Linguistic Research 39(1): 213-233
DOI: 10.17250/khisli.39.1.202203.008

The production of tone 3 by advanced Korean learners of Mandarin*

Yuping Fu · Yong-cheol Lee*****
(Hainan Tropical Ocean University · Cheongju University)

Fu, Yuping and Yong-cheol Lee 2022. The production of tone 3 by advanced Korean learners of Mandarin. *Linguistic Research* 39(1): 213-233. This study examined the differences in tone 3 productions between advanced Korean learners and native speakers of Mandarin by implementing three analyses: acoustic, frequency, and distance metric analyses. Based on the acoustic analyses, the advanced learners exerted more articulatory effort in producing tone 3, regardless of the syllable positions within disyllabic words. Compared to the native speakers, the production of tone 3 by the learners was longer in duration, more intense, and displayed steeper falling and rising pitch movements—these are all salient cues for tone 3 as a dipping pitch contour. The native speakers, however, economized articulatory effort in their tone 3 productions. They did not lower the pitch target of tone 3, nor did they employ the steep falling and rising pitch movements to the same extent as the advanced learners. According to the frequency analyses, the native speakers yielded more variants of half-T3 than the advanced learners in both syllables of the disyllabic words. Finally, the distance metric analyses revealed that the advanced learners' pitch contours were more different from those of the native speakers in syllable 2 than in syllable 1. Consistent with previous work, our results emphasize that the learners' dipping pitch contour for tone 3 stems from the widespread second language pedagogy of Mandarin tones, where tone 3 is predominately described as a dipping tone. Learners, therefore, must understand and become competent using varied tone 3 patterns before attaining native-like or near-native competence. **(Hainan Tropical Ocean University · Cheongju University)**

Keywords Mandarin, tone 3, Korean learners, disyllable words, production

1. Introduction

Mandarin has four lexical tones: high-level (tone 1), rising (tone 2), low/dipping (tone

---

* We thank three anonymous reviewers from *Linguistic Research* for their constructive comments and suggestions regarding this research. This work is a developed and extended version of Chapter 3 from the first author's dissertation.

** First author
*** Corresponding author

3), and falling (tone 4). These four tones are conventionally described as [55], [35], [214], and [51], where the numbers refer to the relative pitch heights from low (1) to high (5) (Chao 1968). Unlike the other tones, tone 3 has three allophonic variants (Lee and Xiong 2021). It surfaces as a low-dipping tone (also known as a full tone 3) in isolation or sentence-final positions and can also be a rising tone when followed by another tone 3 (i.e. T3-T3 → T2-T3). Finally, tone 3 often becomes a low-falling tone [21] (known as a half tone) in connected speech or sentence-medial positions.

Traditionally, a full tone 3 (full-T3) is the base or underlying form of tone 3, and a half tone 3 (half-T3) is a derived form (Chao 1968; Cheng 1973). Despite being a derived form, half-T3 has a wider distribution and occurs more frequently than full-T3. Around 82% of tone 3 occurrences are realized as a half-T3 in connected speech (Chen et al. 1983). Shi and Li (1987) found that only 15% of tone 3 occurrences surface as a full-T3, even in sentence-final positions. This supports Duanmu (2007), who reported that a half-T3 is most frequent in sentence-final positions, and a full-T3 is usually used for emphasis.

It has been documented that acquiring Mandarin tones is challenging for second language (L2) learners, especially those whose native language is non-tonal (Chen 1997; Wiener et al. 2020; Zhang 2014, 2017) because the four lexical tones with different pitch movements co-occurring within a sentence (Wang et al. 2020) must be produced precisely to convey the lexical meaning of each tone, and they must be aligned correctly according to intonation contour types. Of the four tones, tone 3 is the most challenging and is acquired late in both first language (L1) and L2 acquisition (Chen 1997; Hao 2012; Li and Thompson 1977; Lin 1985; Wong et al. 2005). This reflects the complex pitch variations and the three different phonological processes (i.e. full, rising, and low-falling variants) of tone 3. Due to its complexity, tone 3 has often been excluded in previous L2 acquisition studies. Thus, little is known about tone 3 acquisition by learners whose L1 is non-tonal.

To address this gap, this study analyzes tone 3 acquisition among Korean learners of Mandarin by conducting a production experiment with disyllabic tone sequences. To our knowledge, only one study has addressed tone 3 productions on Korean learners of Mandarin whose language proficiency was intermediate-level (Zhang 2014). Thus, it remains unclear how the variants of tone 3 are realized acoustically by advanced Korean learners. Before setting up the research questions of this study, we first provide background information necessary to understand the motivation and aim of this study.

Therefore, this study first briefly describes Korean prosody, then introduces some previous studies of tone 3 in L2 settings, and finally discusses the pedagogical issues of tone 3.

1.1 A brief introduction to Korean prosody

Korean has neither lexical stress nor lexical tones, so Korean pitch variations do not cause a difference in meaning (Jun 1993). To illustrate, consider the word *gong.gan* 'space.' The word's meaning is the same regardless of the different placement of prominence (e.g. *GONG.gan* vs. *gong.GAN*), where capitalization refers to prominence, and a dot syllabifies the word.[1] According to the autosegmental-metrical framework of intonational phonology on Korean (Jun 1998), there are two prosodic units above the word: intonational phrase (IP) and accentual phrase (AP). Within an AP, there are two prosodic templates: HH-LH and LH-LH, in which H is for high and L for low. Only the initial tone differs between the prosodic templates and it differs by the laryngeal feature of its AP-initial segment. When the AP-initial segment is tense or aspirate, the AP begins with H, but otherwise it begins with L. HH-LH or LH-LH is fully realized when there are four or more syllables within an AP, and the AP is realized as LH-LL or HH-LL at the end of a sentence. When there are fewer than four syllables, the second H or the third L (or both) are not realized, thus leading to the following AP patterns: LLH, LL, LHL, LHH, LH, HLL, HLH, HL, HHL, and HH (Jun 2000). These patterns can be roughly expressed by the relative pitch heights from 1 (Low) to 5 (High): LLH (115), LL (11), LHL (151), LHH (155), LH (15), HLL (511), HLH (515), HL (51), HHL (551), and HH (55).

