Linguistic Research 41(3): 391-429
DOI: 10.17250/khisli.41.3.202412.003

The impact of native language on second language rhythm acquisition: Insights from a cross-linguistic and intra-language corpus study*

Sujin Oh** · **Hanyong Park*****
(Soongsil University · University of Wisconsin-Milwaukee)

Oh, Sujin and Hanyong Park. 2024. The impact of native language on second language rhythm acquisition: Insights from a cross-linguistic and intra-language corpus study. *Linguistic Re*search 41(3): 391-429. This study investigates how native language (L1) rhythm structure influences the acquisition of second language (L2) rhythm across 20 different languages and within a single language group. Utilizing diverse corpora of L1 and L2 speech samples from various languages, we explored the rhythmic patterns of syllable-timed (e.g., Korean) and stress-timed (e.g., English) language speakers. Our findings reveal that L2 learners from syllable-timed language backgrounds can achieve rhythm patterns similar to native English speakers, challenging the notion that L1 rhythm disparities inherently disadvantage learners. This supports the Speech Learning Model (SLM), suggesting that learners with rhythmically contrasting native languages may acquire L2 rhythm more effectively than those with similar rhythmic structures. Furthermore, we examined the effect of speech style (reading vs. spontaneous) on rhythm production. Native Korean speakers consistently exhibited higher durational variability in spontaneous speech compared to reading in both L1 and L2, contrary to expectations that educational focus would lead to increased variability in reading. This pattern underscores the potential influence of individual speech traits, suggesting that L1 rhythmic tendencies persist in L2 production. Our study highlights the need for further research into the interplay between L1 and L2 rhythm acquisition and the impact of speech style across diverse language backgrounds. This research contributes to understanding the broader applicability of the SLM and the role of speech style in rhythm production. **(Soongsil University · University of Wisconsin-Milwaukee)**

Keywords rhythm, durational variability, PVI, L2 acquisition, speech style, SLM

* We are grateful to Dr. Jae Yung Song, Dr. Anne Pycha, and anonymous reviewers for their insights that greatly improved the manuscript.
** First author
*** Corresponding author

1. Introduction

In the present study, we investigated the impact of native language (L1) on the production of rhythm in a second language (L2), examining both cross-linguistic and intra-linguistic factors. Specifically, we aimed to determine: (1) whether native speakers of rhythmically contrastive languages (e.g., speakers of syllable-timed languages learning a stress-timed language as L2) can produce rhythm patterns akin to those of the target language, and (2) whether such rhythmic adaptation is consistent across different speech styles within a single language group.

The structure of the paper is as follows: we begin by providing an overview of speech rhythm in the traditional approach, along with empirical evidence from previous studies supporting this concept. We then review the literature on rhythm acquisition and development in L2, as well as research on the effects of speech styles on rhythm production. Next, we present our research questions in detail and discuss the rationale for the current study. Following this, we describe our methodology and report the results of the two studies. Finally, we discuss the main findings and their significance, and conclude with general observations in the last section.

2. Background

2.1 Acoustic metrics on speech rhythm

Speech rhythm has traditionally been regarded as the isochronous recurrence of certain speech units (Pike 1945; Abercrombie 1967). According to this isochrony hypothesis, spoken languages can be broadly categorized into two rhythmic classes: syllable-timed and stress-timed. In syllable-timed languages, such as Spanish, Italian, or French, the intervals between successive syllables occur at regular intervals. Conversely, in stress-timed languages, such as English, German, or Dutch, the intervals between successive stressed vowels occur at regular intervals. However, this approach has been criticized for its lack of empirical support. Since this classification is primarily based on listeners' impressions, researchers have called for empirical evidence to substantiate these claims. Moreover, although perception studies have found that infants can distinguish different rhythm types (Nazzi et al. 1998), researchers argue that acoustic

measurements are necessary to quantify rhythm objectively.

Researchers have introduced various rhythmic metrics to acoustically measure rhythm (Dauer 1983; Ramus et al. 1999; Low et al. 2000; Gibbon and Gut 2001; Grabe and Low 2002; Dellwo 2006). Inspired by Dauer (1983), Ramus and colleagues (1999) proposed quantifying rhythm as a succession of vowels of variable duration alternating with consonants. From Dauer's eight criteria for constructing rhythmic diversity, they focused on two: syllable structure and vowel reduction. They hypothesized that rhythmic differentiation is influenced by whether a language has a simple syllable structure (e.g., a higher ratio of CV structures) or a complex syllable structure (e.g., a higher ratio of CVC or CVCC structures), and whether it exhibits vowel reduction. Ramus et al. (1999) assumed that syllable-timed languages tend to have simpler syllabic structures and less vowel reduction, while stress-timed languages typically have more complex syllabic structures and greater vowel reduction. Based on this assumption, they established three measurements to quantify rhythm:

- %V: The proportion of vocalic intervals within a sentence, calculated as the sum of vocalic intervals divided by the total duration of the sentence.
- ΔV: The standard deviation of the duration of vocalic intervals within a sentence.
- ΔC: The standard deviation of the duration of consonantal intervals within a sentence.

Using these measurements, which provide a quantifiable means to analyze and compare the rhythmic properties of different languages, Ramus and his colleagues analyzed recordings from speakers with eight different language backgrounds: English, Dutch, Polish, French, Spanish, Italian, Catalan, and Japanese. The recordings consisted of short declarative statements that were matched for the number of syllables (ranging from 15 to 19) and average duration (approximately 3 seconds) across the languages. The results showed that the %V and ΔC measurements accurately reflected the traditional rhythm classification. Differences emerged among speakers of stress-timed languages (English, Dutch, and Polish), syllable-timed languages (French, Spanish, Italian, and Catalan), and a mora-timed language (Japanese). Specifically, speakers of stress-timed languages exhibited the highest ΔC values, speakers of syllable-timed languages showed intermediate ΔC values, and speakers of the mora-timed language had the lowest ΔC values. Regarding %V, speakers of the mora-timed language

demonstrated the highest values, followed by speakers of syllable-timed languages, and then speakers of stress-timed languages with the lowest values. These findings suggest that rhythm classification correlates with the durations of vocalic and consonantal intervals, which are related to specific phonological properties of languages. This provides empirical evidence, based on acoustic measurements, that supports the traditional categorization of languages into rhythm classes.

Dellwo (2006) suggested revising Ramus and his colleagues' proposal by taking speech rate into account. He noted that their measurements (%V, ΔV, and ΔC) did not consider the impact of varying speech rates on rhythm measurement. Dellwo and Wagner (2003) and Barry et al. (2003) demonstrated that ΔC and ΔV negatively correlated with speech rate; specifically, slower speech rates resulted in higher ΔC values. Consequently, Dellwo proposed that relative variation should be compared in relation to speech rates rather than relying on absolute variation, ΔC. He introduced a modified measurement, VarcoΔC, which is calculated as follows:

$$Varco\,\Delta = \Delta C \times 100 \div mean\,C$$

C = duration of consonantal intervals

This adjustment aims to normalize the variability of consonantal intervals by accounting for the speech rate, thus providing a more accurate reflection of rhythmic properties across different languages and speaking conditions.

To verify his measurement, Dellwo selected three languages to represent each rhythmic class: German and English as stress-timed languages and French as a syllable-timed language. Speakers of each language were asked to read a text of approximately 80 syllables in their native languages (L1) at different speech rates, ranging from slowest to fastest. The results indicated that VarcoΔC provided better differentiation among the rhythmic classes. Furthermore, the study revealed that some languages, such as German and English, exhibited variation in rhythm as a function of speech rate. Conversely, the rhythm of other languages, such as French, appeared to remain unaffected by changes in speech rate. This suggests that VarcoΔC is a more robust measure for capturing rhythmic distinctions, particularly in accounting for the influence of speech rate.

Another widely used metric for measuring rhythm was proposed by Low et al. (2000). They defined rhythm in terms of durational variability and introduced the

Pairwise Variability Index (PVI) to quantify it. The PVI calculates the durational differences between successive vowels or consonants to measure the variability in the duration of segments within an utterance. The equation for PVI, as adopted from Low et al. (2002), is given as follows:

$$PVI = 100 \times \left[\sum_{k=1}^{m-1} \left| d_k - d_{k+1} \right| \div (m-1) \right]$$

where m is the number of intervals, vocalic or intervocalic, in the text, and d is the duration of the kth interval.

Furthermore, to mitigate the effect of speech rate across speakers, Low et al. (2000) proposed the normalized Pairwise Variability Index (nPVI) metric. This modification was introduced by Grabe and Low (2002) and is calculated as follows:

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

Their preliminary results indicated that stress-timed languages exhibit greater vocalic variability compared to syllable-timed languages. This increased variability is attributed to factors such as vowel quality, specifically the presence or absence of vowel reduction, along with other phonological properties (Low et al. 2000). These findings support the notion that rhythm classification is influenced by the degree of vocalic variability inherent in different language types.

For a more detailed examination of this metric, Low et al. conducted an acoustic analysis of 18 languages, encompassing rhythmically mixed and unclassified languages. The results demonstrated that the PVI values effectively categorized English, Dutch, and German as stress-timed, and French and Spanish as syllable-timed. Stress-timed languages exhibited greater durational variability compared to syllable-timed languages. However, the mixed and unclassified languages did not conform to this traditional classification, instead showing intermediate values that overlapped with the margins of the two rhythmic categories. These findings support Dauer (1983)'s weak categorical distinction hypothesis, which posits that language rhythm can be described by the degree of stress-timing based on the prominence of stress in the language. In other

words, mixed and unclassified languages do not fit neatly into traditional rhythmic classes but can be seen as having varying degrees of stress-timing. Thus, these languages might be better understood as being more or less stress-timed, or more or less syllable-timed, rather than fitting into rigid categories.

All the rhythmic metrics discussed in this section are summarized in Table 1.

