



The realization of English rhythm by Busan Korean speakers*

Wook Kyung Choe**

College of General Education, Dong-A University, Busan, Korea

Abstract

The purpose of the current study is to investigate the realization of speech rhythm in English as spoken by Korean learners of English. The study particularly aims to examine the rhythm metrics of English read speech by learners who speak Busan or the South Kyungsang dialect of Korean. Twenty-four learners whose L1 is Busan Korean and eight native speakers of English read a passage wherein five sentences were segmented and labeled as vocalic and intervocalic intervals. Various rhythm metrics such as %V, Varcos, and Pairwise Variability Indexes (PVI) were calculated. The results show that Korean learners read English sentences with significantly more vocalic and consonantal intervals at a slower speech rate than native English speakers. The analyses of rhythm metrics revealed that when the speech rate was not normalized, Korean learners' English showed more variability in the length of consonantal and vocalic intervals. However, speech-rate-normalized rhythm metrics for vocalic intervals indicated that Korean learners transferred their L1 rhythmic structures (a syllable-timed language) into their L2 speech (a stress-timed language). Overall, the results suggest that Korean learners' English reflects the rhythmic characteristics of their L1. The effect of the learners' L1 dialect on the realization of L2 speech rhythm is also speculated.

Keywords: English rhythm, second language acquisition, rhythm metrics, Korean learners of English, Busan Korean

1. Introduction

All spoken languages show regularity in timing, which is called as speech rhythm. Based on the speech units where the regularity occurs, languages can be divided into 'stress-timed' and 'syllable-timed' languages (Abercrombie, 1967; Pike, 1945). A typical characteristic of a stress-timed language is that the length of the intervals between stressed syllables is equal, while a syllable-timed language is known to have equal length of syllables.

Although the earlier research on speech rhythm provided cate-

gorical distinction between stress- and syllable-timed languages, numerous empirical studies have proved that neither the length of inter-stress intervals in stress-timed languages nor the length of each syllable in syllable-timed languages is always identical. These recent studies have argued a gradient classification of speech rhythm by examining the acoustic characteristics of rhythm using various metrics. These rhythm metrics focused not on the phonological units such as a stress or a syllable but on the phonetic units such as vocalic and consonantal intervals. The studies also aimed to suggest the rhythm metrics which

* This work was supported by the Dong-A University research fund.

** regina1004@dau.ac.kr, Corresponding author

Received 22 November 2019; Revised 10 December 2019; Accepted 11 December 2019

© Copyright 2019 Korean Society of Speech Sciences. This is an Open-Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (<http://creativecommons.org/licenses/by-nc/4.0>) which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

best represent the characteristics of speech rhythm in a certain language. For example, Ramus et al. (1999) investigated the proportion of vocalic intervals within a sentence (%V), the standard deviation of vocalic intervals (ΔV), and the standard deviation of consonantal intervals (ΔC) of 8 different languages. They suggested that %V and ΔC well distinguished the different types of speech rhythms, in that stress-timed languages such as English and Dutch showed lower %V (40.1 for English and 42.3 for Dutch) due to frequent vowel reductions and higher ΔC (53.5 for English and 53.3 for Dutch) due to more complex syllable structures with consonant clusters. However, syllable-timed languages such as French and Spanish showed relatively higher %V (43.6 for French and 43.8 for Spanish) with lower ΔC (43.9 for French and 47.4 for Spanish)*. Since %V and ΔC in Ramus et al. can be affected by speech rate, Dellwo (2006, 2008) suggested VarcoC and VarcoV, in which ΔC and ΔV are divided into the average duration of consonantal intervals and vocalic intervals, respectively. Dellwo (2006) argued that speech-rate-normalized rhythm metric VarcoC better distinguished a syllable-timed language French (about 46) from stress-timed languages English (about 53) and German (about 62.5). In addition, Ling et al. (2000) and Grabe & Low (2002) suggested the raw Pairwise Variability Index (rPVI) and the normalized Pairwise Variability Index (nPVI) to successfully provide the degree of variability in successive measurements. The formula for rPVI and nPVI are as follows. In both formula, m indicate the number of measured intervals (e.g., the number of vocalic intervals for PVI-V), d_k indicates the duration of the k^{th} interval.

$$rPVI = \left[\sum_{k=1}^{m-1} |d_k - d_{k+1}| \right] / (m-1) \quad (1)$$

$$nPVI = 100 \times \left[\sum_{k=1}^{m-1} \left| \frac{d_k - d_{k+1}}{(d_k + d_{k+1})/2} \right| \right] / (m-1) \quad (2)$$

