PROSODIC PARSING:

THE ROLE OF PROSODY IN SENTENCE COMPREHENSION

A Dissertation Presented

by

AMY J. SCHAFER

Submitted to the Graduate School of the
University of Massachusetts Amherst in partial fulfillment
of the requirements for the degree of

DOCTOR OF PHILOSOPHY

September 1997

Department of Linguistics

PROSODIC PARSING:

THE ROLE OF PROSODY IN SENTENCE COMPREHENSION

A Dissertation Presented

by

AMY J. SCHAFER

Approved as to style and content by:

_____
Lyn Frazier, Chair

_____
Charles Clifton, Jr., Member

_____
Kyle Johnson, Member

_____
Elisabeth Selkirk, Member

_____
Ellen Woolford, Department Head
Department of Linguistics

ACKNOWLEDGMENTS

This research was supported in part by Grants HD 18708, HD 07327, and MH 51768 from the NIH.

ABSTRACT

PROSODIC PARSING:

THE ROLE OF PROSODY IN SENTENCE COMPREHENSION

SEPTEMBER 1997

AMY J. SCHAFER, B.A., WASHINGTON UNIVERSITY

Ph.D., UNIVERSITY OF MASSACHUSETTS AMHERST

Directed by: Professor Lyn Frazier

This work presents an investigation of how prosodic information is used in natural language processing and how prosody should be incorporated into models of sentence comprehension. It is argued that the processing system builds a prosodic representation in the early stages of processing, and is guided by this prosodic representation through multiple stages of analysis. Specifically, the results of four sentence comprehension experiments demonstrate that prosodic phrasing influences syntactic attachment decisions, focus interpretation, and the availability of contextual information in the resolution of lexical ambiguity. Two explicit hypotheses of how prosodic structure is used in processing are proposed to account for these effects; one which accounts for effects of phonological phrasing on syntactic processing decisions and a second which accounts for effects of intonational phrasing on semantic/pragmatic interpretation.

Three sources of evidence are provided in support of the central claim that the processor must build and use a prosodic representation from the early stages of processing. First, an experiment on the resolution of prepositional phrase attachment ambiguity demonstrates that syntactic attachment decisions are influenced by the overall pattern of phonological phrasing in utterance, and not simply by prosodic boundaries located at the point of syntactic ambiguity. Thus, the effects of a single kind of prosodic element, at a single level in the prosodic hierarchy, must be accounted for with respect to the larger prosodic structure. A second experiment shows that the interpretation of focus is dependent on both

the pattern of pitch accents in the utterance and the pattern of prosodic phrasing, establishing that different kinds of prosodic elements in the prosodic structure are used jointly in processing decisions. Two additional experiments, one on the interpretation of context-sensitive adjectives and a second on the resolution of within-category lexical ambiguity, demonstrate that phonologically distinct levels of prosodic phrasing have separable effects on language processing.

Taken together, the four experiments suggest that prosody has a much broader role in sentence comprehension than previously recognized, and that models of sentence processing should be modified to incorporate prosodic structure.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

PRELIMINARIES

1.1 Introduction

It has long been recognized that prosodic structure is an important aspect of spoken language. Among other things, prosody can establish focus, distinguish given and new information, disambiguate certain structurally ambiguous strings, and convey emotions and attitudes. The sentences in (1.1) illustrate some of these effects of prosodic structure.

(1.1) Prosodic Signals of Focus, Phrasing, and Attitude:
a. This deli only sells strawberry gelatin salads.
b. The old men and women stayed home.
c. The South College ceilings are flawless!

Sentence (1.1a) can differ in meaning depending on where accent, and therefore focus, falls in the VP. Sentence (1.1b) illustrates a form of structural ambiguity—whether *old* modifies just *men* or the conjoined NP *men and women*—that can be disambiguated through prosodic phrasing (Lehiste, 1973). Sentence (1.1c) would most likely be produced with strong prosodic markers of sarcasm by anyone familiar with the current South College ceilings.

Although such effects of prosody on interpretation are easily observed, and various aspects of prosody have been studied for many years, we still lack a comprehensive explanation of how prosody is constrained by the grammar and interpreted by the language processing system. Of course, some aspects of prosody and their uses are comparatively clear. For example, the prosodic representation developed by Pierrehumbert and her colleagues (see Section 1.2.1) has provided a reliable and widely-accepted account of the range of phonologically well-formed prosodic contours for English. The effects of high pitch accents in the specification of focus are also understood fairly well, both in terms of grammatical constraints and processing effects. We know many of the phonetic, phonological, syntactic, and semantic properties of focus, and we also know something

about the effects of the presence versus of the absence of accent and focus in speech perception (see Chapter 3). Given a sentence such as (1.1a) and a particular pattern of accentuation, we could list the possible focus structures for the sentence, a set of grammatical constraints that allow those focus structures, and perhaps even some of the processing principles that would be involved in building the correct set of linguistic representations for each interpretation of the sentence.

In spite of these advances, other aspects of prosody, such as the connection between prosodic phrasing and syntactic structure, have been much less clear. Since English seems to allow many different patterns of prosodic phrasing for any given syntactic structure (see Figure 1.1 for an example of this), it seems likely that phonosyntactic constraints on prosodic phrasing only partially determine the prosodic structure of a sentence. The selected phrasing appears to be the result of the interaction of multiple factors, such as the kind and placement of pitch accents, the lengths of words and syntactic constituents, the speech rate, the syntactic structure, and the focal structure. A fair amount of optionality in prosodic phrasing may exist even when these factors are fixed. Though it may be apparent for a given syntactic structure that certain prosodic phrasings are acceptable and others are unacceptable (or for a given prosodic phrasing of a string of words that some syntactic representations are possible and others are impossible), we do not yet know the complete set of grammatical constraints which account for these patterns. Further, although we know that two well-formed productions of a sentence which differ in prosodic phrasing can have very different effects on processing (see Section 1.3), explicit general principles of processing which could account for these differences have not been formulated. In fact, experimental studies of prosody have rarely even offered accounts of why the particular contours used in the experiment should lead to processing differences for the syntactic structure that was tested. In short, given a sentence such as (1.1b) and a particular pattern of prosodic phrasing, we have no satisfactory specification of which grammatical constraints and processing principles account for the interpretation(s) of the sentence.

Perhaps because of the many factors that must be addressed when researching prosody, and also because of the great advances that have been made in sentence processing research with visually-presented materials, the dominant models of sentence processing have virtually ignored the role of prosody, leaving us with an impoverished theory of sentence processing. The goal of this work is to address this gap by beginning to lay out a general description and explanation of how prosodic features are employed in sentence comprehension. In particular, this work will investigate the effects of prosodic phrasing on syntactic and post-syntactic processing decisions for American English. Based on the results of four sentence comprehension experiments, I will argue that prosodic phrasing plays a crucial role in sentence comprehension, structuring the input and delimiting the relevant domain for processing decisions at multiple levels of representation.

Specifically, I will present experimental results which show that prosodic phrasing can influence phrase structure construction, focus specification, and the availability of sentential context in the resolution of lexical ambiguity. Additionally, I will demonstrate that phonological phrase boundaries and intonational phrase boundaries have separable effects on sentence processing. I will argue that these effects are best accounted for by assuming that the language processor makes active use of the prosodic representation of a sentence throughout processing, and describe how the relevant prosodic information could be incorporated into one of the dominant models of sentence processing, the Garden Path model.

1.2 Theoretical Assumptions

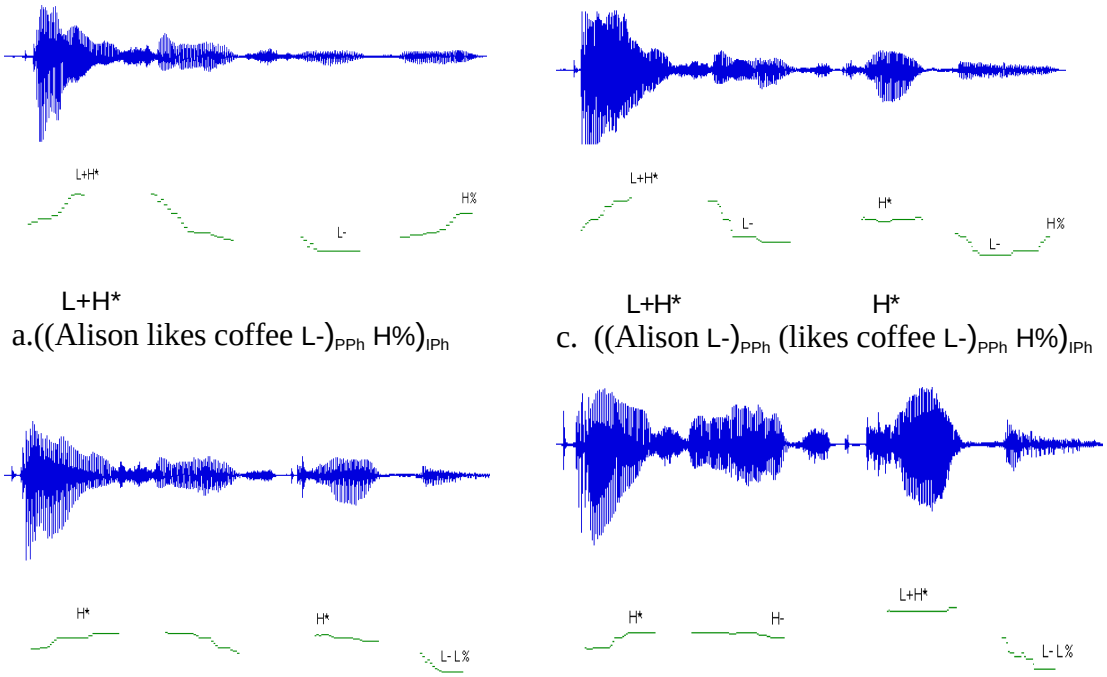
1.2.1 The Prosodic Hierarchy

Throughout this work, I will assume a strictly layered prosodic hierarchy of the sort first described by Pierrehumbert and colleagues and recently incorporated into the ToBI system for prosodic transcription (Pierrehumbert, 1980; Beckman & Pierrehumbert, 1986; Nespor & Vogel, 1986; Beckman & Ayers, 1993). In this system, each English utterance is composed of one or more intonational phrases (IPhs), each of which must end with

either a high boundary tone (H%) or a low boundary tone (L%). Each intonational phrase is composed of one or more phonological phrases (PPhs) (also referred to as intermediate phrases), each of which must end with a high phrase accent (H-) or low phrase accent (L-) and must contain at least one pitch accent, which can be high, low, or bitonal (H*, L*, L+H*, etc.).

For present purposes, I will assume that neither intonational phrases nor phonological phrases are recursive, and that each intonational phrase must be exhaustively parsed into phonological phrases (but see Selkirk (1995b) and Ladd (1986, 1996) for further discussion of these issues). With these specifications, a complete English intonational contour is analyzed as the combination of the local pitch excursions of pitch accents, the roughly linear interpolation between pitch accents (when multiple pitch accents occur in a phonological phrase), the more broadly-realized pitch of phrase accents, which control the pitch from the syllable following the final pitch-accented syllable in a phonological phrase to the end of the phonological phrase, and the locally-realized pitch of boundary tones, which are marked on the final syllable of the intonational phrase.

Figure 1.1 shows four different well-formed intonation contours for the sentence *Alison likes coffee.* The tones of each sentence have been transcribed using the ToBI system, with phonological and intonational phrase boundaries additionally marked by subscripted parentheses. Note the substantial variability in intonation which is possible for this sentence, just from varying a few prosodic elements. The contrast between the final H% boundary tone of (a) and (c) with the L% boundary tone of (b) and (d) is particularly evident in these examples. Closer inspection also reveals the higher excursion used by this speaker for L+H* accents than for H* accents, the level pitch of an H- phrase accent after an H* pitch accent versus the fall in pitch of an L- phrase accent, and the "leftward spreading" of phrase accents, clearly visible in sentence (a), where the L- phrase accent, which is aligned with the final syllable of the phonological phrase at the phonological level of representation, is phonetically realized across all of the final unaccented material.



L+H* H*

a.((Alison likes coffee L-)$_{PPh}$ H%)$_{IPh}$ c. ((Alison L-)$_{PPh}$ (likes coffee L-)$_{PPh}$ H%)$_{IPh}$

H* H* H* L+H*

b.((Alison likes coffee L-)$_{PPh}$ L%)$_{IPh}$ d. ((Alison likes H-)$_{PPh}$ (coffee L-)$_{PPh}$ L%)$_{IPh}$

Figure 1.1. Intonation Contours. For each production, the waveform, pitch track, and text is given. Pitch accents (H*, L+H*), phrase accents (H-, L-), and boundary tones (H%, L%) are transcribed on the pitch track and with the text, and prosodic boundaries are indicated with subscripted parentheses in the text.

In addition to the evidence for prosodic boundaries provided by phrase accents and boundary tones, prosodic boundaries are also typically marked by pausing and/or final lengthening (lengthening of the final syllable of the prosodic phrase) (Klatt, 1975; Scott, 1982; Warren, 1985; Wightman, Shattuck-Hufnagel, Ostendorf, & Price, 1992) and segmental variation (Lehiste, 1960; Scott & Cutler, 1984; Gussenhoven & Rietveld, 1992; Pierrehumbert & Talkin, 1992; Dilley, Shattuck-Hufnagel & Ostendorf, 1996; Fougeron & Keating, 1997). Thus, there are frequently multiple phonetic cues to the presence of a prosodic boundary. However, since the experiments reported here are concerned not so much with the construction of a prosodic representation but with the effects of a prosodic representation on higher levels of analysis, I will often discuss the markings of prosodic boundaries simply in terms of the phrase accents and boundary tones that are associated

with those boundaries. This is not intended to imply that durational or segmental information is unimportant in the identification of prosodic phrase boundaries. Indeed, the reader should keep the various sources of evidence of prosodic structure in mind throughout this work when evaluating attempts to convey, control, or alter a particular prosodic structure.

1.2.2 Psycholinguistic Theory

I will also assume the Garden Path model of sentence comprehension (Frazier, 1978, 1987), a modular, serial, structure-based model of processing. That is, the processing system is divided into multiple separate processing subsystems, each using a specialized computational vocabulary (Frazier, 1990). However, because there is limited overlap in the vocabularies, there is limited interaction between the modules. Within the syntactic module, one structure is built at a time, using structure-based principles such as Minimal Attachment (*Do not postulate any potentially unnecessary nodes*) and Late Closure (*If grammatically permissible, attach new items into the clause or phrase currently being processed*), which stem from the need to assign structure as quickly as possible. Revisions to the first analysis are expected to be the minimal revisions possible, made in response to error signals stemming from other modules or from subsequent information, and are predicted to be more costly if the material which must be revised has already been semantically interpreted.

In particular, I assume that there is a phonological component to the processing system, which, in addition to identifying morphemes and lexical units, builds a prosodic representation for the incoming material. As this output is made available to higher-level components of the processing system, those components can potentially make use of any information which is part of their computational vocabulary. Crucially, I assume that because there are syntactic and semantic/pragmatic constraints which mention prosodic boundary information, this information is part of the computational vocabulary of the syntactic and semantic/pragmatic processing modules. In effect, the prosodic

representation that is constructed by the phonological component is passed on to higher-level modules in the same way that lexical information is made available to them. This system is illustrated in Figure 1.2. The left box shows some of the information we would expect to part of the vocabulary of the phonology module—things like [nasal], mora, foot, and a possible phonosyntactic constraint that aligns the right edges of syntactic phrases and phonological phrases. The right box shows some vocabulary relevant to constituent structure, including the phonosyntactic constraint of right edge alignment that mentions phonological phrase boundaries.

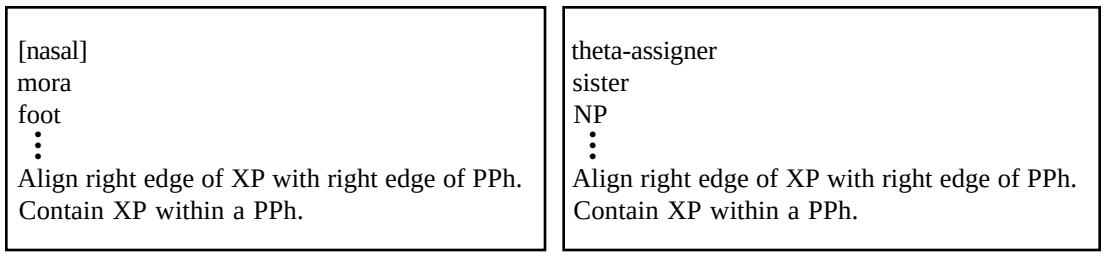


Figure 1.2 Overlap in Prosodic Information Across Processing Modules.

1.3 Previous Research on Prosodic Phrasing and Sentence Comprehension

Although very few studies have looked specifically at the effects of prosodic phrasing on sentence comprehension in English, a larger set of studies have investigated prosodic effects which likely included differences in prosodic phrasing. The studies can be organized into three categories. The first group is primarily concerned with discovering which classes of ambiguities can be disambiguated by prosody. These studies show that at least some kinds of structural ambiguities can be reliably disambiguated by prosodic phrasing differences, but that prosodic manipulations do not appear to be effective cues to the disambiguation of lexical ambiguities. The second group is concerned with just one kind of structural ambiguity, the NP- versus S-comp ambiguity. These studies provide suggestive although not conclusive evidence of the effectiveness of prosodic phrasing in the resolution of this ambiguity. The final group covers another kind of structural

ambiguity, early versus late closure of a subordinate clause, which can clearly be disambiguated by prosodic phrasing.

The next three sections will review the findings from these three groups of studies. In these sections, I will focus on describing the data that must be accounted for in a complete theory of sentence comprehension, by some combination of grammatical and processing constraints, and refrain from discussing specific hypotheses of how the data should be accounted for. Later, in Chapter 2, I will propose a variety of different general hypotheses which might account for the effects of prosodic phrasing on syntactic attachment ambiguities and present results from a new experiment which discriminates among them. As will become evident, simply determining what the relevant data are requires careful consideration of the previous research. The interpretation of most of the studies reported below is limited by one or more factors, such as the use of materials that were produced by speakers who were consciously attempting to disambiguate the sentences and thus may have used exaggerated or unnatural contours, the use of tasks in which the listeners were conscious of the ambiguity and thus may not have engaged in typical processing strategies, or the failure to adequately describe the prosody that was tested. Nevertheless, these studies do establish that prosodic phrasing can have some role to play in sentence comprehension, and thus lay the groundwork for a more detailed consideration of what that role might be.

1.3.1 General Studies on Prosody and Disambiguation

1.3.1.1 Lehiste (1973)

One of the classic studies of prosodic disambiguation is that of Lehiste (1973), who studied the production and identification of fifteen ambiguous sentences including the ones in examples (1.2) and (1.3). The sentences were first read aloud by four speakers of Midwestern American English who had not been told of the ambiguity. Next, the speakers were informed of the ambiguity and asked which meaning they had in mind during their production. The speakers then produced each sentence a second and third time, producing

one token for each interpretation. These three versions of each sentence were presented to listeners with a forced-choice paraphrase selection task, in which the listeners judged which meaning the speaker was intending to convey.

Lehiste found that the sentences which could be successfully categorized by listeners were sentences such as those in (1.2), with syntactic bracketing ambiguities, and not those like (1.3), with only syntactic category ambiguities. Apparently listeners were at least somewhat capable of correctly judging the intended meaning with both original productions and consciously disambiguated productions, although with slightly better success, on average, for the consciously disambiguated ones. Further, the sentences which were successfully categorized showed evidence of differences in prosodic phrasing between the pairs; Lehiste reports segmental and durational effects at the crucial syntactic boundaries. This finding was supported by work of Lehiste, Olive, & Streeter (1976), who tested a subset of Lehiste's materials with the F_0 resynthesized to a constant level and found that sentences with structural ambiguities, but not sentences with lexical ambiguities, could be successfully categorized by manipulating the durations of critical regions, thereby influencing the perception of a boundary.

(1.2) a. The hostess greeted the girl with a smile.
b. The old men and women stayed home.
c. Steve or Sam and Bob will come.
d. The police stopped fighting after dark.

(1.3) a. German teachers visit Greensboro.
b. Visiting relatives can be a nuisance.
c. The shooting of the hunters was terrible.

1.3.1.2 Warren (1985)

Warren (1985) conducted similar studies to those of Lehiste, on five sets of sentences for each of the structural ambiguities given in examples (1.4) - (1.7). In his production task, nine to ten speakers of Southern British English read aloud the experimental sentences from pseudorandomized lists of sentences containing a large number of fillers. Acoustic analyses were performed on the syntactically ambiguous regions marked by underscores in

the examples below. The results showed large differences in the strength of the prosodic boundary at the point of syntactic ambiguity for sentences like those in (1.4), marked by both durational and tonal differences. Similar but somewhat smaller effects were found for sentences like (1.5) and (1.6). Interestingly, syntactically disambiguated sentences like (1.5c) were produced with a slightly stronger prosodic boundary after the first verb (e.g., *learnt*) than their temporarily ambiguous counterparts. Effects for sentences like those in (1.7) were smaller yet, but still showed significant differences between the (a) and (b) versions in the duration of the word and silent interval preceding the PP.

(1.4) a. Before the king rides_his horse_takes ages to groom.
b. Before the king rides_his horse_Ted gives it a groom.

(1.5) a. The actor learnt_the text_and knew his role.
b. The actor learnt_the text_amused the cast.
c. The actor learnt_that the text_amused the cast.

(1.6) a. The women's journal said that the park_acts_mainly as a lunchtime retreat.
b. The women's journal said that the park_acts_meant a violation of rights.

(1.7) a. John broke the clock_with a gold face.
b. John broke the clock_with his bare hands.

Warren also conducted a perception experiment on these materials, with new utterances produced by a single speaker who reportedly showed similar prosodic patterns to the ones just described. In this experiment, the sentences were truncated at the end of the ambiguous region and the resulting fragments were presented to subjects, who chose between written versions of the two original continuations of the sentence. The results showed that subjects were able to correctly match fragments with their original continuations almost 99% of the time for sentences as in (1.4), approximately 85% of the time for the ambiguity exemplified in (1.5), just under 82% of the time for the ambiguity of (1.6), and about 57% of the time for sentences like (1.7). Thus, the large differences in prosodic boundary strength for the first three kinds of ambiguity were highly effective in disambiguating the structures, and the smaller differences found for sentences as in (1.7) were somewhat effective at disambiguating the structure. A second perception experiment

showed that the fragments could still be distinguished, although not as well, when the sentences were resynthesized with the F_0 at a constant level.

1.3.1.3 Price *et al.* (1991)

Lehiste and Warren's findings were further confirmed in a study by Price, Ostendorf, Shattuck-Hufnagel, & Fong (1991). They had four professional FM radio announcers from the Boston area read five sentences each from seven classes of ambiguous sentences presented in disambiguating contexts. The speakers were instructed to use their standard radio style of speaking. Sample sentences from each class appear in Table (1.1).

Table 1.1. Sentences from Price *et al.* (1991) and Percentages of Correct Identification .

Ambiguity	Percentage of Correct Identification	
1. Parenthetical versus Nonparenthetical	86%	Overall
a. Mary knows many languages, you know.	77%	(a)
b. Mary knows many languages you know.	96%	(b)
2. Appositions versus Attached NPs or PPs	92%	Overall
a. The neighbors who usually read, the Daleys, were amused.	91%	(a)
b. The neighbors who usually read the dailies were amused.	92%	(b)
3. Conjoined Main Clauses versus Main + Subordinate Clause	71%	Overall
a. Mary was amazed and Dewey was angry	54%	(a)
b. Mary was amazed Ann Dewey was angry	88%	(b)
4. Tag Questions versus Attached NPs	88%	Overall
a. Dave will never know why he's enraged, will he.	95%	(a)
b. Dave will never know why he's enraged Willy.	81%	(b)
5. Far versus Near Attachment of Final Phrase	71%	Overall
a. I read a [review of nasality] in German.	78%	(a)
b. I read a review of [nasality in German].	63%	(b)
6. Left versus Right Attachment of Middle Phrase	95%	Overall
a. They rose early, in May.	94%	(a)
b. They rose early in May.	95%	(b)
7. Particles versus Prepositions	82%	Overall
a. Why are you [grinding in] the mud?	82%	(a)
b. Why are you grinding [in the mud]?	81%	(b)

Listeners were presented with the ambiguous sentence in isolation from the disambiguating context. They were then asked to select the appropriate context from written versions of the two alternatives. Sentences varied in the degree to which they were successfully matched to the context, as can be seen in Table (1.1). Generally, sentences produced with stronger prosodic boundaries resulted in better discrimination by the listeners. In particular, Price *et al.* found that almost all reliably identified sentences contained intonational boundaries at the major syntactic boundary. They also found that clause boundaries very often coincided with an IPh boundary for these speakers. Overall, Price *et al.* found systematic differences in either the location of the largest prosodic boundary or the relative size of the prosodic boundary at points of syntactic ambiguity. As they did not generally find differences in accent usage or other aspects of the sentences between pairs, these results provide strong evidence of the ability to use prosodic phrasing as a primary means of disambiguation for several kinds of structural ambiguities.

Taken together, these three studies firmly establish the ability of speakers to produce distinct patterns of prosodic phrasing for many kinds of structural ambiguities in English. They also establish the ability of naive listeners to correctly match these prosodically distinct utterances with the intended meaning, at least in cases when the listeners are aware of the ambiguity and there are only two possibilities to choose from. However, they do not necessarily show that prosodic phrasing is used in normal sentence comprehension, in which listeners generally are unaware of the ambiguities in a sentence and do not engage in conscious consideration of alternative meanings. It is consistent with the results of these studies for listeners to use prosodic phrasing information only when they have become conscious of an ambiguity, presumably well after most processing has taken place, and only when there is absolutely no other information to guide the choice between two interpretations.

1.3.2 The NP- versus S-Complement Ambiguity

Research on the use of prosody to resolve NP- versus S-complement ambiguities, although not conclusive, provides some evidence that prosodic phrasing is in fact used during normal sentence processing. The NP- versus S-complement ambiguity has received considerable attention in psycholinguistic research in general as well in psycholinguistic research specifically on prosody. It was one of the structures tested in Warren's work, described above; his example is repeated in (1.8). Like Warren (1985), both Beach (1989, 1991) and Stirling & Wales (1996) found that prosodic cues allowed listeners to successfully discriminate between the (a) and (b) versions of sentences such as (1.9) in forced-choice continuation selection tasks. More importantly, several experiments have tested prosodic effects on this structure using experimental tasks that are generally sensitive to processing loads at early stages of sentence comprehension.

(1.8) a. The actor learnt the text and knew his role.
b. The actor learnt the text amused the cast.

(1.9) a. Jay believed the gossip about the neighbors right away.
b. Jay believed the gossip about the neighbors wasn't true.

1.3.2.1 Marslen-Wilson *et al.* (1992)

Marslen-Wilson, Tyler, Warren, Grenier & Lee (1992) and Watt & Murray (1996) tested sentence fragments like (1.10) in cross-modal naming tasks, in which a visually-presented word that continues the sentence begun by the auditory fragment must be named by the subject. In this task, it is assumed that naming times for the probe word reflect the ease with which the subject can integrate the word into the phrase structure that he or she is in the process of building, and thus should be short when the word is compatible with the current partial phrase structure and long when the word is incompatible with the structure. Marslen-Wilson *et al.* found that naming times for words consistent with an S-complement continuation (e.g., *was*) were as short following a syntactically ambiguous fragment with S-complement prosody as with a fragment disambiguated by the presence of a

complementizer, and were shorter with S-complement prosody than with NP-complement prosody. This result suggests that S-complement prosody was as effective as the presence of a complementizer in disambiguating the auditory fragment to an S-complement parse.

(1.10)	Auditory Fragment:	Visual Word:
	The jury believed the testimony of the last witness	WAS

1.3.2.2 Watt & Murray (1996)

In contrast, Watt & Murray (1996) found significant effects of prosody in judgments of the appropriateness of the visual word as a continuation of the auditory fragment, but no significant effects of prosody on reaction times in two cross-modal naming tasks and three cross-modal lexical decision tasks, all of which used the same auditory materials, replicating a null effect of prosody that they report from Watt (1992) in a mispronunciation detection task on the same auditory materials. However, Watt and Murray failed to find any significant reaction time differences in the experimental items for three out of the five experiments, and the results of the remaining two need not be interpreted as a null effect of prosody but significant effect of syntactic structure.

As Watt and Murray's series of experiments showed that changes in details of the task can produce somewhat different trends in the results, it may be the case that Watt and Murray's tasks were not sensitive enough to reflect processing effects of prosody but Marslen-Wilson *et al.*'s task was. It may also be the case, as Watt and Murray suggest, that subtle effects of prosody do not affect early parsing decisions (for this structure; see below for more on the effects of subtle prosodic boundaries) and that their materials differed significantly from Marslen-Wilson *et al.*'s in the robustness of the prosodic difference. Unfortunately, since neither Marslen-Wilson *et al.* or Watt and Murray describe the prosody used in the experiments, it is not possible to determine from these studies which prosodic differences affect NP- versus S-complement disambiguation and which do not. These studies are nevertheless important, individually, in their use of tasks designed

to test early stages of processing, and important when taken together in raising critical questions of which prosodic manipulations can affect processing decisions for a given structure, when during processing they do so, and whether those manipulations are consistently used by speakers of the language.

