]

Duane Watson & Edward Gibson

Abstract. In this paper, we investigate the relationship between intonational phrase boundaries and syntax in production and comprehension. Watson & Gibson (2004a) proposed that the likelihood of an intonational boundary at a word boundary is a function of 1) the size of the most recently completed constituent and 2) the size of the upcoming constituent if it is not an argument of the most recent head. We explore an alternate hypothesis to (1): the distance of integration between an upcoming word and its attachment site in the sentence influences the likelihood of an intonational boundary. In a production experiment (Experiment 1), we found that speakers' intonational phrasing patterns were consistent with Watson & Gibson's original hypothesis. We then present two studies of sentence comprehension using converging paradigms which test the hypothesis that listeners exploit the relationship between syntax and the production-based intonational boundaries by using a special parsing heuristic. In particular, we propose that listeners prefer not to attach incoming words to lexical heads that are followed by an intonational boundary. This hypothesis is validated in both Experiments 2 and 3.

1. Introduction

It has long been recognized that a relationship exists between the location of intonational boundaries and the location of syntactic boundaries in speech. Much research in this area stems from the literature in linguistics, which has described this relationship in terms of the language users' grammatical knowledge (Truckenbrodt 1999; Selkirk 1984, 1986; Nespor & Vogel 1987; Chomksy & Halle 1968; Hirst 1993). The claim is that a speaker's placement of prosodic boundaries is dependent on a set of grammatical rules or optimality theoretic constraints that restrict the range of grammatical locations for intonational boundaries. A problem with this approach is that boundaries have a much broader distribution than is often acknowledged (Shattuck-Hufnagel & Turk 1996), potentially occurring almost anywhere in a sentence. In addition, placement of intonational boundaries seems to be highly optional such that a given syntactic structure can have a number of equally acceptable intonational phrasings.

In this paper we argue that preferences for intonational phrasing depend in part on processes related to language production and comprehension. In particular, we propose that intonational boundaries are partly a product of planning and recovery processes by the sentence production mechanism and are therefore likely to occur before and after large constituents. Furthermore, we propose that listeners are sensitive to this relationship between intonational phrasing and syntax in production, and use this knowledge to make attachment decisions in comprehension. One novel aspect of this proposal is that the underlying processes involved with preferences in intonational phrasing in production and comprehension are proposed to be different. Thus, different methods are required for studying intonational phrasing acceptability in these different domains.

First we will present work from a production study that expands on a proposal by Watson & Gibson (2004a) and suggests that speakers prefer to place intonational boundaries before and after large syntactic constituents. We propose that this regularity in boundary placement in production is exploited by listeners to make attachment decisions. We test the predictions of this hypothesis in two comprehension studies.

2. Intonational boundaries and production

Researchers such as Cooper & Paccia-Cooper (1980), Gee & Grosjean (1983) and Ferreira (1988) have argued that intonational boundaries are partly a result of processes in production. They developed algorithms that predict likely locations for intonational boundaries within an utterance based on the utterance's syntactic structure. These models are quite good at predicting where intonational boundaries tend to occur in a sentence. However, a drawback to these models is that they tend to have a large number of steps and parameters and are quite complicated.

Watson & Gibson (2004a) argue that these algorithms are successful in so far as they capture two generalities about the role of intonational boundaries in production. The first is that speakers use the time provided by intonational boundaries to plan the upcoming structure of a sentence. The second is that speakers also require time to recover after expending large amounts of resources on producing complex constituents. Watson & Gibson (2004a) operationalize these notions of planning and recovery by assuming that these factors are related to constituent size. Thus, the production mechanism requires more time for planning and recovery before and after long constituents. Consequently, word boundaries that follow or precede large constituents are predicted to be likely locations for intonational boundaries. We formalized this in (1):

(1) The Left/Right Boundary (LRB) Hypothesis: The likelihood of an intonational boundary occurring at a word boundary is a function of:
    (a) the size of the most recently completed syntactic constituent.
    (b) the size of the upcoming syntactic constituent if it is not an argument of the most recent head.

Watson & Gibson (2004a) pointed out that the size of syntactic constituents is an implicit factor in the Gee & Grosjean (1983), Cooper & Paccia-Cooper (1980), and Ferreira (1988) algorithms. Furthermore, formulating an intonational-boundary placement theory in terms of constituent size reduces the number of parameters that are required. Moreover, the syntactic properties of the model are general enough to apply to a wide range of syntactic frameworks.

Evidence from the experimental literature support these hypotheses (Schafer et al. 2001; Snedeker & Trueswell 2003; Price et al. 1991; Ferreira 1991). To take one example, Kraljic & Brennan (2005) analyzed productions in a task where a participant instructed another participant to move objects around on a display. Kraljic & Brennan found that in (2a) speakers were more likely to place an intonational boundary between "dog" and "in", a point that corresponds with the beginning of a relatively long prepositional phrase. In contrast, in (2b), speakers tended to place an intonational boundary between "basket" and "on", a word boundary that coincides with the end of the prepositional phrase "the dog in the basket".

(2) a. Put the dog [in the basket on the star]$_{PP}$.
    b. Put [the dog in the basket]$_{PP}$ on the star.

With the two factors (1) in mind, Watson & Gibson compared the predictions of the LRB against those of Gee & Grosjean (1983); Cooper & Paccia-Cooper (1980); and Ferreira (1988) on nine different syntactic structures, and found that the LRB did as well as the previous theories, and in some instances outperformed the other models in spite of having fewer parameters.

In this paper, we further test one of the components of the LRB, namely that speakers place intonational boundaries at the end of a long constituent. Watson & Gibson (2001) proposed a plausible alternate hypothesis to the claim that speakers place intonational boundaries at the end of long constituents. Namely, that speakers tend to produce an intonational boundary when an incoming word does not attach locally to the pre-boundary word but must be integrated with a head mentioned earlier in the utterance. Speakers might place intonational boundaries before points of long-distance integration because these points are associated with higher complexity. Research in reading suggests that readers slow down when reading words that integrate with a non-local head (Gibson, 1998; Grodner & Gibson, 2005). It is possible that this operation is complex in production as well, and that an intonational boundary at this point helps to alleviate difficulty for the speaker.

