Prosodic Disambiguation of Syntactic Clause Boundaries in Korean

Soyoung Kang and Shari R. Speer
Ohio State University

1. Introduction

This study explored the effects of prosodic boundaries on the comprehension of temporarily and globally ambiguous sentences in Korean. Previous work on English, Japanese, and Korean has demonstrated that prosodic structures can carry critical information about the meaning and/or syntactic structures of spoken sentences (Kjelgaard and Speer 1999; Jun and Oh 1996; Misono et al. 1997; Venditti 1993; among others). These languages provide an interesting contrast in terms of the prosodic and syntactic information available to the listener during sentence comprehension. In contrast to English, Korean and Japanese are head-final, pro-drop languages so that both temporary and global syntactic ambiguities are frequent. For example, in a head-final language, multiple NPs are encountered before the verbs to which they are attached. In addition, syntactic clausal boundaries are marked by a complementizer at the end of the clause so that clause boundaries cannot be detected until this morpheme is encountered. Thus, some or all the NPs encountered are subject to misanalyses in sentences with more than one clause.

The pro-drop nature of Korean also leads to syntactic ambiguity. Unlike English, in which all the arguments of verbs are usually realized phonologically, pro-drop languages allow deletion of verbal arguments. Hence, a string of words can be interpreted in many different ways depending on where those dropped pro are posited. Prosodic phrasal boundaries can play a fundamental role in restricting the possible number of (preferred) interpretations for syntactically ambiguous strings in Korean. Moreover, prosodic disambiguation can be directly tied to prosodic phrasing, because of the distinctly phrasal nature of Korean intonation. In contrast to English and Japanese, Korean lacks lexically-associated

* Special thanks go to Mary Beckman, Psychoids discussion group, and fellow members of Shari Speer’s sentence processing lab at OSU.

prosodic elements, such as stress pattern and accent, which can interact with phrasal prosodic structure in the resolution of structural ambiguity. Thus, in this paper, we will strongly argue that prosodic structure is one of the most important elements in determining the syntactic structures that are built from given strings of words.

The Korean relative clause constructions examined here are potentially ambiguous from the outset. For example, the two-word sentence-initial fragment in example (1) is ambiguous such that the initial nominative marked NP may be the subject of either the immediately following REL-marked verb, or of the sentence-final main verb.  

(1) Cinwuen-ika yupaytangha-n …  
Cinwuen-NOM exiled-REL …

Depending on the nature of head nouns modified by relative clauses, this fragment can be continued as in example (2), which we will call ‘gap type’ or as in example (3), which will be called ‘pro type.’ At the completion of the sentences, example (2) is not ambiguous anymore because the temporary syntactic ambiguity is resolved at the head noun, where it is apparent that the initial NP cannot be the subject of the REL-marked verb. In contrast, sentence (3) remains ambiguous even after it is complete. One of the critical differences for these two examples is that for (2) and (3a), there is a syntactic clause boundary between the initial NP and the REL-marked verb while there is none for (3b). In terms of meaning, this implies that the initial NP in (2) and (3a) cannot be the subject of REL-marked verb whereas it is for (3b).

(2) Cinwuen-ika e_i ywupaytangha-n siin-ul hyeppakhay-sses.  
Cinwuen-NOM e_i exiled-REL poet_i-ACC threaten-PAST  
‘Cinwuen threatened the poet who got exiled.’

(3) Cinwuen-ika ywupaytangha-n sancang-ul konkyekha-yseses.  
Cinwuen-NOM exiled-REL villa-ACC attack-PAST  
a. ‘Cinwuen attacked the villa where (pro) got exiled.’

b. ‘(pro) attacked the villa where Cinwuen got exiled.’

1. The following convention is used to mark grammatical morphemes: NOM refers to a nominative marker, ACC refers to an accusative marker, REL marks a relative clause morpheme, and finally PAST refers to a past tense morpheme.
For this type of syntactically ambiguous structure, we will report the results of two written and two auditory experiments and show that prosodic boundaries are critically used to recover distinct meanings from syntactically ambiguous strings. In the following section, we briefly discuss the intonation model assumed in this study. This is followed by the discussion of four experiments and our conclusion.