1.3.2.3 Nagel *et al.* (1996)

Nagel, Shapiro, Tuller & Nawy (1996) conducted three experiments on the NP- versus S-complement structure. In the first experiment, they tested sentences as in (1.11) with both original tokens (a, c) and cross-spliced tokens (b, d; splicing is indicated by the change in font and the vertical line). They employed a cross-modal lexical decision task, in which subjects made a word/non-word judgment to a visually-presented probe that coincided with the point indicated by the pound sign in (1.11). This task is based on the assumption that because there is a finite amount of processing resources, reaction times to the probe will be longer when processing load is higher, such as when the processor must reanalyze an earlier decision, than when it is lower. Nagel *et al.* found that decision times were longer in the cross-spliced cases, in which the prosodic information presumably conflicted with the syntactic information, than in the original, unspliced cases. However, since any potential effects of prosody are confounded with the potential unnaturalness of the cross-spliced sentences, and they did not perform a naturalness test on the materials, it is possible that the differences in decision times were entirely due to the effect of cross-splicing and not to differences in prosody. And, like the other researchers, Nagel *et al.* do not describe the prosody used in this experiment or provide acoustical analyses.

(1.11) a. The company owner promised the wage increase to # the workers.
b. The company owner promised the wage increase | to # the workers.
c. The company owner promised the wage increase would # be substantial.
d. The company owner promised the wage increase | would # be substantial.

In a second experiment, the sentences used in Experiment 1 were re-recorded by a speaker naive to the purposes of the experiment. Acoustical analyses of the matrix verb

showed that the duration was significantly longer in the S-complement case than in the NP-complement case, and the fall in F_0 from the peak in the verb to the final minima in the verb, as well as any subsequent rise, were significantly larger in the S-complement production. These results match the findings of Warren (1985) reported above of a stronger boundary in the S-complement sentence, and suggest that this speaker generally produced an IPh boundary before the embedded clause (although differences in accentuation could have contributed to both the greater duration and the greater fall in F_0).

Nagel *et al.*'s third experiment tested the comprehension of altered versions of the materials used in Experiment 2. First, the sentences were resynthesized with a level F_0. Then the verb and silent interval preceding the syntactically ambiguous phrase was either lengthened or shortened to match the durations found in the original production of the other structure. This created four conditions: NP-complement with original durations, NP-complement with lengthened verb and pause, S-complement with original durations, and S-complement with shortened verb and pause. These conditions were tested with a cross-modal lexical decision task, as in Experiment 1, but with two probe positions, marked by superscripts in (1.12). The first probe occurred after the altered verb but before the end of the syntactically ambiguous region, and thus should show effects of alteration on performance but not effects of prosody on syntactic resolution. The second probe occurred after the disambiguating word.[1] If prosodic phrasing can guide the resolution of the NP- versus S-complement ambiguity, reaction times to this probe should be shorter for the conditions with original durations than for the conditions with altered durations.

(1.12) a. The company owner promised the wage increase ^{p1} to ^{p2} the workers.
b. The company owner promised the wage increase ^{p1} would ^{p2} be substantial.

[1] Strictly speaking, only the S-complement continuation contained a disambiguating word, since it is still possible to continue the other fragment as an S-complement (e.g., *The company owners promised the wage increase to $10 would please everyone*).

Nagel *et al.* found no significant differences among the reaction times for the first probe, showing that the alteration of the verb and pause did not significantly affect processing load prior to the point of syntactic disambiguation. For the second probe, there was a significant interaction between syntactic structure and alteration. Both structures showed longer times for the altered version of the sentence than for the original version. In conjunction with the null effects seen at the first probe, these results suggest that prosodic phrasing (or at least, one of the acoustic correlates of prosodic phrasing) can significantly affect interpretation before the end of the sentence, presumably at early stages of sentence processing.

Of course, it is possible that the flattening of the F_0 made the materials sufficiently unnatural to prevent the subjects from engaging in natural processing strategies. Further, because the experiment only tested what were presumably natural (original) durations versus unnatural (altered) durations, and did not include comparisons with either syntactically unambiguous sentences or fully ambiguous sentences, it is difficult to determine whether any given condition facilitated processing or interfered with it. Thus, it may not be the case that appropriate prosody can disambiguate this syntactic structure, but only that inappropriate prosody increases processing load.

In summary, the NP- versus S-complement studies show reliable effects of prosodic phrasing differences in production between the two structures, as well as significant discriminability between the two structures in forced-choice continuation selection tasks. They also provide tentative evidence that prosody might affect early stages of processing for these sentences. However, the Marslen-Wilson *et al.* results do not securely establish that the apparent effects of prosody in early stages of processing are due to differences in prosodic phrasing, and they failed to replicate in the Watt and Murray experiments. And, in the Nagel *et al.* study, where an acoustic correlate of prosodic phrasing differed between the two structures, the study does not establish what effect that difference has on processing, or how generalizable the effect would be to sentences with natural F_0 contours.

1.3.3 On-Line Tests of Early versus Late Closure Ambiguities

A particularly informative series of studies on the effect of prosodic phrasing on syntactic attachment comes from the work of Slowiaczek (1981) and Speer *et al.* (1996). The results of these studies provide robust empirical support for the use of prosodic phrasing during early stages of parsing. Moreover, the Speer *et al.* work clearly demonstrates that prosodic phrasing can both facilitate processing, when the prosody is appropriate for the syntactic structure, and interfere with processing, when the prosody is misleading.

The studies focused on early versus late closure ambiguities, as in the sentence pair in example (1.13). Until the disambiguating phrase *is* or *it's* is encountered, both (a) and (b) are consistent with either an interpretation in which *sings* is an intransitive verb and *the song* is the subject of the matrix clause, as in the final interpretation of (a), or with an interpretation with a transitive verb followed by the direct object *the song*, as in the final interpretation of (b).

(1.13) a. When Madonna sings the song it's a hit. (late closure)
b. When Madonna sings the song is a hit. (early closure)

When such sentences are encountered in reading tasks with no comma present to disambiguate the attachment of the NP, there is a general processing advantage for the (a) version.[2] Subjects show an initial preference to attach the temporarily ambiguous NP into the VP of the subordinate clause and must reanalyze the sentence and correct that attachment in the (b) version. Because this effect is accounted for by the principle of Late Closure in the Garden Path model, version (a) is commonly referred to as the "late closure" version of the sentence and (b) as the "early closure" version, where "late closure" refers to

[2] Researchers disagree on whether this effect holds for all verbs or only a subset of them, which types of information can influence the initial attachment decision, and what the proper account of the effect is (see, e.g., Kimball, 1973; Frazier, 1987; Tanenhaus & Carlson, 1989; Marcus & Hindle, 1990; Gibson, 1991; MacDonald, Pearlmutter & Seidenberg, 1994). However, the general preference for version (a) should not be controversial for the materials used in the experiments described here.

the apparent preference of the parser to incorporate material into the syntactic phrase which is currently being processed when allowed by the rules of the grammar, thus "closing" the syntactic phrase late in the string, as opposed to at the earliest point at which it is (potentially) grammatical to do so. Both Slowiaczek and Speer *et al*. tested early versus late closure sentences under various prosodic conditions.

1.3.3.1 Slowiaczek (1981)

Slowiaczek (1981; Caroll & Slowiaczek, 1987) collected sentence comprehension times[3] for four different types of early versus late closure syntactic structures. A sample item for each of the four structures is given in (1.14). Each sentence was tested under four prosodic conditions: the sentence contained only a late prosodic boundary, only an early prosodic boundary, both boundaries, or no boundaries, as shown in (1.15). In theory, this design should allow the effects of an early prosodic boundary and a late prosodic boundary to be evaluated independently. However, most of the conditions were created by cross-splicing materials at locations within a prosodic phrase, and thus may have contained unnatural pitch contours and lengthening patterns.[4] In the sample set of conditions in (1.15), Monaco font is used for strings originally produced with late closure prosody and Helvetica font is used for strings originally produced with early closure prosody. Natural IPh boundaries are indicated by parentheses and an IPh subscript, and capitals signal original main sentence stress. The tonal qualities of the boundaries are not clear from Slowiaczek's description, but the durations she reports for the critical regions suggest that the materials had intonational phrase-level boundaries; Slowiaczek describes the boundaries as natural comma pauses.

---

[3] The subject listened to sentences played over a tape recorder in a quiet room and pressed a button at the end of each sentence, as soon as he or she had understood the sentence. Response times were measured from the beginning of the last syllable of each sentence. If the subject did not understand the sentence, he or she informed the experimenter. Paraphrases were required for 25% of the trials, selected randomly for each subject.

[4] Conditions (a) and (d) were created by recording the complete sentence, splicing out each clause, and resplicing the clauses back together. Other conditions were created by cross-splicing material from conditions (a) and (d) as indicated in (1.15).

(1.14) Sample Items from Slowiaczek (1981):

Group 1:
a. Because her grandmother knitted pullovers, Cathy kept warm in the wintertime.
b. Because her grandmother knitted, pullovers kept Cathy warm in the wintertime.

Group 2:
a. Without a family tradition, parties tend to be less important at Christmas.
b. Without a family, tradition tends to be less important at Christmas.

Group 3:
a. Without the author's permission to sell the book, publicity could be very difficult.
b. Without the author's permission, to sell the book could be very difficult.

Group 4:
a. John was kicking Tim, Paul tried punching, and the teacher had lost her patience.
b. John was kicking, Tim tried punching, and the teacher had lost her patience.

(1.15) Sample Set from Slowiaczek (1981), Group 1:

a. Late closure syntax; late boundary:
(Because her grandmother knitted pullovers)$_{IPh}$ (CATHY kept warm in the wintertime)$_{IPh}$

b. Early closure syntax; late boundary:
(Because her grandmother knitted pullovers)$_{IPh}$ kept Cathy warm in the wintertime)$_{IPh}$

c. Late closure syntax; early boundary:
(Because her grandmother knitted)$_{IPh}$ (PULLOVERS (CATHY kept warm in the wintertime)$_{IPh}$

d. Early closure syntax; early boundary:
(Because her grandmother knitted)$_{IPh}$ (PULLOVERS kept Cathy warm in the wintertime)$_{IPh}$

e. Late closure syntax; both boundaries:
(Because her grandmother knitted)$_{IPh}$ pullovers)$_{IPh}$ (CATHY kept warm in the wintertime)$_{IPh}$

f. Early closure syntax; both boundaries:
(Because her grandmother knitted)$_{IPh}$ pullovers)$_{IPh}$ kept Cathy warm in the wintertime)$_{IPh}$

g. Late closure syntax; no boundaries:
(Because her grandmother knitted (PULLOVERS (CATHY kept warm in the wintertime)$_{IPh}$

h. Early closure syntax; no boundaries:

(Because her grandmother knitted (PULLOVERS kept Cathy warm in the wintertime)$_{IPh}$

Slowiaczek's results showed a clear interaction of syntax and prosody. For the conditions with a single boundary, late closure sentences were significantly faster to

comprehend with a late boundary than with an early boundary and the early closure sentences were significantly slower to comprehend with a late boundary than with an early boundary. This is not surprising; for each syntactic structure, subjects show faster comprehension times for the original (respliced) sentence, which contains a naturally-produced prosodic boundary at the clause boundary, than for the cross-spliced sentence, in which the naturally-produced juncture falls within the clause.

The numerical results from the conditions with both boundaries and no boundaries show the preference for late closure syntax over early closure syntax that is generally found in reading tasks. This pattern was consistent across the four groups of sentences for the Both Boundary condition, but not for the No Boundary condition. However, because splicing was done at points internal to a prosodic phrase, the prosody may not have been properly controlled throughout the stimuli. Thus, as Slowiaczek notes, subjects may not have perceived all of junctures in the stimuli as the presence or absence of a prosodic boundary as intended by the experimenter. In particular, the early boundary in the both boundary conditions may have been less salient than the late boundary, so subjects may have analyzed these items as if they contained only a late boundary. Although it is clear from the conditions with single boundaries that prosody had some effect on comprehension in this study, the potential confounds introduced by the method of stimulus construction make it difficult to judge whether the demonstrated effects were due to the naturalness of the stimuli, the well-formedness of the sentences, or to a true effect of prosodic boundaries on processing decisions.

Despite this limitation, Slowiaczek extended the research on how prosody might disambiguate attachment decisions in two important ways. First, instead of contrasting entire prosodic contours for sets of experimental items, her research focused on a single prosodic element—the comma pause—and investigated the specific effect of that element at particular locations. This method should allow us to evaluate phonological models that claim that the prosodic contour is composed of a discrete set of elements (such as L%

boundary tones) and aid us in developing explicit, detailed hypotheses of how each component of the prosodic contour might affect parsing decisions and interact with other sources of information used during sentence processing.

Second, although prior work such as Lehiste's had shown that listeners could make use of prosody when they were aware of the ambiguity in the sentence, Slowiaczek was the first researcher to employ a task in which subjects were not consciously choosing between two paraphrases of a sentence. Her work (which preceded the work described in the previous section) provided the first evidence that prosody might be used in early, unconscious stages of parsing and at anytime that it is available, as opposed to being used only in conscious deliberation about sentences long after the initial parsing decisions have been made. In the time since Slowiaczek's work it has become increasingly apparent that prosody can exert a strong influence on the interpretation of sentences. Nevertheless, basic questions first addressed in Slowiaczek's work are still being investigated: which prosodic elements affect parsing, where in the string they can do so, and how distinct prosodic events and the relationships among them affect the parser.

1.3.3.2 Speer *et al.* (1996)

Slowiaczek's design was partially replicated in several studies by Speer and her colleagues using the same early versus late closure structure as in Slowiaczek's Group 1 sentences, but with more carefully controlled materials and more sensitive tasks. The Speer *et al.* studies compared early versus late closure sentences under three prosodic conditions—cooperating prosody, in which a prosodic boundary was located at the end of the subordinate clause (i.e., late closure syntax with a late boundary; early closure syntax with an early boundary), conflicting prosody, in which the prosodic boundary was located where the end of the subordinate clause would be for the other structure (i.e., late closure syntax with an early boundary; early closure syntax with a late boundary), and ambiguous prosody, in which the evidence for sentence-internal prosodic boundaries was phonetically neutralized. A sample set for Experiment 1 is given in (1.16).

For their first experiment, the cooperating prosody conditions were produced by a trained speaker with an L- phrase accent and L% boundary tone at the medial clause boundary. Monaco font indicates material originally produced with late closure syntax and prosody, and Helvetica indicates material originally produced with early closure syntax and prosody. These materials were then digitally cross-spliced to create materials with conflicting prosody, as indicated in (1.16). The baseline prosody conditions were produced with accents on the subject of the subordinate clause and the end of the main clause and a fast rate of speech and low F_0 throughout the syntactically ambiguous region, which neutralized the acoustic evidence for any intervening phrase accents or boundary tones. Phonetic measurements confirmed that all of the materials were produced as intended.

(1.16) Sample Set for Speer *et al.* Experiment 1:

Cooperating prosody:
a. (Whenever the guard checks the door L-L%)$_{IPh}$ (it's locked)$_{IPh}$
b. (Whenever the guard checks L-L%)$_{IPh}$ (the door is locked)$_{IPh}$

Baseline prosody:
c. (Whenever the guard checks the door it's locked)$_{IPh}$
d. (Whenever the guard checks the door is locked)$_{IPh}$

Conflicting prosody:
e. (Whenever the guard checks L-L%)$_{IPh}$ (the door (it's locked)$_{IPh}$
f. (Whenever the guard checks the door L-L%)$_{IPh}$ is locked)$_{IPh}$

1.3.3.2.1 Pretest for Experiment 1

The materials were pretested for acceptability to ensure that the prosodic contours used in the cooperating and baseline conditions were equally acceptable with late closure syntax and early closure syntax. Subjects first read and understood a visually-presented version of each sentence, and then listened to the experimental stimulus and evaluated the acceptability of the pronunciation. The results showed an average rate of acceptance of 90% for each condition of the cooperating prosody and the baseline prosody, with no

statistically significant differences among them, and a less than 40% rate of acceptance for the two conflicting prosody conditions.

This pretest provided several important pieces of information. Most importantly, it rules out the possibility that any results in the main experiment are due to differences in naturalness or well-formedness of the materials in the cooperating and baseline conditions. Thus, any significant results from the main experiment can be safely attributed to sentence processing effects. Further, it makes several comparisons possible. First, the late closure syntax can be compared to the early closure syntax within a type of prosody. Thus, if the prosodic boundary in these materials can be used to disambiguate the sentences, the early closure syntax with cooperating prosody should be as easy to process as the late closure syntax with cooperating prosody, in contrast to the case with baseline prosody, in which the absence of prosodic cues should result in the standard processing advantage for the late closure syntax.

Second, conditions which contain the same syntactic structure can be compared across the prosodic conditions, and crucially, this can be done for two prosodies judged to be equally natural off-line. Experiments which only compare cooperating (or appropriate) and conflicting (or inappropriate) prosody for a given syntactic structure do not necessarily show a facilitative effect of prosody, since they can be also be interpreted as showing merely that inappropriate prosody slows processing, for any of a number of reasons. However, the Speer *et al.* design and the results of the pretest allows the potentially facilitative effect of unambiguous prosody to be evaluated by comparing the cooperating and baseline conditions.

Finally, the results of the pretest are a useful piece of evidence in determining the grammatical constraints on the prosody-syntax connection and in separating out grammatical effects from parsing effects. It has long been thought that there may be a grammatical requirement, or at least preference, for a prosodic boundary to fall at the edges of root sentences, and thus at the clause boundary in these materials (Chomsky & Halle,

1968; Downing, 1970, 1973; Bing, 1979; see also Halliday, 1967; Chafe, 1980; Croft, 1995). If such a constraint exists, Speer *et al*.'s pretest results suggest that either it is a violable constraint dominated by constraints which militate against the presence of prosodic boundaries under the conditions found in the baseline sentences, or that the prosodic boundary need not have transparent phonetic instantiation—it may be sufficient to have phonetic input which is merely compatible with the presence of a prosodic boundary, with the phonological representation of the boundary inserted through constraints of the phonology or phonology-syntax interface.

1.3.3.2.2 Experiment 1

Speer *et al.*'s first experiment was a simple end-of-sentence comprehension task, in which subjects listened to the entire sentence and pressed a lever as soon as the sentence was over. The average comprehension times for each of the conditions, measured from the end of the sentence in milliseconds, are presented in Table (1.2). Statistical analyses showed that the times did not differ significantly from each other for the late closure cooperating condition, the early closure cooperating condition, and the late closure baseline condition. The early closure baseline condition was significantly slower than the early closure cooperating condition and the late closure baseline condition. Each of the two conflicting prosody conditions were significantly slower than their baseline counterparts, and the early closure conflicting condition was slower than the late closure conflicting condition.

Table 1.2. End of Sentence Comprehension Times for Speer *et al.* Experiment 1.

	Prosody:		
Syntax:	Cooperating	Baseline	Conflicting
Late Closure	645 ms	703 ms	908 ms
Early Closure	639 ms	780 ms	996 ms

The results show that cooperating prosody can facilitate parsing for the dispreferred early closure syntactic structure but that it does not significantly facilitate the parsing of the preferred late closure structure. They also show that conflicting prosody is disruptive to processing, and provide additional evidence that the processing of ambiguous tokens which resolve to late closure syntax is easier than the processing of ambiguous tokens which resolve to early closure syntax. Thus, the results confirm that general preference for late closure syntax over early closure syntax is found in speech comprehension as well as in reading, and that this difference cannot be attributed to a difference in well-formedness or acceptability (since the pretest verified the matching well-formedness and acceptability of the two baseline conditions). They further establish that disambiguating prosody can override the preference for a late closure structure, as an early closure structure with cooperating prosody is easier to comprehend than an equally well-formed and acceptable token of the sentence with ambiguous prosody, and as easy to comprehend as a late closure structure.

1.3.3.2.3 Experiment 2

In Experiment 1, the measurement of comprehension difficulty did not take place until the end of the sentence, well after the point of disambiguation. As such, it is consistent with prosodic information being used only very late in processing, well after syntactic disambiguation has occurred. Thus, a second experiment (Kjelgaard, 1995; Speer *et al.*, 1996) was conducted using truncated versions of these stimuli in a cross-modal naming task, in which measurements could be collected at the point of disambiguation. In this task, subjects heard sentence fragments which extended through the ambiguous NP. A sample set is given in (1.17). The presentation of the auditory fragment was immediately followed by the presentation of a one-word visual stimulus which was a possible continuation of the sentence. Subjects said the visually-presented word aloud as quickly as possible and then completed the sentence in their own words. Note that, in addition to allowing measurement within the sentence, this task also eliminated the need for splicing in

the conflicting prosody conditions. All six fragments now contained only naturally-produced, unaltered selections of speech.

(1.17) Sample Set for Speer *et al.* Experiment 2:

Auditory Fragment:	Visual Word:
Cooperating prosody:	
a. (Whenever the guard checks the door L-L%)$_{IPh}$	IT'S
b. (Whenever the guard checks L-L%)$_{IPh}$ (the door	IS
Baseline prosody:	
c. (Whenever the guard checks the door	IT'S
d. (Whenever the guard checks the door	IS
Conflicting prosody:	
e. (Whenever the guard checks L-L%)$_{IPh}$ (the door	IT'S
f. (Whenever the guard checks the door L-L%)$_{IPh}$	IS

If prosody is used only very late in the course of processing—perhaps merely to confirm the syntactic structure established at an earlier level of processing—then it might cause effects in an end-of-sentence task that would not be seen in a sentence-internal task like cross-modal naming. However, if prosody is used at an early stage of processing—if, for example, the prosodic structure that is the output of the phonology is the input for syntactic parsing decisions—then the cross-modal naming task should show similar results to the end-of-sentence simple comprehension task.

The results support the early use of prosody: the results from the cross-modal naming task had the same pattern as the results from the end-of-sentence task. Naming times did not differ significantly from each other for the late closure cooperating condition, the early closure cooperating condition, and the late closure baseline condition. The early closure baseline condition was significantly slower than the early closure cooperating condition and the late closure baseline condition. Each of the two conflicting prosody conditions were significantly slower than their baseline counterparts, and the early closure conflicting condition was slower than the late closure conflicting condition. Thus, as before, cooperating prosody facilitated the early closure syntax, conflicting prosody was disruptive

for both syntactic structures, and when the prosody did not disambiguate the syntax, the late closure syntax was easier to process than the early closure syntax. These results establish that prosodic information must be used quite soon after it is encountered, since the prosodic boundary at the end of the subordinate clause affected processing at the beginning of the next clause, at the point of syntactic disambiguation. They further establish that the slower comprehension times seen in the conflicting prosody conditions are not simply artifacts of cross-splicing, but rather the disruptive effect of the conflicting prosody on attachment decisions or grammaticality (or both).

1.3.3.2.4 Experiment 3

Experiments 1 and 2 used materials which contained an intonational phrase boundary in the cooperating and conflicting prosody conditions. These boundaries, analyzed as L-L% in the ToBI transcription system, were marked by significant phrase-final lengthening, pause durations, and lowering of the fundamental. The production of the materials followed natural speech patterns of the speaker, and the materials were judged to be acceptable in the pretest, so it is unlikely that the boundaries were interpreted by the subjects as unnatural or extra-linguistic cues (cf. Watt & Murray, 1996). However, if it is true that only very strong prosodic cues such as IPh boundaries are used during the early stages of parsing (Marcus & Hindle, 1990), then more subtle cues like phonological phrase boundaries should not affect parsing decisions. Thus, a third experiment (Kjelgaard, 1995; Speer *et al.,* 1996) was conducted with pronunciations which employed only phonological phrase boundaries at the critical boundary locations.

Experiment 3 used the same task and design as the previous experiment. However, the conflicting and cooperating conditions now contained only phonological phrase boundaries marked by high phrase accents instead of intonational phrase boundaries marked by low phrase accents and low boundary tones. As before, the materials were pretested for acceptability, and the results showed non-significant differences in acceptability among the non-conflicting conditions. If the subtle phonetic cues of phonological phrase boundaries

are not sufficient to affect parsing, the results for all three prosodies should be similar to the results for the baseline condition in the previous experiment, differing from each other only because of the effect of early versus late closure syntax.

In fact, the results showed the same pattern as in the previous two experiments, with significant effects of prosody in the early closure cooperating condition compared to the early closure baseline condition and in each of the conflicting prosody conditions compared to their baseline counterparts. Thus, for at least some syntactic structures, even acoustically subtle prosodic boundaries can facilitate parsing when they are appropriately located and can interfere with parsing when located inappropriately.

1.3.3.2.5 Dobroth (1996)

Finally, Dobroth (1996) tested another version of these materials in the same cross-modal naming task used in Experiments 2 and 3. Dobroth again used cooperating, conflicting, and baseline prosody. However, in this experiment the only information about the prosodic phrasing came from segmental effects associated with the presence or absence of a prosodic boundary. Dobroth constructed materials as in (1.18), in which the segments at the early closure location contained an unreleased [t] and glottalization on the following vowel, suggesting the presence of a prosodic boundary; a flap, suggesting the absence of a prosodic boundary; or an unreleased [t] followed by a consonant, which provided no information about the presence or absence of a boundary.

(1.18) a. If Pam wants to qui[t W]Andy's team
b. If Pam wants to qui[D] Andy's team
c. If Pam wants to qui[t] Kevin's team

In a set of forced-choice continuation selection tasks, Dobroth replicated a previous finding of Scott & Cutler (1984) that listeners could use segmental information to discriminate between an early closure structure and a late closure structure. However, the cross-modal naming task showed only one significant difference: naming times to words forcing an early closure resolution were significantly longer following fragments with flaps

(conflicting prosody) than following either of the other two fragments (cooperating prosody and baseline prosody). The naming times did not differ significantly from each other when the visual word forced a late closure resolution, and naming times were not significantly different in the early closure cooperating and early closure baseline conditions. Thus, segmental information from the presence of glottalization was not sufficient to facilitate processing for the early closure structure or interfere with processing in the late closure structure, although flapping was effective in interfering with processing in the early closure structure.

1.3.3.2.6 Summary of Speer *et al.* Experiments

Taken jointly, the Speer *et al.* early versus late closure experiments provide a secure base of information about the role of prosody in sentence processing. They show that prosodic phrasing can have a strong effect in sentence processing, at least when it is clearly marked by edge tones or final lengthening. This effect appears to be strong enough to prevent misanalysis in some cases and to interfere with processing in other cases, and it can not be accounted for solely by differences in well-formedness or acceptability. Further, the results suggest that prosodic boundary information is used at an early level of processing, as the effects of cooperating prosody can completely eliminate the evidence of processing difficulty for the early closure structure and are evident at the point of syntactic disambiguation. Moreover, a phonological phrase boundary is sufficient for causing these effects. This is a noteworthy finding, for two reasons. First, phonological phrase boundaries are much more common in the speech stream than intonational phrase boundaries, so if PPh boundaries generally affect syntactic attachment decisions, then effects of prosodic phrasing on syntactic attachment decisions should appear in a wider set of circumstances than if only IPh boundaries affect parsing decisions. I discuss one such case in Chapter 2. Second, the phonetic instantiation of PPh boundaries is much more subtle than the instantiation of IPh boundaries. If PPh boundaries are regularly detected and used by the parser, then our models of sentence processing must make explicit how the

parser gains access to the more detailed phonetic or phonological information that would be needed to identify them.

1.4 Conclusion

The experiments reported in this chapter firmly establish the ability of prosodic phrasing to disambiguate structural ambiguities. Several studies demonstrated the effectiveness of prosodic phrasing as a means of disambiguation for forced-choice continuation or paraphrase selection tasks. There is tentative evidence that prosodic phrasing may aide the resolution of the NP- versus S-complement ambiguity in early stages of processing, and strong evidence that both phonological phrasing and intonational phrasing can be used at early stages of processing, presumably as a guide to initial syntactic attachment decisions.

With this foundation in place, it is now possible to consider carefully the set of hypotheses which could account for the effects of prosodic phrasing on processing decisions. Several views of the role of prosody in parsing can be discarded already. The idea that prosody is *not* used in parsing, or that prosody is used only to facilitate the processing of dispreferred analyses (Pritchett, 1988; Wales & Toner, 1979), is disconfirmed. The Speer *et al.* results demonstrate that prosody can both facilitate processing (as with the early closure cooperating prosody condition) and interfere with processing (as with the two conflicting prosody conditions). The results also disconfirm the idea that prosody is used only very late in processing, after a phrase structure has been constructed on the basis of the non-prosodic information (Pynte & Prieur, 1996). Speer *et al.*'s Experiments 2 and 3 demonstrated that prosody was used as early as the point of syntactic disambiguation. The possibility that only strongly marked prosodic boundaries can have early processing effects (Marcus & Hindle, 1990; Watt & Murray, 1996) is ruled out as well. Speer *et al.*'s Experiment 3 found effects of phonological phrase boundaries and Dobroth found some effect of prosodic boundaries signaled only by segmental information. Whatever the proper account of prosody in sentence comprehension is, it

must be one which recognizes that subtly marked prosodic boundaries can affect processing, and they can do so during early stages of processing.

