Prosody in Spoken Language Processing

Amy J. Schafer and Sun-Ah Jun
Department of Linguistics, UCLA
schafer@ucla.edu, jun@humnet.ucla.edu
Phone: +1-310-206-2808, Fax: +1-310-206-5743

Abstract

Studies of prosody and sentence processing have demonstrated that prosodic phrasing can exhibit strong effects on processing decisions in English. In this paper, we tested Korean sentence fragments containing syntactically ambiguous Adj-N1-N2 strings in a cross-modal naming task. Four accentual phrasing patterns were tested: (a) the default phrasing pattern, in which each word forms an accentual phrase; (b) a phrasing biased toward N1 modification; (c) a phrasing biased toward complex-NP modification; and (d) a phrasing used with adjective focus. Patterns (b) and (c) are disambiguating phrasings; the other two are commonly found with both interpretations and are thus ambiguous. The results showed that the naming time of items produced in the prosody contradicting the semantic grouping is significantly longer than that produced in either default or supporting prosody. We claim that, as in English, prosodic information in Korean is parsed into a well-formed prosodic representation during the early stages of processing. The partially constructed prosodic representation produces incremental effects on syntactic and semantic processing decisions and is retained in memory to influence reanalysis decisions.

1. Introduction

Several studies of prosody and sentence processing have now demonstrated that prosodic phrasing can exhibit strong effects on processing decisions in English (e.g., Kjelgaard & Speer, 1999; Schafer, 1997; Warren, Grabe, & Nolan, 1995). However, very little work on prosodic effects on parsing has been done with other languages. Given that there is variability in prosodic structure and its relationship to syntactic or semantic form across languages (e.g., Beckman & Pierrehumbert, 1986; Venditti, Jun & Beckman, 1996), it could be the case that prosody has very different effects on processing across languages. Thus, just as syntactic processing models can be better evaluated by testing syntactically diverse languages, experimentation on prosodically diverse languages should further our understanding of both the universal and the language-specific relationships between prosodic form and sentence processing decisions.

The goal of this paper is to determine if accentual phrasing in Korean could affect sentence comprehension. Korean accentual phrases are similar to intermediate phrases in English in that they are the level of prosodic structure which is intermediate between phonological words and intonation phrases. A small set of studies on English have provided evidence that intermediate phrase boundaries can affect comprehension (Kjelgaard, 1995; Schafer, 1997; Kjelgaard & Speer, 1999). However, Korean accentual phrases are arguably less phonetically salient than English intermediate phrases, as described further in Section 2. They also tend to differ in span. Korean accentual phrases generally contain fewer syllables and fewer content words than English intermediate phrases. In an analysis of the prosody produced by reading a standard passage, Jun and Fougeron (Jun 1999; Jun & Fougeron, to appear) found that Korean accentual phrases contained an average of 3.2 syllables and 1.2 content words, and thus generally contained a single phonological word. In contrast, with an English version of the passage, Ueyama (1998) found that English intermediate phrases contained an average of 5.3 syllables. Using a different reading text, Ayers (1994) found that English intermediate phrases contained over 3.9 content words on average. Thus, accentual phrases may show smaller effects on comprehension than intermediate phrases for either of two reasons. First, if they are less acoustically salient, they might be less reliably detected during sentence comprehension. Second, if syntactic disambiguation depends on grouping syntactic units into prosodic units, the small size of Korean accentual phrases may mean that they are less commonly associated with disambiguated syntactic structures than English intermediate phrases. Thus, Korean accentual phrases may be a less reliable source of information about syntactic structure than English intermediate phrases.

Nevertheless, Jun (1993, 1994) showed that correspondences exist between accentual phrasing and syntactic phrasing in Korean. She found that speakers produced different accentual phrasings for the Korean equivalent of *black cat's ankle* depending on whether *black* modified *ankle* or *cat*. Speakers grouped *black* with *cat's* to show that *black* modified *cat*, and grouped *cat's* and *ankle* to indicate the other interpretation. We therefore expected there to be some effect of accentual phrasing on sentence comprehension.

Prosodic phrases can group together material which is closely related syntactically or semantically. This "chunking" aspect of prosodic phrasing has been captured explicitly in some models of sentence processing to explain prosodic disambiguation effects. For example, on the basis of observations about English and Japanese, Marcus & Hindle (1990) proposed that material in separate prosodic phrases can be freely combined at later stages of processing, but material within a single prosodic phrase cannot be separated. Such a view of prosodic phrasing effects would most likely predict very similar effects of accentual phrasing in Korean and intermediate phrasing in English.

