Intonation and sentence processing

Shari Speer[†], Paul Warren[‡] and Amy Schafer[°]

† Ohio State University, USA

‡ Victoria University of Wellington, New Zealand

° University of Hawai'i at Manoa, USA

ABSTRACT

This paper summarises recent research concerning the relationship between intonation and the syntactic analysis of sentences. After introductory comments on the nature of intonation, we discuss methodological problems in determining the relationship between syntactic and intonational structure, and the potential dangers of basing claims about this relationship on scripted readings rather than on spontaneous speech. We present some of our own speech production data from the SPOT project, and highlight the variability in the intonational realisation of that data. After discussing the broad question of whether correspondences between syntactic and intonation structure are speaker- or listener-oriented, we review experimental data on the role of intonation in sentence comprehension, and finally discuss the position of intonation in the sentence processing mechanism.

1. INTRODUCTION

Recent reviews of the relationship between prosody and language processing [14, 47] show how the prosodic structure of spoken language has been implicated in production and comprehension at a range of levels of analysis. These include the segmentation of the speech stream into words and the access of word forms from the mental lexicon, the segmentation of larger stretches of speech into syntactic constituents, the determination of linguistic and paralinguistic meaning, and the establishment and maintenance of discourse functions. In this paper we concentrate on the relationship of intonational structure to syntactic structure, focusing on whether and how differences in intonational form are made available for the resolution of syntactic ambiguity. We concentrate here on recent findings, as earlier research has been summarised in the reviews cited above.

Before we discuss the relationship of intonation to sentence processing, we need to comment on what we mean by "intonation", "prosody", etc. Defining these and related concepts solely in terms of the phonetic properties involved proves to be somewhat problematic, since particular prosodic phenomena may be signalled by a mix of different phonetic cues. For instance, marking a syllable as prominent (perceptually), corresponding to some notion of stress or accentuation, may involve aspects of duration and amplitude as well as the height and movement of pitch, and these properties may combine in different ways in different utterance contexts (as well as in different speakers, varieties, languages, etc). Similarly, sentence intonation involves not only tonal patterns over the utterance but also aspects of phrasing and prominence, which also involve durational properties and amplitude. Furthermore, different acoustic properties interact in what listeners perceive as a particular prosodic feature. Thus, pitch is not just cued by the rate of vibration of the vocal folds, but is also affected by factors such as where in the speaker's pitch range the relevant portion of the utterance lies, and other phonetic properties, such as the relative amplitude or duration of the utterance part in question.

These few comments illustrate the richness and complexity of prosodic analysis. This complexity means that objectively measurable properties of utterances such as the duration or mean pitch level of a syllable will often only present a partial picture of the prosody of the utterance. In addition, therefore, it is critically important to describe the utterance in terms of a phonological prosodic system. From the point of view of language processing, the significance of such phonological analysis is that it acknowledges that perceptually relevant distinctions in prosody often involve the combination of more than one acoustic parameter, and that there is context-based interpretation of local phonetic properties. This view acknowledges both the perceptually salient detail and the linguistically relevant contrasts to which this detail contributes.

The main questions which we discuss in this paper are the following:

- How can researchers determine the relationship between syntactic and intonational form in the production of sentences?
- What is the relationship between intonational form and syntactic form in sentence processing?
- What are the relative roles of the sentence production and comprehension processes in determining sentence-level intonational structure?
- How does intonational structure help comprehenders in their processing of sentences?

Our discussion includes reference to recent results from our own research on the SPOT project (e.g. [36, 37, 49]) and to related research on English and other languages.

2. THE LABORATORY SPEECH PROBLEM

There is now a considerable body of research data relating to issues of whether, when, and how listeners use prosodic information to disambiguate syntactically ambiguous utterances (see section 5 below). Typically, such research presents carefully selected utterances, involving a prosodic feature that contrasts minimally in different

conditions, and requires listeners to respond in some manner that will reveal their use of the relevant prosodic information in arriving at a structural interpretation of the utterance. Usually, the utterances in such experiments have been recorded by expert speakers (often the experimenters or their colleagues) who carefully produce the appropriate contrasts. While the research generally shows that listeners do use available prosodic contrasts in their interpretation, the wider relevance of such research has been challenged by claims that the contrasts investigated are a product of laboratory conditions and are not readily found in everyday situations. Thus, it is argued that speakers' awareness of ambiguity is a primary factor that influences the salience of prosodic contrasts in their production of ambiguous sentence materials. In one study [1], little prosodic marking resulted when ambiguous sentences were presented with short paragraph contexts, and speakers' attention was not drawn to the ambiguity, in contrast to clear prosodic disambiguation when speakers read decontextualised ambiguous sentences, using instructions that drew attention to the meanings involved and urged them to realise meaning differences through differences in pronunciation. Such evidence suggests that the use of expert and aware readers in producing materials for comprehension experiments will result in test materials that may be quite different from normal conversational speech, reducing the generalisability of the results.

Some comprehension research (e.g. [48]) has addressed this problem by modelling the experimental materials on recordings of naïve speakers, and by using sentences that contain temporary ambiguities (resolved within the utterance) rather than the global ambiguities used in [1]. This is still not spontaneous speech, though, since it involves reading rather than talking. Readers and talkers have different pragmatic goals, and different processing demands – unsurprisingly a number of researchers have commented on the resulting differences in the prosody of read and spontaneous speech [3, 7, 16]. While using spontaneous speech is desirable, it is extremely difficult to obtain the required minimally contrasting utterances in completely unscripted spontaneous recordings. Some degree of spontaneity might be achieved using role play situations or map tasks [2]. Here, though, there are also typically few constraints placed on the syntactic forms of expression. Other techniques, such as descriptions of node networks [24] or tangram shapes [11, 12], have revealed much about the use of prosody in the planning and repair of utterances and in the construction of collaborative discourse, but these tasks are again not designed for eliciting specific syntactic contrasts.

In our own attempts at resolving or at least side-stepping the laboratory speech issue, we have constructed a task in which a pair of speakers must negotiate the movement of objects around a game board. We elicit multiple renditions of the syntactic structures of interest by requiring participants to use a fixed set of sentence frames and object names to construct instructions, requests and acknowledgements. The two participants have slightly differing boards, and different roles. The "Driver" knows the ultimate goals for the objects, and issues instructions accordingly. The "Slider" moves the objects in response to the Driver's instructions, and knows the locations of bonuses and dangers, but not the goals for the objects. The Slider asks questions about which object to move and provides confirmation that the object has been moved. Both Driver and Slider are aware of basic game rules governing the types of movement for game board objects. Participants quickly become familiar with the utterances available for use under the rules of the game, and learn to produce them fluently and without recourse to printed sentence lists. Game play lasts approximately two hours. The task produces a rich source of data for the study of syntactic ambiguity resolution in a more spontaneous speech situation than that used in traditional sentence reading tasks. Since the utterances include a number of syntactic ambiguities frequently studied in comprehension studies, we can also use our recordings to examine the use of prosodic information in ambiguity resolution during comprehension. What is more, by careful planning of our gameboard layouts, we can construct game situations in which potentially ambiguous utterances are contextually disambiguated and other situations where there is a contextual bias towards one of the meanings.

