Incremental Interpretation of Intonational Phrase Boundaries

Eun-Kyung Lee
(Yonsei University)

Lee, Eun-Kyung. (2014). Incremental Interpretation of Intonational Phrase Boundaries. ***Language Research,*** **50.2, 411-430.**

This study investigates whether intonational phrase boundaries can be used to anticipate upcoming syntactic structure. Participants' eye movements in a visual search task were examined as they listened to auditory instructions that included a globally ambiguous sentence (e.g., "Click on the boy and the girl holding flowers"). An intonational phrase boundary after "girl" creates a bias towards an interpretation in which both the boy and the girl are holding flowers. The presence of an intonational phrase boundary in this location was manipulated. The results showed that there were increased anticipatory fixations to a picture that was directly linked to the interpretation signaled by the intonational phrase boundary even before the next lexical input ("holding") was made available. These data suggest that intonational phrase boundaries by themselves may be sufficient for allowing listeners to predict upcoming syntactic structure.

Keywords: intonational phrase boundaries, attachment ambiguity, parsing, eye-tracking, prosody

1. Introduction

Intonational phrase boundaries, as perceptual breaks in the speech stream, correlate roughly with a change in F0, pre-boundary lengthening and pauses, in English (e.g., Shattuck-Hufnagel 1996; Wagner & Watson 2010). There has been a great deal of recent interest in understanding the role of intonational phrase boundaries in language processing. Work in this area has focused primarily on how intonational phrase boundaries interact with other boundaries in interpretation (Carlson, Clifton & Frazier 2001; Carlson, Clifton & Frazier 2009; Clifton, Carlson & Frazier 2002; Lee 2012), whether speakers disambiguate ambiguous structures for listeners by using intonational phrase boundaries (e.g.,

Kraljic & Brennan 2005; Schafer, Speer, Warren & White 2000; Snedeker & Trueswell 2003), and whether listeners use intonational phrase boundaries in resolving ambiguous structures (e.g., Kjelgaard & Speer 1999; Lee & Garnsey 2012; Marslen-Wilson, Tyler, Warren, Grenier & Lee 1992; Price, Ostendorf, Shattuck-Hufnagel & Fong 1991; Schafer 1997; Speer, Kjelgaard & Dobroth 1996). This study examines the incremental and anticipatory nature of the role intonational phrase boundaries play in on-line sentence comprehension.

The literature on spoken language processing suggests that listeners do not wait until all words in a sentence are encountered. As an utterance unfolds, they instead continuously make predictions about upcoming structure using a wide variety of information sources at the earliest possible moment.

Eye movements to objects in a visual display have been found to be a good indicator of interpretation in online spoken language processing. As eye movements are closely time-locked to referring expressions in the speech stream, fixations to objects in a visual display provide information about real-time speech processing. Previous eye movement studies have shown that verbs (Altmann & Kamide 1999), prenominal adjectives (Eberhard, Spivey-Knowlton Sedivy & Tanenhaus 1995; Sedivy, Tanenhaus, Chambers & Carlson 1999), prepositions (Chambers, Tanenhaus, Eberhard, Filip & Carlson 2002), order-of-mention (Kaiser & Trueswell 2004), referential contexts (Spivey, Tanenhaus, Eberhard & Sedivy 2002; Tanenhaus, Spivey-Knowlton, Eberhard & Sediby 1995) and pragmatic factors (Chambers et al. 2002; Chambers, Tanenhaus & Magnuson 2004) are all used immediately by listeners for generating predictions about upcoming material. For example, Altmann and Kamide (1999) find that listeners anticipate a referent in the visual display as a direct object at the point of the verb i.e., before the direct object is mentioned. On the trial where the sentence "the boy will eat the cake" was played, the listeners shifted their attention quickly to the only edible object 'cake' in the visual display before the word "cake" was encountered. Altmann and Kamide claim that the selectional information at the verb allows listeners to narrow down a set of possible referents in the visual display,

resulting in anticipatory eye-movements.

However, relatively less progress has been made in studying prosodic cues. There have been a handful of studies that have demonstrated the incremental and anticipatory nature of processing pitch accents (Arnold 2008; Dahan, Tanenhaus & Chambers 2002; Ito, Jincho, Minai, Yamane & Mazuka 2012; Ito, Bibyk, Wagner & Speer 2014; Ito & Speer 2008; Watson, Gunlogson & Tanenhaus 2008; Weber, Grice & Crocker 2006). For example, Ito and Speer (2008) show that a contrastive accent on an adjective (e.g., "First, hang the green drum → "Next, hang the ORANGE drum) allows listeners to anticipate a target referent (i.e., orange drum) even before they hear the head noun. Listeners use their knowledge about the relationship between pitch accenting and discourse structure to rapidly predict upcoming material. Although there is a consensus that intonational phrase boundaries are one of many cues listeners use in parallel to determine syntactic structure, however, little is known about whether intonational phrase boundaries *by themselves* can be used predictively in spoken language comprehension.