2. Intonation model

We assume the intonation model of Seoul Korean developed by Jun (1993). According to Jun (1993), each utterance in Seoul Korean is made up of one or more Intonational Phrases (IP henceforth), which in turn consists of one or more Accentual Phrases (AP hereafter). Each AP can contain one or more phonological words and typically is characterized by a final rising tonal pattern (marked by ‘Ha’ in labeling). In addition, depending on the property of initial segments in the initial syllable, an AP can begin with a High (for tense and aspirated segments) or Low tone (for all other segments). An IP, the biggest prosodic unit, ends with one of boundary tones (such as H% or HL%), carrying different pragmatic information. The final syllable of an IP is lengthened, and optionally followed by a pause. When an AP coincides with the end of an IP, the AP rising tone is replaced by one of IP boundary tones.

According to Jun (1993), actual prosodic phrasings are not easy to predict because they depend on several different factors including the syntactic and semantic structures of sentences and the speech rate of speakers, among others.

3. Experiments

3.1. Written study 1

A written norming study was conducted first to investigate participants’ preferred interpretation of sentences like those in (2) and (3). 80 target sentences (40 verb items times two alternating head nouns for gap/pro distinction) were mixed with an additional 10 filler sentences of various syntactic types in a pseudo-randomized list. Participants read each sentence from a questionnaire that contained those target and filler items. Each sentence on the questionnaire was followed by a comprehension question about the sentence, such as “Who got exiled?” along with three choices. The three choices for temporarily ambiguous gap type sentences like (2) included the initial NP (NP1 in the graph), the head noun (NP2), and “can’t tell.” For globally ambiguous pro type sentences such as (3), the three choices given were the initial NP (NP1 in the graph), “I” (NP2; the first person pronoun was used to give one example of a referent of pro), and
“can’t tell.” The reason “I” was chosen as a referent of pro was that without any previous context, we thought that it would be hard to posit someone else as a referent of pro. There were 20 native Seoul Korean speakers, who were recruited by Professor Mira Oh at the Yeojoo Institute of Technology in Korea.

The results showed that for gap types, syntactic bias was so strong that almost no other choices were selected other than the head noun. For pro types, the relatively large number of “can’t tell” responses showed that readers were sensitive to the ambiguity of this structure although there was some bias toward the initial NP. From this, we can see that, in the absence of spoken prosody, there is a clear distinction between gap and pro type sentences in Korean such that gap types are unambiguous whereas pro types are truly ambiguous.

Figure 1  Mean percentage of NP choices in Experiment 1.

3.2. Auditory study 1

We hypothesized that the location of an IP boundary could effectively inform listeners about syntactic structure during sentence comprehension. For temporarily ambiguous gap type sentences like (2), when an IP boundary occurs right after the initial NP, the prosodic phrasal boundary will coincide with the syntactic clause boundary, leading to the correct interpretation of the sentence. When there is no IP boundary after the initial NP, the (lack of) mapping between prosodic and syntactic structures may interfere with sentence comprehension. For pro type sentences (globally ambiguous sentences such as (3)), the presence of an IP boundary will bias listeners toward the interpretation that the first NP is the subject of the matrix verb, while the absence of this early IP will bias them toward an interpretation where the initial NP is the subject of the REL-marked verb.

We conducted an auditory perception study using the materials from the written study. Four verb items out of 40 from the written study were excluded, since they had some specific morphemes attached to the main
verbs, which indicated a particular interpretation. Thus, 72 experimental sentences (36 verb items times two head noun types for gap/pro distinction) and 36 filler items were recorded by a trained phonetician, who was a male native speaker of Seoul Korean. Four conditions were created for a factorial design, with two types of syntax (gap vs. pro) and two types of prosodic phrasing (IP vs. noIP boundary after the initial NP).

Thirty-six experimental items were counter-balanced across four lists so that 9 sentences from each condition occurred in each list. These test sentences were mixed with 36 filler sentences (18 sentences with an IP boundary and 18 with no IP boundary). Filler sentences were intended to produce restrictive and nonrestrictive relative clause interpretations. Items were pseudo-randomized so that no three sentences of the same type occurred consecutively. The wave forms and pitch contours of two example sentences (a gap type) are given below. Figure 2 shows an IP boundary tone (HL%) on the last syllable of the initial nominative marked NP, which is followed by a substantial pause. Figure 3 shows the same sentence pronounced with no IP boundary in this location.

