

Prosodic Cues “and” Syntactic Disambiguation*

Allison R. Blodgett
blodgett@ling.ohio-state.edu

In the literature regarding the relationship between prosody and parsing, three questions remain unanswered or only tangentially addressed. What is the nature of the prosodic information that might differentiate a temporary syntactic ambiguity? How consistently does this information occur in production? Are listeners sensitive to that information? I present the results of a set of production and perception experiments designed to investigate the nature of the prosodic cues that might disambiguate the temporary syntactic ambiguity inherent in coordinate structures. The results suggest that although most speakers fail to provide disambiguating prosody, when they do, they manipulate phrasing around, and accentuation of, the conjunction word *and*. The results further suggest that of the two cues, only the manipulation of phrasing is interpreted as a general cue across speakers. Specifically, the absence of a phrase break is interpreted as a prosodic cue to conjoined NPs.

Prosody is an essential part of any spoken utterance. As the examples in (1) show, prosody helps to define the context or set of contexts within which an utterance may be felicitously produced and understood. In each representation, the black line indicates the relative rise and fall in pitch, while capitalization indicates the “sentence stress,” or the location of the most prominent pitch accent in the utterance.



(1a) The WINDOW'S open. (1b) The window's OPEN. (1c) The window's OPEN?

Even though these three utterances contain the same lexical information, they are not interchangeable. While (1a) can function as the response to the question, *What's*

* This work was supported in part by a Summer Graduate Research Fellowship in Cognitive Science provided by the Center for Cognitive Science at The Ohio State University. I would very much like to thank Mary Beckman and Rick Lewis for their help.

open? (1b) cannot. Conversely, (1b) can function as the response to the question, *What about the window?* while (1a) cannot. Furthermore, the question intonation of (1c) contrasts with the statement intonation of (1a) and (1b). Although it is clear from even this simple example that a relationship exists between the prosody of an utterance and the contexts within which it can be felicitously produced, other relationships involving prosody also exist, including a relationship between prosody and human sentence processing.

The idea that prosody and syntactic parsing are somehow related is not new. For example, work by Lehiste (1973) provided empirical evidence supporting a role for prosody in parsing. After hearing a structurally and semantically ambiguous utterance such as *The old men and women stayed at home*, participants were able to use differences in prosody to indicate reliably whether the speaker had intended the subject of the sentence to be interpreted as "old men and old women" or as "men who are old, and women who are possibly young."

More recently much of the research on prosody and parsing has been focused on determining the point at which prosody influences a syntactic parse. Establishing this point is important because the outcome will provide empirical support for one parsing model over another. If prosody is found to influence initial parsing decisions, then the class of Constraint-based Models (Boland, 1997; MacDonald et al., 1994; McRae et al., 1998) can be argued to provide a better account of human sentence processing behavior. This will be the case regardless of whether prosody is found to influence the generation of initial syntactic structure, or structures, or the selection of one structure from among multiple automatically generated structures. In either case, prosody will simply be adopted into the list of constraints that influence parsing decisions, a list that includes such factors as frequency, thematic information, verb subcategorization, and plausibility. On the other hand, if the influence of prosody is found to be limited to syntactic revisions, then the Garden Path Model (Frazier, 1978; Frazier & Clifton, 1996), which in all its forms excludes a role for prosody in initial structure building, can be argued to provide a better account.

In the work to date, however, no consensus has emerged regarding the relationship between prosody and syntactic parsing. Whereas work on the resolution of temporary syntactic ambiguities by Beach (1991), Marslen-Wilson et al. (1992), and Nagel et al. (1994) suggests that prosody exerts an early influence on parsing, work on

what are some of the exact same syntactic ambiguities by Stirling & Wales (1996) and Watt & Murray (1996) suggests otherwise.

One factor contributing to these mixed empirical results is likely to be the fact that several important questions remain unanswered or only tangentially addressed within this body of work. These questions include the following: What is the nature of the prosodic information that might differentiate a syntactically ambiguous construction? Or asked another way, what are the relevant components of the rhythm and tune that listeners might be using to resolve an ambiguity? Are speakers producing consistent prosodic cues, and if so, what are they? When presented with these cues, are listeners sensitive to them, and if so, is there evidence that those are the cues that are used to resolve the ambiguity? To address these questions, I conducted a set of production and perception experiments using the temporary syntactic ambiguity found at the conjunction *and* in the types of sentences shown in (2). The results of these experiments will suggest that when speakers do provide disambiguating prosody, they do so by manipulating phrasing and accentuation around the conjunction. The results will further suggest that listeners interpret those cues in different ways.

(2a) Mary both sold vegetables and bought seeds at the fair.

(2b) Mary both sold vegetables and seeds at the fair.

It might seem reasonable to expect, at least on the basis of what I intuitively consider to be the prescriptive rules of Standard American English, that the attachment site of *and* in (2a) is, in fact, disambiguated by preverbal *both*, and that (2b) is simply ungrammatical. In other words, when preverbal *both* occurs between a nondual subject and the verb, a prescriptive rule stipulates that it must take scope over conjoined verbs or verb phrases (VPs), as in (2a). The existence of such a rule is supported by the results of a corpus search conducted by Blodgett and Boland (1998). Their analysis of the first 150 tokens of *both* in the Wall Street Journal database of the PennTreebank corpus failed to find any examples of preverbal *both* taking scope within the VP.

Norming data, however, conflict with this apparent prescriptive rule. The results of sentence completion norms collected at The Ohio State University as part of an earlier series of experiments demonstrated that preverbal *both* could, in fact, take scope not only

over conjoined VPs, but over conjoined phrases within the VP as well; approximately 80% of completions were of the form *Warren both read French and Spanish* (Blodgett & Boland, 1998). Thus, when a parser that is incrementally building structure reaches the conjunction in the sentences in (2), as in *Mary both sold vegetables and...*, it is not clear whether *and* must be attached low in the phrase structure tree to conjoin two noun phrases (2b) or higher in the tree to conjoin two verb phrases (2a). In other words, while the combination of a nondual subject and preverbal *both* prevents sentence-level coordination, the occurrence of preverbal *both* does nothing to further disambiguate the attachment site of the conjunction.

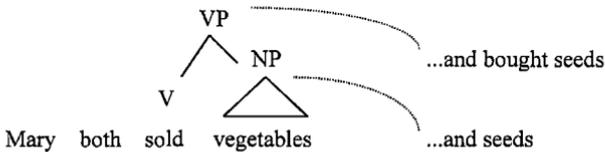


Figure 1. Example of Conjunction Attachment Ambiguity

Although the results of the sentence completions were quite robust, they represent the sole demonstration of the scope alternation of preverbal *both*. Therefore, in order to supplement these original findings and to confirm that the current experimental population also interprets preverbal *both* as taking scope over conjoined VPs and over conjoined NPs embedded within the VP, I collected a set of acceptability ratings on these two constructions. The results of these ratings not only provide additional evidence that the scope of preverbal *both* is not restricted to conjoined VPs, but they further suggest that this alternation is representative of multiple dialects of American English.