The syntactic contrasts that we have investigated include early (1) vs. late (2) closure of constituents [37]; high vs. low attachment of prepositional phrases (PPs; e.g. in (3) the triangle could be an instrument used to move the square or it could be part of a description of a house-like piece, the "square with the triangle") [36]; and filler-gap ambiguities (4, 5) [42].

(1) When that moves the square will...
(2) When that moves the square it...
(3) I want to change the position of the square with the triangle
(4) Which triangle$_i$ do you want to change the position of ____$_i$ this time?
(5) Which triangle$_i$ do you want ____ $_i$ to change the position of the square?

Our data include recordings of eight pairs of Midwestern American English (MAE) speakers and seven pairs of New Zealand English (NZE) speakers, each pair completing as many as four different games, swapping Driver and Slider roles between games. The data have been analysed both acoustically (duration, pausing, F0 patterns) and auditorily (using the English ToBI transcription system [5, 28]). Our auditory analysis includes identification of the position of the strongest prosodic break, in accordance with the categories of the ToBI transcription: Intonational Phrase breaks (IPh) are stronger than Intermediate Phrase breaks (ip) which are in turn stronger than word-level breaks. This produced three main groups of utterances, with the strongest break in the utterance at: 1) the critical syntactic choice point, 2) an opposing point, and 3) at both of these locations. Across sentence types, and for both MAE and NZE speakers, phonetic measures and transcription data consistently indicated that naïve speakers reliably produce patterns of

prosodic phrasing that indicate the syntactic constituent boundaries they intend to communicate. More specifically, speakers tended to produce the strongest break in the utterance at a location that resolved the syntactic ambiguity. The following paragraphs summarise some of our results.

Early/late closure, (1) & (2). Our comparison of early and late closure sentences showed that speakers indicated the resolution of this temporary syntactic ambiguity with prosody even when utterance contexts fully disambiguated the structure ([37], see section 4 below). Phonetic and phonological analyses showed reliable prosodic disambiguation, with the strongest prosodic break in the utterance most likely to occur at the subordinate clause boundary. The verb *moves* was longer and followed by a longer silence when used intransitively as in (1), than transitively as in (2) (mean durations of *moves*: 477ms and 300ms; of silence: 343ms and 6ms). Also the noun *square* was followed by a longer silence in direct object position as in (2) than when it was the subject of the second clause as in (1) (73ms vs. 16ms).

All fluently produced MAE early/late closure sentences were transcribed by two teams of ToBI transcribers. One team transcribed full utterances that included the syntactically disambiguating portion of the sentence. Results indicated that 92% of early closure utterances had the strongest break following *moves*, while 96% of late closure utterances were produced with the strongest break following *square*. Speakers seldom produced the strongest break at a location consistent with the competing syntactic parse (early: 2%; late: 4%), and infrequently produced equal breaks at the two locations (early: 6%; late: 0%). The second team transcribed utterances that had been truncated before the syntactically disambiguating word, thus removing the possible impact of syntactic structure on transcription decisions. This truncation reduced phonetic information available on the final word of the resulting fragments, e.g. boundary information on the clause-final word of late closure utterances. Not surprisingly, there was a corresponding reduction in the number of boundaries transcribed at this location. Still, the pattern of results was strikingly consistent with the full transcriptions: the majority of utterances had the strongest break at the syntactic clause boundary (early: 83%; late: 72%), few had this break at the syntactically misleading location (early: 3%; late: 12%), and few had equal breaks at both locations (early: 14%; late: 16%).

Prepositional phrase (PP) attachment, (3). We also found that our naïve speakers used prosodic phrasing to disambiguate high vs. low attachment of a PP. Our phonetic and phonological analyses showed reliable prosodic disambiguation consistent with previous production results [46] and with models of prosodic disambiguation in sentence comprehension [10, 29]. These predict that sentences like (3) should be biased toward high attachment when the strongest break follows the word *square*. Measures for utterances from our MAE speakers in both Driver and Slider roles revealed a greater combined duration of *square* and following silence for the high-attached form, indicating instrumental use of the triangle (high: 786ms; low: 385ms). Analyses of the NZE utterances gave similar differences (high: 855ms; low: 366ms). Such rhythmic contrasts, particularly at the point immediately preceding the PP, have previously been shown to be a key indicator of the attachment type [13].

To obtain phonological analyses of these utterances, fluent productions of the PP attachment sentences were excised from the game context and presented for transcription without information about the speaker's intended syntax. Transcriptions by 5 trained ToBI transcribers of 13 MAE speakers' Driver and Slider utterances revealed a strong influence of intended syntactic structure on relative boundary strength. The strongest boundary followed *square* in 50% of 113 high attached tokens, as compared to 4% of the 168 low attached tokens. The strongest boundary followed a word other than *square* more often for low than high attachment sentences (81 % vs. 22%). Tied strongest boundaries, with one following *square*, were more likely for high than for low attachment (28% vs. 15%). Three trained ToBI transcribers found a similar pattern of results in their auditory analysis of Driver PP utterances from 6 NZE speakers. The strongest boundary followed *square* for 57% of 72 high attached tokens, but only 6% of the 72 low attached tokens. It followed a word other than *square* more often for low than high attachment (71% vs. 11%). Equally strong strongest boundaries were marked after both *square* and another word more often for high than low attachment (32% vs. 24%).

Filler-gap structures, (4) & (5). In our game task, naïve speakers used prosodic phrasing to differentiate two types of filler-gap *wh*-questions. Previous research comparing similar materials demonstrated prosodic differences at potential gap locations, suggesting that syntactic gaps may be directly indicated in prosodic structure [27]. However, recent findings suggest that when potential gap location and syntactic complement/adjunct structure are manipulated separately, prosodic phrasing reflects 'sense-unit' constituents sensitive to these relations, rather than directly reflecting the location of a syntactic gap [44]. Both of these previous studies used reading tasks to elicit question production, and phonetic measures to evaluate prosodic structure. Our study differs from these in two respects: both of the gap sentences used in our game task have complement *to* clauses; our analysis included duration measures, transcriptions, and the incidence of *wanna*-contraction across one potential gap site. *Wanna*-contraction was more frequent in the MAE data, and more likely in both data sets when there was no gap between *want* and *to*, as in (4) [50]. Phonetic analyses indicated that both groups of speakers consistently lengthened the word preceding the gap location and the following silence. ToBI transcriptions of matched utterance sets showed prosodic constituent breaks most often patterning with the location of major syntactic breaks. While our results need not imply that syntactic gaps are directly pronounced in the phonetic structure of an utterance, it seems likely that prosody-syntax correspondence rules may result in

phonological regularities in naïve speakers' productions of gap sentences, and that these regularities are available to help comprehenders locate gaps in *wh*-questions.

