Production of lexical tones by Southern Min-Mandarin bilinguals

Karen Huang
The University of Auckland, New Zealand

Abstract
This is a preliminary study examining the tonal production of L1 Taiwanese Southern Min (TSM) speakers who are also fluent in Mandarin. Both languages have tone sandhi rules in which certain lexical tones are neutralized in non-XP-final positions. Disyllabic Mandarin and TSM words with different tonal combinations in frame sentences were examined. The results suggest that Mandarin Tone 1, Tone 3 and Tone 4 were assimilated to TSM high level, low falling and high falling tones respectively due to their surface phonetic similarity. L1 TSM speakers can apply Mandarin Tone 3 sandhi without difficulties, but the L1 sandhi rule has affected the production of Mandarin Tone 2. Mandarin Tone 2 (rising tone) was mapped to TSM rising tone (Tone 5) in phrase-final position, but not elsewhere due to the phonological constraints operated through tone sandhi. Mandarin Tone 2 at a non-XP-final position was produced with pitch contours in-between TSM rising and mid level tones, creating a “merged” category. The findings seem to indicate that aside from the phonetic properties, the L2 tones are also influenced by the L1 phonological rules such as tone sandhi in the production of non-native tones, creating context-specific tone production.

Index Terms: lexical tones, second-language speech, Mandarin, Southern Min, bilingualism, tone sandhi

1. Introduction
Studies on the perception of non-native lexical tones by lexical tone users have shown that the L1 tonal acoustic characteristics play an important role, but the phonological system of the L1 or L2 tones can also be a factor [1-9]. For example, Cantonese listeners have difficulties distinguishing the high level tone from the high falling tone in Mandarin because they are allotones in their L1 [1, 8]. Furthermore, for experienced L2 learners, the perceptual assimilation occurs at both phonological and phonetic levels [10]. The L2 learners might acquire the L2 tonal categories and assimilate the L2 tones differently, compared to the naïve listeners who assimilate the L2 tones into their native tones categories based on the phonetic similarity [9].

However, fewer studies were done on the production of non-native lexical tone by lexical tone users. Moreover, for languages with complex tone sandhi rules in which tonal distinctions are neutralized in certain positions, it is unclear how L1 and L2 tonal systems interact. Specifically, how bilingual speakers would produce a sandhi L2 tone, and whether or not L1 sandhi rules would influence the L2 tones.

This study examines how L1 Taiwanese Southern Min (TSM) speakers produce their native language and Mandarin as a second language. Mandarin has four lexical tones, and TSM has seven — including two “checked tones”, which are for syllables end in an oral stop. Table 1 shows the tone values of TSM and Mandarin. There is some documented dialectal variation in TSM, but the tone categories for the non-checked tones can generally be presented as high level, high falling, low falling (low), rising, and mid level. The two TSM checked tones are not relevant to this study because there are no syllables that end in oral stops in Mandarin. As for Mandarin, its four lexical tones are high level, high falling, rising and dipping (low). There are phonetic differences between standard and Taiwan Mandarin, and this study can also examine whether the dialectal differences can be attributed to L1 TSM influences.

Table 1. TSM and Mandarin tone values

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>T1</td>
<td>55&quot;</td>
<td>T1 &quot;55&quot;</td>
<td>T1 &quot;44&quot;</td>
</tr>
<tr>
<td>T2</td>
<td>&quot;51&quot;</td>
<td>&quot;51&quot;</td>
<td>&quot;42&quot;</td>
</tr>
<tr>
<td>T3</td>
<td>&quot;21&quot;</td>
<td>&quot;21&quot;</td>
<td>&quot;214&quot;</td>
</tr>
<tr>
<td>T4</td>
<td>&quot;21&quot;</td>
<td>&quot;21&quot;</td>
<td>&quot;312&quot;</td>
</tr>
<tr>
<td>T5</td>
<td>&quot;24&quot;</td>
<td>&quot;12&quot;</td>
<td>&quot;35&quot;</td>
</tr>
<tr>
<td>T6</td>
<td>&quot;23&quot;</td>
<td>&quot;23&quot;</td>
<td>&quot;323&quot;</td>
</tr>
<tr>
<td>T7</td>
<td>&quot;53&quot;</td>
<td>&quot;53&quot;</td>
<td>&quot;53&quot;</td>
</tr>
<tr>
<td>T8</td>
<td>&quot;85&quot;</td>
<td>&quot;85&quot;</td>
<td>&quot;85&quot;</td>
</tr>
</tbody>
</table>

In Mandarin, T3 becomes T2 when it is before another T3 (Tone 3 sandhi). In TSM, the five non-checked tones form a “tone circle” in the tone sandhi process, in which the citation tones change to another tone in non-XP-final positions as shown in Figure 1 below. Both T1 and T5 merge to T7 in non-XP-final positions. Therefore there is no rising tone in non-final positions in TSM. It remains to be seen whether this distribution asymmetry influences the way TSM users produce the Mandarin rising tone.