In their study, Grabe & Low (2002) examined the PVIs for 18 different languages and confirmed that the prototypical stress-timed languages such as English and German exhibited higher nPVI-V (57.2 for English and 59.7 for German) and rPVI-C (64.1 for English and 54.3 for German) due to frequent vowel reductions and consonant clusters. In contrast, the prototypical syllable-timed languages such as French and Spanish exhibited lower nPVI-V (43.5 for French and 29.7 for Spanish) and rPVI-C (50.4 for French and 57.7 for Spanish) since the lengths of both vocalic and intervocalic intervals in these syllable-timed languages are not likely to vary. However, they also argued that speech rhythm classification is not categorical but gradient, in that the rhythm metrics of some languages showed the PVI values of neither stressed- nor syllable-timed languages.

Different from English—a prototypical stress-timed language, Korean is not considered as a prototype of either stress- or syllable-timed languages (e.g., Arvaniti, 2009, 2012; Jang, 2008, 2009b; Lee et al., 1994). Based on phonological perspectives,

Korean is a syllable-timed language, whose syllable structures are quite simple disallowing consonant clusters. Korean also has neither lexical stresses nor vowel reductions (see Kim et al., 2007; Lee & Song, 2019). However, the acoustic studies indicated that the values of rhythm metrics in Korean were in between a prototypical stress-timed language such as English and a prototypical syllable-timed language such as French (Arvaniti, 2009; Jang, 2009b). For example, Jang (2009a) examined aforementioned rhythm metrics and speech rate of 40 Korean speakers' read speech. He, then, compared the values with those in French, Spanish, Dutch, and English. The results revealed that Korean could be categorized as a syllable-timed language based on higher %V and lower nPVI-V, while the higher VarcoV of Korean was similar to that of a stress-timed language.

This rhythmic distinction among different languages attracted the researchers on the second language (L2) acquisition. That is, studies examined the realization of L2 speech rhythm by learners whose first language (L1) has different rhythmic structures from L2. They assumed the negative transfer of speech rhythm similar to other linguistic elements, so expected that learners' speech might reflect their L1 rhythm characteristics. Despite the inconsistency of rhythm metrics in different studies and some non-prototypical syllable-timing patterns, many researchers have considered Korean as a syllable-timed language. Because of the different rhythmic structures of English and Korean, several studies investigated the realization of English rhythm by Korean learners of English (e.g., Jang, 2009a; Kim, 2008; Kim & Chung, 2016; Lee & Kim, 2005; Lee & Song, 2019; Sa, 2015). Specifically, Jang compared the rhythm metrics of English spoken by native speakers with those spoken by Korean learners. The results showed that the %V, VarcoV, nPVI-V, and rPVI-C values in the native speakers' speech were significantly lower than those in Korean learners' speech. Except for %V, the obtained results were against the expectation, in that Korean learners' English showed larger variability between syllables despite their syllable-timed L1. One possible explanation for the results was more frequent final lengthening at sentence-internal boundaries due to frequent hesitation and/or sentence-internal pauses. That is, since Korean learners frequently hesitate and put sentence-internal prosodic boundaries when speaking L2, they often prolonged the final syllables at the sentence-internal prosodic boundaries. This frequent lengthening of sentence-internal syllables may cause larger VarcoV, nPVI-V, and rPVI-C values despite their L1 rhythmic influence, which showed smaller variability in successive intervals.

The purpose of the current study is to investigate the realization of speech rhythm in English produced by Korean learners, using various rhythm metrics. Especially, the current study aimed to focus on Korean learners whose L1 is Busan and the South Kyungsang dialect of Korean (hereafter, this will be referred to as Busan Korean as a short version) for two reasons. The first is that most of the studies about the acquisition of English rhythm by Korean learners were based

* Although some researchers argued the third type of speech rhythm, mora-timing, the current study will focus only on the two types, stress- and syllable-timing. This is not only because there is only one well recognized mora-timed language, Japanese, but also because the scope of the current study deals with English rhythm spoken by native speakers and Korean learners. Note that Korean can be perceived as mora-timed language as Japanese (Cho, 2004).

on the learners who spoke Seoul Korean, the standard Korean. If we assume the rhythmic structure of L1 is transferred to L2, the rhythmic characteristics of L2 by learners with different dialects of the same L1 could also vary. In fact, previous research indicated the dialectal differences in rhythm metrics such as Singapore Mandarin vs. standard Mandarin (Grabe & Low, 2002; Lin & Wang, 2007) and Singapore English vs. British English (Grabe & Low, 2002).

