

Laurie Maynell
Dissertation Proposal
June 30, 2005

Introduction

A primary question in sentence processing research is this: Does the human sentence processing mechanism use similar strategies regardless of the language being understood, or do facts about languages change the way language comprehension operates? Given that all humans have structurally similar brains and acquire language in similar fashion regardless of the language being acquired, it is reasonable to assume that there are at least a few general, universal principles employed for processing all languages.

One way in which this question has been addressed has been by experimentally examining the resolution of a relative clause attachment ambiguity in multiple languages. An example of the ambiguity is shown in (1) (from Cuetos & Mitchell, 1988).

(1) Someone shot the servant of the actress who was on the balcony.

There are two grammatical interpretations of this sentence. In the first, the relative clause *who was on the balcony* modifies the structurally higher noun *servant*. In the second, the relative clause modifies the structurally lower noun *actress*. The parsing principle of Late Closure (Frazier, 1978; Frazier & Clifton, 1996) predicts that the relative clause should be preferentially interpreted as modifying the lower noun, *actress*. Using a written questionnaire in which people were asked comprehension questions about such sentences, Cuetos and Mitchell (1988) found that attachment preference varied by language: English speakers preferred the higher noun *servant* as the head of the relative clause, whereas Spanish speakers preferred the lower noun *actress* as the head of the relative.

Since then, a number of similar reading studies to examine attachment preferences in several languages have been carried out and they demonstrate that attachment preferences vary from one language to another (Carreiras & Clifton, 1993, Gilboy et al., 1995, Brysbaert & Mitchell, 1996, Gibson et al., 1996, Ehrlich et al., 1999, Fernandez and Bradley, 1999, Miyamoto et al., 1999, Fernandez, 2000; for reviews see Cuetos et al., 1996, Mitchell & Brysbaert, 1998). In reading studies, as summarized in Fodor, (2002 a, b), Croatian, Dutch, Japanese, Spanish and French speakers interpret the relative clause as modifying the higher noun more often, while English, Norwegian, Swedish, Romanian, and Arabic speakers interpret the relative clauses as modifying the lower noun more often.

Such results suggest that the parsing principle of Late Closure might not be a universal parsing principle. Various sentence processing models have been proposed to account for these different attachment preferences in reading (Fodor, 2002 a, b; Mitchell & Brysbaert, 1998), some of which preserve Late Closure as a universal parsing principle.

In addition to studies using written language, this relative clause attachment ambiguity has been investigated using spoken language. The majority of these are studies in language production.

One hypothesis that seeks to connect the apparent differences in attachment preferences across languages found in reading with studies done on spoken language production of such sentences is the Implicit Prosody Hypothesis (IPH) of Fodor (2002 a, b). According to the IPH, "in silent reading, a default prosodic contour is projected onto the stimulus, and it may influence syntactic ambiguity resolution. Other things being equal, the parser favors the syntactic analysis associated with the most natural (default) prosodic contour for this construction" (Fodor, 2002 a, p. 84).

According to Fodor's explanation, the observed cross-linguistic differences in resolving this relative clause attachment ambiguity are a result of the fact that intonation systems differ from one language to another. Since it is not possible to hear inside a reader's head while they read, Fodor assumes that the "default" prosody is identical to the overt prosody a naïve speaker would produce upon his initial read of a sentence (after reading it silently once first). Several production studies (reviewed in Fodor, 2002 b) address what the default prosody of this utterance type is in various languages. Fodor maintains that prosody is secondary to structural requirements on the linguistic output. For example, in these studies it is often observed (e.g. in French, and Croatian) that speakers produce a prosodic phrase boundary prior to long relative clauses, but not prior to short ones. Another interesting example is that in Croatian, a property of the preposition *od* (of), and possibly of Croatian prepositional phrases in general, is that it is preceded by a prosodic phrase boundary in spoken conversation, and this was confirmed in production studies (see Lovrić, et al., 2000, 2001; Lovrić, 2003).

In a further test of the IPH, Jun (2003) collected production data from 4-7 speakers each of seven different languages, both head-initial (English, Spanish, French, Greek) and head-final (Japanese, Korean, Farsi). The goal of this study was to collect different productions of the structures like (1) in several different languages, and to address what the default prosody would be for each. In one pronunciation condition, speakers were asked to read each sentence out loud *without reading it silently first*, and then answer a comprehension question about it. Generally, in all languages tested, if speakers produced a prosodic boundary between the two nouns, the result was a prosodic grouping of the noun nearest the relative clause with the relative clause, and (based on the answer to the subsequent comprehension question) this was meant by readers as attachment to the structurally low noun (e.g. *actress* in (1)). If a boundary was produced between the relative clause and the nearest noun (the structurally lower noun), attachment to the higher noun was intended. This study demonstrates that speakers can use prosodic phrasing both to group words that belong together as well as distinguish different intended meanings in languages that are structurally and prosodically different.

Importantly, my underlying assumptions for this dissertation differ from Fodor's. Although the idea of a language's intonation system influencing attachment preferences in silent reading is not an unreasonable one, I disagree with the assumptions underlying her (and Jun's) testing methods. First, silent prosody is not necessarily identical to overt prosody, and in the absence of a way to hear silent prosody, neither her assumption nor my counter-assumption can be tested in reading studies. Second, producing an initial reading of a sentence without examining it silently at first (as in Jun, 2003) might result in a production of an intonation contour, but this is not necessarily a contour found in natural spoken language, since in natural spoken language, prior planning accompanies language production, and utterances are produced with an intended meaning. In

addition, in such a study, there is potential for repetition priming when many instances of the same structure are read over and over. Third, I assume that there is no such entity as a "default" prosodic contour. Even when speakers are allowed to read a sentence silently at first, before reading it out loud (as described in Fodor 2002 b), speakers have some kind of context in mind when producing the sentence. Prosodic structure conveys meaning, and intonation patterns are felicitous within some contexts, infelicitous in others. Further, there are many influences on the prosody of sentences in context. For example, contrastive stress placed on a word in a sentence in English often results in a boundary following the word, to further highlight the contrastive nature of that word (Beckman, 1996). Finally, while some particular lexical items may call for a prosodic phrase boundary to be produced at a specific location (Lovrić, 2003), my assumption is that prosodic structure is primary in shaping both the intended meaning of a linguistic output and the recovery of its meaning by listeners, as research in spoken language comprehension using this relative clause ambiguity strongly suggests (Schafer et al., 1996, Maynell, 2001, Clifton, Carlson & Frazier, 2002).

A crucial element that contributes to a listener's construction of the mental representation of a spoken utterance is the intonation or prosody of that utterance. In both English and Japanese, prosodic phrasing can indicate the appropriate grouping of words for comprehension (Beckman, 1996). Previous research suggests that the way an utterance like (1) is phrased will help the listener disambiguate it, and that phrasing will work similarly in the two languages. Price et al. (1991) demonstrated that for a high vs. low attachment ambiguity, presence of a prosodic break preceding an ambiguously modifying phrase corresponded to higher rather than lower attachment of that phrase to previous material. This was true for both production and perception studies. When speakers read sentences following a disambiguating context, their strategy for conveying high attachment of a subsequent ambiguously attached phrase was to produce the sentence with an intonation phrase break preceding it. So, for sentence (2) (from Price et al., 1991), a prosodic break preceding the constituent *in German* was produced by speakers when it was meant to be conveyed that the reading was done in the German language (high attachment of the PP at the VP).

(2) I read a review of nasality in German.

By contrast, either a small or no prosodic break at this location was produced when the speaker intended that the review was about nasality in German but not necessarily written in the German language (in this case, the PP would be sister to N' in the noun phrase *nasality in German*). In a subsequent perception study these same recorded sentences were successfully disambiguated by listeners. Price et al. (1991) concluded that presence or absence of an intonational phrase boundary before the ambiguously attached constituent helped listeners to resolve the high vs. low attachment ambiguity.

Building on this result, Maynell (2001) demonstrated the same effect of a prosodic phrase boundary in sentences with a relative clause ambiguity such as (1). When sentences were pronounced with a clear prosodic phrase boundary between the second (lower) noun and the relative clause, listeners interpreted the relative clause as modifying the structurally higher noun (e.g. *servant* in (1)). In the absence of a prosodic phrase boundary, listeners interpreted the relative clause as modifying the structurally lower noun (*actress*).

Finally, an additional production study (Venditti, 1994) showed a similar effect of prosodic phrasing in Japanese. In one experiment, Venditti focused on sentences consisting of matrix and subordinate clauses, which also had adjuncts with ambiguous scope of modification. An example is given in (3).

(3) *kyonen* e_j e_i *anda* *erimaki_i-ga* *nusumareta*.
 Last year knitted scarf_i-NOM was stolen.

‘The scarf that I knitted last year was stolen.’ LOW ATTACHMENT
 ‘The scarf that I knitted was stolen last year.’ HIGH ATTACHMENT

The low attachment interpretation (=attachment of the adjunct *last year* to the subordinate clause *that I knitted*) and the high attachment interpretation (= the adjunct *last year* modifies the matrix clause *the scarf was stolen*) are both possible. In analyzing the prosodic patterns obtained in a production study, Venditti observed a major phrase boundary (intonation phrase) and often a pause (i.e. a region of silence) between the adjunct (*kyonen*) and the relative clause (e_j e_i *anda*) in the high attachment case. This result is consistent with the results found for comprehension of English ambiguous sentences (Price et al., 1991, Maynell, 2001, Clifton et al., 2002) as well as for Jun’s (2003) Japanese production data in which prosodic phrase boundaries were produced following the ambiguously modifying constituent when it was meant to modify the structurally higher candidate for attachment.

Proposed Research

Because of the emphasis on production in previous spoken language studies (and its possible relationship to comprehension in written language), the focus of this dissertation will be on the comprehension of the relative clause attachment ambiguity in spoken language. I will examine this relative clause ambiguity in the head-final language, Japanese, and the head-initial language, English, to critically address the role of prosodic phrasing in sentence comprehension in two structurally different languages. It is my hypothesis that prosodic boundaries will function similarly in the comprehension of this sentence type and in the resolution of its ambiguity in both languages. That is, I aim to show that maintaining a correspondence between prosodic structure and syntactic structure is a universal parsing mechanism; further, such structuring of a sound stream may be a universal cognitive constraint that is applied to (among other sound phenomena) language.

To facilitate such a comparison of parsing in both languages, the sentences used in each language will be translations of those in the other. Example test items are shown in (4) and (5).

(4) The policeman broke the **window** of the **building** [where the thief had placed a mark].

