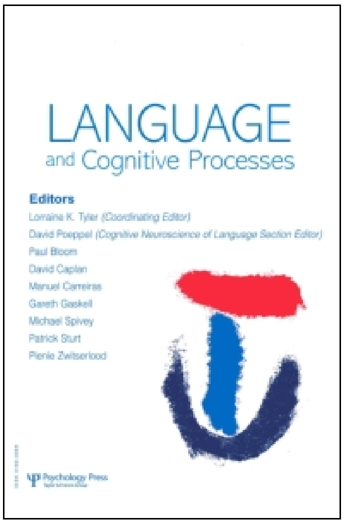

Language and Cognitive Processes
Publication details, including instructions for authors and subscription information:
http://www.informaworld.com/smpp/title~content=t713683153

The implicit prosody hypothesis and overt prosody in English

Sun-Ah Jun[a]
[a] Department of Linguistics, University of California, Los Angeles, CA, USA

First published on: 05 August 2010

To cite this Article Jun, Sun-Ah(2010) 'The implicit prosody hypothesis and overt prosody in English', Language and Cognitive Processes, 25: 7, 1201 — 1233, First published on: 05 August 2010 (iFirst)
To link to this Article: DOI: 10.1080/01690965.2010.503658
URL: http://dx.doi.org/10.1080/01690965.2010.503658

LANGUAGE AND COGNITIVE PROCESSES
2010, 25 (7/8/9), 1201–1233

Psychology Press
Taylor & Francis Group

The implicit prosody hypothesis and overt prosody in English

Sun-Ah Jun
Department of Linguistics, University of California, Los Angeles, CA, USA

This study investigates the validity of the Implicit Prosody Hypothesis (IPH) by examining default phrasing in English, a low attachment language, in overt prosody generated by reading aloud sentences where a complex noun phrase serves as the head of a relative clause (NP1 NP2 RC). The prosodic phrasing of 27 sentences collected from 36 speakers was transcribed by three ToBI-trained labellers. Results show that, counter to the predictions of the IPH, the most common prosodic phrasing was (NP1 NP2)//(RC), which would be expected for high attachment preference languages. This default phrasing was found to be influenced by the length of the RC and by syntactically disambiguating properties of the RC verb (i.e., number agreement) only when the RC was short. It was suggested that the prosody generated in silent reading would not necessarily be the same as the prosody generated in reading aloud, especially when reading without skimming the material in advance. Based on the current results and data from previous studies, various ways to access implicit prosody are proposed.

Keywords: Implicit prosody; Overt prosody; Default phrasing; Read aloud.

Correspondence should be addressed to Sun-Ah Jun, Department of Linguistics, University of California, Los Angeles, CA 90095-1543, USA. E-mail: jun@humnet.ucla.edu

This work was supported by a UCLA Faculty Research Grant. Many thanks to Michael Wagner and five anonymous reviewers for providing insightful suggestions and helpful comments. I would also like to thank Janet Fodor, Kiwako Ito, Shari Speer, and the audience at the Conference on Prosody and Language Processing held at Cornell University in 2008 and to Bruce Hayes, Pat Keating, and other members of the UCLA Phonetics Lab for their suggestions and feedback. Special thanks to Janet Fodor and Anouschka Bergmann for sending me helpful references; to J'aime Roemer, Molly Shilman, and Chad Vicenik for their help in prosodic transcriptions; to Timothy Arbisi-Kelm, Peter Graff, and Matthew Weitz for data collection; to Sameer ud Dowla Khan for proofreading the manuscript; and to Kristine Yu, Megha Sundara, Amy Schafer, and UCLA Statistics Consulting Group for their help in statistics.

http://www.psypress.com/lcp DOI: 10.1080/01690965.2010.503658

INTRODUCTION

Since Cuetos and Mitchell's (1988) seminal work, it has been known that languages differ in the attachment preference of a relative clause (RC) when the head noun is a complex noun phrase (i.e., has two nouns) and the relative clause is not biased towards modifying either noun. For example, in (1), the attachment of the RC, *who was on the balcony*, is ambiguous. It can modify the adjacent, embedded head noun *the actress*, i.e., attach low, or the non-adjacent head noun, *the servant*, i.e., attach high.

(1) *Someone shot the servant of the actress who was on the balcony.*
(Cuetos & Mitchell, 1988)

In processing sentences like (1), high attachment of an RC is preferred by speakers of Dutch (Brysbaert & Mitchell, 1996), French (Zagar, Pynte, & Rativeau, 1997), German (Hemforth, Konieczny, Scheepers, & Strube, 1998), Japanese (Kamide & Mitchell, 1997), Spanish (Cuetos & Mitchell, 1988), Croatian (Lovric, 2000, 2001), Greek (Papadopoulou & Calhsen, 2003), and Hindi (Vasishth, Agnihotri, Fernandez, & Bhatt, 2004), while low attachment of an RC is preferred by speakers of Arabic (Quinn, Abdelghany, & Fodor, 2000), English (Frazier & Clifton, 1996), Norwegian, Romanian, and Swedish (Ehrlich, Fernández, Fodor, Stenshoel, & Vinereanu, 1999; see Fodor, 1998; Fernández, 2003; Augurzky, 2006 for more details[1]).

This apparent cross-linguistic difference has challenged the universalist view of sentence processing. Psycholinguists have hypothesised that there exists a universal, innate mechanism in human sentence processing which guides a parser (e.g., minimal attachment, late closure, or right association by Kimball, 1973; Frazier & Fodor, 1978; Frazier, 1979; Frazier & Rayner, 1988; Fodor, 1998). The main idea of the universal parsing mechanism is to parse efficiently by not building an unnecessary structure and not changing an existing structure at the level higher than the current clause or phrase, i.e., having fewer syntactic nodes and attaching low. This mechanism explains problems in parsing garden-path sentences (Frazier & Fodor, 1978) and predicts the RC in sentence (1) to attach low cross-linguistically.

In order to explain this cross-linguistic variation of RC attachment, several accounts have been proposed, e.g., the frequency-based *Tuning*

[1] The attachment preference is not fixed even in the same language. The preference varies depending on a variety of factors such as the type of adposition within the complex noun phrase (theta-assigning or not), the type of modifiers (adpositional phrase or relative clause), and the early or late stage of processing (e.g., Carreiras & Clifton, 1993, 1999; De Vencenzi & Job, 1994; Brysbaert & Mitchell, 1996). The attachment preference discussed in this paper is confined to the attachment of the RC in the "NP1 *of* NP2 RC" structure, at a later stage of processing.

Hypothesis (e.g., Cuetos & Mitchell, 1988; Mitchell, 1994; Brysbaert & Mitchell, 1996), memory-based *Predicate Proximity/Recency* (e.g., Gibson, Pearlmutter, Canseco-Gonzales, & Hickock, 1996), argument/adjunct distinction and discourse-based *Construal* (e.g., Gilboy, Sopena, Clifton, & Frazier, 1995; Frazier & Clifton, 1996, 1997), and prominence and discourse-based *Anaphoric Binding* (e.g., Hemforth, Konieczny, Scheepers, & Strube, 1998), but all have encountered counter-examples and have had problems in explaining why the same syntactic structure varying only constituent length triggers different attachment preferences (see below for the effect of length in more detail).

More recently, Fodor (1998, 2002) proposed a new account, the prosody-based *Implicit Prosody Hypothesis (IPH)*. She claims that the cross-linguistic difference in attachment preference is due to a cross-linguistic difference in prosody, specifically the default prosodic phrasing of a sentence, i.e., the prosodic phrasing produced in broad/neutral focus condition or phrasing produced without assuming a specific discourse context. The default phrasing was claimed to favour the syntactic analysis associated with it. Since the attachment preference data were collected from processing a written text, i.e., silent reading, the prosody was not overt but implicit. This hypothesis can explain the effect of constituent length on attachment resolution. Numerous experiments on sentence processing, based on off-line questionnaire data, have shown that an RC tends to attach low when the RC is short but high when the RC is long (e.g., Fernández & Bradley, 1999 on Spanish; Quinn, Abdelghany, & Fodor, 2000 on French, English, and Arabic; Lovric et al., 2000, 2001 on Croatian; Pynte & Colonna, 2000 on French; Fernandez, 2003 on Spanish and English; Hirose, 1999 and Jun & Koike, 2003 on Japanese; Wijnen, 2004 on Dutch; Vasishth, Agnihotri, Fernández, & Bhatt, 2004 on Hindi).

Assuming that implicit prosody is equal to explicit or overt prosody, Fodor and her colleagues (e.g., Maynell, 1999; Quinn et al., 2000; Lovric et al., 2000, 2001) examined prosodic phrasing in overt prosody. They claimed that speakers interpret a prosodic break before an RC (in a sequence of NP1 NP2 RC) as a marker of a stronger syntactic boundary, which prompts high attachment. This suggests that speakers of a language with high attachment preference in silent reading tasks would tend to produce a specific pattern of breaks in their overt prosody: a big prosodic boundary between an RC and its adjacent head noun in the default phrasing of a sentence, but a weak boundary between the two head nouns, NP1 and NP2. Since the processing of boundary strength at one point is relative to the strength of the preceding boundary (e.g., Carlson, Clifton, & Frazier, 2001; Clifton, Carlson, & Frazier, 2002; Frazier, Clifton, & Carlson, 2004), it can be hypothesised that, for speakers of high attachment preference languages, a prosodic break between an RC and the adjacent head noun would be *bigger* than the break between the two head nouns, while the opposite relation would be true for speakers of low attachment preference languages.

This hypothesis was tested using Japanese and Korean data in Jun and Koike (2003) and Jun and Kim (2004), respectively. In both Japanese and Korean, the word order is reversed with respect to that of English. A relative clause comes before a head noun, and the "NP1 *of* NP2" structure (e.g., *servant of the actress*) in English is expressed as "NP2's NP1" structure (e.g., *the actress's servant*). Thus, the default prosodic phrasing is expected to be "(RC) (NP2's NP1)" if these languages prefer high attachment of an RC, but "(RC NP2's) (NP1)" if they prefer low attachment. Assuming that reading out-of-the-blue is representative of the default prosody projected in silent reading, Jun and Koike (2003) examined the prosodic phrasing of 48 ambiguous target sentences in Japanese produced by 30 speakers of Tokyo Japanese (randomised with 36 fillers). Speakers were randomly divided into two groups (Skim group vs. Non-skim group) and read the target sentences out-of-the-blue. The Skim group briefly skimmed the sentences before reading and Non-skim group did not. Three weeks later, the same subjects participated in an off-line questionnaire sentence processing experiment.[2] The target sentences differed in the length of RC, both the length and accentedness (i.e., presence or absence of a lexical pitch accent) of the head noun, as well as the location of the target structure. The RC NP2's NP1 structure was located either sentence-initially (i.e., a complex subject phrase) or sentence-medially (i.e., a complex object phrase); the verb was always sentence-final.