Here, we briefly report on a comprehension experiment on prosody and sentence processing in Korean, which are part of an on-going research project on prosodic effects on sentence processing in Korean (see Schafer & Jun, to appear, for more detail). This experiment tested a syntactically ambiguous structure under four conditions of accentual phrasing. On the basis of the results, we will argue that sentence comprehension in Korean, like in English, shows effects of acoustically subtle prosodic phrase boundaries. In the next section, we review the intonational structure of Seoul Korean. We then turn to the description of the experiments.

2. Description of Seoul Korean Intonation

Korean has two prosodic units which are marked by intonation. These are called the "intonation phrase" and the "accentual phrase" by Jun (1993, 1998), adopting the intonation framework developed by Pierrehumbert and her colleagues (Pierrehumbert 1980; Beckman & Pierrehumbert 1986; Pierrehumbert & Beckman 1988; see Ladd (1996) for extensive review). A schematic representation of the intonational structure of Seoul Korean is shown in Figure 1. The intonation phrase (IPh) consists of one or more accentual phrases, and is demarcated by a boundary tone. An IPh boundary tone can be a simple high or low tone (H%, L%) or a combination of high and low tones (e.g. HL%, LHL%, LH%). It is realized on the phrase-final syllable and delivers pragmatic information about the sentence. An IPh-final syllable is subject to final lengthening—such syllables are about 1.8 times longer than IPh-medial syllables (Kim et al., 1997; Korea Telecom, 1996)—and is optionally followed by a pause.

An accentual phrase (APh) in Korean can consist of one or more phonological words (a lexical item followed by case markers or postpositions; written as 'w' in Figure 1), but generally contains only one word. It is defined by phrasal tones demarcating the beginning and the end of the phrase. The underlying tonal pattern of an accentual phrase in Seoul Korean is either Low-High-Low-High (LHLH) or High-High-Low-High (HHLH): thus, THLH in Figure 1. The APh-initial tone is determined by the laryngeal feature of the phrase-initial segment. When the phrase-initial segment is either aspirated or tense, having [+stiff vocal cords] (Halle & Stevens, 1971), an APh begins with an H tone; otherwise it carries an L tone. (For details about this tonal difference, see Jun 1996, 1999a.). Unlike the case for English prosodic phrase boundaries, Korean accentual phrase boundaries are not reliably associated with longer durations.

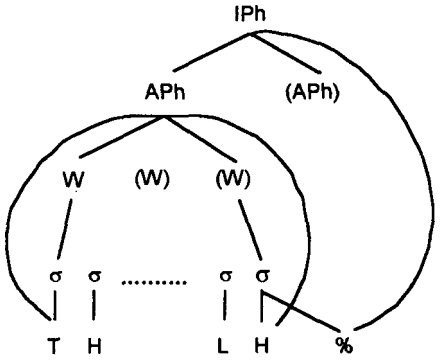


Figure 1. Intonational Structure of Seoul Korean (Jun, 1996). Elements in parentheses are optional. Accentual phrase tones are realized on the first, second, penultimate, and ultimate syllable of the accentual phrase, regardless of how those syllables are parsed into phonological words. IPh: intonation phrase; APh: accentual phrase; w: phonological word; σ: syllable; T: tone, realized as H when the initial segment is aspirated/tense, L otherwise; H: high tone; L:low tone; %: intonation phrase boundary tone.

3. Experiment

3.1 Data and Method

An ambiguous noun phrase, in which a single adjective is followed by a complex noun phrase, is employed in the comprehension experiment. An example is shown in (1). The adjectives in such phrases can readily be interpreted as modifying either the entire complex NP (i.e., the second noun) or only the first noun. This ambiguity is not resolved by syntactic information, although it can be influenced by pragmatic information and prosodic information.

