

Intonational Pattern Frequency of Seoul Korean and Its Implication to Word Segmentation*

Sahyang Kim**

ABSTRACT

The current study investigated distributional properties of the Korean Accentual Phrase and their implication to word segmentation. The properties examined were the frequency of various AP tonal patterns, the types of tonal patterns that are imposed upon content words, and the average number and temporal location of content words within the AP. A total of 414 sentences from the Read speech corpus and the Radio corpus were used for the data analysis. The results showed that the 84% of the APs contained one content word, and that almost 90% of the content words are located in AP-initial position. When the AP-initial onset was not an aspirated or tense consonant, the most common AP patterns were LH, LHH, and LHLH (78%), and 88% of the multisyllabic content words start with a rising tone in AP-initial position. When the AP-initial onset was an aspirated or tense consonant, the most common AP patterns were HH, HHLH, and HHL (72%), and 74% of the multisyllabic content words start with a level H tone in AP-initial position. The data further showed that 84.1% of APs end with the final H tone. The findings provide valuable information about the prosodic pattern and structure of Korean APs, and account for the results of a previous study which showed that Korean listeners are sensitive to AP-initial rising and AP-final high tones (Kim, 2007). This is in line with other cross-linguistic research which has revealed the correlation between prosodic probability and speech processing strategy.

Keywords: frequency, intonation, word segmentation, Accentual Phrase, Korean

I. Introduction

Previous studies have shown that listeners use language-specific prosodic cues for word segmentation (Cutler & Norris, 1988; Cutler & Otake, 1994; Kim, 2007; Welby, 2007, *inter alia*). Listeners' use of certain prosodic cues has been known to be governed by probabilities. For example, 90% of English open-class words that are listed in dictionaries begin with strong syllables (Cutler & Carter, 1987), and English listeners employed a segmentation strategy based on this frequent metrical pattern, *viz.*, trochaic stress pattern (Cutler & Norris, 1988). English listeners were faster at detecting 'mint' in a non-sense word *mintayf* [mɪntef], which was composed of two strong syllables, than in *mintef* [mɪntɛf], which was composed of a strong syllable followed by a weak

* This work was supported by 2008 Hongik University Research Fund.

** Department of English Education, Hongik University.

syllable. Cutler and Norris (1988) suggested that the listeners put a word boundary in front of a strong syllable (in this case, in front of 't' in *mintayf*), such that they have to rebuild the syllable structure of the given string (*mintayf*) in order to detect 'mint'. The detection of the embedded word in *mintef*, in contrast, does not involve such restructuring process because listeners usually do not divide the sequence in front of a weak syllable. Similar results were also found in Dutch. As for input frequency, 90% of Dutch words start with a stressed syllable (Quene, 1992). Just like English listeners, Dutch listeners were faster in detecting words that were embedded in strong-strong stress context than in strong-weak stress context, indicating that they also tend to put a word boundary in front of a strong syllable (Vroomen & de Gelder, 1995). The correlation between the prosodic probabilities and word segmentation is not confined to languages with lexical stress. Warner *et al.* (submitted) performed a corpus study and an on-line perception study on Japanese recently. They collected speech data from eight Japanese speakers, and found that 68% of words in the corpus started with Accentual Phrase initial rising. Japanese listeners employed this cue during word segmentation: They detected words faster when the words occurred with Accentual Phrase rising than when not. Taken together, these studies strongly suggest that listeners are sensitive to prosodic probabilities in their own language, and they use probabilistic knowledge in segmenting words from speech stream.