Ultimately, the language processor must build linguistic structures that are well-formed at every level of representation. Thus, a complete account of prosody's role in sentence comprehension must include a specification of all of the grammatical constraints involving prosody and an explanation of when in the course of processing they are invoked, supplemented by an explanation of how prosody affects processing decisions in situations where the grammar allows options (such as when a prosodic structure is well-formed for two or more syntactic structures). Therefore, for each prosodic effect examined here, I will consider the possible impact of both grammatical constraints and independent processing principles. However, the grammatical constraints on prosodic phrasing are still only partially understood, and very little is known about the effect of prosodic phrasing on general processing strategies, so it will not always be possible in the work presented here to determine when an effect of prosodic phrasing on sentence comprehension necessarily results from the processor's enforcement of grammatical constraints and when an effect depends on independent effects of prosody on the processor. Indeed, our understanding of the role of prosody in sentence comprehension will likely develop from a gradual teasing apart of factors which are clearly based on the grammar from those which are clearly based on performance. Where possible, I have attempted to test cases for which known grammatical constraints and possible processing factors make distinct predictions and can therefore be separated from each other. Elsewhere, I have simply tried to identify the possible effects of each of these influences and left the decision about which factor is most likely responsible to considerations of the overall simplicity of grammatical theory and processing theory.

At a superficial level, this dissertation examines a set of effects of prosodic phrase boundaries on disambiguation. At a deeper level, the work takes one well-supported proposal from grammatical theory—the proposal that prosody is phonologically encoded

by the prosodic hierarchy—and investigates the evidence for, and the implications of, the use of this linguistic structure in sentence processing. That is, this is a study of how sentence comprehension should be affected if prosodic boundaries are interpreted not as an unstructured string of cues, but as part of a well-formed linguistic structure. More broadly, it is a study of what must be included in our models of sentence comprehension when the task is not just to parse sentences but to fully process the prosodically-structured utterances of spoken language.

I will argue that the results of several sentence comprehension experiments strongly support the use of a full prosodic representation in processing, and present three different kinds of evidence in support of this view. In Chapter 2, I begin by showing that several plausible hypotheses about the role of prosodic boundaries in syntactic processing are consistent with the Speer *et al.* results, including ones which require that the processor makes use of the prosodic representation and ones which do not. I then report the results of an experiment which discriminates among these hypotheses and argue that the results show that the processor is influenced not simply by the presence or absence of a prosodic boundary, but by the overall pattern of phonological phrasing. Thus, Chapter 2 provides one piece of evidence for the importance of the prosodic representation in sentence processing by showing that, at a single level of the prosodic hierarchy, the effects of multiple prosodic boundaries are only accounted for by interpreting them as part of a prosodic structure.

In Chapter 3, I present a second piece of evidence for the use of the prosodic representation by showing that prosodic boundary information interacts with other information in the prosodic representation to constraint the interpretation of focus. This finding invites reconsideration of the implication from Lehiste's work that prosodic phrasing can only affect syntactic bracketing ambiguities. In Chapter 4, I directly investigate whether the effects of prosodic phrasing extend to non-syntactic processing decisions by exploring their effect on the resolution of lexical ambiguity. Here, I provide a

third piece of evidence for the processor's sensitivity to the prosodic representation by demonstrating that levels of prosodic phrasing that are differentiated in the prosodic hierarchy for English have separable effects on sentence processing.

Thus, this work will show, firstly, that prosodic phrasing has much broader effects on sentence comprehension than merely being a cue that resolves certain syntactic attachment ambiguities, and secondly, that these effects are best accounted for by assuming that the processor recognizes the prosodic representation of a sentence at early stages of processing and is guided by it across multiple levels of processing. In addition, two explicit hypotheses will be proposed for how the processor is guided by the prosodic representation; one of these will account for the effect of phonological phrasing on processing decisions and the other will account for the effect of intonational phrasing. Of course, the conclusion that the prosodic representation is a necessary and influential structure in processing immediately raises many further questions for research on spoken language processing, such as how that prosodic representation is itself parsed. These questions, and the larger implications of prosodically-informed parsing, will be explored in Chapter 5.

CHAPTER 2

PROSODIC PHRASING AND SYNTACTIC ATTACHMENT DECISIONS

2.1 Introduction

The studies described in Chapter 1 together demonstrate that the prosodic structure of a sentence influences the interpretation of several classes of sentences. But while these studies show that prosody does play some role in sentence comprehension, they do not specify what, exactly, that role is. That is, they do not by themselves establish which effects of prosody on comprehension should be accounted for by grammatical factors (i.e., satisfying well-formedness constraints for the phonological or phonosyntactic structure), which effects should be accounted for by performance factors (e.g., choosing a structure favored by the processor when the grammar creates ambiguity), or what the proper characterization of those factors might be. In fact, with the exception of the Speer *et al.* experiments, in which all aspects of the prosody were controlled, the studies do not even show with certainty which aspect of the prosodic structure is the primary cause of the demonstrated processing effects.

Since every utterance has a prosodic structure, these questions must be addressed by any model of processing that attempts to fully characterize sentence comprehension. Merely determining what the disambiguating prosodic element is for a given structure and testing its effects in on-line experiments constitutes an important step forward at this stage, because it allows explicit hypotheses to be developed. Separating grammatical effects from processing effects should also improve our models of sentence processing considerably. Minimally, confirming that an effect is due to a grammatical constraint should establish that any model of processing must be sensitive to whatever prosodic elements are involved (assuming that the processor must take all grammatical constraints into effect at some point), and might also provide some indication of when that information is likely to be used by the processor. Confirming that an effect is due to a processing factor could provide

strong evidence for or against a class of models. For example, if there is an effect which can only be accounted for in processing through, say, a parallel model, it is crucial to know whether there is a grammatical explanation for the effect or not. Finally, if the effects of prosody in sentence comprehension are attributed to specific effects of prosodic elements on the parser, and incorporating such effects entails revising our processing models, we must have explicit, testable hypotheses of what these prosodic effects are and some evidence that the hypotheses of how prosody affects parsing are correct.

As an illustration of some of the issues that must be resolved in accounting for prosody's effect on sentence processing, consider the temporary NP- versus S-complement ambiguity discussed in Chapter 1 and exemplified in (2.1). Most of the studies reviewed in Chapter 1 suggested that this ambiguity can be resolved by prosody, and several of the researchers reported finding a stronger prosodic boundary after the verb in productions of the S-complement structure than in productions of the NP-complement structure. Nevertheless, the resolution of this ambiguity could easily be influenced by factors other than solely the presence versus absence of a prosodic boundary at the point of syntactic ambiguity, such as the larger pattern PPh and IPh boundaries in the sentence, the placement of pitch accents, and the relationships between these elements and the syntactic structure or information structure of the sentence. For example, adding an IPh boundary to (2.1a) to produce (2.1e) does not merely add a factor which might bias the interpretation toward an S-complement structure; it likely creates an ungrammatical utterance by placing two disjoint constituents in a single intonational phrase (cf. Selkirk, 1984). Thus, the presence of an IPh boundary after *suspected* might not bias interpretation toward an S-complement structure when another IPh boundary follows the NP.

(2.1) a. (Mary suspected her boss immediately)$_{IPh}$
b. (Mary suspected her boss was lying to her)$_{IPh}$
c. (Mary suspected)$_{IPh}$ (her boss was lying to her)$_{IPh}$
d. Mary suspected that her boss was lying to her.
e. *(Mary suspected)$_{IPh}$ (her boss immediately)$_{IPh}$

But suppose that the presence versus absence of an IPh boundary after *suspected* does result in easier processing of (c) than (b). The effect of this boundary could be attributed to a processing advantage for the production with the prosodic boundary (i.e. facilitation of the S-complement structure, so that (c) might be as strongly disambiguated as (d)), or to a processing disadvantage for the production without the prosodic boundary (i.e., the lack of an IPh boundary in (b) facilitates the incorrect NP-complement parse, but the IPh boundary of (c) neither facilitates the S-complement parse nor impedes the NP-complement parse), and several different processing hypotheses could be constructed for each of these potential processing effects. Or, the effect could be attributed to the processor's sensitivity to grammatical constraints, such as the proposed grammatical preference for a clause boundary at the IPh boundary for this structure (Nespor & Vogel, 1986).

Thus, to fully specify how prosody affects sentence comprehension, we must know exactly which prosodic features are being contrasted in experimental items and vary only one factor at a time, as in the Speer *et al.* experiments described in Chapter 1. We must also know the possible grammatical constraints on each of these prosodic features at all potentially relevant levels of the grammar and ensure, as much as possible, that all of the relevant constraints are met equally across the experimental conditions. For example, it will not be very informative to contrast the processing of (2.1a) and (2.1e) if the string *her boss immediately* is ill-formed as a single intonational phrase, and other patterns of intonational phrasing in these sentences could cause violations of constraints on the minimal or maximal size of intonational phrases or constraints against unevenly sized intonational phrases (Gee & Grosjean, 1983, Nespor & Vogel, 1986).

As the discussion of example (2.1) illustrates, there are many factors that may be involved in accounting for the role of prosody in sentence comprehension, and most of these factors are only partially understood. Nevertheless, even without fully understanding all of the grammatical constraints on prosody, by testing the comprehension of sentences in which the prosody is tightly controlled and the grammatical constraints are somewhat

clear we should be able to identify some of the basic effects of prosody in processing and begin to discriminate among competing accounts of the effects and the implications those accounts have for sentence processing models. Such experimentation should also allow us to improve our understanding of the grammar by either winnowing out those effects which need not by covered by grammatical constraints or providing converging evidence for constraints which do appear to be supported by experimental results.

In this chapter, I begin by articulating a range of hypotheses that could account for the results of the Speer *et al.* experiments described in Chapter 1, the one case for which the prosodic manipulations, processing effects, and grammatical constraints are relatively well-defined. I show that these results are compatible with several kinds of hypotheses, including one that attributes the effect to grammatical constraints alone as well as four different processing-based hypotheses. I then present new experimental results which discriminate among these possible accounts and further articulate how prosody affects processing. Based on the new results, I argue that prosodic phrasing can affect parsing in manner that does not seem to be accounted for by grammatical constraints, and further, that the processor must make use of a full prosodic representation, as opposed to merely responding to the presence of a prosodic boundary at a point of syntactic ambiguity. In the final section I consider the implications of these results for current models of sentence processing and grammatical theory.

2.2 Accounting for the Effect of Prosodic Phrasing on Parsing

It has long been noted that prosodic boundaries seem to be required at the edges of certain expressions, including parentheticals, appositives, and initial subordinate clauses (e.g., Downing, 1970, 1973; Bing, 1979). If this is the case, then a grammatical constraint alone could account for the Speer *et al.* results. This grammatical account is formalized in (2.2) as the Clause Alignment Hypothesis.

(2.2) Clause Alignment Hypothesis:
Each edge of a matrix clause must be aligned with the edge of a phonological phrase boundary.

Let us assume that the parser initially attempts late closure in both the early closure sentences and the late closure sentences, following morphosyntactic information and the general preference for late closure. However, under the Clause Alignment Hypothesis this analysis would be immediately blocked, and reanalysis would be initiated, if the prosodic information were inconsistent with the constraint in (2.2). Thus, in the early closure cooperating prosody condition the incorrect late closure analysis would be rejected by the absence of a late prosodic boundary, and reanalysis to an early closure structure would be rapidly confirmed by the presence of the early prosodic boundary. In the baseline conditions the ambiguous prosodic information would be consistent with the presence of a late prosodic boundary, so late closure would be allowed in both the early closure condition and the late closure condition, forcing reanalysis in the early closure condition when the disambiguating syntactic information was encountered later in the signal. In the early closure conflicting prosody condition an incorrect first analysis would be confirmed by the presence of the late prosodic boundary, making reanalysis to the early closure structure difficult. And in the late closure conflicting prosody condition the absence of a prosodic boundary at the clause boundary would interfere with confirmation of the late closure parse.[1]

The Clause Alignment Hypothesis retains the primacy of morphosyntactic information for parsing decisions that is found in most current models of sentence processing, and uses prosody only as a follow-up to the morphosyntactically-based decisions (cf. Pynte & Prieur 1996). Thus, it requires no changes to the global structure of these models.

[1] Note that the Clause Alignment Hypothesis only requires the prosody to be consistent with the presence of a prosodic boundary at the clause boundary, accounting for the difference between the late closure baseline prosody condition (which is consistent with the presence of a late prosodic boundary) and the late closure conflicting prosody condition (which is presumably inconsistent with the presence of a late prosodic boundary).

Further, the change it does require is quite minimal: the recognition that the set of grammatical constraints (all of which, presumably, must be adhered to by the processor in selecting the final interpretation of a sentence) includes a constraint like Clause Alignment. Unlike the next four hypotheses that will be presented, the Clause Alignment Hypothesis requires no additional processing mechanism specifically for prosodic phrasing. As such, it is a very attractive hypothesis. However, even though the tendency for there to be a prosodic boundary at matrix clause boundaries appears to be very strong, it is not necessarily the case that there is a grammatical constraint requiring such a phonosyntactic alignment. Further, it is unclear that grammatical constraints alone will suffice to account for the full range of prosodic phrasing effects in processing. Thus, it is also important to consider processing-based accounts of the effect of prosodic phrasing on syntactic structuring.

The first processing-based hypothesis I will consider builds on the generalization of Price *et al.* that location of the major prosodic boundary in a sentence tended to correspond with the location of the major syntactic boundary in the sentence. Given the evidence for the immediate use of prosodic information, this correspondence is perhaps best cast in processing terms as the Parallelism Hypothesis, in (2.3).

(2.3) Parallelism Hypothesis:
> Upon encountering a phonological phrase boundary, the parser selects, from the set of parses with the minimal amount of syntactic structure, the parse with the largest syntactic boundary at the location of the phonological phrase boundary.

The Parallelism Hypothesis can easily account for the early versus late closure results discussed above. An early prosodic boundary would force the parser to attach the subsequent NP into the upcoming clause instead of attaching it as the object in the current clause, accounting for the facilitation effect for early closure sentences with cooperating prosody and the interference effect for late closure sentences with conflicting prosody. A late prosodic boundary would give positive evidence for late closure and interfere with reanalysis in the early closure syntax condition, and the absence of prosodic boundaries

would allow the parser to follow the default strategy of late closure, accounting for the remaining results.

This hypothesis crucially requires syntactic alternatives to be considered in parallel, and, given the results of the Speer *et al.* experiment with phonological phrase boundaries discussed above, minimally it must do so at each point that a phonological phrase boundary is encountered. If this hypothesis proves to be correct, it will present a strong challenge to serial models of parsing, as it would require parallel structures to be considered at several points in a sentence when processing auditory material, suggesting that either all processing includes substantial amounts of parallel processing or that the processing of visually-presented material differs from the processing of auditorially-presented material.

However, it is not necessary to assume parallel processing or the processing load that would be required by the Parallelism Hypothesis to provide a processing-based account of prosodic phrasing effects. For example, the hypothesis given in (2.4) (cf. Schafer *et al.* 1996; Speer *et al.* 1996) treats prosodic boundaries as simple local cues which could be easily incorporated into a serial parsing model.

(2.4) Prosodic Closure:
A phonological phrase boundary marks the right edge of all syntactic constituents currently being processed which can grammatically be closed at that point.[2]

Prosodic closure would cause the parser to align the right edge of each phonological phrase boundary with the right edge of one or more syntactic boundaries, closing the constituent(s). It is an appealing hypothesis, because it allows both rapid syntactic decisions and minimal processing resources. It too can easily account for the early versus late closure results. When the early prosodic boundary is encountered, the constituents currently being processed are the VP of the subordinate clause and its projections. The prosodic boundary would induce closure of the VP and the subordinate clause, so the only

[2] Note that if the Prosodic Closure Hypothesis only required closure of the lowest constituent currently being processed, as suggested in Speer *et al.*, a late boundary would only close the ambiguous NP. This effect, on its own, would not be sufficient to account for the difference

grammatical attachment for the subsequent NP would be as the subject of the main clause, accounting for both the facilitation in the early closure cooperating prosody condition and the difficulty of the late closure conflicting prosody condition. When a late prosodic boundary is encountered, it would induce closure of the NP, VP, and subordinate clause, strengthening the late closure interpretation through the insertion of the right VP and subordinate clause brackets and accounting for the difficulty of the early closure conflicting prosody condition. As before, when no boundary was present the late closure interpretation would emerge because of other parsing principles.

The three hypotheses discussed so far have treated prosodic effects on syntactic attachment decisions as effects of the alignment of prosodic boundaries and syntactic boundaries. Although this seems plausible, it is not the only possibility. Prosodic Visibility casts prosodic phrasing effects not as edge effects but as domain effects: attachment to a node depends on which prosodically-defined domain contains the node.

(2.5) Prosodic Visibility:

a. The phonological phrasing of an utterance determines the *visibility* of syntactic nodes.
b. Nodes within the phonological phrase currently being processed[3] are more visible than nodes outside of that phonological phrase; visibility is gradient across multiple phonological phrases.
c. In first analysis and reanalysis, attachment to a node with high visibility is less costly in terms of processing/attentional resources than attachment to a node with low visibility.[4]

The Prosodic Visibility Hypothesis is based on the assumption that the output of the phonological component of the language processing system and the input to the syntactic

---

between the early closure syntax with baseline prosody and with conflicting prosody.

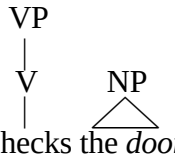
[3] Of course, this is not meant to suggest that syntactic nodes must be part of the phonological representation; it simply means that the syntactic nodes that have been constructed for material within the span of the PPh currently being processed are more visible than nodes constructed for material in other PPhs. This is illustrated below.

[4] Prosodic visibility is a development of proposals in Schafer (1995) and Frazier & Clifton (1995). In addition to claiming that visibility is affected by perceptually-given packages like phonological phrases, Frazier & Clifton argue that recently postulated nodes are more visible than less-recently postulated nodes. For present purposes, I will assume that if visibility is gradient within phonological phrases, the effects are much smaller than the gradiency across phonological phrases.

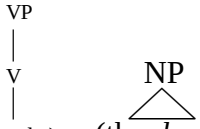
component contains the (partial) prosodic representation that the phonological component has constructed as well as whatever lexical information it has determined. That is, the input to the syntax is not simply a string of words, but rather those words structured into a prosodic representation. The phonological phrasing of this prosodic representation separates the lexical material into prosodically-defined domains of material. These domains then determine the salience of potential attachment sites, on the basis of whether the relevant node is associated with the same phonological phrase as the material the parser is currently processing or is associated to a previous (or subsequent) phonological phrase. Thus, syntactic attachment effects result from the particular syntactic content of the material in a phonological phrase and the overall pattern of phonological phrasing for the utterance.

Using font size to represent visibility, the visibility of the crucial nodes in the Speer *et al.* early versus late closure sentences is shown in examples (2.6) and (2.7), for the point at which the processor encounters *door* or the disambiguating material, respectively.

(2.6) Visibility When *door* is Encountered:



a/f. (Whenever the guard checks the *door*)$_{PPh}$

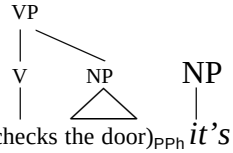


b/e. (Whenever the guard checks)$_{PPh}$ (the *door*

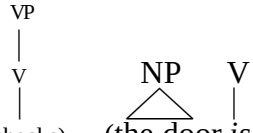
c/d. (Whenever the guard checks the *door*

(2.7) Visibility When *it's* or *is* is Encountered:

Cooperating prosody:

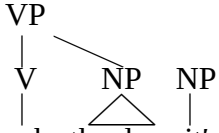


a. (Whenever the guard checks the door)$_{PPh}$ *it's*

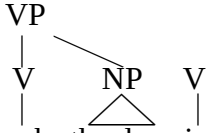


b. (Whenever the guard checks)$_{PPh}$ (the door *is*

Baseline prosody:

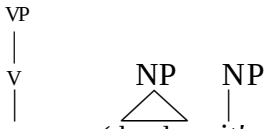


c. (Whenever the guard checks the door *it's*

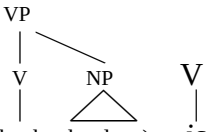


d. (Whenever the guard checks the door *is*

Conflicting prosody:



e. (Whenever the guard checks)$_{PPh}$ (the door *it's*



f. (Whenever the guard checks the door)$_{PPh}$ *is*

In conditions (a) and (c), when no prosodic boundary separates the ambiguous NP from the subordinate verb, the VP node is highly visible to the ambiguous NP because they are in the same phonological phrase. Thus, the NP attaches to the VP node, resulting in a late closure preference. In condition (b), when a phonological phrase boundary separates the verb from the NP, the VP node is no longer in a position of high visibility to the NP, and thus attachment to it should be more difficult and take longer to accomplish. Further, the disambiguating material is encountered immediately, and it occurs in the same phonological phrase as the ambiguous NP, allowing easy reanalysis (if reanalysis is even necessary—it seems likely that disambiguation may occur before the ambiguous NP has been incorrectly attached, for a phrase as short as *the door*) and rapid integration of *is* into the sentence. In condition (d) the high visibility of the subordinate verb to the ambiguous NP results in attachment to the VP and processing difficulty when *is* is encountered. In condition (e) correct attachment of the NP is hindered by the low visibility of the subordinate verb, and in condition (f) incorrect attachment of the NP results from the high visibility of the subordinate verb. Thus, as with the other hypotheses, prosodic visibility can account for the previous experimental results.

Various researchers have made proposals which are similar to Prosodic Visibility in that the prosodic phrasing defines domains of material (Slowiaczek, 1981; Marcus & Hindle, 1990; Pynte & Prieur, 1996). However, in these proposals syntactic processing is determined not so much by the overall pattern of prosodic phrasing, as in Prosodic Visibility, but by whether a phrase which could grammatically be attached in more than one location is grouped with one of the potential attachment sites or not.

Slowiaczek (1981) argues for the parser's use of a prosodic representation on the basis of the experiments described in Chapter 1. However, her model of processing seems to be one in which all syntactic structuring (and possibly some semantic interpretation) is done within a prosodic phrase-based processing unit before the material in that unit can be attached to the larger syntactic phrase marker. In contrast, under Prosodic Visibility the parser does not delay attachments until the end of a prosodic phrase, so material is continually integrated into the partial phrase structure, allowing higher-level processing decisions to be initiated and, presumably, minimizing memory load.

In Marcus and Hindle's model, intonational phrase boundaries separate the string into chunks in the first stage of parsing; these chunks can then be combined (but not split apart) at a later stage of processing, although it is not entirely clear how decisions about combinations are made. Marcus and Hindle assume that no prosodic information other than intonational phrase boundaries is used during initial phrase-structure parsing; phonological phrase boundaries, pitch accents, pitch range, and choice of boundary tone are unavailable at this level in their account. The results of the Speer *et al.* studies, which showed that phonological phrase boundaries affect cross-modal naming times at locations within the sentence, conflict with the claim that phonological phrase boundaries are not used during initial processing, but it is possible that a modified version of this model, in which phonological phrase boundaries are responsible for partitioning the string, could be correct.

Pynte and Prieur (1996) discuss three hypotheses of prosodic phrasing. First, they consider and reject on empirical grounds a model in which prosodic boundaries temporarily block the use of phrase structure rules and thereby temporarily block attachment. Second, they propose that primary attachment decisions may be made through the use of syntactic information alone, and that prosody is only used to aid reanalysis when a need for reanalysis is signaled by non-prosodic information. This proposal would not account for the results presented above. For example, it would not explain the longer reaction times for

late closure syntax with conflicting prosody. Third, they speculate that prosodic phrases may separate the string into constituents which are combined at a later stage, much like the Slowiaczek (1981) or Marcus & Hindle (1990) proposals.

Although the details and theoretical assumptions in these proposals vary, they all share the idea that disambiguation depends on which nodes an ambiguously attached phrase has been grouped with prosodically. For example, in the early versus late closure case, if the NP has been grouped with the verb of the subordinate clause, the parser will assign the late closure structure. If the NP has been grouped with the matrix verb, the parser will assign the early closure structure. In the baseline case, when the grouping allows either attachment, the late closure preference can emerge because of other factors. Thus, this kind of proposal can also capture the Speer *et al.* results. I have formulated the proposal as the Prosodic Chunking Hypothesis, which is given in (2.8).

(2.8) Prosodic Chunking:
Phonological phrases separate the input into domains of material. Attachments are made within a domain before they are made across domains.

These five hypotheses are of course not the only possible accounts of prosody in sentence processing. For example, I have not provided a hypothesis based on constraint-satisfaction models (e.g., Tanenhaus & Carlson, 1989; MacDonald, Pearlmutter & Seidenberg, 1994; Trueswell, Tanenhaus & Garnsey, 1994), as I know of no explicit proposals regarding the use of prosody by researchers working in this framework.[5] Nevertheless, it should be clear that the existing data is compatible with many different kinds of accounts of how prosodic phrasing affects sentence comprehension. In the next

[5] Beach (1989, 1991) argues in support of a fuzzy propositional model for the prosody-syntax relationship based on evidence that durational cues for a prosodic boundary seem to be in a trading relation with intonational cues for a prosodic boundary. (I assume that the integration of durational and intonational information and the postulation of a prosodic boundary take place primarily at the phonetics-phonology interface, consideration of which is beyond the scope of this thesis.) Beach states that the integrated percept of duration and pitch cues influences syntactic interpretation, but does not give an explicit proposal of how it does so (beyond there being some kind of probabilistic relationship between prosodic patterns and syntactic structures).

section, I describe an experiment which discriminates among the five hypotheses I have described above.

2.3 Experiment 1: Prepositional Phrase Attachment

Prepositional phrase attachment ambiguity, like early versus late closure ambiguities, has received considerable attention in the processing literature. An example is given in (2.9).

(2.9) The bus driver angered the rider with a mean look.

The PP in (2.9) can be attached either to the VP, giving a reading in which the driver uses a mean look to anger the rider, or to the NP, in which case the rider has a mean look. Previous psycholinguistic studies have argued for a VP-attachment preference for sentences with PP-attachment ambiguities based on such factors as a structural preference for fewer syntactic nodes (e.g., Frazier, 1978) or a preference for arguments over adjuncts (e.g., Abney, 1989), although others have disputed the generality of a VP-attachment preference, arguing that factors such as frequency, plausibility, or the discourse structure are more important (e.g., Altmann & Steedman, 1988; Taraban & McClelland, 1988; MacDonald, Pearlmutter & Seidenberg, 1994; Spivey-Knowlton & Sedivy, 1995).

Regardless of how strong or consistent the preference for VP attachment may be for these sentences, all of the prosodic hypotheses presented above predict that the resolution of PP-attachment ambiguity should be influenced by the prosodic structure of the sentence. Sentence (2.9), like all sentences, can be produced with a range of well-formed prosodies. Experiment 1 tested the interpretation of sentences with PP-attachment ambiguities with the set of contours indicated in (2.10), which contrast in the presence of phonological phrase boundaries before the direct object and before the prepositional phrase. Although the sentences ranged in strength of VP-attachment bias, the set as a whole showed a 60% bias toward VP attachment in condition (c), which is predicted to be the most neutral production by all of the hypotheses presented above.

(2.10) Sample Set for PP-Attachment Experiment:

a. ((The bus driver angered the rider L-)$_{PPh}$ (with a mean look L-)$_{PPh}$ L%)$_{IPh}$

b. ((The bus driver angered L-)$_{PPh}$ (the rider with a mean look L-)$_{PPh}$ L%)$_{IPh}$

c. ((The bus driver angered the rider with a mean look L-)$_{PPh}$ L%)$_{IPh}$

d. ((The bus driver L-)$_{PPh}$ (angered L-)$_{PPh}$ (the rider L-)$_{PPh}$ (with a mean look L-)$_{PPh}$ L%)$_{IPh}$

The prosodic hypotheses make distinct predictions about the overall pattern of preferred interpretations for this set of pronunciations. Consider first the four processing-based hypotheses. The Parallelism Hypothesis, which forces the parser to select the largest syntactic boundary possible when it encounters a prosodic boundary, predicts a greater number of VP-attachment interpretations for conditions (a) and (d), which have a boundary before the PP, than for the conditions without a boundary at that location. However, the phonological phrase boundary before the NP in conditions (b) and (d) would have no effect under Parallelism, assuming that there is only one grammatical attachment option for the phrase *the rider*.[6] Thus, in (a), the pre-PP boundary should cause a strong bias for VP attachment of the PP. In (b), the pre-NP boundary has no effect, and there is no pre-PP boundary, so the parser is predicted to follow the default preference for VP attachment (perhaps resulting in a weaker bias for VP attachment than in (a)). Since the pre-NP boundary is predicted to have no effect on the interpretation of the sentence, condition (c) is predicted to be identical in its percentage of VP-attachment interpretations to condition (b), and condition (d) is predicted be identical to (a). Because the boundary before the PP only reinforces the preferred interpretation, the (a) and (d) versus (b) and (c) difference predicted

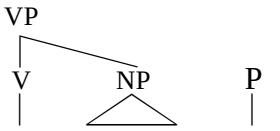
[6] The verbs used in Experiment 1 did not allow continuation with a sentential complement, as the verb *knew* does in a sentence like: *The bus driver knew the rider was looking for a seat.* As currently formulated, the Parallelism Hypothesis does not predict that a prosodic boundary after a verb which can take an NP- or S-complement would lead to the S-complement interpretation, since such a parse would require more syntactic structure than the NP-complement parse. Other formulations of Parallelism, which allowed a wider range of constructions to be considered, might predict an S-complement resolution in such a situation. However, such formulations might also predict a preference for S-adjoined structures like parentheticals any time a prosodic boundary is encountered.

by Parallelism might be small, so Parallelism is consistent with both a significant main effect of a late prosodic boundary or a weak effect of a late prosodic boundary.