The results showed that Japanese speakers prefer high attachment (66%), confirming previous results (e.g., Kamide & Mitchell, 1997), and produced a bigger prosodic break between an RC and the following head noun than that between the two head nouns (i.e., (RC)//(NP2's NP1), called an *early break*) 67% of the time, supporting the predictions of the IPH. Results also showed an effect of RC length in both production and processing. Longer RCs showed more instances of early breaks than shorter RCs, and longer RCs were attached high more often than short RCs. Even though the length of NP1 and NP2 had no effect on attachment or prosodic phrasing, other non-structural factors such as the accentedness of the head noun affected both production and processing. An unaccented word tended to be phrased together with the following word, thus providing more instances of early breaks when NP2 was unaccented, i.e., (RC)//(unaccented NP2's NP1), and

[2] We did this so that the attachment decisions made by our subjects were not influenced by their familiarity with the sentence materials. However, we later discovered that subjects performed very similarly regardless of how well they remembered the material. Thus, in Jun and Kim's experiment on Korean, the attachment data were collected immediately following the reading task.

showed more instances of high attachment. Furthermore, the location of the RC affected the prosodic phrasing of the structure, although only for the Skim group. More instances of *late breaks*, (RC NP2's)//(NP1), were found sentence-medially than sentence-initially. In sum, the experiment revealed that there is a correlation between prosodic phrasing and the processing of RC attachment[3] and that native speakers' performance is influenced by prosodic factors such as the length of a constituent, the accentedness of a lexical item, and the location of the target structure.

A similar experiment was performed on Korean data (32 target sentences, 32 fillers, and 30 speakers) by Jun and Kim (2004) and Jun (2007). As in the Japanese study, subjects were divided into two Groups (Skim vs. Non-skim), read each sentence out-of-the-blue, and performed an off-line questionnaire processing experiment right after the production experiment. The results showed that Korean speakers also preferred high attachment overall (60%), and the early break pattern was more common in their production when analysed in the revised model of Korean intonation (Jun, 2006, 2007), supporting the IPH. As in Japanese, the effect of RC length was found in both production and processing: longer RCs showed more instances of early breaks (70%) than did shorter RCs (34%), and were attached high more often than were shorter RCs (70% vs. 48%). Unlike Japanese, however, Korean data showed the Skim vs. Non-skim Group effect and RC Location effect only in processing; the Skim group chose high attachment more often than the Non-skim group and a sentence-initial RC triggered more high attachment than a sentence-medial RC. In sum, Korean data also supported the IPH by showing a correlation between RC attachment and overt prosodic phrasing of the target structure.

Since both Japanese and Korean prefer high attachment, it would be important to see if a low attachment preference language such as English also shows prosodic patterns predicted by the IPH. The current study reports a production experiment where English sentences containing the target structure, NP1 NP2 RC, are produced in out-of-the-blue reading. Since English speakers tend to prefer low attachment, we expect to find a smaller prosodic break between the RC and the preceding head noun than the break between the two head nouns, i.e., (NP1)//(NP2 RC), in the majority of cases. In other words, the break after NP1 would be bigger than that after NP2, represented in this paper as "NP1 > NP2".

[3] However, a direct mapping between the attachment and prosodic phrasing of an individual sentence was not good (about 63%). Similar results were found in the studies by Bergmann and her colleagues (see the Discussion section). It was suggested in Jun and Koike (2003) that the correlation is the result of group behavior. It may be a by-product of the common default phrasing being (RC)//(NP1 NP2) and high attachment preference in Japanese.

PRODUCTION EXPERIMENT ON ENGLISH: METHOD

Materials

Twenty-seven target sentences differing in the length of the RC (nine Short RC sentences, nine Medium RC sentences, and nine Long RC sentences) and in the bias of the NP (nine No-bias, nine NP1-bias, and nine NP2-bias sentences) were chosen from the literature on English RC attachment: 14 sentences from Fernandez (2003), seven from Carreiras and Clifton (1993), two from Frazier (1990), and four from Deevy (1999, 2000). All target sentences had the structure "Subject Verb *the* NP1 *of the* NP2 RC". Sentences with a short RC had 10–13 syllables before the RC and a 3–5 syllable RC; sentences with a medium RC had 10–14 syllables before the RC and a 7–10 syllable RC; sentences with a long RC had 13–15 syllables before the RC and a 13–18 syllable RC. The bias of the NP was controlled by manipulating number and gender agreement. (Here, the term "bias" should be interpreted as "disambiguated" either syntactically by number or semantically/pragmatically by gender.)[4] It was expected that varying NP Bias would provide the prosodic phrasing data corresponding to high (NP1-bias) or low (NP2-bias) attachment, providing a reference of phrasing when the phrasing of the target sentences in the No-bias condition was examined. Examples of the three NP Bias types by number and gender agreement are given in (2) and (3), respectively, and the prosodic phrasing predicted by the IPH is shown in (4). In each script, there were 19 Number-bias sentences and 8 Gender-bias sentences in each NP Bias condition.[5] A full list of target sentences used in the experiment is given in Appendix 1.

(2) NP Bias types by number agreement (data modified from Fernandez, 2003)
  a. No-bias
     My friend met the aide of the detective that was fired.
  b. NP1-bias
     My friend met the aide of the detectives that was fired.
  c. NP2-bias
     My friend met the aides of the detective that was fired.

[4] Specifically, the number bias means "syntactically disambiguated towards NP1 or NP2" and the gender bias means "semantically or pragmatically disambiguated towards NP1 or NP2". In this paper, the term "bias" is used for convenience of naming. I thank the reviewer who pointed out the ambiguity and possible inappropriateness of the term "bias".

[5] The experiment was not designed to test any effect of bias types on phrasing, so the number of sentences of the two bias types (Number/Gender) is not balanced. The nineteen Number bias sentences include eight Long RC, four Med RC, and seven Short RC sentences, and the eight Gender bias sentences include one Long RC, five Med RC, and two Short RC sentences.

(3) NP Bias types by gender agreement (data modified from Carreiras & Clifton, 1993)
   a. No-bias
      The police arrested the sister of the nursemaid who recently gave birth to twins.
   b. NP1-bias
      The police arrested the sister of the handyman who recently gave birth to twins.
   c. NP2-bias
      The police arrested the brother of the nursemaid who recently gave birth to twins.

(4) NP Bias types and expected prosodic phrasing by the IPH
      No-bias => (NP1)//(NP2 RC)
      NP1-bias => (NP1 NP2)//(RC)
      NP2-bias => (NP1)//(NP2 RC)

Three scripts were made, each containing different bias types of the 27 target sentences: nine sentences per RC Length type and nine sentences per NP Bias type. Therefore, in each RC Length category, there were three No-bias sentences, three NP1-bias sentences, and three NP2-bias sentences.[6] In each script, 27 target sentences were randomised with 30 filler sentences whose structures varied widely and did not include the target structure. The length of the fillers was similar to that of the target sentences: 10 short sentences (12–18 syllables), 10 medium sentences (20–24 syllables), and 10 long sentences (27–35 syllables). The same fillers were used for all scripts.

Subjects

Thirty-six speakers of American English, mostly undergraduate students at University of California, Los Angeles (UCLA), participated in the experiment. Subjects were randomly assigned to one of three groups (Groups 1, 2, and 3) of equal size. The scripts differed by group, so subjects in each group read only one biased version of each target sentence.

[6] The three NP Bias types were generated from the same sentence by changing the gender or number of relevant head noun or the RC. However, the three RC Length type sentences were not related to each other. That is, Long RC sentences were not created by adding more words to Short RC sentences. Therefore, the data is not designed to compare acoustic durations of the target phrase across NP Bias or RC Length conditions.

Procedures

Subjects read each sentence twice in the sound booth at the UCLA Phonetics Lab. They were told to read it at a comfortable speed without skimming it beforehand. Each sentence was written in one line on a page to avoid any effect of input segmentation on reading (Just, Carpenter, & Woolley, 1982; Ferreira & Henderson, 1990; Gilboy & Sopena, 1996; Swets, Desmet, Hambrick, & Ferreira, 2007; Traxler, 2009). No questions on attachment or the content of the sentences were asked. The recordings were digitised at 22 kHz. The first repetition was segmented for further analysis unless disfluent; when the first repetition was disfluent (which occurred about 15% of the time), the second repetition was used. Three ToBI-trained labellers independently transcribed the tones and break indices of each utterance starting at NP1, following the Mainstream American English Tones and Break Indices (MAE_ToBI) conventions for tones (Beckman & Hirschberg, 1994; Beckman & Ayers-Elam, 1997) and a modified ToBI convention for break indices. In the tone tier, the presence of a pitch accent (but not the type of pitch accent) was transcribed on NP1, NP2, and the first prominent word in the RC. When the end of NP1 or NP2 was marked by a phrasal break, the type of phrase accent (i.e., L- or H-) and/or a boundary tone (i.e., L% or H%) was transcribed after NP1 and NP2. Break indices were labelled only after NP1 and NP2. Labellers determined the tonal categories and break indices by listening to the sound files and referring to the waveforms/spectrograms and pitch tracks displayed in *PitchWorks* (Scicon R&D). The break indices used included 1 (phrase-medial-word boundary), 3 (Intermediate Phrase boundary), 4 (Intonation Phrase boundary), 1- (smaller than a word boundary), 3- (juncture between 1 and 3), and 4- (juncture between 3 and 4), following the MAE_ToBI conventions. For the mismatch cases between tones and break indices, however, a new labelling scheme was used: '1m' was used when the degree of juncture matched that of a phrase-medial word boundary but a phrase accent was nevertheless present, and '3m' was used when the degree of juncture matched that of an Intermediate Phrase boundary but with no accompanying phrase accent. When quantifying the break strength, '1m' was treated as '1' and '3m' was treated as '3'. The uncertainty break indices, 1-, 3-, and 4-, were converted to 0.5, 2, and 3.5, respectively. Thus, the break index (BI) 2 in this paper refers to a juncture between BI 1 and BI 3, not an index for the mismatch between a perceived juncture and a tonally marked prosodic boundary as in MAE_ ToBI.

There were a total of 972 sentences (29 sentences × 36 speakers) transcribed. Three labellers independently transcribed different but overlapping portions of the data as shown in Table 1. Labeller 1 transcribed all three types of NP Bias from 24 speakers (eight speakers in each Group, total 648 sentences), and Labeller 2 transcribed all three types of NP Bias from all

TABLE 1
Number of speakers and sentences transcribed by each labeller

	Number of speakers	*Number of sentences in each bias type*	*Total number of sentences transcribed*
Labeller 1	24 speakers	9 No-bias sentences	216
	24 speakers	9 NP1-bias sentences	216
	24 speakers	9 NP2-bias sentences	216
Labeller 2	36 speakers	9 No-bias sentences	324
	36 speakers	9 NP1-bias sentences	324
	36 speakers	9 NP2-bias sentences	324
Labeller 3	36 speakers	9 No-bias sentences	324

36 speakers (total 972 sentences). Labeller 3 transcribed only No-bias items from all 36 speakers (total 324 sentences). Thus, all No-bias type sentences from 24 speakers were transcribed by all three labellers, and NP1- and NP2-bias type sentences were transcribed by Labellers 1 and 2 only.