Methods

Participants. 21 students from an undergraduate linguistics course at The Ohio State University participated as an in-class experiment. An additional 21 students from an undergraduate, introductory psychology course at Rutgers participated for course credit. Responses from 2 nonnative speakers of English were subsequently excluded from the Ohio State pool and are not reported here. All of the remaining respondents were native English speakers.

Materials and procedure. After completing the perception experiment discussed later in this paper, participants turned to a page on which they were asked to rate *Mary both sold vegetables and bought seeds at the fair* and *Mary both sold vegetables and seeds at the fair*. The following three point scale was provided once for each sentence:

GOOD: I would say this and would find it ok if said by others.

ODD: I wouldn't say this but would find it ok if said by others.

BAD: I wouldn't say this and would find it strange if said by others.

Results

The proportion of GOOD, ODD, and BAD responses were tallied for each sentence by school, and the results are shown in Figures 2 and 3. As shown in Figure 2, only a very small percentage of respondents, approximately 10% at each university, rated the conjoined VP construction as BAD or completely unacceptable. This is not surprising, given that this construction is arguably the one that is consistent with the prescriptive rules of Standard American English.

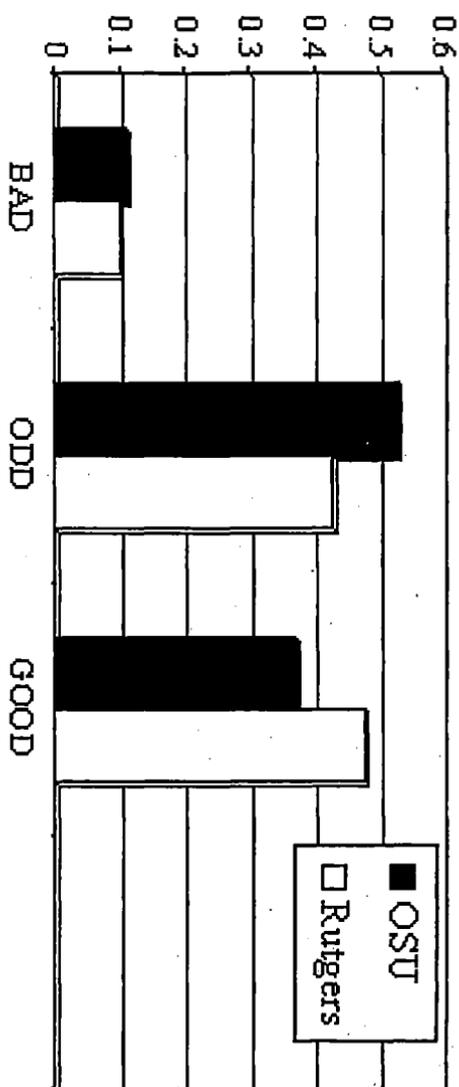


Figure 2. Acceptability Ratings for Conjoined VP

As shown in Figure 3, the conjoined NP construction also received a small percentage of BAD or completely unacceptable ratings, less than 20% at each university. Although this percentage is indeed higher than the comparable ratings for the conjoined VP construction, these responses are still in the minority. Furthermore, nearly half of the respondents at each university (52% at Rutgers and 42% at Ohio State) rated the construction as GOOD or completely acceptable.

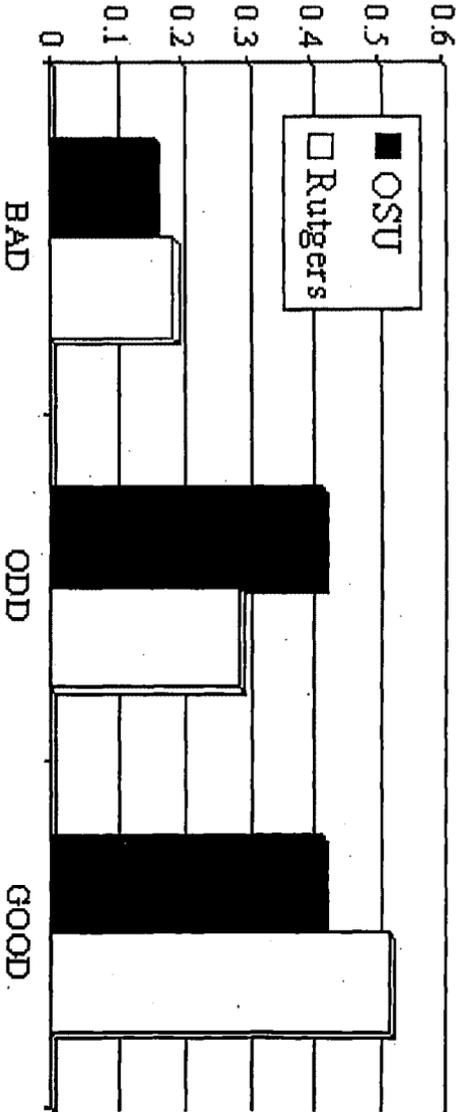


Figure 3. Acceptability Ratings for Conjoined NP

Discussion

The results of these acceptability ratings supplement the findings of the original sentence completions and provide additional evidence that contrary to the prescriptive rules of Standard American English, the scope of preverbal *both* can range over conjoined VPs or over conjoined NPs embedded within the VP.

These results further suggest that the alternation is represented in multiple geographical regions and in multiple dialects of American English. Students at two separate universities, Ohio State and Rutgers, accepted the alternations in scope for preverbal *both*. In terms of geographical location, Ohio State is located in Columbus, Ohio, a city in central Ohio, whereas Rutgers is located in New Brunswick, New Jersey, a city roughly equidistant between Philadelphia and New York City. Ohio State is in the Midwest; Rutgers is in the East. In terms of dialectal regions, Columbus sits at the border between the northern industrial cities dialect area and the Appalachian dialect area. New Brunswick, on the other hand, is situated between the intense linguistic, and social, variation of New York City and the northern border of the mid-Atlantic dialect area. Thus, this phenomenon has a wide geographical and dialectal distribution.

Although these ratings confirm that the current experimental population interprets preverbal *both* as taking scope over conjoined VPs and conjoined NPs, there is an apparent discrepancy between the overwhelmingly number of conjoined NPs obtained in the sentence completions (approximately 80%) and the relatively moderate proportion of GOOD or completely acceptable responses (approximately 50%) for that same sentence type. Why might participants produce conjoined NPs so frequently in sentence completions, but then not admit to using the construction in the acceptability ratings? One reason for the discrepancy might be the nature of the two tasks. Completing sentences is essentially a production task in which participants need only use their implicit knowledge of grammar. Acceptability rating, however, is a metalinguistic task that requires participants to make use of their explicit knowledge as well. Thus, the surprisingly small proportion of GOOD responses might reflect an awareness of the prescriptive rules of English. It might also reflect an inability of participants to think of an appropriate context for a sentence that was presented in isolation and that they themselves did not produce. In any event, the data demonstrate that these sentences are

not merely oddities that arise from the sentence completion task; rather they are acceptable constructions for many speakers of mainstream American English.