1.2 Previous studies on tone 3 in L2 learning

There has been abundant research on the acquisition of tone 3. Here, we describe some of the previous studies to show that acquiring tone 3 is indeed difficult. Chen et al. (2019) used a wug test to evaluate L2 learners' tone 3 productions by recruiting 23 intermediate Mandarin learners from Cantonese and English backgrounds. The results

1 In this study, Korean refers specifically to Seoul Korean, a variety spoken in Seoul and its metropolitan area. Contrary to Seoul Korean, both North and South Kyungsang Korean are pitch-accent dialects, and pitch variations cause differences in meaning in these varieties. For example, *KA.ci* means 'kind,' but *ka.CI* means 'branch.'

revealed that L2 learners' productions of tone 3 were neither accurate nor fluent compared to the native speakers. Specifically, Cantonese learners of Mandarin produced the first half of tone 3 (i.e. 21) with a terminal rise and the second half of tone 3 (i.e. 14) with a low-pitch height, while English learners of Mandarin produced the two contours as relatively flat. This means that regardless of their L1 background, the two groups struggled to produce native-like pitch contours of tone 3.

In large-scale corpus research on tone productions by 305 Mandarin L2 beginners from diverse European backgrounds (Germanic, Romance, and the Slavic languages) (Chen et al. 2016), 300 utterances including 2-5 syllabic words were perceptually transcribed into the corresponding tone categories by 64 native Mandarin speakers. The results showed that tone 3 was the most challenging, with the lowest accuracy (58.8%). Using 80 Mandarin monosyllables, Wang et al. (2003) conducted a two-week tone perceptual training by recruiting American learners with one to two semesters of Mandarin experience. This training showed that tone 3 was comparatively easy to perceive but difficult to produce and improve, which might stem from the peculiarity of the complex pitch contour of tone 3.

There are not many studies on Korean learners of Mandarin in the production of tone 3. In one of the few studies, Zhang (2014) recruited 60 intermediate-level English, Japanese, and Korean learners of Mandarin and evaluated their production of disyllabic tone sequences embedded in carrier sentences. The results revealed that regardless of their L1, the learners had greater difficulty producing half-T3 than full-T3. In subsequent work by the same author (Zhang 2017), similar error patterns were observed even from advanced American learners of Mandarin. In this study, producing half-T3 remained the most challenging, and a large ratio of full-T3 was overused for half-T3. Zhang (2017) attributed these results to a prevalent teaching method that describes tone 3 as a dipping tone by showing that L2 learners prefer using full-T3.

1.3 Pedagogical biases of tone 3

The intrinsic difficulty of tone 3 is its extraordinarily low pitch near the lower limit of a human voice (Chao 1948; Lee et al. 2016). To produce such a low voice, one must adjust their larynx to make a strong and rumbling sound. In classroom settings, therefore, instructors often exaggerate their pronunciation as if tone 3 syllables are stressed (Chin

1987). Besides, most textbooks for L2 beginners simply describe tone 3 as a low-dipping tone [214] (Linge 2011; Zhang 2018). This description of full-T3 gives L2 learners the impression that the rising tail [14] is the most prominent characteristic of tone 3, and thus tone 3 needs to be fully pronounced as a dipping tone.

In L2 learning for Mandarin tones, tone 3—as a dipping tone—has become accepted as 'the norm.' As a result, L2 learners are influenced, often unknowingly, by such pronunciation. Although there is a brief description that tone 3 turns into half-T3 in verbal communication, textbooks for beginners, for example, rarely cover the three allophonic variants of tone 3 (Linge 2011). L2 learners tend to follow the traditional teaching method of tone 3 as full-T3 (Lin 1985; Linge 2011) and are actively encouraged to use the dipping pitch contours of tone 3. Visual gestures are also suggested to help learners produce such a dipping pitch contour. For example, "Drop your chin on your neck and raise your chin when you say the 3rd tone (Tsai 2011: 46)." Although these gestures are advantageous in helping learners produce two-pitch movements—falling and rising—for tone 3 as a dipping tone, this method has resulted in the overproduction of a full-T3 among learners, as found in Zhang (2017).

1.4 The current study

As stated above, this study aims to examine tone 3 productions by advanced Korean learners of Mandarin, with particular attention to whether their tone 3 is half-T3 (21) or full-T3 (214). Before moving on, we should specify the primary difference between Zhang's study (2014) and the current study. Advanced Korean learners of Mandarin were not recruited in Zhang's study, and only intermediate learners with 0.5-1.5 years of experience learning Mandarin participated. To accommodate for this oversight, we recruited advanced learners who had achieved HSK Level 6, and then a native Mandarin speaker interviewed them individually to evaluate whether they could be considered "advanced" learners.

According to the Contrastive Analysis Hypothesis (Lado 1957), prosodic similarities between L1 and L2 help expedite the acquisition of L2 prosody, while the differences between them interfere with the L2 leaner's acquisition of prosody. Based on Lado's hypothesis, Korean learners receive no advantage from producing the two tonal patterns because half-T3 and full-T3 are not corresponding to any possible tone patterns within an AP. With respect to the tone 3 productions by advanced Korean learners of Mandarin, we hypothesize two competing scenarios in their performance of tone 3. On one hand, because L1 interference tends to decrease as L2 proficiency improves (Chen et al. 2012; Liu et al. 2019; Liu and Lee 2021), advanced Korean learners begin to show native-like or near-native performance in producing tone 3. On the other hand, regardless of their improved proficiency, and given the widespread teaching approach of tone 3 as full-T3 in L2 settings, advanced Korean learners continue to overproduce full-T3 relative to half-T3, similar to the advanced American learners of Mandarin in Zhang (2017).