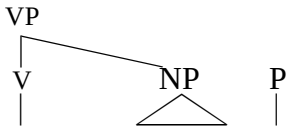
Prosodic Closure also predicts that there should be no effect of the pre-NP boundary of (b) and (d). Under Prosodic Closure, the pre-PP boundary would cause the parser to close the NP and attach the PP to the VP in conditions (a) and (d). However, there are no syntactic constituents which can grammatically be closed at the point of the pre-NP phonological phrase boundary, so the boundary would have no effect. As with Parallelism, Prosodic Closure predicts that conditions (b) and (c) should show the default preference for VP attachment, with identical attachment percentages to each other.

In contrast to the Parallelism Hypothesis and the Prosodic Closure Hypothesis, Prosodic Visibility predicts that the pre-NP phonological phrase boundary will affect the attachment of the PP. According to Prosodic Visibility, the salience of a potential attachment site is highest when it is in the same phonological phrase as the material currently being processed, and lowers with each preceding phonological phrase. The visibility of the relevant nodes in the experimental materials is illustrated in example (2.11).

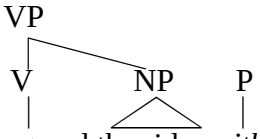
(2.11) Visibility When *with* is Encountered:



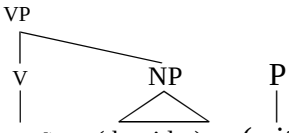
a. (The bus driver angered the rider)$_{PPh}$ (*with*



b. (The bus driver angered)$_{PPh}$ (the rider *with*



c. (The bus driver angered the rider *with*



d. (The bus driver)$_{PPh}$ (angered)$_{PPh}$ (the rider)$_{PPh}$ (*with*

In condition (a), the pre-PP phonological phrase boundary would lower the visibility of all of the preceding nodes, including both the NP node and the VP node, resulting in the default VP-attachment preference for the PP. In condition (b) however, the phrase *the rider*

is in the same phonological phrase as the PP, and so the NP node would be highly visible to the parser at the time when the PP is being processed, while the VP node in the preceding phonological phrase has lower visibility. Thus, the pre-NP phonological phrase boundary is predicted to significantly raise the percentage of NP attachments and lower the percentage of VP attachments. In condition (c) there are no internal phonological phrase boundaries, so the default VP-attachment preference should emerge, as in condition (a). Finally, the pair of internal phonological phrase boundaries in condition (d) would make the NP node less visible than in condition (b) but more visible relative to the VP node than in condition (a). Thus prosodic visibility predicts that the percentage of VP attachments for condition (d) should be between those for (a) and (b).

Interestingly, the Prosodic Chunking Hypothesis predicts that there should be no effect of the pre-PP phonological phrase boundary, because it separates the PP from both of the nodes it could attach to. Thus, Prosodic Chunking predicts equal percentages of attachments for conditions (a), (c), and (d). Prosodic Chunking does predict an effect of the pre-NP boundary, but only in condition (b), in which the pre-NP boundary groups the NP node with the PP. It predicts that there should be no effect of the pre-NP boundary in condition (d), because the pre-PP boundary has separated the PP into its own processing domain.

The grammatically-based Clause Alignment Hypothesis does not apply in Experiment 1. However, it is possible that condition (b) is ungrammatical for the structure with VP attachment of the PP, because the final phonological phrase contains two (disjoint) constituents, the object NP and the VP-modifying PP. Several proposals in the literature rule out prosodic phrases composed of disjoint constituents like these (Selkirk, 1984; Steedman, 1991; Hirst, 1993; Croft, 1995), although it is not clear that whether these constraints should apply to phonological phrases as well as intonational phrases. If these grammatical proposals are correct, and they apply to phonological phrases, then condition

(b) should show a markedly lower percentage of VP-attachment interpretations than the other three conditions, which should not differ from each other.

To summarize, there are three distinct patterns of results that are predicted by the set of hypotheses laid out above. First, Parallelism and Prosodic Closure predict that there should be no effect of the pre-NP boundary, but a small effect of the pre-PP boundary raising the percentage of VP attachments in conditions (a) and (d). Second, Prosodic Visibility predicts a main effect of the pre-NP boundary lowering the percentage of VP attachments in conditions (b) and (d), as well as a small effect of the pre-PP boundary raising the percentage of VP attachments in condition (d) relative to condition (b). And third, Prosodic Chunking predicts an interaction effect, instantiated by markedly lower percentage of VP attachments in condition (b) but equally high percentages of VP attachments in conditions (a), (c), and (d). This third pattern is also predicted by the general hypothesis that all effects of prosodic phrasing on parsing are due to grammatical constraints in conjunction with the specific proposal that a phonological phrase cannot be composed of disjoint constituents. These predictions were tested with an end-of-sentence simple comprehension time experiment in which subjects answered disambiguating questions at the end of each experimental sentence.

2.3.1 Materials

Sixteen sets of materials like those in (2.12) were created. The full list of materials appears in Appendix A. The materials were recorded by a native speaker of English in a sound-attenuated booth. The speaker placed high pitch accents (either H*, !H*, or L+H*) on the subject, verb, direct object, and the object of the preposition in each condition. Phonological phrase boundaries marked by L- phrase accents were placed before the PP, before the NP, in neither location, or in both locations, as marked in (2.12) and shown in the sample pitch tracks in Figure 2.1. An additional boundary was placed after the subject in condition (d) so the phonological phrase containing the direct object would not be markedly shorter than the other phonological phrases in the sentence. Phonetic analyses

verified that the materials were produced as intended and did not significantly vary across conditions except in the placement of phonological phrase boundaries.

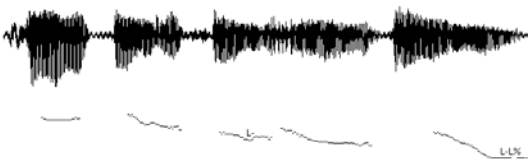
(2.12) Transcribed Sample Set for PP-Attachment Experiment:

H* H* H* H* H*
a. ((The bus driver angered the rider L-)$_{PPh}$ (with a mean look L-)$_{PPh}$ L%)$_{IPh}$

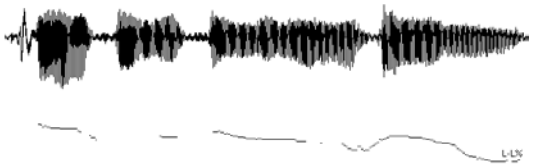
H* H* H* H* H*
b. ((The bus driver angered L-)$_{PPh}$ (the rider with a mean look L-)$_{PPh}$ L%)$_{IPh}$

H* H* H* H* H*
c. ((The bus driver angered the rider with a mean look L-)$_{PPh}$ L%)$_{IPh}$

H* H* H* H* H*
d. ((The bus driver L-)$_{PPh}$ (angered L-)$_{PPh}$ (the rider L-)$_{PPh}$ (with a mean look L-)$_{PPh}$ L%)$_{IPh}$



a. (Paula phoned her friend L-) (in Alabama L-L%)



c. (Paula phoned her friend in Alabama L-L%)

b. (Paula phoned L-) (her friend in Alabama L-L%)

d. (Paula L-) (phoned L-) (her friend L-) (in Alabama L-L%)

Figure 2.1. Sample Pitch Tracks for Experiment 1.

2.3.2 Task

The experimental items were presented to subjects following a Latin-square design, so that each subject heard only one condition from each set of items, with the condition rotating through the items. Each subject received a different randomization of the sixteen experimental items and 92 other sentences which had various prosodic and syntactic

structures. The subject listened to the sentences over speakers in a sound-attenuated booth and pressed a lever as soon as he or she had comprehended each sentence. After each experimental sentence the subject received a visually-presented question such as *Who had a mean look?* The subject responded orally into a microphone and an experimenter, seated outside of the booth, coded the response. Questions also appeared after approximately 30% of the distracter sentences. The entire procedure, including a practice session with six sentences, took less than thirty minutes to complete.

2.3.3 Subjects

Forty-eight undergraduates at the University of Massachusetts participated in exchange for course credit. All subjects reported that they were native speakers of English with normal hearing.

2.3.4 Results

The percentage of responses that reflect VP attachment of the PP for each condition appear in Table (2.1).[7] There was no main effect of the pre-PP boundary ($F_{(1,47)}$=2.6, p<.12; $F_{(1,15)}$=1.6, p<.23), but the percentages of VP-attachment responses were significantly lower when the pre-NP boundary was present ($F_{(1,47)}$=9.5, p<.004; $F_{(1,15)}$=6.8, p<.02). The interaction was non-significant ($F_{(1,47)}$<1; $F_{(1,15)}$=1.1, p<.32). Pairwise tests revealed that there were significantly more VP-attachment responses in condition (c) than in condition (b) ($t_{(47)}$=2.6, p<.01; $t_{(15)}$=3.0, p<.01). The difference

Table 2.1. Percentages of VP-attachment Responses, Experiment 1.

	Pre-PP Boundary	No Pre-PP Boundary
Pre-NP Boundary	52.6% (condition d)	44.3% (condition b)
No Pre-NP Boundary	61.5% (condition a)	59.9% (condition c)

[7] Table (2.1) reports the percentages of VP-attachment responses out of all responses to the open-choice question, which include VP-attachment responses, NP-attachment responses, and other responses (e.g., "Not sure"). The pattern of responses was very similar when other responses was excluded: 65.1% in (a), 47.2% in (b), 63.8% in (c), and 56.5% in (d).

between conditions (a) and (d) was marginal by subjects but did not reach significance by items ($t_{(47)}$=1.8, p<.08; $t_{(15)}$=1.6, p<.18), and the difference between conditions (b) and (d) was marginal by both subjects and items ($t_{(47)}$=1.7, p<.10; $t_{(15)}$=1.8, p<.09).

This pattern of results supports the Prosodic Visibility Hypothesis and disconfirms the other four hypotheses. Only the Prosodic Visibility Hypothesis predicted the main effect of the pre-NP boundary seen in these results. The Parallelism Hypothesis and the Prosodic Closure Hypothesis both fail to predict any effect of the pre-NP boundary, and the Prosodic Chunking Hypothesis and the grammatical account both predict that the pre-NP boundary should only show an effect in condition (b), resulting in a significant interaction.

2.3.5 Discussion

The finding that a pre-NP phonological phrase boundary can affect PP attachment has three important implications for research on sentence comprehension. First, it provides additional evidence that the subtle acoustic cues provided by phonological phrase boundaries are sufficient to cause significant effects on interpretation. This provides another piece of evidence against claims that only unnatural or exaggerated prosody can disambiguate sentences (Watt & Murray, 1996). It also shows that the syntactic processor must be capable of responding to such boundaries, even when they occur in places which are not syntactically ambiguous, lending further support to the claim that the syntactic processor is sensitive to a prosodic representation constructed by a phonological processing module, and is not simply interrupted by overt acoustic events like the long pauses associated with intonational phrase boundaries (Marcus & Hindle, 1990).

Second, previous comprehension studies on prosody and attachment decisions have considered only the effects of the prosodic boundaries that occurred at points of syntactic ambiguity, ignoring the potential effects of prosodic elements located before or after the critical decision point in the string being parsed.[8] Experiment 1 shows that the surrounding

[8] Research done at the same time as Experiment 1, by Pynte & Prieur (1996) on PP attachment in French, employed a similar design to Experiment 1 and found similar effects. Their results will be discussed in Chapter 5, when I consider the effects of prosody on the

prosody matters: the pre-NP boundary was located prior to the decision point about the PP attachment, yet it still affected that attachment decision. Moreover, it affected the attachment decision even when another phonological phrase boundary intervened, in condition (d). Therefore, Experiment 1 demonstrates that the influence of the entire prosodic structure must be taken into account when evaluating the effect of prosody in parsing. It shows that hypotheses like Parallelism, Prosodic Closure, or Prosodic Chunking, which are sensitive to only a portion of the prosodic phrase structure, do not seem to be sufficient to account for the effects of prosodic phrasing in parsing.

Taken together, the finding that phonological phrase boundaries are sufficient to cause effects on syntactic attachment decisions and the finding that these boundaries need not be located at the point of syntactic ambiguity to have effects suggest that the effects of prosody may be far more pervasive than often assumed, as phonological phrase boundaries can occur several times in a typical sentence. Thus, all auditory studies—not just those testing prosodic hypotheses—may need to be much more careful in controlling for the effects of phonological phrase boundaries.

Finally, the fact that the pre-NP boundary affected both condition (d), which was predicted to be prosodically well-formed for either syntactic structure, and condition (b), which was potentially ill-formed for the VP-attachment structure, suggests that there are effects of prosodic phrasing on sentence comprehension that are not accounted for by grammatical constraints alone. Since all of the results from the various experiments on prosody described in Chapter 1 can be plausibly accounted for strictly through the parser's sensitivity to the grammar, Experiment 1 is the first study to provide evidence for independent effects of prosodic phrasing on parsing decisions.

2.4 General Discussion

The experiments described in this chapter and the previous chapter provide considerable evidence that prosodic boundaries affect syntactic attachment decisions. More importantly,

---

processing of languages other than English.

they provide strong evidence that it is really the prosodic structure which affects parsing, not simply the local effect of a prosodic boundary at a point of syntactic ambiguity blocking the syntactically-preferred structure. This implies that the processor must, at some point of processing, build a prosodic representation for each sentence as well as a syntactic representation. Further, given the evidence from Speer *et al.* (1996) that prosodic boundaries affect interpretation at least as early as the point at which syntactic disambiguation is encountered, it seems highly likely that the construction of the (partial) prosodic representation precedes the construction of the (partial) phrase structure, and then actively guides parsing through the effects of a processing constraint like Prosodic Visibility.

Crucial evidence for this conclusion has come from the use of multiple fully grammatical prosodic contours for a single syntactic structure, which allowed the effects of prosodic boundaries in particular locations to be studied independently. Thus, by contrasting cooperating prosody, conflicting prosody, and baseline (ambiguous) prosody, Speer *et al.* were able to prove that prosodic phrasing can both interfere with parsing and facilitate parsing. And by contrasting the presence versus absence of phonological phrase boundaries in two locations, Experiment 1 showed that prosodic boundaries located prior to the point of ambiguity can influence interpretation.

Further, the studies presented in these chapters strongly constrain the possible hypotheses of how prosody affects sentence comprehension. They show that prosodic phrasing information must be used during sentence processing, and that it must first become available at an early stage of processing, most likely before syntactic parsing is initiated. They also show that prosodic information is probably made available as part of a full prosodic representation, and not as isolated cues. Moreover, they show that the parser does not simply use the information in the prosodic representation to adhere to prosody-syntax interface constraints, as Experiment 1 demonstrated that prosodic boundaries can

have effects on parsing that are not accounted for by any currently recognized grammatical constraints.

Finally, this chapter shows that processing-based prosodic effects can be easily incorporated into one of the dominant models of sentence processing through the Prosodic Visibility Hypothesis. As many recent articles on prosody have noted, most of the work in sentence processing has focused on syntactic, and to a lesser degree, semantic and pragmatic processing. While most researchers probably recognize that prosody should be incorporated into sentence processing models, there has been little discussion in the literature of how this should be done. When prosodic disambiguation has been discussed, prosodic boundaries have often been characterized as biasing the parser toward choosing the syntactic option with the larger syntactic boundary at the location of the boundary, implicitly suggesting that parallel models of processing would be necessary to account for prosodic disambiguation effects. However, this chapter shows that prosodic effects can be readily accounted for under the assumptions of a serial model of processing.

2.5 Conclusion

This chapter has provided additional evidence that prosodic phrasing plays an important role in syntactic processing. It has also disconfirmed several initially plausible hypotheses of how prosodic boundaries influence parsing and provided a successful account in the Prosodic Visibility Hypothesis. According to Prosodic Visibility, the only direct effect of phonological phrase boundaries is to define the edges of phonological phrases. However, because phonological phrases determine the visibility of syntactic nodes, phonological phrasing can indirectly influence syntactic processing.

By seeing prosody as a structure composed of phonological phrases instead of as a string of cues, Prosodic Visibility is able to account for prosodic effects that would otherwise seem to be the result of non-local cues (such as the effect of the pre-NP boundary on PP attachment) as well as effects of prosodic boundaries located at syntactically ambiguous points. Thus, Prosodic Visibility covers the disambiguation

effects found in at least two kinds of syntactic structures, including those with early versus late closure ambiguities and those with minimal versus non-minimal attachment ambiguities. Further, Prosodic Visibility does not require attachment decisions that span prosodic boundaries to be delayed until the end of a prosodic phrase has been encountered, as seems to be required in other models which use prosodic phrases to define processing domains.

Because Prosodic Visibility is stated in terms of phonological phrase boundaries, it does not predict that there should be (processing-based) differences in interpretation between sentences with medial intonational phrase boundaries and sentences with only phonological boundaries in medial positions, since the presence of an intonational phrase boundary entails the presence of a phonological phrase boundary. Although this may seem appropriate given the similar results in the Speer *et al.* studies for sentences with medial intonational phrase boundaries and sentences with medial phonological phrase boundaries, the phonological distinction between them alone suggests that they might have separable effects on processing. This issue will be addressed in Chapter 4. But before that, in Chapter 3, I consider one of the implications of a full prosodic representation being passed from the phonological component to other processing components. Namely, if a full prosodic representation is made available to higher levels of processing, then information of different kinds (such as prosodic phrasing information versus pitch accent information) might interact with each other at those higher levels of processing. Thus, we might see prosodic boundaries influencing forms of interpretation based on pitch accent information, or the accentuation pattern influencing forms of interpretation based on prosodic phrasing. Chapter 3 will address the former possibility, and Chapter 5 will consider the latter one.

CHAPTER 3

THE EFFECT OF PROSODIC PHRASING ON FOCUS INTERPRETATION

3.1 Introduction

Although there is ample work from theoretical and experimental linguistics showing the importance of pitch accents in marking given information, new information, and focal structure in English, there has been little work for English relating other aspects of prosody and information structure. Further, there has been very little work testing the effects of prosodic phrasing on processing decisions that do not involve syntactic attachment ambiguities, and the studies that have tested non-structural ambiguities have generally concluded that prosodic phrasing cannot disambiguate such sentences (Lehiste, 1973; Lehiste, Olive & Streeter, 1976; Wales & Toner, 1979). However, it is well-known from grammatical studies of languages other than English that focus can affect the prosodic phrasing of an utterance. In this chapter I will argue, based on the results of a comprehension experiment for English, that the reverse can also hold true: the prosodic phrasing of an utterance can affect the interpretation of focus. Specifically, I will claim that prosodic phrasing blocks the projection of focus, limiting focus to material in the prosodic phrase which contains the focusing pitch accent. I shall refer to this as the *bounded projection* effect. I will further argue that although the bounded projection effect could, in principle, be accounted for by postulating additional grammatical constraints, it is also possible to account for the effect through an independently motivated principle of processing, thus avoiding the need to further complicate either the grammar or the processing system.

The structure of the chapter is as follows: in Sections 3.2 through 3.4 I present background information for the chapter and the bounded projection experiment. In Section 3.5, I consider whether there is an explanation for the results from independent constraints in the grammar. First, I show that the characterization of prosodic phrases as interpretive

units found in Steedman (1991) fails to capture the array of focus possibilities for English. I then argue that while the bounded projection effect could be accounted for through the addition of an Optimality Theoretic constraint on focus projection, doing so only results in an otherwise unnecesary complication of the grammar, as the comprehension facts fall out from an independently motivated parsing principle. Specifically, I argue that the bounded projection effect can be accounted for by the principle of Prosodic Visibility introduced in Chapter 2, but that it is not easily accounted for under different assumptions about the role of prosody in sentence comprehension. Thus, this chapter provides additional evidence for the role of prosodic phrasing in sentence comprehension, establishes that prosodic phrasing can affect focus interpretation as well as bracketing decisions, and lends further support to the Prosodic Visibility hypothesis.

3.2 Psycholinguistic Research on Focus

Previous psycholinguistic studies have suggested that material in focus, whether explicitly signaled by accenting or only anticipated by a question context, is taken as marking the important information in a sentence and may therefore receive more processing attention (Cutler & Foss, 1977; Cutler & Fodor, 1979). In production, new or focused material is generally distinguished from given material through one or more markings of accent, such as longer duration, greater amplitude, or different patterns of pitch excursion (Lieberman, 1967; Fowler & Housum, 1987). In perception, words that have been excised from their context are more easily identified if they originally expressed new information than if they expressed given information (Fowler & Housum, 1987). In sentence comprehension, words marked as presenting given information facilitate the retrieval of previous context better than words marked as presenting new information (Terken & Nooteboom, 1987), and appropriate markings of given versus new information facilitates picture verification (Terken & Nooteboom, 1987) and sentence comprehension times (Bock & Mazella, 1983; Birch & Clifton, 1995). Further, material which could be taken as modifying one of two phrases is more likely to be associated with accented

material than with unaccented material (Schafer, Carter, Clifton, & Frazier, 1996). Thus, the accentual pattern of material can have a significant effect on how that material is processed, as well as the processing of subsequent information. However, although the focus of a sentence is dependent on its accentual pattern, it is not identical to it. As described in the next section, focus can be "projected" from an accented word to include additional, unaccented material.

3.3 Phonological Constraints on Focus

3.3.1 Focus Projection

I will assume that arguments which present new information must be accented (Selkirk, 1984) with an appropriate pitch accent (Pierrehumbert & Hirschberg, 1990), such as an H*. I will also follow Selkirk's theory of focus projection (1984, 1995b). In this theory, accented words are *F-marked*. F-marking can project from internal arguments to heads, from heads to phrasal nodes, and from NP- and *wh*-moved elements to their traces. The focus of a sentence is then defined as the F-marked constituent which is not dominated by any other F-marked constituent. Thus, the production in (3.1), which carries an H* pitch accent on *potatoes*, can have focus as marked by brackets and underlining in (3.1a) - (3.1d), in which the accented word *potatoes* is F-marked, and the F-marking optionally projects to the NP, VP, or sentence.

(3.1) Focus Projection:

a. The farmer delivered some $[\underset{H^*}{\text{potatoes}}]_{FOC}$

b. The farmer delivered $[\underset{H^*}{\text{some potatoes}}]_{FOC}$

c. The farmer $[\underset{H^*}{\text{delivered some potatoes}}]_{FOC}$

d. $[\underset{H^*}{\text{The farmer delivered some potatoes}}]_{FOC}$

3.3.2 Prosodic Phrasing and Focus

There are languages which require that a prosodic phrase be associated with focus. For example, in Chichew^a a broadly-focused VP, is produced with just one prosodic phrase (marked with parentheses), as shown in (3.2a), but utterances of the VP with a narrow, contrastive focus require a prosodic phrase break at the right edge of the focus (Kanerva, 1990), as shown in (3.2b - 3.2d). Other languages which require a prosodic phrase break at the edge of a focused constituent include Bengali (Hayes & Lahiri, 1991), Korean (e.g., Jun, 1993), and Japanese (e.g., Nagahara, 1994).

(3.2) The Effect of Focus on Phrasing in Chichew^a:

a. (Anaménya nyumbá ndí mwalá)
hit house with rock
'He hit the house with a rock'

b. (Anaménya nyumbá ndí [mwalá]$_{FOC}$)
'He hit the house with [a rock]'$_{FOC}$

c. (Anaménya [nyumbá]$_{FOC}$) (ndí mwalá)
'He hit [the house]$_{FOC}$ with a rock'

d. ([Anaménya]$_{FOC}$) (nyumbá) (ndí mwalá)
'He [hit]$_{FOC}$ the house with a rock'

English does not seem to require a prosodic phrase break at either the left or right edge of a focused constituent. Thus the sentence in (3.3), with narrow focus on *delivered*, is well-formed when produced as a single prosodic phrase, as in (a), or when produced with a prosodic break at the left edge of the focus, as in (b); at the right edge, as in (c); or with breaks at both edges, as in (d), as long as other constraints on prosody, such as the one requiring each phonological phrase to contain a pitch accent, are not violated.

(3.3) Phrasing Options for Focus in English:

a. (The farmer [delivered]$_{FOC}$ some potatoes)
L* H* L*

b. (The farmer) ([delivered]$_{FOC}$ some potatoes)
L* H* L*

L* H* L*
c. (The farmer [<u>delivered</u>]$_{FOC}$) (some potatoes)

L* H* L*
d. (The farmer) ([<u>delivered</u>]$_{FOC}$) (some potatoes)

However, example (3.4) shows that prosodic phrasing and focus do interact in English. The response in (3.4), with accents on the subject and object, no accent on the verb, and a prosodic phrase break before the object, is intuitively unnatural as an answer to a VP-focus question. Example (3.5) shows that there is nothing inherently unnatural about the phrasing, as it is fine with an object-focus question.

(3.4) What did the farmer do?
L* H*
(The farmer delivered) (some potatoes)

(3.5) What did the farmer deliver?
L* H*
(The farmer delivered) (some potatoes)

Examples (3.3) through (3.5) show that, in English, focus does not need to be identical in span to the prosodic phrasing, or align with either edge of the prosodic phrasing, but it cannot extend beyond the prosodic phrasing to include unaccented material. These facts suggest the Bounded Projection Hypothesis, given in (3.6), which will be tested in the comprehension experiment described below.

(3.6) Bounded Projection Hypothesis:
F-marking cannot project beyond the prosodic phrase which contains the accented material.

Sample predictions of the Bounded Projection Hypothesis for a sentence produced with an H* accent on the object and no accent on the verb are given in (3.7). In (a), where the F-marking created by accenting *potatoes* does not project, the sentence is predicted to be well-formed. In (b), the F-marking must project to the NP node, but because that node is associated with material within the span of the prosodic phrase containing the accented material, projection is predicted to be possible and the sentence is again predicted to be

well-formed. However, sentences (c) and (d) require projection to syntactic nodes which are not associated with the prosodic phrase containing the F-marked material, such as the V node of *delivered*, so projection is predicted to be blocked and the sentences are predicted to be ill-formed.

(3.7) Sample Predictions of the Bounded Projection Hypothesis:

a. (The L* farmer delivered) (some [H* potatoes]$_{FOC}$)

b. (The L* farmer delivered) ([some H* potatoes]$_{FOC}$)

c. * (The L* farmer [delivered) (some H* potatoes]$_{FOC}$)

d. * ([The L* farmer delivered) (some H* potatoes]$_{FOC}$)

3.4 Experiment 2: Focus Projection

The Bounded Projection Hypothesis was tested in a sentence comprehension experiment in which subjects rated the naturalness of question-answer pairs. The experiment included two kinds of questions, a VP-focus question (e.g., *What did the farmer do?*) and an object-focus question (e.g., *What did the farmer deliver?*). This allowed comparison of the naturalness ratings for question-answer pairs like (3.4) and (3.5), which differ in predicted naturalness for answers with identical prosody.

There were three kinds of answers in the experiment, differing only in prosody. These three answers were fully crossed with the two questions, creating six experimental conditions. A sample set of materials is given in (3.8); the full list of tonally-transcribed materials is provided in Appendix B. The first kind of answer, the *object-phrasing* answer, was as in (3.4) and (3.5). It was composed of two prosodic phrases. The first prosodic phrase contained the subject and verb, and the second prosodic phrase contained the object. The sentence carried an L* accent on the subject, no accent on the verb, and an H* accent on the object. Since there was no accent on the

verb, VP-focus required the projection of F-marking from the accented object to the verb and then to the VP. The Bounded Projection Hypothesis predicts that such projection is impossible in this case, though, because the verb and VP node are in a separate prosodic phrase from the object. However, object focus is predicted to be possible for this prosody, because the F-marking does not need to project beyond the prosodic phrase containing the object. Thus, the Bounded Projection Hypothesis predicts that condition (a), with an object-focus question and an object-phrasing answer, should be judged as natural, but condition (b), with a VP-focus question and an object-phrasing answer, should be judged as unnatural.

(3.8) Sample Set of Materials for Experiment 2:

a. What did the farmer deliver?
L* H*
(The farmer delivered) (some potatoes)

b. What did the farmer do?
L* H*
(The farmer delivered) (some potatoes)

c. What did the farmer deliver?
L* H*
(The farmer) (delivered some potatoes)

d. What did the farmer do?
L* H*
(The farmer) (delivered some potatoes)

e. What did the farmer deliver?
L* L+H*
(The farmer) (DELIVERED some potatoes)

f. What did the farmer do?
L* L+H*
(The farmer) (DELIVERED some potatoes)

In the second kind of answer, the *VP-phrasing* answer, the accents were the same as in the object-phrasing answer: L* on the subject, no accent on the verb, and H* on the object. The VP-phrasing answer differed from the object-phrasing answer only in the prosodic phrasing. The first prosodic phrase contained the subject, and the second

prosodic phrase contained the verb and object. With this prosody, the verb is in the same prosodic phrase as the accented material, so under the Bounded Projection Hypothesis both object-focus and VP-focus should be possible. Thus, the Bounded Projection Hypothesis predicts that condition (c), with an object-focus question and a VP-phrasing answer, and condition (d), with a VP-focus question and a VP-phrasing answer, should both be judged as natural. By including these two conditions any basic differences in judged naturalness between question-answer pairs with VP-focus questions and those with object-focus questions can be separated from any effects of bounded projection.