(1) hyunmyunghan agie appa
wise baby-GEN daddy => 'the [wise baby]'s daddy' *or* 'the wise [baby's daddy]'

Phrases like the one in (1) can be grammatically produced with multiple prosodic patterns, as described in (2). The first pattern, in (2a), is the default prosodic pattern. As noted above, this pattern appears to be quite frequent in Korean, particularly in wide focus ('out of the blue') situations. It commonly occurs with both the interpretation (the adjective modifying only the first noun or the entire complex NP). Patterns (2b) and (2c) are used for purposeful disambiguations toward N1 modification and complex-NP modification (i.e., N2), respectively; they reflect the 'semantic closeness' factor. Pattern (2d) occurs when contrastive focus is placed on the adjective, for either interpretation, following Korean's general pattern of dephrasing following a focused constituent that was described above. Henceforth, we will refer to the pattern in (2a) as the default prosody, to the patterns in (2b) and (2c) as N1-modification prosody and N2-modification prosody, respectively, and to the pattern in (2d) as Adj-focus prosody.

(2) a. (hyunmyunghan]$_{APh}$ (akie)$_{APh}$ (appa)$_{APh}$: Default Prosody (Adj)$_{APh}$ (N1)$_{APh}$ (N2)$_{APh}$
b. (hyunmyunghan akie)$_{APh}$ (appa)$_{APh}$: N1-modification Prosody (Adj N1)$_{APh}$ (N2)$_{APh}$
c. (hyunmyunghan)$_{APh}$ (akie appa)$_{APh}$: N2-modification Prosody (Adj)$_{APh}$ (N1 N2)$_{APh}$
d. (hyunmyunghan akie appa)$_{APh}$: Adj-focus Prosody (Adj N1 N2)$_{APh}$

We expected that constructions such as (2) would show an initial syntactic preference for the structure associated with N1 modification. This expectation is consistent with the classic "Garden Path" depth-first model of sentence comprehension, in which a single structure is built as quickly as possible in all cases of structural ambiguity (e.g., Frazier, 1987), with minimal revisions when necessary (Frazier, 1990; Frazier & Clifton, 1998). It also fits the predictions of Constraint-Based models (e.g., MacDonald, Pearlmutter & Seidenberg, 1994), under the assumption that N1 modification is more frequent in such constructions than N2 modification. We believe this is a reasonable assumption, given that speakers have another, unambiguous option for producing

sentences with N2 modification: They can re-order the items as N1-Adj-N2. Further, previous psycholinguistic work on Japanese has suggested an initial N1-modification preference for similar structures (Inoue & Fodor, 1995; Kamide & Mitchell, 1997).

We constructed two sets of materials, a Main set and a Control set, which were always compatible, initially, with this syntactic preference for the N1-modification structure. Both sets contained two conditions, an N1-bias condition and an N2-bias condition. In the Main set, modification of N2 by the adjective was implausible in the N1-bias condition (ex. 'wise baby's instrument'). In the N2-bias condition, modification of N2 by the adjective was far more plausible than N1 modification (ex. 'wise baby's daddy'). The adjective and N1 never varied across conditions for the Main set; only N2 changed. In the Control set, N2 modification was possible in the N1-bias condition, although less likely than N1 modification (ex. 'stylish model's designer'). In the N2-bias condition, modification of N2 by the adjective was far more plausible than N1 modification, as in the Main set (ex. 'stylish Pope's designer'). The Control items were created by varying N1 between conditions while keeping the adjective and N2 constant. The number of syllables in N1 was matched, within items, for the two conditions. The initial segment of N2 is also matched between the two conditions in order to minimize segmental differences due to coarticulation when listening to the end of N1 for the Cross-model naming task.

For the Cross-model naming task, subjects heard an auditorily presented sentence fragment, which ended with N1. Immediately following the offset of N1, N2 appeared on a computer monitor in standard Korean orthography. Naming times for N2 were collected using *psyscope*. These times were corrected for lexical differences among the visual targets; see Schafer & Jun (to appear) for details. 12 Control items, in two bias conditions, are all presented with default prosody (see (3)) and 24 Main items, all lexically biased toward N2 modification (confirmed by pretesting), are presented under four prosodic conditions (see (4)). 36 subjects, all native speakers of Seoul Korean, participated.