Figure 1: Southern Min Tone Sandhi circle

2. Methods
2.1. Materials
In order to find out the tonal target of each lexical tone, tones in both languages were elicited in different tonal contexts in a frame sentence. For Mandarin tones, 32 meaningful disyllabic words were chosen to elicit 16 (4*4) bi-tonal combinations in the frame sentence qǐng shuō ___  ___ bā cì. ‘Please say __ __ eight times.’. As for TSM tones, 92 meaningful disyllabic words were chosen to elicit 49 (7*7) bi-tonal combinations. All the words chosen here are frequent vocabulary which belong to the colloquial stratum instead of the literary stratum.

The words are written in Chinese characters that approximate the reading of individual syllables with a Mandarin translation listed occasionally. Subjects were asked to put the words in the frame sentence guá kóng __ __ kóng sann pái ‘Please say __ __ say three times.’

2.2. Subjects

This preliminary study includes two male and two female TSM-Mandarin bilinguals, as well as a Mandarin monolingual were recruited. All the bilingual subjects grew up in a TSM speaking family with TSM as their native language. They all acquired Mandarin in school (age 7) and continue to use Mandarin in public domains, and they are fluent in both languages. The Mandarin monolingual does not speak any TSM. All subjects were born and raised in Taiwan.

2.3. Procedures

The subjects were recorded in a quiet room with a Roland R-05 digital recorder with the sample rate of 44100 Hz. They first completed a short questionnaire on their linguistic backgrounds, then read the 92 TSM testing sentences and the 32 Mandarin testing sentences. The subjects were asked to repeat each sentence twice at the speed that they were comfortable with.

2.4. Data Analysis

The rimes of the target disyllabic words were hand labelled and \( f_0 \) measurements extracted at the midpoint of all the 1/10th intervals in Praat. The \( f_0 \) readings in semitone were normalized to z-scores for cross-speaker comparison. The pitch contours shown in the figures below present the normalized rimes (x-axis) in semitone z-score (y-axis).

3. Results

3.1. Lexical tone targets in TSM and Mandarin by bilinguals

3.1.1. High level and high falling tones

The high level (TSM: T1; Mandarin: T1) and the high falling tones (TSM: T2; Mandarin: T4) with varying preceding lexical tones produced by the bilingual subjects are shown in Figure 2. The acoustic results suggest that the high level and high falling tones in both languages are similar in terms of their pitch heights and pitch contours. High level can be characterized as [44] and high falling can be characterized as [53] in both Mandarin and TSM.

Figure 3 shows the falling tone data of the Mandarin monolingual. Compared to Figure 2, the monolingual’s high falling tone started slightly lower. The differences between the monolingual and the bilinguals are more obvious in the first syllable. Unlike the bilinguals, the monolingual’s T1 did not intercept with T4. Instead T4 intercept with T2. The flattening of the bilinguals’ T2 will be discussed in the following section.

Figure 3: High falling tone with varying preceding lexical tone in Mandarin by a Mandarin monolingual speaker

The TSM tone values produced by the speakers match with Cheng’s description [12]. The starting pitch of the high level tone is lower than that of the high falling tone. Furthermore, the degree of the falling produced by the subjects was not as strong as Peng’s description [11]. The Mandarin tonal values of the bilinguals are slightly different from described in Standard Mandarin and Taiwan Mandarin in Table 1, but similar to the recent result [15]. The similarity between TSM and Mandarin high level and high falling tones, along with the differences between the monolingual and the bilinguals, suggests that the L1 TSM speakers assimilate the Mandarin high level and high falling tones to the TSM ones.