The third kind of answer, the *control* answer, contained accents which were inappropriate for the information structure of the sentence, providing an unnatural response to each of the questions. As in the VP-phrasing answer, the subject formed the first prosodic phrase and the verb phrase formed the second one, but the accents in second prosodic phrase differed: in the control answer the verb received an L+H* accent and the object was unaccented. Because the object presents new information in all of the experimental sentences, it was expected that the lack of accent on the object in this contour would cause both the object-focus question-answer pair (condition (e)) and the VP-focus pair (condition (f)) to be judged as unnatural.[1] Because they had VP phrasing, conditions (e) and (f) also served to prevent the subjects from associating unnaturalness with the object-phrasing found in the condition predicted to be unnatural by the Bounded Projection Hypothesis. The predictions for all of the conditions are summarized in Table (3.1).

[1] Condition (e) was also unnatural because of the presence of accent on the verb, as the verb presented non-contrastive given information.

Table 3.1. Predicted Naturalness Judgments, Experiment 2.

	Answer:		
Question:	Object-Phrasing	VP-phrasing	Control Answer
Object-focus	natural (condition a)	natural (condition c)	unnatural (condition e)
VP-focus	unnatural (condition b)	natural (condition d)	unnatural (condition f)

3.4.1 Materials

Twenty-four sets of experimental materials similar to the sample set in (3.8) were constructed, along with twelve filler question-answer pairs and five practice pairs. Eight of the fillers contained subject *who* questions and answers with sentence-initial accent; four fillers contained *where* questions and sentence-final accent. All of the fillers had natural prosody.

The materials were recorded by two native speakers of English in a sound-attenuated chamber using a high-quality microphone and digitized at 11.025 kHz. The questions were produced by a male speaker and the answers were produced by a female speaker so the two voices would sound noticeably different from each other and form a natural dialogue. Each speaker was recorded in a separate session. For each experimental set, two question tokens were recorded (object-focus and VP-focus) and three answer tokens were recorded (object-phrasing, VP-phrasing, and control.)

The pitch accents of the experimental materials was as described above: L* accents on the subject, H* accents on the object and no accent on the verb for the object-phrasing answer and the VP-phrasing answer, and L+H* accents on the verb and no accent on the object for the control answer. The Bounded Projection Hypothesis does not distinguish between phonological phrases and intonational phrases, and intuitions suggest that both phonological phrase boundaries and intonational boundaries block the projection of focus. Thus, the strongest test of the Bounded Projection Hypothesis

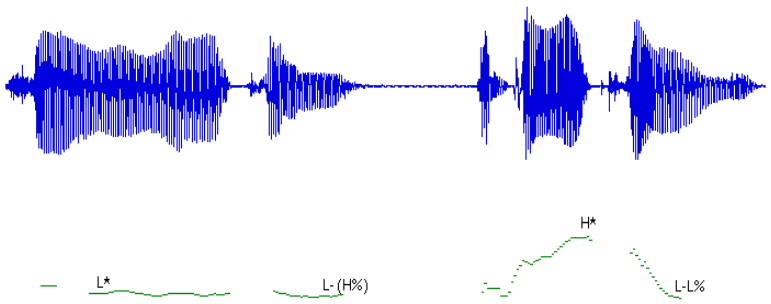
would contrast the placement of phonological phrase boundaries and keep intonational phrasing constant. However, because subjects were to be tested in groups in a regular classroom, intonational phrase boundaries were used. These boundaries were marked by a L- phrase accent and a (somewhat attenuated) H% boundary tone, which is generally an appropriate contour for a sentence which will continue. Phonetic analyses confirmed that the sentences were produced as intended. Sample pitch tracks of the three answers are shown in Figure 3.1 on the next page.

3.4.2 Acceptability Test of Answers

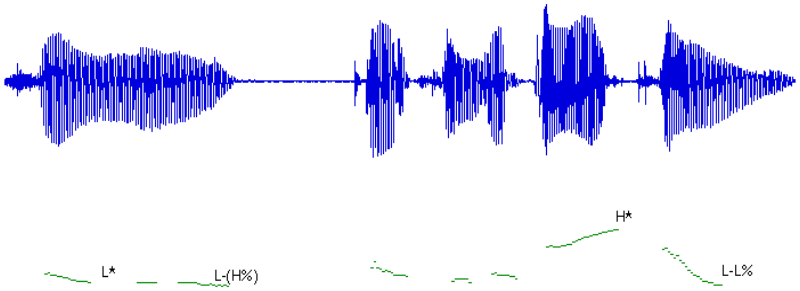
To ensure that none of the answers were prosodically ill-formed, the three answers were presented out of context (without the questions) as part of a separate experiment employing a good/bad grammaticality judgment task. In this task, the subjects listened to each sentence, presented over speakers in a sound-attenuated booth. At the end of each sentence, the subject pressed one lever if the sentence sounded "good" and another level if it sounded "bad." All three answers were judged as "good" over 90% of the time, with no significant differences among them. Thus, any differences in naturalness found in the main experiment should be due to the naturalness of the answer as a response to the question, and not inherent differences in well-formedness among the answers.

3.4.3 Task

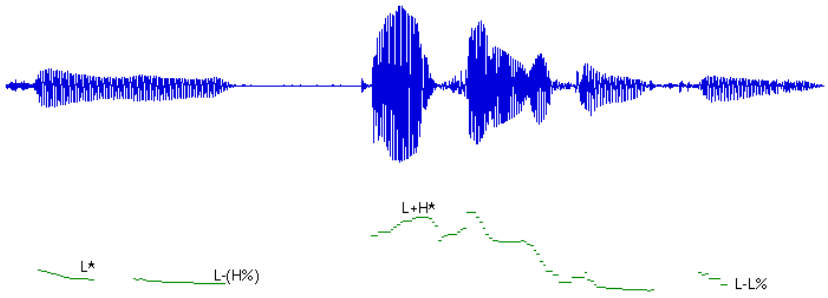
The digitized materials were recorded back to six tapes. Each tape contained five practice pairs and then twenty-four experimental items and twelve fillers, in random order, with conditions assigned by Latin-square. Thus, subjects heard only one condition from each set of items, with the condition rotating through the items. Each answer was spaced one second after the offset of its corresponding question in order to simulate a natural timing for the discourse. Question-answer pairs were separated by four seconds. Subjects, in groups of one or more, listened to the question-answers pairs from one of the six tapes, played over speakers in a quiet classroom. After each



L* H*
a. (Julie opened) (a bakery)



L* H*
b. (Julie) (opened a bakery)



L* L+H*
c. (Julie) (opened a bakery)

Figure 3.1 Sample Pitch Tracks for Experiment 2.

question-answer pair, each subject rated the naturalness of the answer for its question by circling a number from 1 to 5 on a printed answer sheet. The answer sheet did not list the questions or answers. No subject showed difficulty in completing the rating within the four seconds provided between question-answer pairs.

3.4.4 Subjects

Forty-eight undergraduate students from the University of Massachusetts participated in the experiment in exchange for course credit. All subjects reported that they were native speakers of English.

3.4.5 Results

The mean naturalness rating for each condition is given in Table (3.2). Higher numbers reflect greater judged naturalness. As predicted, ratings were lower for the two conditions containing the control answer, which had inappropriate placement of pitch accents, and for the condition (b), in which the object-phrasing answer was predicted to block the focus-projection required for the VP-focus question.

Table 3.2. Naturalness Ratings, Experiment 2.

	Answer:		
Question:	Object-phrasing	VP-phrasing	Control Answer
Object-focus	3.4 (condition a)	3.2 (condition c)	2.3 (condition e)
VP-focus	2.9 (condition b)	3.1 (condition d)	2.6 (condition f)

The ANOVA for conditions (a-d) yielded a significant effect of question ($F_{(1,47)}$ = 22.5, $p < .01$; $F_{(1,23)}$ = 12.0, $p < .01$), no effect of answer, and a significant interaction ($F_{(1,47)}$ = 9.1, $p < .01$; $F_{(1,23)}$ = 5.0, $p < .04$). Planned contrasts showed that the predicted difference between conditions (a) and (b) was significant ($F_{(1,47)}$ = 26.4, $p < .01$; $F_{(1,23)}$ = 24.2, $p < .01$) while the 0.1 difference between conditions (c) and (d) was not significant. This pattern of results matches the predictions of the Bounded Projection Hypothesis and strongly supports it.

3.4.6 Discussion

The results of the bounded projection experiment support several claims about prosodic marking of information status in English. First, the combination of the high

percentage of "good" ratings for the control answer in the grammaticality judgment task and the low ratings for it in the bounded projection experiment provide further experimental evidence that accenting old, non-contrastive information and failing to accent new information is ungrammatical. This is consistent with other studies in the literature on the effect of accent placement on information structure (Bock & Mazzella, 1983; Nooteboom & Kruyt, 1987; Terken & Nooteboom, 1987; Eefting, 1992; Birch & Clifton, 1995).

Second, the significant interaction between focus and phrasing establishes that the prosodic specification of focus is influenced by the pattern of prosodic phrasing in the utterance as well as by the pattern of pitch accents. Third, the lack of significant difference between condition (c), with a object-focus question and a VP-phrasing answer, and condition (d), with a VP-focus question and a VP-phrasing answer, suggests that English neither requires the focused constituent to correspond to a prosodic phrase nor requires the left edge of the focused constituent to be aligned with the left edge of a prosodic phrase. In fact, the ratings are numerically higher in the condition in which focus and prosodic phrasing do not correspond than in the condition in which they do.

Fourth, these results strongly support the Bounded Projection Hypothesis. The pairwise comparisons show that the object-phrasing answer was significantly less natural as an answer to a VP-focus question than an object-focus question, but a VP-phrasing answer was not significantly different in naturalness for the two kinds of questions. That is, the naturalness judgments were significantly lower only for the case in which F-marking had to project outside of a prosodic phrase, as predicted by the Bounded Projection Hypothesis. However, while the bounded projection experiment gives clear evidence that focus interpretation is affected by prosodic phrasing, the results are consistent with either a processing effect of prosodic phrasing on focus interpretation or with the existence of a grammatical constraint on focus and prosodic

phrasing. Thus, in the next section I consider whether the bounded projection effect is best accounted for through a grammatical explanation or a processing one.

3.5 Grammatical Explanations for Bounded Projection

The bounded projection effect could easily result from a constraint of the grammar. I will consider two possible grammatical accounts of the bounded projection effect here. One possibility comes from work of Steedman (1991). Another plausible explanation comes from the addition of a phonological constraint which specifically restricts cross-boundary focus projection.

Steedman (1991) proposes a categorial grammar for prosody which, with slight modification, would correctly rule out the condition predicted to be unnatural by the Bounded Projection Hypothesis. The account would be very much like that for (3.9) (his (53)):

(3.9) I know that ALICE read a BOOK.
But what about FRED? What did HE do?

? (FRED ate) (the BEANS).
L+H* LH% H* LL%
Theme *Rheme*

According to Steedman's grammar of prosody, an L+H* accent combines with a high boundary (L-H%) to create a theme, and an H* accent combines with a low boundary (L- or L-L%) to create a rheme. Unaccented words like *ate* can match any category, but the combination of categories is restricted to block combination across prosodic boundaries. Hence, the relevant interpretation of (3.9) is blocked because *ate* and *beans* cannot combine across the prosodic boundary to form a rheme.

However, Steedman's work only covers a fragment of the possible prosodic contours for English and is highly restrictive: only two kinds of pitch accents (H* and L+H*) and three kinds of end contours (L-, L-L%, and L-H%) are included, H* accents must be followed by either an L- or L-L% end contour, L+H* accents must be followed by an L-H% end contour, and L+H* accents can only be used for themes.

Thus, Steedman's grammar does not provide an interpretation for the control answer, which contained an L+H* pitch accent followed by an L-L% end contour, nor does it provide an interpretation for extremely common contours like the one given in (3.10). More importantly, once the grammar is expanded to account for (3.10) as an appropriate broadly-focused utterance, it would presumably have to allow such an interpretation for (3.11) as well (since nothing prevents *delivered* from matching the category assigned to *farmer*) even though intuitions suggest that (3.11) cannot have such an interpretation. Therefore, Steedman's account does not appear to be sufficient to capture the facts about focus projection for the full range of English intonational contours.

(3.10) (The farmer H* L-) (delivered some potatoes H* L-L%)

(3.11) (The farmer delivered H* L-) (some potatoes H* L-L%)

A more tenable approach is to simply add a constraint similar to the Bounded Projection Hypothesis to the grammar, such as the Optimality Theoretic constraint proposed in (3.12). In the Optimality Theoretic framework of ranked, violable, universal constraints (Prince & Smolensky, 1993), a sentence with VP-focus and object-phrasing would violate a constraint like BOUNDFOCUS. Such a contour should be selected by the grammar only when some other more highly-ranked constraint forces the presence of a prosodic boundary within the focused constituent. Presumably no such constraint exists that would apply to the utterances used in the bounded projection experiment, thus ruling out the object-phrasing answer as an appropriate response to the VP-focus question.

(3.12) BOUNDFOCUS[2]
F-marking cannot project across a prosodic phrase boundary.

The important question, though, is whether it is necessary to postulate such a constraint, either to account for the data or to improve the grammar. It does not seem to be that case that adding the constraint would improve the grammar. In fact, doing so would contradict the strong and interesting proposal put forth by Truckenbrodt (1995) that there is no direct connection between focus and prosodic phrasing.

Truckenbrodt argues that the cross-linguistic effects of focus on phrasing discussed in Section 3.2 can be captured through Optimality Theoretic constraints which connect prominence with focus, prosodic structure, and syntactic structure and connect prosodic structure and syntactic structure but do not directly connect focus and prosodic phrasing. Focus and phrasing patterns such as those found in Chichew^a, according to this account, stem from the interaction of the need for the focused element to be the most prosodically prominent element and general constraints governing the alignment of syntactic phrases and prosodic phrases.

Truckenbrodt's proposal cannot account for the bounded project effect, but as I will show below, it is not necessary to account for the effect through grammatical constraints. And, even if Truckenbrodt's account proves to be incorrect, any time the (universal) set of constraints is increased, the number of possible rankings, and therefore, the difficulty of learning the grammar, is (in principle) increased. Thus, while there does not seem to be evidence against adding some kind of bounded projection constraint to the grammar, neither is there evidence that it is necessary or preferable to do so.[3]

---

[2] Note that the constraint must limit the projection of focus (or the relationship between F-marked nodes), and not the span of focus, as it is possible to present focus in more than one prosodic phrase if each phrase has the appropriate pattern of pitch accents (Birch & Clifton, 1995). Thus, the constraint cannot be a member of the WRAP family of constraints (Truckenbrodt, 1995), which require XPs to be contained within a phonological phrase.

[3] If a constraint like BOUNDFOCUS is part of the grammar, it should be possible to see the interaction of it with other constraints. For example, if a constraint militating against lengthy

3.6 A Processing Explanation for Bounded Projection

It may initially seem unlikely that there could be a parsing explanation for the bounded projection effect. Though the psycholinguistic literature has shown that differences in accent placement can affect the interpretation of information structure, the crucial conditions in the bounded projection experiment did not differ in accentuation. As discussed in Chapters 1 and 2, many studies on prosodic phrasing have shown that prosodic phrasing can be used to disambiguate sentences with different syntactic bracketings, yet the bounded projection sentences all had the same phrase structure. Moreover, the studies on prosody that have examined ambiguities other than bracketing differences (also discussed in Chapter 1) were unable to show reliable effects of prosodic phrasing on discrimination for these kinds of ambiguities.

It is in principle possible that a prosodic boundary could function as a cue to the processor that it has encountered the edge of a focused constituent. Such an effect seems very appealing: it would no doubt be useful for the parser to have access to such information, and it could be viewed as a form of disambiguation of bracketing through prosodic phrasing. Unfortunately, this explanation cannot be correct. It fails to account for the pattern of results found in the NP-focus question conditions, where naturalness ratings were as high for the condition in which the prosodic boundary and the edge of the focused constituent did not coincide as for the condition in which they did.

However, the bounded projection effect falls out directly from the Prosodic Visibility hypothesis introduced in Chapter 2 and repeated in (3.13). Prosodic Visibility predicts that an operation like the projection of F-marking should be more difficult when it involves a low-visibility node. This is exactly what was found in the

---

prosodic phrases dominated BOUNDFOCUS (and the use of additional pitch accents on the focussed material was also blocked by some set of constraints), then the bounded projection effect could disappear for long constituents but remain for short constituents. Such a pattern of data could not be explained by the processing hypothesis given in Section 3.5, and thus would be convincing evidence for a grammatical account of the bounded projection effect.

bounded projection experiment: ratings were lower only in the case when F-marking would have had to project to a low visibility node.

(3.13) Prosodic Visibility
a. The phonological phrasing of an utterance determines the *visibility* of syntactic nodes.
b. Nodes within the phonological phrase currently being processed are more visible than nodes outside of that phonological phrase; visibility is gradient across multiple phonological phrases.
c. In first analysis and reanalysis, attachment to a high-visibility node is less costly in terms of processing/attentional resources than attachment to a low-visibility node.

Because the crucial materials contained unaccented verbs, VP-focus was only possible if F-marking projected from the F-marked direct object to the verb and then to the VP node. In the VP-phrasing answer, the verb and object were part of the same prosodic phrase, and therefore the V and VP nodes were highly visible at the point when the processor would have been processing the object and projecting F-marking. Thus, both NP-focus and VP-focus should have been possible. The results verify that both NP-focus and VP-focus were indeed judged as natural for the VP-phrasing answer.

In the object-phrasing answer, when the verb was not in the same prosodic phrase as the F-marked material, the V and VP nodes would have had low visibility at the point when the parser was processing the object and attempting to project F-marking. The contrast between the high visibility of the V and VP nodes in the VP-phrasing answer and the low visibility of them in the object-phrasing answer is illustrated in example (3.14). As in the PP-attachment situation described in Chapter 2, in which PPs avoided attachment to low-visibility nodes, F-marking should avoid projection (or attachment of F-marking) to low-visibility nodes. Thus, for the object-phrasing answer, NP-focus should have been possible but VP-focus should have been difficult. Again, this is what the results show. Focus is bound within the prosodic phrase which contains the accented material because it is unable to project to the low-visibility nodes of a previous prosodic phrase.

(3.14) Visibility When *potatoes* is Encountered:

Because Prosodic Visibility is already needed to account for the PP-attachment results discussed in Chapter 2, the visibility explanation of the bounded projection effect adds no further mechanism to the grammar or the processing system. Thus, given our current evidence, the visibility explanation should be preferred to the grammatical explanation proposed above.

Assuming that the bounded projection is indeed a processing effect, the results of the bounded projection experiment provide converging evidence for the Prosodic Visibility Hypothesis. Note that the visibility account captures three seemingly disparate effects of prosodic phrasing on processing. First, in the bounded projection experiment, a prosodic boundary constrained the interpretation of focus. Second, in conditions of the PP-attachment experiment which had a late (pre-PP) prosodic boundary, as well as in other cases in the literature (e.g., Price *et al.*, 1991; Speer *et al.*, 1996), the prosodic boundary resulted in the attachment of immediately following material to the higher of two potential attachment sites (e.g., PP attachment to the VP node instead of the NP node.) Third, in the conditions of the PP-attachment experiment with an early (pre-NP) prosodic boundary, the prosodic boundary seemed to have a non-local effect, inducing low attachment of material which did not immediately follow the boundary.

None of the other processing hypotheses considered in Chapter 2 can account for all three of these effects. Particularly limited are those which require some syntactic event (such as the closing of a syntactic phrase) to be cued by the presence of a prosodic boundary. These hypotheses cannot capture the seemingly non-local effect in

the PP-attachment case, they predict general difficulty in the object-phrasing case, since they predict that the processor should attempt to close the VP, and they make no predictions about the effect of prosodic phrasing on focus interpretation.

What appears to be most important in the visibility account is that the parser interprets a prosodic boundary not as some kind of simple cue but as the edge of a prosodic phrase. The prosodic phrasing can then determines visibility, and visibility can then constrain the possibilities for both F-projection and attachment. Thus, an analysis which employs prosodic structure and visibility provides a unified account for the various effects of prosodic boundaries. Further, it does so in a theoretically-constrained fashion: all effects of prosody on sentence comprehension come from either general processing principles (such as visibility) or grammatical constraints on prosodic structure (such as those which provide the prosodic hierarchy.)

3.7 Conclusion

The results of the comprehension experiment presented in this chapter provide strong evidence that prosodic phrasing affects the interpretation of focus in English. The focus of a sentence can span the same material as a prosodic phrase or be contained within a prosodic phrase, but it cannot extend to unaccented material which is outside of the prosodic phrase containing the focusing accent. While this result may suggest the need for a grammatical constraint limiting the projection of focus, I have shown that the pattern can also be accounted for by the independently-motivated processing principle of Prosodic Visibility, avoiding the need to complicate the grammar with a new constraint.

The finding that the interpretation of focus in English is influenced not just by the presence or absence of pitch accents but also by the pattern of prosodic phrasing has important consequences for both formal linguistic research and psycholinguistic research. Regarding formal linguistics, the results point out potential complications for work on focus within syntax and semantics. Whenever a syntactic structure is

obligatorily associated with a prosodic break (e.g., for parentheticals, appositives, gapping structures, etc.), focus should be unable to project beyond these breaks. The projection of focus should also be blocked in sentences which require a prosodic break because of their length (Gee & Grosjean, 1983).[4] Thus, in some sentences it may not be possible to determine whether syntactic factors directly limit focus, because they are confounded with prosodic factors.

As for psycholinguistics, this work shows that the parser must somehow make use of prosodic phrasing information and accent information together, supporting models which make use of the prosodic hierarchy over models which segregate prosodic phrasing from accents (Marcus & Hindle, 1990). This will be discussed further in Chapter 5, when the larger picture of prosody in processing is considered. The results also suggest that the effects of prosodic phrasing on processing might not be as limited as once thought. Although the effect of prosodic phrasing on focus interpretation that I have argued for here is an effect of prosody on syntax (since it depends on the visibility of syntactic nodes), the results show that the effects of prosodic phrasing can extend beyond the basic construction of the phrase marker. Thus, it seems possible that prosodic phrasing might affect other kinds of interpretation as well. The next chapter addresses this question.

[4] Here, I am assuming that the bounded projection effect must always hold. As discussed above, if the effect stems from a violable constraint focus projection could be possible in these situations in a language with the appropriate ranking of constraints.

CHAPTER 4

PROSODIC PHRASING AND INTERPRETATION

4.1 Introduction

Previous experimental work on the role of prosody in sentence comprehension has largely concentrated on only two questions—the role of pitch accents in establishing the information structure of a sentence (by signaling given, new, or focused information) and the effects of various aspects of prosody, but most notably, the effect of prosodic phrasing, in syntactic disambiguation. In this chapter, I consider the role of prosodic phrasing in other kinds of interpretation by keeping syntactic and focal structures constant and examining the effect of prosodic phrasing on lexical interpretation and integration.

Previous studies have also tended to be more concerned with showing that it is (or is not) possible to use prosody to disambiguate sentences than with showing exactly how specific aspects of prosody might affect the processor. Thus, few studies have been explicit about the prosody used in the experiments. Of the studies that have provided a transcription or a detailed description of the prosody, none, to my knowledge, have directly compared or contrasted the effects of phonological phrasing and intonational phrasing by contrasting the two levels within a single experiment. However, their status as distinct levels of prosodic structure might lead one to expect that they could have distinct effects in sentence processing. Based on experimental results reported below, I shall argue that they do.

To the extent that earlier sentence comprehension work has allowed the comparison and contrast of phonological phrases and intonational phrases across experiments, the results have suggested that these two levels of phrasing have similar effects. The Speer *et al.* studies on early versus late closure ambiguities discussed in Chapters 1 and 2 (Kjelgaard, 1995; Speer *et al*., 1996) showed the same pattern of effects for stimuli with phonological phrase boundaries as for stimuli with intonational phrase boundaries. However, this

similarity of behavior for syntactic decisions does not rule out the possibility that phonological phrase boundaries and intonational phrase boundaries could have distinct effects in sentence comprehension for other types of processing decisions. Indeed, if syntactic attachment decisions depend on prosodically-triggered visibility, which is determined by the phonological phrasing, then it is not surprising that phonological phrase boundaries and intonational phrase boundaries could show similar syntactic effects, as the presence of an intonational phrase boundary entails the presence of a phonological phrase boundary.

Thus, if intonational phrase boundaries cause different processing effects from phonological phrase boundaries, the effects may not be evident in cases involving syntactic ambiguities. Therefore, this chapter investigates the possibility that intonational phrases, but not phonological phrases, define semantic/pragmatic processing domains for the language processor, with the presence of an intonational phrase boundary leading to further semantic/pragmatic processing than would be required by a phonological phrase boundary or the absence of a prosodic boundary. In Section 4.2 I summarize the reasons that we might expect a phonological phrase versus intonational phrase distinction in processing and an effect of intonational phrasing on interpretation. I then lay out intuitive evidence suggesting that intonational phrasing can affect the timing of semantic/pragmatic processing, and consequently the interpretation of some sentences. In the following sections I present evidence from experiments on adjective interpretation and the reanalysis of lexical ambiguity in support of the hypothesis that intonational phrases form interpretative units and in support of a processing distinction between the two levels of phrasing. In the final sections I consider the implications of these findings for linguistic theory and psycholinguistic research.

4.2 The Interpretive Domain Hypothesis

Given the differences between phonological phrase boundaries and intonational phrase boundaries, it might be expected that they could have distinct effects on processing. The

two levels correlate with different acoustic patterns, they are phonologically distinct, and they are governed by distinct sets of grammatical constraints.[1] The acoustical differences alone suggest ways in which phonological and intonational phrases might cause different semantic or pragmatic processing effects. Recall that intonational phrase boundaries are associated with a greater degree of final lengthening and longer pauses than phonological phrases, and that intonational phrase boundaries are marked by two edge tones while phonological phrase boundaries are marked by just one. Therefore, an intonational phrase boundary could, in principle, cause different effects in comprehension from a phonological phrase boundary merely because intonational phrase boundaries are in general more acoustically salient than phonological phrase boundaries and thus are presumably more easily detected by the processing system. Or, the two boundaries might cause different effects because intonational phrase boundaries generally allow more processing time before the arrival of new material. That is, there could be an effect of more extensive semantic or pragmatic processing taking place at intonational phrase boundaries than at phonological phrase boundaries, but only because intonational phrase boundaries allow more time for that processing to take place. These acoustically-based hypotheses predict that there should be gradient effects of prosodic phrasing, such that acoustically stronger boundaries would show stronger processing effects, but that phonological phrase boundaries and intonational phrase boundaries would not cause processing effects fundamentally different in type.

A second, more interesting, view draws on the phonological distinction between phonological phrases and intonational phrases and predicts that the two levels of prosodic phrasing could show truly distinct kinds of effects in processing that could not be accounted for by their respective acoustic properties alone. In this view, phonological phrase boundaries would cause the effects on visibility discussed in Chapter 2, so both

[1] It is also possible that intonational phrase boundaries differ from phonological phrase boundaries in psychological status. While intonational phrase boundaries are available for conscious manipulation and can be easily identified by a naive language user, it seems to require some linguistic training to consciously identify or consciously produce phonological phrase boundaries.

sentences with phonological phrase boundaries and sentences with intonational phrase boundaries could show syntactic effects of prosodic phrasing. However, intonational phrase boundaries might cause effects that phonological phrase boundaries would not trigger, not because of a gradient, acoustic difference between them, but because of the categorical, phonological distinction between the two, allowing intonational phrase boundaries to be associated with different grammatical constraints or different processing principles than phonological phrase boundaries. Thus, sentences containing medial intonational phrase boundaries might show prosodic phrasing effects on interpretation that sentences containing only medial phonological phrase boundaries would not show. Further, because of the phonological distinction, an intonational phrase boundary could cause different effects from a phonological phrase boundary even if both boundaries were easily detectable and the two boundaries had matching durational properties.

Considering what is already known about prosodic effects in sentence comprehension, it is reasonable to assume that prosodic phrasing information would be available in the appropriate component of the processing system to affect semantic processing. As I argued in Chapters 2 and 3, there is strong evidence that the prosodic representation for an utterance affects processing decisions beyond the phonological level and at least through the syntactic level, since both pitch accents and prosodic phrasing can influence syntactic structuring and the interpretation of focus. I argued in earlier chapters that the phonological component of the processing system passes both lexical information and the prosodic structure of a sentence to the syntactic component. It is equally plausible—and equally consistent with assumptions about modularity—that this information can be passed again to a semantic or pragmatic component. Indeed, there is ample intuitive evidence that different kinds of pitch accents and edge tones—that is, high pitch accents versus low pitch accents, high versus low boundary tones, and so forth—can have strong effects on the ultimate interpretation of an utterance (e.g., Pierrehumbert & Hirschberg, 1990). Thus, this information must be available to the interpretive components of the processor. And, if

details about different kinds of phrase accents and boundary tones are available to the processor at a post-syntactic level, then information about their presence or absence must be available as well. The question, then, is whether this information can affect interpretive decisions, and if so, how and when it does so.

Little is known yet about semantic and pragmatic processing, so it is difficult to say with certainty what effect prosody could have on interpretive decisions when it is not yet fully clear what those interpretive decisions might be like. Nevertheless, as an initial attempt at addressing the question, I propose that prosodic phrasing affects semantic and pragmatic processing as specified by the Interpretive Domain Hypothesis (IDH), given in (4.1).[2]

(4.1) The Interpretive Domain Hypothesis:
An intonational phrase boundary defines a point at which the processor performs any as yet outstanding semantic/pragmatic evaluation and integration of material within the intonational phrase.