There are theoretical reasons for thinking that it is at least possible that intonational phrase boundaries by themselves could be used to predict upcoming syntactic structure. Studies on intonational phrase boundary production have shown that intonational phrase boundaries tend to occur at the end of syntactic constituents (e.g., Cooper & Paccia-Cooper 1980; Selkirk 1986; Truckenbrodt 1999; Watson & Gibson 2004). Selkirk (1986)'s edge-based mapping rule instantiates this relationship explicitly, stating that edges of intonational phrase boundaries should be aligned with the edges of syntactic constituents. Similarly, Watson and Gibson (2004)'s performance-based theory (the Left/Right Boundary Hypothesis) proposes that intonational phrase boundaries tend to be produced after a complete syntactic constituent. Based on this relationship between intonational phrase boundaries and syntax, the Anti-Attachment Hypothesis (Watson & Gibson 2005) argues that intonational phrase boundaries signal that upcoming information is less likely to be about the most recently processed lexical head by closing it for further attachment. Consider (1).

(1) Tap the frog with the flower.

In the sentence in (1), the prepositional phrase can either modify the preceding direct object noun or serve as an argument of the verb. According to the Anti-Attachment Hypothesis, an intonational phrase boundary after "tap" creates a bias towards a direct object modification reading (low attachment) because the boundary closes off the pre-boundary word "tap" and leads the prepositional phrase to attach to the other possible attachment head ("frog"). Similarly, an intonational phrase boundary after "frog" creates a bias towards interpreting the prepositional phrase as an argument of the verb (high attachment) because the intonational phrase boundary signals that "frog" is unlikely to receive further attachment. If intonational phrase boundaries provide cues to local syntactic structure, it is at least possible that intonational phrase boundaries by themselves are used predictively as soon as they are encountered, especially if the context (or visual display) is suitably constrained.

Although there is very little evidence that intonational phrase boundaries by themselves can generate predictions about upcoming syntactic structure at the point at which they are encountered, previous studies using on-line cross-modal naming tasks (e.g., Kjelgaard & Speer 1999; Marslen-Wilson et al. 1992; Speer et al. 1996; Warren et al. 1995) have shown that intonational phrase boundaries can have an early influence on syntactic decisions in language comprehension. In cross-modal naming tasks, participants are presented with an ambiguous auditory sentence fragment and are then visually presented with a word that disambiguates the sentence. The participant's task is to name the word. The time it takes to do the naming provides a measure of the extent to which the continuation was expected. In these studies, the presence of an intonational phrase boundary is manipulated so that it either conflicts with or is consistent with the syntactic structure associated with the disambiguating visual target. Consider sentence (2).

(2) Whenever the guard checks the door [is / it's] locked.

In sentence (2), the "door" can be interpreted either as a direct object

of the subordinate verb (late closure) or as a subject of the main clause (early closure). Speer et al. (1996) found that when intonational phrase boundaries coincided with clausal boundaries (i.e., after the verb in early closure sentences and after the noun in late closure sentences), naming of the target word ("it" or "it's") was faster compared to a prosodically neutral baseline. When intonational phrase boundaries conflicted with the eventual syntactic interpretation, naming times were slower. These types of data have been used to argue for a very early role for intonational phrase boundaries in parsing. However, because the effect of intonational phrase boundaries was measured at or after the disambiguation point, it is not known whether intonational phrase boundaries by themselves allowed listeners to anticipate a particular structure or whether these effects were due to an interaction between the presence of an intonational phrase boundary and processing the following lexical items.

An eye-tracking study by Snedeker and Trueswell (2003) examined how prosodic boundary information is used by listeners to predict attachment of upcoming prepositional phrases in sentences like (1) (e.g., "Tap the frog with the flower"). The authors found that listeners very rapidly use prosodic boundary information to determine the role of an upcoming prepositional phrase. Shortly after the onset of the direct object noun but before the prepositional phrase was heard, participants showed more fixations to the object with an attribute (e.g., a frog holding a flower) in the condition in which the prepositional phrase was intended to be a modifier than in the condition in which it was intended to be an instrument. An analysis of more fine-grained time windows revealed that the effect came about at around the same time as lexical information about the direct object noun became available. Lexically driven fixations are typically seen around 200ms after word onset (Allopenna, Magnuson & Tanenhaus 1998). Their results showed that the effects of intonational phrase boundaries were marginally reliable in the time window 300-400ms after the onset of the direct object noun and were reliable at the 400-500ms time window. Snedeker and Trueswell (2003) argue that the reason they were not able to observe the effect of the intonational pharse boundary until at the onset of the direct object noun was because the