(5) *KeekaN-wa* [*doroboo-ga* *maaku-o* *tsuketa*] **biru-no** **mado-o** *watta*.
Policeman-NOM *thief-NOM* *mark-ACC* *placed* *building-of* *window-ACC* *broke*.

Like (1), sentence (4) and its translation (5) are ambiguous: in each, the relative clause can be understood as describing either the **window** or the **building**. In English, the relative clause (in brackets) follows the two nouns that it may modify, whereas in Japanese, the relative clause precedes the two nouns. Also, the linear order of the nouns *window* and *building* in Japanese is opposite their order in English. In Japanese, listeners hear the word *building* (with a genitive marker) before they hear the word *window*.

Upon hearing the relative clause in English, the listener must consult her memory for the word being described by it. In Japanese, the relative clause must be retained in memory until the listener encounters the word that it describes later in the sentence. Thus, listeners in both languages ultimately build the same mental representation of the sentence, despite substantial word order differences. To complete comprehension, listeners also must resolve the ambiguity. The auditory memory for these spoken utterances is structured by prosodic phrasing and prominence, so it makes sense that the prosodic structure of the utterances will guide their interpretation by listeners (Speer, Crowder, & Thomas, 1983; Kjelgaard & Speer, 1995).

Three different prosodic phrase conditions will be tested in each language. In one condition, a prosodic phrase boundary will separate the relative clause from the adjacent, structurally lower noun. In a second condition, a prosodic phrase boundary will separate the two possible attachment sites from each other, effectively grouping the structurally lower noun with the relative clause. In a third condition, both prosodic phrase boundaries will bracket the structurally lower noun, separating it from both the relative clause and the structurally higher noun. The prosodic phrase boundary sites (marked by //) are illustrated in (6) and (7) for the critical regions of (4) and (5). The attachment sites are in boldface and the relative clause is in brackets.

(6) the **window** // of the **building** // [where the thief had placed a mark].

(7) [doroboo-ga maaku-o tsuketa] // **biru**-no // **mado**-o watta.
thief-NOM mark-ACC placed // building-of // window-ACC broke.

These prosodic phrase boundary locations have been chosen based on the research described in the previous section, as well as on work done by Clifton et al. (2002), who formulated the Informative Boundary Hypothesis. This hypothesis is paraphrased in (8); note that the hypothesis is formulated based on the head-initial, right-branching language English, since the studies were done in English.

(8) In a structure [A] [B] [C], where [C] can be structurally attached to [A] or [B]:

- a) A prosodic boundary immediately prior to [C] that is large relative to a boundary preceding [B] but not preceding [A] encourages high attachment of [C] to [A].
- b) A boundary immediately preceding [C] that is small relative to a boundary before [B] but not [A] encourages low attachment to [B].

Clifton et al. (2002) tested several ambiguous structures, among them the relative clause ambiguity of interest here. The constituent [A] represents the structurally high noun (*window*),

[B] the structurally low noun (*building*), and [C] the relative clause. When Clifton et al. refer to a boundary as "large," it means "higher in the prosodic hierarchy," so for example intonation phrase boundaries are larger than intermediate phrase boundaries (see Beckman, 1996), and any boundary is larger than no boundary. Clifton et al. (2002) did test boundaries of equal strength for some sentence types, including the relative clause attachment ambiguity, and compared comprehension results to those cases in which one boundary was stronger than the other. Generally, they found that sentences with boundaries of equal strength showed less high attachment than the case described in (8a), and less low attachment than the case described in (8b).

Because I am comparing two languages, I will test prosodic boundaries of equal strength (specifically, intonation phrase boundaries in both languages) in order to establish a baseline result for this syntactic structure in both languages. The intonation system of Japanese is such that manipulating the strength of boundaries is not trivial (for example, see Hirose, 2003), and would require a new set of materials in which lexical accent is carefully controlled in critical regions of the utterances. Additionally, my aim is to find support for a hypothesis about prosodic phrasing that is universal and not language-specific. Note that the Informative Boundary Hypothesis could be rephrased for a head-final language such as Japanese as follows.

(9) In a structure [C] [B] [A], where [C] can be structurally attached to [B] or [A]

- a) A prosodic boundary immediately following [C] that is large relative to a boundary following [B] but not following [A] encourages high attachment of [C] to [A].
- b) A boundary immediately following [C] that is small relative to a boundary after [B] but not [A] encourages low attachment to [B].

For boundaries of equal strength, I would expect Clifton et al.'s results to be reproduced in Japanese. That is, for sentences with two boundaries of equal strength, there should be less high attachment than in the case described in (9a) and less low attachment than in the case described in (9b).

Trying to reword this hypothesis for a head final-language highlights an important consideration. As originally formulated (8), the Informative Boundary Hypothesis essentially states that the boundary prior to the relative clause in English varies in its "effectiveness" based on whether a boundary was encountered *previously* in the utterance between the two possible attachment sites. If the hypothesis is reformulated for Japanese as in (9), one would expect the boundary that follows the relative clause to vary in its effectiveness based on whether an additional boundary is encountered *later* in the utterance between the two possible attachment sites.

However, it is also possible that the linear order of boundaries is what is important: perhaps the second boundary in Japanese (located between the two possible attachment sites) will vary in its effectiveness based on whether the listener encountered a boundary prior to it (immediately following the relative clause). Thus, this third condition which tests both prosodic phrase boundaries together is necessary as a comparison condition to compare a head-initial with a head-final language. If the first boundary encountered in the speech stream controls the outcome

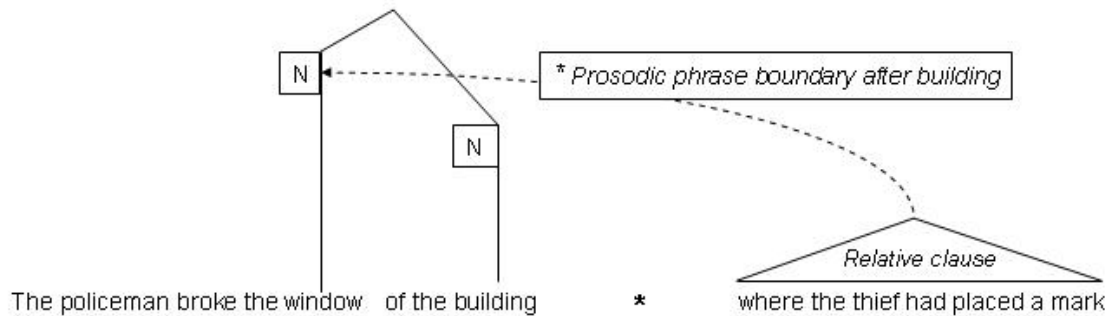
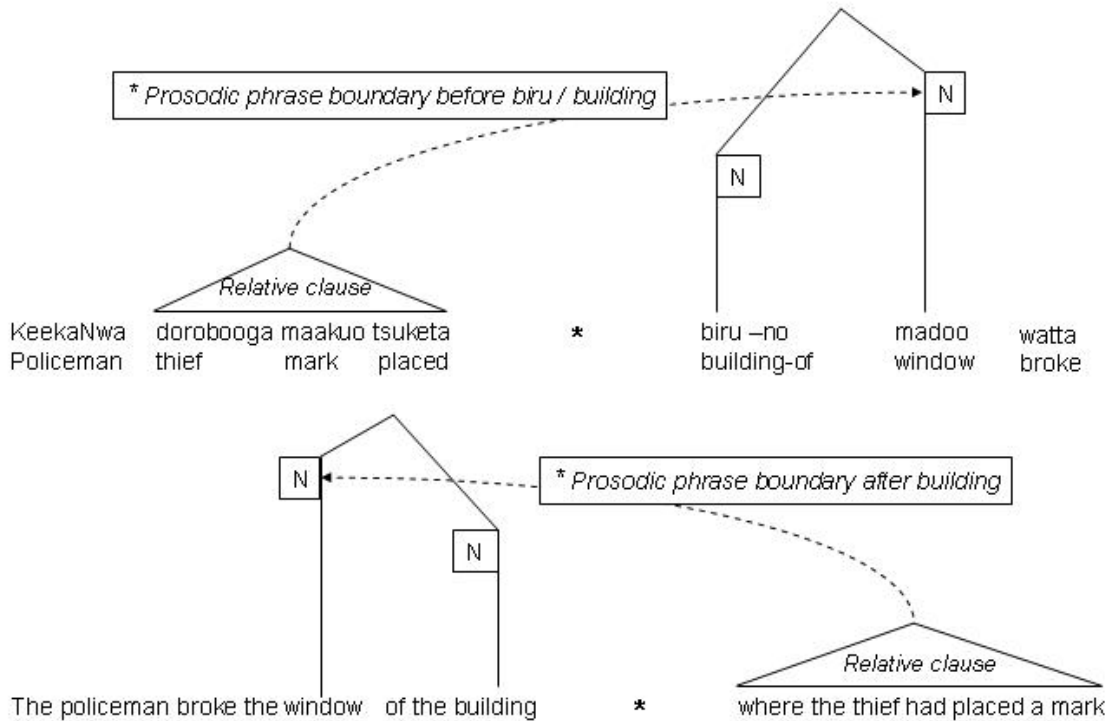
of the parse, then results for English should differ from those for Japanese, since the first boundary encountered will be at a different location in the structure in each language.

