Corpus-based study of durational adjustment in Korean*

Tae-Jin Yoon
(Sungshin Women's University)

Yoon, Tae-Jin. 2015. Corpus-based Study of Durational Adjustment in Korean. *Studies in Phonetics, Phonology and Morphology* 21.2. 279-295. Research in spoken language has shown that an individual speech sound can be produced differently in different structural contexts in spite of presumably the same lexical representation. Cues for a linguistically organized framework can be found in these contextually determined variant phonetic characteristics. This paper examines the patterns of duration adjustment of 'ta', 'ka', and 'neun' in Standard Korean when these segmental sequences occupy different constituents within a word or a phrase. The results indicate that the phonetic manifestation of the utterance boundary is conditioned by linguistically organized structure. They also suggest that the failure of earlier studies to find the effect of phrase-initial lengthening may be due to the greater variation in that position. The findings observed in the study may be a reading characteristic observed in the reading style by male speakers, which warrants detailed phonetic analyses for more speakers of different genders and more speaking styles as well as speaker-dependent variations. (**Sungshin Women's University**)

Keywords: Boundary-related duration, preboundary domain, syllable nucleus, Corpus of Standard Korean

1. Introduction

It is well-known that speech is a continuous signal without consistent demarcation of the words. Listeners must extract the component words from each speech signal in order to understand the speaker's message. Although adult speakers of a language are often assumed to rely heavily on direct identification of known lexical item, infants who have to first learn the lexicon of their language do not have a ready access to it (Jusczyk 1997). The most prominent cue for word segmentation is a silent pause between words if it is present. A silent pause can be used as a reliable cue facilitating segmentation of speech stream (Kim 2004, Seidl and Johnson 2006). Infants were sensitive, not only to the presence of silent pause itself, but also to the location of the pause in the sentence. Experiments in Hirsh-Pasek et al. (1987) illustrate that Infants would listen longer to the utterances, when the pauses coincided with major syntactic phrases, than those that contained mismatched pauses. However, pause itself cannot be a very reliable cue for a word boundary because speakers do not put a pause after every word. Finding words in sentences is made difficult by the frequent absence of

* This work was supported by the National Research Foundation of Korea Grant funded by the Korean Government (NRF-2014S1A5A8013266). Earlier version of the paper has been presented at the 2013 Phonology Conference. I am grateful to three anonymous reviewers, whose comments and suggestions clarified and strengthened the paper.

obvious acoustic markers (e.g., silent pause) at word boundaries. Thus, how listeners find word boundaries in spoken utterances in the absence of consistent acoustic cues has been a long-standing question (Christophe et al. 2003, McQueen and Cutler 2010).

Previous studies provide ample support that various prosodic cues pertaining to both lexical and post-lexical prosodic units are employed in facilitating listeners' word segmentation to a large extent. For example, Echols et al. (1997) observed that, when 9-month-olds were presented with weak-strong-weak nonce words containing a 250ms silent pause either before or after the strong syllable, the infants listened longer to the stimuli in which the pause occurred before the strong syllable. Infants preferred the configuration of speech in which strong syllables were the onset rather than the offset of a unit.

In addition to silent pauses, edges of an utterance are regarded as prestigious position. Seidl and Johnson (2006) report that infants segment words from the edges of an utterance (final and initial positions accompanying adjacent silent pauses) more readily than from the utterance medial position. In addition to the lexical prosodic shape, phrase-boundary cues facilitate speech segmentation. Christophe et al. (2003) found that French listeners (both infants and adults) could discriminate the difference between a bisyllabic unit (C1V1C2V2) extracted from a long word (e.g., [mati] from 'cli**maiti**se') and the same sequence of syllables extracted from two consecutive words (e.g. [mati] from 'panora**ma ty**pique'). For the latter case, a phonological phrase boundary can possibly occur between the two words.

While phrase-final lengthening is well documented (Beckman and Edwards 1990, Edwards et al. 1991, Wightman et al. 1992, Fletcher 2010), convincing evidence for word-final lengthening in non-phrase-final position is hard to obtain (Turk and Shattuck-Hufnagel 2000, Kim 2004). Using an artificial language-learning paradigm, Saffran, Aslin, and Newport (1996) found that word-final lengthening significantly increases English listeners' segmentation ability. The finding appears to provide evidence for word-final lengthening, maintaining that the right-edge of a word be a prestigious position. However, the result needs to be carefully interpreted (Kim 2004). It is often unclear from published reports whether words claimed to show word-final lengthening are in fact phrase-medial, as opposed to phrase-final. If the target word is located at the end of a prosodic constituent like an intonational phrase, the observed lengthening may be associated with phrase-final rather than word-final position. The original vowel in Saffran et al. (1996) is lengthened by 100ms (or 36.1%) from 277ms to 377ms. Given the amount of lengthening, infants may have been sensitive to phrase-final lengthening rather than word-final lengthening (Kim 2004). Pieces of evidence for non-phrase-final word-final lengthening can be found in studies using nonce-words. Nakatani et al.'s (1981) study of reiterant renditions of adjective-noun sequences, like *maMA MA* for 'absurd day' or *MAma ma* for something like ‘sunny day’, showed longer stressed CV syllable durations in word-final position than in non-word-final positions even for the target adjectives, which are not likely to be phrase-final. Beckman and Edwards (1990) also found word-final lengthening for words in phrase-medial position by comparing, for example, the duration of schwa /ə/ that occur in *poppa # pose* (word-final) and that is produced in *pop # oppose* (word-initial). These reports show lengthening of word-final syllables independent of lengthening due to phrase-final position.