The second reason of focusing Busan Korean is due to its phonetic characteristic. Kyungsang province is located in the southeastern part of Korea, and Busan is the biggest city in South Kyungsang area. One of the distinctive phonetic and phonological characteristics of Kyungsang Korean is its lexical tones (e.g., Chang, 2013; Lee, 2009). That is, different from Seoul Korean, Busan Korean speakers successfully differentiate [nun] 'eye' from [nun] 'snow' by using a high tone and a rising tone* respectively. The relevant aspect of these lexical tones in Busan Korean is that the vowel in a syllable with a rising tone is longer than that in a syllable with a high tone (Chang, 2013; Yoshida et al., 2007). Despite the lack of studies on rhythmic characteristics of Busan Korean, we can easily assume that the values in rhythm metrics for vocalic intervals in Busan Korean will be larger than those in Seoul Korean. Therefore, considering L1 transfer in the realization of English rhythm, the learners who speak Busan Korean could show different rhythmic patterns from those who speak Seoul Korean.

2. Methods

2.1. Participants

Eight native speakers of English and 24 Korean learners of English voluntarily participated in the current study. Half of the 32 participants were male, and the others were female. None of the participants are known to have hearing and speech problems.

The native speakers of English were undergraduate or graduate students at the University of Oregon in the U.S., and they were in their 20s to 40s. All of them have been raised and educated only in the U.S., so their English was the mainstream of American English. None of them have reported to fluently speak any language other than English.

The Korean learners of English were all undergraduate students of Dong-A University located in Busan, Korea. They were all in their 20s, and majoring in various fields. All of the learners have been raised and educated only in Busan, Ulsan, or South Kyungsang Province, so all of them speak fluent South Kyungsang Korean though the dialects which their parents used were somewhat diverse. None of the learners have reported to live more than 4 months in English-speaking countries. The English proficiency level of each learner was determined by self-reported TOEIC scores. Except three learners who did not have scores yet, the average TOEIC scores of 21 learners were 543.81, with which we consider them low-intermediate learners.

The speech data of the learners without TOEIC scores were first listened and included only when the author thought that their reading ability was similar to the other learners' one.

2.2. Stimuli and Procedure

Each participant was asked to read an excerpt from a short story from TIME for kids about polar bears, and the whole text is presented in Appendix 1. Since the learners' English proficiency levels were low intermediate, the article for young children was selected. This could make the learners fairly fluently read the text with easy vocabulary and contents to get sufficient speech data without frequent hesitance or disfluency. To gain more natural read speech, all punctuation marks were left as presented in the original text.

Each participant was asked to read the reading material out loud at his/her natural reading pace. All participants were also asked to read the text in mind before reading out loud to make sure they were familiar with the content and the vocabulary. If they had difficulties in pronouncing a certain word, the author read it in isolation to only provide segmental information about the word. When there was a frequent speech error, a self-correction, and/or a hesitance, the participant was asked to re-read the right sentence from the beginning. A few disfluencies or sentence-internal pauses were not interrupted by the experimenter to obtain more natural speech data.

The experiment was separately conducted for the native speakers (hereafter NS) and the learners in Busan, Korea (hereafter BKL). The experiment for the NS group was in a quiet laboratory room in their university. The NS group used Shure ULXS4 wireless receiver and lavalier microphone, and their speech was digitally recorded to a Marantz PMD 660. The BKL group's recording was conducted in a quiet office in their university. These learners used a head-worn microphone (Shure SM35-XLR) and their speech was also digitally recorded to a Marantz PMD 661 MKIII.

2.3. Acoustic measurement

The recorded read speech was analyzed using Praat software (Boersma & Weenink, 2014). For acoustic characteristics of English rhythm, only 5 out of 10 sentences were selected. First, due to their low proficiency, BKL speakers had difficulty in reading sentences with large numbers and a temperature unit although the experimenter taught them how to pronounce the numbers and the unit. Specifically 17 out of 24 learners hesitated, self-corrected, and made mistakes in reading these sentences, so the second, the ninth, and the tenth sentences from the reading material were excluded for further analyses. Except these, the author selected 5 sentences to diverse the length and the grammatical structures of sentences and the following table shows the selected sentences used in acoustic measurement. As in Table 1, Sentence 1 and 2 are simple sentences with the relatively small number of syllables (12 and 11 syllables). Meanwhile, Sentence 3 and 5 are a bit longer and more complex than the first two sentences, in that Sentence 3 has

* Given that the names and the numbers of lexical tones in Busan Korean are still debatable among researchers, here I followed Chang (2013)'s classification of lexical tones.