Figure 2  Wave form and pitch contour for a gap IP sentence.
Phonetic analyses were conducted to make sure that all test sentences were produced with intended prosody. The final syllables of IP final words are usually longer due to phrase-final lengthening, so the final syllable of the initial NP in the IP condition will be longer than the one in the noIP condition. Rather than measuring the last syllable of the initial NP, we measured the duration from the last syllable of the initial NP (‘ka’) until the end of the first syllable of the REL-marked verb and subtracted the duration of any included pause. This procedure allows us to evaluate final syllable duration without determining the location of a boundary between ‘ka’ and the first syllable of the following word when the latter began with a vowel.

2. We thank Mary Beckman for suggesting this.
We also measured the F0 peak in both conditions. Since the initial NP in the IP condition but not in the noIP condition has HL% boundary, the F0 peak in the former condition will be much higher than the latter. On the other hand, since both conditions have the same AP boundary right after the REL-marked verb, the peak in the two will not differ. The following graph confirmed this prediction.

![Graph](image)

**Figure 5** The plot of F0 peak of ‘ka’ against F0 peak of REL-marked verb.

With these materials, 40 native speakers of Seoul Korean with a normal uncorrected hearing at Ohio State University participated in the experiment for a payment of $7. They completed two tasks. First, right after hearing each sentence over the headphones, they indicated whether they understood the sentence or not. Then, they answered the same comprehension question as in the written study about the sentence, such as “Who got exiled?” This appeared on the computer monitor along with the same three choices as in the written study. For temporarily ambiguous gap type sentences like (2), they included the initial NP (NP$_1$ in the graph), the head noun (NP$_2$), and “can’t tell.” For globally ambiguous pro type sentences such as (3), the three choices were the initial NP (NP$_1$ in the graph), someone other than the initial NP (NP$_2$), and “can’t tell.”

As expected, participants responded “understand” most often in the gapIP condition (95.3% of the time), followed by the proNoIP condition (88.7%), the gapNoIP condition (81.7%) and finally the proIP condition (80.1%), indicating that participants successfully understood the sentences in all four conditions. However, the reaction time data revealed a difference in processing difficulty between conditions. Response time was measured from the onset of the sentence sound file to the participant’s button press. Sound file durations were subtracted from the measured time. Figure 6
shows that responses were slowest in the gapNoIP condition (856.3 ms), suggesting that the mismatch between syntactic and prosodic structures impeded processing. This is followed by proNoIP condition (710.2 ms), proIP (640.5 ms) and finally gapIP condition (512.3 ms).

![Figure 6](image1.png) Mean reaction time for understood responses in Experiment 2.

![Figure 7](image2.png) Mean percentage of NP choices in gap types in Experiment 2.

![Figure 8](image3.png) Mean percentage of NP choices in pro types in Experiment 2.
Figures 7 and 8 show the results of comprehension questions. NP1 refers to the initial NP in both gap and pro type sentences. NP2 is an index to refer to the head nouns in gap types, but someone other than the initial NP in pro types.

Statistical analyses were conducted on arcsine transformed percentage data of comprehension question answers. There was a main effect of syntactic sentence type on both subject and item analyses [F1(1,39)=59.72, p<0.01; F2(1,35)=74.89, p<0.01] with gap sentences showing greater consistency of response than pro sentences. Although there was no main effect of prosody in subject and item analyses [both Fs<1], there was a significant interaction of syntax and prosody [F1(1,39)=8.16, p<0.01; F2(1,35)=39.84, p<0.01] and of syntax, prosody and response type [F1(1,39)=20.04, p<0.01; F2(1,35)=37.59, p<0.01]. We will split the discussion into gap and pro types for the ease of exposition.

Figure 7 shows that for temporarily ambiguous gap type sentences, again, the syntactic bias was very strong. In both prosodic conditions, participants overwhelmingly chose the head noun (NP2) as the answer to the comprehension questions. However, a planned comparison showed a significant difference between the number of NP2 responses in the gapIP (97.8%) and gapNoIP (88.4%) conditions [F1(1, 39)=8.72, p<0.01; F2(1,35)=20.1, p<0.01].