Table 1. Rhythm metrics overview

Metrics name	Description	Reference
%V	Proportion of vocalic intervals	Ramus et al. (1999)
ΔV, ΔC	The standard deviation of the duration of vocalic/consonantal intervals	Ramus et al. (1999)
rPVI	Raw pairwise variability index for vocalic or consonantal intervals	Grabe and Low (2002)
nPVI	Nomarlized pairwise variability index for vocalic or consonantal intervals	Grabe and Low (2002)
VarcoΔV, VarcoΔC	Coefficient of variability in the duration of vocalic and consonantal intervals	Dellwo (2006)

2.2 L2 rhythm learning

Although much research has focused on comparing rhythm across languages to provide empirical evidence for the traditional isochrony classification, more recent studies have extended the application of rhythm metrics to nonnative speech (Lin and Wang 2005; White and Mattys 2007; Li and Post 2014; Ordin and Polynskaya 2015; Lee and Song 2019). One area of interest is examining the influence of L1 on the production of L2 rhythm (Low et al. 2000; Gut 2003; Carter 2005; Lin and Wang 2005). These studies have concluded that the characteristics of L1 rhythm transfer to L2 rhythm production. For instance, White and Mattys (2007) investigated the cross-language effect on L2 rhythm production. Speakers of three different L1s—two stress-timed languages (English and Dutch) and one syllable-timed language (Spanish)—were asked to read sentences in both their L1 and L2. Comparisons were made between speakers of stress-timed and syllable-timed languages and between speakers of the two stress-timed languages. The results showed that native Spanish

speakers increased vocalic interval variability when producing English compared to Spanish, while native English speakers decreased variability when producing Spanish compared to their L1. Conversely, native Dutch speakers did not show any change in variability when producing English, and English speakers exhibited no change when producing Dutch. These findings indicate that L1 rhythm structure plays a crucial role in the production of L2 rhythm. However, one limitation of the study is that it only examined a small number of languages (Spanish, Dutch, and English), making it difficult to generalize the findings on L2 rhythm acquisition for learners with other language backgrounds. Therefore, more extensive data is needed for a comprehensive interpretation of the effect of L1 on L2 rhythm production.

Other researchers have examined the relationship between L1 rhythm structure and L2 proficiency in the development of L2 rhythm (Stockmal et al. 2005; Li and Post 2014; Ordin and Polyanskaya 2014, 2015). For instance, Li and Post (2014) compared the rhythm production of Mandarin learners of English and German learners of English at intermediate and advanced proficiency levels with that of native English speakers. Their findings indicated that intermediate learners from both language groups exhibited lower durational variability than advanced learners when speaking L2 English. This suggests that L2 rhythm can be developed and refined as a function of increasing proficiency.

Following Li and Post (2014), Ordin and Polyanskaya (2015) investigated the development of L2 rhythm as a function of the acquisition process. They compared speakers of a rhythmically similar language (German) to the target language, English, with speakers of a rhythmically contrastive language (French) to the target language. These groups included individuals at beginner, intermediate, and advanced proficiency levels. Participants were recorded during an informal interview and a sentence elicitation task. The results showed that both groups developed L2 rhythm as a function of proficiency. However, there was a significant difference concerning their L1. German speakers exhibited native-like durational variability at the advanced level, whereas the French participants did not. The researchers concluded that speakers of an L1 that is rhythmically similar to the target language can develop L2 rhythm to the extent of being native-like. In contrast, speakers of an L1 that is rhythmically contrastive to the target language do not achieve the same level of rhythmic proficiency in L2.

2.3 The role of speech style in L2 rhythm production

One of the controversies surrounding the quantification of rhythm involves the reliability of rhythm metrics (White and Mattys 2007; Arvaniti 2012; Gut 2012; Lee and Song 2019). Previous studies have yielded varying results in quantifying rhythm for the same languages. For instance, as mentioned in Gut (2012), the %V values for British English varied significantly across different studies: 45.7% in Arvaniti (2012), 41.1% in Grabe and Low (2002), 40.1% in Ramus et al. (1999), and 38% in White and Mattys (2007).

The effect of speech styles has been suggested as one of the possible factors contributing to the unreliability of rhythm metrics (White and Mattys 2007; Arvaniti 2012). For instance, Arvaniti (2012) examined the effects of elicitation methods on measuring rhythm in nonnative speech. The study involved rhythm production by eight speakers of six languages (English, German, Italian, Korean, Spanish, and Greek), using three different speech styles: spontaneous speech, reading of *The North Wind and the Sun*, and reading sentences. Metrics such as ΔC, %V, rPVI, nPVI, VarcoC, and VarcoV were calculated for each recording. The results showed that different elicitation types significantly affected all metrics within languages. The values of all the metrics in spontaneous speech were significantly higher than those in the other two reading styles. This finding indicates that rhythm measurements are influenced by speech styles, regardless of the languages spoken. However, the influence of speech styles on rhythm has been under-researched, and to our knowledge, no study has specifically investigated this effect in L2 production. Therefore, further studies are needed to provide a more comprehensive interpretation of the impact of speech styles on rhythm metrics.

2.4 Research questions

The investigation reported here was designed to address how different rhythmic structures in L1 affect the learning of rhythm production in L2. In particular, the following two research questions were addressed:

1. Does L1 rhythm predict L2 rhythm production across languages?
2. Do speech styles affect rhythm production within a language?

To explore these questions, we analyzed how native rhythmic patterns influence the acquisition of rhythm in a second language and examined the impact of various speech styles on rhythm production within the same language.

One significant limitation of previous studies is the insufficient diversity of languages and speakers analyzed. Many studies have focused on only five to seven languages, often involving a small number of participants. To accurately generalize findings across languages, it is essential to collect comprehensive data from a broader range of languages and a larger pool of speakers. To address this gap in the literature, our study revisits the initial research question by examining the rhythm production of non-native speakers from a more diverse set of language backgrounds (20 different languages) and comparing their rhythm production to that of native speakers, both in their L1 and L2.

The non-native participants were categorized into two groups based on their native language types: one group included speakers of stress-timed languages, and the other comprised speakers of syllable-timed languages. English, which is a stress-timed language, was used as the target language for this study. We developed two primary predictions for the research.

First, regarding the comparison between L1 and native production, we anticipated a significant difference in durational variability based on rhythmic class. Specifically, we expected that the rhythm produced by speakers of stress-timed languages would closely resemble that of native English speakers, while the rhythm of syllable-timed language speakers would display distinct differences, aligning with previous research. This prediction was designed to offer additional empirical support for the traditional rhythmic classification approach.

Second, concerning the influence of L1 on L2 production, we predicted that if L1 had no substantial effect on L2 rhythm, there would be minimal differences in durational variability among the groups. In this case, the syllable-timed language speakers would adjust their durational variability to match that of native English speakers. Conversely, if L1 influence proved to be significant, syllable-timed language speakers would demonstrate lower durational variability in L2 English, despite attempts to adapt. This would suggest that the L1 rhythmic structure has a strong impact on L2 rhythm production.

Following up on the first research question, which explored rhythmic patterns across languages, we then shifted our focus to examining non-native rhythm

production within a single language. As discussed earlier, speech styles have been suggested as a potential factor contributing to inconsistencies in rhythmic metrics. Given that previous studies predominantly utilized reading materials to analyze L2 rhythm in non-native speech, it is crucial to investigate whether these patterns are consistent across different speech styles. To address this, we analyzed the effect of speech style on rhythm production. We examined recordings of 30 native Korean speakers producing speech in both L1 Korean (a syllable-timed language) and L2 English (a stress-timed language), using two distinct speech styles: reading and spontaneous speech. The Korean speakers read *The North Wind and the Sun* passage and also retold the story in both languages. We made the following predictions: Regarding durational variability in L2 production, we foresaw that Korean speakers would exhibit increased variability in L2 English compared to their L1 Korean. This expectation was based on the results from Study 1, which indicated that L2 rhythm production tends to align more closely with native English patterns.

For the effect of speech style in L1 production, we predicted that Korean speakers would produce higher durational variability in spontaneous speech than in reading speech in L1 Korean. This prediction was informed by Arvaniti (2012), which suggested that spontaneous speech typically involves greater durational variability than reading speech. In terms of the effect of speech style in L2 production, we hypothesized that Korean speakers would show greater durational variability in reading compared to retelling in L2 English. This hypothesis was based on the idea that many L2 learners, due to their educational experiences, might perform better in a structured reading task than in a spontaneous retelling task. As a result, they would likely demonstrate more effective use of L2 rhythm in reading contexts, leading to increased durational variability in reading relative to retelling. By investigating these predictions, we aim to provide further insights into the impact of speech style on rhythm production and assess the consistency of rhythm patterns in L2 across different speech contexts.

3. Study 1: The role of L1 rhythm in L2 rhythm production

The purpose of Study 1 was to examine whether L1 rhythm influences L2 rhythm production. We aimed to test the following predictions: First, we predicted significant differences in rhythm production between speakers of syllable-timed and stress-timed

languages in their L1. Specifically, we expected speakers of syllable-timed languages to demonstrate lower durational variability compared to their counterparts from stress-timed languages. Second, we hypothesized that there would be no significant difference in L2 rhythm production between speakers of syllable-timed and stress-timed languages. We expected that syllable-timed language speakers would increase their durational variability in L2 production, thereby aligning their rhythm more closely with the patterns of the stress-timed target language. These predictions aim to clarify how L1 rhythmic structures impact L2 rhythm production and whether non-native speakers are able to adjust their rhythmic patterns to better match those of the target language.

In this study, we analyzed recordings from three distinct groups of speakers: a native English group, a group of speakers from stress-timed languages, and a group of speakers from syllable-timed languages. The analysis focused on rhythm production in both L1 and L2 using a specified rhythm metric. The study aimed to investigate the effects of group membership on rhythm production across these two languages.

