Segmental Acoustic Correlates Associated with the Korean Lenis Stops

Hongoak Yun
(Sungkyungkwan University)

Yun, Hongoak. (2013). Segmental Acoustic Correlates Associated with the Korean Lenis Stops. ***Language Research*** **49.1, 73-94.**

The purpose of this study was to investigate which acoustic correlates to stop feature distinction were realized due to the presence of a segmentation boundary and how semantic focus information contributed to acoustic correlates associated with segmental units. A production study was conducted using Korean lenis stops. The closure durations and VOTs of lenis stops were longer when they were in boundary-initial positions than when they were in boundary-internal positions. The VOTs became longer as the position of words in the structural hierarchy of sentences grew higher. The F0s of vowels were lower when they followed boundary-initial lenis stops than when they followed boundary-internal lenis stops. The semantic focus information resulted in longer VOTs of initial lenis stops and lower F0s in vowels following initial lenis stops. These results indicated that talkers reliably indicated the onset of words and phrases in the closure durations and the VOTs of lenis stops and the F0s of their following vowels.

Keywords: Korean lenis stop, segmentation, focus, boundary, segmental unit

One of the characteristics about natural speech sounds is that they are indiscrete, in part, due to the co-articulation with neighboring segments. Nonetheless, studies showed that talkers produce reliable acoustic correlates to signal discrete linguistic unite, so that listeners could successfully perceive continuous speech as a series of discrete linguistic units (e.g., phonemes, words, or phrases) (Liberman, Harris et al. 1957). For instance, the metrical structure of speech, such as moraic information in Japanese (Culterand Otake 1994) and stress distinction between strong and weak syllables in English (Norris, McQueen et al. 1995), cuea syllable boundary. Acoustic correlates (e.g., lengthening, aspiration, and pitch accent) associated with onset syllables often indicate the presence of a word boundary in front of those onset syllables (Lehiste 1972, Gowand Gordon 1995). Word-initial voiceless stop consonants in English tend to be more aspirated than word-internal stops (Trager and Bloch 1941, Lehiste 1960), and pitch accents contribute tolengthening stressed initial consonants and vowels (Klatt 1974). Glottal stops and laryngealized voicing for vowel-initial words and increased aspiration on voiceless stops played explicit physical cues for word onsets (Christie 1974, Lehiste 1960). Intonation patterns across words play an important role in determining the acoustics of segmental units at a phrasal level (Beckman and Pierrehumbert 1986, Jun 1993, Jun 1996, Jun 2002, Jun 2003).

The role of acoustic correlates associated with syllables is important in cueing the position of a segmentation boundary within the stream of sounds, specially, when the representation of the stream is ambiguous. For example, a continuous speech signal of /··· kapaŋ···/ in Korean could be represented ambiguously: 1) One representation is a two-syllable word, /ka.paŋ/ (meaning a *bag* in English), having a syllable boundary within the word. 2) The other representation, /-ka|paŋ/, consists of two syllables with a phrasal (and word) boundary, indicated by '|', which splits a subjective case marker (i.e., /-ka/) from a one-syllable word (i.e., /paŋ/ meaning a *room* in English). In order to distinguish /ka.paŋ/ from /-ka|paŋ/, Korean speakers might produce particular acoustic correlates associated with [k] and [p] in order to indicate one representation over the other, so that Korean listeners can successfully perceive those representations with less ambiguity.

The purpose of this study was to explore such a segmentation phenomenon in Korean. For this purpose, I examined the characteristics of acoustic correlates associated with [p], [t], and [k], so-called lenis stops. More precisely, I closely investigated how acoustic correlates associated with the lenis stops play a role as a segmentation cue to signal the presence of a boundary as a function of the size of hierarchical prosodic structures (Cho and Keating 2001) and the presence of the semantic focus information.

1. Background

The acoustic qualities of Korean stops have, at least, two distinctive characteristics. First, unlike the fact that stop consonants in most languages are categorized in terms of voicing or aspiration (Keating 1984), Korean stops are cross-linguistically unique in that they have a three-way distinction and all stops are voiceless (Keating, Linker et al. 1983, Sohn 1999, Cho, Jun et al. 2002). These stops are categorized as tense stops (i.e., /p*, t*, k*/), lenis stops (i.e., /p, t, k/), and aspirated stops (i.e., /p^h, t^h, k^h/). Second, the Korean lenis stops are of particular interest in that the tenseness and aspiration of the lenis stops exist as a phonemic feature but the voicing differences exist by an intervocalic voicing rule. In this rule, the lenis stops become voiced intervocalically within phonological phrases but remain voiceless in the initial position of a phonological phrase (Sohn 1999). Moreover, fundamental frequencies (hereafter, F0s) of vowels which follow phrasal-initial lenis stops are lower than those of vowels which follow tense or aspirated stops, suggesting that the initial lenis stops have voiceless qualities (Jun 1993, Cho1 996, Jun 1996, Cho, Jun et al. 2002, Kimand Beddor 2002).

What the voicing qualities of lenis stops and the F0s of their following vowels signify is that these acoustic correlates might cue the presence of a boundary in segmentation. For example, the ambiguity of the previous example /⋯kapaŋ⋯/ might be resolved by taking into account the acoustic features of the stops [k] and [p] and their following vowels, [a]. For the two-syllable word case (i.e., /ka.paŋ/) in which a word boundary is present initially, the first stop /k/ (c.f., word-initial stop) remains voiceless but the second stop /p/ (c.f., word-internal stop) becomes voiced intervocalically. As for this representation, its phonetic representation is likely to be [ka.paŋ], in which the F0 of /a/ following /k/ would be lower than that of /a/ following voiced /k/. In contrast, for the other case (i.e., /-ka|paŋ/) in which a boundary separates the two syllables, the first stop /k/ (c.f., word-internal stop) becomes voiced intervocalically[1)] but the second stop /p/ (c.f., word-initial stop) remains voiceless. In this case, the pho-

1) The subject marker /-ka/ immediately binds with a subject, thus the marker and the subject form one unit, resulting in /k/ in /-ka/ is not in the initial position but in the internal position of a phrase.

netic representation for /ka.paŋ/ might be realized as [-ga|paŋ] in which the F0 of /a/ following /p/ will be lower than that of /a/ following voiced /p/. A key point is that the degree of the voicelessness of initial lenis stops and the lowerness of the F0s of vowels following the initial lenis stops indicate where a boundary is present or not.