For example, Watson & Gibson (2004a) found that speakers were more likely to place an intonational boundary before "fired" in (3b) than (3a), and hypothesized that this difference was caused by this word boundary coinciding with the end of a larger subject in (3b) than (3a).

(3) a. The judge who the reporter ignored fired the secretary.
    b. The judge who the reporter for the newspaper ignored fired the secretary.

However, in a right branching language such as English, longer integration distances are correlated with the boundaries of longer completed syntactic constituents. It is possible that speakers place more boundaries before the matrix verb "fired" in (3b) because it must be integrated to the noun "judge" over more material in (3b) than (3a).

One way to distinguish between these hypotheses is to decouple the link between the size of a completed constituent and the distance of integration of an upcoming word. One way of doing this is by manipulating the integration distance while controlling the size of the completed constituent as in (4):

(4) a. Right Branching Subject:
       The secretary of the successful corporate executive was incompetent.
    b. Left Branching Subject:
       The successful corporate executive's secretary was incompetent.

In (4a), the verb "was" undergoes a long distance integration with the head noun of the subject NP "secretary". In (4b) this integration is local. Thus, the integration hypothesis predicts that speakers are more likely to place a boundary before "was" in (4a) than (4b) because of the difference in integration distance. If the recent constituent size hypothesis is correct, than there should be no difference between (4a) and (4b) because the subject NPs are roughly the same size. We test these hypotheses in Experiment 1.

2.1. Experiment 1

2.1.1. Method

2.1.1.1. Participants
A total of 9 pairs of participants from the MIT community participated in the study for $5.00 each.

2.1.1.2. Materials
Sixteen items similar in structure to (4) were constructed for the purpose of this experiment. The items in the local integration condition all consisted of subject head nouns modified by a noun with genitive case that was modified by two adjectives. In the non-local integration

condition, the head noun of the subject was modified by a prepositional phrase that contained a noun modified by two adjectives. These sixteen items were randomly mixed with forty-five unrelated sentences in two counterbalanced lists in a Latin square design.

2.1.1.3. *Procedure and analysis*

In this experiment we used the speaker-listener paradigm developed in Watson & Gibson (2004a). Two participants took part in the experiment. One of the participants, the speaker, was given a list of sentences and was instructed to read each sentence to the other participant, the listener. The speaker was told to read the sentence silently to themselves before speaking until they were confident that the sentence could be produced without errors. The speaker's speech was recorded at 16 Khz.

While listening to the sentence, the listener was presented with a blank computer screen. After the sentence was produced, the listener pressed the space bar, and was visually presented with a yes/no question about the sentence that had just been heard. The speaker did not have access to the questions being asked, and neither the speaker nor the listener was given any feedback on whether they answered the question correctly.

Each production was coded by a listener who was trained in using the ToBI intonational coding system (Silverman et al. 1992) for whether an intonational boundary occurred before the main verb of the sentence. The coder listened for prosodic cues such as pauses, segmental lengthening, and pitch movement to determine whether a boundary was present.

2.1.2. *Results & discussion*

A boundary occurred before the main verb 71.0% of the time in the left branching condition and 73.4% of the time in the right branching condition. This difference was not significant, $F < 1$.

These data are most consistent with the hypothesis that intonational boundaries are a function of the size of the most recently completed constituent rather than the distance of integration of upcoming words. We acknowledge that the main finding here is a lack of a difference between the two conditions, a null result. But there was no suggestion of such an effect, despite potentially substantial integration differences between the two structures. We will therefore assume that the LRB as formulated in (1) is an initially adequate theory of the placement of intonational boundaries in production. In the next section, we discuss how this relationship between syntactic structure and intonational phrasing might be potentially useful to a listener.

3. Intonational boundaries and comprehension

If some aspects of the LRB correctly characterize the relationship between intonational phrasing and syntactic structure in production, it has some interesting implications for how listeners use prosodic information in comprehension. Part of the job of a listener is to determine the syntactic structure of an utterance, and the LRB suggests that intonational boundaries can reveal useful information about syntactic structure.

A great deal of evidence suggests that listeners use intonational boundary information to make decisions in sentence comprehension. Lieberman (1967) observed that differing syntactic surface constructions of globally ambiguous sentences can be distinguished by listeners using intonational boundaries. Consider the sentences in (5), where an intonational boundary is indicated by "//".

(5) a. I'll move on // Saturday.
    b. I'll move // on Saturday.

In (5a), *on* is interpreted as a particle of the verb "move", and the sentence means that the speaker will move to a new topic or idea on Saturday. In (5b), "on" is interpreted as a preposition, and the sentence means that the speaker will be moving his/her house on Saturday. In these two interpretations, the intonational boundary coincides with a major syntactic boundary, the boundary between the verb and the direct object. The location of this syntactic boundary differs in each sentence, and the prosodic boundary clarifies which interpretation is intended. Lieberman (1967) proposed that listeners use intonational boundaries to infer the intended meaning of the sentence.

Lieberman's proposals have been supported by research showing that listeners can use intonational boundary information in globally ambiguous sentences where the interpretations' surface structure differed (Wales & Toner 1979; Lehiste 1973; Lehiste, Olive & Streeter 1976; Cooper & Paccia-Cooper 1980; Warren 1985; Price et al. 1991; Schafer 1997; see Cutler et al. 1997 for a review). Intonational Boundaries can also be used to disambiguate sentences with temporary ambiguities, suggesting that boundary information is used during on-line processing (Marslen-Wilson et al 1992; Warren et al. 1995; Grabe, Warren & Nolan 1994; Speer, Kjelgaard & Dobroth 1996). For example, Speer, Kjelgaard & Dobroth (1996) found that listeners used boundaries to disambiguate ambiguities like (6):

(6) Whenever the guard checks the door [is / it's] locked.

