https://doi.org/10.24303/lakdoi.2019.27.2.87

Learning Syllable-timed Prosody as a Native Speaker of Stress-timed Prosody*

Jong-mi Kim
(Kangwon National University)

Kim, Jong-mi. (2019). Learning syllable-timed prosody as a native speaker of stress-timed prosody. *The Linguistic Association of Korea Journal, 27*(2), 87-107. To see how syllable-timed prosody is learned by a native speaker of a language that uses stress-timed prosody, duration and pitch in the L2 Korean speech of L1 English learners are explored. For analysis, 28 learners were divided into low- and high-proficiency groups, based on native Korean-speaking listeners' evaluation of the learners' read Korean speech data. In addition, 12 native speakers of Korean recorded their own Korean speech as a control group and as model speakers. The pairwise variability of 28 adjacent vowels within each accentual phrase was measured in terms of duration and fundamental frequency. The results showed that learners used both pitch (fundamental frequency) variability and duration variability for prominence acquisition, whereas the native speaker controls used only pitch variability. These results suggest the presence of prosodic transfer of contrastive duration for prosodic acquisition of contrastive pitch in adult language learning of syllable-timed prosody.

Key Words: prosodic transfer, contrastive duration, contrastive pitch, English-speaking learners of Korean, pairwise variability, fundamental frequency, vowel duration, speech rhythm

1. Introduction

Although there are a substantial number of studies on L2 speech learning of contrastive duration or stress-timing, in languages such as English (e.g., Behrman et al., 2019; Deterding,

* This work was supported by the Ministry of Education of the Republic of Korea and the National Research Foundation of Korea (NRF-2016S1A5A2A01024508). An earlier version of this paper was presented at the 20th International Congress of Linguists (ICL20), July 2018, Cape Town, South Africa. Parts of the speech material in this study have been collected by the author during the accent reduction course for native speakers of American English learning Korean at the Center for Korean Studies, University of Hawaii (Kim, 2012). The author is particularly grateful to the volunteer learners in the course.

2001; Gilbert, 2019; Jang, 2008), far fewer studies have discussed how syllable-timedness may be learned in a second language (van Maastricht et al., 2018). The traditional view contrasting stress-timed and syllable-timed speech rhythms is still well received (Ladefoged & Johnson, 2015, pp. 124-126, 261-264), although in a revised form that models it as contrasting the presence and absence of stronger and weaker elements (Nolan & Jeon, 2014, p. 1), instead of aligning speech with an external regularity (Abercrombie, 1967; Pike, 1945). The contrastive elements in question include not only duration but also pitch.

On the basis of this revised definition of speech rhythm, we explore in this paper how syllable-timed prosody is learned. A basic description of the difference between the two rhythms is as follows: some languages such as French, Spanish, Korean, and Chinese can be called syllable-timed languages, in which adjacent syllables have similar vowel duration (Low et al., 2000; Nolan & Jeon, 2014). In contrast, languages such as Thai, Dutch, German, and English could be called stress-timed, in that these languages have heavy stresses and large variations in vowel length (Ladefoged & Johnson, 2015, p. 263). In English, for example, there are short or reduced vowels interspersed with long ones.

1.1 Prosodic differences between Korean and English

To see whether syllable-timedness of speech rhythm is learnable, we chose to investigate L2 Korean learner speech from L1 English speakers. Duration is an important prosodic feature[1] in English, with its variable word stress (Kochanski et al., 2005; Ladefoged & Johnson, 2015, p. 259), while pitch is a primary prosodic feature in Korean, an intonational language (Jun, 2005a). Fundamental frequency (f_0) does not predict prominence in most varieties of English (Kochanski et al., 2005), because the perceptual salience of stressed syllables[2] generally involves increased duration, but not increased pitch. Instead, the pitch is signaled by either lower low tones or higher high tones (Ladefoged & Johnson, 2015, p. 260).[3] On the other hand, fundamental frequency (f_0) or pitch primarily marks accentual phrases in Korean and rises at the end (Jun & Oh, 2000).

1) "Prosodic features" are prominent among the various "prosodic properties" of rhythm, such as pitch, duration, amplitude, and tempo.

2) Stressed syllables are special and more important than the others (Beckman, 1986).

3) Loudness predicts prominence, as evident in the following example: *Did you see my father?* Here, the *fa* in the word *father* becomes prominent by increased loudness, but not by heightened fundamental frequency.

As the English and Korean languages have these sharply contrasting prosodic features, English-speaking learners of the Korean language may plausibly choose one prosodic cue (duration), based on its role in their L1, or another (f_0) derived from their L2. These learners inevitably encounter the prosodic difference between the absence of stresses in their target language and the presence of stresses in their native language, and conversely between the presence of accentual phrases in their target language and their absence in their native language (Jun, 2005b, pp. 434-435). We therefore examine developmental change in their use of L2 prosodic cues in relation to duration and pitch. In particular, we shall observe prosodic transfer in L2 Korean speech by native speakers of English, who are acquiring the different prosodic prominence cues of the target language.

1.2 Research questions

Given the prosodic use differences in duration and pitch between the Korean and English languages, we question whether learners from the L1 English background with alternating duration features, would use this available prosodic cue — duration — to compensate for the unavailable pitch cue in their L2 Korean. That is, will these L1-stress-timed learners use the L2 cue (f_0), the L1 cue (duration), or both to express the target syllable-timed prosody? Given that it is very difficult in general for adults to acquire L2 phonology (Scovel, 1988), we address three research questions intended to render L2 prosody into measurable terms in (1) below: one on transfer, one on acquisition, and one on development.