Crucially, the Interpretive Domain Hypothesis does not claim that all semantic and pragmatic processing must wait until the right edge of an intonational phrase boundary is encountered by the language processor. As assumed in other models of sentence processing, I assume that much of the semantic processing that takes place will occur rapidly and incrementally as each new word is identified in the input and structured into the representation of the sentence. However, the Interpretive Domain Hypothesis does predict that semantic processing which would otherwise take place at a later point in time—perhaps because such processing has not yet been required or because the processor has adopted a

[2] The IDH is somewhat similar to the Selkirk (1984) proposal that intonational phrases must form 'sense units' and the Steedman (1991) proposal that prosodic structure, syntactic structure (in a Categorial Grammar framework), and information structure are isomorphic, in that intonational phrases define some kind of interpretive domain in each proposal. However, the IDH differs from the other two in two important ways. First, the IDH distinguishes between intonational phrases and phonological phrases, while the other two do not. Second, the IDH is solely concerned with what happens at the right edge of an intonational phrase, during processing, and places no constraints on the material that composes the intonational phrase. In contrast, Selkirk's Sense Unit Condition and Steedman's Prosodic Constituent Condition are, fundamentally, prosody-syntax interface constraints, concerned with the well-

local delay strategy—would have to take place when an intonational phrase boundary is reached. Note also that the Interpretive Domain Hypothesis is stated in terms of intonational phrases, not prosodic phrases. I claim that the IDH only applies to intonational phrasing, and not to phonological phrasing. Thus, the IDH predicts a processing distinction between these two levels of prosodic phrasing for certain kinds of processing decisions.

As an illustration of how the Interpretive Domain Hypothesis would apply, consider the resolution of the syntactic category ambiguity found in phrases like *the desert trains*, in which *trains* can be a noun modified by *desert* (as in *The desert trains are dusty*) or a verb with *desert* as its subject (as in *The desert trains hikers to be vigilant*). Frazier and Rayner (1987) have argued on the basis of eye movement data that the parser delays commitment to the syntactic categories for novel phrases like *desert trains* until it receives disambiguating information (but see MacDonald (1993) for an opposing view).[3] Thus, for a string such as *the desert trains are…*, the parser would not assign syntactic categories to *desert* and *trains* until it encountered *are*, which disambiguates the string to the modifier-noun interpretation.

According to the Interpretive Domain Hypothesis, when an intonational phrase boundary follows one of these ambiguous strings, interpretation of the string—and therefore, syntactic category assignment—is required. Intuitions support this prediction; listeners perceive that interpretation of the ambiguous phrase takes place immediately when an intonational phrase boundary falls at the right edge of the string, with the interpretation following the lexical bias of the ambiguous string, but perceive that interpretation is delayed when a phonological phrase boundary immediately follows the ambiguous phrase. Thus, intuitions suggest that there is little difference in processing difficulty between the (a) versions and the (b) versions of sentences (4.2) and (4.3), in which the (a) versions

---

formedness of prosodic phrases. Selkirk's and Steedman's conditions, and their possible roles in sentence processing, are discussed more fully in Chapters 2 and 3.

[3] MacDonald argues from the results of a study contrasting familiar NN-biased phrases such as *tax returns* with novel NV-biased phrases such as *fund returns* against the delay model and

contain an intonational phrase boundary after the ambiguous string and the (b) versions contain a phonological phrase boundary, and, crucially, the sentences continue in a manner which is appropriate for the locally-preferred analysis of the ambiguous string, regardless of whether that local preference is as a noun-verb sequence (as in (4.2)) or as a modifier noun sequence (as in (4.3)).[4]

(4.2) a. (The library holds)$_{IPh}$ (many rare books)
b. (The library holds)$_{PPh}$ (many rare books)

(4.3) a. (The official documents)$_{IPh}$ (were misplaced)
b. (The official documents)$_{PPh}$ (were misplaced)

(4.4) a. (The library holds)$_{IPh}$ (are really useful)
b. (The library holds)$_{PPh}$ (are really useful)

(4.5) a. (The official documents)$_{IPh}$ (the complaints)
b. (The official documents)$_{PPh}$ (the complaints)

However, when the continuation is inappropriate for the inherent bias, as in (4.4) and (4.5), there is a marked difference in processing difficulty between sentences with a medial intonational phrase boundary and those with a phonological phrase boundary, according to intuitions, and a perception of being "garden-pathed" in the cases with intonational phrase boundaries. Thus, (4.4), with a noun-verb bias and a modifier-noun continuation, seems significantly harder to parse in the (a) version, with an intonational phrase boundary, than in the (b) version, with a phonological phrase boundary, and the same is true for (4.5), which has a modifier-noun bias and a noun-verb continuation.

In all of these sentences, the presence of the medial intonational phrase boundary in the (a) versions seems to force interpretation. Because there is no context available to resolve the ambiguity, the interpretation follows the lexical biases. If the continuation fits with that interpretation, intuitions suggest that processing is easy, as in (4.2a) and (4.3a); if it does not, processing is difficult, as in (4.4a) and (4.5a). When only a phonological phrase

---

for a constraint-based model in which non-syntactic factors such as the frequency of modifying noun usage of a word influence the (immediate) resolution of the ambiguity.

boundary is present, interpretation is not forced, so the syntactic category assignment can be delayed until disambiguating information arrives later in the sentence. Thus, processing seems to be relatively easy in all of the (b) cases as there is no case in which the ambiguous material must be reanalyzed because of conflicting information in the sentence continuation. This pattern is exactly as predicted by the Interpretive Domain Hypothesis.

Although examples (4.2 - 4.5) provide good intuitive evidence for an interpretive effect of an intonational phrase boundary on the processor, the construction is not ideally suited for the first experimental test of the Interpretive Domain Hypothesis. The presence of a prosodic boundary immediately after the ambiguous string also seems to have a biasing effect toward an interpretation of the ambiguous string as the subject of the sentence, and therefore toward the modifier-noun interpretation (Kjelgaard, 1995). Thus, an experiment on materials like these would require a large number of conditions to control for the possible syntactic bias of the prosodic boundary and allow the results to be securely interpreted; such an experiment would not be an efficient first test of the Interpretive Domain Hypothesis.[5]

Clearly, what is needed to test the hypothesis is a case in which the syntax does not differ across the conditions, the prosodic phrasing does not interact with any syntactic decisions or preferences, and some aspect of the processing of the material happens sufficiently late in time (either because other factors require delay or because it is necessarily a late stage of processing) for the intonational phrase boundary to have an effect. One case which appears to meet these criteria is the interpretation of context-sensitive adjectives, which will be discussed in the next section. A second case involves the reanalysis of lexical ambiguity, which will be the topic of Section 4.4.

---

[4] These follow the lexical biases collected with a written questionnaire and reported in Kjelgaard (1995).

[5] Recent work has verified the disambiguating effect of IPhs boundaries for NN-biased strings (NV-biased strings were not tested) (Speer, Kjelgaard, O'Bryan & Perry, 1997; Schafer & Speer, 1997; Speer & Schafer, 1997).

4.3 Experiment 3: Context-Sensitive Adjective Interpretation

There is a considerable amount of evidence that the interpretation of lexically ambiguous nouns such as *ruler* or *ball* takes place immediately. For example, cross-modal priming studies have shown that, in general, all meanings of such words are activated initially, and then one meaning is selected within a few hundred milliseconds, either on the basis of the semantic/pragmatic context of the word, if available, or on the basis of the frequency and recency of usage of the different meanings (e.g., Swinney, 1979; Tanenhaus, Leiman & Seidenberg, 1979; Seidenberg, Tanenhaus, Leiman & Bienkowski, 1982; Simpson 1994). However, there is good reason to believe that the full interpretation of adjectives is not as rapid.

Adjectives vary in the manner in which they receive an interpretation. Some adjectives are *intersective*; their interpretation results from the intersection of the meaning of the adjective and the meaning of the noun it modifies. For example, a *pregnant physicist* is someone who is both pregnant and a physicist, and, if she is Swedish and plays the violin, she is also a pregnant Swede and a pregnant violinist. Other adjectives are *subsective*—the adjective selects a subset of the set defined by the head noun. For example, a *junior senator* is a relatively junior member of the set of senators, but not necessarily someone who is inherently junior, and a junior senator who plays the violin is not necessarily a junior violinist. Finally, some adjectives are neither intersective nor subsective: a *former senator* is no longer a senator, an *alleged criminal* may not actually be a criminal, and *faux fur* is not (animal) fur.

In many cases, the interpretation of a prenominal adjective is ambiguous. For example, a phrase like *tall basketball player* readily allows at least two interpretations: (a) ‘a person who is tall for a basketball player’ and (b) ‘a person who is both tall and a basketball player.’ These are adjectives which are intersective, but context-sensitive—they require a context to determine the scale (of height, e.g.) with which the adjective is interpreted (Kamp, 1975; Siegel, 1976). Normally, this scale is determined relative to the context

supplied by the head noun, as happens with the (a) reading, but it can also be influenced by some other context, such as the larger discourse context, allowing the (b) reading of *tall basketball player* (Kamp & Partee, 1995; Partee, 1995; Sedivy, Chambers, Tanenhaus & Carlson, 1997). Thus, *tall basketball player* always picks out the intersection of $tall_{for\ x}$ and *basketball player*, but sometimes it is picking out the intersection of $tall_{for\ basketball\ players}$ and *basketball player* (giving the (a) reading), and sometimes it is picking out the intersection of $tall_{for\ people}$ and *basketball player* (giving the (b) reading).

All subsective adjectives and adjectives which are neither intersective nor subsective (such as *alleged*) must be interpreted with respect to the head noun, and most context-sensitive intersective adjectives seem to be interpreted on the basis of context supplied by the head noun. Therefore, it is likely that adjective-noun sequences present a case of a temporary processing delay in the full interpretation of the adjective.[6] That is, complete interpretation of the adjective does not occur as soon as it is identified, but is delayed until after its head noun has been interpreted. And, the interpretation of context-sensitive adjectives provides a case of semantic ambiguity (of the context used to determine the scale) which presumably is not dependent on a syntactic difference. Thus, the interpretation of adjective-noun sequences with context-sensitive intersective adjectives potentially provides an appropriate first test of the Interpretive Domain Hypothesis.

Recall that the Interpretive Domain Hypothesis predicts that semantic/pragmatic processing which would otherwise happen later in time will be initiated when an intonational phrase boundary is encountered by the processor. Thus, the interpretation of a context-sensitive intersective adjective will take place immediately, according to the Interpretive Domain Hypothesis, when an intonational phrase boundary intervenes between the adjective and head noun. Because the context of the head noun would be unavailable in

[6] For example, Kamp and Partee (1995) have proposed that prenominal adjectives are normally interpreted with respect to the head noun as described by the Head Primacy Principle: *In a modifier-head structure, the head is interpreted relative to the context of the whole constituent, and the modifier is interpreted relative to the local context created from the former context by the interpretation of the head.*

such a situation, the context would have to be supplied by the discourse situation. I will refer to this as the "independent" interpretation and will refer to the interpretation which occurs when the context is supplied by the head noun as the "head-dependent" interpretation. The Interpretive Domain Hypothesis predicts that a higher percentage of "independent" interpretations will occur when an intonational phrase boundary intervenes between a context-sensitive intersective adjective and its head than when no intonational phrase boundary intervenes. Experiment 3 tested this prediction.

4.3.1 Materials

Experiment 3 tested the conditions illustrated in (4.6) with an end-of-sentence comprehension task in which subjects answered disambiguating questions about the interpretation of the adjective after each item (as described below). Each sentence either contained an IPh boundary before the final noun, as in (4.6c) and (4.6d), or no such boundary, as in (4.6a) and (4.6b), and a context-sensitive intersective adjective which either described an atypical characteristic of the head noun, as in (4.6a) and (4.6c), or a typical characteristic, as in (4.6b) and (4.6d).

(4.6) a. Now that he has graduated, Chuck wants to buy an inexpensive Porsche.
b. Now that he has graduated, Chuck wants to buy an expensive Porsche.
c. Now that he has graduated, Chuck wants to buy an inexpensive)$_{IPh}$ Porsche.
d. Now that he has graduated, Chuck wants to buy an expensive)$_{IPh}$ Porsche.

The intonational phrase boundary was marked by a low phrase accent and a high boundary tone (a "continuation rise") to signal non-finality of the utterance and prevent premature end-of-sentence judgments. To ensure that both the adjective and the noun were focused and that the prosodic contour was grammatical both with and without the intonational phrase boundary after the adjective, the adjective received an H* accent and the noun received an L+H* accent in all conditions.[7] The critical tones are transcribed in

[7] This contour created potential phonetic ambiguity in conditions (a) and (b) between a prosodic representation with a low phonological phrase boundary between the adjective and the noun and a representation with no prosodic boundary between the adjective and noun. See the discussion section and Beckman (1996) for further discussion of this ambiguity.

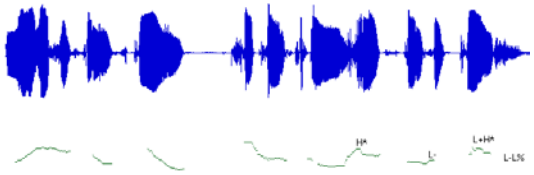
Example (4.7), and sample waveforms and pitch tracks are given in Figure 4.1. A full list of tonally-transcribed experimental items appears in Appendix C.

(4.7) a. Now that he's finished school, Chuck wants to buy an H* (L-) L+H* L-L% inexpensive Porsche.

b. Now that he's finished school, Chuck wants to buy an H* (L-) L+H* L-L% expensive Porsche.

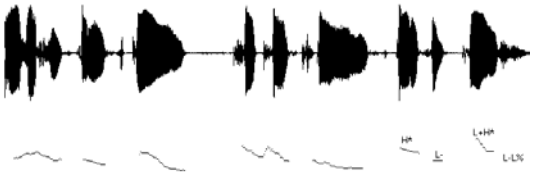
c. Now that he's finished school, Chuck wants to buy an H* L-H% L+H* L-L% inexpensive)$_{IPh}$ Porsche.

d. Now that he's finished school, Chuck wants to buy an H* L-H% L+H* L-L% expensive)$_{IPh}$ Porsche.



a. Now that he's finished school, Chuck wants to buy an inexpensive L-)$_{PPh}$ ((Porsche L-L%))$_{IPh}$

c. Now that he's finished school, Chuck wants to buy an inexpensive L-H%))$_{IPh}$ ((Porsche L-L%))$_{IPh}$



b. Now that he's finished school, Chuck wants to buy an expensive L-)$_{PPh}$ ((Porsche L-L%))$_{IPh}$

d. Now that he's finished school, Chuck wants to buy an expensive L-H%))$_{IPh}$ ((Porsche L-L%))$_{IPh}$

Figure 4.1. Sample Pitch Tracks for Experimental Items, Experiment 3.

The materials were produced by a ToBI-trained native speaker of English in a sound-attenuated booth, recorded onto tape, and digitized at 11.025 kHz. The prosody of the materials was carefully controlled; phonetic measurements verified that it was produced as intended and as transcribed in Appendix C.

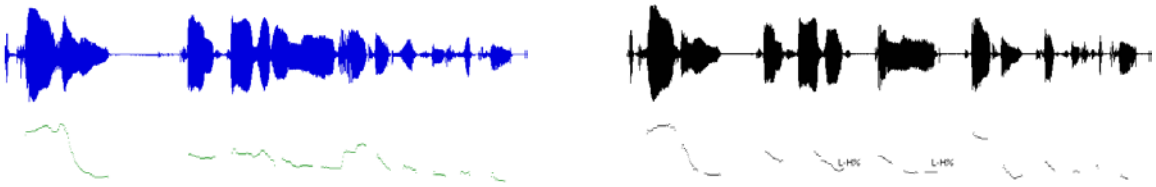
Recall that the (b) and (d) versions of each item contained an adjective which described a typical characteristic of the head noun, while (a) and (c) contained adjectives describing

an atypical characteristic. Subjects were required to choose a paraphrase for each sentence that reflected either the independent interpretation (e.g., 'expensive and a Porsche') or the head-dependent interpretation (e.g., 'expensive for a Porsche'), so the atypical/typical adjective factor allowed a separate, non-prosodic test of the subjects' ability to make this somewhat subtle distinction. Porsches are typically considered expensive, not inexpensive. Therefore, the independent interpretation of the adjective in *expensive Porsche* should be highly plausible, but the independent interpretation of the adjective in *inexpensive Porsche* should be implausible. Thus, if subjects are able to discriminate between the paraphrases appropriately, they should show more independent responses for *expensive Porsche* than for *inexpensive Porsche*. The atypical/typical adjective factor also guarded against the possibility of a prosodic effect being obscured by a very strong pragmatic bias for either the independent interpretation or the head-dependent interpretation with the experimental materials.

As an additional test of subjects' ability to make the independent/head-dependent distinction, a set of eight control sentences was created. A sample item is given in (4.8). The full list of control items is provided in Appendix C.

(4.8) a. Occasionally, John's favorite unmarried uncle takes him to the park.
b. Occasionally, John's favorite)$_{IPh}$ (unmarried)$_{IPh}$ uncle takes him to the park.

In the control set, each sentence contained an adjective-adjective-noun string. In condition (a), the string formed a single intonational phrase. In condition (b), the second adjective was set off as a separate intonational phrase, as indicated in (4.8), which was produced with a reduced pitch range. This low F_0 was matched in the (a) versions by the placement of a low pitch accent on the second adjective, thus keeping the two contours as similar as possible except for the presence or absence of the medial IPh boundaries. Sample waveforms and pitch tracks showing this prosody are given in Figure 4.2.



a. Occasionally, John's favorite unmarried uncle takes him to the park.

b. Occasionally, John's favorite)$_{IPh}$ (unmarried)$_{IPh}$ (uncle takes him to the park.

Figure 4.2. Sample Pitch Tracks for Control Items, Experiment 3.

When the second adjective of the control set was not set off in a separate intonational phrase, it was expected that subjects would tend to interpret the first adjective with respect to the remainder of the phrase (e.g., “favorite of the unmarried uncles”). However, when the second adjective was set off, it was expected that subjects would interpret it as a parenthetical, and therefore interpret the first adjective only with respect to the noun head, interpreting the second adjective as a non-restrictive modifier (e.g., “favorite of the uncles, and also unmarried”). Thus, the control set provided materials for which the subjects would make a judgment about the interpretation of modifiers similar to the judgment required for the experimental set of materials. However, in the control case the interpretive judgment stemmed from the effect of prosodic phrasing on a syntactic distinction (i.e., the interpretive effect was mediated by the parenthetical versus non-parenthetical distinction), while in the experimental case the judgment was hypothesized to depend on the direct effect of the prosodic phrasing on semantic processing. If prosodic phrasing can only affect syntactic decisions, then a prosodic effect should have appeared with the control materials but not with the experimental materials. If prosodic phrasing can affect semantic decisions in addition to syntactic ones, as predicted by the Interpretive Domain Hypothesis, then a prosodic effect should have appeared with both the control materials and the experimental materials.

4.3.2 Task

The items were presented following a Latin-square design, so that each subject heard only one condition for each item, with the condition rotating through the items. Each subject received a different randomization of the sixteen experimental items, eight control items, and 56 other sentences which had various prosodic and syntactic structures. Subjects heard the sentences over speakers in a sound-attenuated booth. At the end of each sentence, the subject marked the sentence as "good" (grammatical) or "bad" (ungrammatical) by pressing one of two levers as soon after the end of the sentence as possible. Following each experimental and control sentence and approximately half of the filler sentences a question such as the one in (4.9) appeared on a computer monitor positioned in front of the subject. The subject selected one of two responses by pressing the corresponding lever.

(4.9) Sample disambiguating question for Experiment 3:

```
                Chuck wants to buy something which is:

inexpensive for a Porsche                    inexpensive and a Porsche
```

4.3.3 Subjects

Forty-eight undergraduates at the University of Massachusetts participated in exchange for course credit. All subjects reported that they were native speakers of English with normal hearing. Two additional subjects participated in the experiment but were excluded from the final analysis because of a mean reaction time for all 80 sentences which was greater than three standard deviations from the mean for all subjects.

4.3.4 Results

The results of Experiment 3 are presented in Tables (4.1) and (4.2) as the percentages of independent responses and Table (4.3) as the reaction times to make the grammaticality judgment. Table (4.1) shows that the subjects were able to make the interpretive judgment for the control items, in which a syntactic difference mediated between the prosodic

difference and the semantic difference. As predicted, the percentage of independent responses is higher for the condition in which the second adjective was set off as a separate prosodic phrase than in the condition in which it was not ($F_{(1,47)}$=5.3, p<.03; $F_{(1,7)}$=3.9, p<.09).[8]

Table 4.1. Percentages of "Independent" Responses for Control Items, Exp. 3.

No Medial IPh Boundaries	Medial IPh Boundaries
32.3% (condition a)	44.3% (condition b)

The results for the experimental materials, presented in Table (4.2), show more independent interpretations for the items with a pragmatic bias for an independent interpretation (b, d) than for items with a head-dependent bias (a, c) ($F_{(1,47)}$=12.8, p<.01; $F_{(1,15)}$=8.2, p<.01), demonstrating that subjects were capable of making the appropriate interpretive judgment for the pragmatic bias of the materials and further showing that they were able to make the interpretive distinction. More importantly, and as predicted by the Interpretive Domain Hypothesis, the presence of an intonational phrase boundary significantly increased the number of independent interpretations. The percentage of independent interpretations is significantly higher for the conditions containing a medial intonational phrase boundary (c, d) than for the conditions without the intervening intonational phrase boundary (a, b) ($F_{(1,47)}$=7.3, p<.01; $F_{(1,15)}$=4.1, p<.063).

[8] The response times for making the grammaticality judgment did not differ statistically for the two conditions of the control sentences. The percentage of items judged as grammatical was 84.9% for the no-boundary condition and 70.8% for the condition with prosodic boundaries. This difference was significant by subjects ($F_{(1,47)}$=10.2, p<.01) but not by items ($F_{(1,7)}$=3.3, p<.12) and presumably reflects either somewhat greater processing difficulty for the parenthetical structure and non-restrictive interpretation or a bias against judging the less formal parenthetical structure as grammatical.

Table 4.2. Percentages of "Independent" Responses for Experimental Items, Exp. 3.

	Prosody:	
Pragmatic Bias:	No Medial IPh Boundary	Medial IPh Boundary
Head-dependent Interpretation	55.7% (condition a)	63.5% (condition c)
Independent Interpretation	68.2% (condition b)	75.5% (condition d)

Additional support for the Interpretive Domain Hypothesis comes from the response times for the grammaticality judgment, given in Table (4.3). According to the IDH, when a prosodic boundary was present after the adjective subjects were forced to construct an independent interpretation of the adjective. When the adjective was a typical description of the noun, such as *expensive* (for *Porsche*), the combination of the independently interpreted adjective and the noun should have formed a highly plausible description, and thus should have been easy to interpret. However, when the adjective was an atypical description of the noun, such as *inexpensive* (for *Porsche*), the combination of the independently interpreted adjective and the noun is somewhat anomalous—it describes something which is both inexpensive (according to a general scale of expense) and a Porsche, not something which is relatively inexpensive for a Porsche. Therefore, this condition should have been more difficult to interpret and should have taken more time to process. As expected if the IDH is correct, subjects took longer to judge sentences as grammatical in condition (c) than in the other conditions, resulting in a significant interaction by items and a marginally significant interaction by subjects ($F_{(1,47)}=3.5$, $p<.07$; $F_{(1,15)}=5.4$, $p<.04$).

The results of Experiment 3 show that the interpretation of context-sensitive intersective adjectives is affected both by the plausibility of the adjective-noun combination and the prosody of the utterance. Assuming that the experimental sentences are unambiguous

Table 4.3. Grammaticality Judgment Response Times for Experimental Items, Exp. 3.

	Prosody:	
Pragmatic Bias:	No Medial IPh Boundary	Medial IPh Boundary
Head-dependent Interpretation	1176 ms (condition a)	1402 ms (condition c)
Independent Interpretation	1186 ms (condition b)	1230 ms (condition d)

syntactically, the prosodic difference could not have caused the effect on semantic interpretation only indirectly, by first affecting syntactic processing. Thus it appears that the prosody directly affected semantic processing. Specifically, assuming that the difference in interpretation is based on the context used in interpreting the adjective, the results suggest that the intonational phrase boundary interfered with the use of the context associated with the head noun. Hence, the results strongly support the Interpretive Domain Hypothesis, which predicted that when an intonational phrase boundary was present, the adjective would be interpreted immediately, before the interpretation of the head noun had taken place, and that therefore the context associated with the head noun would not be available to guide the interpretation of the adjective.

4.3.5 Discussion

There have been many psycholinguistic studies of prosody which have demonstrated effects of intonational phrasing on syntactic decisions, as well as many studies which have shown effects of pitch accent placement and kind on the interpretation of focus or given versus new information, but no other comprehension study, to my knowledge, has demonstrated an effect of prosodic phrasing on interpretive decisions such as these. Thus, Experiment 3 provides the first experimental evidence that intonational phrasing affects semantic/pragmatic processing as well as syntactic processing. Although there are many implications of this finding, I will delay discussion of them until the general discussion section, after I have reported the results of Experiment 4.

The results of Experiment 3 also provide preliminary evidence which specifically supports the Interpretive Domain Hypothesis. That is, the results do not merely show that there is some effect of intonational phrasing on non-syntactic interpretation; they support an explicit hypothesis of how to characterize the role of intonational phrasing in higher-level processing. However, the prosody used in Experiment 3 did not unambiguously contrast intonational phrase boundaries with phonological phrase boundaries, as the prosody in the no-intonational-phrase-boundary conditions was potentially ambiguous between a contour with a phonological phrase boundary and one without a phonological phrase boundary. Thus, it is possible that the interpretive effect demonstrated in Experiment 3 would also be seen for materials containing a clearly marked phonological phrase boundary in the crucial location. Further, it is in principle possible that the presence of the medial intonational boundary introduces a syntactic effect, such as extraposition of the noun or a different internal structure of the AdjP or NP, in addition to the semantic difference in the assignment of context.[9] It is also possible that the conditions with the medial intonational boundary were simply considered to be less natural than the no-boundary conditions, and that the difference in interpretation is an artifact of this difference in naturalness.[10]

---

[9] Note that if the presence of an IPh boundary causes extraposition of the noun, we must still account for why the parser is willing to build extra structure in these cases and how the head-dependent interpretation is blocked by extraposition. Bernstein (1993) has proposed distinct syntactic structures for Romance phrases equivalent to 'poor man' = 'pitiable man' and 'poor man' = 'impoverished man.' While this work establishes that Romance adjective-noun strings can vary in their syntactic structure, it is not clear that this analysis must hold for English. It is also not clear how such an analysis would be expanded to account for the pitiable/impoverished distinction as well as the difference in assignment of context discussed here.

[10] The percentages of "good" judgments for the experimental items were: condition (a), 95.3%; (b), 93.8%; (c), 83.8%; (d) 85.4%. While the conditions containing medial IPh boundaries received a significantly lower percentage of "good" judgments than the conditions without medial IPh boundaries, the percentages were roughly the same as the percentage of "good" judgments for the control sentences without medial IPh boundaries and higher than the percentage for the control sentences with medial IPh boundaries. Note that the experimental items showed an overall bias toward "independent" responses, so if there is an effect of lowered naturalness for the conditions with medial IPh boundaries on the interpretation of the adjective, it is one which strengthens the general bias in the experiment, and not one of only choosing an unusual interpretation for sentences with unusual prosody (cf. Wales & Toner, 1979).

Given the lack of certainty about how the prosodic ambiguity of conditions (a) and (b) is resolved and the possibility of alternative explanations based on syntactic structure or naturalness, Experiment 3 provides only preliminary support for the Interpretive Domain Hypothesis. Stronger support would come from a case which required fewer assumptions about how the materials are interpreted and which clearly contrasted phonological phrase boundaries with intonational phrase boundaries. Experiment 4 was designed to test such a case.

4.4 Experiment 4: Intonational versus Phonological Phrases and Reanalysis

As described above, there is clear psycholinguistic evidence about how lexically ambiguous nouns such as *ruler* or *ball* are processed: all meanings are activated initially, and one meaning is selected within a few hundred milliseconds, either on the basis of disambiguating context, if available, or the frequency and recency of usage of the different meanings. For sentences like (4.10a) and (4.10b), with the lexically ambiguous word *glasses* and no biasing information in the initial clause, it is expected that the most frequent meaning of *glasses* will be selected before the disambiguating information in the second clause is encountered. Assuming that the dominant meaning of *glasses* is generally "spectacles," the second clause of (4.10a) should generally be processed easily, as the information in it would be consistent with the selected meaning. In (b), however, the information in the second clause would frequently conflict with the selected meaning of *glasses*. To resolve this conflict the processor would be forced to reanalyze the interpretation of *glasses* established earlier, resulting in more difficult processing and longer comprehension times for the sentence.

(4.10) a. Although the glasses were ugly, Stacey wore them anyway.
b. Although the glasses were ugly, they held a lot of juice.