Figure 8 shows the results for pro type sentences, which had a standing syntactic ambiguity. We begin the discussion of proNoIP condition before proIP, a more complicated case. In the proNoIP condition, as predicted, NP1 was chosen significantly more often (82.6%) than NP2, which was chosen only 12.55% of the time [Planned comparison F1(1,39)=213.39, p<0.001; F2(1,35)=366.71, p<0.001]. This result suggests that the absence of an IP boundary led to the interpretation where the initial NP was the subject of the REL-marked verb. For the proIP condition, it had been predicted that with an IP boundary right after the initial NP, listeners would put a syntactic clause boundary right after that NP and so would posit pro as the subject of the REL-marked verbs. As predicted, indeed, NP2 was the one chosen most frequently (43.6%). However, there were a considerable number of NP1 responses (38.8%). The planned comparison of the means did not reach significance [F1<1; F2(1,35)=2.4, p>0.1]. This indicates that while the majority of listeners chose NP2 (here, someone other than NP1) as the subject of the REL-marked verb, many participants assigned the initial NP as the subject. A closer look at the materials revealed that a subset of the sentences such as (5) allowed an unexpected alternative interpretation.
(5) Nayong-ika kongpwuha-nun hakkyo-rul pangmwunha-ysse.
Nayong-NOM study-REL school-ACC visit-PAST
‘Nayong visited the school where pro studied.’

For these sentences, even when the initial NP is posited as a subject of the main verb and pro is posited as the subject of the REL-marked verb, the embedded pro can refer back to the initial NP (i.e., there is a reflexive interpretation of pro). This may account for why there were more responses that chose the initial NP as the subject of REL-marked verbs in the IP boundary condition.

The reaction time data in Figure 9 further support this idea; data shown in this graph are response time to answer the comprehension question. Since some items were so strongly biased, not all of pro items had both NP1 and NP2 responses. Therefore, we present an analysis of the subset of items that did receive both NP1 and NP2 responses in both IP and noIP conditions. Mean reaction times to select the initial NP (529.2 ms) vs. someone other than the initial NP (542.1 ms) in proIP condition did not differ, indicating that both are easily interpreted as the subject of the REL-marked verbs. On the other hand, the noIP condition showed a clear advantage for selecting the initial NP (672.4 ms) vs. someone other than the initial NP (1023.4 ms).

![Figure 9](image_url)

Figure 9  Mean response times for pro-type comprehension questions in Experiment 2.

Results of the auditory experiment demonstrate that the location of the prosodic boundary was critical to the process of comprehending syntactically ambiguous sentences. To test our hypothesis about the nature of the inserted pro, we completed two additional experiments using materials constructed to exclude the reflexive interpretation. In the following section, we first report the second written norming study.
3.3. Written study 2

A second written norming study was conducted first to investigate participants’ preferred interpretation of sentences like those in (2) and (3). The method was similar to the first written experiment. 36 test items that have alternating head nouns (for gap/pro distinction) were mixed with additional 36 filler sentences in a pseudo-randomized list where no three sentences of the same type occurred consecutively. Response options differed slightly from those in Experiment 1. After the question, e.g., “Who got exiled?” the three choices for temporarily ambiguous gap type sentences like (2) included the initial NP (NP1 in the graph), the head noun (NP2), and either the initial NP or the head noun (either). For globally ambiguous pro type sentences such as (3), the three choices given were the initial NP (NP1 in the graph), someone other than the initial NP (NP2), and either the initial NP or someone other than the initial NP (either). There were 24 native Seoul Korean speakers from the OSU community, who were paid $7 for participation.

![Graph showing mean percentage of NP choices in Experiment 3.](image)

Figure 10  Mean percentage of NP choices in Experiment 3.

The results showed the exact same pattern found in Experiment 1. For gap types, syntactic bias was so strong that almost no other choices were selected other than the head noun. For pro types, there was again a considerable number of either responses, showing that ambiguity still remains. Based on this, we conducted another auditory listening test, reported in the following section.

3.4. Auditory experiment 2

We conducted the second auditory perception study using the materials from Experiment 3. Experimental and filler items were recorded by a trained phonetician, a female native speaker of Seoul Korean. Four
conditions were created for a factorial design, with two types of syntax (gap vs. pro) and two types of prosodic phrasing (IP vs. noIP boundary after the initial NP). Thirty-six experimental items were counter-balanced across four lists, so there were 9 sentences in each condition in each list. These were mixed with 36 filler sentences (18 sentences with an IP boundary and 18 with no IP boundary) and pseudo-randomized so that no three sentences of the same type occurred consecutively.