display contained a distracter animal that could also be described with a modifier (e.g., a giraffe in pajamas). Thus, in this experiment, even with prosodic boundary information, listeners had to wait until the relevant lexical information was heard to identify the target referent (e.g., a frog with flowers). Thus, like the cross-modal studies discussed above, this design cannot tell us whether an intonational phrase boundary was interpreted immediately at the point of the boundary or whether its interpretation was delayed until after the next lexical word was heard.

The current study examines whether intonational phrase boundaries by themselves can be used for predicting upcoming syntactic structure. This question was tested by creating a context that made it possible for listeners to anticipate the target referents based on their interpretation of an intonational phrase boundary, even before the presentation of a post-nominal modifier. In the experiment, participants were presented with auditory instructions along with a corresponding visual display on a computer screen. Each visual display included two boxes each containing a pair of pictures like those shown in Figure 1. Critical instructions had the form of example (3).

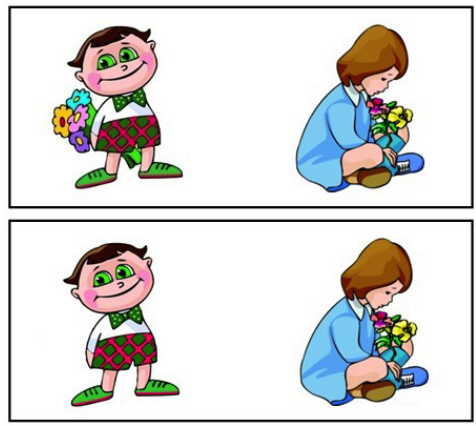

Figure 1. Example visual display.

(3) Click on the boy and the girl holding flowers.

Participants were tasked with clicking one of the two boxes on the screen

that contained the picture pairs. Sentence (3) is ambiguous because the modifier "holding flowers" can attach low to the second noun phrase (i.e., the girl), or high to the entire conjoined noun phrase (i.e., the boy and the girl). Thus, for the low attachment interpretation, a correct response would be for the listener to click on the box in which the girl is holding flowers but the boy is holding nothing. For the high attachment interpretation, the correct response would be for the listener to click on the box in which both the boy and the girl are holding flowers. An intonational phrase boundary before the modifier should create a bias towards a high attachment modification (e.g., Clifton et al., 2002; Price et al., 1991; Watson & Gibson 2005). The presence of an intonational phrase boundary in this position was explicitly manipulated.

When hearing an instruction like (3), it is predicted that upon hearing "Click on the boy", listeners would alternate fixations on the two referents of the first noun (boys) until hearing "and the girl", at which point, they would alternate fixations on the two referents of the second noun (girls). It is during the second noun that acoustic cues for an intonational phrase boundary become available and that participants could potentially anticipate the correct targets from prosodic boundary information. As described above, the ambiguity lies in whether the upcoming description is associated with the referents of both nouns or with only the referent of the second noun. Thus, looks to the referents of the first noun (the boy with flowers) should reflect the listeners' expectation of how the upcoming ambiguous phrase is to be structured. As discussed earlier, understanding whether listeners use intonational phrase boundaries to predict upcoming syntactic structure requires measures that occur at the intonational phrase boundary and before the word that follows it (i.e., holding). If intonational phrase boundaries can be used predictively as a cue, then participant responses should be seen before the post-boundary word is encountered. Thus, a critical measure here is whether the effect of intonational phrase boundaries can be seen at the point of the boundary and before the next word ("holding") is encountered.

2. Method

2.1. Participants

Thirty-six undergraduate students from the University of Illinois at Urbana-Champaign participated in the study for course credit (mean age: 19, 18 males). All participants were native speakers of American English. They had normal or corrected-to-normal vision and no reported hearing impairment.

2.2. Procedure

Participants were seated in front of a computer display wearing a head mounted eye tracker (SR Research EyeLink 2). In the experiment, auditory instructions were presented to participants over speakers simultaneously with a corresponding visual display. The participants' task was to click on pictures in the display. Participants' fixations were recorded from the onset of the auditory instruction until the selection of picture. Two practice items were provided at the beginning of the experimental session.