Experiment 4 tested sentences like (4.10a) and (4.10b) with either a phonological phrase boundary or an intonational phrase boundary at the medial clause boundary. The Interpretive Domain Hypothesis does not predict significant differences in processing

difficulty for a sentence like (4.10a) with a medial intonational phrase boundary versus a rendition with a medial phonological phrase boundary, since regardless of how much interpretation has taken place at that prosodic boundary, the remainder of the sentence will generally be consistent with that interpretation and should therefore be easy to process. However, in sentence (b), reanalysis is often required. Assuming that reanalysis is more difficult for more deeply processed material (e.g., because more inferences about the material have been made), reanalysis times should be longer for material contained in a preceding intonational phrase than for material contained in a preceding phonological phrase, according to the Interpretive Domain Hypothesis, because the material in the preceding intonational phrase has been more deeply processed. Therefore, it should take longer to reanalyze lexical ambiguity when the disambiguating material occurs in a subsequent intonational phrase than when it occurs in a subsequent phonological phrase (and the same intonational phrase) as the ambiguous word. Processing times should be longer for sentences like (4.10b) produced with a medial intonational phrase boundary than for the same sentence produced with a medial phonological phrase boundary.

Sentences like (4.10a) and (b) produced with phonological versus intonational phrase boundaries provide an ideal test of the Interpretive Domain Hypothesis. The processing characteristics of the semantic ambiguity are well-known, the syntactic structure of the materials is uncontroversial, both prosodic contours are known to be well-formed, and it is exceedingly unlikely that the prosodic difference being employed interacts with some other phonological, syntactic, or semantic decision for these sentences in a way that could confound the results.[11] Further, the Interpretive Domain Hypothesis makes clear predictions about how the processing of the sentences should differ: processing should be easy for the sentences which instantiate the preferred interpretation of the ambiguous word under both prosodic conditions, but processing should be significantly harder for the

[11] It is possible that the greater time associated with intonational phrase boundaries (in general) could affect the amount of time available for contextual processing and therefore confound the results. This issue is addressed below.

sentence instantiating the dispreferred interpretation of the ambiguous word when an intonational phrase boundary occurs at the clause boundary than when a phonological phrase boundary occurs there.

4.4.1 Lexical Pretest

To create a set of materials containing lexically ambiguous words strongly biased toward one meaning, a superset of the materials used in the main experiment was normed with a written questionnaire. Thirty-two subordinate clauses containing a lexically ambiguous word in a neutral context were created (e.g., *Although the glasses were ugly*...). These sentence fragments were randomized with 44 other subordinate clause sentence fragments and presented to twenty undergraduates at the University of Massachusetts. For each fragment, the subject completed the sentence by writing down the first completion that occurred to him or her. Then the subject answered a question about the fragment (e.g., *Why were the glasses ugly?)* with the first answer that came to mind. These responses were coded for lexical bias; disambiguation occurred in over 92% of the responses. The sixteen most strongly biased items were chosen for use in Experiment 4. A complete list of tonally-transcribed experimental items and their lexical biases appears in Appendix D.

4.4.2 Makes-Sense Pretest

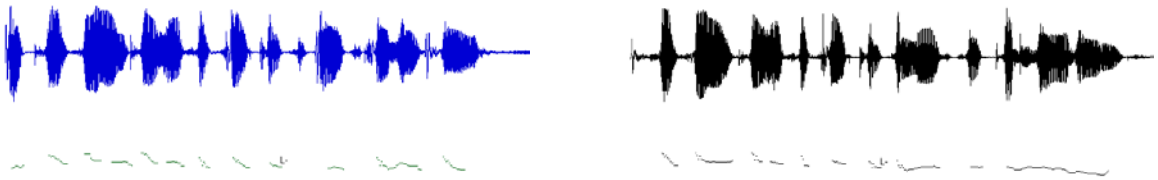
The materials were also pretested to ensure that the full sentence made sense with both the continuation instantiating the preferred meaning of the ambiguous word and the continuation instantiating the dispreferred meaning. The sixteen experimental items were randomized with 32 other sentences, of which half were expected to make sense and half were expected to not make sense (e.g., *While the lawyer was friendly, they contained too much sugar*). These 48 sentences were presented in a Latin-square design, so subjects heard only one continuation for each ambiguous word, in a written questionnaire. The subjects rated each sentence on a scale of 1 (“makes perfect sense”) to 5 (“makes no sense”). The results showed that both continuations were judged as making sense, with an

average rating of 1.18 for the preferred-meaning sentences and 2.08 for the dispreferred-meaning sentences, but 4.56 for the sentences predicted to not make sense.

4.4.3 Materials

Sixteen sets of experimental materials similar to the sample item given in (4.11) were created. Each sentence contained a strongly-biased ambiguous word in an initial subordinate clause and a continuation appropriate for either the preferred (a,b) or dispreferred (c,d) meaning of the ambiguous word. The medial clause boundary coincided with either a phonological phrase boundary marked by a low phrase accent (L-) or an intonational phrase boundary marked by a low phrase accent and a high boundary tone (H%), as can be seen in the sample pitch tracks in Figure 4.3.

(4.11) a. ((Although the glasses were ugly L-)$_{PPh}$ (Stacey wore them anyway))
b. ((Although the glasses were ugly L- H%))$_{IPh}$ ((Stacey wore them anyway))
c. ((Although the glasses were ugly L-)$_{PPh}$ (they held a lot of juice))
d. ((Although the glasses were ugly L- H%))$_{IPh}$ ((they held a lot of juice))



a. Because the anchor was effective L-)$_{PPh}$ the boat stayed in place.

c. Because the anchor was effective L-)$_{PPh}$ the news got good ratings.

b. Because the anchor was effective L-H%))$_{IPh}$ the boat stayed in place.

c. Because the anchor was effective L-H%)$_{IPh}$ the news got good ratings.

Figure 4.3 Sample Pitch Tracks for Experiment 4.

The predictions for the experiment rest on the assumption that a single meaning for the ambiguous word would be chosen before the disambiguating material was encountered, so

that reanalysis would be required in conditions (c) and (d). To guarantee that lexical selection and integration had occurred prior to disambiguation in all sentences, the region from the ambiguous word to the disambiguating material contained sufficient material to be at least 500 milliseconds for each sentence, averaging over 1000 milliseconds. This length also served to minimize the effects of inherent length differences between intonational phrase boundaries and phonological phrase boundaries by eliminating the possibility that because the intonational phrase boundaries allow more processing time, lexical selection could occur for the conditions with intonational phrase boundaries but not for the conditions with phonological phrase boundaries.

The materials were produced by a ToBI-trained native speaker of English in a sound-attenuated booth, recorded onto tape, and digitized at 11.025 kHz. The prosody of the materials was carefully controlled; phonetic analyses verified that it was produced as intended and as transcribed in Appendix D. Crucially, the duration of the silent interval associated with the intonational phrase boundaries was trimmed to 40 milliseconds to minimize durational differences between the two boundary levels; this duration was judged by the experimenter to be the shortest duration necessary to provide natural-sounding intonational phrase boundaries for the sentences.[12] To ensure that potential differences in the production of the ambiguous word or other aspects of the first intonational phrase would not affect the interpretation, the first intonational phrase from the condition with the dispreferred continuation (condition d) was digitally spliced to the second intonational phrase for the preferred continuation (condition b). Thus, the first intonational phrase of conditions (b) and (d) was identical, so that if greater processing difficulty was found in condition (d), it could not be attributed simply to processing difficulty associated with the production of the material in the first intonational phrase. As condition (d) is the condition

[12] Pausing appears to show high variability both within and across speakers (Cooper & Paccia-Cooper, 1980). Thus there seems to be a wide range of acceptable pause durations for intonational phrase boundaries. Note that while the difference in pause duration between the two levels of boundaries was minimized, the materials still presented clear evidence of the boundary distinction through the difference in edge tones.

predicted to be more difficult to process by the Interpretive Domain Hypothesis, splicing from (d) to (b) ensured that if any difficulty in processing or well-formedness arose from the splicing, it would work against the predictions of the hypothesis.[13]

4.4.4 Task

The items were presented following a Latin-square design, so that each subject heard only one condition for each item, with the condition rotating through the items. Each subject received a different randomization of the sixteen experimental sentences and 52 other sentences which had various prosodic and syntactic structures. Subjects heard the sentences over speakers in a sound-attenuated booth. At the end of each sentence, the subject judged whether the sentence "made sense" or not and encoded the judgment by pressing one of two levers as soon after the end of the sentence as possible. A simple comprehension question followed approximately half of the sentences to ensure that subjects were properly attending to the task.

4.4.5 Subjects

Fifty-two undergraduates at the University of Massachusetts participated in exchange for course credit. All subjects reported that they were native speakers of English with normal hearing. An additional subject participated in the experiment but was excluded from the final analysis because of a mean comprehension time for the 68 total sentences which was greater than three standard deviations from the mean for all subjects.

4.4.6 Results

The results are presented in Table (4.4), which shows comprehension times from the end of the sentence for "makes sense" responses.[14] These times were significantly longer

[13] Similar splicing was not done for conditions (c) and (a) as it was judged that coarticulation effects in the materials would create unacceptable splicing artifacts.

[14] Comprehension times from the end of the sentence which were less than 100 ms or greater than 4000 ms and times which were more than three standard deviations from the mean were excluded from analysis. This excluded less than 7% of the data; the majority of these exclusions resulted from premature judgments at the medial clause boundary. Zeros in the data were replaced by using the average of the experiment-wise individual subject or item means. One of the sixteen items was removed from the analysis because of excessively low

for conditions (c) and (d) than for (a) and (b), as expected given that (c) and (d) instantiate the dispreferred reading ($F_{(1,14)}$=10.6, p<.01; $F_{(1,51)}$=15.4, p<.01). Also as expected, the main effect of prosody was not significant, that is, there was no overall effect of an intonational phrase boundary resulting in longer processing times than a phonological phrase boundary.

Table 4.4. Comprehension Times for "Makes Sense" Judgments, Experiment 4.

	Medial Prosodic Boundary:	
Lexical Bias:	Phonological Phrase	Intonational Phrase
Preferred Meaning	730 ms (condition a)	708 ms (condition b)
Dispreferred Meaning	824 ms (condition c)	946 ms (condition d)

Crucially, and as predicted by the Interpretive Domain Hypothesis, the interaction of prosody and lexical preference was significant by items, although it did not reach significance by subjects ($F_{(1,14)}$=5.7, p<.04; $F_{(1,51)}$=2.5, p<.13). Comprehension times were numerically shorter for preferred-meaning sentences with an intonational phrase boundary (condition b) than with a phonological phrase boundary (condition a), but longer for dispreferred-meaning sentences with an intonational phrase boundary (condition d) than with a phonological phrase boundary (condition c); a planned comparison of means for conditions (c) and (d) showed that this difference was significant by items and marginally significant by subjects ($F_{(1,14)}$=5.7, p<.04; $F_{(1,51)}$=3.6, p<.07).

The Interpretive Domain Hypothesis predicted that processing would be easy for the two conditions instantiating the preferred meaning of the lexically ambiguous item, and that the prosodic difference would have no effect on the comprehension times for these two conditions. It also predicted that there would be an effect of prosody in the conditions instantiating the dispreferred meaning, with longer comprehension times in the condition

"makes sense" judgments (apparently due to a dialectal difference in pronunciation for a

containing a medial intonational phrase boundary. This is exactly the pattern of results seen for Experiment 4; therefore, the results support the Interpretive Domain Hypothesis.

4.4.7 Discussion

The results of Experiment 4 show a processing difference between sentences containing medial intonational phrase boundaries and those containing only phonological phrase boundaries medially. This is an important finding for psycholinguistic research, since previous sentence processing research has generally failed to differentiate these two levels of prosodic phrasing and has implicitly suggested that they should have similar effects on processing. Further, when research has discriminated between the two, it has shown that phonological and intonational phrase boundaries can have very similar effects on syntactic processing. Experiment 4 shows that the apparent similarity of behavior of the two levels of prosodic phrasing for syntactic processing does not hold for higher levels of processing.

It may seem that the data do not discriminate between an acoustically-based account of the processing difference between phonological and intonational phrase boundaries and an account that postulates a non-gradient effect based on distinct phonological or psychological categories. However, several arguments eliminate an account of the results as being merely an artifact of durational differences between the two kinds of boundaries. First, the durational differences between productions with medial phonological phrase boundaries and productions with medial intonational phrase boundaries were minimized through digital editing to be as similar as possible without altering the naturalness of the utterance. Second, the region between the ambiguous word and the disambiguating material was quite long, ensuring that even in the phonological phrase boundary condition the ambiguous region allowed ample time for lexical selection and integration of the ambiguous word. Thus, the relatively small difference in duration between prosodic conditions accounted for only a small portion of the ambiguous region. And finally, there was no correlation between the duration of the ambiguous region and the comprehension

---

key word in the sentence).

times for the conditions expected to require reanalysis ($r^2_{Dispref.}$ = .03, p>.3). Thus, the amount of processing time available in the critical region cannot account for the observed pattern of results.

It is also possible, in principle, that the processing difference could have resulted from pragmatic effects of the end contour of the initial prosodic phrase instead of the difference in the level of prosodic phrasing, since the prosodic conditions differed both in the strength of the prosodic boundary and in intonational contour. The low phrase accent of the phonological phrase condition creates a contour with a final fall, while the low phrase accent plus high boundary tone of the intonational phrase condition provides a fall-rise contour. It has been proposed that the shape of the end contour signals pragmatic distinctions such as whether the following prosodic phrase is, in some sense, a continuation of the preceding phrase (e.g., Pierrehumbert & Hirschberg, 1990; Bartels, 1997). Hence, it is possible that the fall versus fall-rise distinction, and not the phonological versus intonational phrase distinction, could cause a processing difference. However, such an account would have predicted greater processing difficulty for the phonological phrase condition, not for the intonational phrase condition. A fall in intonation is generally taken to mark less connection to a following phrase, while a fall-rise pattern is the typical "continuation" contour. If reanalysis is affected by a pragmatic distinction encoded in the end contour, it should be expected that reanalysis would be easier for the fall-rise continuation contour of the intonational phrase boundary condition and harder for the falling contour of the phonological phrase boundary condition. Therefore, the possible pragmatic effects of the end contour do not account for the experimental results. Indeed, as these potential pragmatic effects work against the predictions of the Interpretive Domain Hypothesis, the processing effect of an intonational phrase boundary may actually be stronger than suggested by the results of Experiment 4.

In principle, it might also be possible to account of the results of Experiment 4 by appeal to well-formedness constraints governing the relation between the prosodic contour

and the syntactic structure of the experimental sentences or those governing the relation between the prosodic structure and the semantic structure of the materials. However, the similarity in response for the two conditions instantiating the preferred meaning of the ambiguous word suggests that both a contour with a phonological phrase boundary at the medial clause boundary and a contour with an intonational phrase boundary at the medial clause boundary are well-formed contours, as do the results of the Speer *et al.* work with phonological phrase boundaries at medial clause boundaries discussed in Chapter 1. Further, there was clearly no syntactic difference within the pairs of sentences in Experiment 4, so prosodic effects on syntax could not have be a factor in explaining the results. Moreover, there is no plausible grammatical connection between the prosody used in the experiment and the disambiguation of lexically ambiguous words. Thus, it is extremely unlikely that the effect of the medial intonational boundary on the comprehension time for the condition instantiating the dispreferred meaning of the ambiguous item was due to grammatical considerations.

Because the demonstrated processing difference between the two prosodic boundaries in Experiment 4 cannot be readily attributed to durational differences between the conditions, or to pragmatic effects of the end contour of the initial prosodic phrase, it is reasonable to believe that it is the distinct phonological and psychological status of phonological and intonational phrase boundaries that results in their differing effects on processing. Further, since the results cannot be accounted for by effects of prosody on syntactic attachment decisions or on syntactic or semantic well-formedness, it is reasonable to believe that it is the direct effect (or lack of effect) of the prosodic categories on semantic/pragmatic decisions of the processor that causes the processing difference. Thus, the results of Experiment 4 support three claims: (1) phonological phrases and intonational phrases function as separate categories for the language processor and cause separable processing effects, (2) intonational phrase boundaries affect semantic/pragmatic processing decisions, and therefore prosodic phrasing can affect processing decisions beyond syntactic

attachment decisions, and (3) the effect of intonational phrase boundaries on processing decisions is as described by the Interpretive Domain Hypothesis, such that intonational phrase boundaries, but not phonological phrase boundaries, define a point at which further semantic/pragmatic evaluation and integration occurs.

4.5 General Discussion

In Chapters 1 and 2, I argued that the syntactic component of the parser has access to the prosodic representation, on the basis of evidence that phonological phrase boundaries, as well as intonational phrase boundaries, affect syntactic parsing decisions. The research presented in this chapter provides evidence that the information contained in the prosodic representation must be available to the semantic/pragmatic component (or components) of the language processing system as well. Thus, there is evidence that the prosodic representation is influential across the phonological, syntactic, and semantic/pragmatic levels of representation, structuring the input at each of these levels and delimiting the relevant domain for the processing decisions made at each level.

In particular, Experiments 3 and 4 provide strong evidence for prosodically-defined interpretive processing domains by showing that further interpretive processing takes place at intonational phrase boundaries than at phonological phrase boundaries. Many other processing units have been proposed in the psycholinguistic literature, such as the (roughly) six-word packages of the Sausage Machine (Frazier & Fodor, 1978) and of Frazier (1982), or the syntactic clause (e.g., Fodor, Bever & Garrett, 1974; Flores d'Arcais, 1978); see Levelt (1989) for an extensive list of proposed processing units. However, since the crucial pairs of conditions were identical except for the prosodic phrasing, it cannot be the case that the interpretive effects demonstrated in Experiments 3 and 4 can be reduced to the effect of any of these other processing units. For example, although it is likely that further interpretive decisions also take place at clause boundaries, the results of Experiment 3 show that more interpretive decisions are initiated when an intonational phrase boundary is present than when it is not at a location which is not even a

potential clause boundary. Hence, in the absence of a clause boundary, an intonational phrase boundary marks the edge of an interpretive domain. The results of Experiment 4 show that an intonational phrase boundary also leads to further interpretive processing in the presence of a clause boundary. All of the conditions of Experiment 4 contained a clause boundary at the location of the prosodic boundary, but processing differed between the two conditions requiring reanalysis, showing that an intonational phrase boundary can result in further interpretive processing even when it falls at a clearly marked clause boundary. Thus, neither the results of Experiment 3 nor the results of Experiment 4 can be explained by a model in which clause boundaries are the sole catalysts to further interpretation, and similar arguments can be made about the inability of other proposed processing units in accounting for these results.

I have further claimed that the processor's sensitivity to prosodic phrasing at the semantic/pragmatic level is compatible with a modular system of sentence processing. Frazier (1990) has proposed that a module can only operate on elements that are part of its computational vocabulary, but that there is overlap in the vocabularies of the various modules. We need only accept that intonational phrase boundaries are part of the vocabulary of the semantic/pragmatic component of the processor (as well as the phonological and syntactic components) to accept that the prosodic effects described by the Interpretive Domain Hypothesis are compatible with this view of modularity. Given the evidence reported above that grammatical constraints on semantic and pragmatic interpretation make reference to different kinds of edge tones, we should expect that intonational phrase boundaries are part of the relevant vocabulary and could affect semantic and pragmatic processing. The results of Experiments 3 and 4 can thus be taken as further support for the grammatically-based, tightly-constrained kind of interaction that is defined by Frazier's modular processing system.

There is also reason to believe that a system in which intonational phrase boundaries cue the processor to initiate further interpretive decisions could be very helpful. Recall that

the Interpretive Domain Hypothesis does not require that interpretive operations be delayed until the end of an intonational phrase, rather, it requires that interpretive operations that might otherwise be delayed take place when an intonational phrase boundary is detected. In effect, an intonational phrase boundary signals that upcoming linguistic context is irrelevant to current interpretive decisions, allowing immediate, full interpretation instead of delayed interpretation in cases such as the processing of context-sensitive intersective adjectives, for which upcoming context might otherwise be expected to be important or even necessary for full interpretation. Presumably, reducing processing delay for cases such as these would reduce the processing load and provide greater processing efficiency, at least for cases in which the initial interpretation is correct. Moreover, cases in which intonational phrase boundaries aid interpretive processing decisions could be widespread. For example, if the end of an intonational phrase is interpreted as the end of a clause whenever such an analysis is grammatical, then the processor could regularly use intonational phrasing to more accurately predict whether a potential clause boundary location was an actual clause boundary location. Thus, the processor could close a clause as soon as an intonational phrase boundary was detected (or even plan to close a clause as soon as the location of the intonational phrase boundary could be predicted) instead of having to wait until it had received and partially processed material which would force the postulation of a preceding clause boundary, potentially reducing the need for reanalysis for a large class of early versus late closure ambiguities.

Little is known yet about the time course of semantic and pragmatic processing decisions, and it is beyond the scope of this dissertation to fully specify which interpretive operations take place immediately and which might take place later in the course of processing and thus be sensitive to the intonational pattern of the sentence. Nevertheless, the Interpretive Domain Hypothesis and the experiments presented in this chapter provide an important first step in delimiting how prosody affects interpretive decisions and when some of those decisions must be made. And, even with our limited knowledge of

interpretive processing, the research presented here demonstrates that the prosodic structure of an utterance is a necessary component of a complete model of semantic/pragmatic processing.

4.6 Conclusion

Previous work on prosody has suggested that prosody affects semantic/pragmatic interpretation in two ways: by the assignment of focus, done primarily through the placement of pitch accents, and by marking pragmatic distinctions through the choice of the tone or tones used in pitch accents and edge tones. In this chapter I have argued that prosody also affects the time course of semantic/pragmatic interpretive operations, as described by the Interpretive Domain Hypothesis. Thus, I have argued that prosodic phrasing can affect interpretive decisions and that intonational phrase boundaries and phonological phrase boundaries differ in their effects on the processing system in a manner which cannot be explained solely by acoustic differences, pragmatic effects, grammatical principles, or lexically- or syntactically-defined processing units.

The research presented in this chapter has also provided further evidence in support of grammatical theories of adjective interpretation, as the results of Experiment 3 are best explained by assuming that an adjective is typically interpreted with respect to its head noun, and grammatical theories of prosodic representation, as the results of Experiment 4 are best explained though a categorical distinction between phonological phrase boundaries and intonational phrase boundaries. As for theories of sentence processing, the results provide further evidence for the active use of the prosodic representation throughout the course of processing, affecting decisions from the phonological level up to the semantic/pragmatic model, and provide evidence in support of a specific hypothesis about how one aspect of the prosodic representation—the prosodic phrasing—affects semantic/pragmatic processing. Currently, the dominant models of sentence processing make little mention of the role of prosody in sentence comprehension, so the findings of this chapter and the previous chapters suggest that these models should be revised and re-

evaluated to account for the effects of prosodic structure. Thus, in the next chapter I consider how the specific hypotheses proposed thus far should be integrated into a general model of sentence processing, as well as the extent to which the Prosodic Visibility Hypothesis and the Interpretive Domain Hypothesis can be derived from more general principles of processing.

CHAPTER 5

EXTENSIONS AND CONCLUSIONS

5.1 Introduction

It is not surprising that prosody, important though it is to sentence comprehension, has not yet been adequately described in processing theories. Prosody interacts with multiple levels of the grammar, but is apparently only partially constrained by any given level. For example, the presence of a pitch accent may be required at the phonological level to avoid an unaccented phonological phrase, but the kind of pitch accent (e.g., L* versus H*) is probably constrained by the discourse structure. Similarly, the location of prosodic boundaries seems to be influenced by both focal structure and constituent structure, and thus may not be a reliable signal to the processor of either structure. Even when major functions of prosody are ignored, such as its role in expressing emotion and affect, prosody can still have extremely complex and subtle effects on interpretation. Further, although these effects are likely due to both grammatical constraints on well-formedness and independent processing strategies, refinements of prosodic theory continue to be made, and explicit, general hypotheses of how prosody affects processing have just begun to be carefully specified and tested. Thus, determining how prosody influences sentence comprehension is still very much a bootstrapping problem of identifying secure findings in processing research or grammatical theory and building from there to a more complete understanding of what grammatical structures the processor must build, what ambiguities the processor must resolve, and how and when it does so.

At the beginning of Chapter 2 I argued that the securely established effects of prosodic phrasing on sentence processing were compatible with several different hypotheses of how prosodic phrasing influences processing decisions. Perhaps the simplest of these hypotheses was the proposal that prosodic phrasing plays no special role in processing; that all effects of prosodic phrasing in sentence comprehension are the result of grammatical

constraints. That is, in forced-choice selection tasks, well-formed structures are chosen over ill-formed structures, and in 'on-line' tasks, inputs with a well-formed prosodic structure (for the syntactic or semantic structure that must be constructed) are processed more quickly and easily than inputs that are ill-formed.

Though it is surely true that well-formed structures are processed more easily than ill-formed structures and would be preferred to them in selection tasks, this hypothesis does not suffice to account for the results of three of the four experiments that I presented. Experiment 1 showed that a phonological phrase boundary located before an object NP led to higher percentage of NP attachments of a PP even though the NP attachment was presumably not required by grammatical constraints. Experiments 3 and 4 demonstrated that intonational phrase boundaries lead to further interpretive integration and evaluation of preceding material even though there is no known grammatical constraint that would induce such an effect. In fact, it would be undesirable to require the grammar to directly relate intonational phrasing and the kind of lexical interpretation examined in Experiment 4.

Only the bounded projection experiment can be readily explained through plausible grammatical constraints, and even in that case it is neither necessary nor preferable to do so given what we currently know about cross-linguistic patterns of focus and prosodic phrasing. This is not to say that there is no role for grammatical constraints involving prosody in sentence comprehension—indeed, I assume they are necessary for the construction of the prosodic representation, and it is the existence of grammatical constraints at the syntactic and semantic/pragmatic levels that allows prosodic information to be part of the computational vocabulary of the syntactic and semantic processing modules. However, it appears that grammatical constraints do not account for all of the effects of prosodic phrasing in sentence comprehension. As the experiments presented here show, some of the effects of prosody are most plausibly captured by independent effects of prosodic phrasing information on processing decisions.

The majority of processing theories of prosody have treated prosodic boundaries as locally-detected and locally-interpreted cues to the disambiguation of syntactic ambiguities, which primarily function to block the incorporation of additional material into the constituent currently being processed. Yet as all four experiments have demonstrated, what is crucial to the proper account of prosodic phrasing in parsing is that prosodic boundaries do not function only as indications of the closing edges of syntactic constituents. They can mark the leading edge of constituent, as in the NP- versus S-complement cases described in Chapter 1 or as in condition (b) of Experiment 1, where a phonological phrase boundary was located at the beginning of the direct object. They can also mark the (possible) edge of a focused constituent, as in Experiments 2 and 3, or either edge of a parenthetical, as in the control materials of Experiment 3. And, as Experiments 3 and 4 showed, they can mark the closing edge of something which is most plausibly a processing unit and not necessarily a syntactic unit.

As has often been noted in the literature, prosodic structure and syntactic structure are related, but not isomorphic. The strongest prosodic boundary in a sentence may tend to occur at the strongest syntactic boundary, but it can also be located at a minor syntactic boundary under certain conditions of focus or constituent length. Even when the prosodic structure does reflect the syntactic structure, prosodic boundaries do not always mark the closing edges of syntactic phrases. Thus, hypotheses that limit the effects of prosodic boundaries to cues of syntactic closure are only tenable within accounts that predict that all other effects of prosodic phrasing result from grammatical constraints. Given the range of grammatical constraints that would be necessary to account for the experimental findings presented here, it is hard to imagine a situation in which such a hypothesis would actually be required to account for an effect that was not covered by the grammar.

Clearly, hypotheses like Prosodic Closure or Parallelism are too limited. They only predict effects of prosodic boundaries when they occur at points of syntactic ambiguity, and they only predict effects of prosodic boundaries on syntactic decisions. Hypotheses

like Prosodic Chunking are less limited in that they can predict some of the effects of prosodic boundaries which precede points of syntactic ambiguity and could easily capture the bounded projection effect. Presumably, such approaches could also be extended to capture the interpretive domain effect with the proper specification of PPh effects versus IPh effects. However, chunking hypotheses do not account for the effects of multiple prosodic phrases (such as those seen in Experiment 1), and they seem to entail processing delays by not allowing the sub-structures built within a prosodic phrase to be attached to the larger partial phrase marker until the end of the prosodic phrase has been reached (e.g., attaching syntactic material in the second of two phonological phrases to a syntactic node within the first PPh would be delayed until the end of the second PPh).

In order to capture the broad range of effects that prosodic phrasing can produce in sentence comprehension, a processing theory must do several things. First, it must be able to respond to the overall pattern of prosodic phrasing in a sentence, and not just to a single prosodic boundary located at the point of syntactic ambiguity or immediately prior to it. Second, it must allow different types of prosodic information to interact with each other, for example, allowing prosodic boundaries and pitch accents to interact in the specification of focus. Third, it must distinguish among the various levels of prosodic structure in the prosodic hierarchy, such as distinguishing between the phonological phrase level and the intonational phrase level in English. These three requirements, taken together, provide extremely strong evidence that the processing system must build a prosodic representation at some stage of sentence comprehension and use the information encoded in the prosodic representation to inform interpretive decisions.

