Prosodic effects in parsing early vs. late closure sentences by second language learners and native speakers

Hyekyung Hwang & Amy J. Schafer

Department of Linguistics
University of Hawai'i, Hawai'i
{hyekyung; aschafer}@hawaii.edu

Abstract

The Informative Boundary Hypothesis (IBH: [4]) claims that a prosodic boundary is interpreted relative to preceding boundaries. This study tests predictions of the IBH with Korean learners of English (L2ers) and English native speakers (L1ers) in a prosody experiment on the resolution of an Early vs. Late Closure ambiguity in spoken English sentences. A control experiment assessed and controlled for English morpho-syntactic knowledge in the main experiment. The main experiment presented the syntactically ambiguous portion of sentences in a forced-choice continuation-selection task. The results showed that 1) Korean L2ers at all levels used relative boundary size to disambiguate sentences, like L1ers; 2) intonation phrase boundaries provided stronger evidence for syntactic boundaries than intermediate phrase boundaries, especially for the L2ers; and 3) the IBH's 3-way categorization of relative boundary size – larger/same-size/smaller – appears insufficient for this syntactic structure.

1. Introduction

Many studies have discussed how prosodic boundaries can disambiguate syntactic ambiguities in spoken language. Some consider prosodic phrasing effects to be "domain" effects, with the processor interpreting prosodic boundaries not in isolation, but in relation to the global prosodic contour of an utterance e.g., [3], [4], [8], [9]. Others consider them to be purely local "edge" effects, with the processor interpreting the presence of a large prosodic boundary as a signal not to attach the incoming constituent to the just-encountered syntactic constituent e.g., [10].

This study focuses on effects of the relative size of prosodic boundaries on the resolution of a type of closure syntactic ambiguity in second language (L2) processing, and tests an extension of the Informative Boundary Hypothesis (IBH) – a position in support of "domain" effects – put forward by Clifton et al. [4]. The IBH predicts that for ambiguity as in (1) the effectiveness of a local prosodic boundary and its interpretation are determined by its size relative to the prosodic boundary at earlier relevant positions in the utterance, i.e. whether it is larger, the same size, or smaller than would be expected on the basis of the size of the previous boundaries.

(1) [Susie learned] A [that Bill called] B [after John visited].

In Clifton et al.'s example (1), if the prosodic boundary at position (B), before the ambiguously attached adverbial clause, is larger than a prosodic boundary at position (A), the adverbial clause modifies the higher clause. If a prosodic boundary at (B) is smaller than a boundary at (A), the adverbial clause modifies the lower clause. If two prosodic boundaries of the same size are placed at (A) and (B), attachment of the adverbial clause is determined by other factors. Underlying these predictions is an assumption that prosodic boundaries group lexical material into packages, which determine the preferred interpretation; it is this packaging that is informative.

In the current study, the materials contained a subordinate clause with a verb (*moves*) that can be intransitive or transitive, resulting in a temporary syntactic ambiguity between Early and Late Closure of the clause, as in (2).

(2) EC: **When that moves the square** will ...
LC: **When that moves the square** it'll ...

The beginning of the sentence through *square* can continue as either Early Closure (closing the subordinate clause early at the intransitive verb *moves*) or Late Closure (closing the subordinate clause late at the direct object *the square*). Generally, the Late Closure structure is preferred, which Frazier [5] explained in terms of the Late Closure principle: attach an incoming component to the phrase currently being processed, if grammatically permissible**.** In this set of materials, the IBH's informative packaging assumption leads to the following predictions.

(3) [When that moves] A [the square] B [will … /it'll …]

In (3), any prosodic boundary larger than a word boundary at position (A) should end the first package, at least preliminarily. However, a larger boundary at (B) extends the first package and groups *the square* with the subordinate verb *moves*. Therefore, any prosody with a larger boundary at (B) than at (A) cooperates with the Late Closure interpretation. In contrast, if the boundary at (B) is smaller than the one at (A), *the square* is packaged with the matrix verb *will*, instead of the subordinate verb, supporting an Early Closure interpretation. Finally, if the boundaries at (A) and (B) are the same size, the interpretation should be determined by other factors.