The effect of intonation pattern in Korean word segmentation has been shown in an on-line perception study (Kim, 2007): Korean listeners were sensitive to Accentual Phrase (henceforth AP) initial rising, and they also seemed to be affected by AP final H tone. The results imply that listeners can be benefited from specific AP intonational patterns when words are aligned with the edges of AP. The current study examines whether this implication can be supported by various levels of distributional properties of Korean AP. A corpus study is undertaken in order to obtain a quantitative data of Korean intonation pattern and to find grounds for the previous perception study. Although Korean AP is phonologically defined as LHLH (the initial tone can be H when the initial segment is an aspirated or tensed consonant), many other intonation patterns are possible, according to Jun (2000). Given that there are quite a few varieties of AP tone patterns, a question still remains as to whether speech input delivers enough regularity to be found and exploited by the listeners for speech segmentation. This study obtained speech production data from the recordings of read speech (henceforth Read speech corpus) and two radio shows (henceforth Radio corpus), and collected distributional information from these data. To observe the frequency of various intonational patterns of APs and content words, the data were transcribed using K-ToBI labeling conventions (Jun, 2000). Furthermore, in order to examine whether the frequency of post-lexical intonation could directly influence listeners' word segmentation, the study examined the numbers and locations of content words within the AP. The next section will explain the nature of the corpus adopted for this study in detail.

2. Method

2.1 The corpora

The Read speech corpus used in the current study was originally designed to investigate default phrasing and attachment

preference for relative clauses in speech production (Jun & Kim, 2004). The list contained fifty-six sentences. Among them, thirty-two sentences had one relative clause and two NPs in them (AC-NP1-NP2), while twenty-four sentences did not have this sentential structure. The current study used this production data for the analysis of the corpus. The data employed here were produced by two male and two female native speakers of Seoul Korean, who were graduate students at UCLA at the time of recording. The sentences were read at normal speed, and four speakers each read 56 sentences. Thus, a total of 224 sentences were analyzed for the current study.

The Radio corpus was recorded from two radio programs, *Radio Moodae* 'radio stage' and *Radio Dokseoshil* 'radio reading room', both of which aired on KBS (Korean Broadcasting System) I Radio. *Radio Moodae* ('radio stage') broadcasts radio drama as recorded by voice actors. *Radio Dokseoshil* ('radio reading room') introduces and reviews newly published books. The format of *Radio Dokseoshil* includes an introduction of a novel by the show host, an interview with the author of the novel, and the dramatized presentation of the novel. Thus, the Radio corpus was more similar to spontaneous speech than the Read speech corpus in general, and some part of this data actually contained sentences from spontaneous conversation. The show host, interviewees, and voice actors were all Seoul Korean speakers. Two hundred sentences were extracted from various parts of two one-hour episodes. The sentences that were produced with background music and noise were excluded.

2.2 Analysis

The recorded sentences were transcribed using K-ToBI (Jun, 2000), by the author. The first step for the analysis was to locate AP boundaries and count the number of APs in each corpus. The Read speech corpus had 1592 APs and the Radio corpus had 1493 APs. The corpus was annotated with intonational patterns of APs so as to examine the frequency of the attested tonal patterns. For this analysis, IP-final APs were excluded because AP-final tone is pre-empted by IP-boundary tone when an AP is in the IP-final position. The Read speech corpus had 1203 non-IP-final APs, and the Radio corpus had 906, making the total number of APs used for the intonation frequency analysis 2109. The intonational patterns that were imposed upon content words were also transcribed. Again, only non-IP-final APs were considered for this analysis. Additionally, it should be noted that monosyllabic content words were not included for the study, and that only the intonation pattern of the first two syllables of multisyllabic content words were included. Although monosyllabic nouns and adverbs were not included in the tonal pattern analysis for content words, monosyllabic verbal and adjectival stems were included, because these are bound stems in Korean and hence can compose a prosodic word only when they are combined with bound suffixes.

There were four possible combinations of tone patterns that could be imposed on the initial two syllables of words: level H (HH), level L (LL), falling (HL), and rising (LH). In the tonal transcription, contour tones (LH and HL) included not only when the target syllable shows the lowest F0 point (L) or the highest F0 point (H) (i.e., Figure 1, left panels), but also when the second syllable is in the middle of rising or falling, thus not being specified with a tone (i.e., Figure 1 right panels).