The Production Experiment

The production experiment was designed to investigate two questions. First, what is the nature of the prosodic information that might differentiate a particular temporary syntactic ambiguity? Second, how consistently do speakers produce those prosodic cues? I recorded speakers producing target utterances like those in (2), and then analyzed the prosodic structure of those utterances. On the basis of that analysis, I will argue that most speakers fail to produce consistent prosodic cues, at least in the region preceding any disambiguating lexical information, but when they do prosodically disambiguate conjunction attachment, they manipulate phrasing immediately before the conjunction word, and accentuation of the conjunction, as well.

Prosodic analysis. Although much of the work on the relationship between prosody and sentence processing has focused on the influence of such factors as pitch rise, pitch fall, and syllable duration (Beach, 1991; Marslen-Wilson et al., 1992; Nagel et al., 1994; Stirling & Wales, 1996; Watt & Murray, 1996), such analyses fail to take into account the prosodic structure that these factors help create. In the current study, by contrast, the utterances obtained in the production experiment were prosodically annotated using the ToBI (Tones and Break Indices) system (Silverman et al., 1992).

ToBI is a system for annotating the alignment of tune to text in some varieties of English, including mainstream American English. It is a system for expressing the relative prominences within an utterance, namely those syllables that native listeners perceive as being accented, and the relative groupings of words into intonational phrases. In other words, ToBI is a system for annotating the tune and prosodic structure of an utterance.

A ToBI transcription consists of a recording of the utterance that can be played in its entirety, or in smaller segments, as often as needed, a fundamental frequency contour showing the rises and falls in pitch, and symbolic codings on four tiers. The orthographic tier displays the words in the utterance; the miscellaneous tier provides a space for transcribers to annotate dysfluencies such as coughs; the break index tier displays subjective measures of the relative degree of disjuncture between adjacent words; and the tone tier displays the location and type of particular tones.

ToBI captures those stresses (or accentual prominences), and prosodic groupings (or phrasings), that are not predictable from the dictionary. For example, although native speakers of English know that the first syllable of *window* is stressed, they cannot predict whether *window* will be the most prominent word in a phrase such as *The window is open*. If the phrase is the answer to the question *What about the window?* the word *open* is likely to be most prominent (*The window is OPEN*). In contrast, if the phrase is a response to the command *Open the window*, when the window is already open, *is* is likely to be most prominent (*The window IS open*). It is also possible, although not necessary, to produce the syntactic phrase *The window is open* as two prosodic phrases, with a break between *the window* and *is open*, and with independent prominences on *window* and *open*. It is these sorts of unpredictable pieces of the tune and rhythm that ToBI is designed to express.

A closed set of pitch accents, including such tones as H*, L*, and L+H*, is used to represent the phrase-level stresses (or accentual prominences). Each accent is associated with a particular syllable. There are two levels of phrasing, the intermediate phrase and the larger intonational phrase. One or more pitch accents are nested within an intermediate phrase, and one or more intermediate phrases are nested within an intonational phrase. The right edge of an intermediate phrase is marked by a phrase accent, H- or L-. The phrase accent fills up the space between the last pitch accent and the right edge of the intermediate phrase. The right edge of an intonational phrase is marked by a boundary tone, H% or L%. In addition, the last pitch accent within an intermediate phrase is called the nuclear pitch accent, and it is typically perceived as being the most stressed within that phrase. Every utterance is spoken with at least one intermediate phrase and one intonational phrase. Speakers manipulate these components of pitch accents, phrase accents, and boundary tones, and intermediate and intonational phrases, to produce a tune or particular intonation for an utterance, and these are the components that were used to assess the prosodic structure of the utterances elicited in the production experiment.

Methods

Participants. 5 native speakers of mainstream American English were each paid \$5 for their participation.

Materials and procedure. Speakers were recorded individually, as they produced twelve tokens of conjoined VP sentences (e.g., *Mary both sold vegetables and bought seeds at the fair*) and a matching set of twelve conjoined NP sentences (e.g., *Mary both sold vegetables and seeds at the fair*). A complete list of the target utterances is included in the appendix. Each recording session lasted 20 minutes and was conducted in a sound attenuated room. Speakers wore a Shure SM10A head-mounted microphone and were recorded on a TEAC V-427C stereo cassette deck using a TDK D90 cassette tape.

Prior to recording, each speaker was provided with written instructions that included a discussion of the structural contrast between the two types of sentences that speakers would be producing. In addition, the instructions stated that the experimenter was interested in the *way or ways that native speakers produce these two types of utterances*. Each speaker was prompted by the instructions and by the experimenter to ask for clarification about the uses of *both* or about the task itself; none requested any.

During a brief practice session, the experimenter set the recording level while speakers read contexts and sentences similar in kind to those used during the actual experiment. After the practice session, the utterances were obtained the following way: the experimenter read a two line context followed by the question, *Did you know that?* The speaker then produced the target utterance in response. Both the experimenter and the speaker read from a matching set of index cards on which they were able to see each context, question, and target utterance in its entirety. The index cards were divided into two blocks, with an equal number of conjoined NP and conjoined VP sentences in each block. If block one contained a particular conjoined NP sentence, the matching conjoined VP sentence was in block two. The order of block presentation was alternated between speakers, and for each speaker, the cards within each block were shuffled.

The same context was used for each member of an item pair, and the same question was used to prompt every utterance. This strategy is important because prosodic choices often reflect whether information in an utterance is new to the discourse or already shared by the parties involved (Chafe, 1976; Prince, 1981). By holding the questions and contexts constant, any prosodic differences could be attributed to the syntactic ambiguity in the target utterances, rather than to differences in information status induced by the context.

A total of 120 utterances were elicited (5 speakers x 24 sentences). ESPS/waves+ 5.0 was used to digitize each utterance and to calculate fundamental frequency contours.

Each utterance was digitized at 16 bits with a sampling rate of 16kHz, and each fundamental frequency contour was calculated using “get_f0,” an autocorrelation-based F0-tracking utility program that is part of Entropic’s ESPS/xwaves package. An experienced ToBI transcriber, who was naïve to the design and purpose of the experiment, annotated all of utterances.

The prosodic structure of each of the 120 ToBI transcriptions was analyzed for two types of prosodic cues. The first consisted of any consistent differences in the location of pitch accents, phrase accents, and boundary tones. The second consisted of any consistent differences in the particular tones used within each set of pitch accents, phrase accents, and boundary tones.