3.1 Methodology

3.1.1 Materials

For the current study, we utilized two distinct corpora. The non-native speakers' materials were sourced from the *Multi-talker Corpus of Foreign-accented English* (MCFAE) (Tamati et al. 2001). This corpus includes recordings of words, sentences, and passages of connected speech produced by non-native speakers of American English. Specifically, we focused on the recordings of speakers reading *The North Wind and the Sun* passage in both their L1 and L2 English.

The analysis encompassed recordings from 58 speakers representing 20 different language backgrounds. The languages were categorized into two groups: six stress-timed languages (Arabic, German, Russian, Brazilian Portuguese, European Portuguese, and Thai) and 14 syllable-timed languages (Bengali, Cantonese, French, Gujarati, Hindi, Indonesian, Italian, Korean, Malay, Mandarin, Persian/Farsi, Spanish, Tagalog, and Turkish). To balance the number of speakers across groups, an additional recording was obtained from a Thai speaker and a European Portuguese speaker.

However, the recording from the Portuguese speaker was excluded from the analysis due to noise interference. Additionally, note that the number of speakers in each language group is small due to the limited scope of the corpus used in this study. We hope to address this limitation in future research by using a larger corpus. The detailed classification of these languages is provided in Table 2.

Table 2. The rhythmic classification of languages used in the present study

Language	**Classification**
Arabic	Stress-timed (Abercrombie 1967)
German	Stress-timed (Kohler 1982)
Russian	Stress-timed (Abercrombie 1967)
Portuguese (Brazilian)	Stress-timed (Major 1981, 1985)
Portuguese (European)	Stress-timed (Cruz-Ferreira 1995)
Thai	Stress-timed (Luangthongkum 1977)
Bengali	Syllable-timed (Prahallad and Black 2003)
Cantonese	Syllable-timed (Lin and Wang 2007; Mok and Dellwo 2008)
French	Syllable-timed (Abercrombie 1967)
Gujarati	Syllable-timed (Prahallad and Black 2003)
Hindi	Syllable-timed (Das et al. 2008)
Indonesian	Syllable-timed (Miller 1984)
Italian	Syllable-timed (D'lmperio 2002)
Korean	Syllable-timed (Lee 1982)
Malay	Syllable-timed (Deterding 2011)
Mandarin	Syllable-timed (Lin and Wang 2007)
Persian/Farsi	Syllable-timed (Haghshenas 1978)
Spanish (Latin American)	Syllable-timed (Pike 1945)
Tagalog	Syllable-timed (Gonzalez 1970)
Turkish	Syllable-timed (Schierling 2007)

Table 3 shows the non-native speakers' information consisting of the number of speakers, the average Age of Arrival (AOA), and the average Length of Residency (LOR) in the United States in each language analyzed in the current study.

Table 3. Nonnative subject information from the MCFAE corpus

Language	Number speakers	Mean AOA (range) (years)	Mean LOR (range) (years)
Arabic	2	25 (25-25)	1.83 (0.67-3)
Bengali	4	23 (22-24)	2.52 (1.83-3)
Cantonese	3	19 (17-22)	1.78 (0.42-4.42)
Farsi	2	26.25 (24.5-28)	3.34 (3.17-3.5)
French	3	26 (23-28)	1.03 (0.25-1.42)
German	3	31 (27-35)	3.96 (3.67-4.25)
Gujarati	3	16 (14-18)	5.13 (2.58-7.67)
Hindi	2	23 (21-25)	0.63 (0.5-0.75)
Indonesian	2	31 (26-36)	1 (0.17-1.83)
Italian	3	23 (23-23)	0.17 (0.17-0.17)
Korean	4	24.25 (20-34)	0.92 (0.17-2.5)
Malay	4	20 (18-22)	2.19 (0.92-2.83)
Mandarin	4	19.25 (18-23)	1.94 (0.58-5.0)
Portuguese (Brazilian)	3	23.33 (20-26)	0.68 (0.25-1.42)
Portuguese (European)	1	25.0	5.50
Russian	4	31 (21-42)	2.52 (1.0-3.92)
Spanish	4	24.75 (13-43)	4.23 (2.25-5.67)
Tagalog	2	32 (32-32)	3.33 (3.33-3.33)
Thai	3	29.33 (23-39)	2.33 (0.5-4.83)
Turkish	2	24.75 (24.5-25)	1.33 (1.17-1.5)

The materials for native English speakers were drawn from the *Archive of L1 and L2 Scripted and Spontaneous Transcripts and Recordings* (ALLSSTAR) corpus (Bradlow, n.d.). This corpus includes recordings of native English speakers reading the same passage in English. For the purpose of this study, we utilized recordings from 26 native English speakers, comprising 14 females and 12 males.

3.1.2 Measurement

To examine the influence of L1 rhythmic structure on L2 rhythm production, we utilized the normalized Pairwise Variability Index (nPVI) metric for our analysis.

Despite ongoing debates about the reliability of various rhythmic metrics (Grabe and Low 2002; White and Mattys 2007; Arvaniti 2012; Gut 2012), the Pairwise Variability Index (PVI) has been widely adopted in research comparing rhythm across languages. Grabe and Low (2002) found that PVI measurements are particularly effective for classifying languages into rhythmic categories compared to other metrics. Furthermore, we chose to focus on vocalic interval measurements, as prior studies have demonstrated their higher reliability relative to consonantal interval measurements (White and Mattys 2007; Arvaniti 2012).

To investigate durational variability, which serves as a proxy for rhythm, we measured the duration of vowel and consonant intervals. The recordings of 58 non-native speakers and 26 native English speakers were annotated using Praat software (Boersma and Weenink 2018) by the first author and two research assistants. Each utterance was manually segmented into two categories: vocalic intervals (V) and consonantal intervals (C). This segmentation adhered to the criteria established in previous studies on speech rhythm metrics, such as those by Ramus et al. (1999) and Grabe and Low (2002). Vocalic intervals were measured from the onset to the offset of vowels, irrespective of the number of vowels within an interval. Consonantal intervals were measured from the offset of a vowel to the onset of the following vowel, regardless of the number of consonants present. Measurements were based on the second formants. For glide sounds, we followed the acoustic criteria outlined by Grabe and Low (2002). Specifically, if an initial glide demonstrated a noticeable change in formants leading to the subsequent vowel, it was included in the consonantal interval. Conversely, if the glide did not exhibit such a change, it was considered part of the vocalic intervals. All pauses between intonation phrases were excluded from the analysis.

Following the measurement of durations, we computed the normalized Pairwise Variability Index for vocalic intervals (nPVI-V) for each recording. The nPVI-V metric was chosen to account for variations in durational variability across different speakers. Research has consistently shown that non-native speakers often exhibit slower speech rates in their L2 production compared to their L1 (Lennon 1990; Munro and Derwing 1995, 1998). To mitigate the impact of these speech rate differences on our results, we utilized the normalized nPVI metric, which adjusts for variations in speech rate, thereby providing a more accurate reflection of rhythmic patterns.

3.1.3 Statistical analysis

To test the hypothesis that speakers of syllable-timed languages would exhibit lower durational variability compared to speakers of stress-timed languages in their L1 production, we conducted a one-way independent analysis of variance (ANOVA). In this analysis, *Group* served as the independent variable, while *L1 nPVI-V* was the dependent variable. The *Group* variable included three levels: native English speakers (*NS*), speakers of stress-timed languages (*Stress*), and speakers of syllable-timed languages (*Syllable*).

To evaluate the second prediction—that speakers of syllable-timed languages would increase their durational variability when producing L2—we conducted a similar one-way independent analysis of variance (ANOVA). In this analysis, *Group* was the independent variable, and *L2 nPVI-V* was the dependent variable. The *Group* variable included three levels: native English speakers (*NS*), speakers of stress-timed languages (*Stress*), and speakers of syllable-timed languages (*Syllable*).

3.2 Results

Statistical analyses were conducted using R 3.5.0 (R Core Team 2023) with the assistance of the following packages: *ggpubr* (Kassambara 2023). Table 4 shows the mean nPVI-V values and their standard deviations in each group in both L1 and L2 conditions.

Table 4. The mean nPVI-V and its standard deviation produced by three groups in L1 and L2

Language	Group	N	Mean	S.D.
L1	NS	26	56.0	5.79
	Stress	16	58.1	6.30
	Syllable	42	50.8	6.47
L2	NS	26	56.0	5.79
	Stress	16	60.5	6.33
	Syllable	42	56.8	5.34

Note: NS = native English speakers; Stress = stress-timed language speakers; Syllable = syllable-timed language speakers; N = the number of speakers; S.D. = standard deviation

Figure 1 illustrates the nPVI-V values for L1 production across the three groups, revealing a statistically significant effect of *Group* on average *nPVI-V*, $F(2, 56) = 10.06$, $p < .001$. Post hoc analyses using Tukey's Honestly Significant Difference (HSD) test identified significant pairwise differences. Specifically, there were significant differences between native English speakers (NS) and syllable-timed language speakers (Syllable), with an average difference of 5.15 ($p < .01$), and between stress-timed language speakers (Stress) and syllable-timed language speakers, with an average difference of 7.22 ($p < .001$). No significant difference was found between NS and Stress, with an average difference of 2.07 ($p = .55$). These results indicate that rhythm class significantly influences L1 rhythm production: syllable-timed language speakers exhibited lower durational variability compared to both stress-timed language speakers and native English speakers. Conversely, stress-timed language speakers demonstrated higher durational variability than syllable-timed language speakers, while their variability was comparable to that of native English speakers.

