The Prosodic Resolution of Syntactic/Semantic Ambiguity: An Exemplar-based Account

Hyunah Ahn
(Seoul National University)

This study tests the hypothesis that the ambiguity of a null argument construction in Korean like *Lwummeituka mwusewun ka poayo* (The roommate must be scared/scary) can be prosodically resolved. The null argument construction is tested with two-place psychological predicates such as *mwusewun* (scared/scary) and *kwichanun* (bothered/bothersome). A naturalness rating experiment shows the following: (1) when the NP *Lwummeituka* is the experiencer of the psychological predicate *scare* and the sentence means 'The roommate must be scared,' both an Accentual Phrase (AP) boundary and an Intonational Phrase (IP) boundary are equally accepted between the NP and the VP; (2) when the NP is the stimulus (The roommate must be scary), only an AP boundary is perceived natural. Reaction time and comprehension question accuracy data further display the role played by prosody in ambiguity resolution. The results are discussed within the exemplar framework.

Keywords: Thematic roles, prosodic disambiguation, exemplar-based language processing, double nominative construction, psychological predicates

1. Introduction

Attempts to explain the formation of grammar through the accumulation of exemplars in memory have been made mainly in two independent branches: phonetic exemplar theories and syntactic exemplar theories. The phonetic details that are seemingly irrelevant to the categorical distinction of phonemes were found to influence lexical decisions (Johnson, 1997;

Pierrehumbert, 2000; 2016). Such findings led to the argument that abstract categories emerge as a result of comparing percepts to stored exemplars. As for syntax, word chunks, often referred to as constructions, are units of exemplars. Through repeated uses of such constructions in situations and contexts, meanings are mapped onto forms, and grammar emerges based on the distribution and frequency of constructions via analogies and generalizations (Bybee, 2006; Goldberg, 2003; Tomasello, 2000). There have been recent attempts to unify the two branches of exemplar-based language learning (Hay & Bresnan, 2006; Walsh, Mobius, Wade, & Schutze, 2010), however, studying the interface of the two branches is still a road untrodden by many.

Hay and Bresnan (2006) present data in which words or phrases of higher frequency lead the sound change of New Zealand English. The vowel /æ/ in *hand* is more raised when it refers to a limb than an abstract concept (such as help as in *give a hand*), and /ɪ/ in *give* is more centralized when it takes for a complement an abstract theme (as in *give a hand*) than a concrete object (*give a present*). The results of the study lead to some intriguing questions: What is the unit of an exemplar: A phoneme, a lexical item, or a phrase? And what is stored with an exemplar: Phonetic details, other linguistic contexts and/or extra-linguistic information? All these questions are remaining unanswered and more examples from other languages than English are needed in which linguistic changes result from co-occurrences of two or more factors of no apparent relationship.

The current study aims to contribute to this quest by probing how a distribution of prosodic junctures of different sizes can lead to the syntactic and semantic disambiguation of a null argument construction. The interface of prosody and syntactic structures is a new area to be explored via the framework that views the repetition of episodes and the accumulation of exemplars in memory as the most important factor in language learning and processing. Below is a preliminary study that explores the potential role of episodic memories of prosody that accompanies syntactic structures.

2. Background

The role of prosody in sentence processing has been investigated by many researchers and their studies cover a great variety of topics from attachment ambiguities (Snedeker & Trueswell, 2003; Trueswell, Tanenhaus, & Garnsey, 1994) to discourse comprehension (Ito & Speer, 2008) to reference resolution (Dahan, Tanenhaus, & Chambers, 2002). Among the various branches of studies on prosody in sentence processing, a topic relatively most relevant to the current study will be the resolution of attachment ambiguities. To explore this topic, researchers most commonly use the construction that includes three elements (A, B, C), which can be phrased either as [[A+B] C] or [A [B+C]]. Examples in (1) and (2) show cases of English and Korean, respectively, in which parsing decisions can be made based on the location of prosodic breaks.

(1) Tap the frog with the flower.   From Snedeker and Trueswell (2003)
    a. [[Tap the frog] with the flower]
    b. [Tap [the frog with the flower]]

(2) Hyenmyeng-han aki-uy   appa   From Schafer and Jun (2002)
    Wise-ADJ   baby-GEN dad
    a. [[hyenmyeng-han aki-uy] appa]
      [[the wise baby's] dad]
    b. [hyenmyeng-han [aki-uy appa]]
      [the wise [baby's dad]]

The existing accounts of prosodic disambiguation, however, cannot explain how sentences as in (3) can also be prosodically disambiguated. The sentence can be interpreted as either (3a) 'Cengyen must be frightened' or (3b) 'Cengyen must be frightening.' The sentence seemingly has nothing to do with attachment decisions. Within the sentence, *Cengyennika* and *twuryewun ka poayo* will always be separated into the nominative NP and the VP. Unlike the examples in (1) and (2) where there are three elements A, B, and C, there are only two elements A and B in (3). Therefore, such a phrasing distinction as [[A+B] C] and [A [B+C]] cannot be made in (3).

(3) [[Cengyeni-ka]~NP~ [twuryewu-n ka po-a-yo[1])]~VP~]~S~
    Cengyeni-NOM   frighten-ADJ INF AUX-DEC-HON
    a. 'Cengyen must be frightened.'
    b. 'Cengyen must be frightening.'

Despite the apparent lack of phrasing differences, the sentence can be disambiguated via prosodic phrasing: (3a) sounds natural with either a smaller or a larger break, operationalized by an Accentual Phrase (AP) and an Intonational Phrase (IP), respectively, in K-ToBI (Jun, 2000), but (3b) sounds natural only with an AP. This paper aims to demonstrate [1] that the two readings of a null argument construction in (3) can be distinguished via prosodic junctures of different sizes and [2] the prosodic disambiguation can be accounted for within the framework of exemplar-based language processing.