In this section, we have presented one means of addressing the laboratory speech problem. The SPOT task attempts to balance the need for tight syntactic control with the desire for a more spontaneous speaking style. Ideally, of course, we would investigate of a broad range of speaking styles: from the careful and perhaps self-conscious styles of psycholinguistically informed researchers to the multiplicity of speaking styles found in naïve speakers at various levels of formality, various rates, with various types of discourse participants and communicative goals, and so forth. Sentence processing research on prosody has just begun to investigate some of these types of variation, from the perspective of what speakers might do to aid listeners. We describe this research in section 4. First, though, we present an indication of the amount of prosodic variation we found within and across speakers in the fairly constrained sociolinguistic situation of the SPOT task.

3. MULTIPLE INTONATIONAL FORMS

Although speakers in the game task consistently used prosody to disambiguate these utterances, they varied in the contours used to supply the disambiguation. Within and across speakers, many different pitch accent, phrase accent, and boundary tone combinations were used for the same morphosyntactic structure. For the sequence *change the position of the square with the triangle* in the ambiguous PP attachment utterances, we found 63 distinct patterns on 79 high-attached utterances, and 87 patterns on 101 low-attached utterances. The discourse situation varied for these utterances (see section 4), which could account for the range of tunes. A similar result was found over the sequence *moves the square* in early and late closure utterances ((1) and (2) above). Full-sentence transcriptions showed 25 distinct intonational patterns for 35 early closure utterances, and 22 for 48 late closure utterances. As there was little variation in the discourse sequencing preceding these utterances, or in the discourse situation, this is an intriguing finding. Finally, the filler-gap ambiguities also showed a range of different tunes for the sequences *want (gap) to change the position of the square* and *want to change the position of (gap) this time*. The MAE data showed 27 distinct patterns for 44 early gap tokens, and 32 patterns for 110 late gap tokens. A random sample of 20 early and 20 late gap tokens from the NZE data showed 10 patterns for the early gap sentences, and 7 for the late gap.

Our transcriptions clearly show the many-to-many mapping between prosodic and syntactic structures. We do not find a canonical pronunciation for a particular syntactic form. Instead, the variation in intonation emphasises the fact that exact prosodic form cannot be predicted solely on the basis of morphosyntactic structure. Despite this variation, we consistently found that speakers most often pronounced the strongest sentence-internal prosodic break at a location in syntactic structure that would resolve ambiguity for listeners. So, while phrasal boundary location was largely predictable from syntactic form, the choice of phrase accent and boundary tone type at a phrase boundary was not. The pattern of high, low, and bi-tonal pitch accents across an utterance also showed substantial variation. We speculate that this variation is due to a combination of factors, including individual differences in speech style and attentional factors that varied from subject to subject and game to game.

There are clearly some implications arising from these observations for the question of how intonation is used in sentence processing. If listeners can use intonational distinctions to determine the speaker's intended syntactic structure, then they must be interpreting intonational form at a number of levels, as well as "filtering out" style and attentional factors. Because of the emphasis on laboratory-type speech in the research in this field, this set of problems is particularly under-researched, as will become apparent in our review in the next two sections. This is despite the relevance of such issues for questions concerning the overall architecture of the spoken language processing system, such as when and how prosodic and intonational information is used by the listener [47].

4. SPEAKERS HELPING LISTENERS?

One of the intriguing questions in the study of the relationship of intonation and syntax concerns the functional motivation for patterns of prosodic disambiguation. On the one hand, since many of the prosodic contrasts can be linked to syntactic contrasts, it is conceivable that the prosodic patterns stem from some cognitive mapping between the two structures [38]. For example, one recent study [33] found that speakers produced small but significant differences between two syntactic patterns even when they were (a) not instructed to disambiguate, (b) unaware of the syntactic contrast, (c) reading instead of speaking spontaneously, and (d) dealing with syntactic ambiguities that were pragmatically resolved within the sentence. One pair is shown in (6) and (7), contrasting high and low attachment of the final PP.

(6) The janitor removed the smudges on the weekend.
(7) The janitor removed the smudges on the window.

However, these effects were modulated by the information structure of the sentence. Prosodic cues to attachment diminished with contrastive focus on *janitor* and deaccentuation of the following material, suggesting that some of the variability within and across experiments may be due to non-syntactic influences on prosodic form. The prosodic reflection of phrase structure may be obscured or enhanced by the prosodic reflection of given, new, and contrastive information [45]. Extending the mapping idea, such results might indicate the speaker's mapping of syntactic, semantic, and pragmatic structures onto prosodic form.

On the other hand (and of course this need not preclude a syntax-prosody mapping in production), speakers may produce prosodic contrasts because they are sympathetic to the needs of listeners, i.e. to help disambiguate. They

may do this by systematically producing, in situations that require disambiguation, a particular kind of prosodic element to indicate a particular kind of syntactic boundary. For example, speakers might produce the strongest possible prosodic boundary to indicate the end of a clause whenever the clause boundary location is ambiguous in the speech situation. (See [40] for a discussion of the reliability of potentially infrequent cues to syntactic structure.) Another view [10] is that speakers provide variable yet helpful contrasts in their choice of prosody and listeners interpret prosodic elements as elements chosen to be helpful in indicating the syntactic structure. Rather than there being a fixed interpretation of prosodic elements, speakers choose prosodic elements on the basis of preceding material (and other factors, such as speech rate) and listeners compute the meaning of these elements in relation to that preceding material. (See further discussion in section 5.)

The research literature referred to in section 1 included studies that contrasted conditions where speakers were explicitly asked to disambiguate with those where they were not [1]. Awareness of ambiguity was argued to have a determining effect on prosodic disambiguation. The advantage of our game task in this context is that we can ask a similar research question without instructing the speakers that the utterances are ambiguous or should be disambiguated. In one study [36], we found that the degree of prosodic disambiguation of PP ambiguities did not vary according to the level of ambiguity in the situation. PP sentences (3) were uttered in three types of game context. In ambiguous contexts, it was possible to move either a combined square-and-triangle piece or a square, using a triangle to push it (the game rules did not allow the square to move on its own). In unambiguous contexts, one of the relevant pieces could not be moved, or it made no sense to move it since it was in its target finishing location. In biased contexts, one interpretation was more likely because of the recent history of the game – for instance, a triangle had just been shifted to a space next to a square, setting up a subsequent move in which the triangle was used to move the square. These three types of context offer different constraints on the likelihood of a particular interpretation, yet acoustic data from matched subjects showed no difference in the size of the contrast between high and low attached PPs. (Mean *square*+silence durations for high- vs. low-attachments were: ambiguous 822ms and 367ms; unambiguous 994ms and 439ms; biased 753ms and 330ms).