RESULTS

There was no Group difference in the phrasing, so results are reported by combining data across groups. Only break index data are presented because tonal transcription data show that NP1, NP2, and at least one word in the RC were always pitch accented and the types of phrase accent or boundary tone did not differ by RC Length and NP Bias type.[7] Due to subjectivity in the perception of the degree of juncture by a transcriber, all three labellers did not always agree on a break index. To show the range of variation in the labelling, the break index data of an individual transcriber are shown separately, except for Figure 2. Though the transcriptions do not always match among the labellers, the inter-transcriber agreement rate was nevertheless fairly good. The agreement rate on raw break index data (based on the nine No-bias items from 24 speakers transcribed by all three labellers)

[7] In general, high (H) tones were more common than low (L) tones after NP1: H occurred 70% of the time in Labeller 1's data, 85% for Labeller 2, and 91% for Labeller 3. After NP2, however, H occurred 35% of the time in Labeller 1's data, 54% for Labeller 2, and 75% for Labeller 3. A higher rate of disagreement among the labellers after NP2 is likely due to the difficulty of categorising tones in a reduced pitch range. However, within each labeller, the distribution of H and L tones after NP1 and NP2 was similar across the RC Length and NP Bias types.

was 53–66% using the stringent method employed in Silverman et al. (1992).[8] When the break index was simplified into two categories[9]—a break index marking the edge of a domain smaller than or equal to the size of a word ($BI \leq 1$) vs. that of a domain larger than a word ($BI > 1$)—the agreement rate increased to 72–88%. Specifically, the agreement rate was 79.5% and 87.8% for the BI after NP1 and NP2, respectively, and 71.6% for the relative break strength between NP1 and NP2. The agreement rate on the relative break strength data of the biased items (18 items from 24 speakers, transcribed by two labellers) was 70%.

Break index after NP1 and NP2

Since English speakers have been shown to prefer low attachment, it was predicted, following the IPH, that the break index after NP1 would be, overall, greater than that after NP2. However, the results show that—counter to the prediction of the IPH—the most common phrasing pattern of the target structure in the No-bias condition was a bigger boundary after NP2 than after NP1, i.e., NP1 < NP2.

Figure 1 shows the proportions of $BI \leq 1$ (a word boundary or smaller) and $BI > 1$ (larger than a word boundary) after each NP in the No-bias condition for each labeller. Labeller 1's data are from 24 speakers and Labellers 2 and 3's data are from 36 speakers. As shown in Figure 1, about half of the breaks after NP1 are larger than word size, and about 90% of the breaks after NP2 are larger than word size. That is, the end of NP2 was more often marked by a bigger break than the end of NP1, for all labellers. The difference between the NP1 break pattern and the NP2 break pattern is significant for all labellers: Labeller 1 ($\chi^2 = 72.07$, $df = 1$, $p < .001$), Labeller 2 ($\chi^2 = 55.23$, $df = 1$, $p < .001$), Labeller 3 ($\chi^2 = 83.16$, $df = 1$, $p < .001$). There was no significant inter-transcriber difference when comparing data from the 24 speakers transcribed by all three.

To make sure that the break after NP2 is larger than the break after NP1 in the same sentence, the relative strength of breaks between NP1 and NP2 in each utterance is measured by the ratio of the break after NP2 to the break

[8] This method was used in calculating labeller agreement in English ToBI data, labelled by Silverman and colleagues (1992). Agreement was calculated across all possible pairs of transcribers for each word. For example, four labellers (a, b, c, and d) would produce six possible transcriber pairs (ab, ac, ad, bc, bd, and cd). If three of four transcribers (a, b, and c) agree, only three of six pairs will match (ab, ac, and bc; but not ad, bd, and cd), and this would result in 50% agreement. If two out of three labellers (a and b) agree, only one of three pairs will match (ab, but not ac and bc), resulting in 33.3% agreement.

[9] This simplification was introduced because the great majority of break indices were either 1 or 3; the small number of breaks transcribed with particular labels made the χ^2 test inappropriate for the data in its original form.

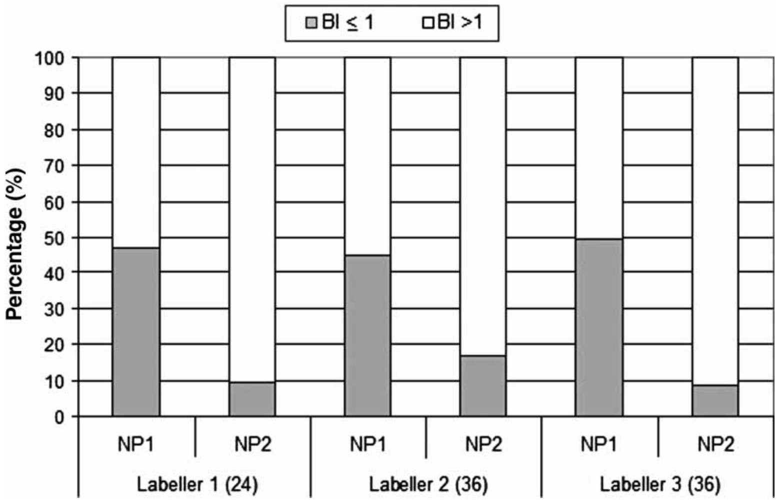

Figure 1. Frequency (in %) of BI ≤ 1 (break index corresponding to a word boundary or smaller) and BI >1 (break index bigger than a word boundary), after each NP in the no-bias condition, for each labeler. The number in parentheses is the number of speakers each labeler's data is based on.

after NP1 (represented as "NP2/NP1"): if the break after NP2 is bigger than that after NP1, the ratio is >1; if the break after NP2 is smaller than that after NP1, the ratio is <1; and if the break after NP2 is the same as that after NP1, the ratio $= 0$. The average ratio of NP2/NP1 was 1.98 for Labeller 1, 1.70 for Labeller 2, and 1.76 for Labeller 3, confirming that speakers generally produced a larger break after NP2 than after NP1 in the same sentence.

The effect of RC length on default phrasing

The length of the relative clause (RC Length) influenced the degree of break after each noun. A stronger effect of RC Length on the degree of juncture was found after NP2 than after NP1. Figure 2 shows the frequency (in %) of raw BI categories found after NP1 and NP2 in each RC Length type in the No-bias condition. Data from the three labellers (24 speakers each) are combined. The observed frequency of each BI category across RC Lengths was marginally significant after NP1 ($\chi^2 = 15.001$, df $= 8$, $p = .057$). Here, the Short RC condition had fewer BI $= 1$ than did the longer RC conditions. On the other hand, the observed distribution of BI categories after NP2 was significantly different across RC Length types ($\chi^2 = 33.981$, df $= 8$, $p < .001$). The Short RC condition had fewer phrase boundaries and more word-size boundaries than did the Long or Medium RC conditions. This pattern is

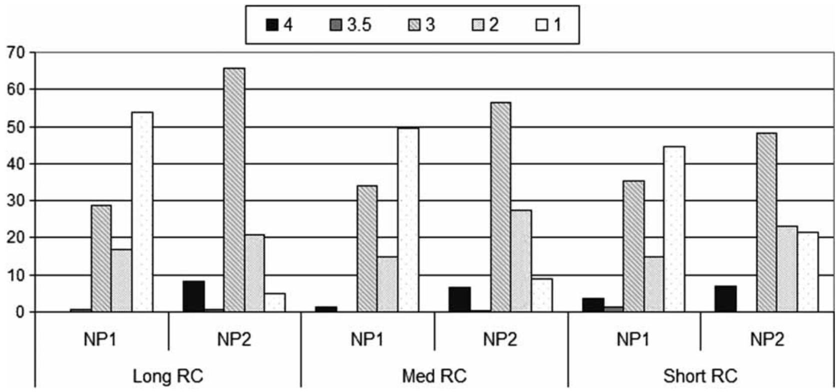

Figure 2. Frequency (in %) of raw BI categories after each NP in the no-bias condition, divided by RC length type. Data from all three labelers are combined (24 speakers per labeler). BI 2 is a juncture between BI 1 (word boundary) and BI 3 (Intermediate Phrase boundary). BI 3.5 is a juncture between BI 3 (Intermediate Phrase boundary) and BI 4 (Intonation Phrase boundary).

TABLE 2
The mean and standard deviation of raw BI values after NP1 and NP2 by RC Length and NP Bias for each labeller

			NP1		*NP2*	
Bias types	*Labellers*	*RC length*	*Mean*	*SD*	*Mean*	*SD*
No-bias	Labeller 1	Long	1.864	0.92	2.921	0.66
		Med	1.944	1.01	2.824	0.80
		Short	2.232	1.10	2.634	0.90
	Labeller 2	Long	1.885	0.91	2.680	0.72
		Med	1.901	0.89	2.554	0.76
		Short	1.922	0.91	2.137	0.92
	Labeller 3	Long	1.835	0.90	2.721	0.69
		Med	1.857	0.93	2.622	0.68
		Short	1.949	0.93	2.556	0.77
NP1-bias	Labeller 1	Long	1.711	0.88	2.936	0.69
		Med	1.992	1.03	2.978	0.68
		Short	1.845	0.94	2.911	0.88
	Labeller 2	Long	1.874	0.87	2.722	0.77
		Med	1.875	0.92	2.690	0.73
		Short	1.833	0.89	2.505	0.89
NP2-bias	Labeller 1	Long	1.697	0.88	2.957	0.76
		Med	1.868	1.08	2.979	0.67
		Short	2.042	1.01	2.528	1.06
	Labeller 2	Long	1.754	0.85	2.725	0.73
		Med	1.792	0.90	2.688	0.73
		Short	2.059	0.89	2.069	0.78

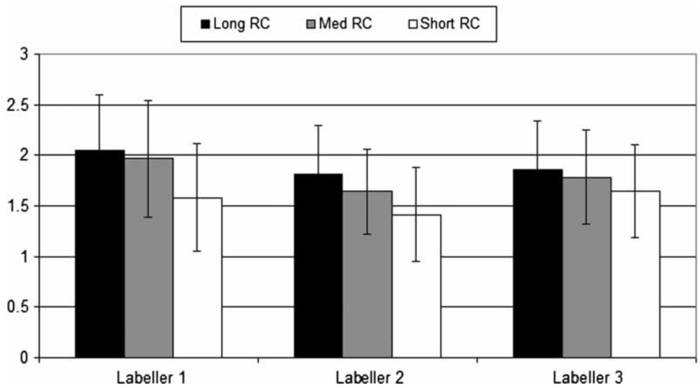

Figure 3. Average ratio of the break after NP2 to the break after NP1 across RC Length types, for each labeler (in the no-bias condition). Error bars reflect standard deviation.

confirmed in the individual labellers' data in Table 2 where the mean and standard deviation of raw BI values after NP1 and NP2 are shown in each RC Length condition.

Again, since what is important in processing is the *relative* strength of boundaries found in the same sentence (e.g., Carlson, Clifton, & Frazier, 2001; Clifton, Carlson, & Frazier, 2002; Frazier, Clifton, & Carlson, 2004), the average ratio of NP2/NP1 for each RC Length, for each labeller, is examined (see Figure 3). All data are from the No-bias condition. Here, the tendency is that the ratio is higher as the RC gets longer. A linear mixed effects model with Speaker as a random effect (model: ratio.np2.np1 $\sim$ group + length + (1|speaker)) shows that, for Labeller 1, the ratio is significantly shorter in the Short RC condition than in the Medium RC condition, and for Labeller 2, the ratio between Medium and Short RC conditions is marginally significant ($p = .052$).[10] The ratio across RC Length types was not significant for Labeller 3. The results of the statistics run in R (Baayen, 2008; Bates & Maechler, 2009) are shown in Table 3.

