How Does Mandarin Learning Experience Modulate Second-Language Learners' Phonological Knowledge of Tone 3 Sandhi in Word Production?

Zhen Qin

Abstract Tone 3 (T3) in Mandarin is one of the most difficult tones for second-language (L2) learners given its variants in different contexts. While previous studies investigated L2 learners' acquisition of T3 in the sandhi context, it remains unclear how L2 learners' learning experience modulates their phonological knowledge of T3 sandhi in producing different types of words. This study used a wug production test to investigate the effect of Mandarin learning experience on the production of the sandhi form (a rising tone) by experienced and inexperienced Korean-speaking learners. The acoustic analyses showed that experienced Korean-speaking L2 learners were better at using their phonological knowledge of T3 sandhi than less experienced learners when producing pseudo and novel words, but not real Chinese words. The experienced learners had higher pitch values and a steeper rising slope for the T3 sandhi form of pseudo words, and higher pitch values for that of novel words, than the inexperienced learners. The findings suggested that learners' increased experience with Mandarin facilitated their use of the phonological knowledge of T3 sandhi. Given learners' difficulties with the T3 sandhi rule, language teachers are suggested to develop teaching materials with different word types to promote the generalization of the rule.

Keywords Tone 3 sandhi · Phonological knowledge · WUG test · L2 korean learners

1 Introduction

Mandarin Chinese (henceforth, Mandarin) is a language which is rapidly gaining in importance on the international scene. Accordingly, it has become commonly taught in foreign language programs outside of Mandarin-speaking areas (e.g., USA) according to the National Council of Less Commonly Taught Languages (NCOLCTL), and an increasingly large number of students are learning Mandarin.

Z. Qin (✉)
Division of Humanities, The Hong Kong University of Science and Technology, Clear Water Bay, Kowloon, Hong Kong
e-mail: hmzqin@ust.hk

C. Yang (ed.), *The Acquisition of Chinese as a Second Language Pronunciation*, Prosody, Phonology and Phonetics, https://doi.org/10.1007/978-981-15-3809-4_9

Importantly, it differs from many other languages in the types of information they use to convey meaning in words: Mandarin uses lexical tones (i.e., pitch movement) to contrast word meanings (e.g., /pā/ "eight" (Tone 1 (T1)), /pá/ "to pull out" (T2), /pǎ/ "to hold" (T3), /pà/ "father" (T4) (Yip 2006)). Learning tones pose difficulties for adult second-language (L2) learners of Mandarin who speak non-tonal languages as their mother tongue (Pelzl 2019; Wiener, Ito, and Speer 2018). Among the four Mandarin tones, T3 is one of the most challenging tones for English-speaking L2 learners (e.g., Hao 2012; Yang 2019). One possible reason is that T3 is involved in a tone sandhi rule with allophonic variations and thus poses great difficulty in learning its different variants depending on the tone contexts (Yip 2006).

Specifically, different from other Mandarin tones, Mandarin T3 is involved in a phonological alternation as follows: T3 (214; the dipping tone) in Mandarin becomes T2 (35; the rising tone) before another T3. This phonological alternation of tones is called the T3 sandhi rule. It can be written as follows:

214 (T3, the citation form) → 35 (T2, the sandhi form) / ____ 214 (T3, the citation form).

The T3 sandhi rule must apply in Mandarin disyllabic words, for instance, *yǔ sǎn* "umbrella" (Yip 2006). The L2 acquisition of T3 in the sandhi and non-sandhi contexts has been researched in several studies (e.g., Zhang 2016, 2017; Zhang and Xie 2020). For instance, Zhang (2017) tested the production of different allophonic variants of T3 as follows: (a) the citation form (214) of T3 in the final position; (b) the half-form (21) of T3 when preceding T1,T2,and T4; and (c) the sandhi form (35) of T3 when preceding another T3. By examining the production of Mandarin disyllabic words by English-speaking learners with different learning experience, the study showed that the learning experience of Mandarin modulated the learners' use of the T3 half-form. Specifically, inexperienced learner groups (beginner, intermediate level) mispronounced a T3 half-form as a citation form of T3, when it preceded T1, T2, and T4, more frequently than the experienced learner group (advance-level). Pronunciation errors were also found for the T3 sandhi form when it preceded another T3 across learner groups. While the findings of the above-mentioned studies suggested an effect of Mandarin learning experience on L2 learners' acquisition of T3 in different forms, the error-based approach (on real Chinese words) of the studies did not allow us to build a deeper understanding of L2 learners' phonological knowledge of T3 sandhi (i.e., the correct application of the T3 sandhi rule). Therefore, acoustic studies using a wug production test are needed to examine the effect of learning experience on L2 learners' use of their underlying phonological knowledge of T3 sandhi in producing different types of Chinese words.