This study investigated how Korean L2ers, as compared with English L1ers, used prosodic cues such as the relative size of prosodic boundaries in the comprehension of spoken English sentences involving temporary Early/Late Closure ambiguities. In our discussion, we will report the results of a morpho-syntax control experiment and a prosody experiment and show that the relative size of prosodic boundaries was critically used to recover the intended interpretation from syntactically ambiguous utterances. The next section of the paper briefly discusses and compares intonation models for English and Korean. The following section turns to the presentation of the two experiments and our conclusions.

2. Intonation model

This study assumes the intonation models developed by Pierrehumbert and Beckman [1], [7] for English and by Jun [6]

for Korean. Both languages employ two levels of prosodic boundaries. At the highest level, each utterance is made up of one or more intonation phrases (IPhs). IPh boundaries have similar phonetic and phonological properties in Korean and English: a boundary tone following the IPh (e.g., H% or L%), lengthening of the final syllable of the IPh, a following silent period, and pitch reset. An IPh consists of one or more intermediate phrases (ips) in English and one or more accentual phrases (APs) in Korean. An ip is marked by one of three phrase tones (H-, L-, or !H-) and final lengthening. It must contain at least one pitch accent, which occurs on the stressed syllable of a prominent word. On the other hand, an AP is typically defined by a final rising tonal pattern (e.g., Ha) and its final syllable is not typically lengthened. It must contain one or more phonological words. Thus, English ips are more tonally variable and carry stronger durational marking of their boundaries than Korean APs.

3. Morpho-syntax experiment

3.1. Method

The morpho-syntax experiment employed a forced-choice continuation-selection task, and was run immediately following the prosody experiment (described below) on each subject. Its purpose was to control interference from L1 (Korean) morpho-syntactic features on L2 (English) morpho-syntactic decisions in the prosody experiment, and also to verify the subjects' proficiency levels. Subjects listened to the initial portion of sentences twice over headphones and chose between two endings visually presented on the left and right sides of a computer screen. The items were randomly presented and the correct choice was balanced for left and right sides of the screen within subjects.

3.1.1. Subjects

One hundred seventy six native Korean L2ers at four levels of English proficiency (40 advanced with immersion experience ("extra-advanced"), 40 advanced without immersion experience ("advanced"), 42 intermediate, 54 beginning) and 20 English L1ers took part in the experiment. The Korean L2ers were recruited in Seoul and the English L1ers were recruited at the University of Hawai'i.

3.1.2. Materials

Twenty five trials for "subject-drop" and twenty trials for "topicalization" were recorded with cooperating prosody in which an IPh occurred at subordinate clause boundaries. These utterances were digitally truncated at the offset of the final segment of a critical word. In order to effectively relate the results of this experiment to those of the prosody experiment, most items contained verbs of movement (e.g., *arrive, move,* or *push*) for the subordinate clause verb, a demonstrative pronoun (e.g., *that*) for the subordinate clause subject, and the third person nominative pronouns (e.g., *he, she,* or *they*) for the main clause subjects as in (4), (5), and (6).

(4) Subject-drop
auditory stimuli — *visual stimuli*
When that arrives (IPh) — will encounter a cookie.
he will encounter a cookie.

(5) Topicalization I (Correct use of topicalization)
When that *arrives* (IPh) she will encounter a cookie.
he will encounter a cookie.

(6) Topicalization II (Topic. preference)
When that *arrives* (IPh) Tom will encounter a cookie.
he will encounter a cookie.

3.2. Results

A Tukey's HSD test showed no significant differences among English L1ers, the extra-advanced group, and the advanced group, but significant differences in other pairs of groups (all p's < 0.01). The results suggest significant improvement in English morpho-syntax as proficiency increased from the beginning to the advanced levels. Moreover, these results verified the proficiency grouping.