As a comparison, we included a further object – a cylinder – which could also be used to move the square, but which did not make up a combined game-piece with the square. The utterance in (8) is therefore unambiguous, unlike the corresponding form in (3):

(8) I want to change the position of the square with the cylinder.

Our durational analyses showed the degree of marking of the high-attached PP did not depend on whether the sentence was ambiguous (3) or not (8). Both had significantly longer *word*+silence durations than the low-attached version of (3), but did not differ from one another (high-attached *triangle* 857ms, high-attached *cylinder* 871ms, low-attached *triangle* 379ms).

Our phonological analysis of these utterances showed prosodic constituent boundaries consistent with syntactic structure, but again no differences due to situational ambiguity level. For 27 high-attached utterances spoken in ambiguous situations, 52% had the strongest break following *square*, 22% after a word other than *square*, and 26% had tied strongest breaks after *square* and another word. Twenty-five high-attached PPs in biased game situations showed a similar pattern, 48% had the strongest break after *square*, 20% after a word other than *square*, and 32% after *square* and other word(s). For 27 high-attached PPs from unambiguous game situations we also found breaks consistent with syntactic structure: 70% had the strongest break after *square*, 15% after a word other than square, and 15% after *square* and other word(s).

These analyses of our production data suggest that speakers produce largely the same prosodic contrasts regardless of the level of situational ambiguity. Recent follow-ups to this work have studied productions by naïve speakers in less complex discourse tasks than the gameboard task, with mixed results. One study [40] used a task in which the speaker uttered a series of commands involving the manipulation of a set of toys, to a listener separated by a screen. Participants did not interact in this task, apart from the speaker asking if the listener was ready. Experimental materials were presented as printed text and toy manipulations acted out by the experimenter. The textual stimulus was then removed and the command produced by the speaker from memory. Speakers received novel text before each utterance and performed minimal manipulations of the toys, a situation which makes this task something more of a reading situation than the game task described above. The results showed an effect of ambiguity on production. When speakers were asked to produce both high and low PP attachments and were shown toy displays consistent with either interpretation, they produced prosodic disambiguation. When speakers in a separate experiment only saw texts and toy scenarios consistent with just one of the interpretations, those who received the high attached materials did not produce significantly different prosodic patterns from those who received the low attached materials.

In contrast, a further study of PP ambiguity using no script whatsoever, but simply relying on participants producing appropriate sentence structures in response to a pictorial display, reports consistent prosodic disambiguation [23]. Although this method produced various sentence forms, roughly half the utterances produced by the participants had the targeted syntactic structure. For these, speakers reliably produced a prosodic distinction between modifier and goal interpretations of *Put the dog in the basket on the star*. These results obtained regardless of the level of situational ambiguity, that is, prosodic boundaries were produced to disambiguate syntactic intention when the

situation was ambiguous (e.g. objects included a dog in a basket, another basket on a star, and another star) or unambiguous (e.g. only a dog and a basket on a star).

5. INTONATION IN SENTENCE COMPREHENSION

Given that speakers often produce prosodic structures that reflect syntactic structures, we now turn to the question of how listeners make use of prosodic information to develop an interpretation of a sentence. Early research on prosody in sentence comprehension focused primarily on showing that prosody does indeed affect some syntactic parsing decisions. Syntactically ambiguous utterances were categorised according to whether or not they could be disambiguated, and reliable phonetic prosodic cues were identified in the disambiguated sentences. The more recent work we focus on here has shifted to investigations of how quickly prosodic information is used during sentence comprehension, the ways in which prosodic effects should be included in comprehension models, and the cross-linguistic comparison of the use of prosody.

Confirmation that the more spontaneous utterances that we have recorded in the SPOT project contain information that listeners can and do use in their sentence analysis comes from a series of off-line forced-choice tasks using the MAE recordings. Early/late closure and PP attachment utterances from our games were presented to naïve listeners in comprehension experiments. Consistent with results from studies of laboratory speech [22, 41, 48], we found that naïve listeners successfully use prosodic cues to categorise syntactically ambiguous materials.

Responses to early/late closure fragments truncated before the disambiguating word were sensitive to the location of the strongest prosodic break. Both when speakers intended an intransitive utterance, with the strongest break after *moves*, and when they intended a transitive utterance with the strongest break after *square,* listeners chose the appropriate completion 80% of the time. Listener performance was poorer when the speaker's intention was inconsistent with the position of the strongest break(s): 47% for intransitives with the strongest break after *square*, and 67% for transitives with the biggest break after *moves*. This pattern of results, with performance around or above chance when location of the break did not coincide with that of the subordinate clause boundary, illustrates an intriguing finding – that the location of the strongest break(s) is not the sole determinant of syntactic categorisation. Clearly the utterances are rich with additional information that guides listeners in their interpretation, presumably including tonal information not fully reflected in our break index analysis.

To test the use of the prosodic distinction in the PP utterances, we presented naive listeners with a large set of utterances like (3), namely all fluent PP utterances from our MAE data. 196 utterances from 15 speakers were presented to 19 listeners in a visually illustrated forced-choice task. Subjects chose between the instrumental meaning and the square+triangle meaning for each ambiguous utterance. The results showed a better than chance identification of the correct meaning: 76% and 64% for high and low attachment.

Taken together, the forced-choice categorisation data for these two sentence types confirm that the consistent prosodic patterns found in phonetic measures and phonological analyses can resolve syntactic ambiguity for naïve listeners. But such studies ask participants to choose between two possible sentence continuations after they have heard an utterance excised from its discourse context, and so may not give a clear indication of how participants use prosodic information during the course of hearing utterances. Some of our earlier studies investigated the immediate use of prosodic phrasing during the comprehension of closure ambiguities in an "on-line" cross-modal naming task. Participants listened to an auditory fragment and then saw a visual probe, which they named as quickly as possible. Example materials from one set of experiments [22, 41] are shown in (9) and (10), similar in structure to our sentence types in (1) and (2). The auditory fragments are underlined. The visual naming targets were the syntactically disambiguating words IS and IT'S. Three types of prosodic structure, co-operating, conflicting, and baseline, were compared for both early (9) and late closure (10) versions of the sentences. Co-operating prosody had the strongest syntactic break at the subordinate clause boundary. Conflicting prosody was created by cross-splicing, so that the strongest break was at the syntactic boundary consistent with the competing syntactic parse. The baseline condition had an identical and grammatical prosody for both early and late closure sentence versions, with contrastive stress on the subordinate clause subject, and deaccenting of the material in the ambiguous region (*checks the door*), giving ambiguous information about prosodic break location. All prosodic structures were confirmed by acoustic and auditory analyses, and relative acceptability of the combined syntactic and prosodic structures was confirmed by naïve listeners.