Jun (1993, 1996) argued that VOTs or F0s associated with lenis stops at a word level did not necessarily indicate the lenis stop voicing. Instead, she observed that the intonational pattern of a phrase, the Accentual Phrase (hereafter, AP), is correlated with the domain of the lenis stop voicing, and proposed that an AP is the domain for the voicing rule. The Lenis Stop Voicing Rule, which Jun named, defines the voicing pattern of lenis stops in a way that the lenis stops are voiced within an AP but not across an AP boundary. Thus, with respect to the voicing of lenis stops, there are no significant acoustic differences between word-initial but AP-internal stops and word-internal and AP-internal stops. This suggests that the voicing quality of lenis stops might not signal a word boundary, unless the word boundary coincides with an AP boundary.

Cho and Keating (2001) provided evidence related to the voicing issue associated with the Korean lenis stops. Their assumption was that phrase-initial consonants are more strengthened (or lengthened) to an articulation than phrase-internal consonants, and that the initial consonants in a higher phrasal prosodic domain are more strengthened than consonants in a lower prosodic domain. As evidence, they observed that Korean consonants /t, t^h, t*, n/ vary in their linguopalatal contact and acoustic correlates (e.g., closure duration, and VOT) as a function of their positions in a prosodic hierarchy. That is, the VOTs of AP-initial stops were longer than those of word-initial (AP-internal) stops, but shorter than those of IP-initial stops. Cho and Keating's study suggests that the presence of a boundary and the type of boundary domain (IP, AP, or word) influence the degree to which consonants are strengthened. If this is so, similar to Jun's prediction, the voicelessness of a word-initial and AP-initial lenis stop can signal the presence of a boundary. However, slightly different from Jun, the voicing qualities of the lenis stops which are located in word-initial but AP-internal positions can signal the presence of a boundary at some degree.

Cho and Keating's proposal on consonant strengthening might become complicated when another type of a strengthening factor (e.g., focus information) is additionally included. For example, if a word or a phrase is accented by using contrastive focus or semantic focus information, the focus information may elicit an additional strengthening effect for the phrase such that the duration of the focused phrase could be lengthened (Klatt 1974, Turk and Sawusch 1997, Junand Lee 1998). If this is so, the longer VOTs of focused initial stops will further facilitate the signals for the presence of a boundary. Note that the strengthening effect due to the focus information is semantic rather than syntactic, whereas the strengthening effect due to the structural information (e.g., domain-initial positions or domain size) is syntactic rather than semantic. It might be an interesting question to ask how the semantic strengthening effect driven by the focus information further contributes to signaling the presence of a boundary in concert with the syntactic strengthening effects driven by the structural information.

In summary, Korean lenis stops are susceptible to the voicing rule that the stops become voiced intervocalically in the middle of a domain but not in the beginning of a domain. As a result, the voiceless quality of lenis stops and the F0s of following vowels can be used to signal whether a boundary of a domain (either a phrase or a word) is present. According to Jun (1993, 1996), the voicing quality of a lenis stop can suggest the presence of an AP boundary, but not a word boundary. Given Cho and Keating's (2001) finding, however, a stop in the word-initial and AP-internal position may survive from being fully voiced and contribute to signaling the presence of a word boundary. Moreover, if a unit beginning with a lenis stop is focused, the strengthening effects by the focus factor may increase the voiceless qualities of the stop by lengthening its VOT. So far, there had been no single study which comprehensively covered these issues simultaneously. The purpose of this study was to investigate which acoustic correlates were produced as lexical or phrasal cues in a potentially ambiguous situation.

2. Experiment

2.1. Participants

Five male and five female talkers of Seoul dialect volunteered in the experiment. All of them were attending the University at Buffalo. Talkers' ages ranged from 25 to 36 years old. They had stayed in an English-speaking country for 28 months on average.[2)]

2.2. Materials

Six quadruples, as presented in (1a-d), were used in this experiment. Each member of the quadruples consisted of a question and an answer. The question sentence was formed as an interrogative pattern (e.g., *Whois in the room*?). The answer sentence was constructed to provide information pertinent to the question (e.g., *My sister is in the room*). Two target syllables (e.g., /ka/ and /paŋ/) were contained in each answer sentence. The first syllable (i.e., /*ka*/)always had a form of a CV in which a velar stop consonant /*k*/ preceded a vowel /*a*/. The second syllable was a form of a CV (i.e., /*ke*/ and /*ku*/) or a CVC (i.e., /*koŋ*/, /*paŋ*/, /*pal*/, and /*pul*/). The consonant in the first syllable was always a velar voiceless stop /*k*/, whereas the initial consonant in the second syllable was a bilabial stop (i.e., /*p*/) or a velar stop (i.e., /*k*/). The experimental stimuli differ in two ways. First, a word boundary, indicated by "|", and an AP boundary, indicated by "{}", were present within the two target syllables, as in (1a) and (1c), or they were absent in the target syllables, as in (1b) and (1d). Second, focus information, indicated by an underline "_", was manipulated, by using a question-answer format. The underlined constituents in answer sentences were designed to be focused because they corresponded to specific information that their corresponding question-sentences inquired (Selkirk1 986, Rooth 1997, Welby 2003). For example, the agent in (1a), *ənni-ka* (meaning sister-NOM), would be produced with focus because it corresponded to the answer for its corresponding question. In this way, the first constituents in sentences

2) In order to avoid the situation that participants' Korean acoustic qualities would be influenced by their English acoustic qualities, participants who stayed in English-speaking countries for more than 3 years were excluded.

like (1a-b) were likely to be focused, while the second constituents in sentences like (1c-d) were likely to be focused. That is, */ka/* was focused but */paŋ/* was not focused in (1a), whereas */ka/* was not focused but */paŋ/* was focused in (1c). Both */ka/* and */paŋ/* in (1b) were focused, but neither */ka/* nor */paŋ/* in (1d) was focused.