(3) Example of a Control item

Auditory Fragment:	Visual Target:	Prosody:	Lexical Bias:
a. ...(selyuntwen)$_{APh}$ (motele)$_{AP}$ 'stylish' 'model's'	DESIGNER	Default	N1 bias
b. ...(selyuntwen)$_{APh}$ (kyohwange)$_{APh}$ 'stylish' 'Pope's'	DESIGNER	Default	N2 bias

(4) Example of a Main item

Auditory Fragment:	Visual Target:	Prosody:	Lexical Bias:
a. ...(hyunmyunghan akie)$_{APh}$	DADDY	N1-mod. prosody	N2 Bias
b. ...(hyunmyunghan akie	DADDY	Adj-focus prosody	N2 Bias
c. ...(hyunmyunghan)$_{APh}$ (akie)$_{APh}$	DADDY	Default prosody	N2 Bias
d. ...(hyunmyunghan)$_{APh}$ (akie	DADDY	N2-mod. prosody	N2 Bias

3.2 Results and Discussion

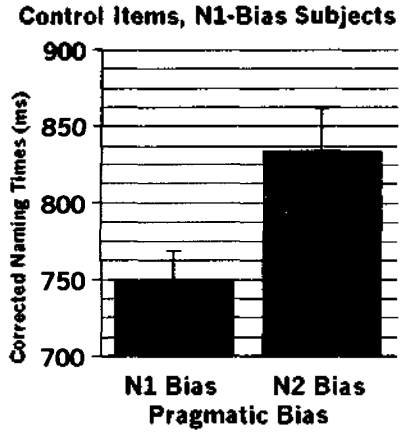
If subjects have an initial, syntactic preference for the structure that supports N1 modification, they would initiate reanalysis when the pragmatic information in N2 is interpreted. (However, reanalysis is not absolutely required: N1 modification is still grammatical, just less plausible than N2 modification.) This prediction of an initial N1 bias can be tested in the naming time of Control items, which are produced in default prosody. Figure 2 shows that the prediction is correct. The naming times are significantly shorter for the N1-bias condition than the N2-bias condition.

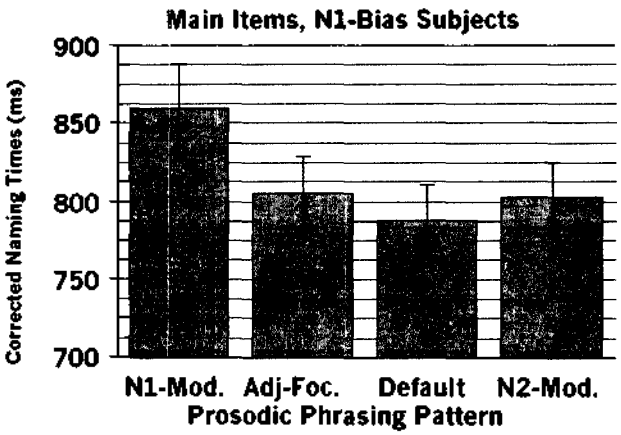
For the naming times of Main items which vary in their prosodic phrasing, we predicted following Schafer (1997) that an APh boundary between the adjective and N1 would impede syntactic construction and semantic interpretation of the N1-modification structure, and that an APh boundary between N1 and N2 would impede syntactic construction and semantic interpretation of the N2-modification structure. Thus, assuming that reanalysis occurs, and a similar percentage of tokens are reanalyzed for each condition, naming times should show the following pattern: N1 modification > Adj-focus > Default > N2 modification.

Because the predictions rest on the assumption that reanalysis takes place, we used the results of the Control items to assess the likelihood that subjects were performing reanalysis for a majority of tokens. The results for the 24 subjects who showed indications of reanalysis are given in figures 2 and 3.

Results of Main items show that the naming time of items with N1-modification phrasing is significantly longer than that of the other conditions, which are not significantly different among themselves. Since the N1-modification phrasing is contradicting the pragmatic grouping, i.e., N2-bias, the slowest naming time of N2 in this condition suggests that prosody modulates processing on-line, showing incremental effects of prosody on interpretation. That is, the existence of an early APh boundary—one occurring prior to the point that N2—can affect interpretation. On the other hand, the naming time of N2-Mod phrasing was not significantly faster than the Default or Adj.-Foc phrasings even though it is cooperating, thus supporting, the pragmatic grouping. This may be due to the lack of phrase boundary cues at the end of N1 due to truncation.

Furthermore, duration measurements of Adjective and N1 show that the naming time difference was not due to the duration difference between the lexical items. The duration of adjective was not significantly different across conditions, and the duration of N1 across conditions was not predictable by the presence or absence of an accentual phrase boundary after N1. In sum, results show that Korean accentual phrasing influences sentence processing even though accentual phrasing is quite subtle acoustically and boundaries are not reliably associated with longer durations.