There is a temporary ambiguity at the point of processing the NP "the door". This NP can be the direct object of "checks" or the subject of the main clause yet to come. In both on-line and off-line experiments, comprehension was facilitated when an intonational boundary occurred at the corresponding clause boundary for the two interpretations: a

boundary after "the door" in the direct object attachment condition and a boundary before "the door" in the main clause subject attachment condition. Understanding was facilitated compared to a baseline condition with neutral prosody. Speer et al. also found that switching the prosodic boundaries for the two interpretations interfered with comprehension relative to the baseline.

These results raise the question of how intonational phrase boundaries are used to signal syntactic structure to a listener. One possibility is raised by applying the LRB in comprehension. A natural consequence of the LRB hypothesis in production is that the presence of an intonational boundary serves as a strong cue not to attach the upcoming word to the last potential attachment site before the boundary. This consequence follows for two reasons. First, according to the first component of the LRB, speakers tend to place an intonational boundary after a completed constituent, especially if the constituent is long. A boundary therefore often indicates that a syntactic constituent is complete and no further attachments should be made to it, resulting in a tendency to not make attachments to a head followed by a boundary. Second, according to the second component of the LRB, speakers tend to place an intonational boundary before a large upcoming constituent, if it is not an argument of the current head. A boundary would therefore suggest that an upcoming constituent is not an argument of the current head. It might still be a modifier, but it is likely not to be an argument, and there is therefore an increased likelihood not to attach this word to the current head. These two consequences of the LRB give rise to the anti-attachment hypothesis in (7):

(7) Anti-Attachment Hypothesis (AAH):
     Listeners prefer not to attach an incoming word to a lexical head that is immediately followed by an intonational boundary.

The AAH has the following implications: 1) the presence of an intonational boundary after a lexical head that is the site of subsequent attachment increases processing difficulty and 2) the presence of an intonational boundary after a lexical head that is not the site of subsequent attachment decreases processing difficulty.

To see how the AAH applies, consider a structure such as (8) where an intonational boundary occurs before the matrix verb "fired".

(8) The judge who the reporter ignored // fired the secretary.

A listener that is aware of the relationship between intonational boundaries and constituency in production could deduce that the boundary serves as a potential cue to being the end of the subject NP and that the upcoming word does not attach to the pre-boundary word "ignored". By using the AAH heuristic, the listener could therefore predict that the upcoming word is the matrix verb and that it attaches to the head subject NP "judge".

The AAH also predicts that intonational phrase boundaries will be useful in syntactic ambiguity resolution as in the globally ambiguous structure in (9).

(9) a. The cop saw the spy // with the telescope.
    b. The cop saw // the spy with the telescope.

In (9a), the AAH predicts that the boundary after "spy" will cue the listener not to attach the ambiguous PP "with the telescope" to the pre-boundary word "spy", so listeners should be biased toward attaching the PP to the verb. In (9b), the boundary after the verb "saw" should bias listeners towards attaching the PP "with the telescope" low to "the spy", because the boundary is a signal not to attach to the verb. Note that, in spite of the boundary following "saw", there is no alternative but to attach the following NP "the spy" to the verb "saw". This illustrates that the presence of a boundary is just one of multiple constraints that determines the most preferred interpretation of an utterance, in a multiple constraint system such as that proposed by MacDonald, Pearlmutter & Seidenberg (1994) and Tanenhaus et al. (1995).

Central to the AAH is the idea that intonational boundaries cause a break between constituents. That is, boundaries create biases that force constituents apart during the parsing process. This contrasts with many of the theories in the literature that claim that intonational phrasing's primary role in comprehension is to group relevant constituents together (Pynte & Prieur 1996; Schafer 1997; Speer, Kjelgaard & Dobroth 1996; Kjelgaard & Speer 1999; Frazier & Clifton 1998). One common claim of many of these researchers is that words occurring within the same intonational phrase undergo syntactic analysis at roughly the same stage of processing, whereas words occurring in different intonational phrases are analyzed at different processing stages. Constructing a syntactic dependency within an intonational phrase is therefore less difficult than constructing a dependency between intonational phrases because the syntactic analysis in the latter case involves words that are processed at different stages. We refer to this claim as the domain hypothesis, in (10):

(10) Domain Hypothesis: The presence of an intervening intonational phrase boundary between two dependent heads results in an increase in processing difficulty.

Evidence for the domain hypothesis comes from work on the influence of intonational phrases on ambiguity resolution. For example, the domain hypothesis accounts for biases towards local attachment in globally ambiguous structures such as (9b) repeated below as (11):

(11) The cop saw // the spy with the telescope.

Of the two potential attachment sites for "with the telescope", only "the spy" is in the same prosodic domain as the PP, thus listeners are predicted

to attach to it. Although a domain theory accounts for this preference, it is not clear how such a theory can account for the bias to attach to the verb when a boundary immediately precedes the PP as in (9a).

There are at least two ways in which the parser's architecture might account for the domain hypothesis. One possibility is that memory or attentional limitations of the sentence processor give rise to the domain hypothesis, such that connecting elements in different intonational phrases requires more memory or attentional resources than connecting elements in the same intonational phrase. This type of theory predicts that the effects of intonational phrasing should affect parsing in both ambiguous and unambiguous sentences. We will refer to this hypothesis as the unambiguous domain hypothesis. A second possibility is that the domain hypothesis is related to a heuristic used by the parser in ambiguous contexts only. This hypothesis will be referred to as the ambiguity-only hypothesis.

Whereas the ambiguity-only hypothesis makes no predictions for the effects of intonational phrasing in unambiguous structures, the unambiguous domain hypothesis predicts that the addition of intonational phrase boundaries to an unambiguous sentence will make the sentence more difficult to understand, because establishing dependency relationships between lexical items across intonational phrases is more difficult than establishing dependency relationships within intonational phrases. Thus the unambiguous domain hypothesis predicts that (12b), which includes one more boundary than (12a), should be more difficult to process than (12a):

(12) a. The detective showed the blurry picture of the diamond to the client.
     b. The detective showed the blurry picture of the diamond // to the client.