The wug test, which has been widely adopted to test the productivity of morpho/phonological alternations in non-Chinese languages (Hayes and Londe 2006; Hsieh 1976; Zuraw 2007), is also good at testing Mandarin speakers' phonological knowledge of T3 sandhi (Zhang and Lai 2010). In a typical wug test, experimental participants are taught novel forms (pseudo words or novel words) in their native language (L1) and then asked to provide morphologically complex forms (e.g., Tagalog infixation; see Zuraw, 2007 for details), using the novel forms as the base. A better application of the morpho/phonological alternations in the novel forms indicates a better phonological knowledge of the alternations. The T3 sandhi rule in Mandarin has been shown to be highly productive in novel forms for native Mandarin-speaking adult participants (Zhang and Lai 2010; Zhang, Xia, and Peng 2015) and child participants (Huang, Zhang, and Zhang 2018; Huang, Zuo, and Zhang 2019). For instance, Zhang and Lai (2010) used a wug test to examine adult native speakers' phonological knowledge of Mandarin tone sandhi patterns. The study aimed to test the synchronic relevance of phonetics by investigating native Mandarin speakers' applications of the T3 sandhi process from *real* disyllabic words, consisting of actual occurring (AO) morphemes, to two types of wug words: *pseudo* disyllabic words consisting of non-occurring sequences consisting of real morphemes, and *novel* words consisting of non-occurring sequences of non-occurring syllables of accidental gaps (AG). The results from Zhang and Lai (2010) showed that native speakers' production of pseudo and novel T3 sandhi words shared a greater similarity of pitch shape with the citation form (214) of T3 in having a lower and later turning point (i.e., a less rising slope) than did their production of real sandhi words (the sandhi form, 25). In pseudo and novel words, the T3 sandhi form, with a less rising slope, was produced more like an underlying T3 than it was in real words. This finding indicates an acoustically "incomplete application" of the T3 sandhi rule when generalizing the phonological knowledge to pseudo and novel words in adults. Zhang and Lai (2010) linked the "incomplete application" of the T3 sandhi rule by native speakers to a weaker phonetic motivation for this type of sandhi pattern in nature.

In a similar fashion, the wug test has also been used to test L2 learners' underlying phonological knowledge of T3 sandhi. For instance, Chen et al. (2019) adopted a similar paradigm to examine the ability to produce Mandarin T3 sandhi by two groups of L2 learners, that is, tonal (Cantonese) speakers and non-tonal (English) speakers. The functional data analysis of normalized pitch values revealed that compared with native speakers of Mandarin, L2 learners showed less accurate production of the T3 sandhi form (25) with lower pitch values and less rising slope, which was attributed to L2 learners' less (acoustically) detailed phonological representations of allophonic variants. However, Cantonese- and English-speaking L2 learners applied the T3 sandhi rule similarly for both real words and non-real (pseudo and novel) words, suggesting that their phonological knowledge of T3 sandhi was equally learned. While the learning experience of Mandarin was balanced between the English-speaking and Cantonese-speaking L2 learner groups, the L2 learners' learning experience was not manipulated. It remains unclear whether, and if so, how Mandarin learning experience modulates L2 learners' use of the phonological knowledge of T3 sandhi in producing real, pseudo, and novel words.

To fill this research gap regarding the effect of Mandarin learning experience and to complement findings of previous L2 studies (Zhang 2017) using an error-based approach, the current study adopted the wug test paradigm to investigate the effect of Mandarin learning experience on L2 learners' phonological knowledge of T3 sandhi. Specifically, experienced versus inexperienced (non-tonal) Korean-speaking adult L2 learners, who were relatively understudied in the previous literature, were compared in their productions of different types of Chinese words (real, pseudo, and novel words). The productions were then acoustically analyzed and statistically modeled to examine their phonological knowledge of T3 sandhi and uncover the underlying mechanism involved.

2 Methods

2.1 Participants

Sixteen native Korean speakers (mean age: 24.6, SD: 4.3, nine females and seven males) who learned Mandarin as the L2 participated in this study. All the L2 learners reported that (Seoul) Korean was their L1, and that both their parents were native Korean speakers. They were all college students studying in Shanghai, China. Additionally, they reported having learned Mandarin after the age of 12 and not having been exposed to tone languages other than Mandarin. They reported normal hearing and no history of speech or language disorders. In compensation for their time, the participants each received the equivalent of ten US dollars.

Crucially, to test the effect of Mandarin learning experience on the phonological knowledge of T3 sandhi, eight L2 learners who had passed HSK 5 were recruited as experienced learners. Another eight L2 learners who had not passed HSK 5 were recruited as inexperienced learners. The L2 learners' Mandarin learning experience is summarized in Table 1. As can been seen from their biographical information, both the experienced and inexperienced learners started learning Mandarin at a similar age, that is, around 20 years old. However, the experienced learner group received Mandarin instruction for a longer time in the classroom and spent a longer time immersed in Mandarin-speaking areas than the inexperienced learner group.

Table 1 Biographical information of experienced and inexperienced Korean-speaking L2 learners of Mandarin

	AOE (year)	**Years of Mandarin Instruction**	**LOR (month)**
Experienced Learner Group ($n = 8$)	20 (3.4)	3.4 (2.1)	39.0 (24.4)
Inexperienced Learner Group ($n = 8$)	21 (3.6)	1.6 (1.0)	8.1 (5.9)

Mean (standard deviation), *AOE* age of first exposure to Mandarin, *LOR* length of residence in Mandarin-speaking areas

2.2 Materials

Following the experimental design of Zhang and Lai (2010) and Chen et al. (2019), a Chinese wug test was conducted with the Korean-speaking L2 learners of Mandarin. The items were disyllabic words consisting of either actual occurring (AO) morphemes such as *měi* with 美 as corresponding Chinese character or accidental gaps (AG) such as *hěi* with no corresponding Chinese character. To test the phonological knowledge of T3 sandhi, three sets of disyllabic test words which carry T3 in the first and second syllables, as listed in Table 2, were constructed as follows: (1) real words: real Chinese disyllabic words (AO-AO); (2) pseudo words: non-occurring sequences consisting of actual occurring morphemes (*AO-AO); (3) novel words: non-occurring sequences of non-occurring syllables of accidental gaps (AG–AG).