Domain initial strengthening is another boundary-induced phenomenon. Fougeron and Keating (1997) observed acoustic durations of the English target consonant /n/ in the left-edge of each of five prosodic domain (U, IP, PP, W, S). They reported that acoustic durations were longer in higher prosodic domains in general. The alveolar nasal aligned with the initial position of Utterance, Intonational Phrase (IP), and Phonological Phrase was the longest in acoustic duration. In Cho and Keating (2001), a consonant's oral articulation was observed to be strengthened and lengthened. They also ascribed the strengthening and lengthening of a consonant to the strong position in the hierarchically layered prosodic domains. The duration of a word-initial consonant constriction is longer than that of a word-medial consonant in the same stress environment. But this word-level effect is less consistent across speakers than initial lengthening at the phrase level (Oller 1973, Jun 1993, Dilley et al. 1996, Fougeron and Keating 1997, Turk and Shattuck-Hufnagel 2000, Fougeron 2001). Jun (1993) proposed ‘a hierarchy of strength of prosodic position’ to account for her finding that voice onset time (VOT) of a Korean consonant is greatest when phrase-initial, next greatest when word-initial but phrase-medial, and least when word- and phrase-medial. Dilley et al. (1996) examined 3709 word-initial vowels produced by professional Radio News announcers and reported that speakers were more likely to glottalize word-initial vowels when those vowels occur at the beginning of a new intonational phrase boundary, with more occurrence of glottalization in full intonational boundaries than in intermediate phrase boundaries. Thus, it is likely that the word-initial consonants are affected by the boundary-induced strengthening. On the other hand, Fletcher (2010) cites Fougeron (1998) who found no evidence of lengthening for the vowel following a phrase-initial consonant in French. Likewise, Byrd (2000) and Cho and Keating (2001) found no consistent evidence of lengthening beyond the phrase-initial consonant. The only effect of the post-consonantal vowel duration is when the vowel occur after /t^{*}/. At this environment, the vowel is observed to be generally longer in a higher prosodic domain (Cho and Keating 2001: 174). These findings lead some researchers to hypothesize that initial lengthening may be localized on initial consonants.

This paper aims to investigate duration adjustments of post-consonantal vowel segments that occur in different phrasal positions in sentences using a large-scale corpus. Experiments with laboratory speech can reveal regular underlying mental patterns that otherwise would remain obscure due to the extreme variability of speech sound durations in connected speech. Whether and how these regularities show up in real speech depends on many factors. As Nooteboom (1997) asserts, “a first reassurance we need is whether similar temporal regularities can be demonstrated for real words and phrases, and whether such regularities can be shown to be part of what language users (implicitly) know about they way words and phrases in their language should sound.” Therefore, the current research aims to find the temporal adjustment at different phrasal positions using a large-scale corpus of connected speech.

It is well-known, as we saw above, that prosodic hierarchy of utterances plays an important role in duration adjustment. However, prosody transcription is a complex task that incorporates the transcriber's auditory impression of prominence and phrasal juncture with visual inspection of the graphical speech display (including at least the pitch track, waveform, and spectrogram), and requires specialized training. It is also a slow task, taking anywhere from 10 to 100 times the duration of the speech recording (Cole and Hasegawa-Johnson 2012), and requires first having a reliable time-aligned word transcription. Instead of using manually transcribed prosodic information, I will analyze the duration of the target segments based on their position in a sentence. These structurally-defined positions are directly related to a loosely defined prosodic structure. Consider the following example in (1).

(1) Chulsu-ka hankaunde iss-neun kagu-lul sa-ss-ta.
Chulsu-Nom in the middle being-AND furniture-ACC buy-PAST-DECL
‘Chulsu bought a piece of furniture which was in the middle’

There are a number of different ways of rendering the sentence in (1), including (2). In (2), the parentheses indicate prosodic phrasing.

(2) a. {Chulsu-**ka**} {han**ka**unde iss-neun **ka**gu-lul sa-ss-ta.}
b. {Chulsu-**ka**} {han**ka**unde iss-neun} {**ka**gu-lul sa-ss-ta.}
c. {Chulsu-**ka** han**ka**unde iss-neun **ka**gu-lul sa-ss-ta.}

The sequence ‘ka’ can occur at the beginning or end of a phonological word[1] within a sentence, and it can be accompanied by adjacent silent pause, as in (2a) and (2b). If a silent pause precedes ‘ka’, as the initial syllable ‘ka’ in ‘kagu-lul’ in (2b), it is likely that ‘ka’ forms a left-edge of either AP or IP. If a silent pause follows ‘ka’, as the final syllable ‘ka’ in ‘Chulsu-ka’ in (2a) and (2b), it is very likely that ‘ka’ coincides with the right-edge of IP, because the right-edge of AP is not followed by silent pause. When no silent

[1] In the Hangul writing system, sets of orthographic phoneme segments are grouped into orthographic syllables, and sequences of orthographic syllables are grouped into orthographic word. Orthographic words in the Hangul system is called ‘*eojeol*’ (K. Yoon 2005). In this paper, we adopt an ‘*eojeol*’ as a unit of (phonological) word.

pause is present adjacent to ‘ka’, the target segments ‘ka’ can occupy word-initial, word-medial, or word-final positions, respectively, depending on how ‘ka’ is aligned with the phonological words. For example, ‘ka’ in ‘hankaunde’ is in word-medial position in (2), ‘ka’ in ‘kagu-lul’ in (2a) is in word-initial position, and ‘ka’ in ‘Chulsu-ka’ in (2c) is in word-final position. Therefore, by analyzing the durational patterns of ‘ka’ in these positions, we can observe indirectly what the durational adjustment will be in the Korean prosodic structure.