(1) a. Question: Who is in the room?
Answer: /ənni-ka/ /paŋ-e/ /is'əjo/
{|ənni-ka|} {|paŋ-e|} {|is'əjo|}
Sister-NOM room-LOC is
(My) sister is in the room.

b. Question: Where is the movie ticket?
Answer: (/p^hyo-ka/) /ənni kapaŋ-e/ /is'əjo/
((|p^hyo-ka|) {|ənni|kapaŋ-e|} {|is'əjo|}
(The movie ticket) Sister-POSS bag-LOC is
(The movie ticket) is in the bag of (my) sister.

c. Question: Where is the sister?
Answer: /ənni-ka/ /paŋ-e/ /is'əjo/
{|ənni-ka|} {|paŋe|} {|is'əjo|}
Sister-NOM room-LOC is
(My) sister is in the room.

d. Question: Isn't the movie ticket in the sister's bag?
Answer: (/p^hyo-ka/) /ənni kapaŋ-e/ /is'əjo/
((|p^hyo-ka|) {|ənni|kapaŋ-e|} {|is'əjo|}
(Themovie ticket) Sister-POSS bag-LOC is
(The movie ticket) is in the bag of (my) sister.

Experimental sentences are normed in various ways. First, to ensure whether focus manipulation created by the experimenter was constructed in a way that it was supposed to be, an off-line norming study was conducted. Forty native Korean speakers participated in this norming study. The participants were asked to underline a constituent of an answer sentence that they would have read with emphas is to answer a given question. Over 90% of responses were consistent with the experimenter's pool of sentence materials. This result suggested that talkers would read constituents with focus as similarly as the norming results showed. Second, to reduce the possibility that sentence production should not be influenced by confounding factors like frequency, the experimental stimuli had to satisfy that the target expressions across conditions had to be equally frequent. The frequency of the one-syllable words and two-syllable words were counted by using Sejong Corpus. Log frequencies of target words were compared across conditions. A paired t-test revealed that log frequencies of six items pairs were not equalized (t (5) = -3.32, $p < .05$) in between one-syllable words (M = 2.79, S.E. = 0.24) and two-syllable words (M = 1.68, S.E. = 0.41). Finally, in order to control the familiarity of a one-syllable word (e.g., /paŋ/ meaninga room) used in (1a) and (1c) and that of a two-syllable word (e.g., /ka.paŋ/ meaning a bag) used in (1b) and (1d), a norming study was conducted.[3] Ten native Korean speakers rated the familiarity of the target expressions when the expressions were presented without context, and when they were presented with context on a 5-point Likert scale (1 corresponding to highly unfamiliar; 5corresponding to highly familiar). Two paired t-tests revealed that the familiarity ratings of six items pairs were equalized between one-syllable words (M = 4.31, $S.E.$ = 0.29) and two-syllable words (M = 4.35, $S.E.$ = 0.16) when the target words were presented without context (t (5) = -.12, p = 1.91), and between one-syllable words (M = 4.17, $S.E.$ = 0.30) and two-syllable words (M = 4.08, $S.E.$ = 0.09) when they were presented with context (t (5) = 0.33, p = 0.76). Moreover, the mean ratings for both conditions were more than 4, suggesting that the words used with or without context were considered fairly familiar. Although one-syllable words occurred more frequently than two-syllable words, all items were equally familiar. Therefore, the six it empairs were included in the production experiment.

2.3. Procedure

The quadruples of six sentence pairs were divided into two blocks: one in which target phrases were focused and the other in which target phrases were not focused. The order of stimuli within a block was randomized across participants. Each talker produced the quadruples of six experimental stimuli, respectively. This resulted in for-

3) /-ka/, a nominative case marker, used in (2a) and (2c), was excluded in the normings for familiarity and frequency control in consideration that this functional marker was highly frequent and familiar.

ty-eight utterances per talker and 480 utterances in total. Each experimental stimulus was presented on an index card on which a question and an answer sentence were written. The answer sentences were highlighted. Talkers were asked to read only answer sentences as if they were answering the corresponding questions.[4] Talkers read the sentences naturally after a brief practice period. If any utterance was judged by the experimenter or talkers to be defective (e.g., speech errors, misplaced pitch accent, etc.), talkers were asked to repeat that sentence. Recordings were done on a Macintosh computer using a microphone. The microphone signal was amplified and digitalized at a 44.1 kHz sampling rate.

2.4. Measurement

In each sentence, the closure durations and the VOTs of the two lenis stops were measured and the F0s for the three vowels which occurred before and after the target lenis stops were measured, as illustrated in Figure 1.[5]

2.4.1. Closure Duration

Acoustic closure durations for all target stops were taken from wave forms and spectrograms. The closure duration of the first stop was named C1 and that of the second stop was named C2. This measure included both voiced and voiceless portions of closure, from the offset of a preceding vowel to the beginning of a release burst. The offset of the vowel was determined by two criteria: 1) when amplitudes dropped and/or 2) when the wave forms and spectrum showed a decrease in high frequency energy. In general, these two measures coincided.

---

4) A reviewer expressed a concern about a methodology in which pragmatic factors like focus information might not be easily realized in the reading context. Note that participants in this study were asked to read target sentences as if they were in a conversation. Although the participants might not produce sentences as naturally as they would do in a real conversation, they did not simply read sentences, at least.

5) Because acoustic qualities indicating the beginning and end of vowels were extremely ambiguous, the accuracy to measure the duration of vowels was not guaranteed. Thus, the total duration of target syllables was not measured. Also, glottal stops were not produced reliably enough for an analysis. The frequency patterns of glottal stops across conditions were not reported.