2.3. Materials

There were sixteen critical sentences (see Appendix A for a list of critical sentences). Each sentence comprised a conjoined noun phrase followed by a modifier as in (3). The two nouns in the conjoined noun phrase were linked to each other by the conjunction "and". Each sentence was produced by a trained female native speaker of American English in two different conditions as shown in (4) below (double slashes indicate the location of the intonational phrase boundary). Intonational phrase boundaries were produced with an L-L% boundary tone in the ToBI (Tones and Break Indices) coding scheme (Beckman & Elam 1997).

(4) a. No Boundary:
Click on the boy and the girl holding flowers.
b. Boundary:
Click on the boy and the girl // holding flowers.

In order to prevent unintended acoustic differences in the region before the word of interest from confounding the results, the words preceding the second noun (e.g., Click on the boy and the) in the Boundary condition were replaced with the corresponding words in the No Boundary condition. This manipulation ensures that the effects reported below are due to acoustic differences in the second noun (e.g., girl), but not by those in the words preceding it (i.e., Click on the boy and the). The F0 plots for the sentences in (4) are displayed in Figure 2 below.

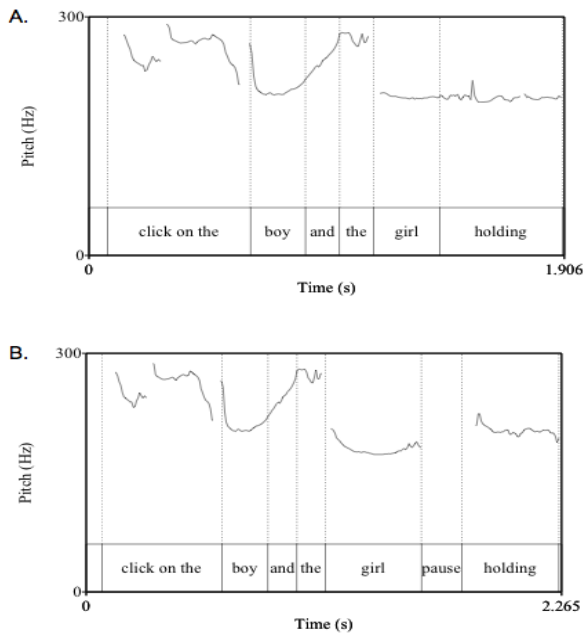

Figure 2. Example F0 plots.

In order to ensure that the speaker produced intonational phrase boundaries as intended, the acoustic properties of the second noun (i.e. girl) were measured, both when the intonational phrase boundary was present and when it was not. Intonational phrase boundaries are typically signaled through lengthening of the pre-boundary word followed by a pause. They also typically include a change in F0 on the pre-boundary word (in this case a drop) (e.g., Shattuck-Hufnagel & Turk, 1996; Wagner & Watson 2010). Consequently, the duration of the noun

was expected to be longer when an intonational phrase boundary was present compared to when it was not. The pre-boundary word was also expected to undergo a drop in F0, compared to its counterpart in the No Boundary condition. Table 1 presents the mean durations of and the mean values of maximum F0 and minimum F0 at the critical noun (i.e., girl). The mean duration of the critical noun was longer when it was preceded by an intonational phrase boundary than when it was not ($t(15) = 13.8$, $p<.01$). There was no reliable difference in the maximum pitch between the No Boundary and the Boundary conditions ($t(15) <1$), but the minimum pitch was reliably lower in the Boundary condition than in the No Boundary condition ($t(15) = 7.3$, $p<.01$). In the Boundary condition, the mean duration of the pause following the pre-boundary word was approximately 200ms. Thus, these acoustic data confirmed the distinction between the Boundary and the No Boundary conditions.

Table 1. Mean Durations of and Mean Pitch Values at the Critical Noun

	No Boundary	Boundary
Duration (ms)	386 (24.3)	580 (28.1)
Max F0 (Hertz)	211 (4.1)	212 (2.5)
Min F0 (Hertz)	170 (1.8)	189 (1.5)

Note. Standard errors are presented in parentheses.

Critical items were rotated through two experimental conditions using a Latin Square design, resulting in two different experimental lists. Each list contained thirty distracter trials in addition to sixteen critical trials. Seven distracter trials included conjoined noun phrases without any post-nominal modifiers (e.g., Click on the ballerina and the diver). In another seven distracters, only the first noun was modified (e.g., Click on the woman holding a cup and the man). There were sixteen distracters that consisted of the same structure as that of critical trials (e.g., Click on the dog and the cow wearing earrings). These trials differed from critical trials in that they were disambiguated by the visual display: the display supported only one interpretation (8 trials: high attachment, 8 trials: low attachment). Half of the distracters were produced with intonational phrase boundaries.