   The ambiguity in (3) comes from the combination of the double nominative construction (DNC) and the null argument features in Korean. The canonical sentence structure that involves a psychological predicate as *twuryewun* takes its experiencer with a dative case and its stimulus with a nominative case. That is, in (4a) *Cengyeni* has a dative case marker *-eykey,* followed by an optional topic marker *-nun,* and *tongsayng* (younger sibling of either sex) is marked nominative with *-i* (or *-ka* for words ending with a vowel). Here, *Cengyen*, the one who experiences the psychological state of "fear" takes a dative case and the one who stimulates the psychological state is the younger sibling, which takes the nominative case.

   The experiencer / stimulus relationship is shifted in (4b) with the opposite case marking. Since the dative and nominative case will dictate who experiences and stimulates a psychological state, the two arguments can be scrambled as long as the case marking remains intact; therefore, (4a') and (4b') still retain the same meaning as (4a) and (4b), respectively.

(4) a. Cengyeni-eykey(-nun)   tongsayng-i   twuryewun   ka po-a-yo
       Cengyen-DAT-TOP   younger sibling-NOM frighten-ADJ INF AUX-DEC-HON
       'To Cengyen, the younger sibling must be frightening.'

---

1) All transliterations of Korean follow the Yale Romanization.

    b. Tongsayng-eykey(-nun)   Cengyeni -ka   twuryewun   ka   po-a-yo
       younger sibling-DAT-TOP   Cengyen-NOM   frighten-ADJ   INF   AUX-DEC-HON
       'To the younger sibling, Cengyen must be frightening.'
    a'. Tongsayng-i   Cengyeni-eykey(-nun)   twuryewun   ka   po-a-yo
       younger sibling-NOM   Cengyen-DAT-TOP   frighten-ADJ   INF   AUX-DEC-HON
       'The younger sibling, to Cengyen, must be frightening.'
    b'. Cengyeni-ka   Tongsayng-eykey(-nun)   twuryewun   ka   po-a-yo
       Cengyen-NOM   younger sibling-DAT-TOP   frighten-ADJ   INF   AUX-DEC-HON
       'Cengyen, to the younger sibling, must be frightening.'

The scrambling of the two arguments, however, becomes difficult when both arguments take the nominative case as in (5); this is the so-called double nominative construction in Korean. When both the experiencer and stimulus arguments take the nominative case, it is the word order that provides cues in assigning thematic roles to the NPs.

   (5) a. Cengyeni-ka   tongsayng-i   twuryewun   ka   po-a-yo
       Cengyen-NOM   younger sibling-NOM   frighten-ADJ   INF   AUX-DEC-HON
       'Cengyen must be frightened of the younger sibling.'
    b. Tongsayng-i   Cengyeni-ka   twuryewun   ka   po-a-yo
       younger sibling-NOM   Cengyen-NOM   frighten-ADJ   INF   AUX-DEC-HON
       'The younger sibling must be frightened of Cengyen.'

The first NP will be the experiencer and the second will be the stimulus. This nominative marking of both arguments creates the ambiguity introduced in (3) since either (or both) of the NPs can be silent as in (6). When one of the two nominative NPs is omitted, it is the discourse context that can clarify which thematic role the overt NP denotes. If the omitted null argument is a stimulus (6a), the overt NP will be the experiencer. If the experiencer is omitted as in (6b), the overt NP will be the stimulus. Throughout the paper, the two interpretations in (6a) and (6b) will be referred to as the experiencer reading and the stimulus reading, respectively.

   (6) Cengyeni-ka twuryewun ka po-a-yo.
    a. Cengyeni-ka (null stimulus) twuryenwun ka po-a-yo
       'Cengyen must be frightened.'
    b. (Null experiencer) Cengyeni-ka twuryewun ka po-a-yo
       'Cengyen must be frightening.'

(7) a. Experiencer [AP] stimulus [AP] verb phrase
    b. Experiencer [IP] stimulus [AP] verb phrase

As for the prosody of the two-argument sentences in (5), a general con-
sensus is expected among native speakers of Korean. Both the boundary
between the experiencer NP and the stimulus NP and the boundary be-
tween the stimulus NP and the VP will at least have the juncture of
Accentual Phrases (AP), a prosodic unit of a phonological word or longer
in Korean prosody (Jun, 2000) (7a). A larger prosodic unit, the
Intonational Phrase (IP), is not mandatory but allowable between the
experiencer NP and the stimulus NP (7b). That is, between the experiencer
and the stimulus NPs, both AP and IP breaks are possible but only AP
breaks are allowed for the juncture between the stimulus NP and the
VP. More importantly, it is hardly conceivable that the juncture between
the stimulus NP and the VP will be larger than that between the experiencer
NP and the stimulus NP.

    Under the assumption that it is phonetic details of exemplars that are
stored in our memory, one can postulate that every time an experiencer
NP is said, either an AP or IP break will be used and the combination
of the thematic role and prosody will be stored as episodes. In the case
where a stimulus NP is pronounced, it will be an AP but not an IP
break that accompanies the NP. Again, the episodes of hearing the stimulus
NP with an AP break will be stored in memory. This distribution of
prosodic junctures then carries over to the null argument construction
in (6), where one of the arguments is omitted. The accumulation of epi-
sodes in which one hears the sentences in (5) with the prosodies in (7)
will lead to the formation of a rule that the experiencer reading (6a) can
be said with either a small or large prosodic break (8) but the stimulus
reading (6b) must be said with a smaller prosodic break (8a). To probe
that postulation, a naturalness rating experiment was conducted; partic-
ipants were asked to rate the naturalness of the construction at issue when
it was manipulated for meaning (6a-b) and prosody (8a-b).