To test the two competing scenarios, we recruited advanced Korean learners of Mandarin and native speakers of Mandarin to directly compare the two language groups, using a production experiment with the stimuli of disyllabic words. Details of the production experiment are described in the next section.

2. Production

2.1 Stimuli

We created 16 tone sequences using disyllabic words (4 tones in syllable 1 × 4 tones in syllable 2). Each tone sequence was combined with seven basic vowels (/i, u, y, ɤ, ɹ̩, ɻ̩, ɑ/, Howie 1976) in each target syllable position. We employed two words for each of the 16 tone sequences, thus leading to 224 target words (16 tone sequences × 7 basic vowels x 2 words). Most target words were chosen directly from the following high-frequency wordlists: *5000 Graded Vocabulary for HSK Outline* (2015) and *General Outline of the Chinese Vocabulary Levels and Graded Chinese Characters* (1992). Four low-frequency words (*tǐ.cí* 'nominal,' *cí.gēn* 'word root,' *zǐ.nǚ* 'children,' and *yǘ.gē* 'fishing song') were included as stimuli due to the difficulty of finding all high-frequency words from the above two lists.

In selecting target words, we followed Zhang's (2018) criteria, with some minor modifications. The stimuli were limited to content words, except for one case (*kě.yǐ* 'can'). Sonorants were maximally used to guarantee smooth pitch tracking. Each target vowel appeared only once in each tone sequence. Table 1 exhibits 16 sample tone sequences with pinyin, diacritics, and English gloss.

Table 1. Sample tone combinations using disyllabic words (Syllable boundaries are marked by a dot)

	__T1	__T2	__T3	__T4
T1__	sī.jī 'driver'	kē.xué 'science'	biān.zhě 'editor'	jī.dàn 'egg'
T2__	dú.shū 'reading'	pí.xié 'leather shoes'	bó.mǔ 'aunt'	fú.wù 'service'
T3__	mǔ.qīn 'mother'	dǎ.zhé 'discount'	yǔ.fǎ 'grammar'	zǐ.xì 'careful'
T4__	qì.chē 'car'	dì.tú 'map'	hàn.yǔ 'Chinese'	kè.hù 'customer'

To reduce potential coarticulatory effects from adjacent tones (Xu 1997), the stimuli were embedded in a carrier phrase, as shown in (1), and two particles—*ge* and *de*—bearing neutral tones occurred before and after the target word. This structure allowed the same contextual effect throughout the stimuli. And inserting the target word in a sentence-medial position prevented the possible disturbance of sentence-initial and final intonation. Chinese characters and their pinyin with tonal diacritics were in the reading lists to avoid pronunciation ambiguity.

2.2 Participants

We recruited two groups of speakers in Cheongju, South Korea. The first group included seven advanced Korean learners of Mandarin (2 males and 5 females; mean age = 26.7, SD = 6.5) who had learned no other tonal languages besides Mandarin, passed HSK Level 6, and resided in China for about 1.5 years (from 2 months to 5 years). The second group consisted of eight native Mandarin speakers (4 males and 4 females, mean age = 24.9, SD = 1.9) who were graduate or undergraduate students at Cheongju University. Group 2 participants came from North China and could speak standard Mandarin fluently. All participants received 10,000 won (about $10 USD) for their participation.

2.3 Recording procedure

Recordings were conducted in a sound-attenuated booth at Cheongju University using a Sennheiser headset microphone and were directly saved onto a laptop computer as

16-bit wave files at 44.1 kHz with Praat (Boersma and Weenink 2020). Participants were seated before a computer monitor while wearing a microphone headset. Before recordings, participants were instructed to carefully read, at a normal speed, with neither emphasis nor emotion. They were also told to repeat themselves when mistakes or unnecessary pauses were detected. PowerPoint slides were used to present the 224 carrier phrases, which were further divided into three quasi-equal blocks (i.e. the first two blocks had 75 phrases, and the last block included 74 phrases). Participants had a practice session with three sample phrases that were not part of the stimuli and then produced the actual stimuli twice. The duration of the recording was about 50 minutes.

We collected 6,720 tone sequences (224 tone sequences × 15 participants × 2 rounds). The following tone sequences combined with tone 3 were extracted for further analysis: T3-T1, T3-T2, T3-T4, T1-T3, T2-T3, and T4-T3 (see Appendix for the details). It should be noted that the T3-T3 sequence was purposely excluded because it undergoes tone sandhi, a phonological change where T3 becomes T2 before another T3. In this study, we excluded the tone sequences for further analysis to achieve consistency because language learners, unlike Mandarin native speakers, may not always realize T3-T3 as T2-T3. Therefore, our final extraction yielded 84 tone sequences (6 tone sequences × 7 basic vowels × 2 disyllabic words). Overall, we obtained 1,176 tone sequences for the advanced learners (84 tone sequences × 7 learners × 2 rounds) and 1,344 sequences for the native speakers (84 tone sequences × 8 speakers × 2 rounds).

2.4 A sketch of pitch contours

To help capture different prosodic realizations of tone 3 between the two language groups, we first sketched sample pitch contours produced by each language group. Figure 1 displays the pitch contours of tone 3 in two positions (syllable 1 and syllable 2). Every word was labeled by hand in every carrier phrase, and the pitch was extracted at ten equidistant points from each labeled word using ProsodyPro (Xu 2013). In Figure 1, the x-axis represents the time-normalized duration, and the y-axis corresponds to the pitch in semitones (st). In this study, the semitone scale was converted from hertz by applying the formula ($st = 12log2x$), where x is the raw value in hertz and the reference value is 1 hertz.

