The effect of phrase length and speech rate on prosodic phrasing

Sun-Ah Jun

UCLA, Dept. of Linguistics, Los Angeles, California, U.S.A

jun@humnet.ucla.edu

ABSTRACT

This paper presents the effect of phrase length and speech rate on prosodic phrasing in Korean. A passage consisting of 188 words was modified to include 220 words without affecting the meaning. Fourteen Korean speakers read the passage at normal and fast rates. Results show that the Korean AP includes 5 or fewer syllables at normal rate, but can include up to 7 syllables at fast rate. However, its effect is adjusted by other factors such as speech rate and semantic factors.

1. INTRODUCTION

Prosodic phrasing has been known to be influenced by several factors such as syntax, focus, old/new information, phrase length and speech rate [1, 2, 3, 7, 16]. Among these, syntax has been known to be a major factor affecting phrasing [12, 14, 15, 17]. A prosodic phrase has been claimed to be predicted by the alignment of the left or right edge of a syntactic constituent with the edge of a prosodic unit, by the head-complement relation and the domain of a maximal projection, or by the branchingness of the constituents.

However, studies on speech production and sentence processing have shown that prosodic phrasing determined by syntax is often overridden by the length of a phrase and speech rate [4, 5, 6, 7, 11]. For example, a predicate and an object noun often form one prosodic phrase, but if the object noun is long, the object tends to form its own prosodic phrase. For the same reason, a subject and an immediately following object, as in Japanese and Korean, can form one prosodic phrase if these two words are not long and at the same time the word following the object is not short. Hirose [5, 6] presents the role of prosodic phrasing in parsing where the phrasing is influenced by the constituent length more strongly than by the syntactic boundary. In a sentence fragment of Japanese, consisting of five or six words (i.e., single or double subject NP + object NP + adverb + verb + dative NP), a major prosodic boundary comes after a double subject NP, thus matching a syntactic boundary. But when the subject is a single NP, it comes after the object NP, mismatching a syntactic grouping. This major prosodic boundary either facilitated or delayed the processing of a sentence depending on whether the prosodic boundary matched the syntactic boundary or not. But the effect of length on phrasing would not be the same if the sentence were uttered fast or slow. It is known that a prosodic phrase tends to include more syllables at a fast rate and fewer syllables at a slow rate

compared to those at a normal rate.

This paper examines the effect of length and speech rate and their interactions on prosodic phrasing in Korean. It is part of an on-going research project investigating the effect of four factors (syntax, focus, length, rate) on prosodic phrasing (The effect of syntax and focus on phrasing is reported in Jun [10]). The results of the current paper will further our understanding of prosodic phonology and help us to build a relevant set of constraints on prosodic phrasing and their rankings.

2. METHODS

Subjects: 14 native speakers of Seoul Korean participated in the experiment. They were in their 20s and 30s and most of them were UCLA students.

Material: A passage about bamboo was taken from a Korean social studies textbook. The passage included 17 long and complex sentences containing 188 words. The passage was modified in such a way that 55 words or phrases were replaced with longer words or phrases while maintaining the meaning as much as possible (ex. [kin] => [kitaran] 'long'; [hunhitul] => [hunhi malhakilul] 'as is often said'). The lengthened version of the text included 220 words. Henceforth, I will call the original version a 'short' version and the lengthened version a 'long' version.

#	#of sylls Short => long	#of target word or phrases	#	#of sylls Short => long	#of target word or phrases
1	2=>3	1	10	5=>6	5
2	2=>7	1	11	5=>7	8
3	3=>4	2	12	5=>8	2
4	3=>5	3	13	6=>7	3
5	3=>6	3	14	6=>8	2
6	4=>5	6	15	6=>9	2
7	4=>6	4	16	7=>8	1
8	4=>7	1	17	7=>9	1
9	4=>9	1	sum		46

Table 1: Change in the number of syllables from the 'short' to the 'long' version and the frequency of each pattern

Table 1 provides the distribution of syllables modified from the short to the long version. It shows that there were 46 target words or phrases. Among these, 11 targets were 'word' targets' and 35 were 'phrase' targets. The 'word' targets include cases where adding more syllables to one word did not affect the phrasing of neighboring words due to the syntactic structure of the sentence. The 'phrase' targets include the cases where adding more syllables to one word may have influenced the phrasing of the neighboring words and the cases where one or more syllables were added to each of the two consecutive words which were closely related in meaning.