(1) Research questions
   a. PROSODIC TRANSFER: Will English-speaking Korean learners use the duration property (as measured by time (*s*)) to express pitch variability in Korean?
   b. PROSODIC ACQUISITION: Will English-speaking Korean learners use the pitch property (as measured by fundamental frequency (Hz)) to express pitch variability in Korean?
   c. PROSODIC DEVELOPMENT: Will English-speaking Korean learners increasingly use the pitch property (as measured by fundamental frequency (Hz)), the duration property (as measured by duration (*s*)), or both to express pitch variability in Korean?

For (a), whether L1-English syllable duration is transferred to express L2 Korean prosody to express pitch variability in L2, the answer can resolve disagreeing views in the literature as

to whether there is a speech rhythm distinction between languages. Pitch variability can be measured by calculating the pitch difference between pairs of adjacent syllables as in duration variability (Low et al., 2000). The negative view is that phonetic reality does not show the so-called stress-timed vs. syllable-timed distinction (Roach, 1982).[4] The positive view is that the rhythm distinctions in native speech are quantifiable, highlighting the difference in their L2 speech rhythm (Ordin & Polyanskaya, 2015).

In (b), we question whether L2-Korean prosodic pitch variability will be acquired by an L1 English speaker, although pitch lends little to prominence in English (Kochanski et al., 2005). The answer to this question will help resolve disagreeing views in the literature as to whether speech rhythm is learnable. The negative view is that stress-timed rhythm is not easy to learn (see van Maastricht et al., 2018 for Spanish learners of Dutch). The positive view is that syllable-timed rhythm is learnable (van Maastricht et al., 2018 for Dutch learners of Spanish). Taking these findings and considerations together, we expect that Korean rhythm will be learnable by English speakers. We will test the validity of this expectation in this study.

In (c), we question whether L1 English speakers do or do not develop pitch variability in L2 Korean. The answer to this question will help resolve disagreeing views in the literature as to what phonetic features are used to measure speech rhythm. Duration has been claimed to measure prominence, while fundamental frequency does not, at least in English (Kochanski et al., 2005). On the other hand, pitch (f_0) conveys rhythmicality for German and Swiss-French listeners (Barry et al., 2009; Cumming, 2011). The research gap we address in (c) is whether fundamental frequency conveys syllable-timed rhythm in Korean.

2. Methods

2.1. Participants

The experimental group comprised 28 native American English speakers who were learning Korean as a foreign language; their participation was voluntary. The control group, on the other hand, comprised nine native Korean speakers. All participants read and recorded a set of 16 sentences twice, for pre- and post-test. The experimental group received weekly lessons, while the control group did not. In addition, 10 native Korean listeners evaluated the proficiency of both the native speakers' and the learners' recorded speech. Also participating were three native

4) Roach (1982) argued that the traditional view of contrasting stress-timed with syllable-timed speech rhythm (Abercrombie, 1967; Pike, 1945) was not supported by aligning speech with an external regularity.

Korean model speakers whose recorded speech was downloadable from the class website for the learners to listen to and repeat afterward, for speech practice. These participants' information is given in Table 1.

Table 1. Speaker information by group

	Number (Male/Female)	Mean age
L2 Korean learners of L1 English speakers	28 (12/16)[5]	20.8
L1 Korean controls	9 (4/5)	33.5[a]
L1 Korean listeners of L2 Korean speech	10 (5/5)	23.5
L1 Korean model speakers	3 (1/2)	26.4

Note. L1 = Native language, L2 = Target foreign language
[a]Approximation based on their career history

The learner subjects indicated that English was their best language at the time of data collection and the language they mainly used during childhood. They had previously learned the Korean language for more than six months, and their language proficiency was high enough to read Korean text written in the Korean alphabet. We did not limit the type or duration of learning experience, which varied from institutional to interactional. We included two heritage learners of low and high proficiency. Two learners, regardless of heritage, had lived in Korea for one or more years.[6] All the research participants were college-educated adults who were native speakers of either Korean or English, depending on their designated roles. None of the Korean native speaker participants possessed regional accents of the southern, northern, or eastern areas.[7]

5) The pitch difference between male and female is normalized in the pairwise variability analysis by dividing individual pitch difference by the average pitch of each pair.

6) One non-heritage learner lived only one year in Korea, and the other, who was a heritage learner, lived in Korea for 10 years. We included both in our analysis, because both met the native English speaker criteria of English being their language of growth and their best language.

7) The dialect of the mid-western area is used in national broadcasting and is usually considered "standard."

2.2 Speech materials

Speech materials were the 16 sentences in (2).[8] These sentences embedded 14 pairs of test vowels (marked by brackets), and four sentences were randomly selected for native listener evaluation of proficiency (marked by asterisks, *).[9] The sentences were chosen from an open-access sentence list (Choo & O'Grady, 2003), to facilitate future replication studies by other researchers. All sentences were presented in the Korean alphabet (shown in appendix) with their English translation, when the participants performed experimental tasks. The transliteration in (2) followed the official Korean language Romanization (Korea Ministry of Culture, Sports and Tourism, 2014).

(2) Speech materials (Presented in the Korean alphabet in the Appendix)

a. g[a]td[a] wasseoyo. 'I went and have come back.'
b. m[eo]g[i] jjalbayo. 'The ink stick is short.'
c. n[u]g[u] teogiya? 'Who pays and treats us?'
d.* geumni bissajo? 'Is interest high?'
e. g[i]g[a] mojarayo. 'He is short of vitality.'
f. g[eu]g[i] sing'geoweoyo. 'The pole is not salty enough.'
g. oneul [eo]d[i] gayo? 'Where are we going today?'
h. b[a]m[e]neun chayo. 'At night, it is cold.'
i. j[a]pb[i]ga deureoyo. 'It requires miscellaneous expenses.'
j. y[eo]g[i]seo seoyo. 'Please stop here.'
k.* huchu jom juseyo. 'Please pass me the pepper.'
l.* j[a]g[o]naseo halge. 'I will do it after sleeping.'
m.* sillihageul g[o]ngb[u]hamnida. 'I study utilitarianism.'
n. b[a]d[a]nmuleun jjajo. 'Sea water is salty, of course.'
o. d[o]s[eo]gwane chaek bilireo gayo? 'Do you go to the library to borrow a book?'
p. anyo, d[o]s[eo]gwane ilhareo gayo. 'No, I go to the library to work.'