Hypotheses

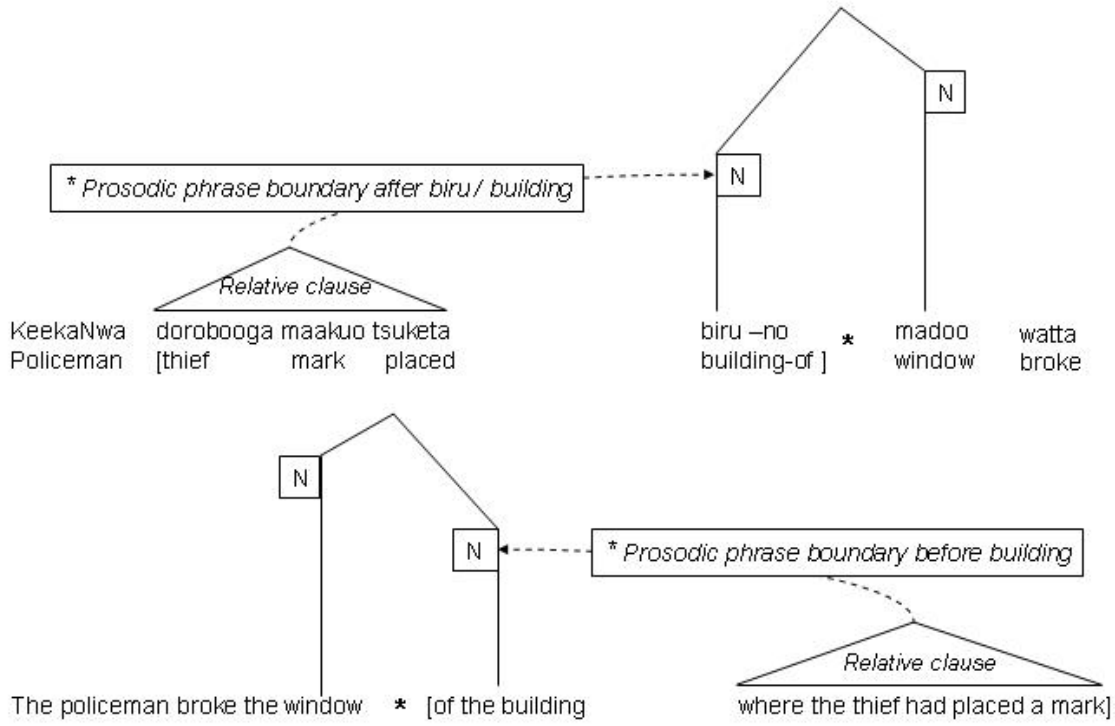
For the condition in which the prosodic phrase boundary separates the relative clause from the adjacent low noun, high attachment of the relative clause (to the non-adjacent, structurally higher noun) is predicted. This is illustrated in Figure 1a. For the condition in which the prosodic phrase boundary groups the low noun with the relative clause, low attachment of the relative clause is predicted. This is illustrated in Figure 1b. These predictions are based on results found in Maynell (2001) and Venditti (1994).

Figure 1. Example English and Japanese Sentences. Partial structures of sentences in Japanese and English, showing ambiguity of relative clause modification and hypotheses for prosodic phrasing.

a. In both languages, a prosodic phrase boundary that separates the relative clause from both potential attachment sites will lead to a high attachment interpretation of the relative clause.



b. In both languages, a prosodic phrase boundary that groups the structurally low noun with the relative clause will lead to low attachment interpretation of the relative clause.



As for the condition with both prosodic phrase boundaries, I expect both boundaries to be important. As demonstrated for English in Clifton et al. (2002), participants are predicted to interpret the relative clause in such sentences as modifying the structurally high noun less often than in the condition with a single prosodic phrase boundary separating the relative clause. At the same time, they are predicted to interpret the relative clause as modifying the structurally low noun less often than the condition in which a single prosodic phrase boundary groups that noun with the relative clause. An analogous result is expected for Japanese: when both prosodic phrase boundaries are present, I predict less high attachment than in the condition in which a single prosodic phrase boundary follows the relative clause, and less low attachment than in the condition in which a single prosodic phrase boundary groups the structurally low noun with the relative clause. In sum, I expect not the linear order of prosodic phrase boundaries to be important, but rather the relationship of those boundaries to the higher level syntactic structure.

Proposed Experiments

Design of Test Items

I have designed and constructed 25 test sentences in English and, with the help of three native speakers of Japanese, their equivalent translations in Japanese. These sentences are listed in Appendix A, and include the example test items in (4) and (5). In keeping with previous written

and spoken language studies in the literature, each sentence contains a relative clause that can modify one of two possible nouns. Thus, each sentence has two possible interpretations.

Because of the constraints involved in working with spoken language, the sentences differ from others in the literature in a couple of ways. First, in the literature on this relative clause attachment ambiguity, the relative clauses are typically formed by moving *wh* from an argument position, as in example (1), and in English they are often headed by "who/which" or "that". In Japanese, there are no overt relative pronouns, so the relative clause *who was on the balcony* in sentence (1) would translate as "balcony-on was." Thus, the relative clause in Japanese would be two words long: the first word would be "balcony" with a post-positional marker, and the second word would be the copula. For this dissertation, locative relatives were originally chosen in order to have relative clauses in Japanese that were three words long, consisting of subject – object – verb (see example (5) and its translation in (4)). The motivation for this was rooted in an idea for a production study, in which participants would be asked to produce these sentences with one or the other intended interpretations. It was necessary to account for the possibility that subjects might not use prosodic phrasing to disambiguate the two meanings of each utterance, and instead use other kinds of pitch movement (for example see Azuma, 1989). In order to see that potential pitch movement clearly, there needed to be enough words over which such putative pitch manipulation could be realized. In the course of designing this dissertation, the focus of which is mainly comprehension, the production study has been omitted (to be included in future research). Locative relatives are expected to behave similarly to the relative clauses tested in the literature: their attachment site, be it the higher or the lower noun, still needs to be resolved by the listener.

Second, in order to be able to show clearly the presence or absence of phrase boundaries (both in the previously planned production study and in the current comprehension study), all words in the critical regions of the utterance in Japanese need to have lexical accent. The words in the critical region are underlined in (10) (which is (5), repeated).

(10) KeekaN-wa [doroboo-ga maaku-o tsuketa] * biru-no * mado-o watta.
Policeman-NOM thief-NOM mark-ACC placed building-of window-ACC broke.

The critical region is the region in which prosodic phrase boundaries will occur (marked by *). Whether or not a word in Japanese has lexical accent contributes to its pitch shape within an intonation phrase. Accentual phrases have at most one lexical accent. Unaccented words are often merged into the same accentual phrase as another accented word.¹ By using only accented words in the critical region, we hoped to avoid this merger in order to produce accentual phrase and intonation phrase boundaries in the appropriate locations very clearly. Also, when there is more than one accentual phrase within the same intonation phrase, they follow predictable down-step patterns (Kubozono, 1993; but see Beckman & Pierrehumbert, 1986), making pitch reset to a higher value at the very beginning of a new intonation phrase easier to see in an F0 contour.

¹ It is also possible to merge accented words into a single accentual phrase given certain focus conditions (see Beckman and Pierrehumbert, 1986). This would have to be considered especially within an experiment that elicits production of such sentences in context. In the current experiments however, test sentences will be recorded in the laboratory by a trained native speaker who is a linguist familiar with Japanese prosody. Thus, such mergers of accented words can be consciously avoided.

A third constraint for Japanese that is not always observed in the literature is that the complex noun phrases need to consist of two inanimate nouns. The post-positional genitive marker *-no* is ambiguous: it can also be a copula. Thus, if one were to translate the phrase "servant of the actress" from (1) into Japanese, it would read "actress-*no* servant" and could potentially mean 'the actress is a servant' (M. Nakayama, personal communication). If both nouns refer to the same entity, there is no longer a disambiguation task to be performed in the parsing of the utterance. Using inanimate noun phrases avoids this confound.

As mentioned above, English sentences are translations of the Japanese sentences, and there are other constraints on the English test items to be observed. Because English has overt relative pronouns, all English relative clauses are introduced with *where* to match the locative relatives in Japanese. For consistency across items, pied piping and stranded prepositions are avoided. Ideally, the two noun phrases that are modifiable by the relative clause should consist of a single head noun each because modification by an adjective can affect the resolution of the relative clause ambiguity (Gilboy et al., 1995). However, in some cases two-word noun phrases are used (for example, *vending machine* (#6), *electric pole* (#7), *gift shop* (#12) and *sushi bar* (#15)). Such phrases are common enough that they may function as single nouns more than as nouns modified by adjectives (or other nouns), but these items will be evenly distributed among lists in case they pattern differently than sentences in which noun phrases contain a single head noun. Article type (*a* vs. *the*) within the two noun phrases has also been shown to affect the ambiguity resolution in such sentences (Gilboy et al., 1995). Here, the definite article will be used throughout, in part to match what had been done to date in the literature and in part because a comparison of article types is beyond the scope of the current study.

Preparation and Analysis of Test Items

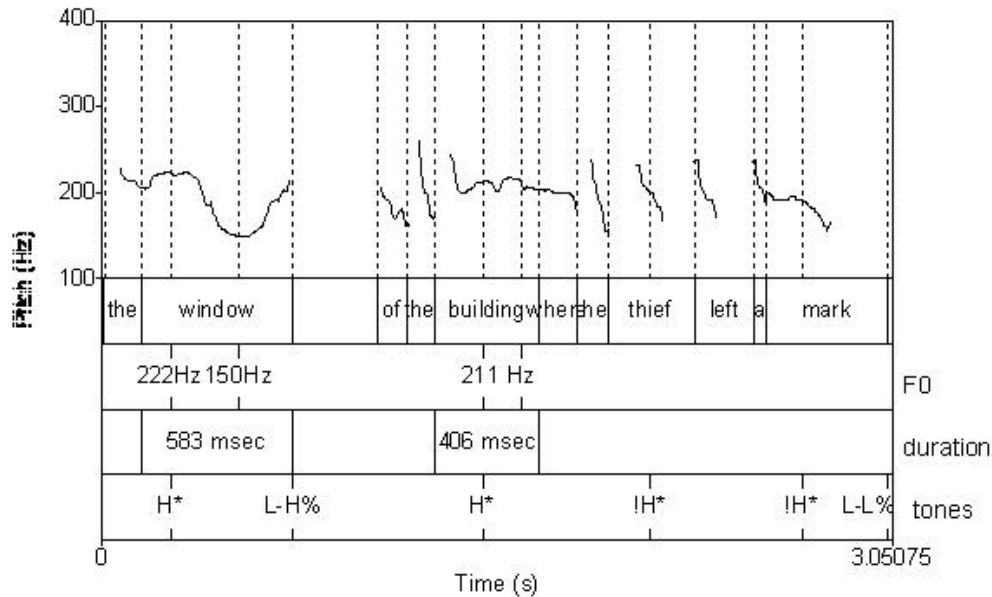
Sentences in each language will be recorded by a native speaker for presentation to participants in a comprehension study (in Experiments 3 and 4, see below). It will thus be necessary to analyze the intonation pattern of all recorded stimuli to ensure consistency in intonation pattern across items. Once sentences are recorded with a prosodic phrase boundary in the designated location(s), it will be necessary to measure fundamental frequency values within the critical regions of each utterance and duration of the words at which the prosodic phrase boundaries occur, in order to show with physical measurements that they were successfully produced. In addition to such measurements, phonological analyses will be conducted using the ToBI system for American English (Silverman, et al., 1992; Beckman & Ayers, 1997) and the J-ToBI system (Venditti, 1997) for Japanese. Both types of analyses are crucial. A phonological analysis using the ToBI systems for describing prosody allows native speaker labelers to indicate their judgments about phonological categories such as pitch accents, phrase accents and boundary tones in English, and accentual phrases and intonation phrase boundaries in Japanese. Measurements of F0 values and word duration are used to support and confirm these native speaker judgments with physical data.