(9) Whenever the guard checks the door is locked.
(10) Whenever the guard checks the door it's locked.

The experiment showed strong sensitivity to the appropriate prosodic marking of clause boundaries. Co-operating prosody conditions showed fastest naming times and conflicting prosody conditions were processed more slowly than baseline conditions. Moreover, while both baseline and conflicting prosody conditions showed longer processing times for early closure sentences as compared to late closure sentences, this 'garden path effect' did not appear in the co-operating conditions. In separate experiments, the same pattern of effects obtained for both intonation phrase breaks followed by substantial silence, and intermediate phrase breaks which had no silent period following the phrasal tones. The lack of processing difficulty for the co-operating prosody conditions and the significant effects in general indicate that prosodic information is used quite early in the parsing process – as early as the appearance of the critical

morphosyntactic disambiguation, IS or IT'S.

A similar study [48] used as stimuli a set of fragments such as the underlined portions in (11) and (12) below, taken from utterances that had prosodic structures modelled on recordings from naïve speakers. The visual target ARE resolved the syntactic ambiguity as in (12). Again there was strong sensitivity to appropriate prosodic marking of clause boundaries. In addition, the study investigated the role of accent placement on words and phrases like *Hong Kong* in the example. In some contexts in English, including the citation context, such items have two full vowels and a late accent, on *Kong*. When the same items are immediately followed by an accented syllable, as in (11), the accent may be placed earlier; as in *HONG Kong PROblems* [15]. When the following accented syllable is separated by a major syntactic and prosodic boundary, as in (12), this stress "shift" does not occur. Our study found that early accent placement could be used as further evidence that the current constituent was incomplete. However, the use of this cue was sensitive to whether early stress is possible even without a following accented item, as in lexicalised shifted forms such as *HEATHrow* for the London airport.

(11) When parliament discusses Hong Kong problems, they take ages to solve.
(12) When parliament discusses Hong Kong, problems are solved instantly.

Additional recent work [9, 26, 30, 31] is consistent with this finding that listeners are sensitive to the placement of pitch accents in their interpretation of structural ambiguities. Initial work in this area has used paraphrase choice tasks to assess the effect of pitch accent type and placement on the recovery of spoken sentence interpretation. In one such study [9], listeners heard utterances such as *Bill took chips to the party and Susan to the game*, ambiguous between a gapping structure (Susan took chips) and a non-gapping structure (Bill took Susan). Participants were more likely to choose a paraphrase that matched the gapping structure if *Bill* and *Susan* carried comparable pitch accents (i.e. showed "prosodic parallelism") than if *chips* and *Susan* carried comparable accents, although the study showed an overall preference for the structurally simpler non-gapping structure. In a separate set of experiments [30], listeners heard utterances such as *I asked the pretty little girl who is cold*. When a pitch accent occurred on the wh- word *who*, listeners were more likely to choose a paraphrase indicating an embedded question interpretation. Without this accent, but with a nuclear accent on the final word of the utterance, *cold*, listeners were more likely to interpret the final wh-phrase as a relative clause modifying the preceding noun. These effects were not modulated by changes in prosodic phrasing such as the insertion of an IPh boundary immediately before the wh-phrase.

Prosodic prominence has been shown to influence syntactic attachment of an ambiguous relative clause in sentences such as *The investigator found the uncle of the businessman who was wanted by the police,* with the relative clause *who was wanted by the police* modifying either the first noun in the direct object noun phrase, *uncle*, or the second, *businessman*. Results showed that when either noun bore a H* pitch accent, the subsequent relative clause was more likely to be interpreted as modifying that noun [31]. A follow-up study extended these findings by examining the effects of a L+H* pitch accent on the noun, in three boundary conditions: no boundary, an IPh boundary after the complex noun phrase, or an ip boundary in this location [26]. The effect of pitch accent was replicated in all three boundary conditions, with listeners more likely to interpret the relative clause as modifying the accented noun than the unaccented noun. Results also showed that for both pitch accent conditions, a boundary after the complex NP consistently biased listeners to interpret the relative clause as modifying the first noun, while the absence of a boundary resulted in more modification of the second noun.

Pitch accent placement has also been shown to influence processing in relation to the assignment of focus relations and given/new status in a conversation. Theories of prosodic influence on focus assignment posit a relationship between pitch accent placement and focus, such that a constituent that contains focused new information will carry a pitch accent, while a constituent that contains old information that is not in focus will not. In these models, the relationship between focus and pitch accent is mediated by syntactic structure, such that an accent on the internal argument of a phrasal constituent can "project", resulting in the assignment of focus to the constituent as a whole [39]. Following a broad focus question (13), focus projection theories predict that (15) and (16) will be felicitous responses, because the verb phrase *teaches math* is new and bears prosodic focus in both (pitch accents are indicated by CAPITALs). However, (17) will not be felicitous, because only the verb *teaches* is prosodically marked as new by the nuclear L+H* accent. Similarly, after the narrow focus question (14), (17) will be felicitous, but (15) and (16) will not.

(13) Isn't Stacy pretty smart?
(14) Isn't Stacy good at math?
(15) Yes, she TEACHES MATH
(16) Yes, she teaches MATH
(17) Yes, she TEACHES math

For English sentence question-answer pairs like these, broad and narrow focus questions such as (13) and (14) respectively were followed by one of three responses that varied in the placement of L+H* pitch accents [6]. Results from acceptability judgement and speeded sentence comprehension tasks indicated that while listeners judged the double-accented (15) as a better answer than (16) to the broad focus question (13), response times to the two answers did not differ, and were faster than those for (17), which was rated as a poor response. Thus, listeners were sensitive to L+H* pitch accent placement in both on- and off-line tasks that involved the evaluation of prosody-based focus assignment in English. However, the response time measure indicated that (15) and (16) are of

comparable processing difficulty, consistent with the predictions of focus projection theories.

On-line studies such as these provide strong evidence that prosodic phrasing and pitch accent placement have immediate effects on structural ambiguity resolution, and that the presence of disambiguating prosodic information at syntactic choice points can preclude the effects of temporary syntactic ambiguity during sentence comprehension. Additional evidence for the immediacy of prosodic effects on syntactic processing comes from recent eye tracking studies of PP attachment in spoken sentences [40]. Listener eye movements to objects were recorded while a speaker pronounced instructions containing ambiguous PP attachments similar to those in (3), e.g. *tap the frog with the flower*. Results showed that the speaker's prosodic phrasing influenced eye movements to objects immediately prior to the onset of the ambiguous PP. When the speaker pronounced a prosodic break following *tap*, listeners looked immediately to an aggregate frog-and-flower object, suggesting that the presence of the early boundary unambiguously indicated an upcoming low-attached PP.