Table 1: *Mean percentages of subject-drop, incorrect topcialization and topic. preference as a function of English proficiency levels.*
** L1: English L1ers; L2-1: beginning L2ers; L2-2: intermediate; L2-3: advanced; L2-4: extra-advanced*

%	L1	L2-4	L2-3	L2-2	L2-1
Occurrence of Subj. drop	4	2	6	27	41
Incorrect use of topicalization	3	3	8	18	34
Preference for topicalization	3	5	10	19	39

Since extra-advanced and advanced L2ers had native-like performance, their results in the prosody experiment should reflect their ability to use prosody for syntactic disambiguation. In contrast, the intermediate and beginning L2ers exhibited interference of Korean morpho-syntactic features. Thus, the interpretation of their results in the prosody experiment is limited by their morpho-syntactic competence.

4. Prosody experiment

4.1. Method

The prosody experiment used a forced-choice continuation-selection task in which ambiguous fragments of 53 Early and 53 Late Closure utterances of (7) were presented over headphones to the same subjects as in the morpho-syntax experiment. On each trial, subjects heard each fragment twice, and then chose between the Early or Late Closure continuations shown in (7). The Early Closure continuations were visually presented on the left of the computer screen and the Late Closure ones, on the right. Presentation order of the items was randomized across subjects.

(7) Auditory stimuli — Visual stimuli
(Lexically ambiguous portion) — *(Disambiguating continuation)*
EC: When that moves the square — will encounter a cookie.
LC: When that moves the square — it'll encounter a cookie.

Since both English and Korean employ two levels of prosodic boundaries, most L2ers were predicted to utilize relative boundary size and resolve the ambiguity at a better-than-chance level. In addition, more proficient L2ers

were predicted to perform better than less proficient ones, from their better knowledge of English morpho-syntax and a postulated better knowledge of prosodic differences between the languages.

4.1.1. Materials

The experiment differed from many listening comprehension experiments in that the materials were more naturally produced than those typically employed, yet limited to numerous pronunciations of one item. They consisted of quasi-spontaneous speech collected in Schafer et al. [9]'s game task. Thirteen native English speakers in their study collectively produced 53 Early and 53 Late Closure utterances of (2), with prosody that varied in boundary sizes (among other prosodic features) in the ambiguous region. Schafer et al. truncated the sentences at the offset of the ambiguously attached NP *square,* analyzed them in the ToBI system [2], and categorized the transcribed set of materials into three types of prosodic patterns, defined by comparing the boundary sizes surrounding the ambiguous NP. The majority of tokens received cooperating prosody (N=77), in which the largest prosodic boundary was placed at the subordinate clause boundary (i.e., at *moves* for Early Closure and at *square* for Late Closure). Nineteen tokens were transcribed with equally sized prosodic boundaries at both *moves* and *square* (hereafter, "same-sized prosody"). Ten tokens were categorized as conflicting prosody, with a smaller prosodic boundary at the subordinate clause boundary than the other critical position. Then, each of three prosodic patterns was sub-categorized into the more fine-grained 9-way distinction in Table 2.

Table 2: *Boundary size patterns in Early Closure.*

3-way distinction	9-way distinction	
	When that moves *(#1)* the square *(#2)* will…	
Cooperating	IPh	wd (word)
	IPh	ip
	ip	wd
Same-size	IPh	IPh
	ip	ip
	wd	wd
Conflicting	wd	IPh
	ip	IPh
	wd	ip

4.2. Results

4.2.1. Global analysis

First, a set of one sample t-tests showed that each group performed above chance (all p's < 0.01). These results suggest that prosody contributed to the L2ers' processing of syntactic ambiguity regardless of proficiency. Next, the percentages of correct responses from each subject were subjected to a repeated measures ANOVA with Syntax (Early vs. Late Closure) as a within-subjects factor and Group (English L1ers vs. extra-advanced vs. advanced vs. intermediate vs. beginning L2 groups) as a between-subjects factor.

The analysis revealed a significant main effect of Group ($F(4, 191) = 25.566$, $p < 0.01$), resulting from increase in correct categorization as proficiency increased. Also, a significant interaction of Syntax with Group was found ($F(4, 191) = 4.091$, $p < 0.01$), which indicates that the two closure structures were treated differently by the five groups of the subjects. As shown in Table 3, L2ers had a stronger bias for Late Closure than English L1ers.