Results

Although there were no consistent differences in the particular types of accents (e.g., H* vs. L+H*, H % vs. L%) used by any of the five speakers, Speaker 4 (S4) did consistently manipulate the presence of a phrase accent and a pitch accent around the conjunction. Importantly, this manipulation occurred prior to the onset of the disambiguating region, the verb or noun of the second conjunct.

	↙ disambiguating region
Mary both sold vegetables and	<u>bought</u> seeds at the fair.
Mary both sold vegetables and	<u>seeds</u> at the fair.

Figure 4. Example of Critical Items Marked for Onset of Disambiguating Region

To make a difference in phrasing, S4 produced a phrase accent after the stressed syllable in *vegetables* in all twelve conjoined VP sentences, but in only one conjoined NP sentence. Thus, there was a contrast between an intermediate phrase boundary following the first conjunct in the conjoined VP constructions (separating *vegetables* and *and*) and no boundary in the conjoined NP constructions.

In addition, S4 accented *and* in ten of the twelve conjoined VP sentences, but always left the conjunction unaccented in the conjoined NP sentences. Thus, there was a contrast between a stressed *and* in the conjoined VP constructions and an unaccented *and* in the conjoined NP constructions. Examples of ToBI transcriptions containing these contrasts of phrasing and accentuation are given in Figures 5 and 6.

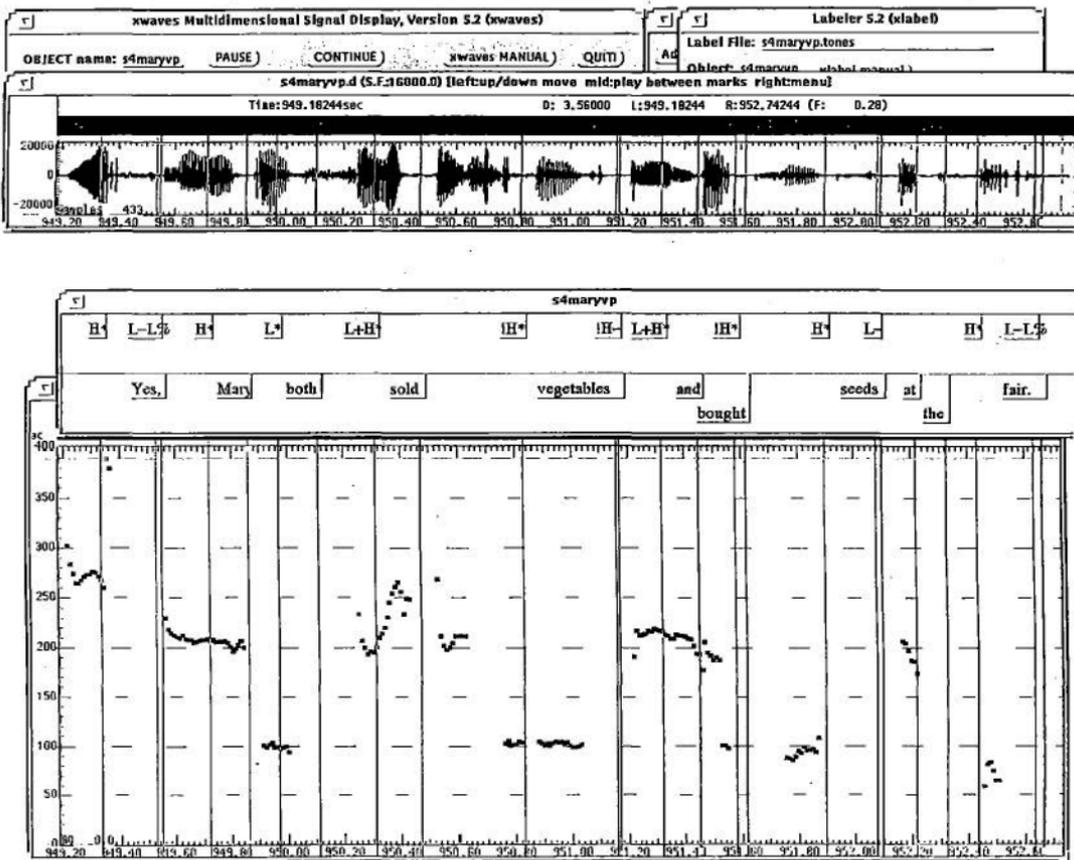


Figure 5. ToBI printout of conjointed VP

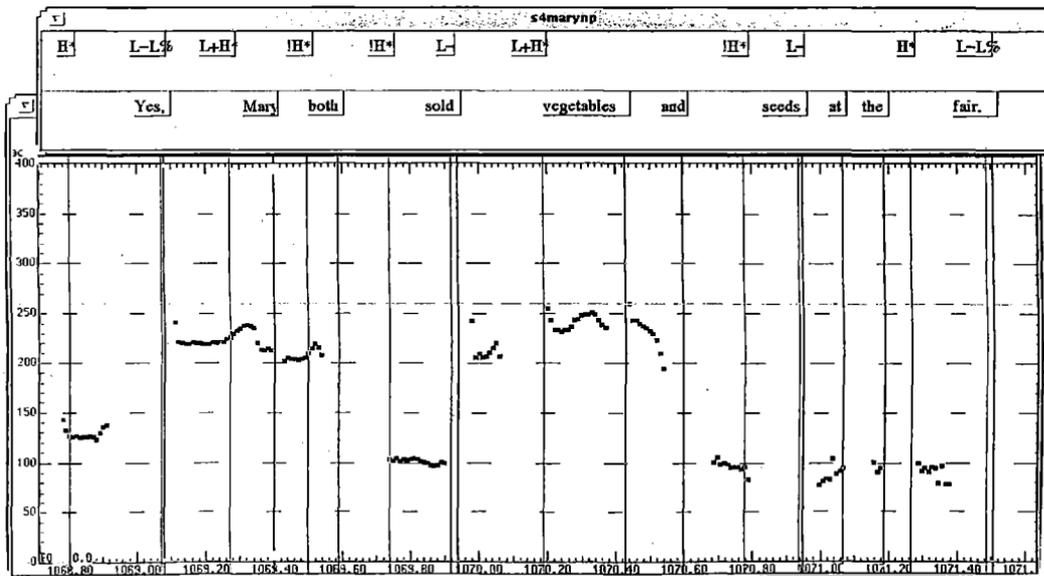
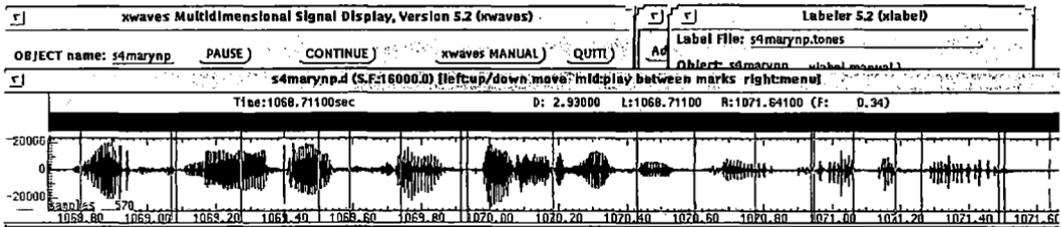


Figure 6.