Note that in (12), the AAH makes the opposite prediction. The boundary after "diamond" should correctly cue non-local attachment of the PP "to the client" to a word further back in the input stream: the verb "showed". Therefore, processing should be facilitated because of the usefulness of the cue that the boundary provides. We test these hypotheses in Experiment 2.

3.1. *Experiment 2*

The goal of Experiment 2 was to test whether the presence of an intonational boundary in certain unambiguous sentence structures would increase or decrease the difficulty of processing the sentence. In this experiment, listeners rated the difficulty of sentences that varied in their intonational phrasing. The materials in Experiment 2 had the form of (13) in a 2 × 2 design, with intonational boundaries at (1) and/or (2):

(13) The detective showed the blurry picture of the diamond (1) to the client (2) who was in an office that was on the fourteenth floor.

The boundary at position (1) precedes the PP "to the client", the second argument of the verb "showed". The boundary at position (2) precedes a relative clause, disambiguating it towards a non-restrictive interpretation.

The crucial manipulation was the presence of an intonational boundary at position (1). The predictions of the theories are as follows. First, the ambiguity-only domain theory makes no predictions for these structures because the test sentences are not ambiguous. Next, the unambiguous domain theory predicts that the intonational phrase boundary at position (1) will increase the overall difficulty of processing (13) because the boundary separates heads that have dependency relationships in different intonational phrases. In particular, the boundary at position (1) separates the verb "showed" from its PP argument "to the client", increasing the difficulty of integration. Finally, the AAH predicts that the intonational phrase boundary at position (1) will lower the complexity of the sentence. A boundary after "diamond" facilitates the non-local attachment of the PP "to the client" to the verb "showed".

The intonational boundary at position (2) served as a control. This intonational boundary was expected to facilitate processing because it disambiguated the relative clause towards a non-restrictive interpretation. Previous results from reading suggest that non-restrictive relatives are easier to interpret in a null context than restrictive relatives (Grodner, Gibson & Watson 2005), possibly because they presuppose a simpler discourse representation (e.g., Crain & Steedman 1985). The items in this experiment were presented without a context, so the non-restrictive interpretation was likely to be the preferred one. Thus, the intonational phrase boundary at position (2) was expected to facilitate processing for discourse reasons, and provided a point of comparison for the effects of an intonational boundary at position (1).

3.1.1. *Method*

3.1.1.1. *Participants*
Forty native English speakers from the MIT community who did not participate in Experiment 1 participated in the study for $5.00 each.

3.1.1.2. *Materials*
The target sentences had the form of (13), in which the verb of the sentence takes two arguments: an NP object, and a PP initiated by the dative preposition "to". The NP object always had an argument PP

initiated by the preposition "of" (e.g., "of the diamond" in (13)). The object of the PP was modified by a relative clause.

The experiment was a 2×2 design, varying the presence of an intonational boundary at the two locations indicated in (13). This resulted in four conditions: two conditions with a boundary at one of the locations indicated in (13), a condition with no boundaries, and a condition with boundaries at both locations. A total of 16 items were used in this experiment along with 30 fillers. The stimuli were presented in four counterbalanced lists in a Latin Square design.

The stimuli were created through digital editing. Each condition was produced and recorded independently. For each item, a control sentence was produced that contained no intonational boundaries. In order to control the prosody among the sentences, the relevant sections of each condition were spliced into the control condition. In particular, the auditory string "of the diamond to the client who was in an office" was extracted from each of the independent productions and spliced into the control. This was done in every condition, including the condition with no prosodic boundaries, to ensure that any differences in difficulty would not be attributed to irrelevant differences in prosody between the conditions or in the splicing itself.

The conditions with intonational boundaries were produced such that the final segment of the intonational phrase was lengthened and was followed by a perceptually salient boundary. The pause between intonational phrases was approximately 200 ms.

3.1.1.3. *Procedure*
The experiment was presented as a questionnaire on a web page in which participants were asked to rate sentences for comprehensibility on a 7-point scale where 1 was easy to understand and 7 was hard to understand. Participants clicked on a link for each item and then listened to the item on headphones. Participants were instructed to listen to each sentence only once. Then they clicked a button below the link to view a question about the sentence to be answered with either a *yes* or *no* on the web page. Four example sentences were presented to the subject. Two of these sentences were relatively comprehensible with normal prosody, and it was suggested that they be rated with a 1 or a 2. One of the sentences was quite difficult and contained two embedded relative clauses: "The dog who the cat who the mouse bit scratched ran away." It was suggested that this sentence be given a rating of 6 or 7. The final example sentence was difficult because it contained an intonational boundary in an unnatural location: "The judge trusted the witness // wouldn't run away." It was suggested that this sentence be given a 6 or 7. No specific reasons were given for why the sentences deserved the suggested ratings, and the syntactic forms of the examples differed from those of the test items. Participants completed the task in 20–25 minutes.

3.1.2. *Results*

The mean difficulty ratings and response accuracies to comprehension questions for the four conditions are presented in Table 1. Sentences with an intonational boundary at position (1) were rated as being significantly easier to understand than sentences without an intonational boundary at that location ($F_1(1,39) = 9.40$, p < .005; $F_2(1,15) = 14.30$, p < .005). Participants also rated sentences with an intonational boundary before the relative clause as being significantly easier than sentences without an intonational boundary at this location ($F_1(1,39) = 11.86$, p < .005; $F_2(1,15) = 11.87$, p < .005). There were no interactions (Fs < 1).

Comparisons between individual conditions also revealed significant differences. Sentences containing intonational boundaries at both locations were rated as easier than sentences containing boundaries at just the clause boundary ($F_1(1,39) = 6.28$, p < .05; $F_2(1,15) = 10.25$, p < .01) or at just the PP position ($F_1(1,39) = 6.29$, p < .05; $F_2(1,15) = 5.19$, p < .05). Sentences containing intonational boundaries at either the PP position or the clause boundary were significantly easier than sentences with no boundary at all ($F_1(1,39) = 4.83$, p < .05; $F_2(1,15) = 4.98$, p < .05) and ($F_1(1,39) = 7.08$, p < .05 ; $F_2(1,15) = 10.46$, p < .01, respectively).