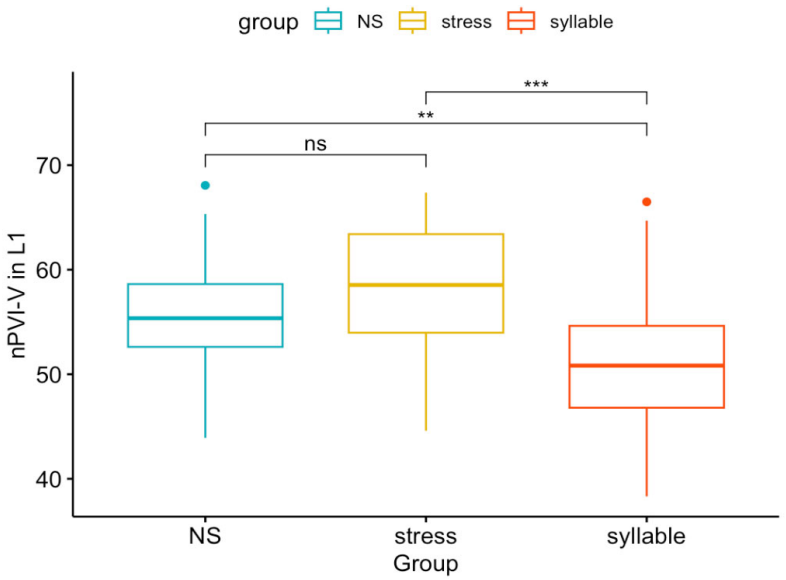

Figure 1. nPVI-V values produced by three groups (*NS* = native English speakers, *Stress* = stress-timed language speakers, *Syllable* = syllable-timed language speakers) in reading *The North Wind and the Sun* in L1 (** = $p < .01$; *** = $p < .001$)

Figure 2 depicts the nPVI-V values for L2 production across the three groups, revealing a statistically significant effect of *Group* on average *nPVI-V*, $F(2, 56) = 3.42$, $p < 0.5$. Tukey's Honestly Significant Difference (HSD) post hoc test identified a significant pairwise difference between native English speakers (NS) and stress-timed language speakers (Stress), with an average difference of 4.51 ($p < .05$). However, no significant differences were found between NS and syllable-timed language speakers (Syllable), with an average difference of 0.76 ($p = .86$), or between Stress and Syllable, with an average difference of 3.75 ($p = .07$). These findings indicate a notable change in durational variability in L2 production. Specifically, both learner groups—regardless of their L1 rhythmic structure—increased their durational variability when producing English. Syllable-timed language speakers achieved a level of variability similar to that of native English speakers. In contrast, stress-timed language speakers exhibited even greater durational variability than native English speakers. This suggests that while both groups adapted their rhythm production in L2, the extent of adaptation varied depending on their L1 background.

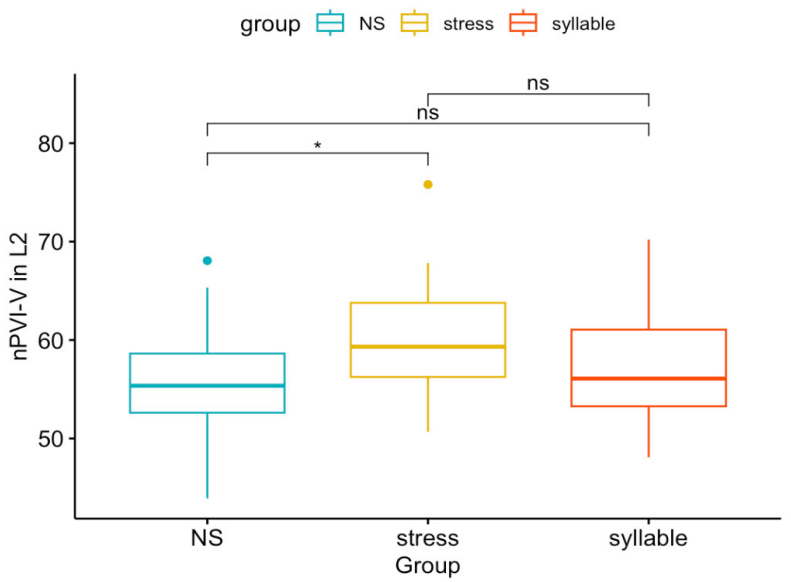

Figure 2. nPVI-V values produced by three groups (*NS* = native English speakers, *Stress* = stress-timed language speakers, *Syllable* = syllable-timed language speakers) when reading *The North Wind and the Sun* in L2 English (* = $p < .05$)

3.3 Discussion

In Study 1, two predictions were made concerning the impact of L1 rhythm on L2 rhythm production: First, for L1 production, we predicted a significant difference in durational variability between speakers of syllable-timed languages and those of stress-timed languages. Specifically, we expected speakers of syllable-timed languages to exhibit lower durational variability compared to speakers of stress-timed languages. Second, for L2 production, we anticipated that there would be no significant differences among the groups. This prediction was based on the expectation that syllable-timed language speakers would increase their durational variability in L2 production to align more closely with the rhythm patterns of native English speakers, thus eliminating any differences between the groups.

The results confirmed the first prediction: syllable-timed language speakers produced lower durational variability than stress-timed language speakers. Additionally, stress-timed language speakers exhibited durational variability levels that were relatively similar to those of native English speakers. These findings align with previous research that supports traditional rhythm classifications, such as those proposed by Grabe and Low (2002) and Ramus et al. (1999).

The second prediction was only partially supported. Syllable-timed language speakers did not differ significantly from native English speakers in L2 production, suggesting that these speakers successfully increased their durational variability to align with English rhythm, differing from their L1 rhythm. Conversely, stress-timed language speakers also exhibited increased durational variability in their English production, contrary to our expectations. Statistical analysis revealed a significant difference in durational variability between native English speakers and stress-timed language speakers, indicating that the latter produced more variability than native speakers. This result diverges from previous studies, which suggested that stress-timed language speakers, such as German learners of English, would produce similar levels of durational variability to native speakers once they achieved a certain proficiency level (Ordin and Polyanskaya 2015).

One possible explanation for this unexpected finding is hypercorrection. According to Eckman et al. (2013), L2 learners may exhibit hypercorrection errors during the final stages of acquiring a target language contrast, particularly when they perceive

a discrepancy between their L1 and L2 phonological systems. Supporting this, Kelly (2022) found evidence of hypercorrection in L2 intonational structures. In the context of rhythm, it is possible that stress-timed language learners overcompensated for perceived differences between their L1 and L2 English rhythm. This overcompensation may have led them to produce less native-like rhythm in their L2, reflecting a misunderstanding of the actual rhythmic difference between their L1 and L2.

4. Study 2: The effect of speech style on L2 rhythm production

The purpose of Study 2 was to examine the impact of speech style on rhythm production in both L1 and L2 contexts. Specifically, the study aimed to test the following predictions: First, we predicted that Korean speakers would exhibit increased durational variability in L2 production. This expectation is based on the assumption that Korean speakers, who are accustomed to the syllable-timed rhythm of their L1, would adapt their rhythm in English to achieve greater durational variability, thus aligning more closely with the stress-timed characteristics of the target language.

Second, we hypothesized that Korean speakers would produce higher durational variability in spontaneous retelling speech compared to reading speech in their L1 production. This hypothesis is grounded in the idea that spontaneous speech often includes more natural variations in rhythm, which could lead to greater durational variability than the more controlled and deliberate rhythm of reading tasks.

Third, we predicted that in L2 production, Korean speakers would produce higher durational variability during reading compared to spontaneous retelling. This prediction stems from the notion that L2 learners might perform better in producing target rhythm patterns within a structured reading context than in the more fluid and variable context of spontaneous speech.

To explore the effect of speech style on rhythm production, recordings of 30 Korean learners of English were analyzed. Each participant performed two tasks in both their L1 Korean and L2 English: reading the passage *The North Wind and the Sun* and retelling the same story. This experimental design enabled us to investigate how rhythm production varies between different speech styles—reading versus spontaneous retelling—across both L1 and L2 contexts.

4.1 Metholdology

4.1.1 Materials

For Study 2, we utilized recordings from 30 Korean speakers, which were sourced from Darcy et al. (2015). While most speakers were reported as native speakers of Seoul Korean, ten speakers identified themselves as speakers of other dialects. To determine if this dialectal variation affected the results, we conducted a *t*-test analysis comparing the rhythm metrics of Seoul Korean native speakers with those of the self-reported dialectal speakers. The analysis revealed no significant differences between the two groups. Consequently, recordings from all 30 speakers were included in the study. As noted by Darcy et al. (2015), all participants were residing in the United States at the time of recording. Detailed information about the speakers is summarized in Table 5.

Table 5. Native Korean speakers' information from Darcy et al. (2015)

	Mean	S.D.	Min	Max
Age of Arrival (years)	25.6	5.4	17.0	41.0
Length of Residence (years)	2.3	1.9	0.2	5.7
Education in US (years)	2.0	1.9	0.0	6.0
L2 use (%)	49	23	10	90
Self-reporting on accentness (1-11)	5.4	2.1	2.0	9.0
Self-reporting on English proficiency (1-7)	4.7	1.1	3.0	6.8

Note: Self-reporting on accentness was self-evaluated on how accent their speech in L2 is by the speakers themselves (1 = very accented, 11 = no accent/native-like); Self-reporting on English proficiency was self-evaluated on their ability of speaking, comprehension, reading, and writing in English by the speakers themselves (1 = very low-proficient; 7 = very high-proficient).

The Korean speakers were recorded performing two tasks: reading the passage *The North Wind and the Sun* in both their native language and in L2 English, followed by a retelling of the passage in both languages. This approach allowed us to analyze rhythm production across different speech styles—reading and spontaneous retelling—in both L1 and L2 contexts.

To assess how Korean speakers' rhythm production compared to that of native

English speakers, we conducted an acoustic analysis using recordings of 26 native English speakers (14 female, 12 male; Mean Age = 19.89 years) from the ALLSTAR corpus (Bradlow n.d.). Since we could not obtain recordings from the same speakers as those used for the Korean speakers, we utilized a comparable but different type of spontaneous material. Instead of retelling *The North Wind and the Sun* passage, the native English speakers described a series of pictures without any time restrictions. We considered that both tasks—retelling a story and describing pictures—share similarities in their spontaneous nature, making them suitable for comparison in terms of rhythm production.