17 syllables and Sentence 5 has 13. Both of them have prepositional phrases: one at the beginning of Sentence 5 and the inserted one for Sentence 3. Sentence 4 is the longest with 28 syllables, and it has quite complex sentence structures such as listing and relative clauses.

Table 1. The stimulus sentences used in acoustic measurements

| Number | Sentence |
|--------|--|
| 1 | The top of the world is a wintry wonderland. |
| 2 | The ocean is surrounded by frozen ground. |
| 3 | There are few people or trees, but to polar bears, the Arctic is home. |
| 4 | Polar bears have thick fur, huge paws and other features that make them well prepared for life in their harsh environment. |
| 5 | In fact, they need the Arctic sea ice for survival. |

In order to obtain rhythm metrics, the read speech was segmented by the author. Both visual information (spectrograms and waveforms) and auditory information were used to segment and label intervals, pauses, and disfluency. Following the previous research on pauses (e.g., Krivokapic, 2007), more than 200 ms of silence was considered as a pause; while a hesitance, a speech error, and other fillers were all considered as disfluency. These pauses and disfluency were noted and labeled, but excluded from further analysis except for the number of sentence-internal pauses. The segmentation criteria for vocalic and intervocalic intervals were followed by Grabe & Low (2002). However, the criteria for syllable boundaries, which are still arguable among researchers, were followed by Lee & Kim (2005), who mostly stuck to Onset Maximization Principle. The durations of segmented vocalic and consonantal intervals were measured using Praat script.

With obtained durations, aforementioned rhythm metrics were calculated: %V, ΔC and ΔV from Ramus et al. (1999); VarcoC and VarcoV from Dellwo (2006, 2008); and PVI values from Grabe & Low (2002). In addition to these rhythm metrics, speech rate in syllable per second—the number of syllables divided by articulation durations (hereafter S-rate), and the number of pauses within an utterance (hereafter NumP) from Jang (2009a) were also calculated for further analysis.

3. Results

From 160 sentences (32 speakers \times 5 sentences), a total of 2,239 vocalic intervals and 2,723 intervocalic intervals were yielded. On average, each speaker in the NS group exhibited 66.88 vocalic intervals and 72.25 consonantal intervals for 5 sentences. The BKL group had more intervals than the NS group, in that the average of 71 vocalic intervals and 89.38 consonantal intervals were yielded per learner. In sum, the Korean learners of English read sentences with more frequent vocalic and consonantal segmentation.

As for the rhythm metrics, Table 2 shows the mean and the standard deviation values of the aforementioned 11 rhythm metrics for the two speaker groups (BKL and NS).

Table 2. The mean and the standard deviation of 11 rhythm metrics for the two speaker group

| Rhythm metrics | Korean learners (BKL) | Native speakers (NS) |
|----------------|-----------------------|----------------------|
| %V | 48.25 (5.63) | 46.71 (4.22) |
| ΔC | 67.77 (20.04) | 54.88 (14.35) |
| ΔV | 75.72 (30.61) | 61.69 (21.15) |
| VarcoC | 48.30 (9.31) | 49.52 (10.84) |
| VarcoV | 52.29 (13.90) | 56.85 (10.68) |
| rPVI-C | 75.06 (25.20) | 57.87 (15.94) |
| nPVI-C | 53.76 (13.51) | 53.47 (12.45) |
| rPVI-V | 85.94 (34.04) | 72.90 (27.63) |
| nPVI-V | 56.00 (12.75) | 66.00 (15.87) |
| S-rate | 3.64 (0.65) | 4.79 (0.60) |
| NumP | 5.62 (2.81) | 1.62 (1.69) |

PVI, Pairwise Variability Index.

The first analysis focused on the effects of group (BKL vs. NS) on the non-normalized rhythm metrics (%V, deltas, and rPVIs) with one-way ANOVAs. The one-way ANOVA tests revealed significant differences between groups on ΔC [$F(1, 158)=14.12, p<.001$], ΔV [$F(1, 158)=7.24, p=.008$], rPVI-V [$F(1, 158)=4.81, p=.03$]. Since the rest two rhythm metrics (%V and rPVI-C) violated homogeneity, the effect of group was tested by Mann-Whitney U tests. The results revealed that only rPVI-C values of the two groups were significantly different [$U=1,343, p<.001$]. These significant differences in non-normalized rhythm metrics are presented in Figure 1.