Another important area in sentence comprehension research concerns how a semantic interpretation of a sentence is developed – i.e., not just a syntactic parse of a sentence, but the integration of the meanings of words and phrases into a sentence-level message. Prosodic phrasing also appears to influence this type of processing. A series of experiments investigated this semantic integration by examining the use of sentence context to bias the interpretation of lexically ambiguous words like *bank* and *glasses*. Listeners show evidence of further semantic processing of ambiguous words and their surrounding sentential material in initial clauses that end with an IPh boundary than in ones that end with an ip boundary, even when durational differences are controlled [18, 29, 35, 51]. Although some integration of the ambiguous words and context must take place immediately, the integration is facilitated by the presence of a following IPh boundary.

The studies reviewed above, like most published studies on intonation and sentence processing, used English materials. Since syntactic and prosodic structures vary considerably between languages, general claims about intonation in sentence processing should also be tested with studies of other languages. Recent work on the influence of IPh boundaries in Korean allows us to examine prosodic effects in a language with substantially different syntactic forms, in that it is a relatively free word order, pro-drop and head-final language. Korean and English have in common that IPh breaks are indicated by phrase final boundary tones and silence, and that these breaks often coincide with syntactic clause boundaries.

Sentences like those in (18) and (19) were presented in a speeded forced-choice task [21]. Until the third word, (18) and (19) are ambiguous between a gap type relative clause, as in (18), and pro-type relative clause, as in both interpretations of (19). While (18) is a temporary ambiguity resolved at the head noun of the relative clause, 'nephew', (19) remains ambiguous through the end of the sentence. Results showed IPh boundaries can resolve both temporary and global syntactic ambiguity in Korean. For temporary ambiguities (18), listeners consistently answered a question such as 'who killed himself?' with 'Mira.' However, these responses were faster when a HL% IPh coincided with the syntactic clause boundary after the initial NP than when there was no such IPh boundary. For global ambiguities (19), a HL% IPh boundary after the first NP biased listeners toward the interpretation that this NP was the subject of the main verb, while the absence of this boundary led to a meaning with the dropped pro as the main verb subject. For ambiguous sentences, response times were faster when listeners chose the interpretation that was consistent with prosodic structure than when they chose the response where prosodic and syntactic structures conflicted.

(18) Mira-ka casalha-n cokha-rul mwutesse.
Mira-NOM e_i suicide-REL nephew-ACC buried
'Mira buried the nephew who killed himself.'

(19) Mira-ka casalha-n cari-rul poasse.
Mira-NOM e_i suicide-REL place-ACC saw
a. 'Mira saw the place where (pro) killed him/herself.'
b. '(pro) saw the place where Mira killed herself.'

These results are consistent with the general pattern laid out above for English: Processing is easiest when there is similarity in prosodic structure and syntactic structure, in this case, an intonational phrase boundary at a clause boundary. Yet the two cases differ in an important respect. In English, comprehension was facilitated by a prosodic boundary at the boundary between two clauses: the end of a subordinate clause and the start of the matrix clause. In Korean, comprehension was facilitated by a prosodic boundary located in the *middle* of the matrix clause and at the start of a relative clause. Thus, the Korean results cannot be explained by a simple parsing strategy that uses a major prosodic boundary as a signal to close a clause.

Korean and English differ intonationally in the structures within the intonational phrase. Korean employs accentual phrases instead of intermediate phrases and has no pitch accents [19, 20]. Accentual phrases tend to be smaller in span than intermediate phrases of English, typically with just one lexical item plus its following case markers or postpositions. They may also be acoustically less salient, since they rarely exhibit final lengthening or pausing at the prosodic boundary. Some flexibility in accentual phrasing exists, however. Sentences containing phrases such as [hjʌnmjʌŋhan akiɬi ap'a] ('wise baby's daddy') were tested in a cross-modal naming task and a fragment completion task [32, 34] that varied accentual phrasing. An accentual phrase boundary was present or absent between the adjective and N1, and between N1 and N2. This provided four groupings, one intended to bias toward each of two possible interpretations ('the daddy of a wise baby', 'the wise daddy of a baby') plus two more neutral groupings, one with neither medial boundaries and one with both medial boundaries. Results showed immediate effects of the presence versus absence of accentual

boundaries on naming times in the cross-modal study and effects of boundary placement on interpretation in the fragment completion study. Consistent with English results above, the postulation of a syntactic unit appeared more likely (and easier to process) given consistent prosodic phrasing. The prosodic information was apparently processed quite rapidly, showing incremental effects on syntactic parsing decisions.

Two findings of a different sort come from research on German. One found evidence of the immediate processing of intonational phrase boundaries in event-related brain potentials [43]. As in an English case described earlier [22], the study found that a major intonational boundary in the appropriate location could resolve a syntactic ambiguity toward the interpretation dispreferred in reading contexts. The second study found that reanalysis of sentence structure is easier when the prosodic structure is compatible with the new interpretation [4].

The relationship between focus, prosodic contour, and comprehension has been investigated in Tokyo Japanese, in an adaptation of the English experiment reviewed earlier [17]. In contrast to optional use of pitch accent in English, Tokyo Japanese has lexically-assigned pitch accent, with accentual phrases obligatorily accented or unaccented depending on lexical content. Focus is conveyed by modifying this lexically determined pitch contour, so that words that follow the focused word exhibit attenuation of their lexically-assigned pitch accent. Despite dramatic differences in these prosodic systems, English and Japanese studies show strikingly similar results. For both languages, findings are consistent with the notion of focus projection – focus on a verbal object was interpreted as conveying focus to the entire verb phrase. Utterances with intonational prominence on just the internal argument were accepted as answers to broad focus questions, and response time did not differ from cases where intonational prominence was on both the internal argument and the verb. As with the English materials, judgement data showed that listeners preferred dual prominence utterances over those with prominence just on the internal argument. Again replicating the English findings, response time measures showed that intonational prominence on the verb was optional in conveying broad focus. One interesting discrepancy between the English and Japanese results was that Japanese listeners accepted utterances with reduced tonal movement for the internal argument as equally good answers to the broad focus question, even though the questions would have been most felicitously answered with intonational prominence on the argument. This difference can be attributed to preserved tonal prominence on internal arguments in the pre-focal position in Japanese. Thus, listeners and speakers of both languages used intonation to establish focal status, but the effects did not obscure language-specific lexical specifications.