With these indirect mapping of a segment within a prosodic hierarchy and a segment within a phrase, we can hypothesize the followings: First, if the initial lengthening is limited to consonantal duration, then it is not likely that we will be observing the lengthening of duration of the syllabic nucleus in post-silence target sequence of the CV form (e.g., ‘ka’). Second, the phrase-final sequences of target segments that is followed by a silent pause is likely to be an IP, thus, we will observe longer nucleus duration due to the phrase-final lengthening effect. If the target sequences that are not followed by any silent pause are indeed a prosodic word-final, and the prosodic word-final coincides many times with AP, then we are not likely to see any AP-induced lengthening effect. What is assumed here is that even though it is possible that IP is signaled by lengthening without silent pause, the occurrence of IP signaled only by phrase-final lengthening is relatively quite limited in its frequency than the occurrence of IP signaled both by phrase-final lengthening and the presence of silent pause or by the presence of silence without phrase-final lengthening. Finally, with regard to word-initial lengthening, if Fougeron’s (1998) findings for French hold for Korean, i.e., if word-initial lengthening affects a word-initial consonant but not the following vowel, then this mechanism predicts that the nucleus duration in the CV-formed target sequence would show no such durational difference between word-/phrase-initial CV sequence and word-medial CV sequence.

2. Feature Extraction from Corpus

The data used for this study was drawn from “A Speech Corpus of Reading-Style Standard Korean,” created around 2003 and distributed by the National Institute of the Korean Language (NIKL) in 2007. The corpus contains 930 sentence types from 19 different well-known short stories and essays read by 120 speakers. Eighty speakers, in his or her 20’s (20 male, 20 female), 30’s (20 male), or 40’s (20 female), read all of the 930 sentences. Speakers 50’s or older (20 male, 20 female) read 404 sentence types. In total, there are 8,622 phrases that consist of 779,300 characters in the syllable form of (C(C))V((C)C), where V can be a monophthong or a diphthong. The sheer size of the data allows researchers to investigate phonetic and phonological variation both within and across speakers. The details of the structure and content of the corpus can be found in T. Yoon (to appear), Yoon and Kang (2014).

The current study is based on data from all 19 stories taken from 10 male speakers in his 20's. In order to study the durational adjustments depending on the position within a phrase, I chose 'ta', 'ka', and 'neun'. The decision is based on the frequency of these three syllables, which suits for statistical analysis for their positional effects. The decision is also based on their role of grammatical markers when they occur at the end of prosodic words such as sentence ending marker for 'ta', nominative marker for 'ka', and either topic marker or adnominal marker for 'neun'. The corpus contains 958 forms of syllable (e.g., 'sis(싯)', 'ryo(료)', 'sip(십)', 'eod(얻)', etc.). Out of the syllabic forms, the most frequently observed monosyllabic morpheme is 'i' (1092 out of 24548 number of syllabic tokens, or 3%), followed by 'ta' (919), 'neun' (689), 'eul' (578), 'go' (536), and 'ka' (517). 'ka' is the sixth most frequent syllabic type and forms a minimal pair with 'ta'. 'neun' is the third most frequent syllabic type and is a closed syllable, unlike an open syllable as 'ka' and 'ta'. Even though 'i' is the most frequent in the corpus, it is not used because of difficulties in segmenting due to other phonological phenomena such as high-vowel devoicing. In sum, out of these frequently occurring syllabic types, I chose 'ka', 'ta', and 'neun' for the current study.

Another reason of choosing these three sequences is due to its role in morpho-syntax. Korean is predicate-final, sharing the typical properties of predicate-final languages, such as Japanese, and very different from languages like English and Chinese. Thus, the predicate (verb or adjective) expression always comes at the end of a clause, whether the clause is a main (matrix) or embedded one. Case, conjunctive, and delimiter particles are all postpositional; an 'auxiliary' predicate follows the 'main' predicate; tense-aspect and modality elements follow the verbal stem in the form of inflectional suffixes; and various sentence or clause types are expressed by a sentence or clause ender which is suffixed to the predicate (H. Sohn 1999). In Korean, morphological marking and syntactic marking are crucial instruments for the structuring of information. For example, '-ka' as a nominative marker denotes the agent of the sentence, as in (3a), and '-neun' in a phrase final position denotes a topic marker and marks the topic of the sentence, as in (4). '-ta' is a neutral sentence-final marker, as in (3) and (4).

(3) a. sonyeo-ka k'oc-eul sa-ss-ta
girl-NOM flower-ACC buy-PAST-DECL
'The girl bought flowers'
b. sonyeon-neun gaeuldug-e anj-ass-ta
boy-TOP brook-LOC sit-PASS-DECL.
'As for the boy, he sat on the brook.'

In addition to topic marker, '-neun' can be used as an adnominal clause ending, representing present tense, as in (4), or complementizer, as in (5). (S. Sohn 1995)

(4) John-i ilk-neun chayk-ul Mary-ka ilk-ess-ta
J.-Nom read-Adn book-Acc Mary-Nom read-Past-Dec.
'Mary read the book that John was reading.'
(5) John-i hakkyo-ey ka-neun] kes-ul Mary-neun al-ass-ta
J. -Nom school-LOC go-COMP thing-ACC M.-Top know-Past-Dec
'May knew that John was going to school.'