The effect of NP bias on default phrasing

It was expected that NP2-biased sentences would show a smaller break after NP2 than after NP1. That is, the phrasing would be (NP1)(NP2 RC) and the ratio of NP2/NP1 would be < 1 because the RC was forced to attach low (or NP2 and RC form one meaning group). Conversely, NP1-biased sentences were expected to show a bigger break after NP2 than after NP1, i.e., the phrasing would be (NP1 NP2)(RC) and the ratio of NP2/NP1 > 1, as the RC was forced to attach high. However, as shown in Figure 4, the average ratio

[10] The item is not included as a random factor because items are not repeated for RC Length.

TABLE 3

Results of a linear mixed effects model for ratio in No Bias condition, with Group and RC Length as fixed predictors and Speaker as a random effect (model: ratio.np2.np1 ~ group + length + (1|speaker)), for each labeller. Non-significant predictors are shaded in grey and a marginally significant predictor is shaded in light grey.

No bias condition				
Predictor	Coef. β	*SE*(β)	*t*	*p*(MCMC)
Labeller 1 ($R^2 = 0.25$)				
Intercept	1.96	0.22	8.88	1e–4
Group(B)	0.04	0.28	0.16	0.85
Group(C)	−0.02	0.28	−0.06	0.95
Length(short)	−0.38	0.17	−2.22	0.028
Length(long)	0.15	0.18	0.86	0.40
Labeller 2 ($R^2 = 0.19$)				
Intercept	1.53	0.14	11.03	1e–4
Group(B)	0.18	0.17	1.10	0.22
Group(C)	0.17	0.17	0.98	0.25
Length(short)	−0.24	0.12	−1.92	0.052
Length(long)	0.17	0.12	1.35	0.18
Labeller 3 ($R^2 = 0.17$)				
Intercept	1.85	0.15	12.24	1e–4
Group(B)	−0.18	0.18	−1.01	0.27
Group(C)	−0.02	0.19	−0.11	0.90
Length(short)	−0.14	0.13	−1.05	0.30
Length(long)	0.10	0.13	0.78	0.43

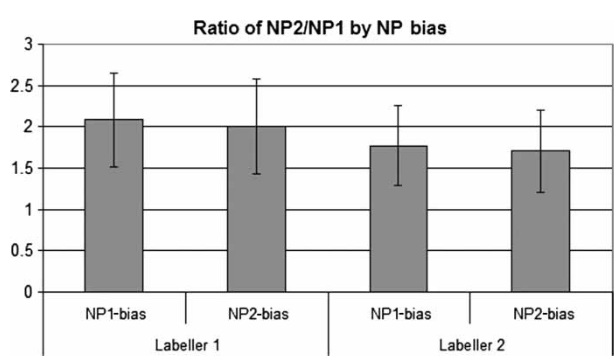

Figure 4. Average ratio of the break after NP2 to the break after NP1 across NP Bias types, for two labelers. Error bars reflect standard deviation.

of NP2/NP1 in each NP Bias condition was > 1 for both labellers (about 2.0 for Labeller 1 and 1.7 for Labeller 2), suggesting that the common default phrasing was (NP1 NP2)(RC) regardless of NP Bias type. A linear mixed effects model with Speaker and Item as random effects (model: ratio.np2.np1 $\sim$ group + np.bias + (1|speaker) + (1|item)) shows that the ratio is not significantly different across NP Bias types for both labellers.

The interaction between RC length and NP bias

However, as shown in Figure 5, the effect of NP Bias varied depending on RC Length for both labellers. The average ratio of break after NP2 to break after NP1 was lower in Short RC than in longer RC conditions for both labellers, but this was true only in No-bias and NP2-bias conditions. (This pattern can be predicted from the raw BI values in Table 2.) In the NP1-bias condition, the ratio was not influenced by RC Length. It seems that the effect of RC Length (i.e., lower NP2/NP1 ratio in Short RC) is over-ridden by the NP1-bias effect (i.e., higher NP2/NP1 ratio).

Table 4 shows the results of a linear mixed effects model for ratio, with Group, RC Length, and NP Bias as fixed predictors and Speaker as a random effect (model: ratio.np2.np1 $\sim$ group + length * np.bias + (1| speaker) for two labellers, combining all three bias conditions. It shows that even though the overall pattern of ratios across NP Bias and RC Length is similar between the two labellers, the NP Bias is not a significant predictor for Labeller 1, while both NP2 Bias and the interaction between NP2 Bias and Short RC were significant predictors for Labeller 2.

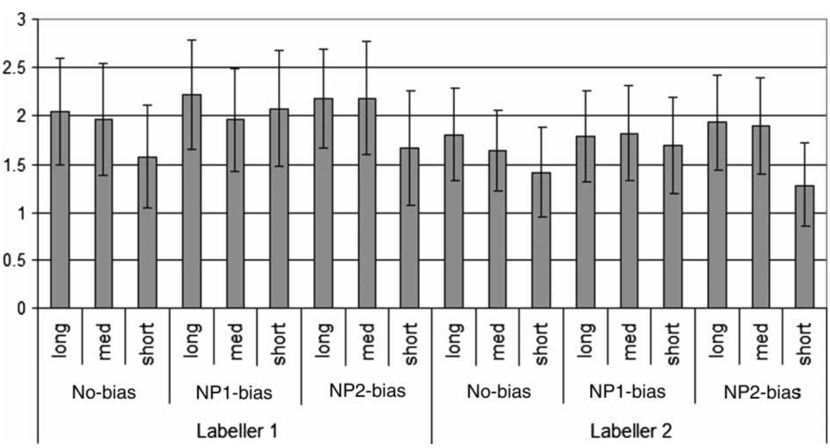

Figure 5. Average ratio of the break after NP2 to the break after NP1 across RC Length in all NP bias conditions for two labelers. Error bars reflect standard deviation.

TABLE 4
Results of a linear mixed effects model for ratio, with Group and RC Length as fixed predictors and Speaker as a random effect (model: ratio.np2.np1 ~ group + length * np.bias + (1|speaker), for two labellers in all three bias conditions. Non-significant predictors and interactions are shaded in grey.

All Bias conditions				
Predictor	Coef. β	*SE*(β)	*t*	*p*(MCMC)
Labeller 1 ($R^2 = 0.25$)				
Intercept	1.91	0.23	8.46	1e–4
Group(B)	0.14	0.28	0.51	0.56
Group(C)	0.02	0.28	0.06	0.93
length(short)	−0.39	0.17	−2.26	0.024
Length(long)	0.15	0.18	0.83	0.40
NP bias(NP1)	0.03	0.17	0.15	0.89
NP bias(NP2)	0.22	0.17	1.27	0.21
Length(short) × NP bias(NP1)	0.45	0.24	1.83	0.069
Length(long) × NP bias(NP1)	0.16	0.25	0.63	0.52
Length(short) × NP bias(NP2)	−0.13	0.24	−0.53	0.60
Length(long) × NP bias(NP2)	−0.04	0.25	−0.18	0.87
Labeller 2 ($R^2 = 0.17$)				
Intercept	1.68	0.14	12.42	1e–4
Group(B)	−0.08	0.15	−0.54	0.58
Group(C)	−0.06	0.15	−0.37	0.72
Length(short)	−0.23	0.13	−1.85	0.07
Length(long)	0.16	0.13	1.31	0.18
NP bias(NP1)	0.18	0.13	1.40	0.17
NP bias(NP2)	0.27	0.13	2.12	0.032
Length(short) × NP bias(NP1)	0.11	0.18	0.62	0.55
Length(long) × NP bias(NP1)	−0.19	0.18	−1.06	0.28
Length(short) × NP bias(NP2)	−0.39	0.18	−2.17	0.033
Length(long) × NP bias(NP2)	−0.15	0.18	−0.81	0.42

The effect of NP Bias was further examined to see if the Bias type—Number or Gender—made any difference. Figure 6 shows the ratio of NP2/NP1 by Number and Gender in only NP1- and NP2-bias conditions. (The number in parenthesis is the number of sentences in each condition. The value in each bar is based on the number of sentences multiplied by eight speakers for Labeller 1 and by 12 speakers for Labeller 2.) The lower ratio of NP2/NP1 in the Short RC and NP2-bias condition shown in Figure 5 is maintained when the NP2 is biased by Number, but to a weaker degree by Gender. A linear mixed effects model for ratio (model: ratio.np2.np1 ~ group + length * np.bias + (1|speaker)) in NP Bias subset analysis shows that the interaction of Short Length and NP2-bias was significant only in

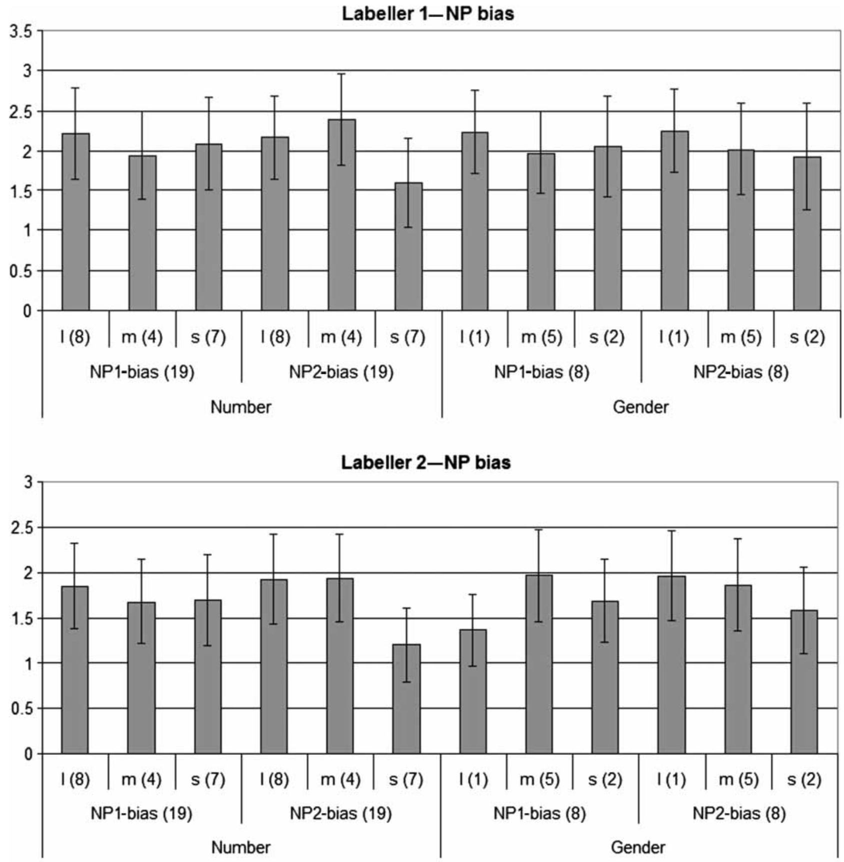

Figure 6. Average ratio of the break after NP2 to the break after NP1 across RC Length in two NP Bias conditions, divided by Number and Gender bias types. Top panel: Labeler 1. Bottom panel: Labeler 2. The number in parenthesis is the number of sentences belonging to each category. Error bars reflect standard deviation.