(8) a. Ambiguous argument [AP] verb phrase
    b. Ambiguous argument [IP] verb phrase

3. Method

The critical construction (9) was fully contextualized so that the thematic role of the overt NP can be clarified as the experiencer (10) or as the stimulus (11) of the psych-predicate. The contextual clarification of the critical sentences is robust and it seems that the mere difference of smaller and larger prosodic breaks, realized as AP and IP junctures between the NP and the VP, is not strong enough to change the interpretation of the sentence altogether.

(9) Lwummeitu-ka   mwusewu-n   ka   pwa-yo
   roommate-NOM   scare-ADJ   INF   must-DEC
   'The roommate must be scared/scary.'

(10) Experiencer context[2]
   a. Yenga's roommate went out in Yenga's shoes without asking.
   b. In the evening, the roommate came home.
   c. Yenga got angry and started screaming.
   d. Lwummeitu-ka mwusewu-n ka pwa-yo ('The roommate must be scared.')
   e. She's being quiet not knowing what to say.
   f. Comprehension Question: Who went out in the other's shoes without asking?
    (a) Yenga   (b) Roommate

(11) Stimulus context
   a. Yenga and the roommate are having a big fight.
   b. The roommate lost her temper and started screaming.
   c. Yenga is walking backwards away from the roommate.
   d. Lwummeitu-ka mwusewu-n ka pwa-yo ('The roommate must be scary.')
   e. (Yenga) went into the bedroom and locked the door inside.
   f. Comprehension Question: Who's retreating into the bedroom?
    (a) Yenga   (b) Roommate

If the postulated distribution of prosodic breaks described above reflects the actual distribution, listeners will find it natural that (10d) is read with either a smaller or larger break (AP or IP) and (11d) is read with a smaller break (AP) only. To test the hypothesis that Korean native speakers will

---

2) All items and comprehension questions were given in Korean. The example sentences and comprehension questions but the critical sentences were translated from Korean to save space.

show such sensitivity to the mapping between the interpretation of the sentence and the prosodic breaks, a naturalness rating experiment was conducted as follows. The prediction was that the naturalness rating of the critical sentence embedded in the experiencer context (10d) will not differ by prosody while that in the stimulus reading (11d) will change by a function of prosody. In other words, the naturalness rating of the experiencer reading (10d) will be the same whether the prosodic break is small or large (AP or IP) while that of the stimulus reading (11d) will be higher with a smaller break (AP) than with a larger break (IP).

Participants. Korean native speakers were recruited from a campus of a major university in the U.S. When they arrived in the lab, they first filled out a language background questionnaire, which was intended to screen non-native Korean speakers, Korean speakers who were not born and raised in Korea or who lived overseas for five years or longer. Data from 41 participants were used for analysis. Participants were paid five dollars for a 30-minute session,[3] which included the language background survey and the main experiment.

Stimuli. Ten psychological predicates, combined with different NPs and inflected with different tenses, were used to create twenty different critical sentences (see Appendix for the complete list of critical sentences). The twenty sentences were contextualized in stories similar to those shown above in (10) and (11) in order for the critical NP to mean either the experiencer or the stimulus. Then, the stories were recorded in a natural story-telling tone by a female native speaker of Korean. When experiencer contexts were recorded, the critical NP was read with an IP boundary tone. When stimulus contexts were recorded, the critical NP was recorded with an AP juncture.

Table 1 summaries the mean lengths of the nominative case marker and the prosodic juncture in sentences read with IP and AP prosodies. It also shows the results of pairwise comparisons between the two prosodic conditions. The mean lengths of the two prosodic junctures were sig-

---

3) The main experiment took about 20 minutes; of which the first five minutes was a recorded instruction and three practice trials. Each participant took a different amount of time to rate and answer each trial and some spent a couple more minutes to ask questions between the instruction and the actual trials.

nificantly different in both measurements. In addition to the length differ-
ences, IP boundaries were also pronounced with distinct IP contours
whereas APs lacked such contours (see Figures 1 and 2).

Table 1. The Mean Lengths (ms) of the Nominative Case Marker and
the Break After the NP

	IP	AP	*t*-value	S.D.	*p*-value
Nominative case marker	420	100	-23.92	50	<.001
Prosodic break after NP	120	20	-4.61	80	<.001

For the critical sentences to be heard via both prosodies, the critical senten-
ces from the two contexts were cropped and spliced into the opposite
context. This way, four different conditions were created: Experiencer
reading with an IP, experiencer reading with an AP, stimulus reading
with an IP, and stimulus reading with an AP. The combination of con-
textualizing each of the twenty critical sentences for two different thematic
roles and reading them with two different prosodies resulted in a total
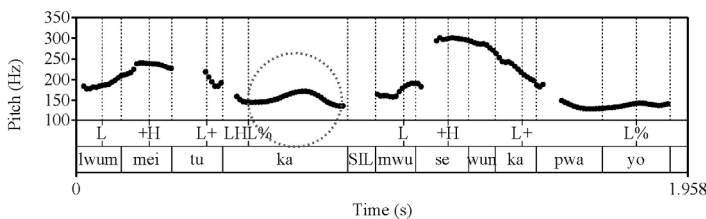of 80 items, out of which four lists were created.