The target syllables 'ta', 'ka', and 'neun' are, as stated above, chosen based on their frequency of occurrence. Figure 1 illustrates the way the target sequence is classified based on their position within a phrase or a word.

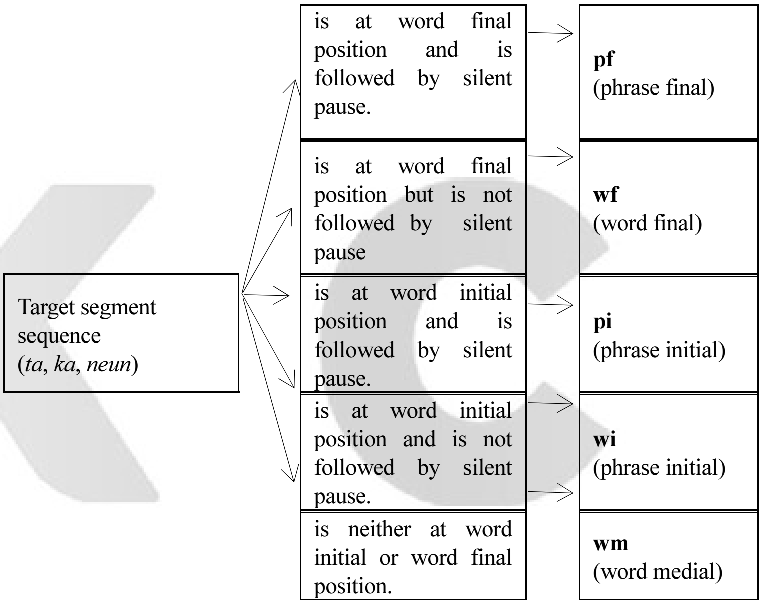

Figure 1. Flow chart indicating the position of target sequence 'ta', 'ka', and 'neun' within a phrase. Note that no instance of 'neun' is observed in phrase initial and word-initial positions.

In Figure 1, if target sequence of 'ta', 'ka', and 'neun' is followed by silent pause, it is classified as 'phrase-final'. If they are preceded by silent pause, those types of target sequences are classified as 'phrase-initial'. If a target sequence of 'ta', 'ka', or 'neun' is neither followed by nor preceded by silent pause, and it is the first syllable in a word or the last syllable in a word, then it is classified as 'word-initial' and 'word-final,' respectively. Otherwise, the instances of the target sequences are referred to as 'word-medial.' To note is that even though it is possible to have a word that begins with 'neun' (e.g., *neun-ta* 'to increase'), no instance of phrase-initial or word-initial 'neun' is observed in the corpus. In all instances, monosyllabic words such as *ta* 'all' which can be both word-initial and word-final were excluded from the analysis. In Table 1, examples of word tokens are given.

Table 1 presents the total number of tokens summed up from 10 male speakers. It is shown that phrase-final 'ta' is the most frequent among different 'ta's. It is mainly due to the fact that the type of the sentences in the reading style of the corpus is declarative, and the sentence ends with 'ta'. As for the frequent occurrence of word-final 'neun', it is mainly due to the function of 'neun' as a topic marker. As for 'ka,' the word-final 'ka' is frequently observed after word-medial 'ka'. When 'ka' occurs word-finally or phrase-finally in the corpus, it often functions as a nominative marker.

Table 1. The total number of tokens summed up from 10 male speakers in the 20's

	'ta'	*'ka'*	*'neun'*
word-initial	342	773	n.a.
word-medial	599	1301	397
word-final	98	706	1818
phrase-initial	44	111	n.a.
phrase-final	1116	120	266
Total	2247	3011	2481

3. Results

The result of the durational adjustment of 'eu' in 'neun' among different positions is presented first, followed by the results of the durational adjustment of the nucleus vowel 'a' in 'ta' and 'ka,' respectively. First, descriptive statistics of 'neun' is given in Table 2 that shows the number of valid cases (N), mean, standard deviation (SD), standard error (SE), and 95% confidence interval (CI) under different boundary condition of 'neun'. The duration measurement is taken at the syllable nucleus of 'neun'.

Table 2. Descriptive Statistics of the nucleus duration in 'neun'

'neun'	*N.*	*MEAN*	*SD*	*SE*	*CI*
Phrase final	266	99.81	36.61	2.24	4.42
Word final	1818	49.48	24.99	0.59	1.15
Word medial	397	42.39	22.07	1.11	2.18

A graphical representation is given in Figure 2, in which a line chart is drawn together with error bars of the syllable nucleus duration in 'neun' under the phrase-final (pf), word-final (wf), and word-medial (wm) positions, respectively[2].

[2] A reviewer expressed a concern that the reported duration in this paper has been extracted from the output of a forced alignment system, which may not be reliable and erroneous.