All the chosen real words and individual character/morpheme for pseudo and novel words were selected from the textbook used for the Mandarin class at Year 1. Thus, the L2 learners in this study were supposed to be familiar with the test items. In order not to reveal the purpose of the experiment, a total of 60 disyllabic filler words were used. The filler words included real words, pseudo words, and novel words in a similar design with the other 15 tonal combinations (T3 + T1; T3 + T2;T3 + T4; T1-T1; T1-T2; T1-T3; T1-T4; T2-T1; T2-T2; T2-T3; T2-T4; T4-T1; T4-T2; T4-T3; T4-T4). In total, there were 12 test words (4 items * 3 word types) and 60 filler words (4 items * 15 tone combinations) with all possible tonal combinations in the test. Before the experiment, each monosyllable used for the test words and filler words was recorded with three repetitions by a female native speaker of Beijing Mandarin. Recordings were conducted in a soundproof room using a microphone linked to a digital recorder. One token for each target monosyllable was chosen from three repetitions by the investigator based on its intelligibility and sound quality.

Table 2 Test words used in the condition of real words, pseudo words, and novel words

Word type	Chinese characters	Chinese pinyin
Real words (AO + AO)	美好 手表 整理 可以	měihǎo shǒubiǎo zhěnglǐ kěyǐ
Pseudo words (*AO + AO)	美朵 手怎 整早 可散	měiduǒ shǒuzěn zhěngzǎo kěsǎn
Novel words (AG + AG)	NA	hěidiǔ cǒusěn sěnduǐ těcǒu

2.3 Procedures

The task was conducted using the paradigm software (Perception Research Systems, Inc. https://www.paradigmexperiments.com/). In each trial, two monosyllables were presented in an auditory mode to the participants with 800 ms in between. Each monosyllable was also presented visually with their characters (if available) and phonetic symbols (pinyin) along with sounds. The participants were instructed to put the two monosyllables they heard together to verbally produce a disyllabic word in Mandarin. They were instructed to speak at a normal speaking rate and could self-correct when necessary. The experiment started with a demonstration session with the investigator demonstrating how to put the two monosyllables together by verbally producing a disyllabic word in Mandarin. The demonstration session was then followed by a practice session involving ten new practice trials, in which the participants were allowed to practice the task before the experimental session. In the experimental session, the order of trial presentation was randomized across participants within one block.

2.4 Measurements

Participants' verbal productions were acoustically coded in Praat (Boersma and Weenink 2018). The fundamental frequency (F0) was extracted at ten equidistant points within each annotated vowel using the ProsodyPro Praat script (Xu 2013). The extracted F0 values were then converted from Hz to semitones with a reference of 50 Hz. The F0 values of each token were z-score normalized against the mean pitch across all tokens for each individual speaker using the following formula: Normalized pitch = Observed pitch – Mean pitch / Standard deviation of pitch (mean pitch and standard deviation of pitch are the grand mean and standard deviations of all tokens per individual participant).

2.5 Data Analysis

Four tokens (one token from the experienced learner group, three tokens from the inexperienced learner group) for disyllabic test words were excluded from the analysis. Two tokens were produced as two isolated monosyllables (syllable interval longer than 300 ms) and the other two tokens were produced with errors (T3 mispronounced as T1 or T4). A total of 188 tokens for disyllabic test words were included in the data analysis (95 tokens from the experienced learner group; 93 tokens from the inexperienced learner group).

The dependent variable for the statistical analyses was normalized pitch values in semitone. The growth curve analysis (GCA) has the intercept for average pitch values, and it also uses the *poly* function to generate two other parameters for pitch shape, the first-order linear polynomial, and the second-order quadratic polynomial. The two polynomials enable us to model participants' normalized pitch shape (curves) over time. According to Mirman (2014), the *intercept* captures the average pitch value with the higher the intercept, the higher the average pitch value; the *linear* polynomial captures the pitch slope with a positive (t) value indicating a rising pitch and a larger value indicating more steepness and vice versa; the *quadratic* polynomial indicates a single-inflection curve of pitch shape with a positive (t) value indicating a concave shape and a larger value indicating a more concave pitch shape and vice versa.

The GCAs were conducted with the *lme4* package in R (Bates et al. 2015). The analyses included the two polynomials (linear and quadratic) modeling pitch shape, condition (real words, pseudo words, and novel words), and group (experienced learners; inexperienced learners) as fixed effects. The effect of condition was dummy coded with real words as baseline, whereas group was contrast-coded (i.e., -0.5 and 0.5). A back-fitting function from the package *LMERConvenienceFunctions* in R (Tremblay and Ransijn 2015) was used to identify the model that accounted for significantly more of the variance than simpler models, as determined by log-likelihood ratio tests; only the results of the model with the best fit are presented, with p values being calculated using the *lmerTest* package in R (Kuznetsova, Brockhoff, and Christensen 2018). All analyses included participant as random intercept and the orthogonal polynomials as random slopes for the participant variable, which allowed the analysis to model a line of a different shape for each individual participant. A larger analysis that tested three-way interaction between the effects of polynomials, condition, and group was conducted to determine whether the two L2 groups differ in their tone production across the three word conditions and justify the main GCAs conducted separately to test the interactions between the effects of polynomials (also, intercept) and group in each condition. The GCAs were conducted separately for each condition with the alpha level being adjusted to 0.017 for each of the three models.