Each sentence in the reading list (2a) through (2d) is composed of five syllables, while the sentences in the list (2e) through (2p) contain a varied number of syllables. The selected pairs of vowels in (2) are the first and second vowels embedded in an accentual phrase within the given sentence in order to control the prosodic environment and to vary number of syllables,

8) The materials have been selected from the speech in a voluntary class, whose purpose was overall pronunciation teaching (Kim, 2012). The learners read one of the comparing words within a sentence in the textbook (see appendix). These audio samples also fulfilled the purpose of this paper, because learners were reading artificial words in fixed phrases, instead of memorized or familiar sentences.

9) The random selection was performed to include two additional sentences in the experimental data (2d, 2k), on the open-access sentence list (Choo & O'Grady, 2003).

syllable weight, and position in the sentence.[10] The test vowel pairs showed a statistically significant rising pitch in the native Korean speech ($p < .05$ by t-test between the f_0 values of the preceding and following vowels).

The phonetic environment of these test words was chosen carefully to ensure reliable measurement; none of the vowels were adjacent to sounds that may induce incorrect duration or pitch measurement. Thus, the liquid and glide sounds [l, r, w, y] were avoided so that the acoustic transition to the adjacent vowel would not cause incorrect duration measurement, while the aspirated sounds [h, p^h, t^h, k^h, c^h] were avoided so that the f_0 values of adjacent vowels would not be heightened. Final lengthening (Turk & Shattuck-Hufnagel, 2007) was avoided as much as possible so that none of these vowel pairs would be followed by an intonational phrase boundary.

If Korean were syllable-timed, and other factors such as segmental environment and number of syllables were equal, then the embedded vowels would have similar duration.[11] For example, the two [u] vowels in (2c) may have isochronic duration, relevant for syllable-timing, but different duration for stress-timing. Vowel reduction could have occurred if it were a stress-timed language such as English. Pitch has also been presented as an indicator of speech rhythm (Cumming, 2011), and all these pairs of adjacent vowels show a statistically significant difference in pitch in the native speaker controls' speech.

The sentences for listening evaluation in (2) were chosen from sentences with varied syntactic structure (question, statement, subject omission) and speech style (formal, informal).

2.3 Procedure

To acquire the data, the native and non-native speeches were first recorded; the recorded speeches were then evaluated by the native listeners to divide the proficiency group into high and low learners as well as native speakers; finally, each group of speech was analyzed for duration and pitch variability.

10) Statistical significance in native speech is an important selection criterion, because individual pitch productions vary. For example, two consecutive syllables may have low pitch (LLH in ToBI transcription convention (Jun, 1998, p. 193)) in an accentual phrase comprising three syllables, as in *b[a]m[e]neun* in sentence (2h).

11) Vowel length may be affected by the presence of stress, a syllable coda, voicing of the following consoants, or the number of syllables (Ladefoged & Johnson, 2015, pp. 107-108).

(3) Procedure

a. Learners first listened to the native Korean speech recordings and repeated after each sentence for practice before pre-test.
b. Learners' read speech was recorded twice (one pre-test recording and another post-test recording) before and after five weekly lessons, with one session per week, on various aspects of Korean pronunciation.
c. Native speech was recorded twice without a lesson.[12)]
d. Ten native Korean-speaking listeners rated the proficiency of recorded L1 and L2 Korean speech on a scale of 1-7.
e. Recorded speech was measured for vowel duration (*s*) and pitch (f_0).

In (3a), learners first listened and repeated after the native speech so that they could be tested on heard forms. In (3b), there were five weeks of practice time during which learners were given the model speech productions of approximately 100 sentences by the three native speaker models.[13)] Lessons were on general and various aspects of Korean pronunciation, including consonants, vowels, phonological adjustments, intonation, and rhythm. Example texts were from two books (Choo & O'Grady, 2003; Shin, 2008). Multiple model speeches were given so that learning could not be based on memorizing a particular speech sample. Reading the prompt list was to control for contextual variation. We chose read speech over conversational speech to ensure that intonation and duration were minimally influenced by conversational or emphatic variation among the speakers. In (3c), native speakers read the sentence list in (2) twice at a time, and the second reading was taken as data to allow the speakers to familiarize the content for natural reading.

In (3d), ten native Korean speakers evaluated the proficiency of the recorded L1 and L2 Korean speech on a 7-point Likert scale (1 = very poor, 7 = definitely native). Both the native controls' and learners' speeches were mixed and randomized to ensure blind testing. The four selected sentences in (2) were played, presented, and rated on a computer screen. A rating session was completed in less than 30 minutes per person, to avoid cognitive burden.

On the basis of the listening evaluation of proficiency, learners were divided into two groups, high and low proficiency as in Table 2.

12) The first and second recordings of native speech were used to compare with the pre-test and post-test recordings of learner speech, respectively.

13) There were 100 model sentences including some comparing sentences and some question-and-answer pairs. All sentences were assigned as homework to students, and some were used as examples in the lessons.