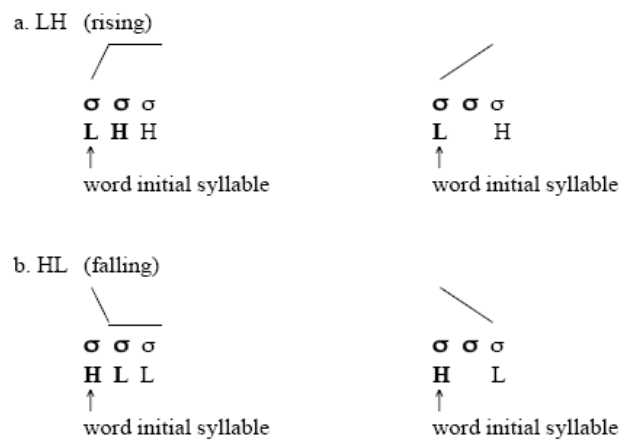


Figure 1. Contour tones on the first two syllables of content words (a. rising, b. falling)

Since the AP-initial tone is determined by the laryngeal feature of the AP-initial segments (Jun, 2000), the data were divided into two groups depending on the property of initial segments: one group had segments that induce H tone in the AP-initial position (i.e., tense and aspirated consonants), and the other group had segments that do not induce H tone in the same position (i.e., the rest of the consonants and vowels). For the sake of convenience, I will refer to the segments that consist of the former group as H tone inducers and those that consist of the latter group as L tone inducers.

3. Results

3.1 AP Intonational Pattern Frequency

Among 2109 non-IP-final APs, 1541 APs started with an L tone inducer (i.e., lenis consonants, nasal consonants, glides, vowels), and 568 APs had an H tone inducer (i.e., tense and aspirated consonants) at the phrase-initial position. The frequency of AP tonal patterns varied depending on what the AP-initial segment was. With L tone inducers, the most frequent pattern was LH, which composed 44% of the data; the second most frequent pattern was LHH (19.9%); the third most frequent pattern was the default tone pattern of LHLH (14.2%). Note also that 88.8% of the APs ended with an AP-final H tone (LH, LHH, LHLH, LLH, HH, HLH, HHLH), and 85.5% of the APs started with an AP-initial rising tone (LH, LHH, LHLH, LHL, LHLL). A small number of AP-initial high tones were observed (4.9%) even though the initial segments were L tone inducers. Although the general tendency was the same, two corpora were a little different with regard to the distribution of the AP tone patterns, in that the Radio corpus showed more variation than the Read speech corpus. Both corpora had all seven tonal patterns starting with an L tone (*viz.*, LH, LHH, LHLH, LLH, LHL, LHLL, LL) and, among them, the three most frequent patterns were LH, LHH and LHLH. However, while 88.9% of the APs had one of the three patterns in the Read speech corpus, only 64.8% of the APs in the Radio corpus appeared with these three

patterns. In addition, the Read speech corpus had three tone patterns with H-initial tone (2.6%: HH, HHLH, HLH), but the Radio corpus had six tone patterns with H-initial tone (7.7%: HH, HHL, HL, HHLH, HLH, HLL). Figure 2 illustrates the distribution of three most-frequent tone patterns (i.e., LH, LHH, LHLH) in the data within each AP syllable count.

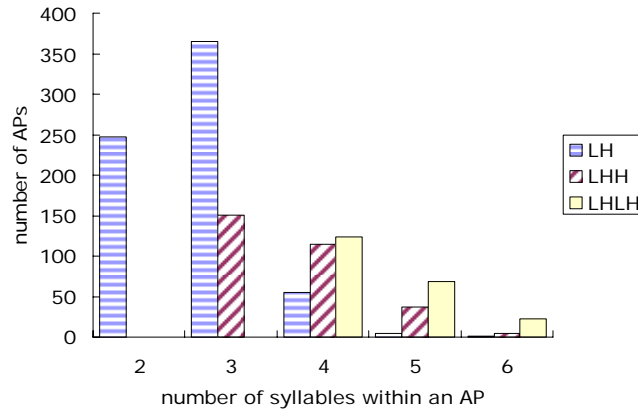


Figure 2. The frequency of AP tonal patterns depending on syllable count (with L tone inducers)

LH was the most frequent pattern when the APs had two or three syllables. The default pattern of LHLH was the most frequent tonal pattern when the APs had more than three syllables, which suggests that the low frequency of the default pattern in the corpus is due to the small number of syllables within the AP. However, it should be noted that the patterns of LH and LHH were still observed, with a non-ignorable proportion, in the APs with four syllables.