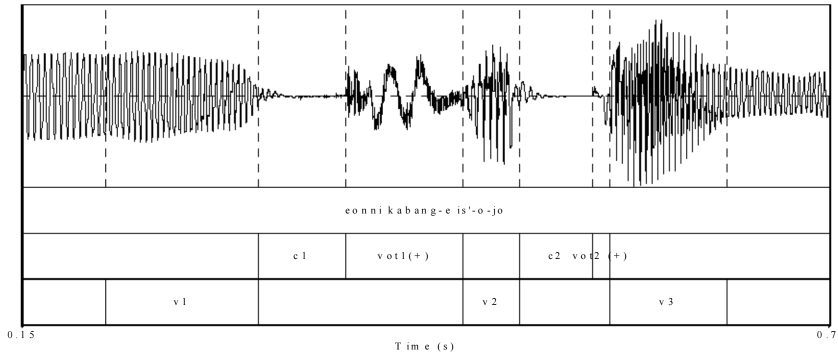


Figure 1. Measurement sample of two stops and vowels.

Note: C1 refers to the closure duration of the first lenis stop and C2 to that of the second lenis stop. VOT 1 refers to voice onset times of the first lenis stop and VOT2 to that of the second lenis stop. V1, V2, and V3 mean the first vowel prior to the first prior to the first lenis stop, the second vowel following the first lenis stop, and the third vowel following the second lenis stop.

2.4.2. VOTs

VOTs for stops were measured from the onset of a release burst until the onset of periodicity (voicing) appeared. If the onset of voicing preceded the stop release, the VOT value was negative. As in the analysis of the closure duration, the VOT of the first stop was named VOT1 and that of the second stop was named VOT2.

2.4.3. F0s

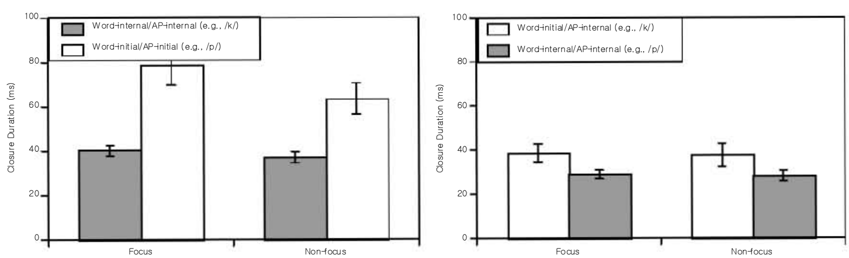
The F0s of three vowels were measured: which immediately preceded the first target stop sound, which immediately followed the first target stop sound and preceded the second stop, and which immediately followed the second target stop sound. The F0s were measured at the onset of voicing and at the midpoint of voicing, and the maximum and minimum F0s of a vowel were also measured, by using the pitch track along with the first harmonic values from a narrow band spectrum. The measurements of a closure duration and a VOT were used to determine the beginning or the end of the vowels. As shown in Figure 1, the end of the first vowel (V1) was determined by the beginning point of the first closure duration (C1). The second vowel (V2) started from the point that the first VOT (VOT1) ended to the point that the second closure duration started (C2). The beginning of the third vowel (V3) was determined by the end of second VOT duration (VOT2). However, the beginning of the first vowels and the end of the third vowels could not be determined in this way. They were determined by the formant patterns represented in a spectrogram. Because all vowels were combined with consonants including nasals and liquids, transitions between constants and vowels were used to split the vowels from their neighboring consonants. Generally, the formants of vowels have good spectral continuity but those of consonants do not, in particular in the end or beginning of nasals or liquids. Thus, the beginnings or ends of the vowels were marked depending on when continuous formant patterns started or ended, respectively.

2.5. Results

Before submitting all data to analyses, some data had to be filtered. One utterance out of 480 utterances was excluded because it was considered to be misarticulated by the experimenter. Because two target consonants and three target vowels were measured in each utterance, there were 958 closure durations, 958 VOTs, and F0s from 1437 vowels, in total. There were sentences with missing data in the analyses of stop consonants and vowels. First, 171 measurements of closure duration and VOT were missed out of 1916 measurements because they lacked the acoustic evidence to reach the condition depicted in Figure 1. These missing data were excluded in the analysis of the closure duration and VOT. This affected about 9% of the data. Second, in the analysis of F0s, 13 out of 1437 vowels were excluded because they were devoiced. This affected less than 1% of the data. Mean measurements for each consonant were submitted to two 2 (Gender) × 2 (Boundary Type) × 2 (Focus) repeated-measures ANOVAs in which participants and items were random variables.

2.5.1. Closure Duration

The results of the closure duration are displayed in Figure 2a-b. First, for the closure duration of the first stop consonant (i.e., C1), a significant main effect of Boundary Type was observed in the analyses of participants and items (F_1 (1, 8) = 14.19, MSe = 65.11, p < .01; F_2 (1, 10) = 13.16, MSe = 83.83, p < .01), indicating that word-initia-land AP-internal stops (M = 38.2 *ms,* S.E. = 4.12) produced longer closure durations than word-internal and AP-internal stops (M = 28.6 *ms,* S.E. = 2.04). Second, the statistical results for the closure duration of the second consonant (i.e., C2) paralleled those of the first consonant. A significant effect of Boundary Type was observed in the analyses of both participants and items (F_1 (1, 8) = 34.81, MSe = 300.35, p < .01; F_2 (1, 10) = 142.88, MSe = 80.12, p < .01), meaning the closure durations of the word-initial and AP-initial stops (M = 71 *ms,* S.E. = 7.19) were longer than those of word-internal and AP-internal stops (M = 38.6 *ms,* S.E. = 2.3). The main effect of Focus was also observed in the analyses of participants and items (F_1 (1, 8) = 6.25, MSe = 136.41, p < .05; F_2 (1, 10) = 80.44, MSe = 12.39, p < .01); the closure durations were longer when the lenis stops were focused (M = 59.43 *ms,* S.E. = 5.25) than when the stops were not focused (M = 50.19 *ms,* S.E. = 4.61). In short, the closure durations were longer when the consonants were in the boundary-initial position than when they were in the boundary-internal position, regardless of whether a boundary was a word or an AP.