Table 2. Proficiency groups of learners as determined by native listeners

Proficiency	Pre-test			Post-test			*N* (speaker)
	Min	Max	Range	Min	Max	Range	
Low-level	1.13	2.30	1.17	1.14	3.35	2.21	16
High-level	2.43	4.23	1.80	2.38	4.86	2.48	12
Native speech	5.39	6.69	1.30	5.17	6.62	1.45	9

As shown in Table 2, there was a clear difference in score between native speech and learner speech; the former was rated between 5.2 and 6.7, while the learner speech ranged from 1.1 to 4.9.[14] Further, the average improvement in learner speech was large (0.43), while that of native speech was very small (0.06). The native listeners' judgement was found to be highly reliable as tested by Cronbach's Alpha (311 valid items for 10 raters; α = .97).

2.4 Analysis

For analysis, both cross-sectional and longitudinal analyses were used to control the time and motivation variables.[15] Duration and fundamental frequency of each vowel were measured to find the prosodic differences between native and learner speech. By analyzing both of these acoustic properties, we examined the features of the native and target prosody: duration for English prosody and pitch for Korean and possibly English prosody (Ladefoged & Johnson, 2015, p. 119; Jun, 2005b, p. 441). The pre-test recordings of learner speech were used for cross-sectional analysis at a specific point in learners' development (i.e., horizontal variation), while each pair of pre-test and post-test recordings of each person for the same stimuli was used for longitudinal analysis across time (i.e., vertical variation). As for the native speech, we used only post-test recordings, to provide a constant reference for various levels and types of learner speech.[16] The differences in duration and f_0 values between the preceding and the following vowels in each pair were computed in terms of

14) Both native and learner speeches were recorded twice to examine whether there is a familiarity effect. Only learner speech had weekly pronunciation lessons between the two tests.

15) The defining feature of a cross-sectional study is that it compares different population groups at a single point in time, while a longitudinal study compares them at different points in time. The time difference may vary between one variable and another, from a week or less up to a decade or more.

16) Learner language is highly variable, as Tarone (1988) demonstrates.

pairwise variability, or the difference divided by mean value (Grabe & Low, 2002); the resulting value was multiplied by 100 to avoid fractional values.[17] Pairwise variability is commonly used in the literature to account for speech rhythm in different languages (e.g., Kim et al., 2007; Ladefoged & Johnson, 2015, p. 263).

3. Results

The results showed that the learners had far more improvement in pitch variability than in duration variability. This will be shown below by comparing pitch and duration values (i.e., Tables 3 and 4, and Figures 1 and 2, respectively).

3.1 Pitch variability

Results for pitch variability of the 14 vowel pairs from (2) are shown in Table 3. For example, the difference in fundamental frequency values of two [u] vowels was measured in (2c) *n[u]g[u] teogiya*.

Table 3. Pairwise variability of pitch (f_0)
(N = 797, $p < .001$, $F(2, 794)$, by ANOVA for the group differences)

Speakers	Pre-test learner speech	Post-test learner speech	Native speech
Pairwise variability: Mean (*SD*)	6.6*** (19.9)	11.3*** (11.3)	19.3*** (12.2)
Number of vowel pairs measured	318	327	152
Percentage of vowel pairs with significant rising pitch ($p < .05$)[18]	57% (=8/14)	93% (=13/14)	100% (=14/14)

Note. ***$p < .001$ by t-test between the pitch values of vowel pairs.

17) The values obtained are expressed as arbitrary units (Low et al., 2000). In the present paper, the results were then multiplied by 100 in order to obtain manageable values as in other studies (Grabe & Low, 2002; Kim et al., 2007; Ladefoged & Johnson, 2015, p. 263). For example, an f_0 pair of 250 Hz and 200 Hz have the value 22.2 of pairwise variability (=100*2(250-200)/(250+200)).

18) The f_0 values of the following vowels are significantly higher than those of the preceding vowels, as revealed by a t-test.

As seen in Table 3, all vowel pairs in native Korean speech show significant contrast in pitch difference in the predicted direction (rising pitch). In native Korean speech, 100% (14/14) of the vowel pairs[19] show statistically significant pitch difference ($p < .05$), a percentage that gradually decreases along with lower proficiency (Native speech > Post-test production of learner speech > Pre-test production of learner speech).

However, the pitch variability of learner speech increased substantially in the post-test (11.3 on average) in comparison with the pre-test (6.6 on average) and became significant in 93% of the 14 vowel pairs, an increase from the pre-test (57%). To compare the learning effect in pitch, we further analyzed the pitch difference in terms of post-hoc comparisons and a repeated measures t-test (i.e., paired t-test).

Post hoc comparisons using the LSD test[20] showed statistically significant differences between all pairs of the three compared groups: between the pre-test and post-test learner speech groups (mean difference = −4.8, 95% CI = −7.5 to −2.0, $p = .001$), between the post-test learner speech and the native speech groups (mean difference = −8.0, 95% CI = −11.4 to −4.6, $p < .001$), and between the pre-test learner speech and the native speech groups (mean difference = −12.7, 95% CI = −16.2 to −9.3, $p < .001$). To see improvement in pitch variability in each speaker, a repeated-measures t-test was conducted that compared the variability values of the pre-test and post-test. The result showed a significant difference in the pitch variability of each speaker for pre-test ($M = 6.6$, $SD = 20.1$) and post-test ($M = 11.6$, $SD = 17.2$); $t(294) = -3.7$, $p < .001$. This indicates that the learner speech developed a bigger pitch difference between the left and right vowel pairs in the post-test. This development was an improvement because the values became closer to the values of native speech.

Next, the result of pitch variability among the different proficiency levels is shown in Figure 1.

19) The number of each vowel pair equals to the number of subjects, because only one recording per person was taken for data in each of the two time points (pretest and post-test).

20) This procedure was named the Least Significant Difference (LSD) to explore all possible pair-wise comparisons of means comprising a factor using the equivalent of multiple t-tests. Our results also showed clear statistical differences in all other post-hoc comparisons we tested, including Tukey's HSD test.