Table 3: *Mean of correct categorization: comparison by English proficiency, prosody, and closure type. * L1: English L1ers; L2-1: beginning L2ers; L2-2: intermediate; L2-3: advanced; L2-4: extra-advanced*

Early Closure % Correct		Ns. of tokens	L1	L2 -4	L2 -3	L2 -2	L2 -1
Coop-erating	IPh-wd	29	92	87	84	74	64
	IPh-ip	7	84	79	73	72	58
	ip-wd	4	76	45	39	43	41
	sub total	40	89	81	77	70	61
Same-size	IPh-IPh	8	76	71	64	62	55
	ip-ip	2	53	56	48	48	44
	sub total	10	71	68	61	59	53
Conf-licting	ip-IPh	2	38	64	45	42	42
	wd-ip	1	80	60	48	55	46
	sub total	3	52	63	46	46	43
Total		53	83	78	72	67	58
Late Closure % Correct		Ns. of tokens	L1	L2 -4	L2 -3	L2 -2	L2 -1
Coop-erating	wd-IPh	9	89	87	88	84	66
	ip-IPh	6	69	77	75	65	57
	wd-ip	22	83	91	88	78	64
	sub total	37	82	88	86	77	63
Same-size	IPh-IPh	3	52	63	54	38	44
	ip-ip	3	58	60	50	62	62
	wd-wd	3	68	75	68	74	64
	sub total	9	59	66	58	58	57
Conf-licting	IPh-ip	4	48	56	54	39	50
	IPh-wd	1	35	50	55	67	52
	ip-wd	2	73	75	69	78	64
	sub total	7	53	60	56	54	54
Total		53	74	80	77	71	61

4.2.2. Detailed analyses

The categorization data from each closure type were reanalyzed by three prosodic patterns – cooperating, same-size and conflicting – and then by sub-types of these three distinctions. According to the IBH (when extended to this structure), categorization should be significantly better in the cooperating boundary patterns than in the same-size and conflicting boundary patterns. Sub-types of these three boundary patterns should show similar results to each other.

The results demonstrated that relative boundary size significantly affected syntactic processing decisions, and that most Korean L2ers were able to use relative boundary size in prosodic disambiguation. A repeated measure ANOVA was performed on categorization data with Syntax (Early vs. Late Closure) and Prosody (cooperating vs. same-size vs. conflicting prosodic patterns) as within-subjects factors and Group (English L1ers vs. extra-advanced vs. advanced vs. intermediate vs. beginning L2 groups) as a between-subjects factor. This ANOVA revealed significant main effects of Prosody ($F(2, 382) = 173.442$, $p < 0.01$) and Group ($F(4, 191) = 16.349$, $p < 0.01$), a significant interaction of Prosody with Group ($F(8, 382) = 4.255$, $p < 0.01$), and a significant interaction of Prosody with Syntax ($F(2, 382) = 7.355$, $p < 0.01$).

Both L1ers and L2ers were most accurate with the cooperating prosodic patterns. The significant interaction of Prosody with Group shows that the three prosodic patterns were used differently by different groups of proficiency and language: the overall performance was better than chance level in the ambiguous and conflicting prosodic patterns for the English L1ers (as in [9]) and extra-advanced groups, but near or below chance for the other groups. These results suggest that variation in edge tones, pitch accents, or pitch range can also influence L1 and L2 syntactic parsing [9]. However, for L2ers, immersion experience seems to be an important factor for being able to use such prosodic cues, since only the extra-advanced group had such experience.

Contra the predictions of the extended IBH, the sub-types of the cooperating/same-sized/conflicting boundary conditions achieved significantly different results from each other. (Since the conflicting boundary conditions provided few tokens, these conditions will be excluded in the following discussion.)

In the Early Closure condition, English L1ers' data were subjected to paired t-tests comparing the percentage correct of sub-types. The t-tests indicated that among the three sub-types of the cooperating pattern, the differences between the IPh-wd and IPh-ip patterns and between the IPh-wd and ip-wd patterns were significant ($p < 0.05$ in each case) but the difference between the IPh-ip and ip-wd patterns did not reach significance. The difference between the two sub-types of same-size prosody – the IPh-IPh and ip-ip patterns – was also significant ($p < 0.05$). Likewise, t-tests were run on the Korean L2ers' data: all three sub-types of cooperating patterns significantly differed from each other, as did the two sub-types of the same-sized pattern (p's always < 0.01). Clearly, neither group's performance was fully captured by the 3-way distinction of prosodic boundary size.