Table 1. The difficulty ratings and response accuracies to comprehension questions for each condition in Experiment 1. Standard errors are provided in parentheses

Experiment 1 Difficulty Ratings		
	Boundary at (2)	No Boundary at (2)
Boundary at (1)	3.85 (.18)	4.29 (.17)
No Boundary at (1)	4.28 (.17)	4.61 (.18)
Experiment 1 Accuracy Ratings		
	Boundary at (2)	No Boundary at (2)
Boundary at (1)	84% (2.7)	80% (2.5)
No Boundary at (1)	80% (2.9)	76% (2.0)

The trends for response accuracy comparisons among the four conditions were all the same as for the difficulty ratings, although none of the response accuracy comparisons quite reached significance. In particular, participants tended to answer more questions correctly when there was an intonational boundary at position (1) ($F_1(1,39) = 3.26$, p = .08; $F_2(1,15) = 1.87$, p = .19), and when there was an intonational boundary before the relative clause ($F_1(1,39) = 3.48$, p = .07; $F_2(1,15) = 2.91$, p = .11).

3.1.3. *Discussion*

The results of Experiment 2 support the anti-attachment hypothesis over the unambiguous domain hypothesis. Inserting an intonational phrase boundary between a head and its non-local argument did not increase the difficulty of processing the sentence, as predicted by the unambiguous domain hypothesis. Instead, the presence of the intonational boundary at position (1) made the sentence easier to process, as predicted by the AAH.

Although the results from this experiment are suggestive, they do not address the question of how intonational phrases are used in on-line sentence comprehension, because the task that was used here was an off-line questionnaire. Experiment 3 was designed to evaluate the AAH and the domain hypothesis using an on-line task.

3.2. *Experiment 3*

This experiment had two goals. The first was to explore whether intonational phrase information is used immediately in the on-line processing of a sentence. The second goal was to see whether the presence of a boundary after a head that receives no future attachments facilitates processing or whether simply the presence of a boundary anywhere in a sentence facilitates processing.

These hypotheses were investigated using a cross-modal lexical decision task (Marslen-Wilson et al. 1992; Kjelgaard & Speer 1999). In this paradigm, the participant is auditorily presented with an initial segment of a sentence, and then must perform a lexical decision on a letter string that could serve as a possible continuation of the segment. The assumption behind this paradigm is that the speed of lexical decision reveals the degree to which the listener expects the continuation, and this expectation is a function of the material the participant hears in the auditory segment of the task. This paradigm allows us to directly explore effects of intonational phrasing by manipulating the presence of intonational phrase boundaries in the initial auditory segment.

The experiment was a 2×2 design crossing the presence of an intonational boundary (present, absent) with the attachment site of a preposition (VP, NP). Materials such as the one in (14) were used in this study:

(14) The museum lent the sculpture… of (NP attachment)/
                                 to (VP attachment)
                     H*
                     H* L%

"H*" and "L%" are symbols from the ToBI coding system that represent high and low F_0 targets (Silverman et al. 1992). H* denotes a pitch accent conveyed by a peak in F_0 and L% denotes a boundary tone, a valley in F_0, which coincides with the end of an intonational phrase. Participants

listened to sentences such as (14), up to the word "sculpture". We manipulated the presence of a boundary (L%) on the word "sculpture". The boundary tone in this instance consisted of lengthening of the word and a drop in pitch. In the no-boundary-tone condition, the word "sculpture" was not lengthened and there was no pitch movement. The participants then had to make a lexical decision on the visually presented words "of" or "to", both of which are possible grammatical continuations of the initial sentence segment. The preposition "of" unambiguously attaches as a PP argument of the noun "sculpture", as in "the sculpture of the deity". The preposition "to" unambiguously attaches as a PP argument of the verb "lent", as in "The museum lent the sculpture to the archaeologist".

The unambiguous domain hypothesis predicts higher reaction times when there is a boundary present in both attachment conditions. In both cases, the prepositions "of" and "to" are not in the same prosodic phrase as their dependent lexical head ("sculpture" and "lent" respectively). Because these heads are less accessible by virtue of being in different intonational phrases, listeners should encounter more difficulty in attaching them than when there is no intonational boundary present.

In contrast to the domain hypothesis, it is possible that intonational boundaries will always facilitate processing in long sentences. This hypothesis is consistent with the results of Experiment 2, where boundaries at two positions resulted in improved acceptability. Such a hypothesis predicts that lexical decision times will be faster in both the NP and VP conditions when an intonational phrase boundary is present than when an intonational boundary is not present.

The AAH makes a different set of predictions from the preceding two theories. The AAH predicts that an intonational boundary after the word "sculpture" should signal non-attachment to this lexical head. Thus, when there is a boundary present, the listener should be faster at performing a lexical decision on the word that does not attach to "sculpture" (the preposition "to") than a word that does (the preposition "of"). When there is no boundary present, the listener should be faster at performing a lexical decision to "of" than to "to". The AAH therefore predicts an interaction between the presence/absence of a boundary and the attaching word. The AAH also predicts effects within each attachment condition. Lexical decision times for "to" should be faster when a boundary is present than when it is not present, because the intonational boundary should facilitate the non-local attachment. In addition, lexical decision times for "of" should be slower when an intonational boundary is present than when it is not present because the intonational boundary should interfere with attaching to the head that precedes the boundary.