Figure 1. Intonational Phrase boundary tone on the nominative marker.
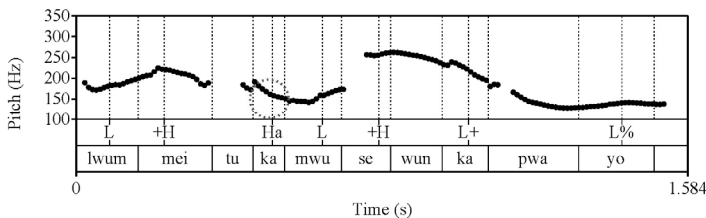

Figure 2. Accentual Phrase on the nominative NP.

Each list had all of the 20 critical sentences and was counterbalanced with items and conditions. That is, in any given list, none of the twenty items were repeated and none of the four conditions were repeated through the same item. This Latin-square design was employed to control the effects of random variables such as items and subjects. Participants' responses to a particular condition can be affected by being exposed to the same item through a different condition in an earlier trial. Unlike in between-item or between-subject designs, the Latin-square design prevents item effects and subject effects from being misinterpreted as the effects of experimental factors.

Procedure. Participants were given recorded instructions and were given time to ask questions as necessary after they listened to the instructions. The recorded stories were presented through the experimental software Psyscope (Cohen, MacWhinney, Flatt, & Provost, 1993) in a self-paced listening format. Participants were presented stories auditorily sentence by sentence. After every sentence, they were instructed to give a rating to the sentence they just heard on its naturalness. The rating was given by pressing one of the six buttons as quickly as possible, which were evenly distanced from each other and from the participant's hand. As soon as they pressed a button for its naturalness on a scale of 1 (Very unnatural) to 6 (Very natural), the next sentence was played. Both their ratings and reaction times (RT) for rating were measured. Because all sentences, not only critical sentences but also all non-critical sentences, were rated, extra filler items were not used. After each story was completed, a comprehension question followed to ensure that participants are paying full attention to each story and that their ratings were based on the thorough understanding of each story. The accuracy of their answers and the reaction time to answer questions were measured.

Statistical analysis. The four measurements (naturalness rating of critical sentences, RT of the rating, comprehension question accuracy, and RT of comprehension question) were analyzed using mixed effects modeling using the lmerTest package (Kuznetsova, Brockhoff, & Christensen, 2017) in R (R Core Team, 2016). Thematic roles and prosody were set as fixed factors and participants and items as random variables.

All four measurements but comprehension question accuracy were ana-
lyzed using linear mixed effects regression. Since correct and incorrect
answers for comprehension questions were coded as 1 and 0, respectively,
a logistic mixed effect model, with only random intercepts included, was
used for the analysis of binary outcomes. For the three other measure-
ments, linear mixed effects models were used and the maximal structure
of random variables were first considered. That is, all models were first
tried with both random intercepts and slopes. Only when a maximal struc-
ture model did not converge, were random slopes reduced step by step
until the model converged.

4. Results

The results confirm the prediction that the naturalness rating of the
experiencer reading will not differ by prosody but that of the stimulus
reading will be higher with an AP juncture. Comprehension question
accuracy data and RT measurements, originally intended to check whether
any participants or items would behave idiosyncratically, show unexpected
but noteworthy patterns. Comprehension question accuracy data indicate
that participants comprehend the experiencer context with an IP break
better than that with an AP break, and RT data show that both critical
sentences and comprehension questions of the experiencer reading take
longer to process.

Naturalness rating of critical sentences. Descriptive statistics of naturalness
rating by thematic role and prosody are provided in Table 2. A linear
mixed effects regression was modeled with a maximal structure of random
intercepts and slopes (Table 3).[4] A significant interaction effect of thematic
role and prosody was observed as well as a significant main effect of
thematic role.

---

4) rating ~ theme * prosody + (1+theme*prosody|item) + (1+theme*prosody|participant).
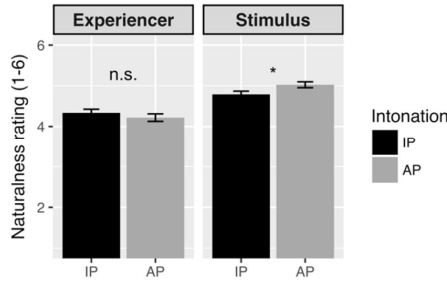
Figure 3. Naturalness rating of critical sentences on 6-point scale.
Note. Error bars indicate mean ± standard error.

Table 2. Naturalness Rating by Thematic Role and Prosody

	Themes	Prosody	Mean rating	SD
Condition 1	Experiencer	IP	4.338	1.403
Condition 2	Experiencer	AP	4.219	1.489
Condition 3	Stimulus	IP	4.792	1.244
Condition 4	Stimulus	AP	5.027	1.157

Table 3. Linear Mixed Effects Rregression of Thematic Roles and Prosody in Naturalness Rating

	Estimate	Std. Error	*t*-value	Pr(>|t|)
(Intercept)	4.595	0.153	30.108	0.000
Thematic roles	-0.315	0.082	-3.835	0.000
Prosody	0.029	0.041	0.718	0.479
Thematic roles x Prosody	-0.088	0.041	-2.156	0.039

A pairwise comparison[5]) shows that the rating of the experiencer reading did not differ by prosody while that of the stimulus reading did, as was predicted. The experiencer-IP combination resulted in the mean rating of 4.34 and the experiencer-AP combination had the mean rating of 4.22. The difference was not statistically significant (β = .12, *S.E.* = .10, *z*-value = 1.13, *p* = .67). However, the stimulus-IP combination elicited the mean rating of 4.79 while the stimulus-AP combination 5.03. The difference

5) A multiple comparison procedure was applied to the lmer model: rating ~ condition + (1|item) + (1|participant).

was statistically significant ($\beta = -.23$, *S.E.* = .10, *z*-value = -2.26, *p* = .047).