Figure 1 features different prosodic patterns between native speakers and advanced learners. When tone 3 occurred in the first syllable of the target word, the advanced speakers produced a falling pitch contour toward the low-pitch target of tone 3, and the pitch subsequently bounced upwards. Yet the native speakers did not display a clear falling pitch contour. Instead, their pitch was almost flat in the first half of syllable 1. A gradual falling pitch appeared, passing by the midpoint, followed by a slight pitch increase toward the end of the syllable. A considerable pitch difference was also observed between the native speakers and learners when tone 3 occurred in the second syllable. The learners produced a very steep falling pitch contour toward the low-pitch target of tone 3, which then rebounded a bit afterward. But the native speakers did not lower the pitch of tone 3 as much as the learners did, nor did their pitch rebound.

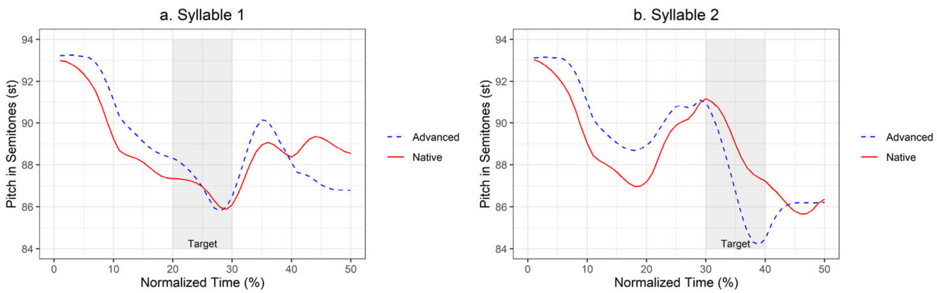

Figure 1. Sample pitch contours of tone 3 in syllable 1 and syllable 2

2.5 Analyses

To identify the different pitch contours of tone 3 between the two language groups, we applied three analyses: acoustic analyses, frequency analyses of half-T3 and full-T3, and a series of distance measurements. In examining the accuracy of L2 tone production, one widely adopted method in the literature was calculating the accuracy rate as perceived by native speakers (e.g. Flege et al. 1995; Hao 2012; Wang et al. 2003; Zhang 2014). For example, Tu et al. (2016) recruited three native Mandarin speakers with a background in phonetics to judge tone recordings produced by Korean learners. Two native speakers evaluated the tone production data in Hao (2012), whereas Zhang (2014) employed one native speaker to judge the data. Instead of drawing a conclusion based solely on the native speakers' perceptual judgment of tone accuracy, we performed three different analyses on the tone 3 productions of L2 learners and native speakers to provide more in-depth information for L2 learners' tone 3 acquisition. What follows is an outline of each of the three analyses in greater detail.

First, with the acoustic analyses, we extracted the following acoustic cues from each of the labeled syllables for tone 3: duration in millisecond (ms), mean intensity in decibels (dB), and four pitch-related parameters in semitones (st), including maximum pitch (MaxP), minimum pitch (MinP), and pitch slopes. Based on the observations in Figure 1, two slopes (slope 1 and slope 2) were calculated. When calculating the slopes, we first specified the turning point of tone 3 in each stimulus by using the MIN function in Excel to determine the lowest pitch point. For example, when MinP occurred at the n^{th} time point, slope 1 was calculated based on the pitch span from the first time point to the n^{th} time point to determine the falling pitch movement of tone 3. Slope 2 was then computed based on the n^{th} time point to the 10^{th} time point within the ten equidistant time points in order to identify the rising pitch movement of tone 3.

Second, for the frequency analyses, we calculated the frequencies of half-T3 and full-T3 as performed by each language group, separately for syllable positions. In this study, we defined the half-T3 as the one whose MinP appeared at the 10th time point within the syllable. When the MinP occurred somewhere in the middle before the 10^{th} time point was reached, we viewed this tone contour as full-T3. Third, we examined three measurements of distance (i.e. Euclidean, Manhattan, Minkowski) that crossed the aggregated pitch contours between the native speakers and the advanced learners in two specific syllable positions (syllable 1 vs. syllable 2). Please refer to Singh et al. (2013) for the formulas of these metrics. These distance metrics enabled us to characterize the differences in pitch contours between the native speakers and the learners.

Using the *lmerTest* package (Kuznetsova et al. 2017) in R (R Core Team 2020), we conducted a series of linear mixed-effects model analysis for the acoustic analyses and a binary logistic regression model for the frequency analyses. The fixed effects, random effects, and dependent variables are tabulated below in Table 2 for the acoustic and frequency analyses. It should be noted that we obtained the *p*-values of each fixed effect and the interaction of the fixed effects by conducting likelihood ratio tests through the ANOVA function of the *lmerTest* package.

Table 2. The fixed effects, random effects, dependent variables of the acoustic, and frequency analyses

Type of analyses	Fixed effects	Random effects	Dependent variables
Acoustic analyses	language group (native, advanced) syllable position (1, 2)	speaker (speakers in each group) round (1, 2) word (84 disyllabic words)	duration mean intensity MaxP MinP Slope 1 Slope 2
Frequency analyses	language group (native, advanced) syllable position (1, 2)	speaker (speakers in each group) round (1, 2) word (84 disyllabic words)	half-T3 (coded as 1) full-T3 (coded as 0)

Finally, for the distance measurements, we directly compared the pitch contours between the two language groups by quantitatively describing the main differences of the observed distance values as performed by the two language groups.