To conclude that the Mandarin learning experience influenced L2 learners' phonological knowledge of T3 sandhi (Chen et al. 2017), the GCA in each word condition must reveal either a main effect of group (interpreted on the intercept) or interactions between group and at least one polynomial (the linear or quadratic polynomial). Specially, a main effect of group indicates the participants' average pitch values are different between the two L2 groups, whereas the interaction between the effects of group and polynomials indicates that the shape (pitch slope indexed by the linear polynomial or concave shape indexed by the quadratic polynomial) of participants' pitch shape is different between the two L2 groups. Given the nature of the wug test, it is predicted that the two L2 groups will show differences in their production of pseudo and novel words, but not necessarily their production of real words which might not tap into L2 learners' underlying phonological knowledge given learners' familiarity with these words.

3 Results

Figure 1 shows normalized pitch values (semitones) of the first syllable in T3 sandhi disyllabic words produced by experienced and inexperienced Korean-speaking learners for real words, pseudo words, and novel words. As illustrated in Fig. 1, the rising slope carried by Syllable 1 indicates that both experienced and inexperienced Korean-speaking L2 learners had correctly applied the T3 sandhi rule to

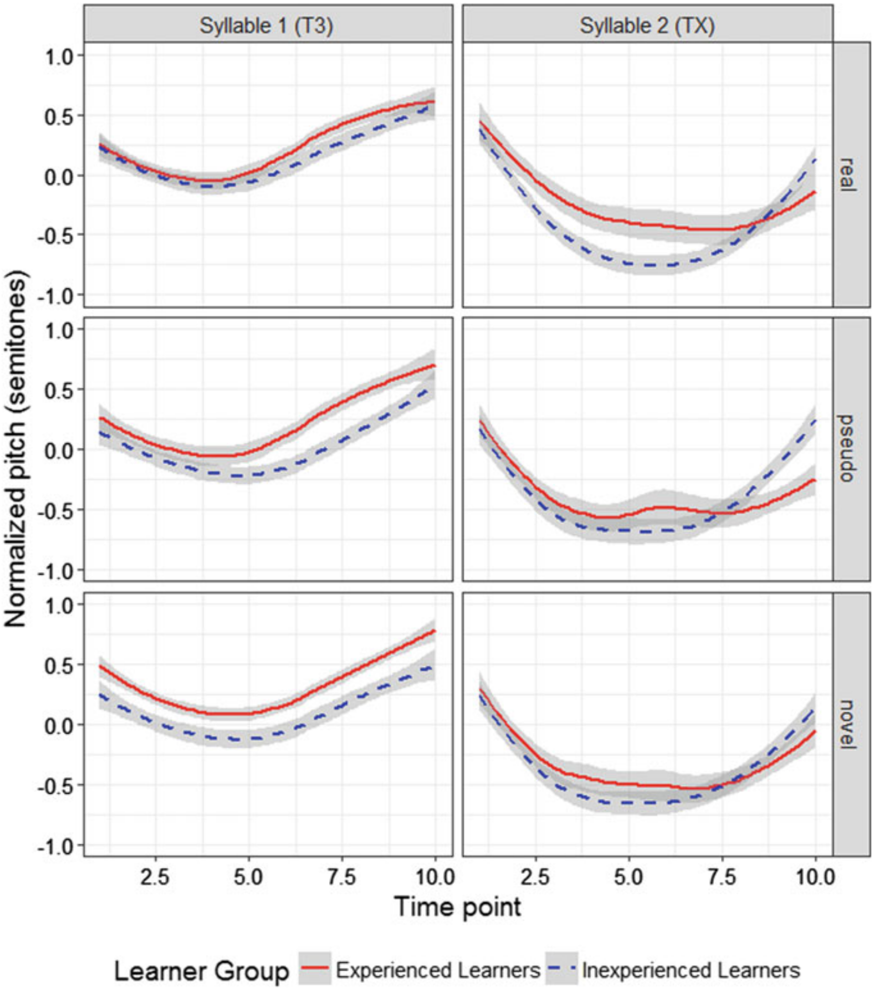

Fig. 1 Normalized pitch values in semitones of Syllable 1 in T3 sandhi words produced by experienced Korean-speaking learners (*red, solid*) and inexperienced learners (*blue, dashed*) in the condition of real words (*top*), pseudo words (*middle*), and novel words (*bottom*); the shaded area represents one standard error above and below the participant mean

real, pseudo, and novel disyllabic words. That is, they changed the tone of the first syllable (T3) to a rising tone (T2) in the T3 sandhi context. However, a visual inspection of Fig. 1 suggests that while the experienced and inexperienced learner groups had a similar pitch shape of Syllable 1 for real words, they had different acoustic realizations of Syllable 1 for pseudo words and novel words.