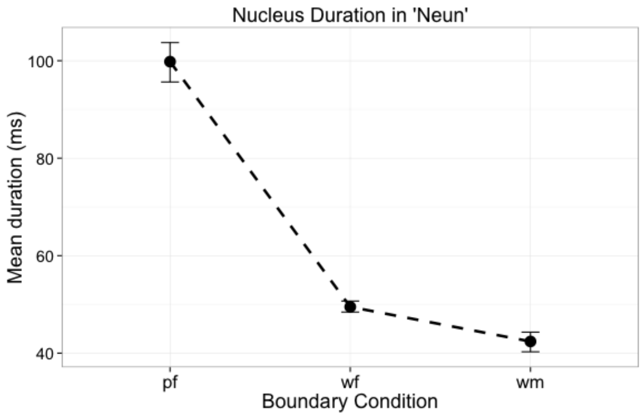

Figure 2. A line chart with error bars of the syllable nucleus duration in ‘neun'

In the figure, the duration of ‘eu’ in ‘neun’ that occurs in phrase final position is significantly longer than those of ‘eu’ that occurs in word-final position or in word-medial position. It is clear from this chart that there are between group differences. Levene's test is conducted to see whether the variance in duration varies across levels of ‘neun’ that occur in different positions in a phrase. The output shows that Levene's test is significant, $F(2, 2478) = 61.821$; $p<.0001$. This means that for these data, the variances are not similar across different boundary conditions. Because the assumption of homogeneity of variance is not met, Welch's F is tested with duration as a dependent variable and the three positions (phrase-final, word-final, and word-medial) as levels of an independent variable. Welch's F makes adjustments for differences in group variances. The output of Welch's F test is $F(2, 530.8)=269$; $p<.0001$, implying that the mean durations do differ significantly across different boundary conditions.

Statistical analysis is applied to find out whether significant difference is observed between different positions. The nucleus duration of ‘neun’ in phrase-final position (M= 99.81; SD= 36.6) is significantly longer than that of ‘neun’ in word-final position (M=49.48; SD=24.9), with a small effect size, $t(302.172)=21.69$; $p<0.001$; Cohen's d=1.424. As for the durational difference between the word-medial ‘eu’ and the word-final one, Welch Two Samples t-test indicates that the nucleus duration in word-final position

Approximately 3000 tokens of vowels were manually corrected and a correlation test was done between the duration of automatically extracted vowels and that of the manually corrected segments. The correlation coefficient turned out to be 0.9182. The high correlation coefficient is partly due to the acoustic model of the forced-alignment system that has been trained on the basis of the corpus as the current study used. The results ensure that the automatically duration measurements are not so erroneous as to minimize the validity of experimental results reported in this paper.

(M=49.48, SD=24.99) is significantly longer than the duration in word-medial position (M=42.39, SD=22.02) with a medium effect size, $t(637.723)=5.66$; $p<0.001$; Cohen's $d=0.31$.

It is hypothesized that the word-final 'neun' may coincide with the right-edge of AP, which tends to be realized as High tone, and that the phrase-final 'neun' with the right-edge of IP. The IP in utterance medial position may have a High tone, which signals that the message is not complete but is continued (Pierrehumbert and Hirschberg 1990). Figure 3 shows mean F0 taken from the nucleus in 'neun' in three different positions within a phrase. In the figure, it is shown that the F0 values under phrase- and word-final positions are higher than those under word-medial position, which is in line with the hypothesis[3].

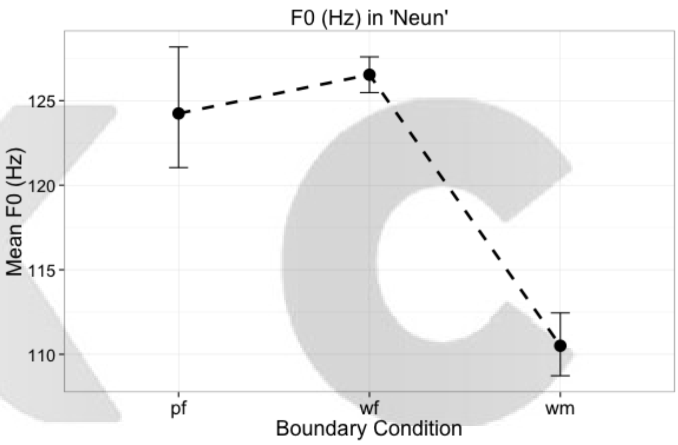

Figure 3. Syllable nucleus duration (left) and F0 (right) in 'neun' in three different positions within a phrase

The duration of syllable nucleus for sequences of 'ta' and 'ka' is measured for each of the five positions. Table 3 presents the descriptive statistics for 'ta' and 'ka,' respectively. Descriptive Statistics of 'ta' and 'ka' that show the number of valid cases (N. of tokens), mean, standard deviation (SD), standard error (SE), and 95% confidence interval (CI) under different boundary condition of 'ta' and 'ka', respectively. The duration measurement is taken at the syllable nucleus of 'ta' and 'ka', respectively.

[3] Welch two sample t-test indicates that no difference is found between the F0 in the phrase final position and the F0 in the word final position ($t(311.8)= -1.172$, $p>0.2$).

Table 3. Descriptive Statistics of vowel duration of ‘ta’ and ‘ka’[4]

		N.	*MEAN*	*SD*	*SE*	*CI*
'ta'	Phrase final	1150	232.80	47.14	1.39	2.72
	Phrase initial	39	95.38	43.82	7.01	14.20
	Word final	98	100.40	47.49	4.79	9.52
	Word initial	341	80.82	35.49	1.92	3.77
	Word medial	437	73.29	28.16	1.34	2.64
'ka'	Phrase final	119	243.53	67.71	6.20	12.29
	Phrase initial	111	74.05	38.64	3.66	7.26
	Word final	706	81.79	40.76	1.53	3.01
	Word initial	770	65.84	28.94	1.04	2.04
	Word medial	1191	63.56	29.84	0.86	1.69

Before conducting statistical analyses, graphical representations are made to visually inspect whether the duration of syllable nucleus show different patterns on the basis of its position within a phrase. Figure 4 shows error bar graphs. The figure on the left side is an error bar graph of the mean duration (in ms.) under 5 different boundary condition of the syllabic nucleus ‘a’ when the nucleus is preceded by the alveolar ‘t’ versus the velar ‘k’. The shaded area is selected and shown again on the right-sided figure. In the figure on the right side, phrase-final duration, which is prominently long, is omitted.