Although there was no specific prediction regarding the interaction of the two variables, the interaction of thematic role and prosody is a reasonable result in that the effect of one variable applies differently to the two different levels of the other variable. That is, prosody makes a difference in rating the naturalness of the stimulus reading sentences while it does not for the experiencer reading sentences.

RTs for naturalness rating. Table 4 summarizes the descriptive statistics for the reaction times (RTs) for naturalness rating. A linear mixed effects model[6)] shows a robust main effect of thematic roles (Table 5). It indicates that participants took significantly longer to rate a critical sentence in the experiencer reading than in the stimulus reading. Unlike the results of naturalness rating, there was no interaction effect observed, and neither the experiencer reading nor the stimulus reading showed a different reaction time by a function of prosody.
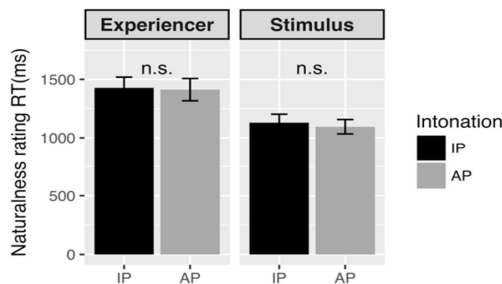

Figure 4. Reaction Time (RT) for naturalness rating of critical sentences. *Note.* Error bars indicate mean ± standard error.

Table 4. Reaction Time (RT) for Naturalness Rating by Thematic Role and Prosody

	Themes	Prosody	Mean RT	SD
Condition 1	Experiencer	IP	1427.76	1475.91
Condition 2	Experiencer	AP	1412.43	1534.40
Condition 3	Stimulus	IP	1130.25	1169.03
Condition 4	Stimulus	AP	1095.66	991.07

6) rating.RT ~ theme * prosody + (1+prosody|item) + (1+theme*prosody|participant).

Table 5. Linear Mixed Effects Regression Model of Thematic Roles and Prosody in RT of Naturalness Rating

	Estimate	Std. Error	*t*-value	Pr(>|t|)
(Intercept)	1266.525	74.525	16.995	0.000
Thematic roles	154.253	50.449	3.058	0.004
Prosody	-12.483	41.402	-0.301	0.765
Thematic roles x Prosody	5.371	38.414	0.140	0.889

Accuracy of responses to comprehension questions. Figure 5 and Table 6 gives an overview of the accuracy rate of comprehension questions. A logistic mixed effects regression with random intercepts for items and participants[7] showed a significant interaction of thematic roles and prosody (Table 7). Also, a pairwise comparison between conditions[8] showed that the accuracy of responses to comprehension questions significantly differed in the experiencer reading by a function of prosody ($\beta = 1.26$, *S.E.*= 0.49, *z*-value = 2.60, *p* = .045) but not in the stimulus reading ($\beta = -0.25$, *z*-value= 0.40, *S.E.* = -0.62, *p* = 0.924).
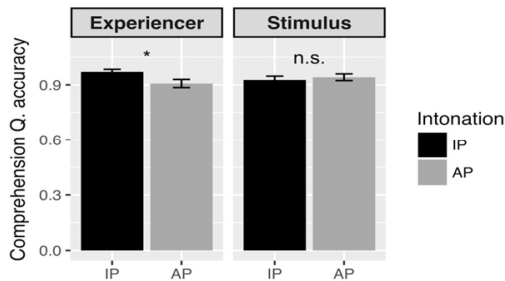
Figure 5. The accuracy ratio of responses to comprehension questions (1=100%).
Note. Error bars indicate mean ± standard error.

---

7) accuracy ~ theme * prosody + (1|item) + (1|participant)
8) A multiple comparison procedure was applied to the logistic regression model: accuracy ~ conditions + (1|item) + (1|participant).

Table 6. The Means and Standard Deviations of Accurate Responses to Comprehension Questions

	Themes	Intonation	Mean accuracy	SD
Condition 1	Experiencer	IP	0.971	0.197
Condition 2	Experiencer	AP	0.907	0.318
Condition 3	Stimulus	IP	0.927	0.294
Condition 4	Stimulus	AP	0.941	0.260

Table 7. Nonlinear Mixed Effects Regression Model of Thematic Roles and Prosody in Accuracy of Comprehension Questions

	Estimate	Std. Error	z-value	$Pr(>|z|)$
(Intercept)	3.0415	0.258	11.804	0.000
Thematic roles	0.1096	0.159	0.689	0.491
Prosody	-0.2535	0.159	-1.595	0.111
Thematic roles x Prosody	-0.3808	0.159	-2.390	0.017

RTs for comprehension questions. Similarly to the RTs for naturalness rating, reaction time to answer comprehension questions also showed a main effect of thematic roles with a longer time needed to answer comprehension questions after hearing a story of the experiencer context.
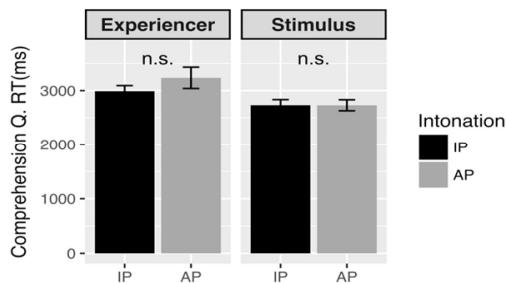

Figure 6. Reaction Time (RT) for comprehension questions.
Note. Error bars indicate mean ± standard error.