Number but not in Gender for both labellers.[11] The results of the statistics are shown in Table 5.

In sum, the NP Bias results suggest that both the location and the type of NP Bias are reflected, to some degree, in the prosodic phrasing of a sentence, but the effect of NP Bias on phrasing is much smaller than the overall preference for the default (NP1 NP2)//(RC) phrasing, and is limited to cases where the RC is short.

---

[11] Results from the linear mixed effects model showed that the effect of Long RC and its interaction with NP1-bias are significant for Labeller 2, but this is not considered meaningful because this is based entirely on one Long RC sentence.

TABLE 5
Results of a linear mixed effects model for ratio in NP bias by Number subset, with Group, RC Length, and NP Bias as fixed predictors and Speaker as a random effect (model: ratio.np2.np1 ~group+length * np.bias + (1|speaker)), for two labellers. Non-significant predictors and interactions are shaded in grey.

NP Bias subset (Number bias)				
Predictor	Coef. β	*SE*(β)	*t*	*p(MCMC)*
Labeller 1 ($R^2 = 0.29$)				
Intercept	1.93	0.28	6.78	1e–4
Group(B)	0.13	0.31	0.42	0.63
Group(C)	0.08	0.31	0.24	0.78
Length(short)	0.03	0.24	0.12	0.90
Length(long)	0.31	0.23	1.32	0.20
NP bias(NP2)	0.39	0.26	1.47	0.15
Length(short) × NP bias(NP2)	−0.82	0.33	−2.49	0.014
Length(long) × NP bias(NP2)	−0.41	0.33	−1.24	0.22
Labeller 2 ($R^2 = 0.19$)				
Intercept	1.82	0.16	11.10	1e–4
Group(B)	−0.22	0.15	−1.43	0.13
Group(C)	−0.06	0.16	−0.36	0.74
Length(short)	7e–4	0.17	4e–3	0.99
Length(long)	0.11	0.16	0.67	0.48
NP bias(NP2)	0.20	0.19	1.09	0.27
Length(short) × NP bias(NP2)	−0.72	0.24	−3.03	2.6e–3
Length(long) × NP bias(NP2)	−0.10	0.23	−0.42	0.66

DISCUSSION

Earlier studies on RC attachment in Japanese and Korean (Jun & Koike, 2003; Jun & Kim, 2004; Jun, 2007) showed that speakers prefer high attachment and the most common default phrasing of the "RC NP2*'s* NP1" structure is (RC)//(NP2*'s* NP1), thus supporting the Implicit Prosody Hypothesis (IPH). Studies such as Quinn et al. (2000) and Lovric et al. (2001) also examined attachment preferences as well as the phonetic realisation of overt prosody (f0 and duration) and found support for the IPH.[12]

However, in the current study, it has been found that the default phrasing of the "NP1 NP2 RC" structure in English—a language often described as

[12] Quinn et al. (2000) and Lovric et al. (2001) estimated phrasing based on phonetic data (i.e., f0 values and the duration of the local target nouns) instead of defining it phonologically. They also interpreted the relation of f0 values the same when determining the phrasing in the three languages (i.e., French, English, and Arabic). Please see Jun (2003, pp. 222–223) for concerns regarding the methodology.

preferring low attachment—did not support the IPH, at least when the default prosody is assessed from out-of-the-blue reading. The most common prosodic phrasing of the structure was (NP1 NP2)//(RC), a pattern predicted to correspond to high attachment. The same result was found in recent studies by Bergmann and her colleagues (Bergmann & Ito, 2007; Bergmann, Armstrong, & Maday, 2008). Bergmann and Ito examined the overt prosody and attachment preferences of English RCs by varying the length of NP1 and an RC. They found a preference for low attachment and (NP1 NP2)//(RC) phrasing regardless of length. This finding was confirmed in Bergmann and colleagues' (2008) study where default phrasing of the NP1 NP2 RC structure was examined in English and Spanish. In their study, 10 native speakers of each language read eight ambiguous target sentences without scanning or practicing the sentences first. After reading each sentence, speakers answered questions on attachment. The tones and the degrees of juncture of each utterance were transcribed by three ToBI-trained labellers, similar to the method adopted in the current study. They found a preference for low attachment in English and a preference for high attachment in Spanish, confirming existing findings, but the most common default phrasing pattern of the two languages was the same, i.e., (NP1 NP2)//(RC). Furthermore, they found that, as in Jun and Koike's (2003) Japanese data, the prosodic break patterns that speakers produced did not influence or match their attachment choices. The distribution of high vs. low attachment was about the same regardless of the default phrasing patterns. The main difference between English and Spanish was in the overall number of low and high attachment choices.

Thus, it seems that the most common default phrasing across several languages is to have a bigger break between an RC and the adjacent head noun than between the two head nouns, i.e., (NP1 *of* NP2)//(RC) or (RC)//(NP2*'s* NP1). This leads one to reconsider the validity of the IPH, and if it is valid, the method of assessing implicit prosody by examining overt prosody, the nature of overt prosody through reading aloud out-of-the-blue, and the special status of the NP1 NP2 RC structure.

Implicit prosody in sentence processing

It is now well known that attachment is influenced by prosodic factors such as constituent length—heavy constituents tend to attach high and light constituents tend to attach low—and that languages differ in their prosodic structures and realisations (e.g., languages whose attachment is shown to be influenced by constituent length: Brazilian Portuguese, Croatian, Dutch, English, French, German, Hindi, Japanese, Korean, and Russian; see Augurzky, 2006, Section 3.2.2.2 for more details).

Recent studies examining the effect of working memory on attachment preference also suggest that the attachment is influenced by implicit prosody. Swets et al. (2007) examined the attachment preference of Dutch and English speakers in off-line questionnaire data after assessing subjects' working memory capacity. They found that, when the test sentences were displayed in their entirety, both English and Dutch speakers showed increased preference for high attachment as their working memory capacities decreased. They interpreted this as the effect of prosody. They suggested that low capacity subjects inserted a prosodic break between NP2 and the RC but that high capacity subjects treated the entire NP1 NP2 RC structure as an integrated unit. When they displayed the test sentences in three segments/lines, i.e., NP1 NP2 // RC // matrix VP, they found that the preference for high attachment greatly increased in both English and Dutch speakers regardless of their working memory capacity. Similar results were found in Traxler's (2007, 2009) on-line eye-tracking studies. When the test sentences were displayed in their entirety (Traxler, 2007), the subjects with higher working memory capacity preferred high attachment, the opposite pattern of Swets et al. (2007), probably reflecting differences in tasks, especially on-line vs. off-line. However, when the test sentences were displayed in two lines broken after NP2 (i.e., NP1 NP2 // RC matrix VP), subjects preferred high attachment regardless of their working memory capacity (Traxler, 2009). All these studies suggest that the line break in a sentence, i.e., segmentation, may have triggered an implicit prosodic break, affecting the attachment preference. (See also Gilboy and Sopena (1996) for the effect of segmentation on the self-paced reading task of attachment processing.)

Therefore, it is clear that implicit prosody affects sentence processing. The question remaining is whether implicit prosody is equal to overt prosody, especially the prosody generated by reading aloud, out-of-the-blue.

Is implicit/silent prosody the same as prosody read aloud?

Fodor and her colleagues hypothesised that implicit prosody computed during silent reading is the same as default explicit/overt prosody and examined the prosody of reading aloud to test the IPH (e.g., Fodor, 2002; Lovric et al., 2000, 2001). Overt prosody generated in reading, especially reading aloud out-of-the-blue, is unique. It is different from the prosody generated by other types of overt prosody such as spontaneous speech or disambiguating prosody. Spontaneous speech produces longer pauses but read speech produces more frequent tonal units or prosodic breaks, resulting in smaller prosodic phrases (Howell & Kadi-Hanifi, 1991; Silverman, Blaauw, Spitz, & Pitrelli, 1992; Ayers, 1994). In spontaneous speech or when speakers are trying to disambiguate an ambiguous structure, speakers generate prosody from the information structure they intend to deliver

(Price, Ostendorf, Shattuck-Hufnagel, & Fong, 1991; Allbritten, McKoon, & Ratcliff, 1996; Schafer, Speer, Warren, & White, 2000; Snedeker & Trueswell, 2003; Schafer, Speer, & Warren, 2005; Snedeker & Casserly, 2010 this issue). Therefore, in auditory sentence processing where disambiguating prosody is employed, listeners actively use the prosodic information in the stimuli—over-riding syntactic biases when the prosody cues otherwise—and avoid garden path (e.g., Kjelgaard & Speer, 1999; Snedeker & Trueswell, 2003; Augurzky, 2006). But in the case of reading aloud out-of-the-blue, the text is already given and the speaker is not trying to emphasise or disambiguate any meaning. That is, reading aloud out-of-the-blue would generate "surface" prosody following only the roughest and simplest prosody-syntax mapping constraints, including phonological weight, i.e., rhythm and length (e.g., a break when incoming material is long; Selkirk, 2000; Watson & Gibson, 2004; Krivokapić, 2007) and structural constraints (e.g., a break before a major syntactic constituent; Selkirk, 1986, 2000; Truckenbrodt, 1999).

On the other hand, silent reading during a processing experiment would not generate surface prosody. In general, in a processing experiment where RC attachment is tested off-line, subjects read the questionnaire silently but reading time is not limited (though they are encouraged to answer as quickly as possible). The subjects have the opportunity to read the text multiple times, and as they are trying to answer questions regarding RC attachment, they are presumably reading at a deeper level.

When reading aloud out-of-the-blue, especially when reading without first skimming the sentence in detail and when the speakers are aware that their reading is recorded, the subjects' focus would be on "fluent reading", trying to make the fewest possible mistakes. In this case, a "safe" way to read, or good "performance" of reading aloud, would be to produce each content word prominently, thus putting pitch accent on every content word, and to put a prosodic break more frequently, following phonological and structural constraints (cf. Hudson, Lane, & Pullen, 2005; Rasinski, 2006).

Furthermore, in no-skim reading aloud, the speakers would not be able to prosodify the entire sentence at the beginning. Instead, they would prosodify a few words, as they read along, within their reading span. This can be evidenced by the weak effect of NP Bias, especially the pattern found in the Short RC, in the current experiment. Compared to Med or Long RCs, Short RCs are more likely to form one phrase with the preceding noun, (NP1)// (NP2 RC). However, as shown in Figure 6, the Short RC cases had lower ratios of NP2/NP1 in the NP2-bias conditions, but only when the bias was by Number, not by Gender. This is probably because the lexical item(s) disambiguating the gender information were further away from the head noun than the RC verb that provides the Number information, and also because semantic/pragmatic information (e.g., whether *a boxer* could be male or female) may require a deeper level of processing than syntactic information (cf. Rayner, Carlson, & Frazier, 1983; Mitchell, Corley, & Garnham, 1992).[13]

When the RC is not short, however, it seems that speakers put a break after NP2 before they comprehend the lexically biasing information in the RC. They do so even when the disambiguating information is close to the head noun, i.e., the RC verb for Number bias. This suggests that the surface reading is constrained by length more than by meaning. When speakers put a break before a long RC that is biased to NP2, they may realise in the later part of the RC that their earlier phrasing was not appropriate. So, it is possible that, in order to compensate for their earlier prosodic mis-phrasing, the speakers perform a sort of prosodic repair late in the RC, e.g., putting focus pitch accent on the disambiguating lexical item. This type of prosodic information was not transcribed in the current study, but a re-examination of the sound files revealed that four out of the twelve speakers who put a bigger break before the long RC indeed produced the biased item with hesitation lengthening, and one speaker produced a rising (L + H*) pitch accent on the biased item.[14] This confirms that, in no-skim reading aloud, speakers produce a break before the RC without fully comprehending the meaning of the whole RC. It is very likely that the speakers could only prosodify a smaller piece of a sentence at a time, constrained by the length and syntax of the sentence. These constraints on surface reading could be so strong that it may over-ride any implicit prosody, if it exists, associated with the target structure. This might explain why the string NP1 NP2 RC was predominantly phrased as (NP1 NP2)//(RC) regardless of RC length or bias in the current study. That is, the sequence of two head nouns contains enough syllables to form its own phrase and the relative clause complementiser signals the beginning of a clause. This implies that the most common default phrasing of NP1 NP2 RC structure across languages would be (NP1 NP2)//(RC), if the structure is produced in reading aloud out-of-the-blue without skimming the material in advance. More research on phrasing in low attachment preference languages is needed to confirm this.