3. Results

3.1. Selection data

There were overall high attachment preferences. Participants clicked on the high attachment pair 85% of the time on average. There were more high attachment responses in the No Boundary condition (87%) than in the Late Boundary condition (83%). A mixed-logit model analysis with Subject and Item as random intercept terms, however, revealed that this difference failed to reach significance (ß = -0.6, z = 1.8, p = .08).

3.2. Fixation data

A challenge in using fixations to investigate the online processing of intonational phrase boundaries is the inherent confound between boundaries and timing. In the condition in which an intonational phrase boundary is present, the pre-boundary word (e.g., girl) is lengthened and followed by a pause. At first blush, aligning fixations at the beginning of the pre-boundary word appears attractive because this is when cues to the presence or absence of an intonational phrase boundary begin to become apparent to a listener. However, the lengthening of the pre-boundary word in addition to the pause that follows it makes it difficult to determine whether differences across conditions are due to the presence of the intonational phrase boundary or simply due to differences in the timing of when listeners hear the next word.

In order to avoid this problem, fixations were aligned at the onset of the word that immediately followed the intonational phrase boundary (e.g., the onset of "holding"). Analyses were conducted on fixations that occur during the region where listeners hear prosodic boundary information, i.e., the negative time regions. These fixations reflect anticipatory effects of intonational phrase boundaries towards the end of the pre-boundary words. Ito and Speer (2008) used the same technique to examine anticipatory effects of contrastive accents.

Figure 3A presents the time course of a proportion of fixations to the modified boy out of the fixations to both boys, for each condition. From

Figure 3A, it is clear that consistent with the selection data, there was an overall preference to look at the modified boy across conditions (a mean proportion is roughly higher than 0.5). The preference for fixating the modified boy over the unmodified boy was greater in the Boundary condition than in the No Boundary condition. Crucially, this difference came about well before the alignment point (i.e., 0ms, the onset of "holding"), suggesting that there were anticipatory effects of intonational phrase boundaries towards the end of the pre-boundary word.

To this end, separate by-subjects and by-items growth curve analyses (Mirman, Dixon & Magnuson, 2008) were conducted for the time window -375-200ms. 200ms was chosen as an endpoint because it takes roughly 200ms to program an eye-movement (Allopenna et al. 1998). The -375ms time point was chosen because the mean duration of the second noun (e.g., "girl") in the condition with no boundaries was approximately 375ms. Thus, the analyses did not include any fixations that were initiated by the next lexical word (i.e., holding). If there are any reliable differences in the pattern of fixations between the Boundary and the No Boundary conditions in this particular region, this should suggest that listeners use boundary information (i.e., a change in FO, a lengthening of "girl", and pausing) for predicting upcoming syntactic structure as soon as it is made available.

The fixation curve was modeled by a third order orthogonal polynomial, which allowed for the examination of the effect of different parameters including the intercept, linear, quadratic, and cubic, on the curvature of the fixation curve. Differences between the No Boundary and the Boundary conditions were captured by differences in each of the parameters of the curve. The effect on the intercept term captures a difference in overall proportion of fixations to the modified boy. The linear term reflects a difference in fixations over time. The quadratic and cubic terms capture a difference between the rise and fall of the fixation curve between conditions. The improvement in model fit after adding each parameter was evaluated using a deviance statistic (i.e., minus 2 times the log-likelihood).

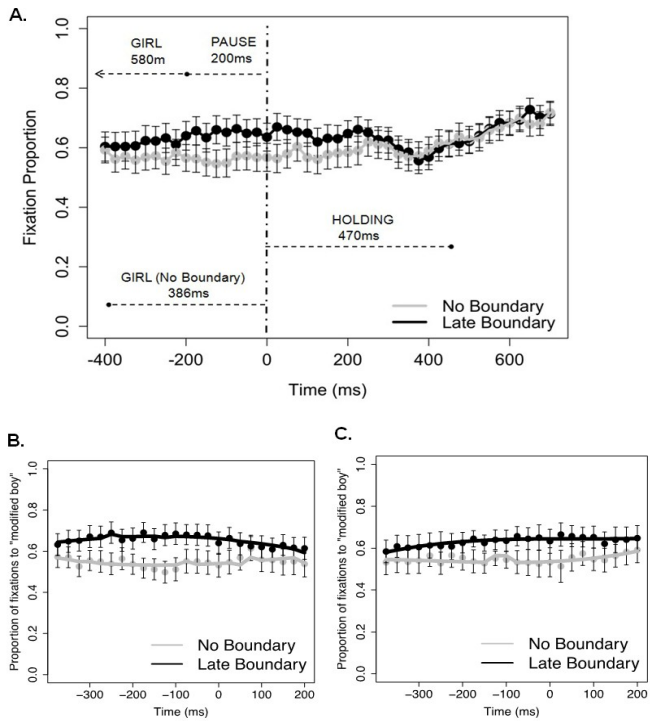

Figure 3. Panel A plots mean proportions of fixations to the modified boy out of the fixations to both boys at a given time point. Samples were taken every 25ms. The utterances were synchronized at the onset of "holding". Dotted lines in panel A represent the mean durations of the words indicated above the line. Panel B and panel C display fixations over -375-200ms, our analysis region, with solid lines representing the fits of the by-subjects growth curve analysis (panel B) and the by-items growth curve analysis (panel C). In all figures, error bars represent standard error.