6. DISCUSSION

The relationships between the syntactic and prosodic structures of sentences and between these and sentence processing are clearly far from straightforward. We started this paper by pointing out that there is a complex interaction of physical properties that makes up what listeners perceive as patterns of intonation and prosody, demanding a phonological approach in addition to a phonetic analysis of relevant parameters. Additionally, we emphasised the difficulty in obtaining recordings of a natural enough character that nevertheless allow us to compare prosodic patterns on minimally-contrastive sentence structures of the type explored in processing research. Research in our own laboratories and elsewhere has identified that there are phonetic and phonological patterns of intonation and prosody that correspond to differences in syntactic analysis, and that although the correspondence could be claimed to be rather loose, it nevertheless persists across different degrees of situational ambiguity. The research also shows that listeners can use these differences to select an appropriate structural analysis of ambiguous utterances, and that both prosodic boundary marking and accent placement facilitate the immediate processing of such ambiguities, as reflected in response time and eye-tracking studies. Studies have shown the use of intonational boundaries and pitch accents in the processing of syntactic and focus assignment ambiguities in a range of languages, but with differences in the implementation of correspondences between prosodic and other linguistic levels of representation, reflecting differences in the character and distribution of the prosodic features involved.

Current research on intonation and sentence processing continues to provide evidence that there is no clear simple and direct correspondence between syntactic and prosodic structure. Intonation properly forms part of prosodic structure, not part of syntactic structure. What, then, is the mechanism by which the intonational structure of utterances is used in sentence processing? A number of possibilities have been proposed (see [29] for more detailed argument). One rather simple mechanism gives primacy to syntactic analysis, claiming that prosodic structuring supplements syntactic analysis by marking the location of clause boundaries, as a sort of spoken punctuation [25]. Some of the evidence cited above is compatible with such a position, but the position is weakened by the observation that any one utterance may have a number of different possible prosodic intonational realisations, including phrasings that are not simply determined by grammatical considerations. Other mechanisms respond more directly to the prosodic constituency of an utterance, for instance by allowing prosodic boundaries to determine selection between a number of alternative syntactic interpretations – the prosodic boundaries noted in closure ambiguities could feasibly operate in this way [8]. An approach that is potentially more sensitive to the hierarchy of prosodic structures appeals to a notion of prosodic visibility [29], with structural decisions about an utterance made on the basis both of the lexical content of the utterance and of how that content is grouped into domains by the prosodic structure provided by the speaker. Prosody thus forms part

of an enriched input to the sentence processing system.

What is the nature of this enriched input and how is it used? Several articles have argued for effects of prosodic boundaries located prior to a syntactic choice point on the subsequent parsing decision [10, 29, 32, 37]. Such findings point to the need either for interpretation of one prosodic element with respect to preceding ones [10] or for incremental effects of prosodic phrasing on syntactic parsing, whether or not there are multiple syntactic options available at the point of the prosodic boundary [29]. They converge with findings showing effects of subtle intonational and prosodic contrasts, and those suggesting that a prosodic structure is available to influence reanalysis, to support the notion that the parsing system builds or somehow encodes a prosodic structure in addition to a syntactic structure.

A number of important questions remain in this area, which we simply list here as indications of future research. What is the effect on structural interpretation of different tonal choices, whether through the role of different pitch accents in focal projection or in attracting attachments, or through that of different edge tones in establishing prosodic phrases? How are the effects of pitch accents in determining parallel interpretative structures weighed up against other functions of accents as markers of focus, information status, etc? What are the effects of pitch range and dynamism? To what extent do different speech styles change the production and comprehension of intonation? Do listeners give less weight to boundaries for attachment decisions when an utterance is produced with multiple prosodic boundaries, for instance for stylistic reasons? How do we best model the small but significant effects of prosody in some structures vs. the large effects of prosody in others? And how do prosodic biases interact with other information sources, such as verb bias and the semantic features of lexical items to determine sentence comprehension?

ACKNOWLEDGEMENTS

This work was supported by NIH research grants DC-00029 and MH-51768, NSF grant 0088175, NZ/USA Cooperative Science Programme grant CSP95/01, and Marsden Fund grant VUW604.

REFERENCES

[1] D.W. Allbritton, G. McKoon and R. Ratcliff, "Reliability of prosodic cues for resolving syntactic ambiguity," *Journal of Experimental Psychology: Learning, Memory, & Cognition,* 22, 714-735, 1996.

[2] A.H. Anderson et al, "The HCRC map task corpus," *Language and Speech,* 34, 351-366, 1991.

[3] G. Ayers, "Discourse functions of pitch range in spontaneous and read speech," *OSU Working Papers in Linguistics,* 44, 1-49, 1994.

[4] M. Bader, "Prosodic influences on reading syntactically ambiguous sentences," in *Reanalysis in sentence processing,* J.D. Fodor and F. Ferreira Eds., Dordrecht: Kluwer Academic Publishers, 1998.

[5] M.E. Beckman and J. Hirshberg, *The ToBI annotation conventions*, Columbus, OH: Ohio State University, 1994.

[6] S. Birch and C.E. Clifton, "Focus, accent, and argument structure: Effects on language comprehension," *Language and Speech,* 38, 365-391, 1995.

[7] E. Blaauw, "The contribution of prosodic boundary markers to the perceptual difference between read and spontaneous speech," *Speech Communication,* 14, 359-375, 1994.

[8] E.J. Briscoe, *Modelling human speech comprehension: a computational approach*, New York: Wiley, 1987.

[9] K. Carlson, "The effects of parallelism and prosody in the processing of gapping structures," *Language and Speech,* 44, 1-26, 2001.

[10] K. Carlson, C. Clifton and L. Frazier, "Prosodic boundaries in adjunct attachment," *Journal of Memory and Language,* 45, 58-81, 2001.

[11] H. Clark and M.F. Schober, "Understanding by addressees and overhearers," in *Arenas of Language Use,* H. Clark Ed., 176-197, Chicago: University of Chicago Press, 1992.

[12] H.H. Clark and D. Wilkes-Gibbs, "Referring as a collaborative process," *Cognition,* 22, 1-39, 1986.

[13] W.E. Cooper and J. Paccia-Cooper, *Syntax and Speech*, Cambridge, MA: Harvard University Press, 1980.

[14] A. Cutler, D. Dahan and W.v. Donselaar, "Prosody in the comprehension of spoken language: A literature review," *Language and Speech,* 40, 141-201, 1997.

[15] E. Grabe and P. Warren, "Stress shift: do speakers do it or do listeners hear it?," in *Phonology and phonetic evidence: Papers in Laboratory Phonology IV,* B. Connell and A. Arvaniti Eds., 95-110, Cambridge: Cambridge University Press, 1995.

[16] P. Howell and K. Kadi-Hanifi, "Comparison of prosodic properties between read and spontaneous speech," *Speech Communication,* 10, 163-169, 1991.

[17] K. Ito, *The interaction of focus and lexical pitch accent in speech production and dialogue comprehension: Evidence from Japanese and Basque,* PhD, Univ. Ill., Urbanna-Champaign.2002.