4.1.2 Measurements

To measure durational variability in the production, we employed the nPVI-V metric, using the same criteria as in Study 1. Among the Korean speakers, five individuals were excluded from the analysis due to excessive hesitation, unfinished utterances, or the production of fragmented series of words.

Since the recordings of native speakers were conducted without a time restriction, some recordings extended to approximately five minutes. To ensure comparability with the Korean speakers' recordings, which were restricted to a specific length, we measured durational variability only in the first minute of the native speakers' recordings.

4.1.3 Statistical analysis

To evaluate the effect of speech style and language spoken on the production of rhythm, we conducted a two-way repeated measures ANOVA. The analysis included two independent variables: *Style* (with two levels: *Read* and *Retell*) and *Language* (with two levels: *L1* and *L2*). The dependent variable was *nPVI-V*, which measures durational variability.

For the comparison with native speakers of English, we conducted a two-way mixed ANOVA in each language condition (i.e., L1 and L2). The *Group* variable (Korean speakers, Native English speakers) served as the between-subjects factor, while the *Style* variable (Read, Retell) was the within-subjects factor.

4.2 Results

Statistical analyses were conducted using R 3.5.0 (R Core Team 2023) with the assistance of the following packages: *ggplot2* (Wickham 2016). The mean nPVI-V and its standard deviation in two speech styles in L1 and L2, respectively, are presented in Table 6.

Table 6. The mean nPVI-V and the standard deviation by speech styles in L1 and L2

Language	Style	N	Mean	S.D.
L1	Read	25	58.1	6.64
	Retell	25	64.2	5.53
L2	Read	25	61.8	5.10
	Retell	25	67.3	6.56

Note: Read = reading speech style; Retell = spontaneous speech style; N = the number of speakers; S.D. = standard deviation

Figure 3 illustrates the nPVI-V values for the Korean speakers across reading and spontaneous speech styles in both L1 and L2. The results reveal a significant main effect of *Language* on durational variability, $F(1, 24) = 7.13$, $p < .05$. Korean speakers exhibited significantly lower durational variability in L1 production compared to L2 production. Additionally, there was a significant main effect of *Style* on durational variability, $F(1, 24) = 32.96$, $p < .001$. This indicates that, regardless of the language spoken, the durational variability differed based on the speech style used, with higher variability observed in spontaneous speech compared to reading style speech. Importantly, there was no significant interaction between *Style* and *Language*, $F(1, 24) = 0.097$, $p = .76$, suggesting that the effect of speech style on durational variability was consistent across both languages.

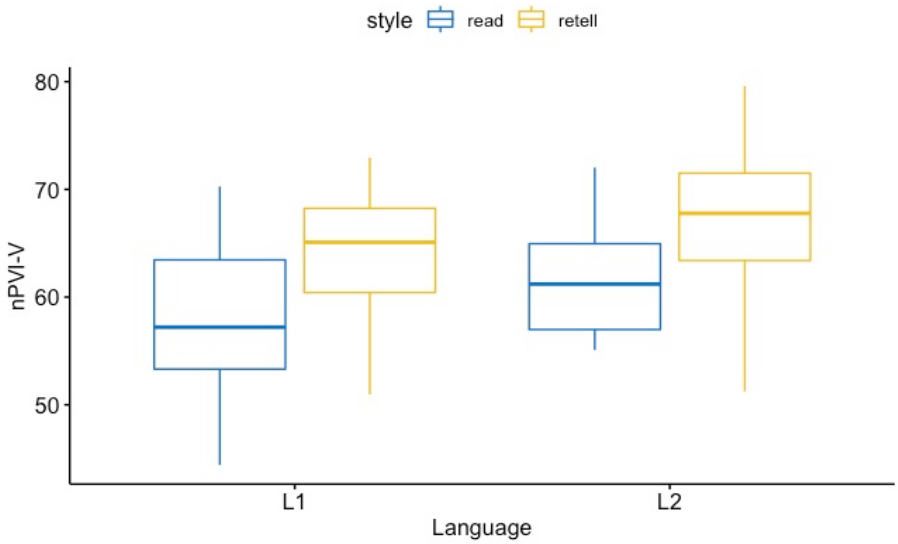

Figure 3. nPVI-V values produced by native Korean speakers in two different speech styles (Read = reading speech style, Retell = spontaneous speech style), in L1 Korean and L2 English

Table 7 and Figure 4 present a comparison of rhythm production between Korean speakers and native English speakers.

In Figure 4, the left panel depicts nPVI-V values for L1 production, where the L1 varies between the English and Korean groups. The right panel shows the nPVI-V values for English produced by both Korean and native English speakers. Note that for the English speakers, English is their native language, not a second language. The values for the English group are provided as a reference for evaluating the values of the Korean group.

The results indicate that, similar to Korean speakers, native English speakers also exhibited higher durational variability in spontaneous speech compared to reading style speech. However, the increase in durational variability from reading to retelling was more pronounced among Korean speakers, with a difference of 6.1 in L1 and 5.5 in L2. In contrast, native English speakers showed a smaller increase, with a difference of 4.9. This suggests that Korean speakers demonstrate a greater disparity in rhythm production between reading and spontaneous speech compared to native English speakers.

Table 7. The mean nPVI-V and the standard deviation by speech styles in L1 and L2 in comparison with native speakers of English

Group	Language	Style	N	Mean	S.D.
KOR	Korean	Read	25	58.1	6.64
		Retell	25	64.2	5.53
	English	Read	25	61.8	5.10
		Retell	25	67.3	6.56
ENG	English	Read	26	56	5.79
		Retell	26	60.9	5.11

Note: KOR = native speakers of Korean (target group); ENG = native speakers of English (control group); Read = reading speech style; Retell = spontaneous speech style; N = the number of speakers; S.D. = standard deviation

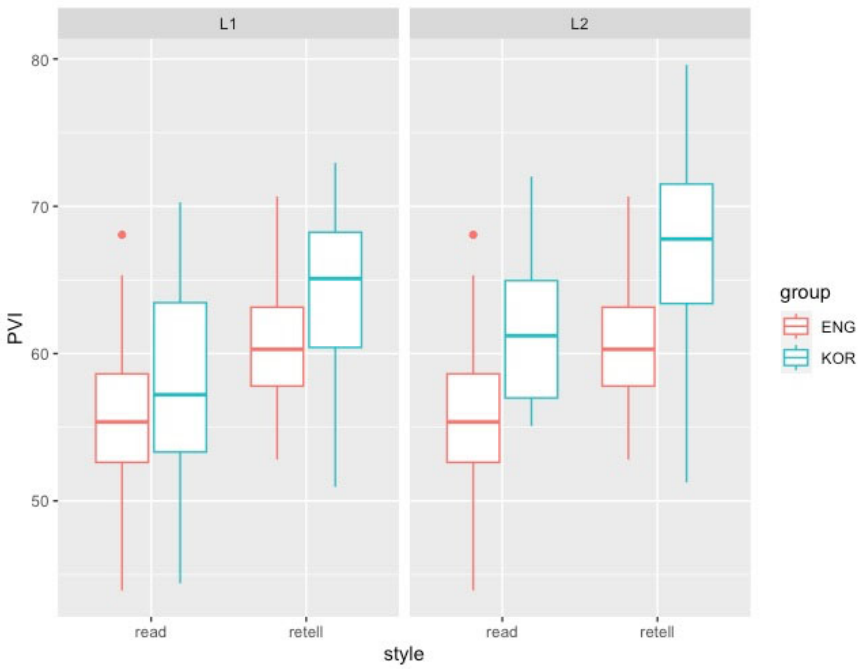

Figure 4. Comparison of nPVI-V values produced by native Korean speakers and native English speakers in two different speech styles (Read = reading speech style, Retell = spontaneous speech style); The left panel shows the results in the L1 production, and the right panel shows the L2 production. Note that the red bar (English group) shows the same data in L1 and L2.

To examine whether there is an interaction between *Group* and *Style* within each language condition, we conducted two separate two-way mixed ANOVAs: one for L1 and one for L2 conditions.

The analysis for the L1 condition revealed a significant main effect of *Style*, $F(49) = 42.71$, $p < .001$. This indicates that both English and Korean speakers produced different levels of durational variability depending on the speech style, with higher variability in the retelling condition compared to the reading condition. However, there was no significant effect of *Group*, $F(49) = 3.79$, $p = .06$. This suggests that the difference in durational variability between reading and retelling was consistent across both groups, meaning that the increase in variability was observed equally among both English and Korean speakers.

In the L2 condition, the analysis revealed a significant main effect of *Style*, $F(49) = 24.18$, $p < .001$, and a significant main effect of *Group*, $F(49) = 27.26$, $p < .001$. This indicates that both groups increased their durational variability in the retelling condition compared to the reading condition. However, there was no significant interaction between *Group* and *Style*, $F(49) = 0.081$, $p = .78$. This lack of interaction suggests that the effect of speech style on durational variability was consistent across both groups.

Notably, the *Group* effect was pronounced in the L2 condition, with Korean speakers showing higher durational variability in English rhythm compared to native English speakers. This finding is consistent with the observation that Korean speakers exhibited increased variability in L2 English compared to their L1 Korean. Given that the nPVI-V values for native English speakers in L2 were the same as those in L1, this suggests that the increased variability observed was specific to the Korean speakers' L2 production.

4.3 Discussion

The aim of Study 2 was to investigate the impact of different speech styles on rhythm production, specifically examining whether syllable-timed language speakers would show variations in rhythmic patterns based on the style of speech. The study focused on two speech styles: reading and spontaneous speech.