Figure 2 shows example fundamental frequency (F0) contours of the critical region of example (4), "...the window of the building where the thief left a mark." The recording shown in Figure 2a was produced with an intonation phrase boundary after *window*, and the recording in 2b was

produced with an intonation phrase boundary after *building*, prior to the relative clause. A ToBI transcription is shown in bottom tier of each TextGrid. In English, an intonation phrase boundary entails the presence of an intermediate phrase boundary (Beckman, 1996; Beckman & Hirschberg, 1994). The intermediate phrase boundary is often seen as either a fall or rise in the F0 contour². In Figure 2a, the word *window* shows a fall in the fundamental frequency in the second syllable, which in ToBI is labeled as a low intermediate phrase accent L-; the subsequent rise is a high boundary tone, H%.

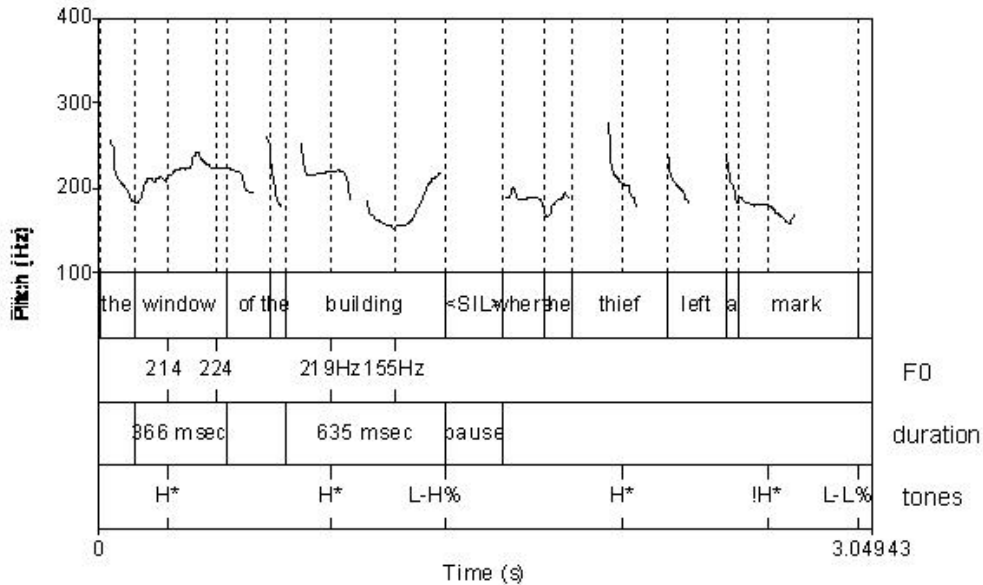
Figure 2. Fundamental frequency contours for two different recordings of the critical region of the utterance "The computer will flash the window of the building where the thief left a mark."

2a. Fundamental frequency contour of the critical portion of the utterance "The computer will flash the window of the building where the thief left a mark." The four-tier TextGrid (made using Praat Sound Analysis Software) shows word boundaries in tier 1, F0 measurements in tier 2, word duration in the third tier, and a ToBI transcription tone labels in tier 4.



² Sometimes the intermediate phrase accent stays level in pitch, but not in these particular examples. When the pitch remains level throughout the intermediate phrase accent (and boundary tone, if present), then the presence of the phrase boundary is determined by phrase-final lengthening.

2b. Fundamental frequency contour of the critical portion of the utterance “The computer will flash the window of the building where the thief left a mark.” The four-tier TextGrid (made using Praat Sound Analysis Software) shows word boundaries in tier 1, F0 measurements in tier 2, word duration in the third tier, and a ToBI transcription tone labels in tier 4.



The F0 minimum in the second syllable of *window* is 150 Hz; the same word in Figure 2b, without the phrase boundary, does not show a similar fall: the F0 minimum of the second syllable is 224 Hz. In addition, the duration of *window* in Figure 2a is 583 milliseconds, whereas the duration of the same word in Figure 2b produced without the phrase boundary is 366 milliseconds. Thus, the presence of a phrase boundary in Figure 2a can be confirmed by measuring F0 values in the relevant region of the utterance, as well as the duration of the final word in the phrase. A similar comparison can be made between the word *building* in Figure 2a, uttered without a phrase boundary, and the same word in Figure 2b, uttered with a phrase boundary. The F0 minimum in the second syllable of *building* in Figure 2a is 211 Hz; compare the F0 minimum of 155 Hz in Figure 2b in the second syllable of *building* produced with a phrase boundary. The duration of *building* in Figure 2a is 406 milliseconds. When produced with an intonation phrase boundary as shown in Figure 2b, the duration is 635 milliseconds.

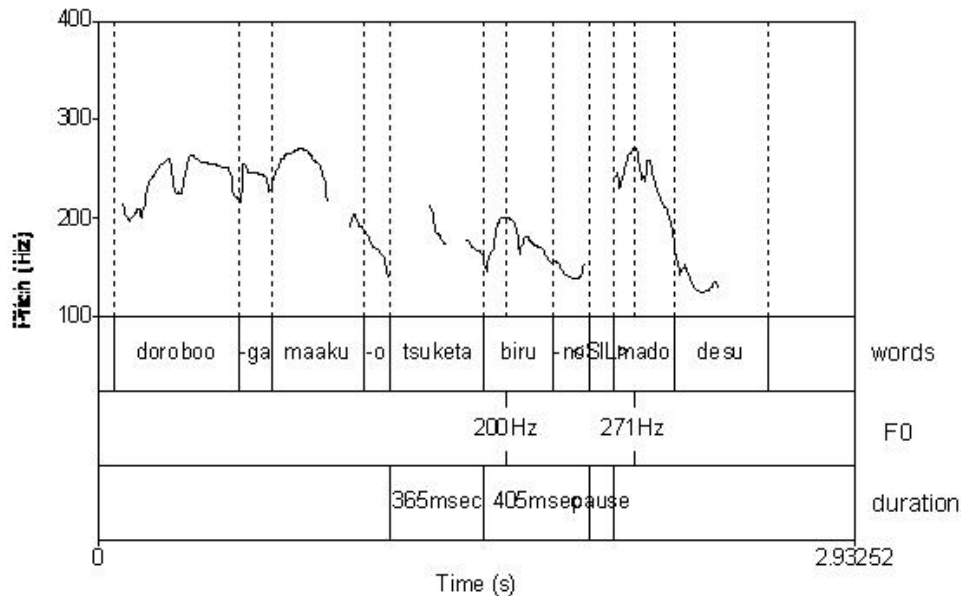
A representation that shows both intonation phrase boundaries in each language is not included here, because having two boundaries will not influence the way each of the boundaries is realized.

In Japanese, the end of an intonation phrase is also often accompanied by a very short duration of silence, and word final lengthening (Beckman & Pierrehumbert, 1986). A reliable way to find the beginning of a new intonation phrase is to look for a resetting of the pitch range to a higher

value than in the previous prosodic phrase. As an example, the critical region of sentence (5), "... doroboo-ga maaku-o tsuketa biru-no mado desu" is shown in Figure 3.³ The maximum F0 value of the word *mado* in Figure 3a is much higher (271 Hz) than the preceding *biru* (200 Hz), which is evidence that a new intonation phrase begins at *mado*. Contrast Figure 3b, in which the F0 maximum for *biru* is 312 Hz, much greater than its F0 in Figure 3a. It is also greater than the F0 maximum in the preceding word *tsuketa* (148 Hz) in Figure 3b, indicating that *biru* begins a new intonation phrase. There can also be a short region of silence (pause) prior to the beginning of a new intonation phrase, as seen in both 3a and 3b. Note also the difference in word length for *tsuketa* produced in the middle of an intonation phrase (365 milliseconds, Figure 3a) and at the end of an intonation phrase (443 milliseconds, Figure 3b). Similarly, *biru-no* has a longer duration when it ends an intonation phrase (405 milliseconds, Figure 3a) than when it does not (313 milliseconds, Figure 3b).

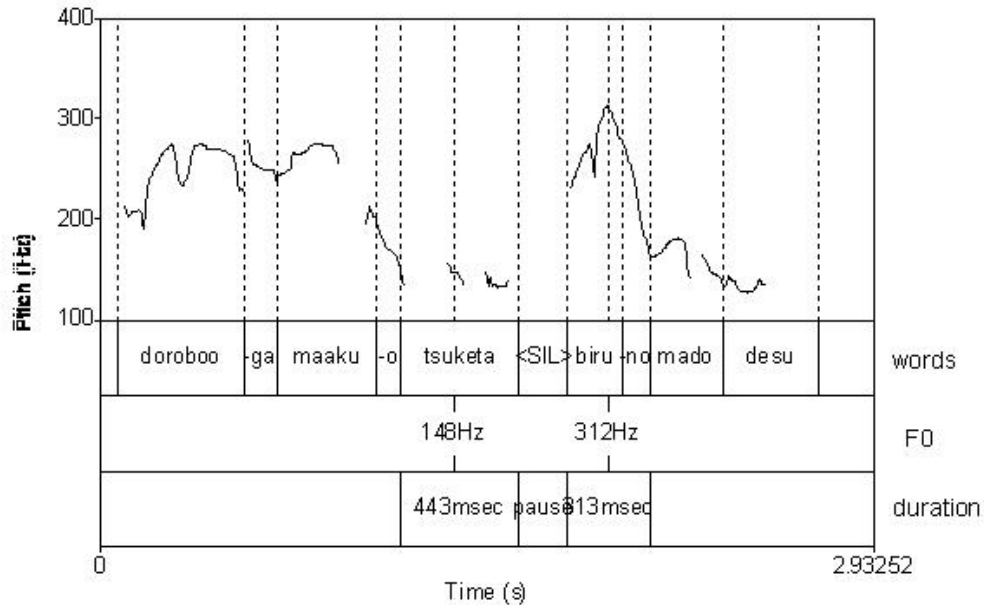
Figure 3. Fundamental frequency contours for two different recordings of the critical region of the utterance "The computer will flash the window of the building where the thief left a mark." in Japanese.

3a. Fundamental frequency contours of the critical portion of the utterance "Tenmetsu-suru no wa doroboo-ga maaku-o tsuketa biru-no mado desu. The three tier TextGrid (made using Praat Sound Analysis Software) shows word boundaries in tier 1, F0 measurements in tier 2 and word duration in the third tier



³ This version of the utterance differs slightly from the example shown in (5). This is because stimuli have to be modified slightly for two of the four experiments proposed. This is explained in subsequent sections.