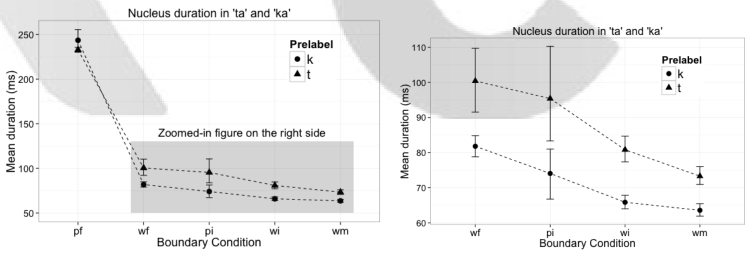

Figure 4. Error bar graph of the mean duration (in ms) under 5 different boundary conditions

The figure on the right illustrates that the duration of the vowel after ‘t’ is significantly longer than the duration of the vowel after ‘k’. It is not known as to why there is a length difference between post-consoantal vowels. It may be the case that the difference is made in order to make the syllable duration more or less constant. Given that the voice onset time (VOT) which reflects the aspiration of the voiceless stops depends on the place of

[4] Tokens longer than 400ms. in duration were eliminated from the analysis. For example, 21 out of 4952 'ta' tokens are excluded.

articulation such that the VOT of velar stops is longer than that of alveolar stops, the vowel after the velar stops may be shorter than the vowel after the alveolar stop. Interesting though it may sound, this question is left for future research because the corpus is not tagged with the VOT events.

A number of planned contrast tests are conducted to examine the difference of the nucleus duration among different phrasal positions. First, the difference of syllabic nucleus ‘a’ among ‘ta’s is examined under different positions. The first planned contrasts test is conducted with word-final ‘ta’ as the baseline level and comparing it with the ‘ta's in other positions. Table 4 shows the results. In Table 4, we can see that the word-final ‘ta’ differs from other ‘ta’s except for the word-initial ‘ta’.

Table 4. Planned contrasts test: Treatment contrast with word-final ‘ta’ as the baseline level[5]

	Estimate	*Std. Error*	*t value*	*Pr(>|t|)*
(Intercept)	100.408	4.912	20.441	$<2e^{-16}$ ***
pf vs. wf	135.623	5.115	26.517	$<2e^{-16}$ ***
pi vs. wf	-5.024	9.206	-0.546	0.585361
wi vs. wf	-18.625	5.572	-3.343	0.000844 ***
wm vs. wf	-27.113	5.435	-4.989	$6.59e^{-7}$ ***

In the table, we can tell that word-final ‘ta’ is shorter than phrase-final ‘ta’ by 135ms, longer than word-initial ‘ta’ by 18ms, and longer than word-medial ‘ta’ by 27ms. Word-final ‘ta’ is longer than phrase-initial ‘ta’ by 5ms, but the difference is not significant ($p>0.5$).

It is, however, unknown whether word-medial ‘ta’ is longer than word-initial ‘ta’ or not. Second planned contrasts test is designed to see whether there is a significant difference between the two regarding the nucleus duration. This test is conducted with word-medial ‘ta’ as the baseline level. As shown in Table 5, significant difference is found in the nucleus duration between word-medial and word-initial ‘ta.’ The nucleus duration in word-initial ‘ta’ is about 8ms longer than that in word-medial ‘ta,’ with statistical significance ($p<0.05$).

[5] The following are significance codes used in the statistical analysis: 0.0001 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’

Table 5. Planned contrasts test: Treatment contrast with word-medial 'ta' as the baseline level

	Estimate	*Std. Error*	*t value*	*Pr(>|t|)*
(Intercept)	73.295	2.326	31.509	$<2e^{-16}$ ***
pf vs. wm	162.736	2.728	59.652	$<2e^{-16}$ ***
pi vs. wm	22.089	8.127	2.718	0.00662 **
wf vs. wm	27.113	5.435	4.989	$6.59e^{-7}$ ***
wi vs. wm	8.488	3.511	2.418	0.01570 *

In a similar way, two planned contrasts tests are conducted for the nucleus duration of 'ka'. The resulting tables are given in Table 6 and Table 7, respectively.

Table 6. Planned contrasts test: Treatment contrast with word-final 'ka' as the baseline level

	Estimate	*Std. Error*	*t value*	*Pr(>|t|)*
(Intercept)	81.799	1.498	54.609	$<2e^{-16}$ ***
pf vs. wf	163.201	3.930	41.528	$<2e^{-16}$ ***
pi vs. wf	-7.745	4.064	-1.906	0.0568
wi vs. wf	-14.412	2.072	-6.956	$4.32e^{-12}$ ***
wm vs. wf	-17.390	1.890	-9.20	$<2e^{-16}$ ***

Table 7. Planned contrasts test: Treatment contrast with word-medial 'ka' as the baseline level

	Estimate	*Std. Error*	*t value*	*Pr(>|t|)*
(Intercept)	64.409	1.152	55.896	$<2e^{-16}$ ***
pf vs. wm	180.591	3.812	47.379	$<2e^{-16}$ ***
pi vs. wm	9.645	3.950	2.442	0.0147 *
wf vs. wm	17.390	1.890	9.202	$<2e^{-16}$ ***
wi vs. wm	2.978	1.838	1.620	0.1053

Tables 6 and 7 collectively demonstrate that the only pairs that do not show differences are word-initial 'ka' and word-medial 'ka' ($p>0.1$) and phrase-initial 'ka' and word-final 'ka' ($p>0.05$). Note that in the case of word-initial 'ta' and word-medial 'ta', the difference in the nucleus duration is statistically different at the level of $p<0.05$.