To determine whether the two L2 groups differ in their tone production across the three word conditions and justify the main analysis of three GCAs conducted in each condition as illustrated in Fig. 1, a larger GCA was performed on the normalized pitch values (semitone) of Syllable 1 in T3 sandhi disyllabic words with the effect of condition (real words, pseudo words, and novel words; baseline: real words), group (experienced learners; inexperienced learners; contrast-coded), and two polynomials (linear and quadratic; baseline: linear) modeling pitch shape as fixed factors. The results of this GCA with the best fit are presented in detail in Table 4 of the Appendix. The GCA with the best fit on the normalized pitch values included the linear and quadratic polynomials, condition, group, and the interactions between the linear polynomial and other factors. Importantly, there was a three-way interaction between the linear polynomial, condition, and group (for the pseudo words condition). The main GCAs were therefore performed on the effects of the linear polynomial and group separately for real words, pseudo words, and novel words, as illustrated in Fig. 1.

Table 3 presents the results of the main GCAs which included an effect of linear polynomial, group (contrast-coded), and their interaction in each condition. Of the results of the GCA on real words in Table 3, the significant positive *t* value for the linear polynomial indicates that L2 learners' normalized pitch for the real words had a rising pitch slope. There was no significant effect of group (experienced: 0.19 semitone; inexperienced: 0.16 semitone). Crucially, there was no significant interaction effect between the linear polynomial and group. The results of the model on real words suggest that the experienced and inexperienced learner groups had their T3 sandhi form of Syllable 1 produced as a rising tone with no difference in average pitch values and pitch slope.

For the results of the GCA on pseudo words, the linear polynomial was not significant. However, the significant positive *t* value for the group effect (experienced: 0.29 semitone; inexperienced: 0.05 semitone) indicates that the T3 sandhi form of the experienced learner group had higher pitch values than that of the inexperienced learner group for pseudo words. Crucially, the significant positive *t* value for the interaction between the linear polynomial and group indicates that the T3 sandhi form of the experienced learner group had a steeper rising slope than that of the inexperienced learner group for pseudo words. The results of the model on pseudo words suggest that the experienced learner group had their T3 sandhi form of Syllable 1 produced with higher pitch values and a steeper rising slope than the inexperienced learner group.

For the results of the GCA on novel words, the linear polynomial was not significant. However, the significant positive *t* value for the group effect (experienced: 0.29 semitone; inexperienced: 0.14 semitone) indicates that the T3 sandhi form of the experienced learner group had higher pitch values than that of the inexperienced

Table 3 Growth curve analyses on normalized pitch values (semitone) of syllable 1 in the T3 sandhi context for real words, pseudo words, and novel words

Condition	Effect	Estimate	t	p
Real words	(Intercept)	0.409	6.736	**< 0.001**
	Polynomial			
	Linear	0.380	3.211	**< 0.01**
	Group	0.010	0.226	0.821
	Polynomial × Group			
	Linear	0.322	2.280	0.023
Pseudo words	(Intercept)	0.276	4.485	**< 0.001**
	Polynomial			
	Linear	− 0.177	− 1.945	0.069
	Group	0.117	3.921	**< 0.001**
	Polynomial × Group			
	Linear	0.272	2.880	**< 0.01**
Novel words	(Intercept)	0.339	4.759	**< 0.001**
	Polynomial			
	Linear	− 0.088	− 0.874	0.389
	Group	0.111	2.612	**< 0.01**
	Polynomial × Group			
	Linear	− 0.222	− 1.650	0.099

$\alpha = 0.017$, significant results are in bold, real words: $n = 640$ observations, pseudo words: $n = 620$ observations, novel words: $n = 620$ observations

learner group for novel words. There was no significant interaction effect between the linear polynomial and group. The results of the model on novel words suggest that the experienced learner group had their T3 sandhi form of Syllable 1 produced with higher pitch values, without difference in pitch slope, than the inexperienced learner group.

To summarize the experienced and inexperienced groups of Korean-speaking L2 learners, both correctly pronounced a rising tone for Syllable 1 of the real words; however, their phonological knowledge of T3 sandhi in pseudo words and novel words differed. Specifically, the experienced learner group had higher pitch values and a steeper rising slope for pseudo words, and higher pitch values for novel words, than the learner group with less learning experience.

4 Discussion

The present study examined the effect of Mandarin learning experience on Korean-speaking L2 learners' phonological knowledge of T3 sandhi when producing real, pseudo, and novel Chinese words. The results of the acoustic analysis indicated that the Korean-speaking L2 learners with more learning experience were better at using the phonological knowledge of T3 sandhi than less experienced learners in producing pseudo and novel words, but not real words. In the text below, we first discuss the effect of Mandarin learning experience on the L2 learners' better use of phonological knowledge of tones and then turn to the discussion of L2 pedagogical implications which could be applied to the Mandarin language classroom.

First, the finding regarding the Mandarin learning experience is consistent with previous studies on the L2 acquisition of Mandarin tones in isolation (Hao 2018; Sun 2012) and in tonal contexts (He and Wayland 2010; Yang 2011). While T3 was often mispronounced as T2 by L2 learners as the two tones are neutralized in the T3 sandhi context (Hao 2012; Yang 2011), an increased learning experience could help L2 learners reduce pronunciation errors of the T3 half-form (21) in the non-sandhi context (Zhang 2017). Complementing the existing studies of L2 tone acquisition, our findings supported with acoustic evidence further suggested that an increased learning experience also facilitated the L2 learners' use of their phonological knowledge in the sandhi context, with the experienced learners producing the T3 sandhi form (35) of pseudo and novel words more accurately than the inexperienced learners. Compared with inexperienced L2 learners of Mandarin, experienced L2 learners might have been more exposed to natural tonal variants in the Mandarin-speaking environment and thus should have more robust representations of lexical tones (Qin, Tremblay, and Zhang 2019).