3. Results

3.1 Acoustic analyses

Figure 2 exhibits the aggregated mean of the six acoustic parameters, separately for language groups and syllable positions. The results show that the native speakers and advanced learners clearly differentiated all the acoustic cues in the productions of tone 3, except for MaxP and MinP in syllable 1. The productions of tone 3 were, however, affected by syllable positions. When tone 3 occurred on the first target syllable, the advanced learners increased both the duration and intensity of tone 3 by exerting greater effort in producing tone 3. In comparison, the native speakers produced cues with reduced duration and intensity for both syllable positions. When tone 3 appeared on the second syllable, the native speakers did not lower the pitch of tone 3 as much as the advanced learners did—this is implied from the higher MaxP and MinP values of this group. As for the slopes, the advanced learners produced a lower value for slope 1 and a higher value for slope 2, revealing that the advanced learners seemed to include more tokens of full-T3 in the production of tone 3. In contrast, the native speakers showed the opposite patterns for slopes, producing a higher value for slope 1 and a lower value for slope 2. What follows is a statistical approach for determining the significance of the interaction effects between the language group and syllable position for each fixed effect.

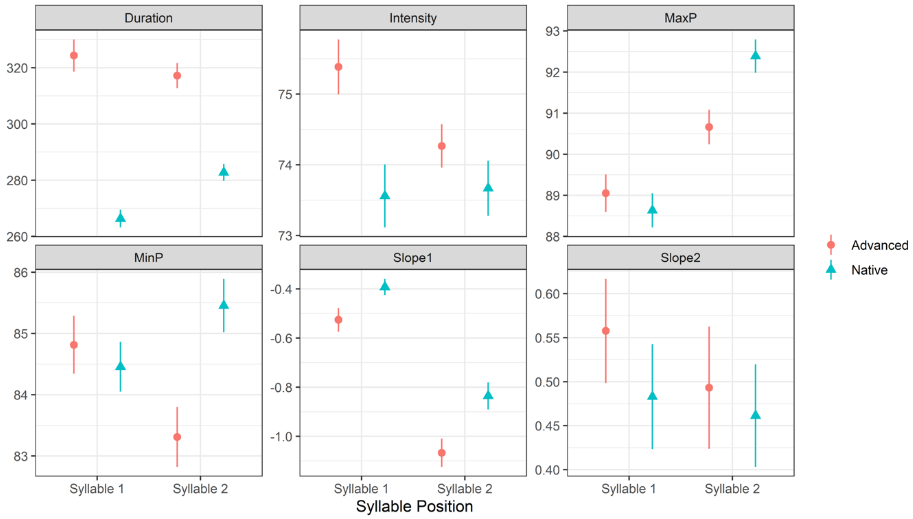

Figure 2. The aggregated mean of duration, intensity, MaxP, MinP, slope 1, and slope 2 of each language group in two syllables. Points refer to mean values, and error bars signify 95% confidence intervals

Statistical analyses confirmed our visual observations because the interaction effects between language group and syllable position are highly significant for all the acoustic cues, except for slope 2 (duration: $X^2 = 659.24$, $df = 3$, $p < 0.001$; intensity: $X^2 = 118.77$, $df = 3$, $p < 0.001$; MaxP: $X^2 = 77.66$, $df = 3$, $p < 0.001$; MinP: $X^2 = 77.51$, $df =3$, $p < 0.001$; slope 1: $X^2 = 56.41$, $df = 3$, $p < 0.001$; slope 2: $X^2 = 2.13$, $df = 3$, $p = 0.545$). These results suggest that the two language groups produced tone 3 differently depending on the position of this tone within disyllabic words. This motivated us to assess the effect of the language group for each acoustic cue, separated by syllable positions, except for slope 2 yielding no significant interaction effect between language group and syllable position.

Table 3 shows that, with two exceptions (MaxP and MinP in syllable 1), the effect of language group is significant for all parameters in both syllables. The effect of language group on both duration and intensity is also significant for all syllable positions, showing that the advanced learners produced significantly increased duration and intensity of tone 3 than the native speakers in each syllable position. For MaxP and MinP, there was no significant difference between the two groups in the first syllable. The native speakers, however, produced a significantly higher value of MaxP and MinP in the second syllable. This suggests that, unlike the advanced speakers, the native speakers did not lower the pitch target of tone 3. The effect of language group on slope 1 was significant for all syllables, showing that the advanced learners had steeper falling pitch movements for both syllables. The effect of language group on slope 2 was also significant for both syllables, but the direction of the pitch movement was different. The advanced learners produced steeper rising pitch movements in the second syllable of the disyllabic words. Combining the statistical results of both slope 1 and slope 2 implies that, compared to the native speakers, the advanced learners tended to produce tone 3 as full-T3.

Table 3. The effect of language level for the five acoustic cues in each syllable position (df = 1)

	Syllable 1		Syllable 2	
	X^2	p-value	X^2	p-value
Duration	420.43	<0.001	236.83	<0.001
Intensity	84.66	<0.001	28.50	<0.001
MaxP	0.088	0.767	64.08	<0.001
MinP	0.036	0.850	62.18	<0.001
Slope 1	18.70	<0.001	25.44	<0.001

3.2 Frequency analyses

Let's now look at whether the two language groups realized tone 3 as either half-T3 or full-T3. Table 4 gives the frequencies of half-T3 and full-T3 performed by each language group, separately for syllable positions. When tone 3 occurred on the first target syllable, the native speakers produced half-T3 (32.1%) more than the advanced learners (26.7%). The same is true when the target tone 3 was on the second syllable. But we observed more tokens for half-T3 in the second syllable, regardless of the language groups. The native speakers revealed half-T3 about 46% of the time, while the rate of the half-T3 dropped to about 35% in the advanced learner group. The statistical results confirmed our observation because there is an interaction effect between language group and syllable position (X^2 = 20.94, df = 3, p < 0.001), meaning that the second target syllable has significantly more variants of half-T3. With the main effect of language group, the frequencies of half-T3 are significantly higher in the native group than in the learner group in both syllables (1st syllable: X^2 = 5.38, df = 1, p = 0.02; 2nd syllable: X^2 = 13.87, df = 1, p < 0.001).