Finally, silent reading is also faster than reading aloud (e.g., Taylor, 2006; Rasinski, 2006). As shown in the studies of the effects of speech rate on prosody (e.g., Rietveld & Gussenhoven, 1987; Caspers & van Heuven, 1993; Jun, 1993; Fougeron & Jun, 1998; Zu, Li, & Li, 2006), speakers reorganise

[13] It is therefore also possible that the higher ratio in the NP1-bias by Gender and Short RC condition (Figure 6) is not because speakers processed the gender-bias meaning but because they were influenced by the syntax factor, i.e., a break before RC. In either way, the result on Gender bias should not be given much significance in the current study because the result is based on very small data.

[14] I would like to thank one of the reviewers who directed my attention to the timing of prosodic planning relative to the head noun and bias types and the possibility of prosodic repair.

the prosodic structure of a sentence at faster speech rates. The phrase breaks found at normal speech rates are either reduced or removed at faster speech rates, and this reduction or removal depends on the strength of the break. It is possible that the phrase break before an RC is weak in low attachment languages such as English, and is thus easily deleted or weakened at faster speech rates, while the pre-RC break in French and other high attachment languages is strong and remains even in faster speech rates.[15] More research is needed to explore this possibility.

Accessing implicit prosody in processing

If reading aloud generates "surface" prosody, how could we access the deeper level implicit prosody? One way might be to let the reader be familiar with the written material before reading so that s/he could read it with a deeper level of understanding of the material. Bookheimer and colleagues (Bookheimer, Zeffiro, Blaxton, Gaillard, & Theodore, 1995) claimed that silent reading and oral reading invoke different cognitive processes. Based on the regional cerebral blood flow data using positron emission tomography (PET), they showed that subjects utilise an inferior temporal pathway in silent reading, which is very similar to that involved in object naming (name an object given a visual input), and in which there is direct access to the lexical entry. When reading a word aloud (given an orthographic input), however, subjects engage a superior temporal-inferior parietal route, involving the sequential interpretation of visual elements into phonological sounds, relying on sensorimotor and auditory feedback to monitor oromotor output dynamically. That is, unlike silent reading, reading a word aloud does not activate the brain area responsible for the semantic interpretation but activates the brain area for motor control. This suggests that prosody generated during silent reading may reflect the semantic and pragmatic meaning of the sentence but prosody generated during oral reading might

[15] One of the reviewers asked if the break after the RC in Japanese and Korean (from Jun & Kim, 2004; Jun & Koike, 2003) is bigger than the break before RC in English (the current study). In the prosodic structure of both English and Korean, the most common prosodic break between the RC and the adjacent noun is an Intermediate Phrase (ip) boundary. However, we cannot directly compare the realisations of ip in these two studies because the length and location of the RC is different between the two languages. If we compare the phonetic realisation of the ip in general between the two languages, the English ip has more stable phonetic cues than the Korean ip, as the former normally involves pre-boundary lengthening, phrase accent tone, and pitch reset, while the latter is marked by an optional boundary tone, pitch reset, and minor pre-boundary lengthening. For Japanese, the corresponding break is an Intonation Phrase (IP) boundary in the J_ToBI model, which combines the Intermediate Phrase and Utterance of Beckman and Pierrehumbert's (1986) model. Since the percentage of subcategories (i.e., ip-type or Utterance-type) of Japanese IP is not available in Jun and Koike, we cannot make direct comparisons between English and Japanese, either.

not. Therefore, it is possible that the prosody generated in reading aloud only after thorough comprehension of the sentence is closer to the implicit prosody employed in resolving attachment ambiguities, whether the processing task is on-line or off-line.[16] In fact, this might have been what happened in Fodor and her colleagues' work on overt prosody and the IPH (e.g., Croatian in Lovric et al., 2000, 2001; French, English, and Arabic in Quinn et al., 2000). They asked the subjects to read sentences silently for comprehension first and then read aloud, after which they let subjects answer questions on RC attachment. Allowing the subjects to read for comprehension before reading aloud might be the reason why their phrasing data differed from those found in Bergmann and her colleagues' studies as well as the current study where the subjects were told to read aloud without skimming the sentence first. An experiment is in progress to investigate the effect of reverse ordering (answering attachment questions before reading aloud) on default phrasing.

Another way to access implicit prosody might be to examine on-line processing data using event-related potential (ERP). ERP can reveal an on-line record of processing in a high temporal resolution and provide insights into the brain's activity. Recent ERP studies (e.g., Steinhauer, Alter, & Friederici, 1999; Steinhauer & Friederici, 2001) have shown that the Closure Positive Shift (CPS) component is correlated with a clause boundary cross-linguistically. Thus, observing the CPS component while subjects process an RC attachment may reveal the implicit prosody.

As in other on-line studies on processing using self-paced reading or eye-tracking methodologies (e.g., de Vincenzi & Job, 1993; Kamide & Mitchell, 1997; Kamide, Mitchell, Fodor, & Inoue, 1998; Brysbaert & Mitchell, 2000; Fernandez, 2003; Miyamoto, Nakamura, & Takahashi, 2004), ERP studies on RC attachment in German (Augurzky, 2006) and French (Pynte, Colonna, & Gola, 2003, cited in Augurzky, 2006) found an earlier preference for low attachment and a shift to high attachment when parsing the latter part of a sentence. This shift in attachment preference from low to high through the course of parsing a sentence seems to happen across languages regardless of the attachment preference found in off-line processing studies. This suggests that processing is not guided by a single principle such as late closure or implicit prosody, but by multiple principles in multiple stages of processing. It seems that in silent reading the early part of a sentence is guided by the structural information such as Late Closure and prosodic information, but the processing at the end of a sentence is guided by

[16] Silent/implicit prosody employed in the off-line questionnaire task would be the same as that in the on-line eye-tracking task. But this prosody might be different from that which is employed in the on-line self-paced reading task: in self-paced reading, the reading material is constrained by segmentation chunks, which affect prosody.

pragmatic, thematic, and discourse information. Further research is needed to investigate how many different stages of processing exist and when different parsing mechanisms—including prosody—are integrated.

CONCLUSION

The current study has shown that—counter to the predictions made by the Implicit Prosody Hypothesis—the most common prosodic phrasing of the "NP1 NP2 RC" structure in English, a low attachment preference language, was (NP1 NP2)//(RC), the prosodic phrasing expected to occur with high attachment preference languages. Since this result was based on overt prosody—specifically reading aloud out-of-the-blue—the question of whether this type of overt prosody is equal to the implicit prosody found in silent reading was discussed. It was suggested that the prosody generated in silent reading would not necessarily be the same as the prosody generated in reading aloud, especially when reading without skimming the material in advance. In no-skim reading aloud, speakers seem to prosodify or plan the phrasing of a few words at a time, influenced by the constituent length and syntax of the sentence. The phrasing showed sensitivity to meaning bias only when the relative clause was short and the disambiguating information was near the head noun. Based on these data and data from previous studies, various ways to access implicit prosody such as reading aloud after thorough comprehension of the sentence or reading at a fast rate were suggested.

Since most studies supporting the IPH have investigated the default phrasing of high attachment preference languages, we need to examine more low attachment preference languages to confirm the current findings. Furthermore, we need to examine the default phrasing of various structures in reading aloud to see how the phrasing is influenced by phonological and syntactic constraints. Finally, we need to investigate factors affecting the default phrasing, ways to access implicit prosody, and the attachment difference between on-line and off-line processing experiments.

REFERENCES

Allbritten, D., McKoon, G., & Ratcliff, R. (1996). Reliability of prosodic cues for resolving syntactic ambiguity. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *22*, 714–735.

Augurzky, P. (2006). *Attaching relative clauses in German: The role of implicit and explicit prosody in sentence processing*. MPI Series in Human Cognitive and Brain Sciences 77. Leipzig, Germany.

Ayers, G. (1994). Discourse functions of pitch range in spontaneous and read speech. *OSU Working Papers in Linguistics*, *44*, 1–49.

Baayen, R. H. (2008). *Analyzing linguistic data: A practical introduction to statistics using R*. Cambridge, UK: Cambridge University Press.

Bates, D., & Maechler, M. (2009). lme4: Linear mixed effects models using S4 classes. R package version 0.999375-32. Retrieved from http://CRAN.R-project.org/package=lme4

Beckmann, M. E., & Ayers-Elam, G. (1997). *Guidelines for ToBI Labeling*. Unpublished ms. Version 3. Ohio State University.

Beckman, M. E., & Hirschberg, J. (1994). The ToBI annotation conventions. Unpublished ms. Ohio State University, Columbus, OH.

Bergmann, A., Armstrong, M., & Maday, K. (2008). Relative clause attachment in English and Spanish: A production study. *Proceedings of Speech Prosody 2008*, Campinas, Brazil.

Bergmann, A., & Ito, K. (2007). *Attachment of ambiguous RCs: A production study*. Talk given at the 13th annual conference on architectures and mechanisms for language processing (AMLaP), Turku, Finland, Aug, 24–27.

Bookheimer, S. Y., Zeffiro, T. A., Blaxton, T., Gaillard, W., & Theodore, W. (1995). Regional cerebral blood flow during object naming and word reading. *Human Brain Mapping*, *3*, 93–106.

Brysbaert, M., & Mitchell, D. C. (1996). Modifier attachment in sentence parsing: Evidence from Dutch. *Quarterly Journal of Experimental Psychology*, *49A*(3), 664–695.

Brysbaert, M., & Mitchell, D. C. (2000). The failure to use gender information in parsing: A comment on van Berkum, Brown, and Hagoort (1999). *Journal of Psycholinguistic Research*, *29*, 453–466.

Carlson, K., Clifton, C. Jr., & Frazier, L. (2001). Prosodic boundaries in adjunct attachment. *Journal of Memory and Language*, *45*, 58–81.

Carreiras, M., & Clifton, C. Jr. (1993). Relative clause interpretation preferences in Spanish and English. *Language and Speech*, *36*, 353–372.