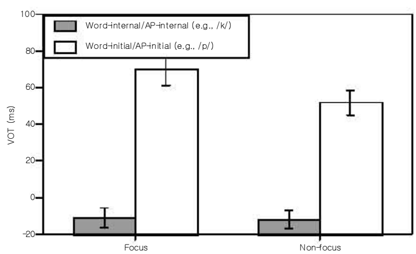


(a) (b)

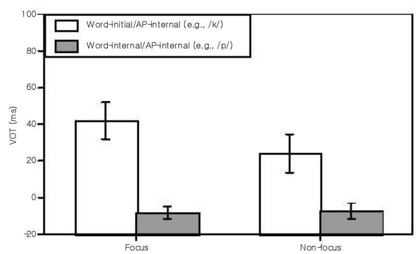
Figure 2. Means of closure duration (ms) of lenis stops for /-ka|paŋ/ type in Figure (2a), and means of closure duration (ms) of lenis stops for /ka.paŋ/ type in Figure (2b).

2.5.2. VOTs

The results of VOTs are displayed in Figure 3a-b. For the VOTs of the first consonants (i.e., VOT1), an interaction effect between Boundary Type and Focus was marginal in the analysis of participants (F_1 (1, 8) = 5.27, MSe = 168.64, p = .05) but significant in the analysis of items (F_2 (1, 10) = 9.01, MSe = 94.44, $p < .05$), suggesting that the VOTs of focused lenis stops became longer when they were in a boundary-initial position. Focus seems to have had no influence on VOT for boundary internal lenis stops. For further analyses, main effect analyses were examined. A significant main effect of Boundary Type was observed in the analyses of participants and items (F_1 (1, 8) = 30.36, MSe = 547.05, $p < .01$; F_2 (1, 10) = 66.69, MSe = 311.74, $p < .01$); the VOTs of word-initial and AP-internal stops (*M* = 33.05 *ms,* S.E. = 9.49) were longer than those of word-internal and AP-internal stops (*M* = -7.71 *ms,* S.E. = 3.79). A main effect of Focus was marginally significant in the analysis of participants (F_1 (1, 8) = 4.47, MSe = 160.35, $p = .07$) but significant in the analysis of items (F_2 (1, 10) = 7.25, MSe = 111.91, $p < .05$), suggesting that the lenis stops showed longer VOTs when they were focused (*M* = 16.9 *ms,* S.E. = 6.14) than they were not focused (*M* = -7.71 *ms,* S.E. = 6.88). Overall, the VOT durations were longer when the lenis stops were in a boundary-initial position than when they were in a boundary-internal position.



(a)



(b)

Figure 3. Means of VOTs (ms) of lenis stops for /-ka|paŋ/ type in Figure (3a), and means of VOTs (ms) of lenis stops for /ka.paŋ/ type in Figure (3b).

For the VOTs of the second lenis stops (i.e., VOT 2), an interaction between Boundary Type and Focus was significant only in the analysis of items (F_2 (1, 10) = 7.03, MSe = 106.51, $p < .05$), meaning that the VOTs of focused lenis stops lengthened only when they were in a boundary-initial position. For further analyses, main effects of each factor were tested. As in the first VOT analysis, a significant effect of Boundary Type was observed in the analyses of participants and items (F_1 (1, 8) = 238.50, MSe = 218.77, $p < .01$; F_2 (1, 10) = 330.59, MSe =

181.25, $p < .01$). Similar to the first VOT pattern, the VOTs of word-initial and AP-initial stops ($M = 60.79$ *ms,* S.E. = 5.71) were longer than those of word-internal and AP-internal stops ($M = -11.44$ *ms,* S.E. = 5.15). Also, the main effect of Focus was significant only in the analysis of items (F_2 (1, 10) = 26.13, MSe = 43.84, $p < .01$); focused stops ($M = 29.43$ *ms,* S.E. = 5.81) showed longer VOTs than non-focused stops ($M = 19.92$ *ms,* S.E. = 5.63).

To compare the VOTs of the lenis stops by a structural hierarchy as examined in Cho and Keating (2001)'s study, 4 (Stop Type: word-initial and AP-internal (e.g., /k/ in /ka.paŋ/), word-internal and AP-internal (e.g., /p/ in /ka.paŋ/), word-initial and AP-initial (e.g., /p/ in /-ka|paŋ/), word-internal and AP-internal (e.g., /k/ in /-ka|paŋ/) × 2 (Focus: Focus vs. Non-focus) factorial ANOVAs were conducted in which participants and items were random variables. The results of VOT analyses by the structural hierarchy are illustrated in Figure 4. The factorial ANOVAs revealed that an interaction between Stop Type and Focus was only significant by items (F_2 (3, 88) = 2.75, MSe = 537.95, $p < .05$). However, the interaction occurred in the pair of focused or non-focused word-internal and AP-internal stops which was not the interesting pair for this study. The main effects of Stop Type were observed in both analyses of participants and items (F_1 (3, 72) = 48.68, MSe = 23889.17, $p < .01$; F_2 (3,88) = 142.41, MSe = 27893.24, $p < .01$). The main effects of Focus were marginal in the analysis of participant (F_1 (1, 72) = 3.29, MSe = 1615.88, $p = .07$), but significant in the analysis of items (F_2 (1,88) = 9.92, MSe = 1942.88, $p < .01$). Importantly, the VOTs were numerically ordered by a structural hierarchy. The VOTs of the word-initial and AP-initial stops were significantly longer than those of the word-initial and AP internal stops both when they were focused (t *(18)* = -2.22, $p < .05$) and when they were not focused (t *(18)* = -2.3, $p < .05$). The VOTs of the word-initial and AP-internal stops were longer than those of the two word-internal and AP-internal stops, when they were focused (t (18) = 4.84, $p < .01$; t (18) = 2.77, $p < .05$) and when they were not focused (t (18) = 3.2, $p < .05$; t (18) = -8.0, $p < .01$). These results replicated Cho and Keating (2001)'s study.