Figures 2 and 3: Corrected naming times for Control and Main items, 24 subjects who showed evidence of N1 bias for the Control items.

4. Conclusion

The Korean results mirror the effects of intermediate phrasing on interpretation in English (Schafer, 1997). We conclude that in both languages prosodic information is parsed into a well-formed prosodic representation during the early stages of processing. The partially constructed prosodic representation produces incremental effects on syntactic and semantic processing decisions and is retained in memory to influence reanalysis decisions.

However, while the cross-modal naming task is useful in showing on-line processes, it does not directly indicate which interpretation subjects choose for our ambiguous structure. It is possible that subjects differed across conditions in the percentages of sentences which they reanalyzed. That is, they may have done more reanalysis in the N2-modification condition than the Default condition, and so forth. A direct measure of the final interpretation of the ambiguous structure is given in Schafer & Jun (to appear).

Selected References

Beckman, M. and J. Pierrehumbert. 1986. Intonational structure in Japanese and English. *Phonology Yearbook* 3. 255-309.

Frazier, L. 1987. Sentence processing: A tutorial review, in M. Coltheart, ed., *Attention and performance*. 559-586.

Frazier, L. and C. Clifton. 1996. *Construal*. Cambridge, MA: MIT Press.

Frazier, L. and C. Clifton. 1998. Sentence reanalysis, and visibility, in Fodor & Ferreira, eds. *Reanalysis in Sentence Processing*.

Inoue, A. and J.D. Fodor. 1995. Information-paced parsing of Japanese, in Mazuka & Nagai, eds., *Japanese Sentence Processing*. Hillsdale, NJ: Lawrence Erlbaum. 9-64.

Jun, S.-A. 1993. *The Phonetics and Phonology of Korean Prosody*. Doctoral dissertation, The Ohio State University.

Jun, S.-A. 1996. Influence of microprosody on macroprosody: A case of phrase initial strengthening. *UCLA WPP* 92. 97-116.

Jun, S.-A. 1998. The Accentual Phrase in the Korean prosodic hierarchy. *Phonology* 15. 189-226.

Jun, S.-A. 1999. Phonetic and phonological tone tier in Korean ToBI. Workshop on Intonation: Models and ToBI Labeling, a satellite meeting of the Fourteenth meeting of the International Congress of Phonetic Sciences, San Francisco, August 1999.

Jun, S.-A. and C. Fougeron. A phonological model of French intonation, To appear in A. Botinis, ed., *Intonation: Models and Applications*. Kluwer.

Kamide, Y. and D.C. Mitchell. 1997. Relative clause attachment: Nondeterminism in Japanese parsing. *Journal of Psycholinguistic Research* 26. Kjelgaard, M.M. and S.R. Speer. 1999. Prosodic facilitation and interference in the resolution of temporary syntactic closure ambiguity. *Journal of Memory and Language* 40. 153-194.

MacDonald, M.C., N.J. Pearlmutter, and M.S. Seidenberg. 1994. The lexical nature of syntactic ambiguity resolution. *Psychological Review* 101.

Marcus, M. and D. Hindle. 1990. Description theory and intonation boundaries, in G. Altmann, ed., *Cognitive Models of Speech Processing: Computational and Psycholinguistic Perspectives*. Cambridge, MA: MIT Press. 483-531.

Schafer, A.J. 1997. *Prosodic Parsing: The Role of Prosody in Sentence Comprehension*. Ph.D dissertation, U. of Massachusetts.

Schafer, A.J., and S.-A. Jun. To appear. Effects of Accentual Phrasing on Adjective Interpretation in Korean, in M. Nakayama, ed., East Asian Language Processing, Stanford, CSLI.

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

Venditti, J.J., S.-A. Jun, and M.E. Beckman. 1996. Prosodic cues to syntactic and other linguistic structures in Japanese, Korean, and English, in J. Morgan and K. Demuth, eds., *Signal to Syntax: Bootstrapping from Speech to Grammar in Early Acquisition*. Mahwah, NJ: Lawrence Erlbaum. 287-311.

Warren, P., S.R. Speer, A.J. Schafer, S.D. White, and J. Kneale. 1999. Situational constraints on the prosodic resolution of syntactic ambiguity. *Proceedings of the Fourteenth ICPhS*, San Francisco, 1-7 August 1999, 1301-1304.