Carreiras, M., & Clifton, C. Jr. (1999). Another word on parsing relative clauses: Eyetracking evidence from Spanish and English. *Memory and Cognition*, *27*(5), 826–833.

Caspers, J., & van Heuven, V. J. (1993). Effects of time pressure on the phonetic realization of the Dutch accent-lending pitch rise and fall. *Phonetica*, *50*, 161–171.

Clifton, C. J., Carlson, K., & Frazier, L. (2002). Informative prosodic boundaries. *Language and Speech*, *45*, 87–114.

Cuetos, F., & Mitchell, D. C. (1988). Cross-linguistic differences in parsing: Restrictions on the use of the late closure strategy in Spanish. *Cognition*, *30*, 73–105.

De Vincenzi, M., & Job, R. (1995). An investigation of late closure: The role of syntax, thematic structure, and pragmatics in initial and final interpretation. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *21*(5), 1303–1321.

Deevy, P. L. (1999). The Comprehension of English subject-verb agreement. Appendix L: Word by word self-paced reading study materials: Agreement in subject relative clauses. Doctoral dissertation, University of Massachusetts.

Deevy, P. L. (2000). Agreement checking in comprehension: Evidence from relative clauses. *Journal of Psycholinguistic Research*, *29*(1), 69–79.

Ehrlich, K., Fernández, E. M., 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, New York, March 18–20.

Fernández, E. M. (2003). *Bilingual sentence processing: Relative clause attachment in English and Spanish*. Amsterdam, The Netherlands: John Benjamins.

Fernández, E. M., & Bradley, D. (1999). *Length effects in the attachment of relative clauses in English*. Poster presented at the 12th annual CUNY conference on human sentence processing, New York.

Fernández, E. M., Bradley, D., Igoa, J. M., & Teira, C. (2003). *Prosodic phrasing in the RC-attachment ambiguity: Effects of language, RC-length, and position*. Paper presented at the 9th AMLAP (http://amlap.psy.gla.ac.uk/programme/node29.html), Glasgow, August 25–27.

Ferreira, F., & Henderson, J. (1990). The use of verb information in syntactic parsing: A comparison of evidence from eye-movements and segment-by-segment self-paced reading. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *16*, 555–568.

Fodor, J. D. (1998). Learning to parse. *Journal of Psycholinguistic Research*, *27*(2), 285–319.

Fodor, J. D. (2002). Prosodic disambiguation in silent reading. *NELS*, *32*, 113–32.

Fougeron, C., & Jun, S.-A. (1998). Rate effects on French intonation: Prosodic organization and phonetic realization. *Journal of Phonetics*, *26*, 45–69.

Frazier, L. (1979). *On comprehending sentences: Syntactic parsing strategies*. Doctoral dissertation, University of Connecticut.

Frazier, L. (1990). Parsing modifiers: Special purpose routines in the human sentence processing mechanism? In D. A. Balota, G. G. Flores d'Arcais, & K. Rayner (Eds.), *Comprehension processes in reading* (pp. 303–330). Hillsdale, NJ: Lawrence Erlbaum.

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

Frazier, L., & Clifton, C. Jr. (1997). Construal: Overview, motivation, and some new evidence. *Journal of Psycholinguistic Research*, *26*, 277–295.

Frazier, L., Clifton, C. Jr., & Carlson, K. (2004). Don't break, or do: prosodic boundary preferences. *Lingua*, *114*(1), 3–27.

Frazier, L., & Fodor, J. D. (1978). The sausage machine: A new two-stage parsing model. *Cognition*, *6*, 291–325.

Frazier, L., & Rayner, K. (1988). Parameterizing the language system: Left- vs. right- branching within and across languages. In J. A. Hawkins (Ed.), *Explaining language universals*. Oxford, UK: Basil Blackwell.

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

Gilboy, E., & Sopena, J.-M. (1996). Segmentation effects in the processing of complex noun pronouns with relative clauses. In M. Carreiras, J. E. Garcia-Albea, & N. Sebastian (Eds.), *Language processing in Spanish* (pp. 191–206). Hillsdale, NJ: Erlbaum.

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

Hemforth, B., Konieczny, L., Scheepers, C., & Strube, G. (1998). Syntactic ambiguity resolution in German. In D. Hillert (Ed.), *Syntax and semantics: A cross-linguistic perspective* (pp. 293–312). San Diego, CA: Academic Press.

Hirose, Y. (1999). *Resolving reanalysis ambiguity in Japanese relative clauses*. Unpublished doctoral dissertation, CUNY Graduate Center, New York.

Hirschberg, J., & Avesani, C. (2000). Prosodic disambiguation in English and Italian. In A. Botinis (Ed.), *Intonation: Analysis, modelling, and technology* (pp. 87–96). Dordrecht: Kluwer Academic.

Howell, P., & Kadi-Hanifi, K. (1991). Camparison of prosodic properties between read and spontaneous speech material. *Speech Communication*, *10*(2), 163–169.

Hudson, R. F., Lane, H. B., & Pullen, P. C. (2005). Reading fluency assessment and instruction: What, why, and how? *The Reading Teacher*, *58*(7), 702–714.

Jun, S.-A. (2003a). Prosodic phrasing and attachment preferences. *Journal of Psycholinguistic Research*, *32*(2), 219–249.

Jun, S.-A. (2003b). The effect of phrase length and speech rate on prosodic phrasing. *Proceedings of the international congress of phonetic sciences,* Barcelona, Spain: 483–486.

Jun, S.-A. (2006). Intonational Phonology of Seoul Korean Revisited. *Japanese-Korean Linguistics 14* Stanford: CSLI. [Also printed in *UCLA Working Papers in Phonetics* 104: 14–25]

Jun, S.-A. (2007). The intermediate phrase in Korean intonation: Evidence from sentence processing. In C. Gussenhoven & T. Riad (Eds.), *Tones and tunes: Studies in word and sentence prosody*. Berlin: Mouton de Gruyter.

Jun, S.-A., & Kim, S. (2004). Default phrasing and attachment preferences in Korean. *Proceedings of INTERSPEECH-ICSLP*, Jeju, Korea.

Jun, S.-A., & Koike, C. (2003). *Default prosody and RC attachment in Japanese*. Talk given at the 13th Japanese-Korean Linguistics Conference, Tucson, AZ. [Published in *Japanese-Korean Linguistics* 3, 41–53, CSLI, Stanford, in 2008]

Just, M. A., Carpenter, P. A., & Woolley, J. D. (1982). Paradigms and processes in reading comprehension. *Journal of Experimental Psychology: General*, *111*, 228–238.

Kamide, Y., & Mitchell, D. C. (1997). Relative clause attachment: Non-determinism in Japanese Parsing. *Journal of Psycholinguistic Research*, *26*, 247–254.

Kamide, Y., Mitchell, D. C., Fodor, J. D., & Inoue, A. (1998). *Relative clause attachment ambiguity: Further evidence from Japanese*. Poster presented at the 11th Annual CUNY Conference on Human Sentence Processing, Rutgers.

Kimball, J. (1973). Seven principles of surface structure parsing in natural language. *Cognition*, *2*, 15–47.

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.

Krivokapiç, J. (2007). Prosodic planning: Effects of phrasal length and complexity on pause duration. *Journal of Phonetics*, *35*, 162–179.

Ladd, D. R., Faulkner, D., Faulkner, H., & Schepman, A. (1999). Constant "segmental anchoring" of F0 movements under changes in speech rate. *Journal of the Acoustical Society of America*, *106*, 1543–1554.

Lovric, N., Bradley, D., & Fodor, J. D. (2000). *RC attachment in Croatian with and without preposition*. Poster presented at the AMLaP Conference, Leiden.

Lovric, N., Bradley, D., & Fodor, J. D. (2001). *Silent prosody resolves syntactic ambiguities: Evidence from Croatian*. Presented at the SUNY/CUNY/NYU Conference, Stonybrook, NY.

Maynell, L. A. (1999). *Effect of pitch accent placement on resolving relative clause ambiguity in English*. Poster at the 12th CUNY Conference, New York.

Mitchell, D. C. (1994). Sentence parsing. In M. Gernsbacher (Ed.), *Handbook of Psycholinguistics* (pp. 375–409). New York: Academic Press.

Mitchell, D. C., Corley, M. M. B., & Garnham, A. (1992). Effects of context in human sentence parsing: Evidence against a discourse-based proposal mechanism. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *18*(1), 69–88.

Miyamoto, E. T., Nakamura, M., & Takahashi, S. (2004). Processing relative clauses in Japanese with two attachment sites. In K. Moulton & M. Wolf (Eds.), *Proceedings of North Eastern Linguistics Society 34* (pp. 441–452). Amherst: GLSA.

Papadopoulou, D., & Calhsen, H. (2003). Parsing strategies in L1 and L2 sentence processing: A study of relative clause attachment in Greek. *Studies in Second Language Acquisition*, *25*, 501–528.

Price, P., Ostendorf, M., Shattuck-Hufnagel, S., & Fong, C. (1991). The use of prosody in syntactic disambiguation. *Journal of the Acoustical Society of America*, *90*, 2956–2970.

Pynte, J., & Colonna, S. (2000). Decoupling syntactic parsing from visual inspection: The case of relative clause attachment in French. In A. Kennedy, R. Radach, D. Heller, & J. Pynte (Eds.), *Reading as a perceptual process* (pp. 529–547). Oxford, UK: Elsevier.

Pynte, J., Colonna, S., & Gola, C. (2003). *Can an "unnatural" prosody help parsing?* Poster presented at the 9th AMLaP, Glasgow, August 25–27.

Quinn, D., Abdelghany, H., & Fodor, J. D. (2000). *More evidence of implicit prosody in silent reading: French, English and Arabic relative clauses.* Poster presented at 13th annual CUNY conference on human sentence processing.

R Development Core Team (2009). *R: A language and environment for statistical computing.* R Foundation for Statistical Computing, Vienna. ISBN 3-900051-07-0. Retrieved from http://www.R-project.org

Rayner, K., Carlson, M., & Frazier, L. (1983). The interaction of syntax and semantics during sentence processing: Eye movements in the analysis of semantically biased sentences. *Journal of Verbal Learning and Verbal Behavior*, *22*, 358–374.

Rayner, K., & Liversedge, S. P. (2004). Visual and linguistic processing during eye fixations in reading. In F. Ferrreira & J. Henderson (Eds.), *The interface of language, vision and action: Eye movements and the visual world* (pp. 59–104). Hove, UK: Psychology Press.

Rasinski, T. (2006). Issues and trends in literacy: Reading fluency instruction: Moving beyond accuracy, automaticity and prosody. *The Reading Teacher*, *59*(7), 704–706.

Rietveld, A. C. M., & Gussenhoven, C. C. (1987). Perceived speech rate and intonation. *Journal of Phonetics*, *15*, 273–285.

Schafer, A., Speer, S., Warren, P., & White, S. (2000). Intonational disambiguation in sentence production and comprehension. *Journal of Psycholinguistic Research*, *29*, 169–182.

Schafer, A., Speer, S., & Warren, P. (2005). Prosodic influences on the production and comprehension of syntactic ambiguity in a game-based conversation task. In M. K. Tanenhaus & J. C. Trueswell (Eds.), *Approaches to studying world situated language use* (pp. 209–225). Cambridge: MIT Press.