Table 8. Reaction Time (RT) for Comprehension Questions

	Themes	Intonation	Mean RT	SD
Condition 1	Experiencer	IP	2990.395	1470.395
Condition 2	Experiencer	AP	3236.273	2814.613
Condition 3	Stimulus	IP	2731.405	1473.208
Condition 4	Stimulus	AP	2730.259	1457.184

Table 8 summarizes the mean RTs for the four conditions and their standard deviations, and Table 9 shows the output of the linear mixed effects regression model of thematic roles and prosody.[9] Just as participants took longer to rate critical sentences of experiencer reading, they also took longer to answer comprehension questions of experiencer contexts. The *p*-value of .004 indicates that the main effect of thematic role was statistically significant in time taken to answer comprehension questions.

Table 9. Linear Mixed Effects Regression Model of Thematic Roles and Prosody in RT for Comprehension Questions

	Estimate	Std. Error	*t*-value	Pr(>|t|)
(Intercept)	2922.04	142.43	20.516	0.000
Thematic roles	189.28	64.03	2.956	0.004
Prosody	58.96	79.47	0.742	0.462
Thematic roles x Prosody	62.16	63.89	0.973	0.334

In sum, the naturalness rating data confirmed the prediction between the thematic roles and prosody while the other measurements presented unexpected patterns. In addition, comprehension question accuracy data resulted in the opposite of what was predicted of naturalness rating: the accuracy of comprehension questions was higher when an IP break was given in the experiencer context but no prosodic distinction was made in the stimulus context. Finally, RT data for both naturalness rating and comprehension questions indicated that overall, the processing of the experiencer reading takes longer than the stimulus reading. The next section discusses the implications of these results in detail.

---

9) cq.RT ~ theme * prosody + (1+prosody|item) + (1+theme*prosody|ID).

5. Discussion

The hypothesis was laid out in the beginning of the paper that when exemplars are accumulated such that the experiencer NP is accompanied by either an AP or IP break but the stimulus NP is said only with an AP break in the double nominative construction, it will lead to the formation of a rule that a null argument construction as in (9) must be said with an AP break for the overt NP to denote the stimulus. This hypothesis on the mapping between prosody and sentence meaning led to the prediction that the naturalness rating of experiencer reading sentences will not differ by prosody but that of stimulus reading sentences will. And the data from the naturalness rating experiment bore out the prediction. The naturalness ratings of the experiencer reading were very close to each other between the two prosodic conditions, and those of the stimulus reading were significantly different from each other by a function of prosody. Stimulus reading sentences were rated much higher when it was said with an AP break.

The data from the experiment, however, also displayed some unexpected effects. First, a main effect of thematic role was observed in the naturalness rating, rating RTs, and comprehension question RTs. The main effect of thematic role in the naturalness rating is surprising since both interpretations of the ambiguous construction are allowed. Rather, one might predict a higher rating for the experiencer reading since it is a much more common interpretation when the ambiguous sentence is presented out of context (Ahn, 2014). As for reaction times, although no explicit prediction was made with regard to how long it will take to rate critical sentences or answer comprehension questions, it would have accorded with the prediction for the naturalness rating if the reaction times had also changed by a function of prosody only in the stimulus condition.

Also unexpected was the significant difference of accuracy in responses to comprehension questions between the experiencer-IP condition and the experiencer-AP condition. The significant interaction of thematic role and prosody solidifies the fact that the accuracy difference by a function of prosody was observed in the experiencer reading but not in the stimulus

condition. It means prosody made a difference only in the experiencer reading but not in the stimulus reading. This is the opposite pattern from what was observed in the naturalness rating, which is very surprising because the prediction was that prosody will not make a difference in processing the experiencer reading and will do so in the stimulus reading. Although the reaction times and comprehension question accuracy were measured to ensure that participants are paying their full attention to each trial, the analysis results of those measurements have important implications in how the critical construction is processed within contexts and are worthy of in-depth discussions.

First of all, a significant main effect of thematic role was observed in naturalness rating (see Table 3). The significantly lower rating of the experiencer reading is surprising but should not be interpreted as participants' dispreference of the construction per se. Rather, it should be related to the shift of focus in the experiencer context, which was not observed in the stimulus context. For example, in both (10) and (11), the two main characters of the story, *Yenga* and *the roommate* are introduced in the very first sentence. Then, the second sentence is about the roommate, which means the roommate is put on the center of the discourse. This means that the focus of the story is on the roommate in the second sentence, but it moves to *Yenga* in the third sentence. When the fourth, critical sentence follows, there is a focus shift in the experiencer context yet again. But in the stimulus context, the focus of the story remains on *Yenga*. That is, the critical sentence in (10d) is about the roommate's feeling but the critical sentence in (11d) is about Yenga's feeling. The last sentences continue the focus of the previous sentence (critical sentence) in both contexts. The last sentence of the experiencer context is about the roommate and that of the stimulus context is *Yenga*, a focus of the discourse continued from the third sentence.

This difference in terms of focus shift and continuity is related to processing costs of each reading, which is related to the distribution of pronouns according to the centering theory (Gordon, Grosz, & Gilliom, 1993; Grosz, Joshi, & Weinstein, 1995). At any given point in a discourse, there is always one referent that is the current focus (or the focus in the immediate

past) and a group of candidates to be the next focus of the discourse. They are called the backward-looking center and forward-looking centers, respectively (Gordon et al., 1993). When the same focus of the discourse is continued from one sentence to the next, it means the backward-looking center continues and it is realized as an unstressed pronoun in English or a null pronoun in Korean. When the focus of the discourse is shifted to one of the candidates in the forward-looking center, a definite NP (an overt NP in Korean) is required to indicate that the discourse now shifts its focus to a different referent.