TobI printout of conjoined NP

As shown in Table 1, none of the other speakers manipulated phrasing and accentuation around the conjunction as consistently as S4. Looking only in the last row in Table 1, which summarizes the frequency of a phrase break occurring before *and* and the frequency of an accent occurring on *and* across speakers, there seems to be a strong tendency for conjoined VP constructions to be produced with an intermediate phrase break just before *and*. There does not seem to be any similarly strong constraint on phrasing in conjoined NP constructions. There also do not seem to be any constraints on the accentuation of *and* in either conjoined VPs or conjoined NPs.

Phrasing: intermediate phrase boundary before <i>and</i>			Accentuation: <i>and</i> is accented		
	Conj VP (n=12)	Conj NP (n=12)		Conj VP (n=12)	Conj NP (n=12)
S1	10	11	S1	6	10
S2	12	6	S2	1	1
S3	12	12	S3	10	8
S4	12	1	S4	10	0
S5	12	9	S5	12	12
total (prop.)	58 (0.97)	39 (0.65)	total (prop.)	39 (0.65)	31 (0.52)

Table 1. Number of Conjoined VPs and Conjoined NPs As Produced by Each Speaker That Show the Prosodic Cue of Phrasing Before *and* and the Prosodic Cue of Accentuation on *and*

Discussion

Of the five speakers who participated in the production experiment, only one of them (S4) consistently produced potentially disambiguating prosodic cues. This speaker manipulated phrasing around, and the accentuation of, the conjunction word *and*.

By inserting a phrase accent between the conjunction *and* and the preceding noun in the conjoined VP construction, S4 produced an intermediate phrase boundary between the first conjunct and the conjunction, thereby dividing the conjuncts into at least two

parts as schematized in Figure 7. In these constructions, the larger syntactic unit that contained the two conjuncts did not fall within a single intermediate phrase.

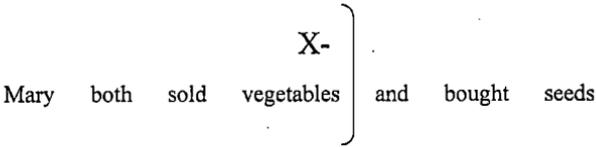


Figure 7. Example of a Phrase Accent Marking the Right Edge of an Intermediate Phrase and Dividing the Conjuncts

In contrast, by not placing an intermediate phrase break in the same location in the conjoined NP construction, it remained possible for the two conjuncts to be produced within the same intermediate phrase. In fact, S4 did just this in eleven of the twelve conjoined NP constructions. The larger syntactic unit containing the two conjuncts occurred within a single intermediate phrase.

By accenting *and* in almost all of the conjoined VP sentences, S4 made the conjunction perceptually more salient as compared to its unaccented match in the conjoined NP sentences. This is not to say, however, that *and* was perceived as the most stressed syllable in its intermediate phrase, for it is the last pitch accent in an intermediate phrase, the nuclear accent, that is typically perceived as the most stressed. Although the pitch accent associated with *and* was nuclear in only two of the ten cases, accenting *and* in the conjoined VP sentences may have been a way for S4 to clearly mark the onset of a new intermediate phrase after the first noun.

The results of the production experiment suggest that although manipulating phrasing and accentuation around *and* might disambiguate coordinate structures prosodically, most speakers fail to consistently provide those cues, even when they are aware that prosody is at issue. Why might it be that only one speaker out of five consistently produced disambiguating prosody? There are several possible explanations. First, at least for mainstream American English, there is no one-to-one relationship between syntax and prosody; it is possible to produce a given syntactic structure with multiple tunes. Thus, it could be that the variety of tunes obtained in the production experiment simply reflects the variety of naturally occurring tunes. Second, the instructions given to the speakers may have been too subtle. In other words, only S4 produced consistent prosodic cues because only S4 was able to understand the task

properly. Third, it is possible that the task was not conducive to eliciting strong prosodic cues, either because speakers knew that the experimenter was already aware of what they were going to say, or because as speakers read their index cards, they failed to read far enough ahead to plan disambiguating prosody adequately. In a more natural elicitation task, disambiguating cues might be more strongly produced. Finally, the attachment ambiguity of *and* is a temporary one, and in the presence of disambiguating lexical information, the need for prosodic disambiguation might be obviated.

Although only one speaker out of five consistently produced disambiguating prosodic cues, the results of the production experiment do provide some answers to the questions raised at the beginning. What is the nature of the prosodic information that might differentiate a syntactically ambiguous construction? Those components include phrasing and accentuation. Are speakers consistently producing prosodic cues? With the exception of S4, the answer is "no." Nonetheless, there are several interesting follow-up questions to be pursued in the perception experiment. Not only is it possible to ask if listeners are sensitive to the cues of phrasing and accentuation that S4 manipulated, but is it also possible to ask two new questions. One, to what extent are the phrasing around, and the accentuation of, the conjunction word *and* independent or related prosodic cues? Two, if the cues are independent, do they signal the same or different information?

The Perception Experiment

Even though there was a great deal of variety in the utterances that the speakers in the production experiment produced, one speaker (S4) did consistently provide potentially disambiguating prosodic cues. Furthermore, those cues occurred prior to the onset of lexical information that resolved the temporary syntactic ambiguity in question, namely the attachment site of *and* in conjoined VP and conjoined NP constructions. Thus, the consistency with which S4 produced disambiguating prosodic cues makes it possible to address the third and fourth questions raised at the beginning of this paper: Are listeners sensitive to prosodic information that might differentiate a temporary syntactic ambiguity, and if so, is there evidence that those are the cues that are used to resolve the ambiguity? In addition, it is possible to raise two new questions that evolved out of the production experiment. To what extent do the phrasing around, and the accentuation of, the conjunction word *and* represent independent or related prosodic cues? If the cues are independent, do they signal the same or different information?

With respect to these new questions, there are several reasons to expect that the absence of a phrase break before *and* is the primary prosodic cue, while the accentuation of *and* serves a secondary purpose. It is apparent from the totals in Table 1 that there is a strong tendency for conjoined VP constructions to elicit an intermediate phrase boundary after *vegetables*. However, because the presence of a phrase accent is “required” for conjoined VPs but “optional” for conjoined NPs, the only way in which that cue could be informative, is for its absence to signal a conjoined NP.¹ Thus, the absence of a phrase accent is likely to signal the presence of conjoined NPs, independently of the accentual status of *and*. Furthermore, the absence of a phrase break is likely to be a general cue to conjoined NPs, as opposed to a speaker specific one, for two reasons. First, the constraint on “obligatory” phrasing in conjoined VPs and “optional” phrasing in conjoined NPs extends across all five speakers. Second, work by Schafer & Speer (1998) and Kjelgaard & Speer (1999) also suggests that intermediate phrases play a privileged role in syntactic parsing.