As for the intonational pattern frequency of APs with H tone inducers, the most frequent pattern was HH (37.5%), followed by the HHLH pattern (18.7%) and then the HHL pattern (15.7%). These three tones account for 71.8% of the APs. The two corpora shared the three most frequent patterns, with the same ranking (HH > HHLH > HHL). A small portion of the AP-initial L tones was observed (6.7%), despite of the existence of H tone inducers in the AP-initial position. An AP-final H tone was frequently observed in APs with H tone inducers, as well: 71.1% of the APs had H tone at the AP-final position. An AP-final L tone was less common (28.9%), but the proportion of these patterns was twice as high as that found in the APs with L tone inducers (11.1%). This seems to be due to the ‘see-saw effect’ (Jun, 1996a), which describes speakers’ tendency to avoid a sequence of two identical tones. Again, as shown in Figure 3, the data revealed that the occurrence of the default pattern (HHLH) was frequent when the APs had more than four syllables. HH was the most frequent pattern in bi- and tri-syllabic APs.

All in all, the results show that the same trend is observed with both L tone inducers and H tone inducers. That is, the tones that were most frequently realized are the AP-initial tone (either L or H) followed by a H tone, and the AP-final tone (H).

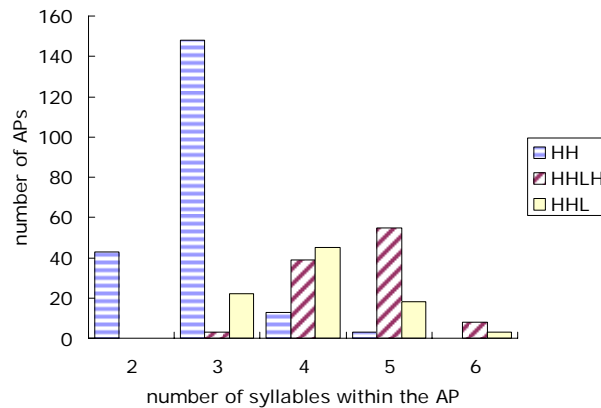


Figure 3. The frequency of AP tonal patterns depending on syllable count (with H tone inducers)

3.2 Content Word Initial Tone Patterns

Among the 2017 multisyllabic content words, 1471 had an L inducer as a word-initial segment. 1309 words out of 1471 appeared in AP-initial position (i.e., a word-initial segment aligns with the onset of an AP); the rest appeared in AP-medial position (i.e., a word initial segment does not align with the onset of an AP, and the word is located in the middle of an AP). As aforementioned, the word-initial tones were classified into four categories (LH, LL, HH, HL). In AP-initial position, 88% of multisyllabic content words occurred with a rising tone (LH) pattern. This tendency was stronger in the Aead speech corpus, where 91.5% of words carried an initial rising tone. In the Radio corpus, 82.9% of words started with a rising tone. The distribution of tone patterns that were realized on the first two syllables of AP-initial multisyllabic words with L inducers is illustrated in Figure 4.

Among the words with an H inducer, 502 words occurred in AP-initial position while 44 occurred in AP-medial position. As shown in Figure 5, 74% of multisyllabic content words with an H inducer had a level H tone (HH) at the beginning of word, and 23% of them had a falling tone (HL). This tendency was a little stronger in the Aead speech corpus, where 75.5% of words occurred with the level H tone, than the Radio corpus, where 70.9% of words occurred with a level H tone. Rising tone (LH) was not common with the words that started with an H inducer.