4. Discussions and Conclusion

The paper investigated the durational adjustments of post-consonantal vowels at the edges of words and in the middle of words in various phrasal positions. Specifically, the temporal adjustment of two vowels 'a' [a] and 'eu' [ɨ] was observed from a reading-style corpus of standard Korean. The vowel ‘a’ was preceded by either ‘k’ (as in [ka]) or ‘t’ (as in [t]), and the vowel ‘eu’ [ɨ] was flanked by ‘n’ (as in ‘neun’ [nɨn]). The duration of the syllabic nucleus (e.g. ‘eu’ in the case of ‘neun’) or syllabic rime ('eun’ [ɨn] in the case of ‘neun’) was analyzed. These sequences of CV (i.e., ‘ka’ and ‘ta’) or CVC (i.e., ‘neun’) were measured as they occur in the middle of sentences. There were five positions that target sequence could occur within a phrase: phrase-final, phrase-initial, word-final, word-initial, and word-medial. The results of the analyses using a reading-style Korean corpus support the view that phonetic manifestation of duration is conditioned by hierarchically organized phrasal structure: In general, the higher the target sequence occupies in the phrasal structure, the longer the domain-final vowel (Phrase-final > Word-final > word-medial). Also in general, the target sequence occurring at the final position tends to be longer than the counterpart occurring at the initial position (Phrase-Final > Phrase-Initial; Word-Final > Word-initial). As for the word-medial and word-initial duration, preceding consonant seems to play a role in that the nucleus duration in word-medial and word-initial ‘ta’ was significantly different, the duration in the same environment for ‘ka’ did not show statistically significant difference. This result seems to be in line with Cho and Keating (2001), in which only the vowels occurring after $/t^*/$ are generally longer in higher prosodic domains.

These findings have implications to the prosodic organization. The Phrase-final position defined in the paper tends to coincide with prosodic phrase final position. The word-final position (or *eojeol*-final position) in the paper tends to coincide with AP-final position. It is interesting to note that word-final duration is longer than word-medial position. The lengthened word-final duration may indicate that there is an AP-induced lengthening in Korean. Initial strengthening was controversial as to whether the domain of strengthening and lengthening is restricted to the initial consonant or not. The finding in this study indicates that the post-consonantal vowel is influenced by the type of pre-vocalic consonants. The vowel duration is lengthened if preceded by an alveolar consonant, but is not if preceded by a velar consonant. The reason for the asymmetric behavior remains to be found.

Statistical approaches to a large-scale database are complementary to theoretical and experimental studies of different processes in the production and perception of speech. Such statistical studies of speech durations have shown a number of systematic regularities in the temporal organization of speech, sometimes confirming, at other times contradicting earlier findings. For example, Crystal and House (1988) note that "certain traditional predictions may not hold for informal connected (read) speech signals." (cf. Umeda 1977). It is not the case that statistical approaches are comprehensive and fine-grained. Sometimes due to the lack of available computational resources or tools, fine-grained analyses are not applicable to large-scaled corpus at hand. Turk and Shattuck-Hufnagel (2000) report six mechanisms that can potentially cue the location of word boundaries: (1) Word-final lengthening and (2) word-initial lengthening for boundary-related effects, (3) accentual lengthening for the effect of pitch accent, (4) polysyllabic shortening and (5) syllable-ratio equalization for changes based on the organization of syllables into larger constituents, and (6) closed syllable vowel shortening especially for vowels preceding voiceless consonants. In this study using large-scaled corpus, some locations (i.e., word-final and word-initial duration, and closed syllable vowel shortening) were possible to be examined, but other position-related durational properties could not be examined due to the complexity of extracting those features from a corpus without manual annotation. Advancement in computational tools and resources may make it possible to conduct more fine-grained analysis.

REFERENCES

BECKMAN, MARY E. and JAN R. EDWARDS. 1990. Lengthenings and shortenings and the nature of prosodic constituency. In John Kingston and Mary E. Beckman (eds.). *Papers in laboratory phonology I: between the grammar and physics of speech* 152-178. Cambridge: Cambridge University Press.

BYRD, DANY. 2000. Articulatory vowel lengthening and coordination at phrasal junctures. *Phonetica* 57, 3-16.

CHO, TAE-HONG and PATRICIA A. KEATING. 2001. Articulatory and acoustic studies on domain-initial strengthening in Korean. *Journal of Phonetics* 29, 155-190.

CHRISTOPHE, ANNE, ARIEL GOUT, SHARON PEPERKAMP and JAMES MORGAN. 2003. Discovering words in the continuous speech stream: The role of prosody. *Journal of Phonetics* 31, 585-598.

COLE, JENNIFER and HASEGAWA-JOHNSON MARK. 2012. Corpus phonology with speech resources. In Abigail Cohn, Cécile Fougeron and Marie K. Huffman (eds.). *Handbook of Laboratory Phonology* 431-440. Oxford, UK: Oxford University Press.

CRYSTAL, THOMAS H. and ARTHUR S. HOUSE. 1988. Segmental durations in connected-speech signals: Current results. *Journal of Acoustical Society of America* 83, 1553-1573.