3b. Fundamental frequency contours of the critical portion of the utterance "Tenmetsu-suru no wa doroboo-ga maaku-o tsuketa biru-no mado desu. The three tier TextGrid (made using Praat Sound Analysis Software) shows word boundaries in tier 1, F0 measurements in tier 2 and word duration in the third tier



A representation that shows both boundaries in the same sentence is not included here, because there is little difference in how each boundary would be realized. The F0 maximum for *biru* would be greater than for the preceding word *tsuketa*, and *mado* would not be down-stepped with respect to *biru* as it is in Figure 3b. The F0 maximum for *mado* in the two-boundary condition would be at least as great if not greater than the F0 of *biru*. A ToBI transcription will be performed by a native Japanese speaker trained in the J-ToBI system for transcribing Japanese prosody (Venditti, 1997), though such a transcription is not shown in Figure 3.

Experiments 1 and 2

Experiment 1, in English, and Experiment 2, in Japanese, will specify the extent of ambiguity in the materials. Preliminary studies completed in September, 2000, established the effectiveness of the experimental method and the appropriateness of the sentence structures in Japanese. Native speakers (30-40 per language) will judge the plausibility of each interpretation of text versions of the ambiguous sentences. Each sentence will be presented with two possible interpretations in written form in counter-balanced lists, distributed among 50 control sentences. Participants will use a five-point scale to indicate their preferred interpretation of the utterance. Using another numerical scale, they will then rate the plausibility of each interpretation. Results will be evaluated statistically by analysis of variance using participants and items as random variables. Results will also be used to calibrate the ambiguity of individual items.

Though the sentences are syntactically ambiguous, it is possible that some will show a preferred interpretation in the results from the first two experiments. However, these sentences will still be used in Experiments 3 and 4. My assumption is that even if one interpretation of a sentence is preferred over another in these off-line reading studies, one of the intonation patterns produced in Experiments 3 and 4 will facilitate the recovery of the less preferred interpretation (Kjelgaard & Speer, 1999).

Experiments 3 and 4

Experiments 3 (in English) and 4 (in Japanese) are listening experiments, using eye movement as a measure of comprehension. Their purpose is to evaluate the influence of intonation on the recovery of the intended meaning of ambiguous test items. Participants (50-60 per language) will wear a head-mounted eye tracking device (ASL Model 501 with eye-head integration) which will monitor their eye movements as they listen to sentences. Eye movements provide a fine-grained measure of comprehension that is closely time-locked to the language input (cf. Chambers et al., 2004; Dahan & Tanenhaus, 2004).

For the eye tracking task, the stimuli will be modified to suit the task but the attachment ambiguity will remain intact. Test items (4) and (5) are reprinted below as (11) and (12).

(11) The policeman broke the window of the building where the thief had placed a mark.

(12) keikaN-wa doroboo-ga ma'aku-o tsuke'ta bi'ru-no ma'do-o watta.

Policeman-NOM thief-NOM mark-ACC placed building-of window-ACC broke.

The corresponding modified stimuli for the eye-tracking study are shown in (13) and (14).

(13) The computer will flash the window of the building where the thief had placed a mark.

(14) Tenmetsu-suru no wa doroboo-ga ma'aku-o tsuke'ta bi'ru-no ma'do desu.

What will flash thief mark placed building-GEN window is

These modified stimuli will be recorded and analyzed as described above in "Preparation and Analysis of Test Items." The boundary conditions are shown for examples (11) and (12) in examples (15-20).

(15) The computer will flash the window **|A|** of the building where the thief had placed a mark.

(16) Tenmetsu-suru no wa doroboo-ga ma'aku-o tsuke'ta bi'ru-no **|A|** ma'do desu.

What will flash thief mark placed building-GEN window is

(17) The computer will flash the window of the building **|B|** where the thief had placed a mark.

(18) Tenmetsu-suru no wa doroboo-ga ma'aku-o tsuke'ta **|B|** bi'ru-no ma'do desu.

What will flash thief mark placed building-GEN window is

(19) The computer will flash the window **|A|** of the building **|B|** where the thief had placed a mark.

(20) Tenmetsu-suru no wa doroboo-ga ma'aku-o tsuke'ta **|B|** bi'ru-no **|A|** ma'do desu.
What will flash thief mark placed building-GEN window is

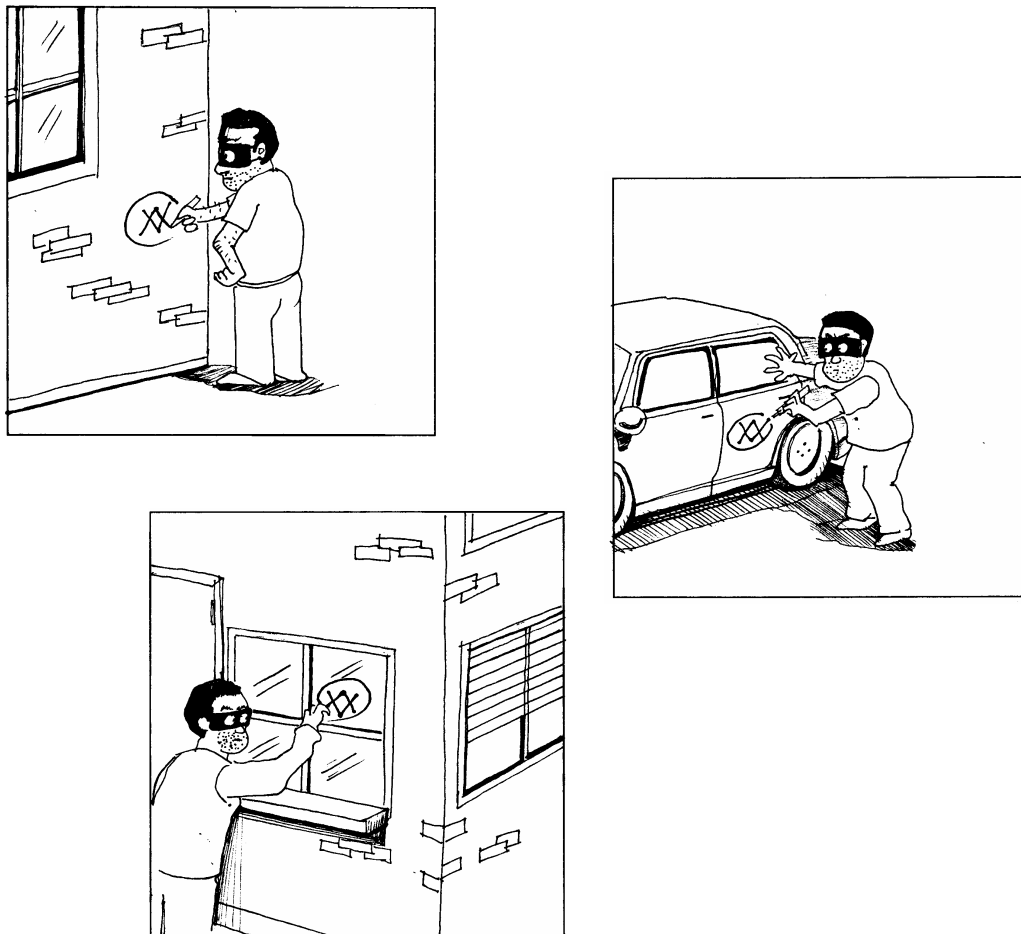
Boundary condition A, shown in (15) and (16), separates both possible attachment sites from each other, prosodically grouping the relative clause with the structurally lower noun. Boundary condition B, shown in (17) and (18), separates the relative clause from both potential attachment sites, grouping the relative clause by itself in its own intonation phrase. Boundary condition C, shown in (19) and (20), combines both conditions A and B, placing the structurally lower noun within its own intonation phase.

Method

On each trial, participants will hear a single sentence played over headphones, and E-Prime software (Psychology Software Tools, Inc., www.pstnet.com) will be used to present both the auditory stimulus and simultaneously three picture choices on a computer screen. Listeners will be told that something on the computer screen will flash, but that on some trials, no flashes will occur. On these trials, listeners will be instructed to click on the picture that *should have* flashed. Although listeners will receive no explicit instruction to look at the pictures, they will move their eyes to the picture described in the sentence in order to monitor for the flash and if necessary, to click on the picture. Participants' eye movements are expected to reflect how they are using intonation patterns to understand ambiguous test sentences. In the interpretation of the test sentences, different intonation patterns should lead to different interpretations, realized as participants look to different pictures.

For experimental trials, two pictures will correspond to possible interpretations of the relative clause, and one distracter picture will contain a subset of elements of the test sentences. Because the same set of pictures will be used for both languages, it will be necessary for all three pictures to be candidates for the flash early in each sentence, so that no one picture is eliminated right away when the beginning of the critical region of each utterance is heard. Specifically, each picture must show the first noun (e.g. *window*, as in (13)), which occurs first in the region of interest in the English sentence, and each picture must show the relative clause's property, which occurs first in the region of interest in the Japanese sentence (*the thief left a mark*, as in (14)). As an example, the three-picture set for sentences (13) and (14) is shown in Figure 4. First, note that all pictures have windows, and all pictures have thieves leaving marks on something. One picture shows a building on which a thief leaves a mark on a window (high attachment of the relative clause), and one picture shows a building with a window, but the thief is shown leaving a mark on the building - not on the window (low attachment of the relative clause). The distracter picture for this trial shows a car on which the thief leaves a mark somewhere on the car (but not on one of its windows). Thus, the distracter picture includes both a window and a thief leaving a mark, as do the other two pictures, but it differs from the other pictures in that it shows no building.

Figure 4. Example display of three pictures corresponding to sentences (13) and (14). The picture in the upper left corner shows the low attachment interpretation of the relative clause, in which the thief is leaving a mark on the building, but not the window. The bottom picture shows the high attachment interpretation, in which the thief leaves the mark on the window itself. The right hand picture is a distracter picture, which includes the structurally high noun, *window*, as well as the action described in the relative clause, *the thief left a mark*. Note that the distracter picture is analogous to a low attachment interpretation of the relative clause, because the thief is not leaving the mark on the window of the car.



Once participants hear the structurally lower noun (*building/biru*, see (13-14)), they should eliminate the distracter picture with the car and have their choice narrowed to the two remaining pictures. Which picture they ultimately click on will depend on the intonation pattern used in the production of the test sentence, as described previously. Distracter pictures for all trials will be balanced between a high attachment and a low attachment interpretation, so for experimental trials, 12 distracter pictures will show an analogous high attachment interpretation and 12 will show an analogous low attachment interpretation. The pictures will be arranged in a quasi-

circular fashion as shown in Figure 4. The target picture will rotate from one square to another in a fixed random order with each new trial.