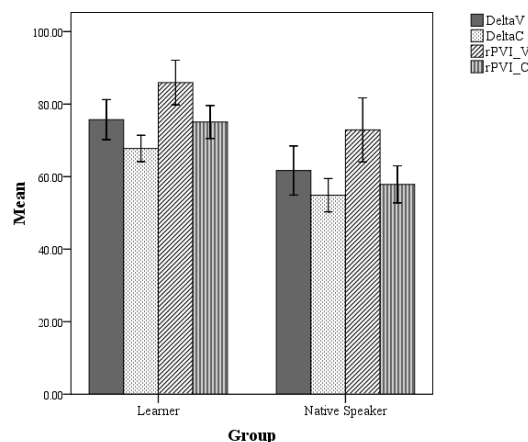


Figure 1. Mean value of ΔV , ΔC , rPVI-V, and rPVI-C as a function of groups (BKL vs. NS) with 95% confidence interval error bars.

Figure 1 shows that the non-normalized rhythm metrics for the NS group's English were lower than those for the BKL group's English. In other words, without considering the differences in their speech rate, both the vocalic and the intervocalic intervals in BKL's speech was more variable than those in NS's speech. Given that a stress-timed language usually shows higher variability in vocalic and consonantal intervals due to consonant clusters and vowel reductions, the obtained patterns are beyond expectation considering the learners' syllable-timed L1. On the other hand, these patterns were consistent with those in the previous research. For example, Jang (2009a) reported significantly higher rPVI-C for Korean learners' English, and then argued that these unexpected results were because of

frequent phrase-final lengthening at sentence-internal pauses for learners. That is, as the learners often hesitated and/or chunked utterances into smaller prosodic phrases when reading L2 texts, they frequently lengthened the last segments before sentence-internal pauses, which made the intervals more variable than those in native speakers' speech.

The results in Figure 1 bring us to the next analysis the speaker group effect on speech rate and the number of pauses. One-way ANOVAs yielded a significant speech-rate differences [$F(1, 158)=97.96, p<.001$] and number-pause differences [$F(1, 30)=14.29, p=.001$] between BKL and NS. Specifically, native speakers' speech was significantly faster than Korean learners, in that the Korean learners produced nearly 4 syllables per second, while the native speakers did 5 syllables per second as in Table 2. In addition, the learners put sentence-internal pauses more frequently than the native speakers (nearly 6 sentence-internal pauses for 5 sentences). The results about speech rate and the number of pauses were consistent with Jang (2009a)'s ones and well represent the L1 and L2 distinction, showing significantly slower and frequently chunked L2 speech.

To investigate the group effect on rate normalized rhythm metrics, one-way ANOVAs with Group as an independent variable and Varcos and nPVIs as dependent variable were conducted. These test revealed the significant Group effect only on nPVI-V [$F(1, 158)=16.24, p<.001$] as shown in Figure 2.

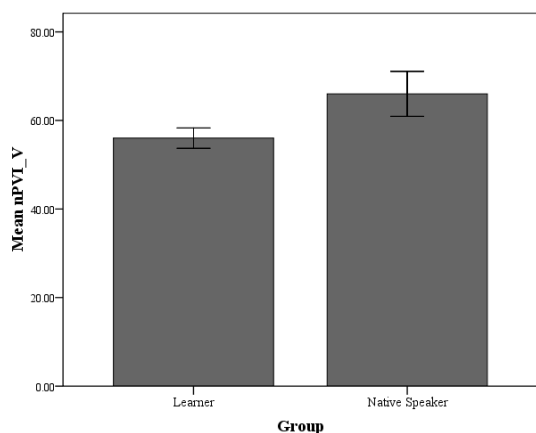


Figure 2. Mean value of nPVI-V as a function of groups (BKL vs. NS) with 95% confidence interval error bars.

Figure 2 indicated that the variability in vocalic intervals for native speakers' English was significantly greater than that for Korean learners' English. This suggests that compared with native speakers, Korean learners' vowels were produced in relatively regular length when speech rate differences were considered. This obtained pattern corresponds to the expectation: since the learners' native language is a syllable-timed language with little variability in vowels and consonants, the English spoken by Korean learners is expected to reflect their native language rhythm structures.

A previous study, Kim (2008) also showed similar patterns as the current nPVI-Vs. She noted that Korean learners' nPVI-V was 59, which was significantly lower than the native speakers' value 78. The result in Kim's study is also consistent with the expectation as the current study. One noted difference

between Kim's and the current result is relatively lower nPVI-V value in Kim's study. Though direct comparison is not legitimate due to different experimental settings, the difference might arise from the dialects of the learners' L1 (i.e., Seoul Korean for Kim vs. Busan Korean for the current study). This issue will be further examined later in the discussion.