Table 4. Frequencies of half-T3 and full-T3 classified by syllable positions and language groups

Syllable Position	Language Group	Half-T3	Full-T3	Total
Syllable 1	Advanced	157 (26.7%)	431 (73.3%)	588 (100%)
	Native	216 (32.1%)	456 (67.9%)	672 (100%)
Syllable 2	Advanced	277 (35.3%)	507 (64.7%)	784 (100%)
	Native	408 (45.5%)	488 (54.5%)	896 (100%)

3.3 Distance measurements

Table 5 establishes the distance measurements of the target pitch contours between the native speakers and the advanced learners in each syllable position (syllable 1 vs. syllable 2), using three distance metrics: Euclidean, Manhattan, and Minkowski. The results of the three distance metrics reveal that the values of the Euclidean, Manhattan, and Minkowski distances were 1.31, 3.37, and 1.03 in syllable 1. The values for these distance metrics increased considerably in syllable 2: 7.58, 22.30, and 5.41 for the Euclidean, Manhattan, and Minkowski distances. Although the three distance metrics showed some variance in each syllable position, the results suggest that the pitch contours between the two language groups were more similar in syllable 1 than for those produced in syllable 2. The values shown in Table 5 do not tell the whole story about the two language groups' different pitch contours. But given the comparison of the acoustic and frequency analyses shown above, the advanced speakers produced more markedly different pitch contours in syllable 2 than in syllable 1, compared to the native speakers.

Table 5. The values of the Euclidean, Manhattan, and Minkowski distances for each syllable position

Syllable Position	Distance Metrics		
	Euclidean	Manhattan	Minkowski
Syllable 1	1.31	3.37	1.03
Syllable 2	7.58	22.30	5.41

4. Discussion and conclusion

This study used disyllabic tone sequences embedded in carrier phrases to compare the different tone 3 patterns between advanced Korean learners and Mandarin native speakers. The acoustic analyses revealed that the advanced learners produced steeper pitch slopes for tone 3 and increased duration and intensity compared to the native speakers. Still, the native speakers did not lower the pitch of tone 3 to the extent displayed by the advanced learners. Tone 3 productions by native speakers were also shorter, less intense, and featured more gradual slope changes than those of the advanced learners. According to the frequency analyses, the native speakers produced more tokens of half-T3 in both target syllables. And finally, the three distance measurements showed that the advanced learners' pitch contours were more markedly different from those of the native speakers in syllable 2 than in syllable 1.

One of the remarkable differences between the native speakers and advanced learners in their tone 3 productions was their different slope patterns. Unlike the advanced learners who produced more dipping pitch contours regardless of syllable positions, the native speakers performed differently depending on where tone 3 appeared within disyllabic words. In syllable 1, the native speakers produced a virtually flat tone up to the midway time points and then lowered the pitch just before the end of syllable 1. In syllable 2, they did not clearly lower the pitch like the advanced learners did. Instead, their pitch contour of tone 3 merged with the following neutral tone. These results suggest that the native speakers economized their efforts when tone 3 was successfully deciphered. In contrast, the advanced learners seemed to take more articulatory efforts to achieve clear articulation. Furthermore, as for the different slope patterns between the native speakers and the learners, we hope that both language teachers and learners will realize several allotones (i.e. variants) of tone 3 at the surface realizations and that language learners will particularly practice producing tone 3 in various phonetic environments within a sentence for them to attain fluency.

This study shows that the duration and intensity of tone 3 are considerably greater for the advanced learners than native speakers, despite the syllable positions within disyllabic words. One question that arises from this study is why the advanced learners produced increased duration and intensity of tone 3. This result may be attributable to the byproducts of producing more tokens of full-T3. The advanced learners consumed extra time and energy generating dipping pitch patterns, first with their pitch lowering and then raising afterward. That is, the two different pitch movements of full-T3 required extra vocal effort, resulting in concomitantly increased duration and intensity.

It is also worth questioning why advanced Korean learners produced tone 3 as a dipping tone instead of half-T3. As outlined in Section 1, this phenomenon likely stems from the widespread pedagogy of Mandarin tones as L2. This aligns with the textbooks commonly used by beginners that generally describe tone 3 as a dipping tone. While L2 learners are explicitly taught to produce full-T3, other configurations of tone 3, like half-T3, are covered briefly or not at all (Chen 1973; Chin 1987). Furthermore, when instructors spoke with learners or corrected their errors, they tended to slow down and produce tone 3 with an emphatic and lengthened contour (Lin 1985). Consequently, L2 learners imitate and memorize the way instructors produced tone 3 with increased duration and intensity. Specifically, one participant in this study mentioned her teacher's constant correction of her tonal errors, particularly tone 3 (in personal communication).

One may wonder why the native speakers produced more tokens of full-T3 than expected. As mentioned in the Introduction, about 15% of tone 3 productions are realized as full-T3 in sentence-final positions (Shi and Li 1987). In this study, however, the native speakers often produced full-T3 in both syllables (1st syllable: 67.9%, 2nd syllable: 54.5%). This high ratio of full-T3 is likely attributable to our experimental design because we embedded our disyllabic target words into carrier phrases and each carrier phrase, as a type of lab-read speech, appeared in the middle of a laptop screen. Therefore, this layout likely encouraged the speakers to articulate tone 3 clearly, thus leading to many occurrences of full-T3.

Although this study reveals several meaningful results on the different prosodic patterns of tone 3 between the native speakers and the learners, there is certainly room for improvement. In particular, there are three considerations worth examining. First, it is worth examining how language learners produce tone 3 that interacts with tone coarticulation. For example, a tone sequence with tone 3 + tone 2 (da3 zhe2) differs from tone 3 + tone 4 (zi3 xi4) because tone 2 starts with relatively low pitch, while tone 4 triggers high pitch. That said, tone 4, compared to tone 2, is more likely to attract a rising pitch toward the end of tone 3. The detailed findings on the coarticulation effects of tone 3 serve as useful sources for language teachers and learners to enhance their understanding of tone 3. Second, future research should investigate L2 tone productions for more natural speech materials, such as one produced with spontaneous, casual speech, or another speech with interactive turn-taking. Finally, different L2 advanced learners in this study showed some variance in the production of disyllabic tone sequences. This needs to be examined in future research by recruiting more speakers, thereby leaving it a limitation of this study.