Table 1 shows the distribution of stimuli across three lists. Each participant will hear each stimulus only once and only in one boundary condition. Each participant hears all boundary conditions by hearing 8 items per boundary condition.

Table 1. Distribution of Stimuli Across Three Lists

Item #	List 1	List 2	List 3
1	Bdry condition A	B	C
2	Bdry condition B	C	A
3	Bdry condition C	A	B
4	A	B	C
5	B	C	A
6	C	A	B
7	A	B	C
8	B	C	A
9	C	A	B
10	A	B	C
11	B	C	A
12	C	A	B
13	A	B	C
14	B	C	A
15	C	A	B
16	A	B	C
17	B	C	A
18	C	A	B
19	A	B	C
20	B	C	A
21	C	A	B
22	A	B	C
23	B	C	A
24	C	A	B

The stimuli will be distributed among control sentences. Table 2 shows the overall experimental design, including control sentences.

Table 2. Design of Comprehension Experiment with Eye-Tracking as Dependent Measure

Flash	Items (Syntax)	Intonation	Stimuli vs. Controls
36 items no flash	24 experimental sentences, syntactically ambiguous	8 sentences per boundary condition x 3 boundary conditions	24 experimental sentences
	12 control sentences, syntactically unambiguous	12 mixed intonation patterns	50 control sentences
38 items with flash	38 control sentences (12 w/critical intonation: syntactically unambiguous; 26 w/mixed intonation: 24 syntactically ambiguous and 2 unambiguous)	12 critical intonation patterns (3 boundary conditions x 4 controls each) 26 mixed intonation patterns	

Control sentences will be similar in structure to test items. Specifically, they will all contain a relative clause ambiguity just as the test items do, but some will be structurally ambiguous and some will not be ambiguous. To counterbalance the trials both auditorily and visually, the following criteria will be met (please refer to Table 2):

Auditory Counterbalancing:

- To prevent test items from standing out, 12 unambiguous control sentences will be presented with pictures sets that do not flash, resulting in 36 total non-flashing trials. These 12 sentences will be unambiguous so that the task of clicking on the appropriate picture can be easily completed by the participants. The intonation patterns of these sentences will be mixed.
- To prevent critical intonation patterns (boundary conditions A, B, or C) from predicting that the corresponding trial will not have a flash, it will be necessary to use critical intonation patterns for 12 of the control sentences that contain a flash in their picture sets. (Combined with 24 experimental items, this results in a total of 36 items, slightly less than half of the total number of items, that use the critical intonation patterns). To distribute the three boundary conditions among 12 control sentences evenly, 4 will have boundary condition A, 4 will have boundary condition B and 4 will have boundary condition C. It is crucial that the flashing picture in these controls does not correspond to the predicted result all of the time, so it will be necessary to make sure that these control sentences are semantically and plausibly unambiguous. Specifically, for boundary condition A, I predict a low attachment interpretation. Thus, for the 4 control trials using boundary condition A, at least 2 of the pictures sets will have a flash in the picture that shows a high attachment interpretation. For boundary condition B, I predict a high attachment

interpretation, so at least 2 of those 4 control trials using boundary condition B will have a flash in the picture that shows a low attachment interpretation.

- Critical intonation patterns will not predict syntactic ambiguity. Therefore, of the 26 control sentences with mixed intonation and flashing picture sets, 24 of the sentences will be ambiguous.

Visual Counterbalancing:

- Slightly less than half of all trials (36/74) will not include a flashing picture. So, any picture set has a (slightly more than) 50% percent chance of having a picture that includes a flash. This means that any picture set has a (slightly less than) 50% chance of requiring a mouse click to be made by the participant.
- The experimental trials are such that I expect 8 low attachment interpretations (boundary condition A) and 8 high attachment interpretations (boundary condition B) to be selected by participants. The other 8 interpretations are unknown. Of the 50 remaining control sentences, 12 sentences will not have a flashing picture set, 38 will.
- Of the 12 unambiguous control sentences that do not have a flashing picture set (all with mixed, non-critical intonation patterns), 6 will lead the participant to click on the picture that corresponds to a high attachment interpretation and 6 will lead the participant to click on the picture that corresponds to the low attachment interpretation.
- Of the 38 control sentence that have flashing picture sets, 12 sentences will have critical intonation pattern; 6 answers will correspond to the picture showing a low attachment interpretation and 6 to the picture showing high attachment interpretation.
- The remaining 26 control sentences will have mixed (and non-critical) intonation patterns. 13 of these will have picture sets in which the picture corresponding to the low attachment interpretation flashes, and the other 13 will have picture sets in which the picture corresponding to the high attachment interpretation (recall that 24 of these 26 sentences will be ambiguous).

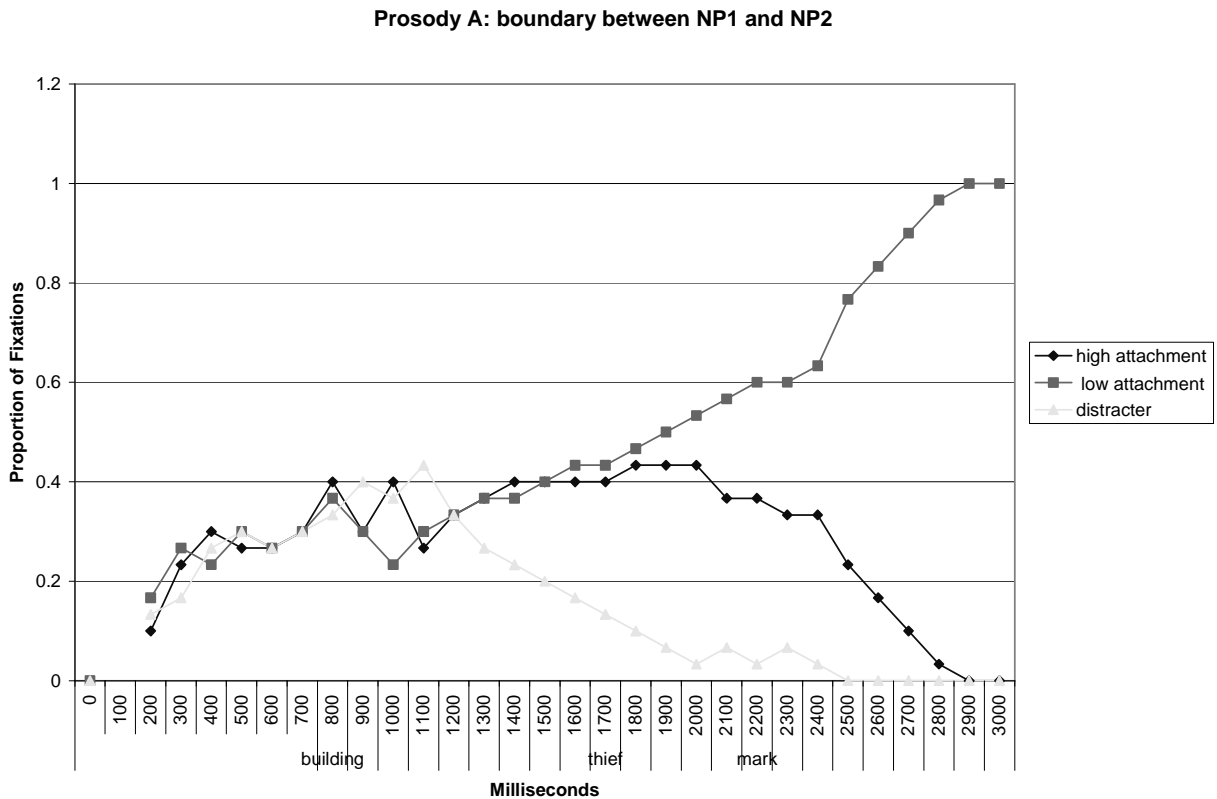
Example Eye Movement Data

Eye movements are expected to reflect a listener's interpretation of the sentence that they are hearing. Figure 5 shows one example graph illustrating the predicted proportion of eye fixations in response to sentence (13) in one prosody condition collapsed over all putative participants. It shows what participants are predicted to do when the intonation phrase boundary separates the two possible attachment sites for the relative clause from each other, thus grouping the structurally low noun *building* with the relative clause.

The X-axis begins at the offset of the word *window* (see Figure 2a for approximate duration in milliseconds of the subsequent regions of the utterance). It takes approximately 200 msec for a listener to launch a saccade upon hearing a word (Dahan & Tanenhaus, 2004, and references cited therein, p. 500), so the sample graph shows that measurements are taken beginning at 200 milliseconds following the offset of *window*. Initially, all windows in all three pictures are potential candidates for the possible flash, so listeners should show an equal likelihood of looking at any of the three pictures shown in Figure 4. Not until participants hear the word *building* between 600-1000 milliseconds later does the distracter picture, in which the window of a car is depicted, become eliminated as a candidate. Because of the boundary following *window*, I predict that participants will look to the low attachment picture before they hear the final word

mark. The prosodic grouping should be a signal to listeners that whatever modification is unfolding in the relative clause is describing the structurally lower noun. Ultimately, participants should click on the picture depicting low attachment of the relative clause, in which the thief is leaving the mark on the building, and not on the window.