Last, the effect of different sentences was investigated. Since the current study selected 5 sentences varying the total length and syntactic complexity, different rhythmic patterns might occur. However, the effect of sentences was only tested for nPVI-V and rPVI-C not only because these showed significant differences in the current study, but also because these two metrics were often considered to well represent the types of speech rhythm (e.g., Grabe & Low, 2002; Sa, 2015). Table 3 below shows the results from one-way ANOVAs with Group and PVIs for each of 5 sentences.

Table 3. Results of one-way ANOVAs with Group on nPVI-V and rPVI-C

| Sentence number | nPVI-V | rPVI-C |
|-----------------|--------|--------|
| 1 | | ** |
| 2 | * | |
| 3 | | * |
| 4 | ** | * |
| 5 | * | |

* $p<.05$, ** $p<.001$.

Table 3 shows that the rhythm metrics which represent the distinctions between L1 and L2 speech varied in different sentences. The stimulus sentences were not well-controlled to investigate which linguistic characteristics determine the best predicting rhythm metrics, but one of the obvious findings is that the longer and the more complex a stimulus sentence, the more rhythm metrics can successfully distinguish learners' speech from native speakers' one.

4. Discussion and Conclusion

The current study was designed to investigate the acoustic characteristics of speech rhythm in English spoken by Korean learners. Specifically, the study focused on the realization of English rhythm by learners who speak Busan or the South Kyungsang dialect of Korean. To examine the rhythmic characteristics of Busan Koreans' English, the study used various rhythm metrics such as Varcos and PVIs. The four main findings are as follows: (a) Korean learners read English sentences with significantly more vocalic and intervocalic intervals than native speakers, (b) when speech rate was not normalized, Korean learners' English showed more variability in the length of consonant and vowel intervals, (c) speech-rate-normalized rhythm metrics for vocalic intervals indicated that Korean learners transferred their L1 rhythmic structures into their L2 speech, and (d) Korean learners read English sentences at slower speech rate as well as frequently put sentence-internal pauses. These findings will be discussed with respect to the concept of L1 transfer of the realization of L2 speech rhythm. In addition, the possible explanation for some unexpected results will be speculated. Last, but not least, the effect of learners' L1 dialect on the realization of L2 rhythm

will also be discussed.

Overall, the current study revealed that the Korean learners' English reading showed typical L2 speech characteristics, and some of these were because of their L1 influence. The negative L1 transfer was noted in the average number of intervals and the less variable vocalic intervals (nPVI-V). The result that the Korean learners read sentences with more vocalic and consonantal intervals may due to the different syllable structures of English and Korean. As mentioned before, Korean is considered as a syllable-timed language mainly because of its syllable structure, which does not allow consonant clusters. Considering the level of English proficiency for the learners in the study (low-intermediate), they might have difficulties in pronouncing more than one consonant in syllable onset or coda positions. To overcome these difficulties, the learners could often insert vowels inside of consonant clusters to produce with simple syllable structures as their native language. While segmenting and labeling the learners' speech, this trend was noted, and the learners' frequent vowel epenthesis can be a possible reason for the average number of vocalic and intervocalic intervals.

The finding that the nPVI-V for the Korean learners was significantly lower than that for the native speakers is another good evidence of the negative L1 influence. In other words, since the learners' L1 is a syllable-timed language, in which the variability in the length of vocalic intervals was lower, this feature influenced in the L2 speech and exhibited lower nPVI-V for the learners' speech. The learners' vowel epenthesis strategy for consonant clusters also provides another reason for lower nPVI-V. That is, if we assume learners with a syllable-timed L1 tend to put vowels (in similar length) between consonants and to make additional syllables mostly with CV structures, this vowel epenthesis may add the regularity in the length of vocalic intervals of L2 speech.

However, this assumption of the vowel epenthesis provides contradictory explanation for higher rPVI-C values for learners. If learners always break consonant clusters by adding vowels, the variability in consonantal intervals for learners must be the same or even smaller, and so the learners' rPVI-C needs to be the same or lower than that for native speakers. The current finding, however, showed the opposite direction. Therefore, we need another speculation for the learners' higher rPVI-C, and here we will adopt Jang (2009a)'s idea. It is natural that L2 learners' speech—especially with lower proficiency levels—shows frequent disfluency and/or sentence-internal pauses to produce their L2 speech with smaller chunks (e.g., Choe, 2016). This feature was evident both in Jang's and the current studies with the significant differences on speech rate and the number of pauses. If we assume that learners plan their speech and make a sentence into smaller but "proper" intonational phrases followed by sentence-internal pauses, the learners should lengthen the final syllable of those frequent intonational phrases. The learners' strategy to chunk the sentences into their own intonational phrases was well captured during segmentation and labeling procedure. That is, many of the learners' sentence-internal pauses were quite grammatical (e.g., There are few people or trees // but to polar bears // the Arctic is home.) instead of sudden stopping due to hesitance and/or mistakes. In sum, we can argue that in many cases, the Korean learners in the current study read the stimulus sentence with more than one intonational

phrase, which then elicits more frequent final lengthening than native speakers. These lengthened syllables may result in not only the higher rPVI-C, but also the higher ΔV , ΔC , and rPVI-V for the learners' speech.