DILLEY, LAURA, STEFAINE SHATTUCK-HUFNAGEL and MARI OSTENDORF. 1996. Glottalization of word-initial vowels as a function of prosodic structure. *Journal of Phonetics* 24, 423-444.

ECHOLS, CATHARINE H., MEGAN J. CROWHURST and JANE B. CHILDERS. 1997. Perception of rhythmic units in speech by infants and adults. *Journal of Memory and Language* 36, 202-225.

EDWARDS, JAN R., MARY. E. BECKMAN and JAN FLETCHER. 1991. The articulatory kinematics of final lengthening. *Journal of Acoustical Society of America* 89, 369-382.

FLETCHER, JANET. 2010. The prosody of timing: Timing and rhythm. In

William Hardcastle, John Laver and Fiona E. Gibbon (eds.). *The Handbook of Phonetic Sciences* (2nd edition), 523-602. Oxford: Wiley-Blackwell.

FOUGERON, CÉCILE. 1998. *Variations articulatoires en début de constituants prosodiques de différents niveaux en Français*. PhD Dissertation. Université Paris III-Sorbonne Nouvelle.

______________. 2001. Articulatory properties of initial segments in several prosodic constituents in French. *Journal of Phonetics* 29, 109-135.

FOUGERON, CÉCILE and PATRICIA A. KEATING. 1997. Articulatory strengthening at edges of prosodic domains. *Journal of the Acoustical Society of America* 101, 3728-3740.

HIRSH-PASEK, KATHY, DEBORAH G. KEMLER NELSON, PETER W. JUSCZYK, KIMBERLY W. CASSIDY, BENJAMIN DRUSS and LORI KENNEDY. 1987. Clauses are perceptual units for young infants. *Cognition* 26, 269-286.

JUN, SUN-AH. 1993. *The Phonetics and Phonology of Korean Prosody*. Unpublished PhD Dissertation. The Ohio State University.

JUSCZYK, PETER W. 1997. *The Discovery of Spoken Language*. Cambridge, MA: The MIT Press.

KIM, SAHYANG. 2004. *The Role of Prosodic Phrasing in Korean Word Segmentation*. PhD Dissertation. UCLA.

MCQUEEN, JAMES M. and ANNE CUTLER. 2010. Cognitive processes in speech perception. In William Hardcastle, John Laver and Fiona E. Gibbon (eds.). *The Handbook of Phonetic Sciences* (2nd edition), 489-520. Oxford: Wiley-Blackwell.

NAKATANI, LLOYD H., KATHLEEN D. O'CONNOR and CARLETTA H. ASTON. 1981. Prosodic aspects of American English speech rhythm. *Phonetics* 38, 84-106.

NOOTEBOOM, SID. 1997. The prosody of speech: Melody and rhythm. In William Hardcastle and John Laver (eds.). *The Handbook of Phonetic Sciences.* 640-673. Oxford: Basil Blackwell Limited.

OLLER, D. KIMBROUGH. 1973. The effect of position in utterance on speech segment duration in English. *Journal of the Acoustical Society of America* 54, 1235-1247.

PIERREHUMBERT, JANET and JULIA HIRSCHBERG. 1990. The meaning of intonation in the interpretation of discourse. In Philip Cohen, Jerry Morgan and Martha E. Pollack (eds.). *Intentions in Communication,* 271-311. Cambridge, MA: MIT Press.

SAFFRAN, JENNY R., RICHARD N. ASLIN and ELISSA L. NEWPORT. 1996. Statistical learning by 8-month-old infants. *Science* 274, 1926-1928.

SEIDL, AMANDA and ELIZABETH K. JOHNSON. 2006. Infant word segmentation revisited: Edge alignment facilitates target extraction. *Developmental Science* 9, 565-573.

SOHN, HO-MIM. 1999. *The Korean Language*. Cambridge: Cambridge University Press.

SOHN, SUNG-OCK. 1995. *Tense and Aspect in Korean*. Hawaii: University of

Hawaii.
TURK, ALICE E. and STEFANIE SHATTUCK-HUFNAGEL. 2000. Word-boundary-related duration patterns in English. *Journal of Phonetics* 28, 397-440.
UMEDA, NORIKO. 1977. Consonant duration in American English. *Journal of Acoustical Society of America* 61, 846-858.
WIGHTMAN, COLIN W., STEFANIE SHATTUCK-HUFNAGEL, MARI OSTENDORF and PATTI J. PRICE. 1992. Segmental durations in the vicinity of prosodic phrase boundaries. *Journal of Acoustical Society of America* 91, 1707-1717.
YOON, KYUCHUL. 2005. *Building a Prosodically Sensitive Diphone Database for Korean Text-To-Speech Synthesis System*. PhD Dissertation. The Ohio State University.
YOON, TAE-JIN. to appear. A corpus-based study on the layered duration in Standard Korean. In Michael Kenstowicz, Ted Levin and Ryo Masuda (eds.). *Japanese/Korean Linguistics* 23. California: CSLI Publications.
YOON, TAE-JIN and KANG, YOONJUNG. 2014. Monophthong analysis on a large-scale speech corpus of read-style Korean. *Phonetics and Speech Sciences* 6, 139-145.

Tae-Jin Yoon
Department of English Language and Literature
Sungshin Women's University
2, Bomun-ro 34da-gil, Seongbuk-gu, Seoul,
South Korea 02844
e-mail: tyoon@sungshin.ac.kr

received: July 17, 2015
revised: August 16, 2015
accepted: August 17, 2015