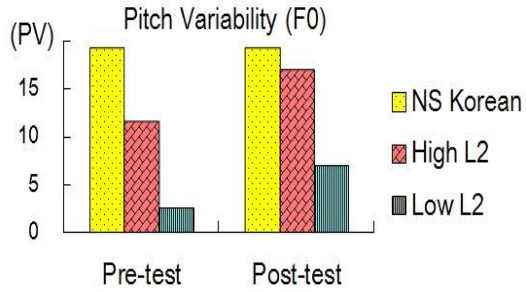

Figure 1. Pitch variability of word pairs within accentual phrases in native and learner speech of Korean. The pairwise variability of the high- and low-level learners increases with time and proficiency level (N=470 in pre-test; 479 in post-test).

Figure 1 shows pitch difference between the adjacent vowels in pairs within accentual phrases in learner speech of different proficiency levels before and after instruction as compared to that in native control speech. The difference in pitch variability increases substantially as learner proficiency does in the pre-test (cross-sectionally) and increases after instruction in the post-test (longitudinally). All learner groups made smaller distinctions in pitch compared with native controls (2.5 and 7.0 for Low L2 group in pre- and post-tests; 11.6 and 16.9 for High L2 group in pre- and post-tests; and 19.3 for native speech controls of Korean). The pitch variability was significantly different among the three proficiency groups, as indicated by an ANOVA of both the pre-test and the post-test recordings (mean variability 19.3, 11.6, 2.5 for pre-test; 19.3, 16.9, 7.0 for post-test; $F(2, 467) = 38.5$, $p < .001$ in pre-test; $F(2, 476) = 29.6$, $p < .001$ in post-test). To examine whether this difference arose from both levels of learner speech, we conducted the following post-hoc comparisons.

Post hoc comparisons using the LSD test showed significant differences between native speech and both level of learner speech in pre-test (mean difference = 7.7, 95% CI = 3.7 to 11.7, $p < .001$ for high-level learners; mean difference = 16.8, 95% CI = 13.0 to 20.6, $p < .001$ for low-level learners) and low-level learner speech in the post-test (mean difference = 12.3, 95% CI = 9.0 to 15.7, $p < .001$), but not between natives and high-level learners in the post-test (mean difference = 2.4, 95% CI = −1.2 to 6.0, $p = .20$). This indicates that learner groups after practice produced a noticeable and improved difference in pitch height between the vowel pairs to the level of native speech—strengthened substantially from the pre-test (2.5 to 7.0 for Low L2; 11.6 to 16.9 for High L2; both of which values became closer toward the native speech value of 19.3).

3.2 Duration variability

The results for duration variability in the 14 vowel pairs from (2) are shown in Table 4.

Table 4. Pairwise variability of duration (*s*)
(N = 855, $p < .001$, $F(2, 852)$, by ANOVA for the group differences)

Speakers	Pre-test learner speech	Post-test learner speech	Native speech
Pairwise variability: Mean (*SD*)	26.9*** (58.7)	24.1*** (53.6)	−.1 (54.1)
Number of vowel pairs measured	348	352	155
Percentage of vowel pairs with significant duration difference ($p < .05$)	64% (=9/14)	57% (=8/14)	43% (=6/14)

Note. ***$p < .001$ by a *t*-test between the duration values of vowel pairs.

As seen in Table 4, the duration variability of vowel pairs in native Korean speech is near zero (−.1 on average), as predicted under syllable-timing duration. On the other hand, vowel pairs in learner speech show significant contrast in duration variability in the predicted direction (the pair of a short vowel followed by a long vowel) in the pre-test (26.9 on average) and slightly less in the post-test (24.1, on average). The duration variability was significantly different in the ANOVA ($p < .001$) of the three comparing groups: learner speech in pre-test, learner speech in post-test, and native speech. This duration difference, which was smaller than the pitch difference in Table 3, may show a non-significant learning effect in terms of post-hoc comparisons and a repeated measures *t*-test in the following.

Post hoc comparisons using the LSD test found clear statistical differences[21] between the native speech and both groups of learner speech (mean difference = −28.1, 95% CI = −38.7 to −17.5, $p < .001$ for pre-test; mean difference = −25.3, 95% CI = −35.9 to −14.8, $p < .001$ for post-test), but not between the pre- and post- learner speech (mean difference = 2.8, 95% CI = −5.5 to 11.1, $p = .51$). To see the improvement in duration variability for each learner, a repeated-measures *t*-test was conducted that compared the variability on the pre-test and post-test. The result showed insignificant difference in the duration variability for pre-test (M = 26.8, SD = 59.1) and post-test (M = 23.7, SD = 53.9) conditions: $t(328)$ =

21) These clear statistical differences were also shown in all the other post-hoc comparisons we tested including Tukey's HSD test.

1.14, p = .26. This indicates that the learner speech did not develop syllable-timed duration by decreasing the durational difference in the post-test production.

Next, the results for duration variability among the different proficiency levels (Native, High, Low) are shown in Figure 2.

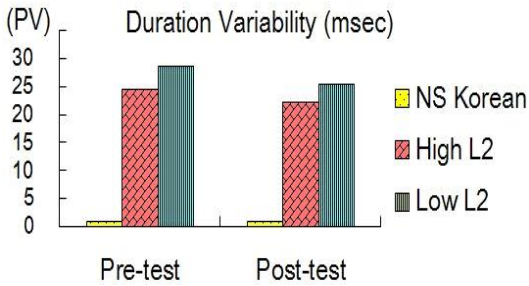

Figure 2. Duration variability of word pairs within accentual phrases in native and learner speech of Korean. The pairwise variability of the high- and low-level learners insignificantly decreases with time and proficiency level (N=503 in pre-test; 507 in post-test).