Growth curve analyses performed over -375-200ms revealed a reliable effect of intonational phrase boundaries on the intercept term (by-subjects: ß = 0.1, t (35) = 3.2, p<.01; by-items: ß = 0.1, t (15) = 2.2, p<.05), indicating that the presence of an intonational phrase boundary led to more fixations to the modified boy than its absence. The effect of intonational phrase boundaries on the quadratic term was also reliable (by-subjects: ß = -0.2, t (1498) = 4.3, p<.01; by-items: ß = -0.1, t (684) = 3.3, p<.01), reflecting that an increase in fixations to the modified boy occurred earlier in the Boundary condition than in the No Boundary condition (See Table 2 for the results of model comparisons).

Table 2. Mean Results of Model Comparisons in Growth Curve Analyses on Fixations to the Modified Boy out of Fixations to both Boys in the 375-200ms Window.

by-subjects			
Term	-2LL	Δ*D*	*p*
Base	1409.2		
Intercept	1419.2	10.0	**<.01**
Linear	1419.4	0.2	>.1
Quadratic	1439.1	19.7	**<.01**
Cubic	1442.1	3.0	<.09
by-items			
Term	-2LL	Δ*D*	*p*
Base	1441.2		
Intercept	1445.8	4.6	**<.05**
Linear	1445.9	0.1	>.1
Quadratic	1457.9	12.0	**<.01**
Cubic	1460.5	2.6	>.1

Taken together, there was a reliable difference between the Boundary and the No Boundary conditions in the pattern of fixations to a referent that was directly linked to syntactic attachment decisions. Critically, the effect was observed even before the next word was heard. This provides evidence that listeners may predict upcoming syntactic structure based on a signal that an intonational phrase boundary provides about syntactic structure, as soon as the boundary is encountered.

4. Discussion

In the present study, a highly constrained visual display was used to determine whether an intonational phrase boundary preceding an ambiguous modifier creates an immediate bias towards high attachment even before the modifier is encountered. The results showed that there were greater preferences to fixate at the modified boy as a function of an intonational phrase boundary. Crucially, this effect occurred in the temporal regions in which eye movements responding to the next lexical input

had not yet been initiated. The data suggest that like other sources of information (e.g., Altmann & Kamide 1999, Dahan et al. 2002; Eberhard et al. 1995; Ito & Speer 2008; Sedivy et al. 1999; Tanenhause et al. 1995), intonational phrase boundaries are used predictively at the earliest possible moment in on-line spoken language comprehension.

The Anti-Attachment Hypothesis (Watson & Gibson 2005) proposes that intonational phrase boundaries play a role in syntactic processing by signaling closure of an immediately preceding lexical head. In the current study, the presence of an intonational phrase boundary signaled that the preceding head (e.g., girl) was an unlikely site for further attachment, leading listeners to expect the modifier to attach to the other possible attachment site (i.e., the entire conjoined NP). However, note that in the type of construction used here, the low attachment site (e.g., the girl) is embedded within the high attachment site (i.e., the entire conjoined NP). This raises a question of how upcoming material is allowed to attach to both nouns when an intonational phrase boundary closes the second noun for further attachment. Given that the effect of the late boundary (i.e., the intonational phrase boundary before an ambiguous phrase) was comparable to that observed in other ambiguous structures, the Anti-Attachment Hypothesis would need to assume that the two attachment sites are simply two independent attachment alternatives. An intonational phrase boundary after the second noun (e.g., girl) is likely to close the possibility that an upcoming modifier attaches to that noun. However, the other attachment alternative is available for further attachment, which consequently leads an upcoming phrase to modify the second noun as part of the conjoined NP.