Second, the results of T3 sandhi production across different word types found for the experienced and inexperienced L2 learner groups are also in line with developmental studies testing children's phonological knowledge of T3 sandhi (Huang, Zhang, and Zhang 2018; Huang, Zuo, and Zhang 2019). With an increased exposure to natural tonal variants of T3 and a stronger phonological/morphological awareness, a developmental trajectory emerged with older children's production of the T3 sandhi form becoming more adult-like than younger children's production. In a similar vein, a learning trajectory was found in this study with the experienced learners' production of the T3 sandhi form, specifically for pseudo and novel words, becoming more acoustically accurate than the inexperienced learners' production.

The effect of Mandarin learning experience revealed for pseudo and novel words, but not for real words, is presumably attributed to the nature of Mandarin T3 sandhi. As Zhang and Lai (2010) found, adult native Mandarin speakers differently pronounced the T3 sandhi form between real words and non-real words (including both pseudo and novel words), with the T3 sandhi form for non-real words having a less rising slope than that for real words. Thus, even native speakers may experience some difficulty using their phonological knowledge of T3 sandhi, which is not as phonetically motivated as other T3 variants (i.e., T3 half-form) (Zhang and Lai 2010). It is not surprising that, like adult native Mandarin speakers, L2 learners also had difficulty applying the rule when producing the pseudo words and novel words with less experienced L2 learners having greater difficulty than their more experienced counterparts. Another possibility is that L2 learners' lexical knowledge, such as their familiarity with spoken forms of real words, may have assisted them in applying the T3 sandhi rule successfully for real words instead of other words regardless of their Mandarin learning experience. To tease apart the effect of tone sandhi nature from L2 learners' word familiarity, future studies may consider testing the (experienced and inexperienced) L2 learners' phonological knowledge of different types of tone sandhi (less phonetically motivated tone sandhi such as T3 sandhi variant vs. more phonetically motivated tone sandhi/coarticulation such as T3 half-form variant) using both real and non-real words (see Chen et al. 2017 for an example).

Interestingly, the experienced learner group had higher pitch values and a steeper rising slope than the inexperienced learner group for pseudo words, suggesting a better use of their phonological knowledge of T3 sandhi. However, they showed higher pitch values than their inexperienced counterparts only for novel words. One plausible explanation for the different effects of Mandarin learning experience between pseudo words and novel words is that both experienced and inexperienced learners shared a difficulty applying the T3 sandhi rule in novel words (not in pseudo words) given the novelty of the items. On the other hand, the experienced learners' increased exposure to Mandarin tonal input and larger vocabulary size might have resulted in a greater explicit awareness of novel words, which was presumably indexed by higher pitch values in their production of the T3 sandhi form (see Huang et al. 2019 for similar findings). To corroborate the plausibility, further studies are required to recruit native speakers as reference to investigate whether the T3 production of experienced learners is more acoustically native-like in terms of average pitch values and a rising slope than that of inexperienced learners (Chen et al. 2019).

This research not only deepens our understanding of the mechanism underlying L2 learners' production of the T3 sandhi form but also provides pedagogical implications for Mandarin language teachers. The results of our current research showed that inexperienced Korean-speaking L2 learners of Mandarin had greater difficulty than their experienced counterparts in generalizing their phonological knowledge of T3 sandhi from production of real words to that of pseudo and novel words. This difficulty for inexperienced L2 learners was possibly attributed to a less robust representation of T3, which has different tone variants. As Zhang (2017) pointed out, the citation form (214) of T3 is often taught to L2 learners first and thus might be treated as a default form of T3 by L2 learners. In contrast, both the sandhi form (35) and the T3 half-form (21) are not introduced in detail in the classroom setting. Accordingly, quite a few L2 Mandarin teaching practitioners also assume that the citation form (214) of T3 is the primary form and thus treats other variants of T3 as "unnatural" (Sun 1997). As a result of classroom instruction, L2 learners who have limited exposure to tonal variants will not have an explicit awareness of the differences between T3 variants, and they will be less likely to apply the T3 sandhi rule to contexts other than familiar contexts (e.g., real words). Mandarin teaching practitioners are thus suggested to treat T3 as a special case given its difficulty for L2 learners and then integrate detailed pronunciations of its sandhi form and other variants in the L2 teaching curriculum (see the chapter by Jiang Liu).

To help L2 learners overcome their difficulties using their phonological knowledge of T3 sandhi, Mandarin language teachers are specifically encouraged to develop L2 teaching materials that focus on the application of the T3 sandhi rule, preferably using not only real words but also pseudo/novel words (Zhang 2017, 2018). One approach to deal with L2 learners' difficulty would be to have them complete intensive training in a laboratory setting, in which learners are required to learn to pay more attention to the differences between the T3 sandhi form and other variants in their production as well as perception of real and non-real (pseudo and novel) words (Li, Yang, and Chen 2018). Furthermore, to assist L2 learners in building a robust representation of Mandarin tones, acoustically variable tonal stimuli produced by different speakers (female and male), in different tonal contexts (sandhi and non-sandhi contexts), and in different types of words (real and non-real words), can be used in such a training paradigm (see the chapter by Yingjie Li). This high-variability training paradigm would initially improve L2 learners' ability in distinguishing tonal variants (the citation form vs. the sandhi form) in different contexts (Chang and Bowles 2015; Liu and Zhang 2016; Wang et al. 1999; Wang, Jongman, and Sereno 2003). And it would result in a more robust representation of tonal categories in the long term and ultimately a more efficient use of Mandarin tones for L2 learners.