[18] S.R. Jackson, T.G. Bever and D.J. Townsend, "Effects of intonation and subordinating conjunction on lexical ambiguity resolution," poster presented at Sixteenth Annual CUNY Conference on Human Sentence Processing, 2003.

[19] S.-A. Jun, *The Phonetics and Phonology of Korean Prosody: Intonational Phonology and Prosodic Structure*, New York: Garland, 1996.

[20] S.-A. Jun, "The accentual phrase in the Korean prosodic hierarchy," *Phonology*, 189-226, 1998.

[21] S. Kang and S.R. Speer, "Prosody and clause boundaries in Korean," in *Speech Prosody 2002,* I. Marlien Ed. Aix-en-Provence 2002.

[22] M.M. Kjelgaard and S.R. Speer, "Prosodic facilitation and interference in the resolution of temporary syntactic closure ambiguity," *Journal of Memory and Language,* 40, 153-194, 1999.

[23] T. Kraljic and S. Brennan, "Use of prosody and optional words to disambiguate utterances," paper presented at 43rd annual meeting of the Psychonomic Society, Kansas City, Missouri, 2002.

[24] W.J.M. Levelt and A. Cutler, "Prosodic marking in speech repair," *Journal of Semantics,* 2, 205-217, 1983.

[25] M. Marcus and D. Hindle, "Description theory and intonation boundaries," in *Cognitive models of speech processing: psycholinguistic and computational perspectives,* G.T.M. Altmann Ed., 483-512. Cambridge, MA: MIT Press, 1990.

[26] L.A. Maynell, "Prosodic Effects on Relative Clause Attachment," poster presented at the Thirteenth Annual CUNY Conference on Human Sentence Processing, La Jolla, CA, 2000.

[27] H.N. Nagel, L.P. Shapiro and R. Nawy, "Prosody and the processing of filler-gap sentences," *Journal of Psycholinguistic Research,* 23, 473-485, 1994.

[28] J.B. Pierrehumbert & J. Hirshberg, "The meaning of intonational contours in the interpretation of discourse", *in Intentions in Communications*, P. Cohen, J. Morgan, and M. Pollack, Eds, 271-311. MIT, 1990.

[29] A.J. Schafer, *Prosodic Parsing: The Role of Prosody in Sentence Comprehension,* Doctoral dissertation, University of Massachusetts, 1997.

[30] A.J. Schafer, K. Carlson, C. Clifton and L. Frazier, "Focus and the interpretation of pitch accent: Disambiguating embedded questions," *Language and Speech,* 43, 75-105, 2000.

[31] A.J. Schafer, J. Carter, C. Clifton and L. Frazier, "Focus in relative clause construal," *Language and Cognitive Processes,* 11, 135-163, 1996.

[32] A.J. Schafer and S.-A. Jun, "Accentual phrasing and syntactic disambiguation in Korean," poster presented at Seventh Conference on Laboratory Phonology, Nijmegen, The Netherlands, 2000.

[33] A.J. Schafer and S.-A. Jun, "Effects of focus on prosodic reflections of phrase structure in American English," paper presented at Prosody in Processing Workshop, Utrecht University, Utrecht, Netherlands, 2001.

[34] A.J. Schafer and S.-A. Jun, "Effects of accentual phrasing on adjective interpretation in Korean," in *Sentence Processing in East Asian Languages,* M. Nakayama Ed., 223-255, Stanford: CSLI Publications, 2002.

[35] A.J. Schafer and S.R. Speer, "Prosodic phrasal effects on the resolution of within-category lexical ambiguity," poster presented at the Eleventh Annual CUNY Conference on Human Sentence Processing, New Brunswick, NJ, 1998.

[36] A.J. Schafer, S.R. Speer and P. Warren, "Prosodic influences on the production and comprehension of syntactic ambiguity in a game-based conversation task," in *Approaches to Studying World-Situated Language Use: Bridging the Action and Product Traditions,* J.C. Trueswell and M.K. Tanenhaus, Eds., forthcoming.

[37] A.J. Schafer, S.R. Speer, P. Warren and S.D. White, "Intonational disambiguation in sentence production and comprehension," *Journal of Psycholinguistic Research,* 29, 169-182, 2000.

[38] E.O. Selkirk, *Phonology and syntax: the relation between sound and structure*, Cambridge, MA: MIT Press, 1984.

[39] E.O. Selkirk, "Sentence prosody: Intonation, stress and phrasing," in *Handbook of phonological theory,* J. Goldsmith, Ed., 550-569. Oxford: Blackwell, 1994.

[40] J. Snedeker and J.Trueswell, "Using prosody to avoid ambiguity: Effects of speaker awareness and referential context," *Journal of Memory and Language*, to appear.

[41] S.R. Speer, M.M. Kjelgaard and K.M. Dobroth, "The influence of prosodic structure on the resolution of temporary syntactic closure ambiguities," *Journal of Psycholinguistic Research,* 25, 247-268, 1996.

[42] S.R. Speer, A.J. Schafer and P. Warren, "Prosodic disambiguation in naïve speakers' production of temporary syntactic ambiguity," paper presented at The University of Southern California Speech Production Conference, Los Angeles, CA, 1999.

[43] K. Steinhauer, K. Alter and A.D. Friederici, " Brain potentials indicate immediate use of prosodic cues in natural speech processing," *Nature Neuroscience,* 2, 191-196, 1999.

[44] K. Straub, C. Wilson, C. McCullum and W. Badecker, "Prosodic structure and wh-questions," *Journal of Psycholinguistic Research,* 30, 379-394, 2001.

[45] J.J. Venditti and J. Hirschberg, "Intonation and discourse processing," paper presented at 15th International Congress of Phonetic Sciences, Barcelona, Spain, 2003.

[46] P. Warren, *The Temporal Organisation and Perception of Speech,* Doctoral dissertation, University of Cambridge, 1985.

[47] P. Warren, "Prosody and language processing," in *Language Processing,* M.J. Pickering Ed, 155-188. Hove: Psychology Press, 1999.

[48] P. Warren, E. Grabe and F.J. Nolan, "Prosody, phonology and parsing in closure ambiguities," *Language and Cognitive Processes,* 10, 457-486, 1995.

[49] P. Warren, A.J. Schafer, S.R. Speer and S.D. White, "Prosodic resolution of prepositional phrase ambiguity in ambiguous and unambiguous situations," *UCLA Working Papers in Phonetics,* 99, 5-33, 2000.

[50] P. Warren, S. Speer and A. Schafer, "wanna-contraction and prosodic disambiguation in US and NZ English," paper presented at Australian Linguistics Society, Melbourne, 2000.

[51] P. Welby, "Effect of two levels of intonational phrasing on sentence processing," paper presented at Prosody in Processing, Utrecht, 2001.