In the current study, we found no reliable effects of the intonational phrase boundary in the off-line selection data even though there were effects of the intonational phrase boundary in the on-line fixation data. The difference was numerically in the opposite direction, although it did not reach significance. The discrepancy between the on-line and offline measures may be attributable to the fact that while on-line processing is driven by just information that intonational phrase boundaries provide about local syntactic structure, off-line decisions may be modulated not just by boundary information, but also by other factors including struc-

tural biases, meta-linguistic consideration, and lexical information available later in a sentence. Given that there was an overall preference for high attachment in the current data, one possibility is that listeners may have weighted an intonational phrase boundary before the post-nominal modifier less strongly because it created a bias that matched overall attachment preferences. The discrepancy might have been because listeners were affected by default attachment preferences to a greater extent during post-interpretive processes than during on-line processing. The dissociation between the on-line and off-line data, of course, could be just due to chance, but if there is a real discrepancy between off-line and on-line use of intonational phrase boundaries, future work will need to unpack the differences.

Furthermore, the paradigm introduced in this paper opens up the possibility of examining intonational phrase boundaries at the moments at which they are heard. In future research, this paradigm can be extended to examine what acoustic aspects of intonational phrase boundaries drive the predictive effects found here, and whether the same effect may generalize across different types of prosodic boundaries.

Although it has been known for quite some time that intonational phrase boundaries affect syntactic parsing (e.g., Shattuck-Hufnagel & Turk 1996; Wagner & Watson 2010), this paper provides a step towards understanding the incremental nature of the role intonational phrase boundaries play in parsing.

Appendix A. Experimental Stimuli

1. Click on the boy and the girl holding flowers.
2. Click on the girl and the boy holding balloons.
3. Click on the man and the boy holding books.
4. Click on the woman and the boy holding lollipops.
5. Click on the policeman and the doctor holding pom-poms.
6. Click on the nurse and the waitress holding bells.
7. Click on the man and the woman holding bags.
8. Click on the tourist and the soldier holding candles.
9. Click on the elephant and the bug holding weights.
10. Click on the cat and the bird holding binoculars.
11. Click on the camel and the hippo carrying gifts.
12. Click on the bear and the turtle wearing sunglasses.
13. Click on the pig and the lion wearing headphones.
14. Click on the dog and the cat wearing ribbons.
15. Click on the parrot and the shark wearing gloves.
16. Click on the duck and the kangaroo wearing shoes.

References

Allopenna, P., Magnuson, J., & Tanenhaus, M. (1998). Tracking the time course of spoken word recognition using eye movements: Evidence for continuous mapping models. *Journal of Memory and Language*, 38, 419-439.

Altman, G., & Kamide, Y. (1999). Incremental interpretation at verbs: Restricting the domain of subsequent reference. *Cognition*, 73, 247-264.

Arnold, J. (2008). THE BACON not the bacon: how children and adults understand accented and unaccented noun phrase. *Cognition*, 208, 69-99.

Beckman, M. E., & Elam, G. A. (1997). Guidelines for ToBI labelling, vers 3.0 [manuscript]: Ohio State University.

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

Carlson, K., Clifton, C., & Frazier, L. (2009). Nonlocal effects of prosodic boundaries. *Memory & Cognition*, 37, 1014-1025.

Chambers, C. G., Tanenhaus, M. K., Eberhard, K. M., Filip, H., & Carlson, G. N. (2002). Circumscribing referential domains during real-time language comprehension, *Journal of Memory and Language,* 47 30-49.

Chambers, C. G., Tanenhaus, M. K., & Magnuson, J. S. (2004). Actions and affordances in syntactic ambiguity resolution. Journal of Experimental Psychology: Learning. *Memory & Cognition*, 30, 687-696.

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

Cooper, W. E., & Paccia-Cooper, J. (1980). Syntax and Speech. Cambridge, Mass: Harvard University Press.

Dahan, D., Tanenhaus, M. K., & Chambers, C. G. (2002). Accent and reference resolution in spoken language comprehension. *Journal of Memory and Language*, 47, 292-314.

Eberhard, K. M., Spivey-Knowlton, M., Sedivy, J., & Tanenhaus, M. K. (1995). Eye movements as a window into real time spoken language comprehension in natural contexts. *Journal of Psycholinguistic Research*, 24, 409-436.

Frazier, L., & Clifton, C. (1998). Sentence reanalysis and visibility. In J. D. Fodor & F. Ferreira (Eds.), Reanalysis in Sentence Processing. Dordrecht: Kluwer.

Ito, K., Bibyk, S. A., Wagner, L, & Speer, S. (2014). Interpretation of contrastive pitch accent in six-to eleven-year-old English-speaking children (and adults). *Journal of Child Language*, 41, 84-110.

Ito, K., Jincho, N., Minai, U., Yamane, N., & Mazuka, R. (2012). Intonation facilitates contrast resolution: evidence from Japanese adults and 6-year olds. *Journal of Memory and Language*, 66, 265-284.