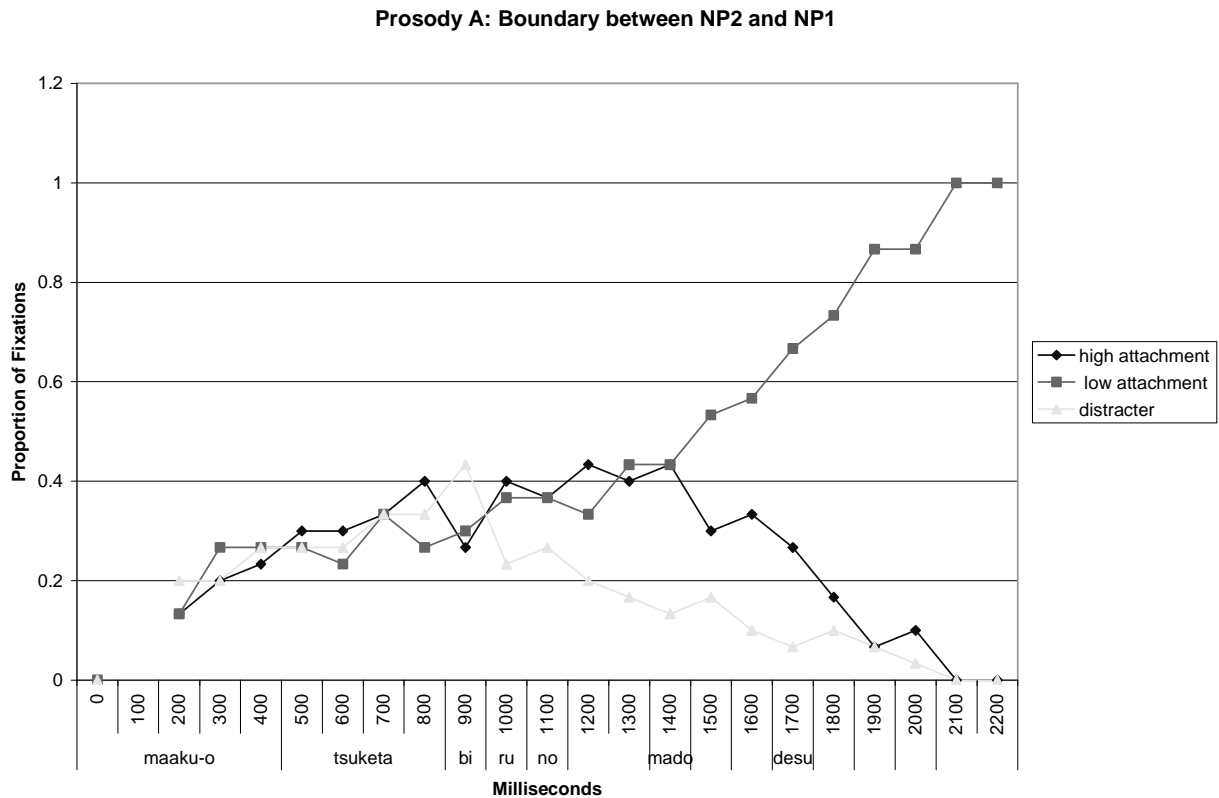
Figure 5. Example graph showing predicted proportions of fixations with respect to time for the English sentence (13) under prosody condition A, in which a low attachment interpretation of the relative clause is predicted.



Another example graph in Figure 6 illustrates the predicted proportions of eye fixations in response to the Japanese sentence (14) in the analogous prosody condition. It shows what participants are predicted to do when the intonation phrase boundary separates the two possible attachment sites for the relative clause from each other in Japanese, grouping the structurally low noun *biru* (*building*) with the relative clause. The graph begins at the offset of the word *doroboo-ga* (*thief*), or at the onset of the word *maaku-o* (*mark*) (see Figure 3a for approximate duration in milliseconds for the subsequent regions of the utterance). Measurements are taken beginning at 200 milliseconds after the onset of the word *maaku-o* (*mark*). Because these words are the first two words of the relative clause, and the relative clause's action is illustrated in all three pictures (see Figure 4), all three pictures are potential candidates for the possible flash, and participants are equally likely to look at any of the three. Upon hearing the word *biru-no* (*building*), there should be fewer looks to the distracter picture as participants eliminate it as a choice. Finally, upon hearing the word *mado* (*window*) at the beginning of a new intonation

phrase, participants should eliminate the picture in which the thief leaves a mark on the window, because the prosodic boundary separates it from the previous candidate for modification by the relative clause.

Figure 6. Example graph showing predicted proportions of fixations with respect to time for the Japanese sentence (14) under prosody condition A, in which a low attachment interpretation of the relative clause is predicted.



Implications

The proposed experiments compare directly how speakers of two languages employ intonation to recover sentence meaning. An important feature of this research is the direct comparison of processing in two languages, one head-initial, one head-final, using an on-line measure. Changes in eye movement patterns are predicted to reflect listeners' interpretation of an utterance resulting from the presence of prosodic boundaries at specific locations. The eye-tracking methodology presents an advantage over other off-line tasks commonly used to evaluate the interpretation of relative clause attachment ambiguities in that eye-movements are closely time-locked to language input during comprehension (Dahan & Tanenhaus, 2004), and it should be possible to watch parsing unfold over time during the visual search.

The research described here is expected to show that maintaining a correspondence between prosodic structure and syntactic structure can account for the integration of meaning in a mental representation of an utterance regardless of word order. Such a mechanism is potentially a cross-linguistically applicable language comprehension strategy that humans share. Indeed, such

structuring of sound as it is perceived and applied to the analysis of a concurrent visual world may be a more general cognitive ability applicable to all types of sound, including language.

There are two general goals behind this research. The first is to add to the knowledge about prosody and its role in spoken language comprehension. The second goal is to begin to generate a single model of spoken language processing by first finding support for the hypothesis stated previously in two structurally different languages. To date, there is no complete model of spoken language comprehension that includes intonation.

By contrast, there are models for comprehension of written language. Models that capitalize on verb sub-categorization information and thematic roles in head-initial languages in which the verb is encountered early in a sentence are difficult to apply to Japanese and other head-final languages. Alternate parsing models for head-final languages have been proposed (for a review, see Butler, 1994). These models are useful in that they focus on the incrementality of processing and the information that becomes available as each word is heard or read in an utterance (for example see Kamide & Mitchell, 1999). For head-final languages such as Japanese, this information is contained in case marking of arguments and in sub-categorization properties of the verb, though unlike in head-initial languages, verb information is theoretically not usable in head-final languages until quite late in processing.

However, it is not just linear word order that is important, but also higher level structure; native speakers of any language have knowledge of both. Prosody shapes the higher level structure. In the languages studied here, prosodic phrases can package single syntactic constituents in a similar way: for example, in both languages, if low attachment of the relative clause is intended, then the structurally low noun can be grouped with the relative clause as a single syntactic unit and a single prosodic unit.

A single model that covers more than one class of languages (head-initial vs. head-final) is desirable because it is simpler. While my assumption is that prosodic structure guides the syntactic parse in spoken language, it is possible that prosodic phrasing will function differently in both languages. Such a result would leave us where we are now, proposing different processing models for different language types. On the other hand, even if prosodic phrasing functions as predicted, there is still room for variability between English and Japanese, and other principles to account for this variability, such as the Tuning Hypothesis (Cuetos et al., 1996) or a constraint ranking hypothesis (Gibson et al., 1996). A single universal model could still exist with such hypotheses or principles to account for some variability between languages.

References

- Azuma, J. (1989). Acoustic factors determining the comprehension of a Japanese ambiguous sentence: In the case of the Himeji dialect. *Studies in Phonetics and Speech Communication* **3**. Kinki Society for Phonetics.
- Beckman, M. E. (1996). The parsing of prosody. *Language and Cognitive Processes*, **11**, 17-67.
- Beckman, M. E. & Ayers, G. (1997). Guidelines for ToBI labeling. Manuscript, Columbus, OH: The Ohio State University.
- Beckman, M. E. & Hirschberg, J. (1994). The ToBI annotation conventions. Manuscript, The Ohio State University, Columbus, OH.
- Beckman, M. E. & Pierrehumbert, J. (1986). Intonational structure in Japanese and English. *Phonology Yearbook*, **3**, 255-309.
- Brysbaert, M. & Mitchell, D. C. (1996). Modifier attachment in sentence parsing: Evidence from Dutch. *Quarterly Journal of Experimental Psychology*, **49A**, 664-695.
- Butler, H. Y. (1994). Processing Japanese and Korean. Ph.D. dissertation, the Ohio State University, Columbus, OH.
- Carreiras, M. & Clifton, C. (1993). Relative clause interpretation preferences in Spanish and English. *Language and Speech*, **36**, 353-372.
- Chambers, C. G., Tanenhaus, M. K. & Magnuson, J. S. (2004). Actions and affordances in syntactic ambiguity resolution. *Journal of Experimental Psychology: Learning, Memory, and Cognition* **30**(3), 687-696.
- Clifton, C., Carlson, K. & Frazier, L. (2002). Informative prosodic boundaries. *Language and Speech* **45**: 87-114.
- Cuetos, F. & Mitchell, D. (1988). Cross-linguistic differences in parsing: restrictions on the issue of the Late Closure Strategy in Spanish. *Cognition*, **30**, 73-105.
- Cuetos, F., Mitchell, D. C. & Corley, M. M. B. (1996). Parsing in different languages. In M. Carreiras, J. Garcia-Albea & N. Sebastian-Galles (Eds.), *Language Processing in Spanish* (pp. 145-187). Hillsdale, NJ: Erlbaum.
- Dahan, D. & Tanenhaus, M. K. (2004). Continuous mapping from sound to meaning in spoken-language comprehension: Immediate effects of verb-based thematic constraints. *Journal of Experimental Psychology: Learning, Memory, and Cognition* **30**(2), 498-513.
- Ehrlich, K., Fernandez, E., Fodor, J. D., Stenshoel, E., & Vinereanu, M. (1999). Low attachment of relative clauses: new data from Swedish, Norwegian and Romanian. Poster

presented at the 12th Annual CUNY Conference on Human Sentence Processing, March 18-20, 1999, New York, NY.

Fernandez, E. (2000). Bilingual sentence processing: Relative clause attachment in English and Spanish. Ph.D. Dissertation, City University of New York. New York, NY.

Fernandez, E. & Bradley, D. (1999). Length effects in the attachments of relative clauses in English. Poster presented at the 12th Annual CUNY Conference on Human Sentence Processing, March 18-20, 1999, New York, NY.

Fodor, J. D. (2002a). Psycholinguistics cannot escape prosody. In *Proceedings of the Speech Prosody 2002 Conference*, Aix-en-Provence, (pp. 83-88).

Fodor, J. D. (2002b). Prosodic disambiguation in silent reading. In M. Hirotani (Ed.), *Proceedings of the Northeast Linguistic Society (NELS)*, **32**. Amherst, MA: Graduate Linguistic Student Association.

Frazier, L. (1978) *On comprehending sentences: Syntactic parsing strategies*. Unpublished doctoral dissertation, University of Connecticut.

Frazier, L. & Clifton, C., Jr. (1996). *Construal*. Boston: MIT Press.

Gibson, E., Pearlmutter, N., Canseco-Gonzalez, E., & Hickok, G. (1996). Recency preference in the human sentence processing mechanism. *Cognition*, **59**, 23-59.

Gilboy, E., Sopena, J-M., Clifton Jr., C. & Frazier, L. (1995). Argument structure and association preferences in Spanish and English complex NPs. *Cognition*, **54**, 131-167.

Hirose, Y. (2003). Recycling prosodic boundaries. *Journal of Psycholinguistic Research* **32**: 167-195.