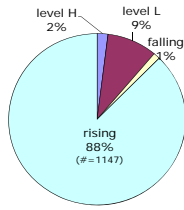


Figure 4. Tonal patterns of AP-initial

multisyllabic content words
(with L tone inducers)

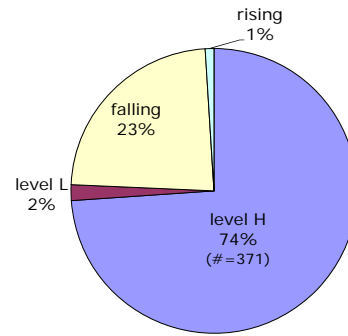


Figure 5. Tonal patterns of AP-initial

multisyllabic content words
(with H tone inducers)

Segments with different laryngeal features do not seem to strongly induce specific tones in AP-medial position as they do in AP-initial position (Jun, 1996). However, there is still a possibility that laryngeal features of word-initial segment still affect phonetic tone realization. Thus, this study will report the AP-medial words which had the L tone inducers as initial segments separately from those which had the H tone inducers as initial segments.

Table 1. Tonal patterns of AP-medial multisyllabic content words

	Rising	Level H	Falling	Level L	TOTAL
L tone inducers	89	28	28	17	162
Number of words (%)	(55%)	(17%)	(17%)	(10%)	(100%)
H tone inducers	9	13	14	8	44
Number of words (%)	(20%)	(30%)	(32%)	(18%)	(100%)

Table 1 shows the tone patterns on the first two syllables of AP-medial multisyllabic words with L tone inducers and H tone inducers. 56% of the words with L tone inducers started with a rising tone in AP-medial positions. Separate analysis of the two corpora revealed that 65.6% of the words started with a rising tone in the Read speech corpus, while only 41.7% of the words started with a rising tone in the Radio corpus. Further observation showed that 69 out of 90 words that started with a rising tone in AP-medial position had their word offsets aligned with AP-final syllables, indicating that the rising pattern observed in this position is the realization of AP-final rising tone (e.g., the second LH in the LHLH tone pattern). The number of AP-medial words with an H tone inducer was very small (n=44), but more than half of the words occurred either with level H tone (HH) or with falling tone (HL). Word-initial rising tone was not commonly observed in this position, unlike the pattern observed with the words

with an L tone inducer.

Overall, regardless of the location of words within the AP, 84% (1236 out of 1471 words) of multisyllabic content words with an L inducer start with a rising tone pattern (LH), and about 70% (384 out of 546 words) of those with an H inducer start with a level H tone pattern (HH) in the current data.

3.3 The Average Number and Location of Content Words within the AP

The average number of multisyllabic content words within the AP was calculated based on 3085 APs (including IP-final APs) from two types of corpora. The average number of content words within the AP was 1.14. The Read speech corpus and the Radio corpus did not differ much when these numbers were counted separately: the average number of content words was 1.14 for the Read speech corpus and 1.13 for the Radio corpus. This is similar to the result of Jun and Fougeron's (2000) study, which reported that the average number of content words was 1.2. The distribution of data, in terms of the number of content words within the AP, is shown in <Figure 6>.

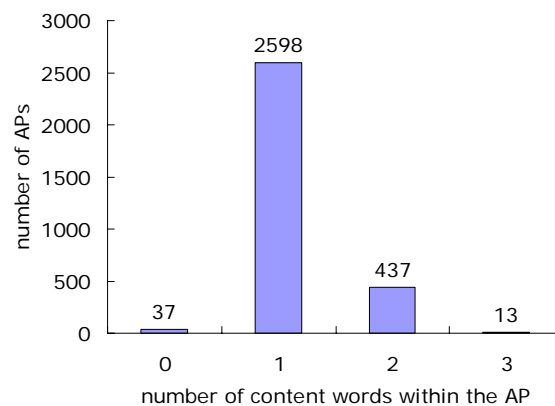


Figure 6. Frequency distribution: number of multisyllabic content words within the AP
(Based on 3085 APs from both corpora)

The majority of APs had one content word (84.2%). Only a small number of APs contained three content words ($n=13$) or just had function words without any content word (i.e., 0 content words in Figure 6, $n=37$), and no AP with more than three content words was found.

The co-occurrence frequency of AP onset and content-word onset was extremely high. The non-IP-final APs in this corpus contained 2343 content words, and among them, 2076 words occurred in AP-initial position (88.6%) and 267 words appeared in AP-medial position (11.4%). This means that the vast majority of multisyllabic content words appeared in AP-initial position in the corpus. Once monosyllabic nouns and adverbs were excluded from the data set, there were 2017 multisyllabic content words, with 1811 of these located in AP-initial position (89.8%) and 206 located in AP-medial position (10.2%).