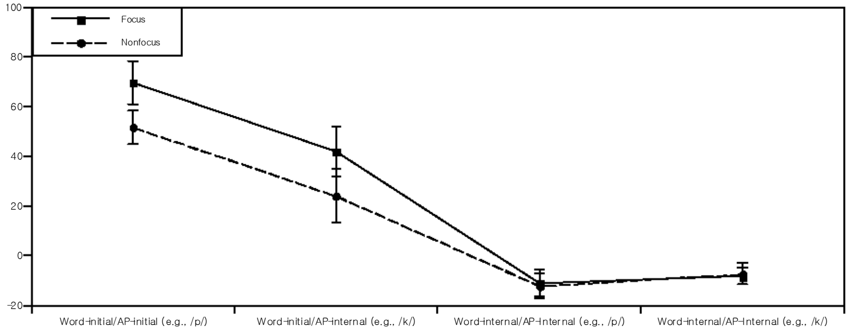


Figure 4. VOT hierarchy of four lenis stops.

2.5.3. F0s

Four F0s measurements of all conditions across three vowels are depicted in Figure 5. Figure 5 indicates the F0 patterns of the three target vowels, indicated by V1, V2, and V3, by connecting the average F0 of each vowel. The patterns of maximum F0s, minimum F0s, and F0s measured at the mid position were similar, but the pattern of F0s measured at the onset position was not. For this paper, the statistics of F0s measured at the mid position were reported, as depicted in Figure 5.

1) At the first vowel (e.g., /i/ in /ən.ni/) prior to the first stop lenis stop, the main effect of Boundary Type was observed in the analyses of participants and items (F_1 (1, 8) = 15.41, MSe = 93.76, $p < .01$; F_2 (1, 10) = 14.81, MSe = 113.76, $p < .01$). The F0s of vowels preceding word-internal and AP-internal lenis stops (M = 215.41 *Hz,* S.E. = 8.05) were lower than those preceding word-initial and AP-internal stops (M = 227.43 *Hz,* S.E. = 8.55). The Focus effects were also observed in the analyses of participants and items (F_1 (1, 8) = 6.94, MSe = 79.83, $p < .01$; F_2 (1, 10) = 59.96, MSe = 10.43, $p < .01$), showing that the F0s in the Non-focused condition (M = 217.7 *Hz,* S.E. = 7.99) were lower than those in the focused condition (M = 225.15 *Hz,* S.E. = 8.56).

2) At the second vowel (e.g., /a/ in /ka/) following the first stop (e.g., /k/), a significant interaction between Boundary Type and Focus was observed in the analyses of both participants and items (F_1 (1, 8) = 5.43, MSe = 128.56, $p < .05$; F_2 (1, 10) = 17.89, MSe = 47.9, $p < .01$). The F0s of vowels following word-initial and AP-internal stops were lower when they were focused (M = 169.18 *Hz,* S.E. = 6.68) than when they were not focused (M = 175.9 *Hz,* S.E. = 8.77), whereas, the F0s of vowels following word-internal and AP-internal stops were

higher when they were focused (*M* = 217.75 *Hz,* S.E. = 8.98) than when they were not focused (*M* = 207.77 *Hz,* S.E. = 7.98). As a following step, analyses for main effects were conducted. The significant main effect of Boundary Type was observed in the analyses of participants and items ($F_1(1, 8) = 42.07$, MSe = 384.53, $p < .01$; F_2 (1, 10) = 114.01, MSe = 169.49, $p < .01$). The F0s of vowels were lower when they followed word-initial and AP-internal stops (*M* = 172.54 *Hz,* S.E. = 7.49) than when they followed word-internal and AP-internal stops (*M* = 212.76 *Hz,* S.E. = 8.30).

3) At the third vowel (e.g., /a/ in /paŋ/) following the second stop consonant (e.g., /p/), the F0s by Boundary Type were significantly different in the analyses of both participants and items (F_1 (1, 8) = 43.16, MSe = 65.47, $p < .01$; F_2 (1, 10) = 32.32, MSe = 104.18, $p < .01$), indicating that the F0s of vowels were lower when they followed word-initial and AP-initial stops (*M* = 168.46 *Hz,* S.E. = 4.46) than when they followed word-internal and AP-internal stops (*M* = 185.27 *Hz,* S.E. = 6.16). Also, the F0s of vowels in the Non-focus condition (*M* = 175.43 *Hz,* S.E. = 5.27) were marginally lower than those in the Focused condition (*M* = 178.3 *Hz,* S.E. = 5.23) in the analysis of participants (F_1 (1, 8) = 5.11, MSe = 16.1, $p = .05$) but not in the analysis of items (F_2 (1, 10) = 2.38, MSe = 27.0, $p = 1.54$).

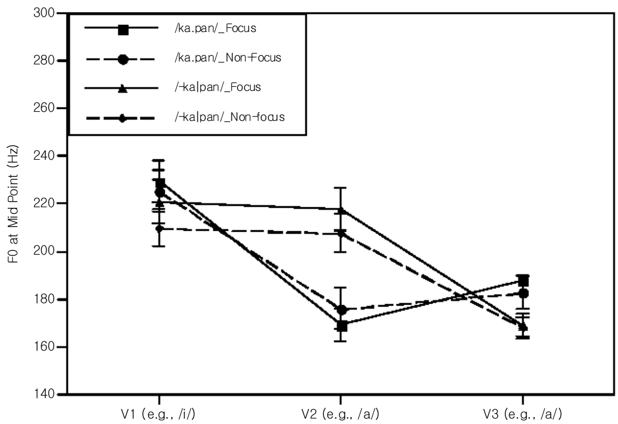


Figure 5. Means of F0s (Hz) at mid position across three vowels.
Note. V1 refers to the first vowel prior to the first lenis stop, V2 refers to the second vowel following the first lenis stop, and V3 refers to the third vowel following the second lenis stop.