So far, we have discussed the main findings considering the rhythm characteristics of the learners' L1 in general. Now let us move our perspectives to the dialect of the learners' L1. Since no study have yet investigated the rhythmic characteristics of Busan Korean using rhythm metrics, the following discussion is based on the previous research on the phonology of Busan Korean in general. As mentioned above, Busan Korean is known to have lexical tones. Despite some disagreement about the category and the type of tones, it is reported that a syllable with a so-called rising tone as in [nun] for 'snow' is longer than that in the syllables with other tones as in [nun] for 'eye'. That is, Busan Korean speakers, different from other dialects of Korean, may be used to manipulating the length of Korean vowels coupled with their lexical tone, and so may show greater variability in the length of their L1 vowel production. Given that the learners' L1 rhythmic characteristics influence the realization of rhythm in their L2 speech, we can possibly expect higher PVI-Vs for learners with Busan Korean than learners who speak other dialects. Again, though it is difficult to directly compare the current results with the results in the previous research, the nPVI-V in English speech by Busan Korean speakers was higher than the nPVI-V by Seoul Korean speakers in Kim (2008). In addition, considering the levels of English proficiency for the learners in the current study (low intermediate) and those in Kim (upper intermediate or advanced), this higher nPVI-V could support the idea of the dialectal influence. Specifically, only as for the variability in the length of vowel intervals, Busan Korean speakers could produce more native-like English than Seoul Korean speakers even though Busan Korean speakers' English proficiency level was lower. We could argue that the more variability in vowel length for the learners of Busan Korean may result from their ability to manipulate vowel lengths in their L1 speech.

Last, the findings of the current study suggest several directions to future research. Especially, considering that the scope of the current study is to illustrate the realization of English speech rhythm by Busan Korean speakers, the study has limitation of examining the actual effect of the learners' dialects on the L2 rhythm acquisition. Therefore, in order to test the hypothesis that learners with different dialects of the same L1 show different rhythmic patterns in their L2 speech, the future studies need to investigate (a) rhythm metrics of Korean speech produced by speakers with different dialects, and (b) the realization of English rhythms by Korean learners who speak different dialects. By comparing rhythm metrics in both L1 and L2 speech produced by different dialectal groups, it will be possible to determine whether or not the rhythmic characteristics in different dialects of the same L1 are transferred and which rhythmic features more successfully distinguish the dialectal differences.

In conclusion, the current study aimed to investigate how learners of English who speaks Busan Korean realize speech rhythm in their English reading. The analyses of rhythm metrics suggest that Korean learners' English reflect the rhythmic characteristics of not only their native language—Korean as a

syllable-timed language, but also specific rhythmic characteristics of the dialect of their native language.