Figure 2 shows duration variability between adjacent vowels in pairs within accentual phrases in learner speech of different proficiency levels before and after instruction as compared to that in native control speech. Duration variability decreases slightly as learner proficiency decreases in pre-test (cross-sectionally 28.8 for High L2 vs. 24.6 for Low L2) and decreases slightly after instruction for post-test (longitudinally 24.6 to 22.4 for Low L2, 28.8 to 25.5 for High L2). Each proficiency group showed significantly different duration variability in the ANOVA of the pre-test recordings as well as in the post-test recordings (mean variability −.1, 24.6, 28.8 for pre-test; −.1, 22.4, 25.5 for post-test; F(2, 500) = 12.1, p < .001 in pre-test; F(2, 504) = 11.1, p < .001 in post-test). This difference arose solely from the distinct native speech value and despite the similar learner values, as evident in the following post-hoc comparisons.

Post hoc comparisons using the LSD test found statistical differences between the native speech and all levels of learner speech in both the pre-test (mean difference = −24.8, 95% CI = −37.6 to −12.0, p <.001 for high-level learners; mean difference = −28.9, 95% CI = −41.1 to −16.8, p <.001 for low-level learners) and the post-test (mean difference = −22.4, 95% CI = −34.5 to −10.5, p <.001 for high-level learners; mean difference = −25.7, 95% CI = −37.0 to −14.3, p <.001 for low-level learners), but not between the high- and low-level learners in either the pre-test (mean difference = −4.2, 95% CI = −16.3 to 8.0, p = .50), or the post-test (mean difference = −3.2, 95% CI = −14.5 to 8.2, p = .59). This indicates that

learner speech did not achieve significantly better syllable-timed duration, despite the increased proficiency level (to high level) and practice time (for post-test).

4. Discussion

This study provides interesting insight into how one prosodic property (duration) cannot be improved in language learning, while another property (f_0) is well learned in L2 Korean produced by L1 speakers of English. Thus, our results answer in the negative our main question in the very first paragraph of this paper: Is syllable-timedness of speech rhythm learnable? To be specific, the research questions in (1) are answered below in the order of prosodic transfer, prosodic acquisition, and prosodic development.

4.1 Prosodic transfer

The research question (1a) on prosodic transfer is answered affirmatively, because our English-speaking Korean learners used high duration variability (Table 4, Figure 2) to express high pitch variability in Korean (Table 3, Figure 1). This affirmative answer suggests that L1 duration prosody in English is transferred to express L2 pitch prosody in Korean. This answer supports the positive view on the speech rhythm distinction between languages in the literature (Pike, 1945; Nolan & Jeon, 2014), standing against the negative view (Roach, 1982).

In Figure 2, every learner proficiency group demonstrated substantial duration variability on the post-test. This intensive use of duration variability was also found at a bigger scale in the L2 speech of the learners on the pre-test. We therefore conclude that English-speaking Korean learners used a duration feature (length variability) from their L1 English prosody to compensate for imperfect acquisition of another prosodic feature (pitch variability) needed in the L2 Korean rhythm. However, the L1 English prosodic feature (duration) lends little in the target Korean prosody (the pairwise variability is near zero in native Korean speech in Table 4).

This finding of *prosodic transfer* from stress-timing to syllable-timing is new, to the best of my knowledge. Learner speech in English is more stress-timed than native Korean speech. However, it should be noted that a previous research by the author (Kim, 2005, 2017) on the learner speech of reversed direction from syllable-timing to stress-timing, is in line with this new finding, in that L1 Korean learners made greater distinction of pitch (but not duration) in L2 English after practice than native speakers did.

4.2 Prosodic acquisition

The research question in (1b) on prosodic acquisition is also answered affirmatively, because our English-speaking Korean learners used pitch to express pitch variability in Korean, as shown in Figure 1. This affirmative answer indicates that L2 Korean prosody of pitch variability is acquired by L1 English speakers, although this prosodic property (f_0) lends little to prominence in their L1 (Kochanski et al., 2005). This answer does not support either the positive view (speech rhythm, in terms of duration, is easy to learn, as in Dutch learners of Spanish syllable-timed rhythm in van Maastricht et al., 2018) or the negative view (speech rhythm, in terms of duration, is difficult to learn as in Spanish learners of Dutch stress-timed rhythm in van Maastricht et al., 2018), because syllable-timed rhythm of Korean was learned in terms of the pitch property, not duration. Overall, the expectation that the syllable-timed rhythm of Korean is learnable turned out to be correct, although the prosodic feature was different.

This finding of *rhythm learnability by pitch property* is also new, as far as this author knows. The syllable-timed rhythm of Korean in our data was learned in terms of the pitch property by speakers of a stress-timed language. However, it should be noted that previous studies by Nolan and Jeon (2014) and Jun (2005b) are in line with this new finding, as they suggest that pitch is also an important property of speech rhythm.

4.3 Prosodic development

The research question in (1c) on prosodic development is answered partially affirmatively, because our English-speaking learners of Korean produced pitch variability increasingly close to native speech level, as shown in Figure 1, but failed to reduce the duration variability to express syllable-timing in Korean, as shown in Figure 2. This partially affirmative answer indicates that L1 English speakers do develop pitch variability in L2 Korean but do not reduce their duration variability to assimilate with Korean native speech.

This answer supports the more recent view of prosodic features to measure speech rhythm in the literature, in that fundamental frequency (f_0) reflects rhythmicality, as in German and Swiss-French listeners (Barry et al. 2009; Cumming 2011). In fact, a repeated-measures t-test has shown us the significant improvement in pairwise pitch variability up to the native speech level (Table 3) along with insignificant improvement in pairwise duration variability (Table 4). We therefore conclude that prosodic development in the acquisition of L2 Korean is measured in terms of fundamental frequency, but not duration. In other words, fundamental frequency is an additional prosodic cue to prominence.