Because of the nature of the task, the AAH is compatible with other related patterns of data. In particular, the no-boundary condition actually *does* contain a cue indicating a prosodic boundary: the break

in the auditory presentation. Thus the no-boundary condition contains one cue that indicates a boundary – the break in presentation – and other cues that indicate no boundary – the lack of segmental lengthening and pitch movement on the final word before the break in presentation. There are therefore conflicts between the prosodic cues and the attachment site in the no-boundary conditions. In contrast, there are no conflicting prosodic cues in the boundary conditions. In the boundary conditions, the prosodic cues unambiguously signal non-local attachment, according to the AAH. Thus the predictions of the AAH are clear in the two boundary conditions: The VP attachment boundary condition should have faster RTs than the NP attachment boundary condition. Because of the conflicting prosodic cues in the two no-boundary conditions, the RTs for these conditions may fall between the RTs for the boundary conditions, or they may pattern with one of them. In particular, it is possible that people may have difficulty with the lexical decision task if *any* prosodic cues conflict with the attachment site. If this is the case, then the two no-boundary conditions may pattern with the boundary/local attachment condition, because in all three cases, the intonational cues do not unambiguously support the target attachment site. It is only in the boundary/non-local attachment condition that the intonational cues are fully consistent with the attachment site, so this may be the only case in which RTs decrease. Crucially, the AAH predicts an interaction between the boundary and attachment conditions in all of these possibilities.

3.2.1. *Method*

3.2.1.1. *Materials*
Twenty items with the syntactic structure in (14) were constructed for this experiment. In all of the items, the target word for lexical decision was "of" for the local NP attachment condition and "to" for the non-local VP attachment condition. All of the NP direct objects in the items could potentially take a PP argument headed by the preposition "of", and none allowed a PP headed by "to". All of the verbs allowed for PP argument headed by "to" but none took a PP headed by "of" as an argument. Thus, the attachment site for the target preposition was unambiguous.

   A speaker who was trained in the ToBI coding system (Silverman et al. 1992) produced the boundary and the no-boundary conditions individually along with a baseline condition with no intonational boundaries. The boundary condition was produced with an H* pitch accent on the final word in the speech segment and ended in a L% boundary tone. The boundary tone on the final word was signaled through lengthening. No pause was included at the end of the utterance. The no-boundary condition was produced with only an H* on the final word. Both conditions were spliced at the end of the word "sculpture" in (14).

In order to control for irrelevant differences in prosody, the final word in the speech segment (in this case "sculpture") was extracted from each of the relevant conditions and spliced into the corresponding position in the baseline utterance. In order to reduce clicks, the splicing occurred at the zero crossing in the waveform closest to the final word so that in some conditions, additional words were spliced into the baseline. However, the spliced words were the same in both conditions for a given item.

Sixty fillers were included with the twenty experimental items. The sixty fillers had the same syntactic structure as the test items. Every sentence contained an unmodified definite subject followed by a verb that was followed by a direct object and a prepositional phrase. In half of the fillers, the PP modified the direct object, and in the other half, it modified the verb. The prepositions "to" and "of" were used as lexical decision targets only in the experimental items. The non-word lexical decision targets in the fillers were English words that had one letter changed so that the word did not constitute a real word in English (e.g. FER instead of FOR). In addition, the point in the sentence at which the lexical decision task occurred varied across fillers, so that subjects could not predict the lexical decision point.

3.2.1.2. Procedure
Participants performed the experiment at a PC. They were given headphones and were told that they would be listening to utterances that would end at a random point mid-sentence. Immediately following the speech segment, a word would appear on the screen, and they were to indicate whether the word was a real word in English by pressing a computer key for "yes" and another key for "no". Participants were told to answer as quickly and as accurately as possible. They were also told to listen to the sentence, as it would help them in their task. Participants were also given five practice items that were similar in structure to the fillers, so that they would grow accustomed to the task.

3.2.1.3. Participants
Forty participants from the MIT community participated in the experiment. They were given $5 for their participation.

3.2.2. Results
Incorrect responses constituted 3.1% of the data set and were not included in the analyses. The remaining data were trimmed at 1000 ms to reduce the effects of spurious outliers, eliminating 4.7% of the reaction times. The mean response times for the four conditions are presented in Table 2.

Table 2. Lexical decision times in msec for the conditions in Experiment 2. Standard errors are presented in parentheses

	Local Attachment (*of*)	*Non-Local Attachment* (*to*)
Intonational Boundary	583 (17)	555 (16)
No Intonational Boundary	587 (16)	598 (18)

An ANOVA revealed a main effect of boundary ($F_1(1,39) = 6.52$, $p < .05$; $F_2(1,19) = 3.547$, $p = .07$), such that RTs were faster when there was a boundary present than when there was no boundary present. RTs did not differ according to attachment site ($F1 = 1.43$, $F_2 = 0.19$). Finally, there was a significant interaction between the boundary and attachment conditions ($F_1(1,39) = 7.78$, $p < .01$; $F_2(1,19) = 5.18$, $p < .05$), as predicted by the AAH.

Individual condition comparisons were conducted to further test the predictions of the AAH. In the boundary conditions, RTs to non-locally attached elements were faster than locally attached elements ($F_1(1,39) = 9.05$, $p < .01$) although this difference was only marginal in the items analysis ($F_2(1,19) = 4.26$, $p = .05$). There was no significant difference between the attachment conditions in the no-boundary condition (Fs $<$ 1). Within the non-local VP attachment conditions, reaction times were significantly faster when a boundary was present than when a boundary was not present $F_1(1,39) = 16.98$, $p < .001$; $F_2(1,19) = 7.23$, $p < .05$. Within the local NP attachment conditions, the presence of a boundary had no significant effect on RTs (Fs $<$ 1).

3.2.3. *Discussion*

The results are not compatible with the predictions of the domain hypothesis. The domain hypothesis predicted that a boundary would hinder performance for both conditions, but contrary to this prediction, a boundary aided performance overall, especially in the VP conditions. The results are also not predicted by the boundary facilitation hypothesis. According to this hypothesis, the presence of a boundary should have facilitated the lexical decision task for both the NP and VP attachment conditions. Although the boundary did facilitate the VP attachment condition, it did not facilitate the NP attachment condition.