3.1.2. Low falling tone

As shown in Figure 4, both TSM Tone 3 and Mandarin Tone 3 show a low falling contour, reaching from -1 to -2 semitone z-score. The tone value can therefore be described as “21”. It is different from what was described in Table 1 because the tone is elicited in a sentence instead of in isolation.

The pitch contours in Mandarin are uneven because of the creaky voice quality. About 30% of targeted Tone 3 syllables in this position for both languages were discarded due to the creaky voice. Many remaining syllables also had missing values due to the voice quality. There were no significant differences regarding the percentage of creaky voice. However, there were fewer targeted Mandarin syllables. Therefore the effect of the missing values was more obvious.

Figure 4: Low falling tone with varying preceding tones in TSM and Mandarin by bilingual speakers

The result suggests that the bilingual speakers applied Tone 3 Sandhi in Mandarin. According to Myers and Tsay [16], Taiwan Mandarin speakers tend to process the sandhi
rule categorically instead of phonetically, and acoustic studies show that there is no significant difference between the sandhi Tone 3 and Tone 2. However, in this study we find that at the 45% of the first syllable, the difference between Mandarin sandhi Tone 3 and Tone 2 are statistically significant (t(20)=3.02, p=.007). Nevertheless, the segment of the tested syllable was not controlled. The differences might be a result of the segmental influences.

3.1.3. Rising tone

Figure 5 shows TSM Tone 5 (rising) and Tone 7 (mid level) produced by the bilingual subjects, and Mandarin Tone 2 (rising) produced by the bilingual subjects and the Mandarin monolingual. The rising tones are in the second syllable, phrase-final position. Our results show that Mandarin Tone 2 produced by the monolingual has a shallower pitch fall followed by a steeper pitch rise, but Mandarin Tone 2 produced by the bilinguals has a dipping pitch contour “323”, as documented, which is similar to their TSM Tone 5. On the other hand, TSM Tone 7 has flatter pitch contours with a slightly higher pitch.

Due to Southern Min tone sandhi, TSM Tone 5 (rising) does not occur in non-XP-final positions. Figure 6 compares the bilinguals’ Mandarin Tone 2 (rising) and TSM Tone 7 (sandhi Tone 1 and Tone 5) at the first syllable. The figure show that the bilingual did not assimilate Mandarin Tone 2 to TSM Tone 7, but the degree of rising seems limited.

As illustrated in Figure 3 and part of Figure 5, the pitch contours of Mandarin Tone 2 produced by the monolingual have a much steeper rise at both the first and the second syllables. To calculate the degree of pitch rises, the pitch changes from 45% to 85% of the syllables were examined. A two-way ANOVA shows that it was significantly influenced by both speaker group (F(1, 128)=5.22, p=.001) and its position (F(1, 128)=14.87, p=.001). Post hoc tests reveal that Mandarin T2 produced by the monolingual has a steeper rise than T2 produced by the bilinguals (p<.001). Furthermore, Mandarin T2 at the second syllable (M-XT2) is steeper than Mandarin T2 at the first syllable (M-T2X) when they are produced by the bilingual (p<.001), but the pitch rises of the monolingual’s M-XT2 (M=0.98, SD=0.23) are similar to M-T2X (M=0.93, SD = 0.22) (p=0.98). The monolingual has a steeper pitch rise for T2 everywhere, while Mandarin Tone 2 produced by the bilinguals is context-sensitive.

To further examine the rising/mid tones of the bilingual speakers, the pitch rises were compared factoring the four bilingual subjects and the tone categories. Five tone categories were compared: M-T2X, M-XT2, TSM T5 and T7 at Syllable 2 (SM-XT5, SM-XT7), and TSM Tone 7 at Syllable 1 (SM-T7X). A two-way ANOVA yields a significant main effect on subject (F(3, 391)=34.3, p<.001) and tone category (F(4, 391)=42.57, p<.001). The interaction between subject and tone category is also significant (F(12, 391)=5.44, p<.001). Post hoc comparison using Tukey HSD test reveals that the degree of pitch rises of M-XT2 is significantly greater than SM-XT7 (p<.001), but the differences between M-XT2 and SM-XT5 are not significant (p=0.91). The statistical analysis supports the observation that the bilingual speakers produced Mandarin Tone 2 with a similar pitch contour as TSM Tone 5. On the other hand, at the first syllable, M-T2X rises more than SM-T7X, and the differences are statistically significant (p<.001). This suggests that the bilinguals did not assimilate Mandarin Tone 2 to TSM Tone 7 at this position. Nevertheless, the degree of pitch rise for M-T2X was not as large as M-XT2 (p<.001). The bilinguals produce in-between pitch contours at non-XP-final positions.
merging with Tone 7, but her M-XT2 is steeper than SM-XT5 and SM-XT7, suggesting that she created a new category for M-XT2. Interestingly, F16’s M-T2X is still flatter than M-XT2 ($p=.02$). Her M-T2X was not as flat as SM-T7X, but the difference was not statistically significant. Her Mandarin Tone 2 was not assimilated to TSM tones, but it still has a similar context-sensitive variation. As for M13, his M-XT2, M-T2X, SM-XT5 all share a similar slight pitch rise, and his SM-XT7 and SM-T7X share a slight pitch fall. The differences between the two groups were statistically significant. This suggests that for M13, Mandarin Tone 2 was assimilated to TSM Tone 5 regardless of its position.