Procedure: Subjects were asked to read the two versions of the passage (short, long) at a normal and a fast rate. All utterances were digitized, and the prosodic phrasing of each utterance was determined by the intonation pattern of the utterance and the degree of juncture between words based on Jun's model of the intonational phonology of Seoul Korean and K-ToBI transcription conventions [7, 8, 9]. According to Jun's model, Seoul Korean has two prosodic units higher than a word. They are an Intonation Phrase and an Accentual Phrase. The Intonation Phrase includes one or more Accentual Phrases and is marked by a boundary tone and phrase final lengthening, and is optionally followed by a pause. The degree of juncture matching the end of this unit is represented as a break index '3' in K-ToBI. The Accentual Phrase (AP) is larger than a word but often includes only one content word [13]. It is marked by both a phrase initial and phrase final rising tone, but not by phrase final lengthening nor a following pause. The degree of juncture matching the end of an AP is represented as a break index '2'. Finally, the break index for an AP-medial word boundary is '1'. When there is a mismatch between the tonally defined unit and the perceived degree of juncture, a break index 'm' is added to the break index representing the degree of juncture. Thus, a break index '1m' refers to the case where the degree of juncture after a certain word is like the AP-medial word boundary but the tonal pattern suggests that the word is AP final. A break index '2m' refers to the case where the degree of juncture is equivalent to a break index 2 (the default boundary between APs), but there is no tonal marker showing an AP boundary.

Every word boundary in each utterance was labeled using a break index. We compared the break indices for every word between the short and long versions and between normal and fast rates. For the comparison, mismatch cases were included in a category based on the degree of juncture. That is, '1m' was combined with '1' data, and '2m' was combined with '2' data.

3. RESULTS AND DISCUSSION

Results show that, as found in Schafer and Jun [13], each content word tends to form one Accentual Phrase. However, when two adjacent words are closely related in meaning such as an adjective and a noun or a sequence of nouns close to a derived compound, the two words tend to form one Accentual Phrase only if the phrase is shorter than 6

syllables. The phrase tends to form two APs when a phrase is longer than 5 syllables at a normal rate and when it is longer than 7 syllables at a fast rate. Three patterns were found in the data regarding the relation between the number of syllables in the target word/phrase between short and long versions (see Table 1) and its impact on the prosodic phrasing.

3.1 WHEN THE PHRASE HAS 5 OR FEWER SYLLABLES IN THE *LONG* VERSION

This includes the types of syllable changes shown as #1, #3, #4, and #6 in Table 1. There were 4 word targets and 7 phrase targets, resulting in 11 targets. All target words/phrases formed a maximum of two APs in both versions. Therefore, for each type of target, the number of APs that increased due to the lengthening refers to the number of speakers who produced one AP in the short version and produced two APs in the long version. Table 2 shows the number of AP increases due to lengthening in both normal and fast rate. '-4' means that four speakers produced the target phrase in two AP in the short version and in one AP in the long version. '11' means 11 speakers produced the target phrase in one AP in the short version and in two APs in the long version.

Target phrase		# of words	# of syllables	# of AP increases	
#	id	Short=>long	short => long	Normal rate	Fast rate
1	8-1	2wd=>2wd	3=>5	0	4
2	8-2	2wd=>2wd	5=>5	1	5
3	11-1	1wd=>2wd	3=>5	1	1
4	12-1	2wd=>2wd	3=>4	1	1
5	15-3	1wd=>2wd	4=>5	0	0
6	17-1	2wd=>2wd	4=>5	0	3
7	18-3	1wd=>2wd	2=>3	0	3
8	19-1	2wd=>2wd	4=>5	-4	-3
9	19-2	1wd=>1wd	3=>4	0	0
10	7-3	2wd=>2wd	3=>5	7	11
11	13-1	2wd=>2wd	4=>5	8	2
12	18-4	2wd=>2wd	4=>5	10	6

Table 2: No. of AP increases in the 'long' version in normal and fast rates when the target in the long version has 5 or fewer syllables. Data combined across 14 speakers.

Table 2 shows two patterns of AP increase. The first 9 target phrases show that the number of APs in the long version did not change much from the short version, especially in the normal rate. This was true whether the target was one word or two words in the short version. This

suggests that when the target increased from one word to two words, the phrasing did not change when the phrase was shorter than 6 syllables. When the target was two words in the short version, speakers produced two APs even when each word was very short, 1-3 syllables, and the two words were semantically close as in Adj+N (e.g. 8-2: (kin)(jangte)=>(kitaran)(jangte) 'long'+'stick' → 'long stick') or a monosyllabic relative clause modifying a noun (e.g. 19-1: (ttan)(juksunun)=> (ttanen) (juksunun) 'to pluck'+'bamboo shoot' → 'bamboo shoot that's been plucked'.). This was an unexpected result if we consider the phrasing only from the number of syllables. What seems to be involved here is that the meaning of the monosyllabic modifier is not easily recoverable if it forms one AP with the following noun (e.g. ?(kin jangte)). It turns out that a monosyllabic modifier forms one AP with the following noun only when the sequence 'mod+N' is frequently used and the meaning is clear or when the modifier provides new information and the noun provides old information (e.g., (hayan jip) 'white house'). In fact, when the modifier is lengthened, some speakers produced the modifier and the noun in one AP probably because the meaning of the modifier is more easily recoverable when it is two syllables long. This is illustrated by the negative number of APs in the target phrase #8 (19-1) above.