References

- Abercrombie, D. (1967). *Elements of general phonetics*. Edinburgh, UK: Edinburgh University Press.
- Arvaniti, A. (2009). Rhythm, timing and the timing of rhythm. *Phonetica*, 66, 46-63.
- Arvaniti, A. (2012). The usefulness of metrics in the quantification of speech rhythm. *Journal of Phonetics*, 40(3), 351-373.
- Boersma, P., & Weenink, D. (2014). Praat: Doing phonetics by computer (Version 5.3.66) [Computer Program]. Retrieved from <http://www.praat.org/>
- Chang, S. E. (2013). Effects of fundamental frequency and duration variation on the perception of South Kyungsang Korean tones. *Language and Speech*, 56(2), 211-228.
- Cho, M. H. (2004). Rhythm typology of Korean speech. *Cognitive Process*, 5(4), 249-253.
- Choe, W. K. (2016). Acquisition of prosodic phrasing and edge tones by Korean learners of English. *Phonetics and Speech Sciences*, 8(4), 31-38.
- Dellwo, V. (2006). Rhythm and speech rate: A variation coefficient for delta C. In P. Karnowski, & I. Szigei (Eds.), *Language and language processing: Proceedings of the 38th Linguistic Colloquium* (pp. 231-241). Frankfurt, Germany.
- Dellwo, V. (2008). The role of speech rate in perceiving speech rhythm. In *Speech Prosody 2008, Proceedings of Speech Prosody 2008 series* (pp. 375-378). Campinas, Brazil.
- Grabe, E., & Low, E. L. (2002). Durational variability in speech and the rhythm class hypothesis. In Gussenhoven, C., & Warner, N. (Eds.), *Laboratory phonology 7*, (pp. 515-546). Berlin, Germany: Mouton de Gruyter.
- Jang, T. Y. (2008). Speech rhythm metrics for automatic scoring of English speech by Korean EFL learners. *Malsori*, 66, 41-59.
- Jang, T. Y. (2009a). Automatic assessment of non-native prosody using rhythm metrics: Focusing on Korean speakers' English pronunciation. *Proceedings of the 2nd International Conference on East Asian Linguistics*.
- Jang, T. Y. (2009b). Rhythm metrics of spoken Korean. *Language and Linguistics*, 46, 169-185.
- Kim, H. K., & Chung, H. (2016). An analysis of the rhythm and reduced vowels of Korean adult learners of English. *Studies in Foreign Language Education*, 30(3), 109-132.
- Kim, J. M., Flynn, S., & Oh, M. (2007). Non-native speech rhythm: A large-scale study of English pronunciation by Korean learners. *Studies in Phonetics, Phonology and Morphology*, 13(2), 219-250.
- Kim, S. A. (2008). Rethinking the dichotomy between syllable-timed vs. stress-timed languages with particular reference to Korean L1 speakers' English. *The Journal of Studies in Language*, 24(3), 473-493.
- Krivokapic, J. (2007). Prosodic planning: Effects of phrasal length and complexity on pause duration. *Journal of Phonetics*, 35(2), 162-179.
- Lee, D. M. (2009). *The loanword tonology of South Kyungsang Korean* (Unpublished doctoral dissertation). Indiana University, Bloomington, IN.
- Lee, H. B., Jin, N. T., Seong, C. J., Jung, I. J., & Lee, S. M. (1994). An experimental phonetic study of speech rhythm in Standard Korean. *In 3rd International Conference on Spoken Language Processing* (pp. 1091-1094). Yokohama, Japan.
- Lee, H. Y., & Song, J. (2019). Evaluating Korean learners' English rhythm proficiency with measures of sentence stress. *Applied Psycholinguistics*, 40(6), 1363-1376.
- Lin, H., & Wang, Q. (2007). Mandarin rhythm: An acoustic study. *Journal of Chinese Language and Computing*, 17(3), 127-140.
- Lee, O. H., & Kim, J. M. (2005). Syllable-timing interferes with Korean learners' speech of stress-timed English. *Speech Sciences*, 12(4), 95-112.
- Ling, L. E., Grabe, E., & Nolan, F. (2000). Quantitative characterizations of speech rhythm: Syllable-timing in Singapore English. *Language and Speech*, 43, 377-401.
- Pike, K. L. (1945). *The Intonation of American English*. Ann Arbor, MI: University of Michigan Press.
- Ramus, R., Nespore, M., & Mehler, J. (1999). Correlates of linguistic rhythm in the speech signal. *Cognition*, 73(3), 265-292.
- Sa, J. J. (2015). Rhythm metrics approach on English speech by Korean learners of English. *Journal of British and American Studies*, 35, 235-254.
- Yoshida, K., Yoon, J., & Kim, H. J. (2007). Production and perception of word prosody in three dialects of Korean. *Proceedings of 16th International Congress of Phonetic Sciences* (pp. 1,169-1,172). Saarbrücken, Germany.

- **Wook Kyung Choe**, Corresponding author
Professor, College of General Education, Dong-A University
37 Nakdong-Daero 550 beon-gil,
Saha-gu, Busan 49315, Korea
Tel: +82-51-200-1022
Email: regina1004@dau.ac.kr
Areas of interest: Phonetics, Phonology, L2 acquisition

Appendix 1.

The top of the world is a wintry wonderland. Icebergs float in the cold Arctic Ocean. In winter, the temperature often falls to -30°F and the sun never rises. The ocean is surrounded by frozen ground. There are few people or trees, but to polar bears, the Arctic is home.

Polar bears have thick fur, huge paws and other features that make them well prepared for life in their harsh environment. In fact, they need the Arctic sea ice for survival. But climate change is causing larger and larger areas of summer sea ice to melt. Experts say that if warming patterns continue, the Arctic could be free of summer sea ice by 2050. That may cause two-thirds of the world's 20,000 polar bears to be gone by then too.