This result for prosodic development supports previous findings in the literature showing that adult learners can learn a new speech rhythm (Kim et al., 2007; Kim, 2017), with varying degrees of difficulty in acquisition of different prosodic features. Our study adds to these findings in that speech rhythm may develop in terms of the pitch property, with which an L1 prosodic property of duration variability (as in English) may interplay to learn syllable-timed speech rhythm (as in Korean).

5. Conclusion

To address the goal of our investigation, we conclude that syllable-timedness of speech rhythm is learnable in terms of pitch property. Isochronic duration may not be learned by an L1 speaker of a stress-timed language, but pitch variability may be learned to the perfect level of native speech.

Overall, pairwise variability of prosodic properties shows that native speakers of English use both duration and pitch properties to express prominence or speech rhythm in L2 Korean, while L1 Korean speakers use only the pitch property to express speech rhythm. The learner speech shows significant improvement in pitch variability but insignificant improvement in duration variability. In other words, only f_0 variability is used correctly in L2 speech to accommodate syllable-timed speech rhythm.

Our findings imply two theoretical points also raised by previous studies. The first is that there is a speech rhythm contrast between variable stress and variable tone. Our results in Table 4 have shown that Korean has indistinct duration variability among adjacent syllables, supporting the assertion that Korean rhythm is syllable-timed (Nolan & Jeon, 2014). On the other hand, our results in Table 3 have shown that Korean has fixed pitch variability in a predictable direction. This result means that Korean has variable tone. Therefore, disagreement on whether there is stress-timed or syllable-timed rhythm in given languages (Roach, 1982) may be resolved by interpreting the contrast in terms of contrastive prominence (Nolan & Jeon, 2014; Ladefoged & Johnson, 2015, pp. 125-126).

The second point is that prominence in learners' Korean speech is not determined by duration alone (Kochanski et al., 2005), but by pitch as well (Figure 1). Native Korean speech creates prominence using pitch only (Table 3). This supports the traditional view that there is a typological difference between languages based on whether prominence is primarily determined by pitch or by duration.

The current study may have two limitations that should be mentioned. First, the study did not address pitch or duration variation depending on syllable numbers or segmental contexts. Second, the amount of data was limited only to 14 pairs of adjacent vowels. This limitation is evidence of the difficulty of collecting a highly specific rare set of data: learners' pronunciation data for two discrete time points among a homogeneous group of L1 American English and L2 Korean speakers. As the focus of this study was on pairwise variability, there is a possibility that dissimilar evaluations would have arisen if the focus had been on consecutive rhythm in various contexts. Despite these limitations, the findings we gained from this study are well substantiated with statistically significant results.

References

Abercrombie, D. (1967). *Elements of general phonetics*. Edinburgh, UK: Edinburgh University Press.

Barry, W., Andreeva, B., & Koreman, J. (2009). Do rhythm measures reflect perceived rhythm? *Phonetica* 66, 78-94. DOI: 10.1159/000208932

Beckman, M. E. (1986). *Stress and non-stress Accent.* (Netherlands Phonetic Archives No. 7). Dordrecht, Netherlands: Foris. (Additional printings, 1992, 2012, by Walter de Gruyter.)

Behrman, A., Ferguson, S. H., Akhund, A., & Moeyaert, M. (2019). The effect of clear speech on temporal metrics of rhythm in Spanish-accented speakers of English. *Language and Speech*, 62(1), 5-29. DOI: 10.1177/0023830917737109

Choo, M., & O'Grady, W. D. (2003). *The sounds of Korean: A pronunciation guide.* Honolulu: University of Hawaii Press.

Cumming, R. E. (2011). The language-specific interdependence of tonal and durational cues in perceived rhythmicality. *Phonetica, 68*, 1-25. DOI: 10.1159/000327223

Deterding, D. (2001). The measurement of rhythm: a comparison of Singapore and British English. *Journal of Phonetics*, 29, 217-230.

Gilbert, J. B. (2019). An informal account of how I learned about English rhythm. *TESOL Journal* e00441. DOI: 10.1002/tesj.441

Grabe, E., & Low, E. L. (2002). Durational variability in speech and the rhythm class hypothesis. *Papers in Laboratory Phonology,* 7, 515-46.

Jang, T.-Y. (2008). Speech rhythm metrics for automatic scoring of English speech by Korean EFL learners. *Malsori*, *66*, 41-59. The Korean Society of Phonetic Sciences and Speech Technology.

Jun, S.-A. (1998). The accentual phrase in the Korean prosodic hierarchy. *Phonology*, 15(2), 189-226.

Jun, S.-A. (2005a). Korean intonational phonology and prosodic transcription. In S.-A. Jun (Ed.), *Prosodic typology: The phonology of intonation and phrasing* (pp. 201-229). Oxford: Oxford University Press.

Jun, S.-A. (2005b). Prosodic typology. In S.-A. Jun (Ed.), *Prosodic typology: The phonology of intonation and phrasing* (pp. 430-458). Oxford: Oxford University Press.

Jun, S.-A., & Oh, M. (2000). Acquisition of second language intonation. *The Journal of the Acoustical Society of America, 107*(5), 2802-2803. DOI: 10.1121/1.429024

Kim, J.-M. (2005). Stress assignment rules in Korean English. *Studies in Phonetics, Phonology and Morphology, 11*(2), 247-258.