The presence of an accent on *and*, on the other hand, is “optional” for both types of constructions. Therefore, its presence and/or absence is uninformative, and it is unlikely that the accentual status of *and* acts by itself as a cue to syntactic structure across speakers. It is possible, however, that listeners are sensitive to strategies employed by specific speakers, in which case the accentual status of *and* might act as a signal to syntactic structure when it is used consistently by a specific speaker, such as S4. A second, and more likely possibility, is that accenting the conjunction when it immediately follows a phrase accent (as S4 consistently did) represents a strategy for clearly marking the onset of a new intermediate phrase.

The perception experiment was designed to investigate whether or not listeners could perceive the prosodic cues that S4 produced and use those cues to identify sentence fragments as conjoined NPs and conjoined VPs. Listeners were not expected to be able to identify fragments from the other four speakers because their fragments contained neither the phrasing and accentuation cues that S4 had so consistently provided, nor additional prosodic cues. If listeners were actually able to identify the fragments from

¹ I thank Rick Lewis for pointing this out to me. Unfortunately, this makes it impossible to use these constructions to test the influence of prosody in parsing, since conjoining NPs is already the default within both Constraint-based models and the Garden Path model.

speakers other than S4, then that would suggest that there were prosodic cues that had been overlooked. Finally, it was possible to investigate the extent to which the accentuation of, and the phrasing around, the conjunction word *and* represented independent or related prosodic cues. This was done by assessing identification rates to fragments that contained different combinations of these cues.

The results suggest that the manipulations of phrasing and accentuation around *and* are indeed perceptible prosodic cues. The results further suggest that the absence of a phrase break acts as a cue to conjoined NPs, and that this cue is not restricted in its interpretation to the one speaker who produced it the most consistently.

Methods

Participants. The 42 undergraduates who participated in the acceptability ratings participated first in the perception experiment. Responses from the same 2 nonnative speakers of English were again excluded from the Ohio State pool and are not reported here. All remaining respondents were native English speakers.

Materials and procedure. 118 sentence fragments of the form *Mary both sold vegetables and...* were created by splicing each of the 120 utterances that had been collected and digitized for the production experiment. Two conjoined VP utterances, one from Speaker 3 and one Speaker 5, were discarded due to dysfluencies within the fragment. The splice point for each fragment was chosen with two goals in mind, to keep as much of the offset of *and*, and as little of the onset of the following word, as possible. The fragments were randomized and recorded to a TDK D90 cassette tape. Fragments were separated from one another by 3 seconds of silence, a tone, and another 3 seconds of silence. A double tone separated every 10 fragments. The 20 minute tape was then presented to listeners, in groups and individually, on a portable cassette deck.

The fragments were presented in a forced-choice task; listeners had to decide whether each fragment belonged to a conjoined NP or a conjoined VP utterance. Listeners had one of two types of response sheets. Half of the listeners received "whole sentence" response sheets that presented the original sentences in their entirety. The other half received "partial sentence" response sheets that provided two versions of the fragment, one followed by NOUN and the other followed by VERB. This was done to address two concerns.

First, it is important to recognize that there are two types of information provided in each fragment, intonational and segmental. Since the last word in every fragment was *and*, it was likely to be coarticulated with the following word. Listeners who actually knew the word following *and* might be able to make their decisions based on segmental information rather than on intonational information. If listeners were indeed using coarticulatory cues to the following word's initial segments to make their decisions, as opposed to intonational cues to the prosodic parse of the sentence fragment, then they should be able to use that information in both conditions. Thus, there should be higher identification rates for conjoined VP constructions, as well as for conjoined NP constructions, from listeners with whole sentence response sheets.

The second reason for using two types of response sheets stemmed from the overwhelming preference for conjoined NP sentences that had been evident in the sentence completion data. This raised the possibility that there might be an overall bias to identify fragments as conjoined NPs. Because the whole sentence response sheets eliminate the need to think of a likely completion for a given fragment, bias effects should only be evident among those listeners with partial sentence response sheets.

The order of noun option first or verb option first was held constant for each individual response sheet but split among each set of response sheets.

Results

The proportion of each speaker's conjoined VP and conjoined NP fragments that were identified correctly are summarized in Table 2. As expected, listeners performed best at identifying the fragments produced by S4, the one speaker who consistently manipulated phrasing and accentuation around *and*, and performed poorly for the other 4 speakers who failed to produce consistent prosodic cues.

Proportion of Fragments Correctly Identified		
Speaker	Conj VP	Conj NP
S1	0.56	0.47
S2	0.45	0.55
S3	0.47	0.55
S4	0.60	0.62
S5	0.42	0.56

Table 2. Proportion of Fragments Correctly Identified by Speaker and Fragment Type

The proportion of fragments identified as conjoined VP by subjects (F1) was submitted to a 2 (phrase type: conjoined NP or conjoined VP) x 5 (speaker) repeated measures ANOVA. Response sheet (whole sentence or partial) and order (noun option first or verb option first) were treated as between subjects variables.² The same proportion of responses was assessed by items (F2)³ using a 2 (response sheet) x 2 (order) x 5 (speaker) repeated measures ANOVA. The results showed a main effect of phrase type [$F(1,36) = 11.598, p \leq 0.002$; $F(1,9) = 11.605, p \leq 0.008$], a main effect of speaker [$F(4,144) = 5.676, p \leq 0.001$; $F(4,36) = 4.766, p \leq 0.003$], and a speaker x phrase type interaction [$F(4,144) = 8.931, p \leq 0.001$; $F(4,36) = 11.435, p \leq 0.001$].

Looking at the data in Table 2, the positive responses to S4 seem likely to be contributing heavily to the speaker x phrase type interaction. To investigate this possibility, two additional ANOVAs were conducted. The first included only responses to S4 and the second excluded responses to S4.

The results of the S4-only ANOVA showed a strong main effect of phrase type [$F(1,36) = 25.884, p \leq 0.001$; $F(1,9) = 63.143, p \leq 0.001$]. Also in the items analysis, there was a reliable main effect of order [$F(1,9) = 5.661, p \leq 0.04$], and a marginal phrase type x list interaction [$F(1,9) = 4.457, p \leq 0.06$].

The results of the ANOVA that excluded S4 showed a main effect of speaker [$F(3,108) = 7.609, p \leq 0.001$; $F(3,27) = 7.181, p \leq 0.001$], and as expected, no speaker x

² Because there were 19 Ohio State listeners, but 21 from Rutgers, school was not included as a variable.