At fast rate, however, the number of APs slightly increased in the long version when the target phrases were a monosyllabic modifier and a noun. This was because some speakers produced a one-AP target phrase in the short version, but two APs in the long version. This suggests that a semantic constraint, i.e., recoverability of the meaning of the monosyllabic modifier, is weakened at fast rate.

On the other hand, the last 3 target phrases show that speakers increased several APs at both rates. This is because speakers produced the target phrase (two words) in one AP in the short version and in two APs in the long version. In these cases, each target includes a monosyllabic noun whose meaning is almost empty (they are called an 'incomplete' noun in Korean grammar and have a generic meaning like 'a thing, a place, or a way'). For example, the target #12 (18-4) is (jaran kosul)=>(jarn)(juksonul) 'grew-mod'+'thing-Acc' (the thing which is grown) => 'grew-mod'+'bamboo-Acc.' (the bamboo which is grown). This suggests that a semantic constraint is stronger than the length constraint.

3.2 WHEN THE PHRASE HAS 6 OR MORE SYLLABLES IN THE *SHORT* VERSION

This includes the types of syllable changes shown as #13-16 in Table 1. Table 3 shows the phrasing data for each target type. All targets had two words in the short version and formed a maximum of two APs in both versions.

When a target phrase had 6 or more syllables in the short version, it generally formed two APs in both rates. Therefore, not much change in the accentual phrasing happened in the long version. The target #17 (18-2) was sometimes produced as one AP in the short version but two in the long version (i.e., (himtulki / ttemune) =>(himtulojiki)

(ttemune) 'to become hard-reason-why' → '(because) it gets harder'. But at fast rate, speakers did not change the phrasing of this phrase between short and long versions.

Target phrase		# of words	# of syllables	# of AP increases	
#	Id	Short=>long	short => long	Normal rate	Fast rate
13	10-2	2wd=>2wd	6=>7	0	1
14	14-2	2wd=>2wd	6=>7	1	4
15	14-3	2wd=>2wd	6=>7	0	1
16	20-1	2wd=>2wd	6=>7	0	1
17	18-2	2wd=>2wd	6=>8	6	1
18	7-2	2wd=>2wd	6=>9	3	2
19	21-1	2wd=>2wd	6=>9	0	0
20	15-4	2wd=>2wd	7=>8	0	1

Table 3: No. of AP increases in the 'long' version in normal and fast rate when the target in the short version has 6 or more syllables. Data combined across 14 speakers.

3.3 WHEN THE PHRASE HAS 5 OR FEWER SYLLABLES IN THE *SHORT* VERSION

This category includes the rest of the syllable types in Table 1. This type showed the greatest increase in the number of APs between the short version and the long version. This is expected because a phrase with 5 or fewer syllables often forms one AP and a phrase with 6 or more syllables often forms two APs. The tendency was weakened at fast rate because some speakers still produced 6 or 7 syllable long phrases in the long version in one AP. This suggests that the maximum number of syllables forming one AP increases from 5 at a normal rate to 7 at a fast rate.

Table 4 shows the number of AP increases when the target phrase in the short version has 5 syllables. Except for two cases, all targets had two words in the short version. However, most speakers produced the target phrases in one AP in the short version because the phrases contained either a monosyllabic incomplete noun or a sequence of two nouns whose meaning is like a derived compound noun.

The last 3 phrases in Table 4, on the other hand, show no increase in APs in normal rate. This is because speakers read the short version in two APs. The reason is that, for target #34 (16-2), the first word is a monosyllabic adjective, and as mentioned in section 3.1, speakers often produced the adjective in one AP, not including the following noun (i.e., (khun)(tenamuro) => (khotaran)(tenamuro) 'large'+ 'bamboo -with' → 'with a large bamboo'. For the other two, whose first word was 2 or 3 syllables long, the whole phrase was often produced in one Intonation phrase (IP). It was found that an IP final word often forms one AP even when it is short. These phrases were not produced in one IP

at fast rate, so the number of APs increased at fast rate.