Due to this constraint, the experiencer contexts had to be created such that the overt NP of the critical sentence was different from the focus of the previous sentence. This meant that the experiencer reading was conflated with a focus shift. The stimulus contexts, on the other hand, had to maintain the focus of the previous sentence as a null NP, which, also led to the conflation of the stimulus reading and focus continuity. Grosz et al. (1995) also notes that focus shift requires a higher processing load, which implies that the experiencer reading critical sentences in the current study were inherently more costly to process than the stimulus counterparts.

Further evidence that experiencer reading is more difficult to process than stimulus reading is found in the main effect of thematic role in the reaction times (RT) to give a rating to the naturalness of the critical sentences. Figure 4 shows that it took significantly longer for participants to rate the experiencer reading than the stimulus reading.

Meanwhile, prosody did not play a role in comprehension accuracy in the stimulus context. This testifies that when cognitive processors are not burdened, they notice the unnaturalness of a mismatching prosody (IP with stimulus reading) but comprehend the contents despite the unnatural intonation. That is, the stories used in the experiment clarified the thematic roles of each critical sentence unequivocally. Therefore, even if a participant finds an IP boundary in the stimulus reading sentence unnatural, the perception of unnaturalness will not lead to the re-interpretation of the sentence. In other words, as long as the ambiguous construction is embedded in the stimulus context, an IP break alone will not make a participant interpret the NP to be the experiencer. If the same

mechanism is applied to the experiencer reading, whichever prosodic break is used to read the sentence should not make a difference in understanding the meaning of the critical sentence.

However, the analysis of the comprehension questions clearly indicates that participants' understanding of the stories improved when an IP break accompanied the experiencer-reading critical sentence. This function of prosody in the accuracy of responses to comprehension questions indicates that, to compensate for the aforementioned processing difficulty of the experiencer reading, participants make use of the fact that an IP is allowable only for the experiencer reading. In other words, processing the experiencer reading stories takes a higher processing cost and the presence of an IP boundary makes it clear that the critical NP should be interpreted as the experiencer since a stimulus NP will not be accompanied by a larger prosodic break.[10]

In the initial prediction, an IP break was expected not to make any difference in the experiencer reading in comparison to an AP break, which was confirmed in the naturalness rating task. The significantly higher accuracy in the experiencer-IP condition over the experiencer-AP condition, therefore, should be related to the processing difficulty induced by the focus shift in the experiencer condition. That is, when the processing burden increases in the experiencer condition, an IP break can be a clarifying cue that the overt NP is the experiencer since there can only scarcely be exemplars where the stimulus NP is said with an IP break; hence, the higher accuracy in the experiencer-IP condition.

In brief, the present study shows that a sentence that seemingly has nothing to do with attachment ambiguities can have its ambiguity resolved by prosodic phrasing. If a repository of exemplars where thematic roles co-occur with different prosodic phrasing is not assumed, the observed mapping between the interpretations of the null argument construction and different prosodic breaks will hardly be explained.

---

10) An anonymous reviewer mentioned that the comprehension questions are not related to the interpretation of the critical sentences. However, the questions were related to events that incurred psychological states in each story; therefore, the response accuracy of the questions indicates not only the degree to which participants could understand the stories in general but also the extent to which the prosodic differences hinted on the interpretation of the critical sentences.

6. Future Research

The current study showed that participants are sensitive to the combinations of meaning (thematic roles) and form (prosody) that they are not familiar with, which sheds light on the possibility that the unit of an exemplar may go beyond a lexical item. The different prosodic phrasings investigated in this study are composed of two parts. One is the duration of the nominative marker *-ka* at the end of the critical NPs, and the other is the lengths of the pauses after the NP. Even when only the NP is considered, it already goes beyond the unit of a lexical item since it is composed of a noun and a subject marker *-ka.* If the duration of the pause between the NP and the VP is considered, it is definitely at the phrasal level.

However, there are limitations to be overcome to provide clearer answers to questions raised by Hay and Bresnan (2006). First, the inherent conflation of experiencer and stimulus conditions with focus shift and focus continuity, respectively, prevents the results from being directly interpreted with regard to the role played by exemplars. Secondly, the distribution of different prosodic phrasings were predicted based on the native speaker's intuition, an empirical investigation into the issue will have to be via speech corpora and production studies. Finally, stronger evidence is needed that the ambiguous construction out of context can be disambiguated with different prosodic phrasings. The naturalness ratings alone, as in the present study, are not enough to make a strong argument. Future research should address such limitations and examine whether the said construction, presented out of context, will be interpreted differently based on prosodic manipulations.

References

Ahn, Hyunah. (2014). Reference in Discourse: The case of L2 and Heritage Korean. In M. Kenstowicz, T. Levin, & R. Masuda (Eds.), *Proceedings of the 23rd Japanese/Korean Linguistics Conference*. Stanford: CSLI Publications.

Bybee, Joan L. (2006). From usage to grammar: The mind's response to

repetition. *Language* 82.4, 713-733.

Cohen, Jonathan D., MacWhinney, Brian, Flatt, Matthew, and Provost, Jefferson. (1993). PsyScope: A new graphic interactive environment for designing psychology experiments. *Behavioral Research Methods, Instruments, and Computers* 25.2, 257-271.

Dahan, Delphine., Tanenhaus, Michael K., and Chambers, Craig G. (2002). Accent and reference resolution in spoken-language comprehension. *Journal of Memory and Language* 47, 292-314.