We formulated the following three predictions. First, we predicted that Korean

speakers would exhibit increased durational variability in their L2 English production compared to their L1 Korean production. This expectation is based on the assumption that Korean speakers, accustomed to a syllable-timed rhythm in their L1, would adapt their rhythm in English to better align with the stress-timed characteristics of the target language. Second, we anticipated that Korean speakers would produce higher durational variability in spontaneous retelling speech compared to reading speech in their L1 Korean. This prediction is grounded in the idea that spontaneous speech generally involves more natural rhythm variations, leading to greater durational variability than the more controlled rhythm of reading tasks. Third, we hypothesized that Korean speakers would produce lower durational variability in spontaneous retelling speech compared to reading speech in their L2 English. This expectation arises from the notion that L2 learners, having been trained predominantly in structured reading contexts, might perform better in producing target rhythm patterns in reading than in spontaneous retelling. These predictions aim to explore how rhythm production changes with speech style and to determine whether these changes align with or deviate from existing patterns observed in both native and non-native speech production.

The findings of Study 2 largely supported our predictions. Firstly, Korean speakers exhibited higher nPVI-V values in their L2 English production compared to their L1 Korean production, aligning with the results from Study 1. This confirms that Korean speakers increased their durational variability when producing English rhythm, contrasting with their native Korean rhythm. Additionally, the results showed that Korean speakers produced higher durational variability in the retelling speech compared to the reading speech in their L1 Korean. This indicates that speech style does affect rhythmic production, with spontaneous speech eliciting greater durational variability.

However, contrary to our expectations, Korean speakers also produced higher durational variability in the retelling condition compared to the reading condition in their L2 English production. This suggests that the tendency to increase durational variability in spontaneous speech is a consistent pattern across both L1 and L2 productions, regardless of the language spoken.

In summary, while the first two predictions were supported, the results revealed a consistent pattern where Korean speakers showed higher durational variability in spontaneous speech compared to reading, regardless of whether they were speaking

in L1 or L2. These findings will be discussed in more detail in the next section to explore their implications and how they align with existing theories on rhythm production and speech style.

5. General discussion

The current study aimed to extend our understanding of L2 rhythm acquisition by addressing the limitations of previous research, which often focused on a narrow range of languages. By incorporating a broader array of language backgrounds, we sought to identify generalized patterns in L2 rhythm acquisition. Our main findings are as follows: Firstly, both syllable-timed and stress-timed language speakers showed an increase in durational variability when producing rhythm in L2 compared to L1. This finding highlights that speakers of both rhythmic types adjust their rhythm patterns when acquiring a new language. Secondly, native Korean speakers, representing a syllable-timed language, exhibited greater durational variability in spontaneous speech compared to reading speech, irrespective of the language used. This result indicates that the influence of speech style on rhythm production is consistent across both L1 and L2, suggesting that the effect of speech style on durational variability is a robust phenomenon. These findings contribute to our understanding of rhythm production by demonstrating that both language background and speech style play significant roles in L2 rhythm acquisition.

Regarding the first finding, the distinction in rhythm production across groups in L1 was consistent with the traditional rhythmic classifications. Specifically, the data revealed significant differences in nPVI-V values between the groups. Syllable-timed language speakers exhibited lower durational variability compared to both stress-timed language speakers and native English speakers, demonstrating a distinct rhythm pattern. Conversely, stress-timed language speakers produced a level of durational variability comparable to that of native English speakers, reinforcing previous research on L1 rhythmic classification. This alignment with traditional rhythmic categories confirms that the rhythmic patterns observed in the study fit well within established theoretical frameworks.

If L2 rhythm learning had occurred, we would expect the rhythmic distinction between L1 and L2 to disappear. However, the results showed that this distinction

persisted, but in an unexpected manner. Syllable-timed language speakers did increase their durational variability in L2, achieving levels similar to those of native English speakers. Interestingly, stress-timed language speakers also increased their durational variability, surpassing the variability levels of native speakers.

A notable aspect of our findings is the increased durational variability observed in syllable-timed language speakers. Although previous studies (e.g., Li and Post 2014; Ordin and Polyanskaya 2015) have documented increases in durational variability among syllable-timed language speakers, these studies often found that such speakers still produced significantly lower variability compared to native English speakers. In contrast, our study shows that syllable-timed language speakers achieved durational variability comparable to that of native English speakers. This suggests that syllable-timed language speakers are capable of overcoming the limitations imposed by their L1 rhythm structure to align with the durational variability of L2 English. Moreover, our study's inclusion of a diverse set of 20 languages across different language families provides additional support for the notion that syllable-timed language speakers' adaptability in L2 rhythm production is a widespread phenomenon. This broader scope strengthens the evidence that the ability to adjust durational variability in L2 production is not limited to specific languages but is a more universal trait.

One unexpected finding from our study was that stress-timed language speakers also showed an increase in nPVI-V values in their L2 production, despite their L1 nPVI-V values being relatively similar to those of native English speakers. In other words, these speakers altered their rhythmic structure in L2 production, even though such a change was neither necessary nor beneficial. This result contradicts previous studies (e.g., Lin and Wang 2005; Ordin and Polyanskaya 2015) which suggested that non-native speakers benefit from learning L2 rhythm when their L1 rhythm is similar to the target language by maintaining their PVI values in L2.

Our findings imply that having an L1 rhythm that differs from the L2 rhythm is not necessarily disadvantageous. This conclusion aligns with the Speech Learning Model (SLM; Flege 1995), which posits that L2 sounds dissimilar to existing L1 categories are often acquired more effectively than those similar to L1 sounds because they are more perceptually distinct. Following this model, our study suggests that learners are more likely to acquire a 'new' L2 rhythmic structure when it differs from their L1 rhythm, rather than using the same rhythmic structure for both L1 and L2.

This extends the SLM's application beyond segmental sounds to rhythm, as supported by other studies (Zhang et al. 2008; Park 2013). Our results indicate that the SLM's claims about 'similar' and 'new' L2 sounds can be applied to L2 rhythm learning, suggesting that the distinctiveness of the L2 rhythm from the L1 rhythm facilitates its acquisition.

Despite the alignment of the general trajectory of L2 rhythm learning with the SLM, the observed increase in durational variability among stress-timed language speakers in L2 production warrants further investigation. Several factors could elucidate this phenomenon. Firstly, the relatively smaller sample size of stress-timed language speakers compared to native English and syllable-timed language speakers might have influenced the results. A reduced sample size could introduce variability that skews the findings, especially given that the number of languages in the stress-timed group was less diverse. This discrepancy could potentially affect the generalizability and reliability of the observed rhythmic patterns. Additionally, variations in phonological structures among stress-timed languages might play a significant role. Dauer (1983) argued that the degree of stress-timing can vary depending on the phonological intricacies of a language, such as consonant clusters and diphthong usage. Stress-timed languages with more complex phonological structures might exhibit different rhythmic behaviors compared to those with simpler structures. This phonological variance could contribute to the unexpected findings regarding increased durational variability in L2 production. Another plausible explanation is the hypercorrection hypothesis. According to Eckman et al. (2013) and Kelly (2022), L2 learners might overcompensate in their production to create a more discernible difference from their L1. This tendency towards hypercorrection could explain why stress-timed language speakers, despite having similar rhythmic patterns in their L1 as native English speakers, exhibited increased durational variability in their L2. Such hypercorrection might reflect a misunderstanding or overemphasis on rhythmic differences between L1 and L2. Given these possibilities, future research should delve deeper into these factors to better understand the nuances of rhythm production across different language backgrounds. By examining larger and more diverse samples, exploring phonological structures in greater detail, and considering the impact of hypercorrection, we can achieve a more comprehensive understanding of how rhythm is acquired and adapted in L2 learning. These inquiries will be crucial in refining our theoretical models and practical approaches to language instruction

and acquisition.

Turning our attention to the second research question, Study 2 aimed to investigate the impact of speech style on rhythm production, particularly in L2 contexts. Building on the findings from Study 1, which demonstrated that syllable-timed language speakers (such as Koreans) increased their durational variability when producing an L2 stress-timed language (like English), Study 2 sought to determine if this pattern would hold true across different speech styles within the same language group. The focus of Study 2 was to assess whether Korean speakers exhibited consistent rhythm production changes when shifting between two distinct speech styles: reading and spontaneous speech. This investigation is grounded in existing literature that has explored the influence of speech style on rhythm production (Thomas and Carter 2006; White and Mattys 2007; Mok and Lee 2008; Arvaniti 2012). However, the current study distinguishes itself by comparing these two speech styles with a substantial number of speakers, offering a more comprehensive examination of how speech style affects rhythm production. In Study 2, we compared rhythm production in reading speech versus spontaneous speech among Korean speakers, focusing on how each style influences durational variability. The study considered the effects within a single language group (Korean) and its impact on both L1 and L2 rhythm production. By involving a large number of speakers, the study aimed to provide a robust analysis of how different speech styles affect rhythm production.

The results from Study 2 contribute to our understanding of how speech style influences rhythmic patterns in both native and non-native contexts. They offer insights into whether Korean speakers, when producing rhythm in English, show consistent patterns across different speech styles, and how these patterns compare to their production in Korean. As observed in Study 1, there was a notable increase in durational variability from L1 Korean (a syllable-timed language) to L2 English (a stress-timed language) productions. Native Korean speakers demonstrated distinct patterns in their rhythm production depending on the speech style used. Specifically, they exhibited higher durational variability in spontaneous speech compared to reading speech. This finding aligns with previous research indicating that native Korean speakers produce more variable rhythm patterns in spontaneous speech than in reading (Mok and Lee 2008; Arvaniti 2012). A similar pattern was evident in the native English speakers' productions as well. The results revealed that native English speakers also showed greater durational variability in spontaneous speech compared to reading

speech. This consistency across both language groups suggests a general trend where speakers produce higher durational variabilities in spontaneous speech relative to reading, irrespective of the language spoken. However, these conclusions are based on the data from the current study and should be validated with a larger and more diverse sample of speakers from various language backgrounds. Future research could provide further insights into whether this pattern holds true across different languages and speech styles.