The sub-type analyses for Early Closure indicated other effects as well. An IPh at position # 1 (after *moves*) attracted higher correct categorization than an ip at the same position. This suggested stronger disambiguating effects for IPhs than ips and the importance of absolute boundary size at the major syntactic boundary. When an ip was the largest boundary at the subordinate boundary Korean L2ers' performance was extremely poor (42% correct on the average, as in Table 3). This implies that the ip boundaries did not sufficiently disambiguate the syntax to Early Closure, although they could be detected and used by the L2ers (as discussed below). However, since the results were based on only four ip tokens, which were not controlled for choices of pitch accents and phrase tones, further study is needed with more controlled materials.

In the Late Closure conditions, for both Korean L2ers and English L1ers paired t-tests showed highly significant differences between all pairs of sub-types of cooperating and same-size prosodic patterns ($p < 0.01$ in each case) except the wd-IPh versus wd-ip patterns. Thus, these results provided further evidence that larger/smaller/same-size categorization of prosodic boundaries is not sufficient to account for subjects' performance.

The Late Closure condition demonstrated several other effects. First, categorization for the wd-wd pattern was significantly better than for the IPh-IPh and ip-ip patterns, and categorization between the ip-IPh and wd-wd patterns did not significantly different from each other. These results suggest stronger effects of boundary #1 than boundary #2, presumably because boundary #1 occurs at the first syntactic choice point, demonstrating an incremental processing effect, or possibly from the effects of truncation on the perception of boundary #2. Second, the wd-IPh pattern achieved a significantly higher percentage correct than the ip-IPh pattern, as did the wd-ip pattern than the wd-wd pattern. This suggests that Korean L2ers could detect and use ips to some extent (although again further study is needed with more controlled materials). Finally, the non-significant difference between the wd-IPh and wd-ip patterns likely indicates an effect of the Late Closure preference.

5. Conclusion

The results from a forced-choice continuation selection experiment showed that cooperating prosody was more disambiguating than same-size and conflicting prosody for most subjects. This suggests that the L2 syntactic parser, like the L1 syntactic parser, is sensitive to the complete prosodic representation and interprets a local prosodic boundary with respect to it.

However, the patterns in these data suggest that the 3-way distinction of prosodic boundary size is too coarse. IPhs determined subordinate clause boundaries more strongly than ips (especially for Korean L2ers). Thus, the absolute prosodic boundary size in critical positions also seems to serve as a good predictor of syntactic disambiguation. In addition, contrasts in the first prosodic boundary produced stronger effects than contrasts in the second one. In short, the results suggest effects of both absolute and relative boundary size, resulting in more fine-grained prosodic distinctions.

5. References

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

[2] Beckman, M.E.; Ayers, G., 1997. *Guidelines for ToBI labeling*. Ms. Columbus, OH: Ohio State University.

[3] Carlson, K.; Clifton, C.; Frazier, L., 2001. Prosodic boundaries in adjunct attachment. *Journal of Memory and Language*, 45, 58-81.

[4] Clifton, C.; Carlson, K.; Frazier, L., 2002. Informative prosodic boundaries. *Language and Speech*, 45, 87-114.

[5] Frazier, L., 1978. *On Comprehending Sentences: Syntactic Parsing Strategies*. Ph.D. dissertation, University of Connecticut.

[6] Jun, S.-A., 1993. *The Phonetics and Phonology of Korean Prosody*. Ph.D. dissertation, The Ohio State University.

[7] Pierrehumbert, J., 1980. *The Phonetics and Phonology of English Intonation*. Ph.D. dissertation, MIT.

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

[9] Schafer, A.J.; Speer, S.R.; Warren, P.; White, S.D., 2000. Intonational disambiguation in sentence production and comprehension. *Journal of Psycholinguistic Research*, 29,169-182.

[10] Watson, D.; Gibson, E., 2004. The relationship between intonational phrasing and linguistic structure in language production. *Language and Cognitive Processes*, 19, 713-755.