3. Discussion

3.1. Which Acoustic Correlates Do Talkers Produce as Physical Cues to Signal the Presence of a Boundary for Listeners?

The acoustic correlates (i.e., closure duration, VOT, and F0) were realized differentially as a function of the presence of a boundary.[6] As predicted, boundary-initial lenis stops were realized to be voiceless (i.e., longer closure duration and VOT), but boundary-internal lenis stops were voiced (i.e., shorter closure duration and VOT). As for /-ka|paŋ/, the lenis stop voicing rule proposed by Jun (1993, 1996) was supported. That is, when a lenis stop was word-initial and AP-initial, the stop (e.g., /p/ of /paŋ/) was realized as voiceless. When a lenis stop was word-internal and AP-internal, the stop (e.g., /k/ of /-ka/) was realized as voiced. However, Jun's voicing rule did not seem to be fully applied for /ka.paŋ/. When a lenis stop was word-initial and AP-internal, the stop (e.g., /k/ of /ka/) was nearly voiceless and when a lenis stop was word-internal and AP-internal, the stop (e.g., /p/ of /paŋ/) was voiced.

Recall Jun's (1993, 1996) lenis stop voicing rule that lenis stops are voiced within an AP but not at the beginning of an AP. This rule means that the voicing quality of lenis stops can signal the presence of an AP boundary but not a word boundary. However, the results of this study provide counter evidence such that a word-initial stop, even though it is in the middle of an AP, tends to be voiceless and could cue the presence of a word boundary, in particular when they were focused. On one hand, the results of this study provide evidence that the domain for the lenis stop voicing might not be an AP. On the other hand, as Jun (1993, 1996) discussed, focus may contribute to transforming a boundary prior to the word-initial and AP-internal stops from a word to an AP, by lengthening the segmental duration of the boundary. If this is so, when word-initial but AP-internal stops are focused, they tend to be voiceless. The question of what should

6) Note that participants produced all target sentences at a casual speed as if they were talking. As a reviewer pointed out, the acoustic qualities that were observed in this study might not occur if participants produce sentences at a rapid speed. This is true, in part. The voicing qualities observed at a causal-rate speed disappeared in a preliminary study in which participants were asked to produce sentences at a rapid speed.

be the domain for lenis stop voicing in Korean needs further studies.

Importantly, note that the VOT data of this study replicated the results of Cho and Keating (2001). The lenis stop which occurred at the initial position of a word and AP boundary showed the longest VOTs , followed by the VOTs of word-initial and AP-internal stops. The VOTs of word-internal and AP-internal stops were the shortest, and these stops appear to have been produced as voiced.

Second, the F0s of vowels which followed lenis stops were lower when the stops were in the initial position of a boundary than when they were in the internal position of a boundary. At the second vowel (e.g., /a/ in /ka/ or/-ka/), F0s of /a/ following a word-initial and AP-internal lenis stop (i.e., /k/in /ka/) were lower than those of /a/ following a word-internal and AP-internallenis stop (i.e., /k/ in /-ka/). Similarly, at the third vowel (e.g., /a/ in /-paŋ/ or /paŋ/), F0s of /a/ following a word-initial and AP-initial stop (i.e., /p/ in /paŋ/) were lower than those of /a/ following a word-internal and AP-internal stop (i.e., /p/ in /-paŋ/). These results suggest that the low F0 can function as an index which can signal that there is a boundary prior to the preceding stop, regardless of whether the boundary is a word or an AP.

In sum, the closure duration and VOT results suggest that the voicing qualities of lenis stops could cue the presence of a boundary at both phrasal and word levels for listeners. In addition the acoustic correlates associated with lenis stop consonants, the F0 of vowels following the stops were realized differently depending on the presence and location of a boundary. These acoustic correlates were produced not only in the word level but also in the phrasal level.

3.2 How Does the Focus Information Influence the Acoustic Correlates Associated with Lenis Stops, in Particular When a Boundary is Present?

The focus information facilitated the acoustic correlates of lenis stops unless the overall segmentation in a sentence was disturbed. In the analysis of VOTs between word-initial and AP-internal stops, and word-initial and AP-initial stops, no interaction was observed between Focus and Boundary Type. Word-initial and AP-initial stops showed longer VOTs than word-initial and AP-internal stops, regardless of the focus information. These results suggest that the focus information did not significantly contributed to forming a prosodic structure associated with lenis stops. Moreover, the results of this study indicate that the syntactic (or structural) information influenced by the boundary type is more dominant than the semantic information influenced by the focus factor. Jun (2002) has shown that the syntax information operated more strongly than the focus information. She reasoned that talkers might try not to combine a focused word with a following word if the following word belonged to a different syntactic phrase. In this way, syntactic structure could be clearly preserved. A similar manner of speaking was also observed in this experiment. The talkers in this experiment seemed to try avoiding possible ambiguity that may have happened if the focus had been emphasized over the structural information. In sum, the focus information plays an important role in segmentation by facilitating or suppressing acoustic correlates associated with lenis stops in order to maximize signaling the presence of a boundary. However, the influence from the focus information is less dominant than that from the structural information.

4. Conclusion

This study aimed to show that acoustic correlates associated with Korean lenis stops would be realized differently depending on the presence of a boundary in the word and/or phrase levels of an utterance. The results indicated that the voicing quality of lenis stops, and the F0s of the vowels following initial lenis stops were produced differently which may signal the presence of a word and/or phrase boundary to the listener. Furthermore, focus information was likely to maximize the acoustic correlates related to lenis stops, by lengthening lenis stops and by lowering the F0 of the vowels following lenis stops at the initial position of a boundary. The results showed that the focus information was not as dominant as the structural information. In conclusion, this study showed that the acoustic qualities were consistently correlated with the position of word and phrasal boundaries. The results from this production study do not guarantee that listeners will also exploit the same acoustic qualities as important cues to identify the presence of a boundary. A future perception study will investigate which physical cues listeners might use in perceiving the presence of a boundary.