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

Selkirk, E. (2000). The interaction of constraints on prosodic phrasing. In M. Horne (Ed.), *Prosody: Theory and experiment (studies presented to Gösta Bruce)* (pp. 231–262). Dordrecht: Kluwer Academic.

Silverman, K., Beckman, M., Pitrelli, J., Ostendorf, M., Wightman, C., Price, P., Pierrehumbert, J., & Hirschberg, J. (1992). ToBI: A standard for labeling English prosody. *Proceedings of international conference on spoken language processing*, *2*, 867–870.

Silverman, K., Blaauw, E., Spitz, J., & Pitrelli, J. (1992). A prosodic comparison of spontaneous speech and read speech. *Proceedings of international conference on spoken language processing*, *2*, 1299–1302.

Snedeker, J., & Casserly, E. (2010). Is it all relative? Effects of prosodic boundaries on the comprehension and production of attachment ambiguities. *Language and Cognitive Processes*, *25*, 1234–1264.

Snedeker, J., & Trueswell, J. C. (2003). Using prosody to avoid ambiguity: Effects of speaker awareness and referential context. *Journal of Memory and Language*, *48*, 103–130.

Steinhauer, K., Alter, K., & Friederici, A. D. (1999). Brain potentials indicate the immediate use of prosodic cues in natural speech processing. *Nature Neuroscience*, *2*, 191–196.

Steinhauer, K., & Freiderici, A. D. (2001). Prosodic boundaries, comma rules, and brain responses: The closure positive shift in ERPs a universal marker for prosodic phrasing in listeners and readers. *Journal of Psycholinguistic Research*, *30*, 267–295.

Swets, B., Desmet, T., Hambrick, D. Z., & Ferreira, F. (2007). The role of working memory in syntactic ambiguity resolution: A psychometric approach. *Journal of Experimental Psychology: General*, *136*, 64–81.

Taylor, S. E. (2006). *Fluency in silent reading*. Huntington Station, NY: Taylor Associates/ Communications, Inc. Retrieved from www.readingplus.com

Traxler, M. J. (2007). Working memory contributions to relative clause attachment processing: A hierarchical linear modeling analysis. *Memory and Cognition*, *35*, 1107–1121.

Traxler, M. J. (2009). A hierarchical linear modeling analysis of working memory and implicit prosody in the resolution of adjunct attachment ambiguity. *Journal of Psycholinguistic Research*, *38*, 491–509.

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

Vasishth, S., Agnihotri, R. K., Fernández, E. M., & Bhatt, R. (2004). Noun modification preferences in Hindi. In *The Proceedings of Seminar on Construction of Knowledge*, Vidya Bhawan Education Resource Centre, Udaipur, India, pp. 160–171.

Watson, D., & Gibson, E. (2004). The relationship between intonational phrasing and syntactic structure in language production. *Language and Cognitive Processes*, *19*, 713–755.

Wijnen, F. (2004). The implicit prosody of jabberwocky and the relative clause attachment riddle. In H. Quené & V. van Heuven (Eds.), *On speech and language. Studies for Sieb G. Nooteboom* (pp. 169–178). Utrecht: Landelijke Onderzoeksschool Taalwetenschap. (LOT Occasional Series 2).

Zagar, D., Pynte, J., & Rativeau, S. (1997). Evidence for early-closure attachment on first-pass reading in French. *Quarterly Journal of Experimental Psychology*, *50A*(2), 421–438.

Zu, Y., Li, A., & Li, Y. (2006). Speech rate effects on production features. *Report of Phonetic Research*: 141–144. Beijing, P. R. China: Phonetics Laboratory, Institute of Linguistics, Chinese Academy of Social Sciences.

Appendix 1

Code: F03: Fernandez (2003), CC93: Carreiras and Clifton (1993), Fr90: Frazier (1990), D99: Deevy (1999), D00: Deevy (2000); #a: Neutral, #b: NP1-biased, #c: NP2-biased; 'mod': modified; G: gender biased

Group I

9 Short RC

1a. My friend met the aide of the detective that was fired. (F03)
2b. Rob met the interpreter of the ambassadors that was skinny. (F03mod)
3c. Roxanne read the reviews of the poem that was unfinished. (F03)
4a. Patricia saw the teachers of the students that were in class. (F03)
5b. Lisa couldn't find the refills of the pen that were on sale. (F03)
6c. The doctor called in the son of the pretty nurse who hurt herself (Fr90), G
7a. Jane had inspected the printers of the computers that were stolen (F03)
8b. Peter met the boyfriend of the hostess who was a boxer (CC93 modified), G
9c. John spoke to the secretaries of the lawyer who was out of town. (F03mod)

9 Med RC

1b. The journalist interviewed the daughters of the hostage that were about to exit the airplane. (F03mod)
2c. Someone shot the servants of the actress who was on the balcony. (Fr90)
3a. The hotel director didn't want to see the guides of the tourists that were waiting at the reception desk. (F03)
4b. The police arrested the sister of the handyman who recently gave birth to twins. (CC93), G
5c. The children followed the grandfather of the girl who was wearing a long black skirt. (CC93), G
6a. Jane wrote a story about the uncle of the milkman who was known as a real gentleman. (CC93mod)
7b. Everybody ignored the stepfather of the nun who had a long silky white beard. (CC93 modified), G
8c. An armed robber shot the guard of the actress who had just divorced her husband. (CC93), G
9a. The chef couldn't find the cover of the pan that was in the cupboard on the left. (F03)

9 Long RC

1a. John was excited to meet the niece of the actor who was recently starring in a very successful play. (D00)
2b. Officer Phillips questioned the friend of the housekeepers who was quickly cleaning the motel room where the murder had taken place. (D99)
3c. Dr. Lee conferred with the daughters of the crash victim who was apparently suffering from severe shock. (D99)
4a. The curious reporter talked to the son of the delegate that was watching television in the lobby for the whole afternoon. (F03)
5b. The tall receptionist greeted the clients of the lawyer that were waiting in the conference room in the north hall. (F03)
6c. Kelly argued with the cousins of the farmer who was thoughtlessly fertilizing the fields on windy days. (D99)
7a. Yesterday afternoon I saw the daughter of the madwoman who was trying out some new eye shadow and purple lipstick. (CC93)
8b. Nobody noticed the bodyguards of the ambassador that were talking to the photographer at the party. (F03)
9c. The student read the revisions of the manuscript that was on the list of readings required for the final exam. (F03)

Group II

9 Short RC

1b. My friend met the aide of the detectives that was fired. (F03)
2c. Rob met the interpreters of the ambassador that was skinny. (F03)
3a. Roxanne read the review of the poem that was unfinished. (F03)
5c. Lisa couldn't find the refill of the pens that were on sale. (F03)
4b. Patricia saw the teacher of the students that was in class. (F03)
6a. The doctor called in the sons of the pretty nurses who hurt themselves. (Fr90)
7b. Jane had inspected the printer of the computers that was stolen. (F03)
8c. Peter met the girlfriend of the host who was a boxer. (CC93 modified), G
9a. John called the secretary of the lawyer who was out of town. (F03)

9 Med RC

1c. The journalist interviewed the daughter of the hostages that were about to exit the airplane. (F03 modified)
2a. Someone shot the servant of the actress who was on the balcony. (Fr90)
3b. The hotel director didn't want to see the guide of the tourists that was waiting at the reception desk. (F03)
4c. The police arrested the brother of the nursemaid who recently gave birth to twins. (CC93), G
5a. The children followed the mother of the girl who was wearing a long black skirt. (CC93)
6b. Jane wrote a story about the uncle of the laundress who was known as a real gentleman. (CC93), G
7c. Everybody ignored the stepmother of the monk who had a long white silky beard. (CC93), G
8a. An armed robber shot the maid of the actress who had just divorced her husband. (CC93)
9b. The chef couldn't find the cover of the pans that was in the cupboard on the left. (F03)

9 Long RC

1b. John was excited to meet the niece of the actors who was recently married to a Hollywood star. (D00)
2c. Officer Phillips questioned the cousins of the housekeeper who was quickly cleaning the motel room where the murder had taken place. (D99)
3a. Dr. Lee conferred with the daughter of the crash victim who was apparently suffering from severe shock. (D99)
4b. The curious reporter talked to the sons of the delegate that were watching television in the den for the whole afternoon. (F03)
5c. The tall receptionist greeted the client of the lawyers that were waiting in the conference room in the north hall. (F03)
6a. Kelly MacDonough argued with the wife of the farmer who was thoughtlessly fertilizing the fields on windy days. (D99)
7b. Yesterday afternoon I saw the daughter of the madman who was trying out some new eye shadow and purple lipstick. (CC93), G
8c. Nobody noticed the bodyguards of the ambassador that was talking to the photographer at the party. (F03)
9a. The student read the revisions of the manuscripts that were on the list of readings required for the final exam. (F03)

Group III

9 Short RC

1c. My friend met the aides of the detective that was fired. (F03)
4c. Patricia saw the teachers of the student that was in class. (F03)
7c. Jane had inspected the printers of the computer that was stolen. (F03)
2a. Rob met the interpreter of the ambassador that was skinny. (F03)
5a. Lisa couldn't find the refills of the pens that were on sale. (F03)
8a. Peter met the boyfriend of the host who was a boxer. (CC93 modified)
3b. Roxanne read the review of the poems that was unfinished. (F03)
6b. The doctor called in the son of the pretty nurse who hurt himself. (Fr90), G
9b. John called the secretary of the lawyers who was out of town. (F03)

9 Med RC

1a. The journalist interviewed the daughters of the hostages that were about to exit the airplane. (F03)
2b. Someone shot the servant of the actresses who was on the balcony. (Fr90)
3c. The hotel director didn't want to see the guide of the tourists that were waiting at the reception desk. (F03)
4a. The police arrested the sister of the nursemaid who recently gave birth to twins. (CC93)
5b. The children followed the mother of the boy who was wearing a long black skirt. (CC93), G
6c. Jane wrote a story about the aunt of the milkman who was known as a real gentleman. (CC93), G
7a. Everybody ignored the stepfather of the monk who had a long silky white beard. (CC93)
8b. An armed robber shot the maid of the actor who had just divorced her husband. (CC93), G
9c. The chef couldn't find the covers of the pan that was in the cupboard on the left. (F03)

9 Long RC

1c. John was excited to meet the nieces of the actor who was recently starring in a very successful play. (D00)
2a. Officer Phillips questioned the cousin of the housekeeper who was quickly cleaning the motel room where the murder had taken place. (D99)
3b. Dr. Lee conferred with the daughter of the crash victims who was apparently suffering from severe shock. (D99)
4c. The curious reporter talked to the son of the delegates that were watching television in the den for the whole afternoon. (F03)
5a. The tall receptionist greeted the clients of the lawyers that were waiting in the conference room in the north hall. (F03)
6b. Kelly MacDonough argued with the cousin of the farmers who was thoughtlessly fertilizing the fields on windy days. (D99)
7c. Yesterday afternoon I saw the son of the madwoman who was trying out some new eye shadow and purple lipstick. (CC93), G
8a. Nobody noticed the bodyguards of the ambassador that was talking to the photographer at the party. (F03)
9b. The student read the revisions of the manuscript that were on the list of readings required for the final exam. (F03)