As for how the information in the prosodic representation is used by the processor, the studies presented in Chapters 2 through 4 suggest that prosodic phrasing information is used to delimit the most important local context for various kinds of processing decisions. That is, phonological phrases delimit the most important nodes for syntactic attachment decisions, and intonational phrases delimit the most important material for higher-level

interpretive decisions. The Prosodic Visibility Hypothesis and the Interpretive Domain Hypothesis manage to effectively account for the effects of phonological phrasing and intonational phrasing in defining processing domains. Prosodic Visibility captures the gradient effects of multiple phonological phrase boundaries on phrase marker construction exhibited in Experiment 1 without requiring processing decisions to be delayed until the ends of phonological phrases. Further, it also accounts for the bounded projection effect of Experiment 2, obviating the need for an additional grammatical constraint which would apparently apply solely to the projection of focus. Similarly, the Interpretive Domain Hypothesis accounts for the differential use of upcoming context seen in Experiments 3 and 4, in which the processor behaves as if any (non-obligatory) information that has not been encountered by the time an intonational boundary has been reached is not likely to be relevant for any remaining interpretive decisions. However, these hypotheses stipulate the effects of prosody in processing, rather than deriving them from general principles. Before adopting them as axioms, it is worth considering whether they might be specific instantiations of more general sentence processing strategies. It is also necessary to examine whether the effects they predict hold in languages other than English. And finally, if these hypotheses should be incorporated into sentence processing models, it is crucial to know whether they can be readily incorporated into any of the current models, and how their addition might alter the predictions of sentence processing models. This chapter provides preliminary consideration of these issues.

5.2 Tests of Prosodic Disambiguation in Other Languages

Speech processing involves the interaction of grammatical constraints (at multiple levels), processing-specific effects of prosody, and the more general processing effects that can be seen in studies on visually-presented materials. Because of the complexity these interactions create, I have focused in this work on the effect of prosody in the comprehension of English. Nevertheless, if prosodic structure guides processing decisions in all languages, it should be possible to see the effects of prosodic phrases on early

processing decisions in languages other than English. Unfortunately, the evidence about prosodic phrasing effects in other languages is even more limited than the evidence that exists for English.

It is certain that English is not the only language in which differences in prosodic phrasing can allow listeners to discriminate between two syntactic structures for a single word string. As part of a larger demonstration that prosodic structure is related to syntactic structure, but does not fully correspond to it, Nespor & Vogel (1986) conducted a systematic study of prosodic phrasing in Italian, with the goal of distinguishing the class of structures which could be discriminated by prosodic phrasing from the class of structures which could not.[1] In a forced-choice paraphrase selection task they presented subjects with 78 sentences instantiating ten different kinds of syntactic or lexical ambiguities in Italian. They found that subjects could reliably discriminate between sentences in which the syntactic differences correlated with differences in prosodic phrasing (including cases in which the prosodic phrasing differed at the phonological phrase level, the intonational phrase level but not the phonological phrase level, or at both levels simultaneously), but could not discriminate between ambiguous sentences which were not associated with differences in prosodic structure (i.e., sentences with lexical category ambiguities that were not predicted to affect the prosodic phrasing and sentences with within-category lexical ambiguities).

Nespor and Vogel argue from their results that prosodic phrases provide the initial domains for syntactic structuring. However, the cases in which prosodic phrasing affected discrimination were cases in which the prosodic structure was predicted to be well-formed for only one of the syntactic structures. Thus, their results, much like the results from

[1] It is crucial to note here that while I have been identifying prosodic phrases solely on the basis of their phonetic properties (i.e., evidence of phrase accents, boundary tones, final lengthening, silent intervals, pitch reset, or segmental effects), Nespor & Vogel argue that prosodic phrasing is primarily determined by syntactic structure, with some restructuring based on constituent length and rate of speech. Thus, what they define as a phonological phrase or an intonational phrase may differ slightly from the phonological phrases or intonational phrases described elsewhere in this dissertation.

English of Lehiste (1973), Warren (1985), and Price *et al.* (1991), do not necessarily show early effects of prosody on initial syntactic attachment decisions. Although they are certainly consistent with this claim, they are also consistent with a situation in which the initial syntactic structure is built solely on the basis of non-prosodic information, and then revised if necessary to reflect phonosyntactic constraints at a later stage of processing.

The evidence to date from Japanese is similarly inconclusive. Misono, Mazuka, Kondo, & Kiritani (1997), also using a forced-choice paraphrase selection task, found that listeners could discriminate between Japanese sentences that differed in the placement of an intonational phrase boundary, even in cases where the sentence was strongly biased pragmatically toward the other reading. Venditti & Yamashita (1994) found that listeners could discriminate between prosodic phrasing indications of the end of a simplex Japanese sentence versus the beginning of a relative clause. Thus, as in Italian, prosodic phrasing is clearly used at some stage of processing, but it is not yet possible to determine exactly when and how it is used.

Stronger evidence for the early use of prosodic phrasing in parsing decisions comes from work on PP attachment in French by Pynte & Prieur (1996), which was very similar in design to Experiment 1 (and conducted concurrently). Pynte and Prieur performed a series of experiments on French sentences such as the ones given in (5.1), which were produced with either a single prosodic boundary located before the direct object (as in condition (b) of Experiment 1) or with two boundaries, one before the object and the other before the PP (similar to condition (d) of Experiment 1). However, based on their description of F_0 minima and maxima, syllable durations, and pauses, it is likely that Pynte and Prieur manipulated intonational phrase boundaries instead of the phonological phrases boundaries of Experiment 1.

(5.1) a. Les espions informent % les gardes (%) du complot.
'The spies inform the guards of the conspiracy'

b. Les espions informent % les gardes (%) du palais.
'The spies inform the guards of the palace'

c. L'étudiant choisit % un appartement (%) avec soin.
'The student chooses an apartment with care'

d. L'étudiant choisit % un appartement (%) avec balcon.
'The student chooses an apartment with a balcony'

Pynte and Prieur found that the one-boundary productions either slowed reaction times in a word detection task for sentences biased toward VP-attachment (a, c) or facilitated reaction times for sentences biased toward NP-attachment (b, d) relative to the two-boundary production, as would be predicted by Prosodic Visibility. However, because their materials probably contained intonational phrase boundaries instead of phonological phrase boundaries, it is more likely that the prosodic conditions that phrased the object NP with a VP-attached PP violated a phonosyntactic constraint. This is supported by the results of their acceptability pretest, which found lower acceptability ratings for those cases predicted to be ill-formed by Selkirk (1984). Thus, Pynte and Prieur's results are consistent with Prosodic Visibility, but they do not provide unambiguous evidence in support of it. However, their work does support the general proposal that the information encoded in the prosodic structure should influence early processing decisions in all languages. Even if their results come solely from the processor's use of well-formedness constraints, they show that prosodic phrasing information is made use of, in some fashion, in tasks which do not require conscious consideration of syntactic ambiguities. Further, the effects of the prosodic information is available prior to the end of the sentence and final prosodic phrase, influencing decisions made at the point that the biasing word is encountered.

In summary, most of the research on prosodic phrasing in other languages, like the initial work on prosodic phrasing in English, shows only that prosodic phrasing has some effect on interpretation. While the research on Italian, Japanese, and French provides no evidence against the specific proposals of Prosodic Visibility or the Interpretive Domain Hypothesis, and the research on French provides some evidence for the early use of

prosodic structure, the cross-linguistic results to date are still compatible with a variety hypotheses about the role of prosodic phrasing in processing.

However, research on languages other than English has provided some information about the use of phonological constituents other than phonological phrases in comprehension. Notably, Cutler and colleagues have argued that in word segmentation listeners rely on whatever rhythmic structure characterizes their language (Cutler, Mehler, Norris & Segui, 1992); i.e., listeners use strong-weak patterns in English (Cutler & Norris, 1988) and Dutch (Vroomen, van Zon & de Gelder, 1996), syllabic units in French (Cutler, Mehler, Norris & Segui, 1986), and moraic units in Japanese (Cutler & Otake, 1994).

It is not possible to consider here the range of evidence for and against the effects of various phonological distinctions in lexical access and retrieval, but the word segmentation findings highlight the importance of testing for the effects of phonological units across a range of language types. If the effects of higher-level prosodic structure are similar to the effects of rhythmic structure, then the relationships between different levels of prosodic phrasing and different kinds of processing effects may be somewhat different across languages. That is, the specific constraints that are active in a language, including both constraints on the exact form of the prosodic structure for the language and constraints relating prosodic structure to syntactic, semantic, or pragmatic structure could strongly influence the processor's favored mapping between prosodic structure and the other structures. Thus, it is critical that we examine prosodic phrasing effects in a range of languages so that we can properly characterize any underlying processing principles governing prosodic phrasing. Of course, if the effects captured by Prosodic Visibility and the Interpretive Domain Hypothesis can be derived from more general processing strategies, it should also be possible to predict the potential differences in prosodic phrasing effects on processing from the languages' different grammars.

5.3 On Deriving Prosodic Effects

In principle, there are at least three ways in which the effects of prosodic phrases on sentence comprehension could be derived from other, more general, processing principles. One possibility is that all of the effects are due to timing differences: because both phonological phrases and intonational phrases are associated with additional time (from final lengthening and/or silent intervals at the end of the phrase), they might simply allow more time for syntactic nodes preceding prosodic boundaries to become less salient (in the case of phonological phrases) or for further interpretation of material preceding the boundary to take place (in the case of intonational phrases).

Based on the evidence that is currently available, this explanation does not seem sufficient to account for all of the effects of prosodic phrases. Recall that Dobroth (1996) found some effect of prosodic boundaries signaled by allophonic variation, with no durational differences in the materials. In the research presented here, Experiment 4 showed a processing difference between PPh boundaries and IPh boundaries even though the timing difference between the two kinds of boundaries was minimized and the results exhibited no reliable correlation between the length of the ambiguous region (which contained the prosodic boundary) and comprehension time for the conditions requiring reanalysis. Finally, Speer & Kjelgaard (submitted) compared sentences with short and long pause durations by manipulating the silent intervals associated with hesitation pauses and prosodic boundaries. They found no effect of pause duration on sentence comprehension times for a speeded grammaticality judgment task or an end-of-sentence comprehension task. However, in their study there was generally only one grammatical syntactic attachment available for the lexical item following the pause. Thus, it is still conceivable that durational differences could affect processing decisions when they occur at points of attachment ambiguity. Nevertheless, while further study is necessary to show conclusively that prosodic phrasing effects are not solely dependent on durational effects, it seems unlikely that timing is the primary explanation for them.

A second possibility is that prosodic and other linguistic units that can be easily identified from the sensory input form perceptual units, and affect processing because all such perceptual units function as preliminary processing domains. This explanation seems more tenable than the durational explanation. There is clear evidence from a range research areas that rhythmic groups in general and prosodic phrases in particular form perceptual units (e.g., Martin, 1967, 1968; Suci, 1967; Johnson, 1970; Tyler & Warren, 1987; Jusczyk *et al.*, 1992; van Nice, 1994 Morgan, 1996;). Indeed, the notion of perceptually-based units motivated the Prosodic Visibility Hypothesis given here and the more general Visibility Hypothesis of Frazier & Clifton (1995). However, recognizing that prosodic phrases form perceptual units, even with the plausible assumption that phonological phrases and intonational phrases form different types of perceptual units, does not explain why phonological phrases and intonational phrases should cause the specific differences in processing shown here. Before we can derive Prosodic Visibility and the Interpretive Domain Hypothesis directly from the effects of perceptual units, we need a performance-based explanation of why intonational phrases seem to correlate with semantic/pragmatic processing decisions but phonological phrases do not.

The third possibility is that the effects of prosodic phrasing stem from the partial grammatical information that is available to the processor at intermediate points in the sentence. Presumably, if the processor is building and using a prosodic structure, it must continually incorporate material into a partial sentence structure that is both prosodically well-formed and syntactically well-formed. As is the case for (purely) syntactic parsing, the prosodic constraints of the grammar may allow more than one well-formed structure at intermediate points in the string; at such a point, the processor must choose which structure (or structures) to build. Thus, we should expect that some prosodic effects in processing might result from the selection of whatever kind of structure is predicted to be preferred at points of grammatical ambiguity. For example, the most frequent or contextually most

expected phonosyntactic structure might be selected in constraint-based models, and the simplest phonosyntactic structure should be selected in the Garden-Path model.

In cases of syntactic closure ambiguities, this explanation of how Prosodic Visibility could be derived seems fairly straightforward. Consider the case of a phonological phrase boundary at the right edge of an optionally transitive verb in a sentence-initial subordinate clause. This boundary could result from a constraint aligning the right edge of the verb phrase (or the clause) with the PPh boundary. Alternatively, the processor could assume that the boundary results from the need to separate off a particularly lengthy object NP into a second PPh. It seems very reasonable that the former phonosyntactic structure would be considered simpler, since it entails less syntactic and phonological structure (and in particular, less otherwise unmotivated structure), and that the parser could construct it more quickly than the latter structure; it also seems reasonable that this structure could be the more frequent one. Therefore, an account based on minimal (or favored) phonosyntactic structure could likely capture the early versus late closure effects seen in the Speer *et al.* experiments.

However, it is difficult to see how an this kind of explanation could account for the results of the PP-attachment experiment of Chapter 2. In particular, it does not seem to explain the lower percentage of VP attachments in the condition with both a pre-NP boundary and a pre-PP boundary compared to the condition with just the pre-PP boundary or the condition with no internal prosodic boundaries. It is also not clear how the differential effects of phonological phrases boundaries and intonational phrase boundaries on interpretive decisions identified in Chapter 4 could be derived simply from the assumption that the minimal linguistic structure is assigned. However, we are still defining the prosodic constraints of the grammar, and we are still in the early stages of psycholinguistic research on prosody. As our understanding of the grammar of prosody develops and as additional experimental results are accumulated, this explanation of Prosodic Visibility will need to be re-evaluated. Whether it ultimately proves to be true or

not, its existence underscores the need for very precise predictions and measurements of when different kinds of information become available to the processor and how that information interacts with various processing decisions. In particular, it underscores the importance of considering phonological and phonosyntactic well-formedness at each point in the input in addition to well-formedness at other levels of representation.

Whatever the deeper explanation is for its effects, the role of prosodic phrasing in sentence comprehension seems to one of balancing and informing both the tendency of the syntactic parser to continue to incorporate material into the constituent currently being processed (with the minimum number of nodes) and the need to perform interpretive decisions as quickly as possible. The prosodic structure underspecifies the syntactic and semantic relationships in the material, but by separating material that is not closely related syntactically or semantically and keeping together material which should in some way be interpreted together, it apparently creates an influential preliminary structure for many kinds of processing decisions. Consequently, the possible effects of prosodic structure should be considered carefully in any sentence processing research in which it could vary across conditions—including research with visually-presented materials, as recent work has suggested that prosody can affect comprehension even when it is constructed by the reader during silent reading on the basis of syntactic or semantic information (Bader, 1994, 1996; Gilboy & Sopena, 1996). Furthermore, to accurately describe and explain sentence comprehension processes, we will evidently require a very detailed account of the processes that take place within and across multiple levels of representation. Thus, in the next section I consider some of the predictions of incorporating a full prosodic representation into the processing system and some of the basic questions that must be addressed by a model which does so.

5.4 Incorporating Prosodic Structure into Processing Models

The ways in which prosody is expected to influence processing strongly depends on the prosodic representation that is assumed. If prosody is viewed as a global sentence tune

which cannot be analyzed into subcomponents, we should expect it to affect processing only after the particular tune has been identified, presumably at or near the end of the sentence, and hence fairly late in the course of processing. If prosody is treated as an unstructured string of elements, we might predict early processing effects of these elements, but we should expect to see only local effects of them. A prosodic boundary located at or immediately prior to a point of ambiguity might affect an attachment decision, but a boundary which precedes the local one should not.

If prosody is analyzed as a set of phonological elements organized into a structure, the set of predictions is much larger. One prediction of having a prosodic structure was verified in Chapter 2, where the PP-attachment experiment showed that the overall pattern of phonological phrasing influenced the resolution of attachment ambiguities. That is, even when considering the effect of just one kind of prosodic element, the phonological phrase boundary, the effect of any particular boundary must be determined with reference to the prosodic structure in which it is embedded. Another prediction is that contrasting levels of the prosodic structure could cause separable effects on processing. This was confirmed in Chapter 4 by Experiments 3 and 4, which showed separable effects of phonological phrase boundaries and intonational phrase boundaries.

Assuming that the form of the prosodic structure is somewhat constrained, then there should be cases when it is possible for the processor to predict parts of the prosodic structure in advance. For example, in (5.1), if the processor is building a Pierrehumbertian prosodic representation, the processor could anticipate the lack of phonological phrase boundary after *telephoned* as soon as it determines that the main-stressed syllable is unaccented, since a phonological phrase without a pitch accent is ill-formed in this theory of prosodic structure.

(5.1) (When the manager H* L-)$_{PPh}$ (telephoned...

Therefore, if the processing model happens to predict that the parser would otherwise consider both the transitive and intransitive analysis of *telephoned*, or would initially prefer the intransitive analysis of this verb, it might use this information as early evidence against the intransitive parse. In fact, there is already some evidence for the ability of listeners to anticipate prosodic form. Grosjean (1983; Grosjean & Hirt, 1996) found that English-speaking subjects (but interestingly, not French-speaking subjects) could reliably predict the amount of time to the end of the sentence on the basis of prosodic cues at the end of a word which potentially ended the sentence.

Example (5.1) also illustrates two more predictions of having a prosodic structure instead of a string of elements. First, the absence of a prosodic element, such as the absence of a pitch accent on *telephoned*, could influence processing by precluding a particular analysis. Second, different types of prosodic information, such as pitch accents and edge tones, could interact with each other. We have already seen one piece of evidence for this kind of interaction in the bounded projection experiment of Chapter 3. There, prosodic boundaries influenced effects primarily based on pitch accents. The reverse may also hold true. Warren, Grabe & Nolan (1995) present evidence that pitch accent placement can affect the resolution of early versus late closure of a subordinate clause. Specifically, they presented auditory fragments such as in (5.2) in a cross-modal naming task, in which subjects named a visual word which was compatible with only the early closure analysis of the auditory fragment. The auditory fragments were produced with either an early or a late prosodic boundary (indicated by parentheses) and either shifted or unshifted stress on the penultimate word of the fragment, as indicated by capitalization.

(5.2) Auditory Fragment:	Visual Word:
a. (Whenever parliament discusses Hong KONG) (problems	ARISE
b. (Whenever parliament discusses HONG Kong) (problems	ARISE
c. (Whenever parliament discusses Hong KONG problems)	ARISE
d. (Whenever parliament discusses HONG Kong problems)	ARISE

In conditions (a) and (c), *Hong Kong* was produced with citation-form stress, so that the primary stress fell on the second syllable. In conditions (b) and (d), it was produced with stress shifted to the first syllable. When a string such as *Hong Kong problems* is part of a single phonological phrase, shifting the stress of *Hong Kong* avoids the clash of two adjacent stresses from the final syllable of *Hong Kong* and the initial syllable of *problems*. However, when the two words are in separate prosodic phrases, stress shift is unnecessary, and probably only well-formed if the stressed syllable is contrastively focused. Thus, if the processor is building a prosodic structure for the input, the presence of stress on the first syllable of *Hong Kong* should signal either that there is additional material coming within the phonological phrase (and this material is initially stressed) or that *Hong* is constrastively focused. Thus, stress shifting could influence the processor's expectation of an upcoming prosodic boundary.

As expected, they found shorter naming times for the (a) and (b) versions, with a cooperating prosodic boundary, than for the (c) and (d) version, with a conflicting prosodic boundary. They also found that, across their entire set of materials, naming times were numerically shorter, but not significantly shorter, for the conditions with unshifted stress than for the conditions with shifted stress. However, using a set of ratings from five subjects, they performed a post-hoc separation of the items into a set for which contrastive focus was likely (e.g., *navy-blue*, which could easily be contrasted with *sky-blue*) and a set for which contrastive focus was less likely (e.g., *routine*). Analyses of these sub-groups revealed that naming times were significantly shorter in the unshifted stress conditions than in the shifted stress conditions for the second set of items, in which the contrastive focus interpretation was unlikely. Further testing will be needed to separate out such factors as whether subjects merely have longer naming times when the phonological structure is ill-formed (cf. Tyler & Warren, 1987) from the possibility that the processor can exploit shifted stress as a predictor of a late phonological boundary and hence a late syntactic boundary. But regardless of whether the effect is purely phonological or

phonosyntactic, it is more evidence of the importance of recognizing prosody as part of a well-formed prosodic structure in sentence processing research.

As these many predictions illustrate, incorporating prosodic structure into processing models is desirable on purely theoretical grounds because doing so generates many interesting and testable hypotheses about sentence processing. It is equally desirable empirically because of the wide range of data that can be more readily accounted for by assuming a prosodic structure than by assuming some other representation of prosody. However, the inclusion of a full prosodic representation in processing increases the number of structures which must be built by the processor. If prosodic structure is used in processing, our models must be able to specify how it is constructed. Moreover, if the method for building the prosodic structure differs from the method for building other linguistic structures, there should be a principled explanation of why the two methods differ. Thus, consideration of the relative success of various processing models in parsing prosodic structure (at the phonological level) may supply an additional means of choosing among possible models of sentence processing. Here, I will consider how the parsing of prosody might fit into just one model, the Garden Path model.

Just as there are temporary and standing ambiguities in linguistic material at the syntactic level, there can be more than one mapping from the rises, falls, and various rhythmic patterns of an utterance to the phonological representation of its prosodic structure. Beckman (1996) discusses several cases of this kind of phonological ambiguity in Japanese and English, such as the ambiguity found in Experiment 3, where a high-low-high sequence could be analyzed as H* L- H* or H* L+H* (among other things). Although they were not designed as a test of how prosodic structure is built, the Speer *et al.* experiments that have been discussed throughout this work may be interpreted as one source of evidence for how ambiguous structures are parsed at the phonological level.

Recall that Speer *et al.* contrasted unambiguous conflicting or cooperating prosodic structures with a prosodically ambiguous baseline prosody. This baseline prosody was

produced with a contrastive, L+H* pitch accent on the subject of the subordinate clause and an H* pitch accent on the predicate of the matrix clause. The material between these two accents was deaccented, and phonetically ambiguous between a phonological structure with an L- phrase accent aligned with the right edge of the subordinate verb and one with an L- phrase accent aligned with the right edge of the following NP. A sample set from the cross-modal naming experiment employing phonological phrase boundaries is given in (5.3).

(5.3) Auditory Fragment:	Visual Word:
Cooperating prosody:	
a. (When Roger leaves the house H-)$_{PPh}$	IT'S
b. (When Roger leaves H-)$_{PPh}$ (the house	IS
Baseline prosody:	
c. (When Roger leaves the house	IT'S
d. (When Roger leaves the house	IS
Conflicting prosody:	
e. (When Roger leaves H-)$_{PPh}$ (the house	IT'S
f. (When Roger leaves the house H-)$_{PPh}$	IS

Presumably, the extended low F_0 of the deaccented region provides sufficient evidence to indicate to the processor that an L- phrase accent must be present in the prosodic structure. However, the processor does not appear to align that phrase accent (and postulate a phonological phrase boundary) at the first point at which it is possible to do so. If it did, it would postulate a phonological phrase boundary immediately after the subordinate verb. In this case, the baseline prosody and the early closure cooperating prosody would acquire the same phonological phrasing (although with different phrase accents), so we should expect to see similar processing behavior for the two conditions. In fact, the early closure baseline prosody exhibits evidence of processing difficulty, while the early closure cooperating prosody does not. Naming times were significantly shorter in both the late closure baseline prosody condition and the early closure cooperating prosody condition than in the early closure baseline prosody condition, and times did not differ significantly between the early closure cooperating prosody condition and the late closure

baseline prosody condition. Thus, in phonological parsing as in syntactic parsing, the processor appears to follow a late closure strategy of continuing to incorporate material into the current (phonological) phrase.

Moreover, this effect is not accounted for by assuming that a syntactic preference to attach the NP to the subordinate verb overrides a phonological early closure preference. Consider the contrast between the early closure conflicting prosody case, in which an unambiguous phonological phrase boundary was located after the NP, and the early closure baseline prosody case. In both cases, there is only one grammatical attachment of the following, disambiguating, word. With the baseline prosody, the phonetic information is consistent with the presence of a phonological phrase boundary after the NP, but the processor evidently does not postulate a phonological phrase boundary in this location either: naming times were longer in the early closure conflicting prosody condition than in the early closure baseline prosody condition.

More extensive research, on a range of structures, will be necessary before the apparent similarity of phonological parsing and syntactic parsing seen in the Speer *et al.* results can be confirmed. For the present, though, it appears that the use of prosodic structure can be easily incorporated into the Garden Path model of sentence processing. First, although several kinds of prosodic phrasing effects have been demonstrated in this dissertation, they all conform to the view of modularity presented in Frazier (1990). Second, all of the secure effects of prosodic phrasing in early comprehension processes can be accounted for through serial processing with the recognition that structure-based processing includes prosodic structure in addition to other linguistic structures and the addition of Prosodic Visibility and the Interpretive Domain Hypothesis. Although these hypotheses cannot yet be derived from more general processing strategies, the perceptually-based account and the minimal structure account each provide a promising way in which they might be. And finally, based on the very limited evidence currently available, I have speculated that no

additional principles are required to account for how the prosodic structure is parsed at the phonological level.

5.5 Conclusion

Although the idea that prosodic structure is necessary for sentence comprehension may seem obvious, it was only about twenty years ago that Pierrehumbert introduced the proposal that intonational contours could be analyzed as a structured sequence of high and low tones. Many other developments in prosodic theory, such as the distinction between phonological phrases and intonational phrases, are even more recent. As I observed above, it is very difficult to account for the kinds of effects presented here if the intonational contour is not broken down into smaller prosodic elements like pitch accents, phrase accents, and boundary tones. Thus, it is probably natural that processing models have centered on morphosyntactic information and have made little or no mention of prosody.

Here, I have attempted to address this gap by posing what seemed to be the most fundamental question for prosodically-enriched processing: how (if at all) are the various distinctions encoded by the prosodic representation used by the processor? What I have provided here is (unsurprisingly) only a partial answer to the question. I have argued that all prosodic elements are interpreted by the processor with respect to the prosodic structure that contains them, and presented various sources of evidence for this hypothesis. Thus, I have presented empirical facts which argue against the view, implicit in much of the psycholinguistic literature, that prosodic boundaries are used only to block certain syntactic attachments and pitch accents are used only to mark information status, with no interaction between these two types of prosodic information. I have also shown that the effects of prosodic phrasing go beyond effects on syntactic attachment decisions, affecting both the interpretation of focus and higher-level semantic/pragmatic processing, and proposed two explicit hypotheses of how prosodic phrases affect processing decisions, the Prosodic Visibility Hypothesis (for phonological phrasing) and the Interpretive Domain Hypothesis (for intonational phrasing).

However, there are many prosodic distinctions that have not been addressed. To name just two, the difference between high and low phrase accents or high and low boundary tones surely has a significant effect on at least some processing decisions (see Grabe, Warren & Nolan (submitted) for some intriguing preliminary evidence of this), so explicit hypotheses about their particular effects will need to be developed and tested. Similarly, the use of variations in pitch range has not yet been accounted for in sentence processing research, but this too is crucial to a full explanation of how prosodically-structured sentences are processed, as observed by Marcus & Hindle (1990) and suggested by the control sentences of Experiment 3.

In addition to identifying all of the effects of the various contrasts found within an intonational contour, there are many theoretical issues left to resolve. In fact, even the nature of the prosodic structure requires further consideration. Issues such as how to naturally encode pitch range variations must be addressed, as well as alternatives to the prosodic theory assumed here, such as the proposal of Ladd (1986, 1996) that prosodic structure is recursive. Within psycholinguistic theory, we must evaluate how prosodic structure can be included in models other than the Garden Path model. Further, the many questions of how and when the processor makes use of different sources of information, including prosodic information, must be carefully addressed in each model of processing, and specific proposals of how particular sentential structures are processed must be expanded to encompass any relevant prosodic information.

As we have seen, a well-formed prosodic structure does not necessarily disambiguate the syntactic structure of a sentence. Many syntactic structures may be possible for any given prosodic structure, and many prosodic structures are possible for any given syntactic structure. If the goal of psycholinguistic research is limited to syntactic parsing, prosodic structure, with all of its variability, may not seem very helpful. For example, Allbritton, McKoon & Ratcliff (1996) had subjects read aloud paragraphs containing syntactically ambiguous sentences preceded by disambiguating information. Presumably, the

appropriate syntactic structure for the ambiguous sentence was made apparent to the subject by the preceding context. Yet Allbritton *et al.* found that subjects consistently disambiguated the syntactic structure only when they were informed of the ambiguity in the target sentences and instructed to use disambiguating pronunciations.

If prosody is expected to be an invariant cue of syntactic structure, this result is quite discouraging. However, if prosody is recognized as a complex structure that is subject to both phonological well-formedness constraints and constraints relating prosodic structure to (multiple) other linguistic structures, this result is quite interesting. It suggests that speakers might be adept at marking important discourse contrasts like given versus new information with contextually-appropriate strategies. That is, speakers of English might normally signal that a word or phrase expresses given information by deaccentuation and dephrasing (i.e., reduction of the prosodic phrase structure), but also be able to employ low pitch accents and richer prosodic phrase structure when the syntactic structure is recognized as ambiguous.

As sentence processing research broadens from primarily studying isolated sentences to more complete accounts of how language is produced and understood in natural discourse contexts, I suspect we will discover that prosodic structure is far more predictable, and therefore more constraining, or even predictive, than it seems today. Yet even if we find that prosody is never fully predictive of syntactic, semantic, or discourse structure, prosodic structure appears to be a necessary component of any model of natural language processing, and an extremely interesting avenue for future psycholinguistic research.