In summary, this study has addressed the crucial differences in tone 3 patterns between advanced Korean learners and native Mandarin speakers. The native speakers economized their articulatory effort in producing tone 3, thus neither lowering the pitch of this tone nor displaying steep falling and rising pitch movements just like the advanced learners. Language teachers should first identify these core differences between the learners and native speakers and then design pedagogical strategies to enable L2 learners to understand and imitate native speakers' tone 3 patterns in different syllable positions. This will enable learners to understand varied tone 3 patterns and hopefully reach native-like or near-native performance levels for tone 3.

References

Boersma, Paul and David Weenink. 2020. *Praat: doing phonetics by computer*. https://www.fon.hum.uva.nl/praat/

Chao, Yuen-Ren. 1948. *Mandarin primer: An intensive course in spoken Chinese*. Cambridge, MA: Harvard University Press.

Chao, Yuen-Ren. 1968. *A grammar of spoken Chinese*. Berkeley, CA: University of California Press.

Chen, Meng-Yuan, Zhu Zhu, and Ji Liu. 1983. Disansheng de xingzhi jiqi jiaoxue [On the nature of tone 3 and its pedagogy]. Paper presented at the *Annual Conference of China Association for Teaching Chinese as a Foreign Language*. Beijing, June.

Chen, Nancy F., Darren Wee, Rong Tong, Bin Ma, and Haizhou Li. 2016. Large-scale characterization of non-native Mandarin Chinese spoken by speakers of European origin: Analysis on iCALL. *Speech Communication* 84: 46-56.

Chen, Qing-Hai. 1997. Toward a sequential approach for tonal error analysis. *Journal of the Chinese*

Language Teachers Association 32(1): 21-39.

Chen, Shirley. 1973. The third tone and see-saw pairs: A field technique. *Journal of the Chinese Language Teachers Association* 8(3): 145-149.

Chen, Si, Yunjuan He, Ratree Wayland, Yike Yang, Bei Li, and Chun Wah Yuen. 2019. Mechanisms of tone sandhi rule application by tonal and non-tonal non-native speakers. *Speech Communication* 115: 67-77.

Chen, Ying, Susan Guion-Anderson, and Yi Xu. 2012. Post-focus compression in second language Mandarin. *Proceedings of Speech Prosody 2012*, 410-413.

Cheng, Chin-Chuan. 1973. *A synchronic phonology of Mandarin Chinese*. The Hague: Mouton.

Chin, Tsung. 1987. Half-third first: On the nature of the third tone. *Journal of the Chinese Language Teachers Association* 22(1): 87-101.

Duanmu, San. 2007. *The phonology of standard Chinese*. Oxford: Oxford University Press.

Flege, James Emil, Munro J. Murray, and Mackay Ian R. A. 1995. Effects of age of second-language learning on the production of English consonants. *Speech Communication* 16(1): 1-26.

Hao, Yen-Chen. 2012. Second language acquisition of Mandarin Chinese tones by tonal and non-tonal language speakers. *Journal of Phonetics* 40(2): 269-279.

Howie, John M. 1976. *Acoustical studies of Mandarin vowels and tones*. New York: Cambridge University Press.

Jun, Sun-Ah. 1993. *The phonetics and phonology of Korean prosody*. PhD Dissertation. The Ohio State University.

Jun, Sun-Ah. 1998. The accentual phrase in the Korean prosodic hierarchy. *Phonology* 15(2): 189–226.

Jun, Sun-Ah. 2000. K-ToBI (Korean ToBI) labelling conventions, Version 3. *The Korean Journal of Speech Sciences* 7: 143–169.

Kuznetsova, Alexandra, Per B. Brockhoff, and Rune H. B. Christensen. 2017. lmerTest package: tests in linear mixed effects models. *Journal of Statistical Software* 82(13): 1-26.

Lado, Robert. 1957. *Linguistics across cultures: Applied linguistics for language teachers*. Ann Arbor, MI: University of Michigan Press.

Lee, Ok Joo and Yan Xiong. 2021. Distribution of the Mandarin vowels in typological perspective. *Linguistic Research* 38(2): 329-363.

Lee, Yong-cheol, Ting Wang, and Mark Liberman. 2016. Production and perception of tone 3 focus in Mandarin Chinese. *Frontiers in Psychology* 7:1058.

Li, Charles N. and Sandra A. Thompson. 1977. The acquisition of tone in Mandarin-speaking children. *Journal of Child Language* 4(2): 185–199.

Lin, William C. 1985. Teaching Mandarin tones to adult English speakers: Analysis of difficulties with suggested remedies. *RELC Journal* 16(2): 31-47.

Linge, Olle. 2011. Teaching the third tone in Standard Chinese. *LUP Student Papers*. https://lup.lub.lu.se/luur/download?func=downloadFile&recordOId=1979823&fileOId=1979824.

Liu, Jun and Yong-cheol Lee. 2021. Focus prosody by Korean learners of English. *Linguistic*

Approaches to Bilingualism, 1-30. Amsterdam: John Benjamins. https://doi.org/10.1075/lab.20031.liu.

Liu, Jun, Yi Xu, and Yong-cheol Lee. 2019. Post-focus compression is not automatically transferred from Korean to L2 English. *Phonetics and Speech Sciences* 11: 15-21.

R Core Team. 2020. *A language and environment for statistical computing*. Vienna: R Foundation for Statistical Computing.