³ Two items were excluded from analysis because of an uneven number of cells. Because two speakers had produced dysfluent fragments, two items were missing a speaker/phrase type entry.

phrase type interaction. By subjects, there was a marginal speaker \times order interaction [$F_1(3,108) = 2.404, p \leq 0.072$]. By items, there was a list \times order interaction [$F_2(1,9) = 5.335, p \leq 0.046$] and a phrase type \times list \times speaker interaction [$F_2(3,27) = 4.031, p \leq 0.017$].

The results suggest that listeners are in fact able to use the prosodic cues provided by S4 to identify her conjoined VP and conjoined NP fragments. This analysis is further supported by a set of post-hoc *t*-tests that were based on the total proportion of conjoined NP and conjoined VP fragments identified correctly, by 40 listeners, for each speaker. The only statistically significant response belonged to S4 [$t(79) = -4.169, p \leq 0.001$]; the responses to the other four speakers were nonsignificant.⁴

The absence of any reliable effects of response sheet in the primary ANOVA suggests that listeners were indeed relying on intonational information, and not segmental information, to make their decisions. In further support of this interpretation, listeners with whole sentence response sheets did not perform better at identifying both types of fragments.

There is some suggestion, however, that listeners are biased to identify fragments as conjoined NPs. As shown in Table 2, listeners are more likely to identify fragments as conjoined NPs when those fragments are produced by S2, S3, and S5. These are three of the four speakers who failed to produce consistent prosodic cues. Additionally, listeners with partial sentence response sheets were more likely to identify both types of fragments as conjoined NPs.

In order to assess the relationship between the phrasing and accentuation around *and*, responses to groups of fragments that contained various combinations of those two prosodic cues were analyzed. Table 3 presents a summary of different subsets of fragments and the proportion of conjoined VP identifications for each of those different subsets. For example, across all five speakers, ninety five fragments contained a phrase accent between *and* and the preceding object noun; twenty three fragments did not. Sixty four fragments contained both a phrase accent between *and* and the preceding object noun, and a pitch on *and*; nineteen fragments contained neither.

⁴ It is surprising that the results from a signal detection theory analysis suggest that the conjoined NP prosody and conjoined VP prosody of S4 are not discriminable ($D\text{-prime} = 0.534, \beta = 1.013, \text{criterion} = 0.291$). I am not sure at this point how to reconcile these findings with the ANOVAs and *t*-tests.

Excluding S4, fifty eight fragments contain a pitch accent on *and*; thirty six fragments did not.

Proportion of Subsets of Fragments Identified as Conjoined VPs (n = number of tokens)					
	all speakers			all speakers except S4	
	with tonal cue	without tonal cue		with tonal cue	without tonal cue
phrase break before <i>and</i>	0.50 (n = 95)	0.38 (n = 23)	phrase break before <i>and</i>	0.48 (n = 82)	0.39 (n = 12)
accent on <i>and</i>	0.50 (n = 68)	0.44 (n = 50)	accent on <i>and</i>	0.47 (n = 58)	0.46 (n = 36)
phrase break and accent	0.50 (n = 64)	0.38 (n = 19)	phrase break and accent	0.48 (n = 54)	0.37 (n = 8)

Table 3. Proportion of Subsets of Fragments Identified as Conjoined VPs

The data presented in Table 3 provide evidence that listeners interpret the absence of a phrase break before *and* as a general prosodic cue; interpretation of this cue is not restricted to its most robust source, S4, but rather extends across speakers. The proportion of conjoined VP identifications decreases whenever a phrase accent is omitted. This occurs even when responses to S4 are excluded from analysis, as shown on the right hand side of the table.

Although there also seems to be a slight effect of the absence of accentuation, the effect is not particularly robust, and it disappears when S4's fragments are excluded. Furthermore, there is no evidence that the absence of an accent on *and* combines with the absence of a phrase break to further decrease the number of conjoined VP identifications. Thus, these data provide evidence that across speakers, listeners do not interpret either the presence or absence of an accent on *and* as a cue to syntactic structure.

Discussion

The results of the perception experiment provide evidence that listeners are able to perceive the prosodic cues produced by S4 and use those cues to identify her conjoined

VP and conjoined NP fragments. The data also suggest that listeners show a slight bias to identify fragments as conjoined NPs, particularly in the absence of prosodic cues.

Although S4 manipulated phrasing and accentuation around the conjunction word *and*, these two potentially disambiguating prosodic cues were interpreted by listeners in different ways. In terms of phrasing, its absence was interpreted as a cue to conjoined NPs, across speakers. This finding accords well with the differences in distribution of phrasing obtained in the production experiment, and with evidence from other experimental work regarding the privileged role that intermediate phrases play in parsing (Schafer & Speer, 1998; Kjelgaard & Speer, 1999).

In contrast, the presence or absence of an accent on *and* did not act as a general cue to syntactic structure, and based on its distribution in conjoined NPs and conjoined VPs, it could not have. With respect to the speaker who consistently manipulated the accentual status of *and*, S4, accenting the conjunction in the conjoined VP sentences may have been a way of clearly marking the onset of a new intermediate phrase. However likely this may be, I cannot rule out the possibility that the accentual status of *and* was interpreted by listeners as a speaker specific cue to syntactic structure. Not only were listeners able to identify S4's conjoined NPs, most likely on the consistent absence of a phrase break in those constructions, but they were also able to identify her conjoined VPs. Since the presence of a phrase break cannot act as a cue to conjoined VPs, some other prosodic cue must have been providing that information. One possibility then, is that listeners were sensitive to the way in which S4 manipulated the accentuation of *and*, and they reserved the interpretation of accented *and* for only her utterances.

Of course, it is not possible from these data to know whether listeners were using only accented *and* to detect S4's conjoined VPs, or whether the relevant prosodic cue was some other event or combination of events. For example, whenever S4 used accented *and*, she also produced a phrase break before it. Thus, speakers might have been sensitive to the combined effect of a phrase break preceding the accented conjunction. Another possibility, and one that I am actively investigating, is that listeners were sensitive to a particular tonal contour that S4 produced over the first VP conjunct. This contour, a prenuclear L+H* immediately followed by a downstepped nuclear H* (!H*), can be seen in the ToBI transcription in Figure 5. This pattern holds promise for explaining the apparent VP bias shown in listener responses to S1, as well.

In sum, even though S4 consistently manipulated both phrasing and accentuation around the conjunction, only the absence of a phrase break is interpreted as a reliable prosodic cue to the syntax of coordinate structures, independent of the speaker who produced it. The absence of a phrase break signals conjoined NPs, across speakers. In the case of the accentual status of *and*, it is clear that no such general interpretation exists. Furthermore, the extent to which it acts as a cue to other prosodic structures, such as the onset of a new intermediate phrase, or acts as a speaker specific cue to syntactic structure remains an area for future research.