Target phrase		# of words	# of syllables	# of AP increases	
#	Id	Short->long	short => long	Normal rate	Fast rate
21	5-2	2wd->2wd	5=>6	7	4
22	15-1	2wd->2wd	5=>6	5	3
23	6-2	2wd->2wd	5=>6	4	4
24	15-2	2wd->2wd	5=>6	6	3
25	4-1	2wd->2wd	5=>6	13	10
26	9	1wd->2wd	5=>7	6	2
27	16-1	2wd->2wd	5=>7	5	2
28	21-2	1wd->2wd	5=>7	13	11
29	5-1	2wd->2wd	5=>7	5	7
30	18-1	2wd->2wd	5=>7	3	1
31	22-1	2wd->2wd	5=>8	8	12
32	17-2	2wd->2wd	5=>8	12	14
33	6-1	2wd->2wd	5=>7	0	8
34	16-2	2wd->2wd	5=>7	0	5
35	2-1	2wd->2wd	5=>7	0	4

Table 4: No. of AP increases in the 'long' version in normal and fast rate when the target is 5 syllables long in the short version and is 6 or more syllables long in the long version. Data combined across 14 speakers.

Eleven other target phrases whose short version had 4 or fewer syllables and whose long version had 6 or more syllables showed more AP increases than those with 5 syllables in the short version. This is because a phrase shorter than 5 syllables was more often produced in one AP than was a 5 syllable phrase.

When the phrasing was compared across lengths and rates, two patterns were found depending on the size of the target. The number of APs in the target 'words' (1wd => 1 or 2wd) in the long version increased more at normal rate than at fast rate (72 vs. 55), but in the target 'phrase' (2wd => 2wd), it decreased at the normal rate compared with the fast rate (146 vs. 156). The reason is that target 'words' form one AP in the short version for both rates, but in the long version, they form two APs at normal rate but one AP at fast rate. For target 'phrases', they form two APs in the short and long versions at normal rate, while at fast rate, they formed one AP in the short version but 2 APs in the long version. In sum, though the length of a phrase affects prosodic phrasing in Korean, its influence is adjusted by other factors such as speech rate and semantic factors. Data

from more languages should be investigated to generalize the current findings.

REFERENCES

[1] A. Cutler, D. Dahan, & W, van Donselaar, "Prosody in the Comprehension of Spoken Language: A Literature Review," *Language and Speech,* vol. 40, pp. 141-201, 1997.

[2] E. Delais-Roussarie, *Pour une approche parallèle de la structure prosodique,* Thèse de Doctorat, Univ. Toulouse le Mirail, 1995.

[3] F. Ferreira, "The creation of prosody during sentence production," *Psychological Review*, vol. 100, pp. 233-253, 1993.

[4] P. Gee & F. Grosjean, "Performance structures: A psycholinguistic and linguistic appraisal," *Cognitive Psychology,* vol. 15, pp. 411-458, 1983.

[5] Y. Hirose, *Resolving Reanalysis Ambiguity in Japanese Relative Clauses,* Unpublished doctoral dissertation, CUNY Graduate Center, N.Y., 1999.

[6] Y. Hirose, "Recycling prosodic boundaries," *Journal of Psycholinguistic Research,* in press.

[7] S.-A. Jun, *The Phonetics and Phonology of Korean Prosody,* Doctoral dissertation, The Ohio State University, Columbus, Ohio, 1993 [Published in 1996 by Garland, New York].

[8] S.-A. Jun, "The Accentual Phrase in the Korean prosodic hierarchy," *Phonology,* vol. 15(2), pp. 189-226, 1998.

[9] S.-A. Jun, "K-ToBI (Korean ToBI) labeling conventions - Version 3," *The Korean Journal of Speech Sciences*, vol. 7(1), pp. 143-169, 2000.

[10] S.-A. Jun, "Syntax over Focus" in the *Proceedings of ICSLP*, John H. L. Hansen & B. Pellom, Eds., pp. 2281-2284, Denver, CO., 2002.

[11] N. Lovrić, D. Bradley, and J. D. Fodor, "Silent prosody resolves syntactic ambiguities: Evidence from Croatian," Paper presented at the SUNY/ CUNY/NYU Conference, Stonybrook, NY., 2001.

[12] M. Nespor and I. Vogel, *Prosodic Phonology,* Dordrecht: Foris, 1986.

[13] A. Schafer and S.-A. Jun, "Effects of Accentual Phrasing on Adjective Interpretation in Korean," in *East Asian Language Processing,* M. Nakayama Ed., Stanford, CSLI. pp.223-255, 2002.

[14] E. Selkirk, "On derived domains in sentence phonology," *Phonology Yearbook*, vol. 3, pp.371-405, 1986.

[15] E. Selkirk, "The interaction of constraints on prosodic phrasing", in *Prosody: Theory and Experiment. Studies presented to Gosta Bruce*, M. Horne, Ed., pp. 219-255. Dordrecht: Kluwer Academic Publishers, 2000.

[16] S. Shattuck-Hufnagel and A. Turk, "A prosody tutorial for investigators of auditory sentence processing," *Journal of Psycholinguistic Research,* vol. 25(2), pp. 193-247, 1996.

[17] H. Truckenbrodt, "On the Relation between Syntactic Phrases and Phonological Phrases," *Linguistic Inquiry*, vol. 30(2), pp. 219-255, 1999.