The results are roughly as predicted by the AAH. First, the interaction between the boundary and attachment conditions was as predicted by the AAH: people were fastest at lexical decision when there was a boundary on the non-local VP attachment. The pattern of data best fits the AAH under the additional assumption that any conflict between prosodic cues and attachment site leads to a ceiling effect in RTs. The only condition in which the prosodic cues are consistent with the attachment site is the

boundary/non-local VP attachment condition, so RTs are fastest here. In the other three conditions, some prosodic cues conflict with the attachment site, and so RTs are slower for all three. In the boundary/local NP attachment condition, all the prosodic cues suggest a non-local attachment, but the attachment is local. In the no-boundary/non-local VP attachment condition, the lack of segmental lengthening and pitch movement on the final word in the auditory presentation suggests local attachment, but the attachment is non-local. In the no-boundary/local NP attachment condition, the cue consisting of the break in presentation suggests non-local attachment, but the attachment is local.

4. General discussion

In this paper, we have argued that intonational phrase boundaries tend to occur before and after long syntactic constituents and are partly the result of resource processing demands on language production. A consequence of this relationship is that intonational phrase boundaries can provide useful parsing information for a listener. We hypothesized that listeners interpret intonational phrases boundaries as marking the end of a constituent, and prefer not to attach incoming heads to words that are followed by an intonational phrase boundary.

We have argued that these preferences in comprehension and production result from preferences in performance rather than speakers' linguistic knowledge. To compare the predictions of the AAH against at least one competence theory of intonational phrasing, Watson & Gibson (2004b) conducted an experiment testing the predictions of the AAH and Selkirk's (1984) Sense Unit Condition (SUC).

Selkirk argues that intonational phrases correspond to semantic units called Sense Units. Sense Units are defined as groups of words that engage in a semantic relationship with each other. If the constituents within an intonational phrase do not form a sense unit, the sentence is ungrammatical. Formally, Selkirk defines the Sense Unit Condition as in (15).

(15) The Sense Unit Condition of 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 either (a) or (b) is true of the semantic interpretation of the sentence:
   a. C_i modifies j (a head)
   b. C_i is an argument of C_j (a head)

For sentences such as the ones in (16), the SUC makes clear predictions.

(16) a. Baseline
      The artist gave a portrait of the president to the manager on Wednesday.

b. Boundary after attachment site
The artist gave a portrait (1) of the president to the manager on Wednesday.
c. Boundary after non-attachment site
The artist gave a portrait of the president (2) to the manager on Wednesday.

The SUC predicts that the boundary at (1) should make the sentence more difficult because it creates an intonational phrase consisting of "of the president", "to the manager", and "on Wednesday". Because these prepositional phrases do not participate in a semantic relationship with each other, they violate the sense unit condition. The SUC also predicts that the boundary at (2) in (16c) should increase difficulty because the intonational phrase "to the manager on Wednesday" also does not form a sense unit.

The AAH makes different predictions. A boundary at (1) is predicted to increase difficulty because it incorrectly signals that "portrait" receives no upcoming attachment. The predictions of the SUC and the AAH differ in their predictions for the boundary at (2). The AAH predicts that this boundary should facilitate processing because it correctly signals that "president" receives no upcoming attachments.

In an auditory survey, Watson & Gibson (2004a) found that listeners' ratings were consistent with the AAH and not the SUC. Sentence (16c) was rated as easier than (16a) and (16b). Sentence (16b) was rated as more difficult than (16a) numerically, but this difference was not significant. The results from this experiment are problematic for the Sense Unit Condition, and lend support to a performance-based theory such as the AAH.

The data presented here suggest that a performance-based theory of prosodic phrasing is capable of accounting for a large amount of English data. The AAH gives a better account of listener preferences than domain processing theories and the Sense Unit Condition, both of which argue that the grouping of words into intonational phrases influence sentence acceptability. In contrast, the AAH is based on the notion that intonational boundaries are useful to listeners in so far as they signal a break between syntactic units.

In general, the work here suggests that different syntactic factors play a role in intonational phrasing in production and comprehension and that these are best accounted for by the LRB and the AAH. The fact that there appears to be a need for two different processing theories of intonational phrasing in these two language domains is not surprising given that the processing goals in production and comprehension are very different. This work suggests that future studies investigating the relationship between syntax and prosodic phrasing must look at preferences in both production and comprehension to give a full account of the relationship between syntactic and prosodic structure.

References

CARLSON, K., CLIFTON, C. & FRAZIER, L. 2001. Prosodic boundaries in adjunct attachment. *Journal of Memory and Language* 45, 58–81.

CHOMSKY, N. & HALLE, M. 1968. *The sound pattern of English*. New York: Harper and Row. Reprinted in 1991 by MIT Press.

COOPER, W. E. & PACCIA-COOPER, J. 1980. *Syntax and speech*. Cambridge, MA: Harvard University Press.

CRAIN, S. & STEEDMAN, M. 1985. On not being led up the garden path; the use of context by the psychological parser. *Natural language parsing: Psychological, computational, and theoretical perspectives*, eds. D. R. Dowty, L. Karttunen & A. Zwicky, 320–358. Cambridge University Press.

CUTLER, A., DAHAN, D. & VAN DONSELAAR, W. 1997. Prosody in the comprehension of spoken language: A literature review. *Language and Speech* 40, 141–201.

FERREIRA, F. 1988. *Planning and timing in sentence production: The syntax-to-phonology conversion*. University of Massachusetts, Amherst, MA. Unpub. diss.

FERREIRA, F. 1991. Effects of length and syntactic complexity on initiation times for prepared utterances. *Journal of Memory and Language* 30, 210–233.

FRAZIER, L. & CLIFTON C., JR. 1998. Sentence reanalysis and visibility. *Reanalysis in sentence processing*, eds. J. D. Fodor & F. Ferreira, 143–176. Dordrecht: Kluwer.

GEE, J. P. & GROSJEAN, F. 1983. Performance structures. A psycholinguistic and linguistic appraisal. *Cognitive Psychology* 15, 411–458.

GIBSON E. 1998. Linguistic complexity: Locality of syntactic dependencies. *Cognition* 68, 1–76.

GRABE, E., WARREN, P. & NOLAN, F. 1994. Prosodic disambiguation of coordination structures. Paper presented at the Eighth Annual CUNY Conference on Human Sentence Processing. Tucson, Arizona, March 16–18.