In addition to the observed increase in durational variability from L1 Korean to L2 English, we also identified a similar trend in the L2 production of native Korean speakers, which was contrary to our initial prediction. We had anticipated that Korean speakers would demonstrate greater durational variability in reading compared to spontaneous speech in their L2, given their extensive education in English. The educational system in Korea typically emphasizes reading and listening skills over speaking and writing, primarily to prepare students for exams and standardized tests. This focus on comprehension rather than production might lead one to expect that Korean speakers would show enhanced rhythm patterns in reading tasks, which are more aligned with their educational experiences. However, the results showed that Korean speakers exhibited higher durational variability in spontaneous speech than in reading speech, even in their L2 English production. This suggests that the expected advantage in reading tasks due to educational training did not materialize. Instead, the data reveal that the Korean speakers' rhythmic production in spontaneous speech matched the pattern observed in their L1, contrary to the expectation based on their educational focus. This finding indicates that despite the educational emphasis on reading and listening, the natural rhythm patterns produced in spontaneous speech might still prevail, highlighting the complexity of how rhythm production is influenced by both language and speech style. Future research could further explore this discrepancy to better understand the factors affecting rhythm production in L2 contexts.

Despite our initial expectations, the results indicated that Korean speakers displayed a consistent pattern across both languages. Specifically, they produced higher durational variability in spontaneous speech (retelling in this study) compared to reading tasks, regardless of whether they were speaking in their native Korean or their second language, English. This unexpected finding suggests that the rhythm pattern associated with speech style was maintained across L1 and L2 productions.

This consistent pattern might be attributed to individual traits that persist regardless of the language spoken. Analyzing individual recordings revealed that fluent readers were generally able to retell stories more smoothly, whereas less fluent readers exhibited more stuttering, hesitations, and final lengthening during retelling. This observation was consistent across both languages: speakers who displayed disfluency in their L1 also showed similar traits in their L2 production. If individual traits are indeed robust and influence speech patterns in both L1 and L2, it is reasonable to expect that the durational variability related to speech style differences would remain consistent across languages. This aligns with Bradlow et al.'s (2017) findings, which showed that L1 speaking rate is a strong predictor of L2 speaking rate, regardless of the language spoken. Their study found a high correlation between speaking rates in L1 and L2, suggesting that speakers who talk quickly in their native language are likely to do the same in their second language, and vice versa for slower speakers. In a similar manner, we speculate that Korean participants who exhibited high durational variability in spontaneous speech and low variability in reading tasks in their L1 are likely to exhibit the same pattern in their L2. This pattern might be amplified in spontaneous speech due to its less controlled nature compared to reading tasks, leading to even greater differences in durational variability. Thus, the enduring nature of individual speaking traits across languages could explain why the effect of speech style on durational variability was similar in both L1 and L2 for the Korean speakers.

In addressing the classification of Korean within the traditional rhythm typology, it's essential to recognize the ongoing debate and evolving perspectives on this issue. Historical classifications have often positioned Korean as a syllable-timed language, characterized by regularity in syllable durations. However, recent research has challenged this rigid classification, suggesting that Korean exhibits a blend of phonological features that defy simple categorization. Seong (1995) highlighted that Korean encompasses both stress-timed characteristics, such as final lengthening and the occurrence of tap sounds, and syllable-timed features, like a simple syllabic structure. This complexity has led to the argument that Korean does not fit neatly into either rhythmic class. Mok and Lee (2008) further illustrated this point by showing that the classification of Korean's durational variability can vary significantly depending on the metrics employed. For instance, while metrics such as %V and PVI tend to classify Korean as syllable-timed, other metrics like ΔC suggest a more intermediate position between stress-timing and syllable-timing. The debate extends beyond Korean,

with broader implications for rhythm classification. Some scholars argue against a rigid three-class system, proposing instead that languages should be understood along a continuum or within intermediate categories. Nespor (1990) and Dauer (1983, 1987) argue that languages might not fit neatly into distinct stress-timed or syllable-timed categories but rather fall somewhere along a spectrum. For example, Polish, despite being labeled as stress-timed, does not exhibit vowel reduction, whereas Catalan, considered syllable-timed, does feature vowel reduction. Dauer (1987) proposed that all languages exhibit varying degrees of syllable-timing or stress-timing, suggesting a more nuanced approach to classification. In light of these perspectives, Korean might be viewed as occupying an intermediate position in rhythm classification, as suggested by Nespor's continuum model. Alternatively, it could be characterized as syllable-timed in reading speech contexts and more stress-timed in spontaneous speech contexts, aligning with Dauer's view of variable rhythmic timing based on speech style. This nuanced understanding underscores the need for further research to clarify how rhythm timing interacts with different speech styles and to refine our understanding of rhythm classification across diverse languages.

6. Conclusion

The current study explored the influence of a native language on the production of L2 rhythm, with a focus on both cross-linguistic and within-language comparisons. Two primary findings emerged from the research. First, our analysis, which included an extensive corpus of L1 and L2 productions from a diverse range of languages, revealed that speakers of syllable-timed languages produced L2 rhythm with durational variability more similar to native English speakers than those from stress-timed languages. This finding challenges the traditional view that learners from rhythmically dissimilar languages are inherently disadvantaged in acquiring L2 rhythm. Instead, it suggests that speakers of syllable-timed languages can achieve L2 rhythm patterns comparable to native speakers, despite the inherent rhythmic differences between their native and target languages. This result enhances our understanding of L2 speech learning and supports the broader application of the Speech Learning Model (SLM; Flege 1995). According to the SLM, learners are generally more successful at acquiring new L2 speech sounds that are distinct from their L1 categories rather than those

that are similar. Our findings extend this principle to rhythm acquisition, demonstrating that learners from rhythmically contrasting language backgrounds (syllable-timed) may actually acquire L2 rhythm more effectively than those from rhythmically similar backgrounds (stress-timed). This supports the idea that the process of acquiring L2 rhythm may benefit from the contrast between L1 and L2 rhythmic structures, rather than being hindered by it.

Another significant finding of this study is that rhythm patterns from L1 persist into L2 production. Contrary to our expectation that Korean speakers would exhibit lower durational variability in spontaneous speech compared to reading in L2—due to their educational focus on reading rather than speaking—the results showed a consistent pattern: higher durational variability in spontaneous speech than in reading, both in L1 and L2. This persistence of L1 rhythmic characteristics in L2 production suggests a strong influence of individual traits. Much like speech rate, where individual speaking styles in L1 have been shown to carry over into L2 (Bradlow et al. 2017), rhythmic patterns in L1 might reflect unique aspects of an individual's speech style that are retained across languages. Although this interpretation is speculative and not yet supported by direct evidence, it aligns with findings from related studies. For instance, Kartushina and Frauenfelder (2014) found that speakers with greater precision in their L1 productions tended to exhibit higher accuracy in L2 category production. Similarly, Oh and Park (2023) demonstrated that native Korean speakers who produced high vowels with greater compactness in Korean also showed higher accuracy in producing high vowels in L2 English. These observations suggest that rhythmic properties, along with other speech characteristics, could serve as valuable indicators of individual differences among L2 learners. Future research investigating these rhythmic properties may offer further insights into how individual speech patterns influence L2 acquisition.

In conclusion, our findings reveal that native speakers of a rhythmically contrastive language to the target language, such as syllable-timed language speakers learning a stress-timed language as an L2, are capable of acquiring rhythm that closely resembles that of native speakers. This adaptation in L2 rhythm production is consistent across various speech styles within a single language group. While the generalizability of these results warrants further exploration, this study supports the SLM's assertion about the ease of acquiring 'new' L2 segments, extending this claim to L2 rhythm learning. Additionally, our research contributes to the growing body of evidence demonstrating

that speech style influences rhythm production, a trend that appears to hold across different languages. We hope this study encourages further research into these important aspects of L2 rhythm acquisition.

References

Abercrombie, David. 1967. *Elements of general phonetics.* Edinburgh: Edinburgh University Press.

Arvaniti, Amalia. 2012. The usefulness of metrics in the quantification of speech rhythm. *Journal of Phonetics* 40(3): 351-373.

Barry, William J., Bistra Andreeva, Michela Russo, Snezhlna Dimitrova, and T. Kostadinova. 2003. Do rhythm measures tell us anything about language type? *Proceedings of the 15th International Congress of Phonetic Sciences (ICPhS-15),* 2693-2696.

Boersma, Paul and David Weenink. 2018. Praat: Doing phonetics by computer [Computer program]. Version 6.0.37. Retrieved 14 March 2018 from http://www.praat.org/.

Bradlow, Ann R. n.d. *ALLSSTAR: Archive of L1 and L2 scripted and spontaneous transcripts and recordings*. Retrieved from https://speechbox.linguistics.northwestern.edu/#!/?goto=allsstar.

Bradlow, Ann R., Midam Kim, and Michael Blasingame. 2017. Language-independent talker-specificity in first-language and second-language speech production by bilingual talkers: L1 speaking rate predicts L2 speaking rate. *The Journal of the Acoustical Society of America* 141(2): 886-899.

Cruz-Ferreira, Madalena. 1995. European Portuguese. *Journal of the International Phonetic Association* 25(2): 90-94.

Darcy, Isabelle, Hanyong Park, and Chung-Lin Yang. 2015. Individual differences in L2 acquisition of English phonology: The relation between cognitive abilities and phonological processing. *Learning and Individual Differences* 40: 63-72.

Das, Tanusree, Latika Singh, and Nandini C. Singh. 2008. Rhythmic structure of Hindi and English: New insights from a computational analysis. *Progress in Brain Research* 168: 207–272.