Forty-four native speakers of Seoul Korean participated in the experiment for a payment of $7 each. The method was the same as that in Experiment 2, with the same response types used in Experiment 3.

Statistical analyses were conducted on the arcsine transformed data of response choices on the comprehension question. Gap types were less ambiguous, resulting in a main effect of sentence type [F1(1, 43)=57.85, p<0.001; F2(1,35)=80.41, p<0.001]. In addition to a main effect of prosody [F1(1,43)=7.56, p<0.01; F2(1,35)=7.19, p<0.05], there was an interaction of sentence type and prosody [F1(1,43)=25.92, p<0.001; F2(1,35)=33.87, p<0.001] and an interaction of sentence type, prosody and response choices [F1(1,43)=12.5, p<0.001; F2(1,35)=13.28, p<0.001].

For the ease of exposition, we again split the discussion into gap and pro type sentences. Figure 11 shows the results of comprehension questions for gap type.

![Figure 11](image)

Figure 11  Mean percentage of NP choices in gap types in Experiment 4.

Again, the syntactic bias was very strong; in both prosodic conditions, participants overwhelmingly correctly chose the head noun (NP2) as the answer to the comprehension questions. However, a planned comparison showed a significant difference between the number of NP2 response in the IP (95.71%) and noIP (75.52%) gap conditions [F1(1,43)=74.26, p<0.001; F2(1,35)=60.65, p<0.001]. This difference suggests that the mismatch
between syntactic and prosodic structures misled listeners, creating some “garden path” misanalysis effect for this type.

Figure 12 shows the results for pro type sentences, which had a standing syntactic ambiguity. Again, a planned comparison between different NP choices was conducted. In contrast to the results of Experiment 2, the difference between the number of NP1 (20.95%) and NP2 (52.25%) choices was significant in the proIP condition on both subject and item analyses [F1(1,43)=44.35, p<0.001; F2(1,35)=43.99, p<0.001]. This means that as predicted, the IP could lead to the interpretation where pro is posited as the subject of the REL-marked verb. Planned comparisons showed significantly more NP1 choices (59.1%) than NP2(11.1%) choices in the proNoIP condition [F1(1,43)=124.38, p<0.001; F2(1,35)=140.73, p<0.001], and more NP1 choices in the proNoIP condition than in the proIP condition [F1(1,43)=83.03, p<0.001; F2(1,35)=87.13, p<0.001]. This indicates that, in the absence of an IP boundary, listeners were most likely to interpret the initial NP as the subject of the REL-marked verb. Finally, there were significantly more NP2 choices in proIP than in proNoIP [F1(1,43)=74.26, p<0.001; F2(1,35)=83.93, p<0.001].

![Figure 12](image_url)

Figure 12  Mean percentage of NP choices in pro types in Experiment 4.

The results from the second auditory experiment clearly demonstrate the effects of prosodic boundaries on resolving syntactic clausal boundaries. It was shown that the IP boundary could mislead the parser toward a wrong syntactic interpretation, causing a “garden path” effect on temporarily ambiguous sentences. Furthermore, from the results of pro type sentences, we can see that this prosodic boundary even can lead listeners to build a structure where a phonologically null element is posited as an argument of embedded predicates. This provides further evidence of the robustness of prosodic information in sentence comprehension, since this type of structure is not often encountered or interpreted in the written form in the absence of previous disambiguating context.
4. Conclusion

We reported the results from four experiments that examined the effects of intonation phrase boundaries on the resolution of ambiguous relative clause constructions in Seoul Korean. The findings of numerous studies that showed robust effects of prosodic structure on sentence comprehension in English were confirmed for a language that is typologically totally different from English. Prosodic boundary information in Korean was shown to be critical in understanding both temporarily and globally ambiguous sentences. Furthermore, the results from globally ambiguous sentences suggest that given a string of words with multiple possible syntactic attachment sites, prosodic boundaries can have a major impact on leading the parser to build a certain syntactic structure, even before encountering the relevant verb information. The results from a temporarily ambiguous structure indicate that prosody can influence processing, even when the overall sentence ambiguity can be resolved by syntactic information. For a future study, it will be interesting to examine the role of prosodic boundaries on the same syntactic structure in Japanese, which shares the same syntactic properties with Korean but differs in lexical contribution to prosodic phrasing.

References