Ito, K., & Speer, S. (2008). Anticipatory effects of intonation: Eye movements during instructed visual search. *Journal of Memory and Language*, 58, 541-573.

Kaiser, E., & Trueswell, J. (2004). The role of discourse context in the processing of a flexible word-order language. *Cognition*, 94, 113-147.

Kjelgaard, M., & Speer, S. (1999). Prosodic facilitation and interference in the resolution of temporary syntactic closure ambiguity. *Journal of Memory and Language*, 4, 153-194.

Kraljic, T., & Brennan, S. (2005). Prosodic disambiguation of syntactic structure: For the speaker or for the addressee? *Cognitive Psychology*, 50, 194-231.

Marslen-Wilson, W. D., Tyler, L., Warren, P., Grenier P., & Lee, C. (1992). Prosodic effects in minimal attachment. *The Quarterly Journal of Experimental Psychology*, 45A, 73-87.

Lee, E.-K. (2012). Local and global effects of prosodic boundaries: Evidence from eye movements. *Language Research,* 48, 349-363.

Lee, E.-K., & Garnsey, S. M. (2012). Do contrastive accents modulate the effect of intonational phrase boundaries in parsing? *Lingua*, 122, 1763-1775.

Mirman, D., Dixon, J. A., & Magnuson, J. S. (2008). Statistical and computational models of the visual world paradigm: Growth curves and individual differences. J*ournal of Memory and Language*, 59, 475-494.

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

Schafer, A. (1997) Prosodic parsing: the role of prosody in sentence comprehension. PhD Dissertation, Univ. of Massachusetts, Amherst, MA: GLSA.

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

Shattuck-Hufnagel, S., & Turk, A. E. (1996). A prosody tutorial for investigators of auditory sentence processing. *Journal of Psycholinguistic Research*, 25, 193-247.

Sedivy, J. C., Tanenhaus, M. K., Chambers, C. G., & Carlson, G. N. (1999). Achieving incremental semantic interpretation through contextual representation. *Cognition*, 71, 109-147.

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

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

Speer, S., Kjelgaard, M., & Dobroth, K. (1996). The influence of prosodic structure on the resolution of temporary syntactic closure ambiguities. *Journal of Psycholinguistic Research*, 25, 247-268.

Spivey, M., Tanenhaus, M. K., Eberhard, K. M., & Sedivy, J. C. (2002). Eye movements and spoken language comprehension: Effects of visual context

on syntactic ambiguity resolution. *Cognitive Psychology*, 45, 447-481.
Steinhauer, K., Alter, K., & Friederici, A. D. (1999). Brain potentials indicate immediate use of prosodic cues in natural speech processing. *Nature Neuroscience,* 2, 191-195.
Tanenhaus, M. K., Spivey-Knowlton, M. J., Eberhard, K. M., & Sedivy, J. C., (1995). Integration of visual and linguistic information in spoken language comprehension. *Science*, 268, 1632-1634.
Truckenbrodt, H. (1999). On the relation between syntactic phrases and phonological phrases. *Linguistic Inquiry*, 30, 219-255.
Wagner, M., & Watson, D. G. (2010). Experimental and theoretical advances in prosody: A review. *Language and Cognitive Processes*, 25, 905-945.
Watson, D. G., & Gibson, E. (2004). The relationship between intonational phrasing and linguistic structure in language production. *Language and Cognitive Processes*, 19, 713-755.
Watson, D. G., & Gibson, E. (2005). Intonational phrasing and constituency in language production and comprehension. *Studia Linguistica*, 59, 279-300.
Watson, D. G., Gunlogson, C. A., & Tanenhaus, M. K. (2006). Online methods for the investigation of prosody. In S. Sudhoff, D. Lenertov, R. Meyer, S. Pappert, P. Augurzky, I. Mleinek, N. Richter, J. Schlieer (Eds.), Methods in Empirical Prosody Research, 259-282. Walter de Gruyter: New York.
Watson, D. G., Gunlogson, C. A., & Tanenhaus, M. K. (2008). Interpreting pitch accents in on-line comprehension: H* vs. L+H*. *Cognitive Science*, 32, 1232-1244.
Weber, A., Grice, M., & Crocker, M. (2006). The role of prosody in the interpretation of structural ambiguities: A study of anticipatory eye movements. *Cognition*, 99, B63-B72.

Eun-Kyung Lee
Department of English Language and Literature
Yonsei University
50 Yonsei-ro, Seodaemun-gu, Seoul 120-749, Korea
E-mail: eunkyunglee@yonsei.ac.kr

Received: June 29, 2014
Revised version received: July 26, 2014
Accepted: July 29, 2014