Dauer, Rebecca M. 1983. Stress-timing and syllable-timing reanalyzed. *Journal of Phonetics* 11(1): 51-62.

Dauer, Rebecca M. 1987. Phonetic and phonological components of language rhythm. *Proceedings of the International Congress of Phonetic Sciences* 11: 447-450.

Dellwo, Volker. 2006. Rhythm and speech rate: A variation coefficient for delta C. *Language and language-processing: Proceedings of the 38th Linguistics Colloquium,* 231–241.

Dellwo, Volker and Petra Wagner. 2003. Relationships between speech rate and rhythm. *Proceedings of the International Congress of Phonetic Sciences* 15: 471-474.

Deterding, David. 2011. Measurements of the rhythm of Malay. *Proceedings of the International Congress of Phonetic Sciences* 17: 576-579.

D'Imperio, Mariapaola. 2002. Italian intonation: An overview and some questions. *International Journal of Romance Linguistics* 14(1): 37-69.

Eckman, Fred R., Gregory K. Iverson, and Jae Yung Song. 2013. The role of hypercorrection in the acquisition of L2 phonemic contrasts. *Second Language Research* 29(3): 257-283.

Flege, James E. 1995. Second language speech learning: Theory, findings, and problems. In Winifred Strange (ed.), *Speech perception and language experience: Issues in cross-language research,* 233-277. Timonium, MD: York Press.

Gibbon, Dafydd and Ulrike Gut. 2001. Measuring speech rhythm. *Proceedings of European Conference on Speech Communication and Technology* 7: 91–94.

Gonzalez, Andrew. 1970. Acoustic correlates of accent, rhythm, and intonation in Tagalog. *Phonetica* 22(1): 11-44.

Grabe, Esther and Ee Ling Low. 2002. Durational variability in speech and the rhythm class hypothesis. In Carlos Gussenhoven and Natasha Warner (eds.), *Laboratory phonology* volume 7, 515-546. Berlin; New York, NY: De Gruyter Mouton

Gut, Ulrike. 2003. Prosody in second language speech production: The role of the native language. *Fremdsprachen Lehren und Lernen* 32: 133–152.

Gut, Ulrike. 2012. Rhythm in L2 speech. In Dafydd Gibbon, Daniel Hirst, and Nick Campbell (eds.), *Rhythm, melody and harmony in speech: Studies in honour of Wiktor Jassem, special edition of Speech and Language Technology 14/15*, 83-94. Poznań: Polskie Towarzystwo Fonetyczne.

Haghshenas, Ali Mohammad. 1978. *Avashenasi [Phonetics].* Tehran: Agah Press.

Jeon, Hae-Sung. 2015. Prosody. In Lucien Brown and Jaehoon Yeon (eds.), *Handbook of Korean linguistics*, 41-58. New York, NY: Wiley-Blackwell.

Kartushina, Natalia and Ulrich H. Frauenfelder. 2014. On the effects of L2 perception and of individual differences in L1 production on L2 pronunciation. *Frontiers in Psychology* 5: Article 1246.

Kassambara, Alboukadel. 2023. Ggpubr: 'Ggplot2' based publication ready plots. R package version 0.6.0. https://rpkgs.datanovia.com/ggpubr/.

Kelly, Niamh. 2022. Change across time in L2 intonation vs. segments: A longitudinal study of the English of Ole Gunnar Solskjaer. *Languages* 7(3): Article 210.

Kohler, Klaus. 1982. Rhythmus im Deutschen [Rhythm in German]. *Arbeitsberichte, Institut für Phonetik der Universität Kiel* 19: 89-106.

Lee, Ho-Young and Jieun Song. 2019. Evaluating Korean learners' English rhythm proficiency with measures of sentence stress. *Applied Psycholinguistics* 40(6): 1363-1376.

Lee, Hyunbok. 1982. Hangugeo lideumui eumseonghagjeog yeongu [A phonetic study on Korean rhythm]. *Malsori* 4: 31-48.

Lennon, Paul. 1990. Investigating fluency in EFL: A quantitative approach. *Language Learning*

40(3): 387–417.
Li, Aike and Brechtje Post. 2014. L2 acquisition of prosodic properties of speech rhythm: Evidence from L1 Mandarin and German learners of English. *Studies in Second Language Acquisition* 36(2): 223-255.
Lin, Hua and Qian Wang. 2005. Vowel quantity and consonant variance: A comparison between Chinese and English. Presented at *Between Stress and Tone conference (BeST)*. Leiden, The Netherlands: International Institute for Asian Studies. June 16-18.
Lin, Hua and Qian Wang. 2007. Mandarin rhythm: An acoustic study. *Journal of Chinese Linguistics and Computing* 17(3): 127-140.
Low, Ee Ling, Esther Grabe, and Francis Nolan. 2000. Quantitative characterizations of speech rhythm: Syllable-timing in Singapore English. *Language and Speech* 43(4): 377-401.
Luangthongkum, Theraphan. 1977. *Rhythm in standard Thai*. PhD Dissertation. University of Edinburgh.
Major, Roy C. 1981. Stress-timing in Brazilian Portuguese. *Journal of Phonetics* 9(3): 343-351.
Major, Roy C. 1985. Stress and rhythm in Brazilian Portuguese. *Language* 61(2): 259-282.
Miller, M. 1984. On the perception of rhythm. *Journal of Phonetics* 12(1): 75-83.
Mok, Peggy and Sang Im Lee. 2008. Korean speech rhythm using rhythmic measures. Presented at the *18th International Congress of Linguists (ICL18)*. Seoul, South Korea. July 21-26.
Mok, Peggy and Volker Dellwo. 2008. Comparing native and non-native speech rhythm using acoustic rhythmic measures: Cantonese, Beijing Mandarin and English. *Proceeding of Speech Prosody* 4: 423-426.
Munro, Murray J. and Tracy M. Derwing. 1995. Foreign accent, comprehensibility, and intelligibility in the speech of second language learners. *Language Learning* 45(1): 73–97.
Munro, Murray J. and Tracy M. Derwing. 1998. The effects of speaking rate on listener evaluations of native and foreign-accented speech. *Language Learning* 48(2): 159–182.
Nazzi, Thierry, Josiane Bertoncini, and Jacques Mehler. 1998. Language discrimination by newborns: Toward an understanding of the role of rhythm. *Journal of Experimental Psychology: Human Perception and Performance* 24(3): 756-766.
Nespor, Marina. 1990. On the rhythm parameter in phonology. In Iggy M. Roca (ed.), *Logical issues in language acquisition*, 157-176. Dordrecht: Foris.
Oh, Sujin and Hanyong Park. 2023. L1 category precision hypothesis in L2 production: Korean learners' English front vowels. *The Journal of the Acoustical Society of America* 153(3): A342-A342.
Ordin, Mikhail and Leona Polyanskaya. 2014. Development of timing patterns in first and second languages. *System* 42: 244-257.
Ordin, Mikhail and Leona Polyanskaya. 2015. Acquisition of speech rhythm in a second language by learners with rhythmically different native languages. *The Journal of the Acoustical Society of America* 138(2): 533-544.
Park, Hanyong. 2013. Detecting foreign accent in monosyllables: The role of L1 phonotactics.

Journal of Phonetics 41(2): 78-87.

Pike, Kenneth L. 1945. *The intonation of American English.* Ann Arbor, MI: University of Michigan Press.

Prahallad, Kishore S. and Alan W. Black. 2003. Unit size in unit selection speech synthesis. *Proceeding of Eighth European Conference on Speech Communication and Technology*, 1317-1320.

Ramus, Franck, Marina Nespor, and Jacques Mehler. 1999. Correlates of linguistic rhythm in the speech signal. *Cognition* 73(3): 265-292.

Schiering, René. 2007. The phonological basis of linguistic rhythm: Cross-linguistic data and diachronic interpretation. *STUF-Language Typology and Universals* 60(4): 337–359.

Seong, Cheoljae. 1995. *Hangugeo lideumui silheomeumseonghagjeog yeongu : Sigan gujowa gwanlyeonhayeo [The experimental phonetic study of the standard current Korean speech rhythm: With respect to its temporal structure].* PhD Dissertation. Seoul National University.

Stockmal, Verna, Dace Markus, and Dzintra Bond. 2005. Measures of native and non-native rhythm in a quantity language. *Language Speech* 48(1): 55–63.

Stockwell, Robert P. 1957. *A contrastive analysis of English and Tagalog.* Los Angeles, CA: University of California Los Angeles.

Tamati, Terrin N., Hanyong Park, and David B. Pisoni. 2011. The development of a new corpus of foreign-accented English. Presented at the *2011 VLSP New Tools and Methods for Very Large-Scale Phonetics Research Workshop.* Philadelphia, PA: University of Pennsylvania. January 29-31.

Thomas, Erik R. and Phillip M. Carter. 2006. Prosodic rhythm and African American English. *English World-Wide* 27(3): 331-355.

White, Laurence and Sven L. Mattys. 2007. Calibrating rhythm: First language and second language studies. *Journal of Phonetics* 35(4): 501–522.

Wickham, Hadley. 2016. *Ggplot2: Elegant graphics for data analysis.* Retrieved from https://ggplot2.tidyverse.org.

Zhang, Yanhong, Shawn L. Nissen, and Alexander L. Francis. 2008. Acoustic characteristics of English lexical stress produced by native Mandarin speakers. *Journal of the Acoustical Society of America* 123(6): 4498–4513.

Sujin Oh
Lecturer
Department of English Language and Literature
Soongsil University
369, Sangdo-ro, Dongjak-gu
Seoul 06978, Korea
E-mail: sujinoh971@ssu.ac.kr

Hanyong Park
Associate Professor
Department of Linguistics
University of Wisconsin-Milwaukee
Milwaukee, WI 53201, USA
E-mail: park27@uwm.edu

Received: 2024. 08. 26.
Revised: 2024. 11. 20.
Accepted: 2024. 11. 21.