Shi, Pei-Wen and Ming Li. 1987. Quansansheng de shiyong he yudiao dui disansheng de yingxiang [Full tone 3 use and effect of intonation on tone 3]. *Selected Papers of the Second International Seminar on Chinese Language Teaching*, 153-159.

Singh, Archana, Avantika Yadav, and Ajay Rana. 2013. K-means with three different distance metrics. *International Journal of Computer Applications* 67(10): 13-17.

Tsai, Rachel. 2011. Teaching and learning the tones of Mandarin Chinese. *Scottish Languages Review* 24: 43–50.

Tu, Jung-Yueh, Hsiung Yuwen, Jih-Ho Cha, Min-Da Wu, and Yao-Ting Sung. 2016. Tone production of Mandarin disyllabic words by Korean learners. In Jonathan Barnes, Nanette Veilleux, Stefanie Shattuck-Hufnage, and Alejna Brugos (eds.), *Proceedings of Speech Prosody 2016*, 375–379.

Wang, Ting, Jun Liu, Yong-hun Lee, and Yong-cheol Lee. 2020. The interaction of tone and prosodic focus in Mandarin Chinese. *Language and Linguistics* 21(2): 331-350.

Wang, Yue, Allard Jongman, and Joan A. Sereno. 2003. Acoustic and perceptual evaluation of Mandarin tone productions before and after perceptual training. *The Journal of the Acoustical Society of America* 113(2): 1033-1043.

Wiener, Seth, Marjorie KM Chan, and Kiwako Ito. 2020. Do explicit instruction and high variability phonetic training improve nonnative speakers' Mandarin tone productions? *Modern Language Journal* 104(1): 152-168.

Wong, Puisan, Richard G. Schwartz, and James J. Jenkins. 2005. Perception and production of lexical tones by 3-year-old, Mandarin-speaking children. *Journal of Speech, Language, and Hearing Research* 48(5): 1065–1079.

Xu, Yi. 1997. Contextual tonal variations in Mandarin. *Journal of Phonetics* 25(1): 61-83.

Xu, Yi. 2013. ProsodyPro — A tool for large-scale systematic prosody analysis. In Dan Brigitte Bigi and Daniel Hirst (eds.), *Proceedings of Tools and Resources for the Analysis of Speech Prosody*, 7-10.

Zhang, Hang. 2014. The third tone: Allophones, sandhi rules and pedagogy. *Journal of the Chinese Language Teachers Association* 49(1): 117-145.

Zhang, Hang. 2017. The effect of theoretical assumptions on pedagogical methods: A case study of second language Chinese tones. *International Journal of Applied Linguistics* 27(2): 363-382.

Zhang, Hang. 2018. *Second language acquisition of Mandarin Chinese tones: Beyond first-language transfe*r. Leiden/Boston: Brill.

Appendix
Stimuli used for this study

Mandarin Chinese	Pinyin	Tone sequence	Mandarin Chinese	Pinyin	Tone sequence
编者	bianzhe	1-3	补习	buxi	3-2
吃苦	chiku	1-3	此时	cishi	3-2
方法	fangfa	1-3	打折	dazhe	3-2
发展	fazhan	1-3	等于	dengyu	3-2
钢笔	gangbi	1-3	改革	gaige	3-2
机场	jichang	1-3	解答	jieda	3-2
开始	kaishi	1-3	紧急	jinji	3-2
科普	kepu	1-3	可能	keneng	3-2
思考	sikao	1-3	朗读	langdu	3-2
污染	wuran	1-3	起床	qichuang	3-2
虚伪	xuwei	1-3	体词	tici	3-2
因此	yinci	1-3	小时	xiaoshi	3-2
英语	yingyu	1-3	语言	yuyan	3-2
中午	zhongwu	1-3	指责	zhize	3-2
伯母	bomu	2-3	打印	dayin	3-4
词典	cidian	2-3	讽刺	fengci	3-4
词语	ciyu	2-3	景色	jingse	3-4
罚款	fakuan	2-3	举办	juban	3-4
合法	hefa	2-3	考试	kaoshi	3-4
合理	heli	2-3	可靠	kekao	3-4
集体	jiti	2-3	恐怕	kongpa	3-4
局长	juzhang	2-3	满意	manyi	3-4
啤酒	pijiu	2-3	米饭	mifan	3-4
食品	shipin	2-3	努力	nuli	3-4
田野	tianye	2-3	跑步	paobu	3-4
王子	wangzi	2-3	使用	shiyong	3-4
无比	wubi	2-3	体育	tiyu	3-4
牙齿	yachi	2-3	仔细	zixi	3-4
本科	benke	3-1	报纸	baozhi	4-3

产区	chanqu	3-1	翅膀	chibang	4-3
反思	fansi	3-1	大米	dami	4-3
卡车	kache	3-1	电子	dianzi	4-3
烤鸭	kaoya	3-1	地理	dili	4-3
可惜	kexi	3-1	地铁	ditie	4-3
老师	laoshi	3-1	汉语	hanyu	4-3
母亲	muqin	3-1	课本	keben	4-3
首都	shoudu	3-1	密码	mima	4-3
手机	shouji	3-1	宁可	ningke	4-3
许多	xuduo	3-1	跳舞	tiaowu	4-3
已经	yijing	3-1	物理	wuli	4-3
指挥	zhihui	3-1	玉米	yumi	4-3
子孙	zisun	3-1	自己	ziji	4-3

Yuping Fu
Associate Professor
School of Foreign Languages
Hainan Tropical Ocean University
No. 1, Yucai Road, Sanya City
Hainan Province, 572022 China
E-mail: fuyuping1969@foxmail.com

Yong-cheol Lee
Assistant Professor
Department of English Language and Literature
Cheongju University
298 Daesung-ro, Cheongwon-gu,
Cheongju, 28503 Korea
E-mail: soongdora@gmail.com

Received: 2022. 01. 23.
Revised: 2022. 02. 24.
Accepted: 2022. 03. 10.