To the best of our knowledge, the present study is one of the first to examine the effect of Mandarin learning experience on L2 learners' phonological knowledge of T3 sandhi in word production. The findings suggest that experienced Korean-speaking L2 learners were better in using their phonological knowledge of T3 sandhi than less experienced learners in producing pseudo and novel words, but not real words. These findings shed light on L2 learners' underlying mechanism of producing tones in the sandhi context and provide pedagogical implications for Mandarin teaching in the classroom setting. More importantly, the study sparks interest in questions regarding the different types of tone sandhi and the native likeness of the tone sandhi production for further research.

Acknowledgements We thank Dr. Haifeng Qi at the Shanghai International Studies University for her involvement in the earlier stages of the project and her assistance in data collection. We also thank the editor and the reviewer for their insightful comments on this research.

Appendix

(See Table 4).

Table 4 Growth curve analyses on normalized pitch values (semitone) of Syllable 1 in the T3 sandhi context

Effect	Estimate	*t*	*p*
(Intercept)	0.176	3.223	**< 0.01**
Polynomial			
Linear	0.750	6.026	**< 0.001**
Quadratic	0.626	8.664	**< 0.001**
Group	0.010	0.246	0.806
Condition (Pseudo)	– 0.046	– 0.256	0.798
Condition (Novel)	0.034	1.574	0.116
Polynomial × Group			
Linear	0.322	2.482	**0.013**
Polynomial × Condition (Pseudo)			
Linear	– 0.224	– 2.416	**0.016**
Polynomial × Condition (Novel)			
Linear	– 0.266	– 2.868	**0.004**
Group × Condition (Pseudo)	0.112	1.913	**0.051**
Group × Condition (Novel)	0.089	1.522	0.128
Polynomial × Group × Condition (Pseudo)			
Linear	– 0.548	– 2.961	**0.003**
Polynomial × Group × Condition (Novel)			
Linear	– 0.041	– 0.221	0.825

$\alpha = 0.05$, significant results are in bold,$n = 1880$ observations

References

Bates, D. M., Maechler, M., Bolker, B., & Walker, S. (2015). lme4: linear mixed-effects models using S4 classes. *Journal of Statistical Software*, *67*(1), 1–48. http://dx.doi.org/https://doi.org/10.18637/jss.v067.i01

Boersma, P., & Weenink, D. (2018). *Praat: doing phonetics by computer [Computer program]. Version 6.0.43*. Retrieved 8 September 2018.

Chang, C. B., & Bowles, A. R. (2015). Context effects on second-language learning of tonal contrasts. *Journal of the Acoustical Society of America, 138*(6), 3703–3716. https://doi.org/10.1121/1.4937612.

Chen, S., He, Y., Wayland, R., Yang, Y., Li, B., & Yuen, C. W. (2019). Mechanisms of tone sandhi rule application by tonal and non-tonal non-native speakers. *Speech Communication, 115,* 67–77. https://doi.org/10.1016/j.specom.2019.10.008.

Chen, S., He, Y., Yuen, C. W., Li, B., & Yang, Y. (2017). Mechanisms of tone sandhi rule application by non-native speakers. *Proceedings of the Annual Conference of the International Speech Communication Association, INTERSPEECH*, *Stockholm*, *Sweden*, 1760–1764. https://doi.org/https://doi.org/10.21437/Interspeech.2017-143

Hao, Y. C. (2012). Second language acquisition of Mandarin Chinese tones by tonal and non-tonal language speakers. *Journal of Phonetics, 40*(2), 269–279. https://doi.org/10.1016/j.wocn.2011.11.001.

Hao, Y. C. (2018). Second Language Perception of Mandarin Vowels and Tones. *Language and Speech, 61*(1), 135–152. https://doi.org/10.1177/0023830917717759.

Hayes, B., & Londe, Z. C. (2006). Stochastic phonological knowledge: The case of Hungarian vowel harmony. *Phonology, 23*(1), 59–104. https://doi.org/10.1017/S0952675706000765.

He, Y., & Wayland, R. (2010). The Production of Mandarin Coarticulated Tones by Inexperienced and Experienced English Speakers of Mandarin. *Proceedings of the 5th International Conference on Speech Prosody*, *Chicago*, *United States*, 1–4.

Hsieh, H. I. (1976). On the unreality of some phonological rules. *Lingua, 38*(1), 1–19. https://doi.org/10.1016/0024-3841(76)90038-3.

Huang, X., Zhang, G., & Zhang, C. (2018). A Preliminary Study on the Productivity of Mandarin T3 Sandhi in Mandarin- speaking Children. *Proceedings of the 6th International Symposium on Tonal Aspects of Languages, Berlin, Germany*, 88–92.

Huang, X., Zuo, Y., & Zhang, C. (2019). Seven-year-olds reach an adult-like productivity in the application of Mandarin tone sandhi. *Proceedings of the 19th International Congress of Phonetic Sciences, Melbourne, Australia.*, 3125–3129.

Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2018). lmerTest Package: Tests in Linear Mixed Effects Models. R package version 3.0–1. *Journal of Statistical Software*. https://doi.org/https://doi.org/10.18637/jss.v082.i13

Li, B., Yang, Y., & Chen, S. (2018). Perceptual Evaluation of Mandarin Tone Sandhi Production by Canton-ese Speakers before and after Perceptual Training. *Proceedings of the 32nd Pacific Asia Conference on Language, Information and Computation*, *Hong Kong*, 358–366. https://www.aclweb.org/anthology/Y18-1041

Liu, J., & Zhang, J. (2016). The effects of talker variability and variances on incidental learning of lexical tones of Asian Languages and Literatures , University of Minnesota , USA 2 . Incidental Learning — Video Game. *Proceedings of the 5th International Symposium on Tonal Aspects of Languages*, *Buffulo*, *United States*, 23–27.

Mirman, D. (2014). *Growth curve analysis and visualization using R*. CRC Press.

Pelzl, E. (2019). What makes second language perception of Mandarin tones hard? *Chinese as a Second Language, 54*(1), 51–78. https://doi.org/10.1075/csl.18009.pel.

Qin, Z., Tremblay, A., & Zhang, J. (2019). Influence of within-category tonal information in the recognition of Mandarin-Chinese words by native and non-native listeners: An eye-tracking study. *Journal of Phonetics, 73,* 144–157. https://doi.org/10.1016/j.wocn.2019.01.002.

Sun, K.-C. (2012). *The role of lexical tone in L2 Mandarin spoken word recognition*. Unpublished Ph.D. thesis. Department of Linguistics, the University at Buffalo, State University of New York.

Sun, S. (1997). *The development of a lexical tone phonology in American adult learners of standard Mandarin Chinese*. University of Hawaii at Manoa.

Tremblay, A., & Ransijn, J. (2015). LMERConvenienceFunctions: A suite of functions to back-fit fixed effects and forward-fit random effects, as well as other miscellaneous functions. R package version 2.1. *Comprehensive R Archive Network*.

Wang, Y., Jongman, A., & Sereno, J. A. (2003). Acoustic and perceptual evaluation of Mandarin tone productions before and after perceptual training. *Journal of the Acoustical Society of America, 113*(2), 1033–1043. https://doi.org/10.1121/1.1531176.

Wang, Y., Spence, M. M., Jongman, A., & Sereno, J. A. (1999). Training American listeners to perceive Mandarin tones. *Journal of the Acoustical Society of America, 106*(6), 3649–3658. https://doi.org/10.1121/1.428217.

Wiener, S., Ito, K., & Speer, S. R. (2018). Early L2 Spoken Word Recognition Combines Input-Based and Knowledge-Based Processing. *Language and Speech, 61*(4), 632–656. https://doi.org/10.1177/0023830918761762.

Xu, Y. (2013). ProsodyPro - A tool for large-scale systematic prosody analysis. *Proceedings of Tools and Resources for the Analysis of Speech Prosody*, 7–10.

Yang, C. (2011). *The acquisition of Mandarin prosody by American learners of Chinese as a foreign language (CFL)*. Unpublished Ph.D. Thesis. Graduate Program in East Asian Languages and Literatures, Ohio State University.

Yang, C. (2019). The effect of L1 tonal status on the acquisition of L2 Mandarin tones. *International Journal of Applied Linguistics (United Kingdom), 29*(1), 3–16. https://doi.org/10.1111/ijal.12223.
Yip, M. (2006). Tone: Phonology. *Encyclopedia of Language & Linguistics*. https://doi.org/10.1016/b0-08-044854-2/00054-7.
Zhang, C., Xia, Q., & Peng, G. (2015). Mandarin third tone sandhi requires more effortful phonological encoding in speech production: Evidence from an ERP study. *Journal of Neurolinguistics, 33,* 149–162. https://doi.org/10.1016/j.jneuroling.2014.07.002.
Zhang, H. (2016). Dissimilation in the second language acquisition of Mandarin Chinese tones. *Second Language Research, 32*(3), 427–451. https://doi.org/10.1177/0267658316644293.
Zhang, H. (2017). The effect of theoretical assumptions on pedagogical methods: A case study of second language Chinese tones. *International Journal of Applied Linguistics (United Kingdom), 27*(2), 363–382. https://doi.org/10.1111/ijal.12132.
Zhang, H. (2018). Current trends in research of Chinese sound acquisition. In *The Routledge handbook of Chinese second language acquisition* (pp. 217–233). Taylor and Francis. https://doi.org/10.4324/9781315670706.
Zhang, H., & Xie, Y. (2020). Coarticulation effects of contour tones in second language Chinese. *Chinese as a Second Language Research, 9*(1), 1–30. https://doi.org/10.1515/caslar-2020-0001.
Zhang, J., & Lai, Y. (2010). Testing the role of phonetic knowledge in Mandarin tone sandhi. *Phonology, 27*(1), 153–201. https://doi.org/10.1017/S0952675710000060.
Zuraw, K. (2007). The role of phonetic knowledge in phonological patterning: Corpus and survey evidence from tagalog infixation. *Language, 83*(2), 277–316. https://doi.org/10.1353/lan.2007.0105.