General Discussion

The combined results of the production and perception experiments provide evidence that speakers can in fact consistently produce prosodic cues as a means of disambiguating the temporary syntactic ambiguity inherent in coordinate structures, although most of them fail to do so, at least in the region preceding disambiguating lexical information. The results also suggest that speakers are likely to make use of phrasing and accentuation around the conjunction when asked to provide disambiguating prosody, although there seems to be some variation in the ways that listeners interpret those two cues. In the case of phrasing, listeners are sensitive to the general constraint that the presence of a phrase break is optional only in the case of conjoined NPs. In the case of accentuation, no such constraint exists, and thus, the accentual status of *and* cannot and does not act as a general cue to syntactic structure.

The role of accentuation highlights one of the more interesting, and ongoing, research questions raised by this data. Given the general bias to identify fragments as conjoined NPs, and given that the informativeness of phrasing is limited to the absence of a phrase break signaling conjoined NPs, what accounts for the large proportion of correct VP identifications for S4 and for what seems to be a VP bias in the case of S1? As mentioned above, one possibility is that listeners were sensitive to S4's manipulation of the accentual status of *and*. Another possibility, and one that I am actively investigating is that listeners were sensitive to a particular tonal contour, or section of tune. This tonal contour has possible ramifications for work on *association with focus* (Jackendoff, 1972; Rooth, 1992) and the scope ambiguity of preverbal *both*.

References

- Beach, C. M. (1991). The interpretation of prosodic patterns at points of syntactic structure ambiguity: evidence for cue trading relations. *Journal of Memory and Language*, **30**, 644-663.
- Blodgett, A. & Boland, J. E. (1998). Understanding *both* in coordinate structures: The OSU subject pool both surprised Allison and Julie! Poster presented at the Eleventh Annual CUNY Conference on Human Sentence Processing, New Brunswick, March, 1998.
- Boland, J. E. (1997). The relationship between syntactic and semantic processes in sentence comprehension. *Language and Cognitive Processes*, **12**, 423-484.
- Chafe, W. (1976). Givenness, contrastiveness, definiteness, subjects, and topics. In Charles N. Li, ed., *Subject and Topic* (pp. 25-55). New York: Academic Press.
- Frazier, L. (1978) On comprehending sentences: Syntactic parsing strategies. Doctoral Dissertation, University of Connecticut.
- Frazier, L. & Clifton, C. (1996) *Construal*. Cambridge, MA: The MIT Press.
- Jackendoff, R. (1972). Focus and presupposition. In *Semantic interpretation in generative grammar* (pp. 229-278). Cambridge, MA: MIT Press.
- 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.
- MacDonald, M., Pearlmutter, N., & Seidenberg, M. (1994) Lexical nature of syntactic ambiguity resolution. *Psychological Review*, **101** (4), 676-703.

Marslen-Wilson, W. D., 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.

McRae, K., Spivey-Knowlton, M., & Tanenhaus, M. (1998) Modeling the influence of thematic fit (and other constraints) in on-line sentence comprehension. *Journal of Memory and Language*, **38** (3), 283-312.

Nagel, H. N., Shapiro, L. P., & Nawy, R. (1994). Prosody and the processing of filler-gap sentences. *Journal of Psycholinguistic Research*, **23** (6), 473-485.

Prince, E. (1981). Toward a taxonomy of given – new information. In Peter Cole, ed., *Radical Pragmatics* (pp. 223-255). New York: Academic Press.

Rooth, M. (1992). A theory of focus interpretation. *Natural Language Semantics*, **1**, 75-116.

Schafer, A. J. & Speer, S. R. (1998). Effects of prosodic phrasing on the resolution of within-category lexical ambiguity. Poster presented at the Eleventh Annual CUNY Conference on Human Sentence Processing, New Brunswick, March, 1998.

Silverman, K., Beckman, M., Pitrelli, J., Ostendorf, M., Wightman, C., Price, P., Pierrehumbert, J., & Hirschberg, J. (1992). TOBI: a standard for labeling English prosody. *Proceedings of the 1992 International Conference on Spoken Language Processing*, **2**, 867-870, Banff, Canada.

Stirling, L. & Wales, R. (1996). Does prosody support or direct sentence processing? *Language and Cognitive Processes*, **11**, 193-212.

Watt, S. M. & Murray, W. S. (1996). Prosodic form and parsing commitments. *Journal of Psycholinguistic Research*, **25** (2), 1996.

Appendix

Listed below are the materials from the production and perception experiments. Conjoined NP and conjoined VP sentences have been collapsed; the second verb is marked in parentheses. A “/” marks the end of the fragment used in the perception experiment. An “*” marks the two dysfluent conjoined VPs that were excluded from the perception experiment.

1. Ohio’s vegetable farmers always do a lot of business at the state fair, but last year it was more than farmers. Columbus city gardeners participated, too. Did you know that?

Yes, Mary both sold vegetables and / (bought) seeds at the fair.

2. Last summer there was a power outage during one of the worst heat waves ever. People went crazy. Did you know that?

*Yes, looters both destroyed stores and / (stole) cars during the blackout.

3. David and his friends like getting together at Barley’s on Friday nights. Last week they ended up staying for several hours. Did you know that?

Yes, the gang both had sandwiches and / (drank) beer at the brewpub.

4. The ranger really enjoyed his post at Yellowstone National Park. He was in pretty close contact with the wildlife. Did you know that?

Yes, the ranger both tracked elk and / (saw) bear throughout the park.

5. The choir met in the church basement to begin learning their first number. They were there for hours. Did you know that?

Yes, the choir both whistled choruses and / (sang) verses during the song.

6. The gardener woke up early each day ready to work in her greenhouse. She easily spent several hours there every day. Did you know that?

Yes, the gardener both selected herbs and / (chose) flowers during the morning.

7. The student enjoyed last quarter's English class. It was the first one he had taken at OSU. Did you know that?

Yes, the student both learned stories and / (memorized) poems from the book.

8. The coach of the little league team was always at his busiest at the beginning of the season. He had several responsibilities. Did you know that?

Yes, the coach both picked players and / (held) equipment for the team.

9. Christopher worked for a landscaping company last Friday. He said it was hard work. Did you know that?

Yes, Christopher both carried rocks and / (brought) tools to the site.

10. The maid complained about her job to anyone who would listen. She worked her fingers to the bone. Did you know that?

Yes, the maid both scrubbed dishes and / (did) floors all by hand.

11. The officer did a pretty thorough job looking into the robbery at the high school. He took notes on everyone. Did you know that?

*Yes, the officer both investigated parents and / (met) teachers before the assembly.

12. William always had a good time playing construction site at the vacant lot. He would play there for hours. Did you know that?

Yes, William both shoveled dirt and / (threw) stones just for fun.