Goldberg, Adele E. (2003). Constructions: a new theoretical approach to language. *Trends in Cognitive Science* 7.5, 219-224.

Gordon, Peter C., Grosz, Barbara J., and Gilliom, Laura A. (1993). Pronouns, names and the centering of attention in discourse. *Cognitive Science* 17, 311-347.

Grosz, Barbara J., Joshi, Aravind K., and Weinstein, Scott. (1995). Centering: A framework for modeling the local coherence of discourse. *Computational Linguistics* 22.2, 203-225.

Hay, Jennifer and Bresnan, Joan. (2006). Spoken syntax: The phonetics of giving a hand in New Zealand English. *The Linguistic Review* 23.3.

Ito, Kiwako and Speer, Shari R. (2008). Anticipatory effects of intonation: Eye movements during instructed visual search. *J Mem Lang* 58.2, 541-573.

Johnson, Keith. (1997). Speech perception without speaker normalization. In K. Johnson & J. W. Mullennix (Eds.), *Talker variability in speech processing*. San Diego: Academic Press.

Jun, Sun-Ah. (2000). K-ToBI (Korean ToBI) Labelling Conventions: Version 3. *Speech Sciences* 7, 143-169.

Kimball, Amelia E., Cole, Jennifer, Dell, Gary S., and Shattuck-Hufnagel, Stefanie. (2015). Categorical vs. Episodic memory for pitch accents in English. In the Scottish Consortium for ICPhS 2015 (Ed.), *Proceedings of the International Congress of Phonetic Sciences*. Glasgow: UK: The University of Glasgow.

Kuznetsova, Alexandra, Brockhoff, Per B., and Christensen, Rune H. B. (2017). "lmerTest Package: Tests in Linear Mixed Effects Models." *Journal of Statistical Software* 82.13, 1-26.

Pierrehumbert, Janet B. (2000). Exemplar dynamics: Word frequency, lenition, and contrast. In J. Bybee & P. Hopper (Eds.), *Frequency effects and the emergence of linguistic structure*. Amsterdam: John Benjamins.

_____ (2016). Phonological representation: Beyond abstract versus episodic. *Annual Review of Linguistics* 2, 33-52.

R Core Team. (2016). R: A language and environment for statistical computing.

R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.

Schafer, Amy J. and Jun, Sun-Ah. (2002). Effects of accentual phrasing on adjective interpretation in Korean. *Sentence processing in East Asian Linguistics*, 223-255.

Snedeker, Jesse and Trueswell, John. (2003). Using prosody to avoid ambiguity: effects of speaker awareness and referential context. *Journal of Memory and Language* 48, 103-130.

Tomasello, Michael. (2000). First steps toward a usage-based theory of language acquisition. *Cognitive Linguistics* 11.1/2, 61-82.

Trueswell, John C., Tanenhaus, Michael K., and Garnsey, Susan M. (1994). Semantic influences on parsing: Use of thematic role information in syntactic ambiguity resolution. *Journal of Memory and Language* 33.3, 285-318.

Walsh, Michael, Mobius, Bernd, Wade, Travis, and Schutze, Hinrich. (2010). Multilevel exemplar theory. *Cognitive Science* 34.4, 537-582.

Hyunah Ahn
Department of English Language and Literature
Seoul National University
1 Gwanak-ro, Gwanak-gu, Seoul 08826, Korea
E-mail: prosodygal@snu.ac.kr

Appendix
The complete list of critical sentences in the experiment

1. *Yengswuka kekcengtoyna poayo*   Yengswu must be worried/worrisome.

2. *Halapecika kekcengtoysina poayo*   Grandpa must be worried/worrisome.

3. *Chinkwuka komawuessna poayo*   The friend must have been thankful/merciful.

4. *Chinkwuka komawunka poayo*   The friend must be thankful/merciful.

5. *Nampyeni kwichahnassna poayo*   The husband must have been bothered

                                         The husband must have been bothersome.

6. *Cokhaka kwichahnunka poayo*   The nephew must feel bothered/be bothersome.

7. *Atuli taptaphanka poayo*   The son must be frustrated/frustrating.

8. *Pwuini taptaphanka poayo*   The wife must be frustrated/frustrating.

9. *Sensayngnimi mwusewunka poayo*   The teacher must be scared/scary.

10. *Lwummeituka mwusewunka poayo*   The roommate must be scared/scary.

11. *Acessika pankawuessna poay*   The man must have been pleased to meet (Ø).

                                         The man must have been pleasing to meet.

12. *Taypaykisaka pankawunka poayo*   The delivery guy must be pleased to meet (Ø).

                                         The delivery guy must be pleasing to meet.

13. *Ttoswunika pwurewunka poayo*   Ttoswun must be envious/envied.

14. *Yenaka pwurewunka poayo*   Yena must be envious/envied.

15. *Atuli carangsurewunka poayo*   The sun must be proud.

                                         (Someone) must be proud of the son.

16. *Ayini carangsurewunka poayo*   The lover must be proud.

                                         (Someone) must be proud of the lover.

17. *Caywonika cinkurewunka poayo*   Caywon must feel creepy (about something)

                                         Caywon must be creepy (to someone)

18. *Minswuka cinkurewuessna poayo*   Minswu must feel creepy (about something)

                                         Minswu must be creepy (to someone)

19. *Tonsayngi hansimhanka poayo*   Younger sibling must find (someone) pathetic.

                                         Younger sibling must be pathetic.

20. *Cengmmissika hansimhanka poayo*   Cengmi must find (someone) pathetic.

                                         Cengmi must be pathetic.