References

Beckman, Mary E., and Pierrehumbert, Janet. (1986). Intonational Structure in Japanese and English. *Phonology Yearbook* 3, 255-309.

Cho, Taehong. (1996). *Vowel correlates to consonant phonation: an acoustic-perceptual study of Korean obstuents.* MA thesis, University of Texas at Arlington.

Cho, Taehong., Jun, Sun-Ah, and Ladefoged, Peter. (2002). An acoustic and aerodynamic correlates of Korean stops and fricatives. *Journal of Phonetics* 30, 193-228.

Cho, Taehong and Keating, Patricia A. (2001). Articulatory and acoustic studies on domain-initial strengthening in Korean. *Journal of Phonetics* 29, 155-190.

Christie, William M. (1974). Some cues for syllable juncture perception in English. *Journal of the Acoustical Society of America* 55, 819-821.

Culter, Ann and Otake, T. (1994). Mora or phoneme? Further evidence for language-specific listening. *Journal of Memory and Language* 33, 824-844.

Fougeron, Cecile. (2001). Articulatory properties of initial segments in several prosodic constituents in French. *Journal of Phonetics* 29, 109-136.

Fougeron, Cecile, and Keating, Patricia. A. (1997). Articulatory strengthening at edges of prosodic domains. *Journal of the Acoustical Society of America* 101, 3728-3740.

Gow, David W. and Gordon, Peter. C. (1995). Lexical and prelexical influences on word segmentation: Evidence from priming. *Journal of Experimental Psychology: Human Perception and Performance* 21, 344-359.

Hayashi, Wendy, Hsu, Chai-Shune, and Keating, Patricia A. (1999). Domain-initial strengthening in Taiwanese: a follow-up study. *UCLA Working Papers in Phonetics* 97, 152-156.

Hsu, Chai-Shune and Jun, Sun-Ah. (1998). Prosodic strengthening in Taiwanese: syntagmatic or paradigmatic? *UCLA Working Papers in Phonetics* 96, 69-89.

Jun, Sun-Ah. (1993). The Phonetics and Phonology of Korean Prosody, *Unpublished Ph. D. dissertation*. Columbus: The Ohio State University.

Jun, Sun-Ah. (1996). *The Phonetics and Phonology of Korean Prosody: intonational phonology and prosodic structure*. New York: Garland Publishing Inc.

Jun, Sun-Ah. (2000). K-ToBI (Korean ToBI) labelling conventions *UCLA Working Papersin Phonetics*.

Jun, Sun-Ah. (2002). *Syntax over focus.* Paper presented at the International Conferenceon Spoken Language Processing (ICSLP), Denver, Colorado.

Jun, Sun-Ah. (2003). Prosodic phrasing and attachment preferences. *Journal of Psycholinguistic Research* 32.2, 219-249.

Jun, Sun-Ah. (2006). Intonational Phonology of Seoul Korean Revisited In V. T. J. and K. Jones (Eds.), *Japanese-Korean Linguistics* (Vol. 14). Stanford: Center for the Study of Language and Information.

Jun, Sun-Ah, and Lee, Hyuckjun. (1998). *Phonetic and phonological markers of-contrastive focus in Korean.* Paper presented at the 5th ICSLP, Sydney, Australia.

Keating, Patricia A. (1984). Phonetic and phonological representation of stop consonant voicing. *Language* 60, 286-319.

Keating, Patricia A., Linker, Wendy, and Huffman, Mary. (1983). Patterns in allophone distribution for voiced and voiceless stops. *Journal of Phonetics* 11, 277-290.

Kim, Mi-Ryoung and Beddor, Patrice. S. (2002). "The contribution of consonantal and vocalic information to the perception of Korean initial stops." *Journal of Phonetics* 30, 77-100.

Klatt, Dennis. H. (1974). The duration of[s] in English words. *Journal of Speech and Hearing Research* 17, 51-63.

Lehiste, Ilse. (1960). An acoustic-phonetic study of internal open juncture. *Phonetica* 5, 1-54.

Lehiste, Ilse. (1972). The timing of utterance and linguistic boundaries. *Journal of the Acoustical Society of America* 51, 2018-2024.

Liberman, A. M., Harris, K. S., Hoffman, H. A., and Griffith, B. C. (1957). The discrimination of speech sounds within and across phoneme boundaries. *Journal of Experimental Psychology* 54, 358-368.

Lieberman, Philip. (1967). *Intonation, Perception, and Language*. Cambridge, MA: The MIT Press.

Norris, Dennis, McQueen, James. M., and Cutler, Ann. (1995). Competition and segmentation in spoken word recognition. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 21, 1209-1228.

Rooth, Mats. (1997). Focus. In S. Lappin (Ed.), *The Handbook of Contemporary Semantic Theory*, 271-297. Oxford, UK: Blackwell Publishers.

Selkirk, Elisabeth. (1986). On derive domain in sentence phonology. *Phonology Yearbook* 3, 371-405.

Sohn, Homin. (1999). *The Korean Language*. Cambridge: Cambridge University Press.

Trager, George. L. and Bloch, Bernard. (1941). The syllabic phonemes of English. *Language* 17, 233-246.

Welby, P. (2003). Effects of pitch accent position, type, and status on focus projection. *Language and Speech* 46.1, 53-81.

Hongoak Yun
Interaction Science Research Center
Sungkyungkwan University
Jongno-gu, Myungnyun-dong 3-ga, Seoul, 110-745
E-mail: oaktreepark@gmail.com

Received: March 10, 2013
Revised version received: April 8, 2013
Accepted: April 14, 2013