4. Discussion and Conclusion

This study conducted prosodic analysis of two kinds of corpora, in order to obtain distributional data of Korean APs. The current study found speakers' robust use of specific tonal patterns within APs. When AP-initial consonants are L tone inducers, the most frequent AP tone patterns were LH, LHH, and LHLH. When AP-initial consonants are H tone inducers, HH, HHLH, and HHL were most frequent. In addition, more than 85% of multisyllabic content words begin with a rising tone when there is an L inducing segment in AP-initial position; 74% of multisyllabic content words with H tone inducers are produced with HH tone in the AP-initial position. In general, AP-final H tone is frequently observed (more than 85% for L inducers and more than 70% for H tone inducers), which suggests that this pattern may help listeners' word segmentation by successfully marking AP boundaries. The data further showed that the average number of content words within the AP was 1.14, and that about 90% of multisyllabic content words appear in AP-initial position. This suggests that the detection of AP could directly help the detection of content words from speech stream, at least in Korean.

The obtained distributional data not only provide a better understanding for the production of Korean prosody, but also have implication on speech processing. The number and the location of content words within an AP suggests that the Korean AP could an appropriate unit for modulating lexical access, because the average number of content words contained within an AP is not more than two, and because there is a high probability of co-occurrence rate of AP onset and content word onset. Further, prosodic probabilities, that is, the high frequency of AP-initial rising pattern (with L-inducers only) and AP-final H tone provide grounds for the previous on-line perception study (Kim, 2007), which showed Korean listeners' active use of those cues in segmenting multisyllabic content words with L-inducers from speech stream. Therefore, the results of this study add to the body of cross-linguistic research which has revealed the correlation between prosodic probability and speech processing strategy.

References

- Cutler, A. & Carter, D. 1987. "The predominance of strong initial syllables in the English vocabulary." *Computer Speech and Language* 2, 133-142.
- Cutler, A. & Norris, D. 1988. "The role of strong syllables in segmentation for lexical access." *Journal of Experimental Psychology: Human Perception and Performance* 14, 113-121.
- Cutler, A. & Otake, T. 1994. "Mora or phoneme? Further evidence for language-specific listening." *Journal of Memory and Language* 33, 824-844.
- Jun, S. 1996. *The Phonetics and Phonology of Korean Prosody: Intonational phonology and prosodic structure*, New York, NY: Garland Publishing Inc.

- Jun, S. 2000. "K-ToBI (Korean ToBI) labelling conventions: Version 3," *Speech Sciences* 7, 143-169.
- Jun, S. & Fougeron, C. 2000. "A Phonological Model of French Intonation." In A. Botinis (ed.) *Intonation: Analysis, Modeling and Technology*, Kluwer Academic Publishers, 209-242.
- Jun, S. & Kim, S. 2004. "Default phrasing and attachment preference in Korean", *Proceedings of the INTERSPEECH-ICSLP*, Jeju, Korea.
- Kim, S. 2007. "The role of post-lexical intonation patterns in Korean word segmentation." *Speech Sciences* 14 (1), 37-62.
- Quene, H. 1992. "Integration of acoustic-phonetic cues in word segmentation." In M. Schouten (ed.), *The Auditory Processing of Speech: From Sounds to Words*, 349-355.
- Uroomen, J. & de Gelder, B. 1995. "Metrical segmentation and lexical inhibition in spoken word recognition." *Journal of Experimental Psychology: Human Perception and Performance* 23, 710-720.
- Warner, N., Otake, T. & Takayuki, A. *Submitted*. "Intonational Structure as a Word Boundary Cue in Japanese."
- Welby, P. 2007. "The role of early fundamental frequency rises and elbows in French word segmentation." *Speech Communication* 49, 28-48.

received: April 30, 2008

accepted: June 2, 2008

▲ Sahyang Kim

Dept. of English Education, Hong-ik University

Mapo-Ku Sangsoo-Dong 72-1, Seoul 121-791

Tel: +82-2-320-1848

E-mail: sahyang@hongik.ac.kr