GRODNER, D., GIBSON, E. & WATSON D. 2005. The influence of contextual contrast on syntactic processing evidence for strong interaction in sentence comprehensis. *Cognition* 95, 275–296.

GRODNER, D. & GIBSON, E. in press. Consequences of the serial nature of linguistic input. *Cognitive Science* 29, 261–291.

HIRST, D. 1993. Detaching intonational phrases from syntactic structure. *Linguistic Inquiry* 24, 781-788.

KJELGAARD, M. M. & SPEER, S. R. 1999. Prosodic facilitation and interference in the resolution of temporary syntactic closure ambiguity. *Journal of Memory and Language* 40, 153–194.

KRALJIC, T. & BRENNAN, S. E. 2005. *Prosodic disambiguation of syntactic structure: For the speaker or for the addressee*? *Cognitive Psychology* 50, 194–231.

LESHISTE, I. 1973. Phonetic disambiguation of syntactic ambiguity. *Glossa* 7, 107–122.

LEHISTE, I., OLIVE, J. & STREETER, L. 1976. Role of duration in disambiguating syntactically ambiguous sentences. *Journal of the Acoustical Society of America* 60, 1199–1202.

LIEBERMAN, P. 1967. *Intonation, perception and language*. Cambridge: MIT Press.

MACDONALD, M. C., PEARLMUTTER, N. J. & SEIDENBERG, M. S. 1994. The lexical nature of syntactic ambiguity resolution. *Psychology Review* 101, 676–703.

MARSLEN-WILSON, W. S., TYLER, L. K., WARREN, P., GRENIER, P. & LEE, C. S. 1992. Prosodic effects in minimal attachment. *The Quarterly Journal of Experimental Psychology*, 45A, 73–87.

NESPOR, M. & VOGEL, I. 1987. *Prosodic phonology*. Dordrecht: Foris.

PRICE, P. J., OSTENDORF, M., SHATTUCK-HUFNAGEL, S. & FONG, C. 1991. The use of prosody in syntactic disambiguation. *JASA* 90, 2956–2970.

PYNTE, J. & PRIEUR, B. 1996. Prosodic breaks and attachment decisions in sentence parsing. *Language and Cognitive Processes* 11, 165–191.

SCHAFER, A. J. 1997. *Prosodic parsing: The role of prosody in sentence comprehension*. University of Massachusetts Doctoral Dissertation. Amherst: GLSA.

SCHAFER, A. J., SPEER, S. R., WARREN, P. & WHITE, S. D. 2001. *Prosodic influences on the production and comprehension of syntactic ambiguity in a game-based conversation task*. Fourteenth Annual CUNY Conference on Human Sentence Processing, Philadelphia, PA.

SELKIRK, E. 1984. *The relation between sound and structure*. Cambridge: MIT Press.

SELKIRK, E. 1986. On derived domains in sentence phonology. *Phonology Yearbook* 3, 371–405.

SELKIRK, E. 2000. The interactions of constraints on prosodic phrasing. In *Prosody: theory and experiment*, ed. M. Horne, 231–261. Dordrecht: Kluwer.

SHATTUCK-HUFNAGEL, S. & TURK, A. 1996. A prosody tutorial for investigators of auditory sentence processing. *Journal of Psycholinguistic Research* 25, 193–247.

SILVERMAN, K., BECKMAN, M., PIERREHUMBERT, J., OSTENDORF, M., WIGHTMAN, C., PRICE, P. & HIRSCHBERG, J. 1992. ToBI: A standard scheme for labeling prosody. In: *Proc. Second International Conference on Spoken Language Processing (ICSLP)*, Banff, 867–879.

SNEDEKER, J. & TRUESWELL, J. 2003. Using prosody to avoid ambiguity: Effects of speaker awareness and referential context. *Journal of Memory and Language* 48, 103–130.

SPEER, S. R., KJELGAARD, M. M. & DOBROTH, K. M. 1996. The influence of prosodic structure on the resolution of temporary syntactic closure ambiguities. *Journal of Psycholinguistic Research* 25, 247–268.

TANENHAUS, M. K., SPIVEY-KNOWLTON, M. J., EBERHARD, K. M. & SEDIVY, J. E. 1995. Integration of visual and linguistic information in spoken language comprehension. *Science* 268, 1632–1634.

TRUCKENBRODT, H. 1999. On the relation between syntactic phrases and phonological phrases. *Linguistic Inquiry* 30, 219–255.

WALES, R. & TONER, H. 1979. Intonation and ambiguity. *Sentence Processing: Psycholinguistic studies presented to Merrill Garrett*. eds. W. E. Cooper & E. C. T. Walker, Hillsdale, NJ: Lawrence Earlbaum Associates, 135–158.

WARREN, P. 1985. *The temporal organization of speech*. Cambridge University. PhD diss.

WARREN, P., GRABE, E. & NOLAN, F. 1995. Prosody, phonology and parsing in closure ambiguities. *Language and Cognitive Processes* 10, 457–486.

WARREN, T. & GIBSON, E. 2002. The influence of referential processing on sentence complexity. *Cognition* 85, 79–112.

WATSON, D. & GIBSON, E. 2001. *Linguistic structure and intonational phrasing*. Paper presented at the 14th Annual CUNY Conference on human sentence processing in Philadelphia, PA.

WATSON, D. & GIBSON, E. 2004a. The relationship between intonational phrasing and syntactic structure in language production. *Language and Cognitive Processes* 19, 713–755.

WATSON, D. & GIBSON, E. 2004b. Making sense of the sense unit condition. *Linguistic Inquiry* 35, 508–517.

Duane Watson
Meliora 494
Brain & Cognitive Sciences
University of Rochester
Rochester, NY 14627-0268
U.S.A.
dwatson@bcs.rochester.edu

Edward Gibson
Department of Brain and Cognitive Sciences
MIT
Cambridge, MA 02139
U.S.A.
egibson@mit.edu