Jun, S. (2003). Prosodic phrasing and attachment preferences. *Journal of Psycholinguistic Research* **32**: 219-249.

Kamide, Y. & Mitchell, D. C. (1999). Incremental pre-head attachment in Japanese parsing. *Language and Cognitive Processes*, **14**, 631-662.

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

Kubozono, H. (1993). The organization of Japanese prosody. Tokyo: Kurosio Publishers.

Lovrić, N. (2003). *Implicit prosody in silent reading: Relative clause attachment in Croatian*. Doctoral dissertation, CUNY Graduate Center.

- Lovrić, N., Bradley, D., & Fodor, J. D. (2001, September). Overt prosody and ambiguity resolution in silent reading. *Poster presented at the 7th Annual Conference on Architectures and Mechanisms in Language Processing (AMLaP)*, Saarbrücken, Germany.
- Lovrić, N., Bradley, D., & Fodor, J. D. (2000). RC attachment in Croatian with and without preposition. *Poster presented at the 6th Annual Conference on Architectures and Mechanisms in Language Processing (AmLAP)*, Leiden.
- Maynell, L. A. (2001) Prosodic effects on relative clause attachment in English. Unpublished Manuscript, the Ohio State University, Columbus, OH.
- Mitchell, D. C. & Brysbaert, M. (1998). Challenges to recent theories of crosslinguistic variation in parsing: Evidence from Dutch. *Syntax and Semantics* **31**, *Sentence Processing: A Crosslinguistic Perspective*, pp. 313-335.
- Miyamoto, E. T., Gibson, E., Pearlmutter, N. J., Aikawa, T., & Miyagawa, S. (1999). A U-shaped relative clause attachment preference in Japanese. *Language and Cognitive Processes*, **14**, 663-686.
- Price, P. J., Ostendorf, M., Shattuck-Hufnagel, S. & Fong, C. (1991). The use of prosody in syntactic disambiguation. *Journal of the Acoustic Society of America*, **90**, 2956-2970.
- Schafer, A. J., Carter, J., Clifton Jr., C. & Frazier, L. (1996). Focus in relative clause construal. *Language and Cognitive Processes*, **11**, 135-163.
- Silverman, K., Beckman, M. E., Pitrelli, J., Ostendorf, M., Wightman, C., Price, P., Pierrehumbert, J., & Hirschberg, J. (1992). ToBI: A standard for labeling English prosody. *Proceedings of the 1992 International Conference on Spoken Language Processing*, **2**, 867-870.
- Speer, S. R., Crowder, R. G., & Thomas, L. M. (1993). Prosodic structure and sentence recognition. *Journal of Memory and Language* **32**, 336-358.
- Venditti, J. J. (1994). The influence of syntax on prosodic structure in Japanese. In J. J. Venditti (Ed.), *Papers from the Linguistics Laboratory, Ohio State Working Papers in Linguistics*, **44**, 191-223.
- Venditti, J. J. (1997). Japanese ToBI labeling guidelines. In K. Ainsworth-Darnell & M. D'Imperio (Eds.), *Papers from the Linguistics Laboratory, Ohio State University Working Papers in Linguistics*, **50**, 127-162.

Appendix A: Test Sentences in Japanese and English

The romanized Japanese version of each sentence is given, followed by a literal translation and the final English version to be used in the experiments. Literal translations sometimes had to be minimally modified to meet constraints for English stimuli described in “Design of Test Items.”

1. yoomu'iN-wa keikaN-ga inu'-o oimawa'shita eiga'kaN-no ge'NkaN-o ha'ita.

Literal: The janitor swept the entrance of the movie theater where the officer chased after the dog.

Use: The janitor swept the entrance of the theater where the officer chased after the dog.

2. kodomo'tachi-wa one'esaN-ga eho'N-o yo'Nda hi'roba-no be'Nchi-o obo'eteita.

Literal: The children remembered the bench of the playground where the {young woman/big sister} read a picture book.

Use: The children remembered the bench of the playground where the lady read a picture book.

3. yakuniN-wa oNna'-ga hana'-o i'keta ho'teru-no ro'bii-o kenashita.

Literal: The government official spoke badly of/put down/found fault with/verbally trashed the lobby of the hotel where the woman arranged flowers.

Use: The government official disparaged the lobby of the hotel where the woman arranged flowers.

4. koojo'ochoo-wa ha'NniN-ga giNkoogo'otoo-o takura'Nda so'oko-no kaNrini'Nshitsu-o kaisooshita.

Literal: The factory boss renovated the manager's office of the warehouse where the criminal planned the bank robbery.

Use: The factory boss renovated the office of the warehouse where the criminal planned the bank robbery.

5. shiikuga'kari-wa shi'choo-ga ri'boN-o tsu'keta ko'ara-no atama'-o na'deta.

Literal: The zookeeper pet the head of the koala bear onto which the mayor put the ribbon.

Use: The zookeeper pet the head of the koala where the mayor tied the ribbon.

6. kakari'iN-wa sho'ojo-ga ko'ora-o katta ki'osku-no jidoohaNba'iki-o hojuushita.

Literal: The "person-in-charge" stocked the vending machine of the kiosk where the girl bought a coke.

Use: The serviceman stocked the vending machine of the kiosk where the girl bought a coke.

7. keNsa'tsukaN-wa hahaoya-ga oka'shi-o kakushita ki'cchiN-no shokki'dana-o shira'beta.

Literal: The prosecutor inspected the dish cabinet of the kitchen where the mother hid sweets.

Use: The prosecutor inspected the cabinet of the kitchen where the mother hid sweets.

8. shiha'iniN-wa ko'kku-ga niku'-o kusara'seta choori'shitu-no reizo'oko-o shi'meta.

Literal: The owner closed the refrigerator of the professional kitchen where the cook let the meat spoil.

Use: The owner closed the refrigerator of the kitchen where the cook let the meat go bad.

9. tsuukooniN-wa otoko'-ga biiru'biN-o nagetsu'keta denshinba'shira-no po'sutaa-o miageta.

Literal: The pedestrian looked at the poster on the electric pole where the man threw the beer bottle.

Use: The pedestrian examined the poster of the electric pole where the man threw the beer bottle.

10. ootoo'-wa ka'shu-ga sa'iN-o ka'ita shiidi'i-no ke'esu-o yogoshita.

Literal: The younger brother stained the case of the CD where the singer wrote the autograph.

Use: The younger brother stained the case of the CD where the singer wrote her autograph.

11. ha'kase-wa kaga'kusha-ga sa'Npuru-o atsu'meta kokuritsuko'oeN-no ike'-o otozureta.

Literal: The doctor visited the pond in the (national) park where the scientist collected samples.

Use: The doctor visited the pond of the park where the scientist collected samples.

12. keibi'iN-wa kaNko'okyaku-ga haNdoba'ggu-o oto'shita bijutsu'kaN-no gihutosho'ppu-o teNkeNshita.

Literal: The security guard checked the gift shop of the gallery where the sightseeing visitor dropped her purse/handbag.

Use: The security guard checked the gift shop of the gallery where the tourist dropped her purse.

13. shoosetsuka-wa ki'sha-ga sukya'Ndaru-o aba'ita ba'a-no chika'shitsu-o kono'Nda.

Literal: The novelist liked the basement of the bar where the journalist disclosed the scandal.

Use: as is.

14. buchoo-wa shiNnyuusha'iN-ga shi'iru-o haga'shita tiipo'tto-no pa'kkeeji-o kakuniNshita.

Literal: The manager checked the package of the teapot which the new employee removed the sticker from.

Use: The manager checked the package of the teapot where the new employee peeled off the sticker.

15. terebiba'Ngumi-wa hito'bito-ga re'tsu-o tsuku'tta depa'ato-no sushi'ya-o hihanshita.

Literal: The TV program criticized the sushi bar of the department store where people formed a line (lined up).

Use: The TV program criticized the sushi bar of the department store where people were forming a long line.

16. sense'e-wa se'eto-ga ana'-o ho'tta ka'daN-no sumi'kko-o mitsumeta.

Literal: The teacher stared at the corner of the garden where the student dug a hole.

Use: The teacher scrutinized the corner of the garden where the student dug a hole.

17. ne'ko-wa oba'asaN-ga mo'ofu-o ka'keta so'fa-no semo'tare-o hikka'ita.

Literal: The cat scratched the back of the sofa where the grandmother hung a blanket.

Use: as is.

18. pi'ero-wa kuma'-ga mae'asi-o ka'keta saNri'Nsha-no migiha'Ndoruo-o sasaeta.

Literal: The clown supported the right handle of the tricycle where the bear hung his paws.

Use: The clown supported the handle of the tricycle where the bear hung his paws.

19. maneejaa-wa sakko-ga i'Ntabyuu-o u'keta re'sutoraN-no tokubetsu'shitsu-o yoyakushita.

Literal: The {manager/agent} reserved the special room at the restaurant where the writer received an interview.

Use: The agent reserved the room of the restaurant where the writer gave an interview.

20. shuushuuka-wa chookokuka-ga hanamo'yoo-o ho'tta dainiNgute'eburu-no ashi'-o zessaNshita.

Literal: The collector praised the leg of the dining table where the sculptor carved flower figures.

Use: The collector admired the leg of the table where the sculptor carved flower figures.

21. keikaN-wa doroboo-ga ma'aku-o tsuke'ta bi'ru-no ma'do-o watta.

Literal: The policeman broke the window of the building where the thief had placed a mark.

Use: as is.

22. shuza'ihaN-wa yuuboku'miN-ga ka'mi-o oga'Nda yamao'ku-no i'do-o sagashia'teta.

Literal: The TV crew found the well deep in the mountains where the nomads worshipped gods.

Use: The TV crew found the well of the mountain where the nomads worshipped gods.

23. hi'sho-wa uNsoogyo'osha-ga koNpyu'utaa-o surika'eta koojo'o-no jimusho'-o tsukito'meta.

Literal: The secretary found the administrative office of the factory where the mover secretly exchanged the computer.

Use: The secretary found the office of the factory where the mover secretly exchanged the computer.

24. uNte'Nshu-wa shachoo-ga ta'ima-o kaku'shita yunyu'usha-no do'a-o migaita.

Literal: The driver polished the door of the imported car where the president hid the marijuana.

Use: The driver polished the door of the Mercedes where the president hid the marijuana.

25. shooneN-wa kurasume'eto-ga me'mo-o noko'shita no'oto-no ka'baa-o hikisaita.

Literal: The boy tore the cover of the notebook where the classmate left a memo.

Use: as is.