Kim, J.-M. (2012). An accent reduction course for native speakers of American English learning Korean. *Studies in Phonetics, Phonology and Morphology, 18*(1), 27-50. DOI: 10.17959/sppm.2012.18.1.27

Kim, J.-M. (2017). Prosodic transfer of speech rhythm: Substituting fundamental frequency for vowel duration in Korean learners' English speech. *Studies in English Language and Literature, 43*(4), 181-203. DOI: 10.21559/aellk.2017.43.4.10

Kim, J.-M., & Park, K.-M. (2018, July). Speaking a syllable-timed language as a native speaker of a stress-timed language. Paper presented at the 20th International Congress of Linguistics, Cape Town, South Africa.

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.

Kochanski, G., Grabe E., Coleman, J., & Rosner, B. S. (2005). Loudness predicts prominence: Fundamental frequency lends little. *Journal of the Acoustical Society of America, 118*(2), 1038－1054.

Korea Ministry of Culture, Sports and Tourism. (2014). *Romanization of Korean*. (Notice of Regulation, No. 2000-8, No. 2014-42). Retrieved March 28, 2019 from http://www.mcst.go.kr/english/koreaInfo/language/romanization.jsp

Ladefoged, P., & Johnson, K. (2015). *A Course in phonetics* (7th edition). Boston, MA: Cengage Learning.

Low, E. L., Grabe, E., & Nolan, F. (2000). Quantitative characterizations of speech rhythm: Syllable-timing in Singapore English. *Language and Speech, 43*(4), 377-401.

Nolan, F., and Jeon, H.-S. (2014). Speech rhythm: a metaphor? *Philosophical Transactions of the Royal Society B* 369: 2013396. 1-10. DOI: 10.1098/rstb.2013.0396

Ordin, M., & Polyanskaya, L. (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.

Pike, K. L. (1945). The intonation of American English. *University of Michigan Papers in Linguistics I*. University of Michigan, Ann Arbor.

Roach, P. (1982). On the Distinction between 'Stress-timed' and 'Syllable-timed' languages. In D. Crystal (Ed.), *Linguistic controversies* (pp. 73-79). London: Edward Arnold.

Scovel, T. (1988). *A time to speak: A psycholinguistic inquiry into the critical period for human speech*. New York: Newbury House.

Shin, S. (2008). *Read and speak Korean for beginners (1st ed.).* McGraw-Hill.

Tarone, E. (1988). *Variation in interlanguage*. London: Edward Arnold.

Turk, A. E., & Shattuck-Hufnagel, S. (2007). Multiple targets of phrase-final lengthening in American English words. *Journal of Phonetics, 35*(4), 445-472.

van Maastricht, L., Krahmer E., Swerts, M., & Prieto, P. (2018). Learning direction matters: A study on L2 rhythm acquisition by Dutch learners of Spanish and Spanish learners of Dutch. *Studies in Second Language Acquisition,* 1-35. DOI:10.1017/S0272263-118000062

Appendix

The experimental script in (2) was presented in the Korean alphabet with minimal pairs of words as the following. The list is selected from Choo and O'Grady (2003) and mixed in 100 sets of sentences that learners read and 25 sets of sentences that native speakers read. The recordings were made for general pedagogical use and not specifically for this particular research. Translations into English for the used sentences are in (2).

a. (가/갔)다 왔어요. *(ga/gat)da wasseoyo.*
'I (went halfway, went) and have come back.'
b. (먹/목)이 짧아요. *(meok/mok)-i jjalba-yo.*
'(Ink stick, neck) is short.'
c. 누구 (덕/턱)이야? *nugu (deok/teok)-iya.*
'(Thanks, treat) to whom is that?'
d. (금이/금리) 비싸죠? *(geumi/geumni) bissajo?*
'Is (interest, gold) high/expensive?'
e. (기/키)가 모자라요. *(gi/ki)-ga mojarayo.*
'He is short of (vitality, height).'
f. (극/국)이 싱거워요. *(geuk/guk)-i singgeoweo-yo.*
'(Pole, soup) is not salty enough.'
g. 오늘 어디 가요? *oneul eodi gayo?*
'Are we going somewhere today?/Where are we going today / Let's go somewhere today. (Yes-no question/Wh-question/Request)'

h. 밤에는 (자/차)요. *bameneun (ja/cha)-yo.*
'At night, it is (sleeping, cold).'
i. (자비/잡비)가 들어요. *(jabi/japbi)-ga deureoyo.*
'It requires (personal, miscellaneous) expenses.'
j. 여기서 (사/서)요. *yeogiseo (sa/seo)-yo.*
'Please (buy, stop) here.'
k. (후추/우추) 좀 주세요. *(huchu/uchu) jom juseyo.*
'Please pass me the pepper. (Correct/Incorrect h-deletion)'
l.* 자(고/구) 나서 할게. *ja-(go/gu) naseo halge.*
'I will do it after sleeping (Normal/Friendly) speech.'
m.*(실리/심리)학을 공부합니다. *(silli/simni)hag-eul gongbuhamnida.*
'I study (utilitarianism/psychology).'
n. (바다/바닷)물은 짜죠. *(bada/badat)mul-eun jjajo.*
'Sea water is salty, of course. (Incorrect/Correct Compounding)'
o. 도서관에 책 빌리러 가요? *doseogwone chaek (billi/biri)-reo gayo?*
'Do you go to the library to borrow books? (Long/Short *l*-sound)'
p. 아뇨, 도서관에 일하러 가요. *anyo, doseogwane (ilha/ira)-reo gayo.*
'No, I go to the library to work. (Clear/Casual *l*-sound)'

Jong-mi Kim
Professor
Department of English Language and Literature
College of Humanities, Kangwon National University
1 Kangwondaehak-gil, Chuncheon-si
Gangwon-do 24341, Korea
Phone: +82-33-250-8150
E-mail: kimjm@kangwon.ac.kr

Received on May 1, 2019
Revised version received on June 22, 2019
Accepted on June 30, 2019