4. Discussion

Overall, the data presented here suggested that the L1 TSM speakers assimilate the Mandarin tones into their native lexical tone categories. Mandarin high level and high falling were assimilated into TSM high level and high falling, as supported by the acoustic evidence. Mandarin Tone 3, the dipping tone, was assimilated to the low falling tone. This is not surprising because Mandarin Tone 3 rarely shows the complete dipping contour unless it is in a citation form or an utterance-final position. Our data here is elicited in a frame sentence, in which a dropping contour is not expected. Studies also have shown Thai learners of Mandarin assimilated Mandarin Tone 3 with a dipping variant to the Thai low falling tone, while naive Thai listeners assimilated it to the Thai rising tone [9]. This finding can also help explain why Tone 3 in Taiwan Mandarin tends to be low-falling even in a citation form [17-21], which suggests that L1 TSM speakers assimilate Mandarin Tone 3 (dipping ~ low falling) to TSM Tone 3 (low falling), and tend to apply the assimilation context-free.

The bilingual speakers seemed to produce Mandarin Tone 3 sandhi without any difficulties. In TSM, it is possible to have consecutive low falling tones (TSM T7-T3 → T3-T3). Therefore the bilingual most likely applied Mandarin Tone 3 sandhi as a phonemic rule instead of a phonetic process [16].

The results of the Mandarin rising tone (T2) show that at phrasal-final position, it seemed to be assimilated to TSM rising tone (T5) instead of TSM mid tone (T7). TSM T5 is a low rising tone and Mandarin T2 is a high rising tone. Nevertheless, the bilinguals still tend to map Mandarin T2 to TSM T5 instead of T7. Literature also found that tone language users are more sensitive to $f_0$ contours than $f_h$ height. Therefore L2 tones are more likely to be assimilated to L1 tones with similar $f_0$ contours [9].

The result also seems to indicate context-sensitive productions of Mandarin Tone 2. Except for M13, all the subjects produced a flattened Mandarin Tone 2 at non-XP-final positions. This suggests that for M13, Mandarin Tone 2 has been acquired at the phonemic level. M13 was able to produce the similar rising contour for Mandarin T2 regardless of its position. On the other hand, the fact that TSM rising tone does not occur in non-XP-position has influenced the other subjects’ production of the Mandarin rising tone. This suggests that the process of assimilation was at the context-sensitive allophonic level, as hypothesized by the Speech Learning Model (SLM) [22]. Another piece of evidence to support that these context-sensitive differences are the result of L1 phonology rather than articulatory constraints is that TSM speakers are relatively aware that the rising tone is not permitted in non-XP-final positions. Of all the TSM tone sandhi rules, T5 → T7 is more productive compared to the others when applying Wug tests [23-25].

For these speakers, the in-between pitch rise suggests that Mandarin T2 at non-XP-final position was likely not assimilated to TSM T7, nor was tone sandhi (TSM T5→ T7) applied to it. In the terms of the Perceptual Assimilation Model – Suprasegmental, PAM-S [2], at phrasal-final position, Mandarin T2 and any other Mandarin tone are likely to be perceived as belonging to ‘Two-Categoriy’, but at the non-XP-final positions, they are likely to be perceived as belonging to ‘Uncategorized-Categorized’, with no perceived overlap. The ‘Uncategorized’ Mandarin Tone 2 at the non-XP-final position is produced in in-between pitch contours. It demonstrates that they created a “merged” phonetic category. The similar in-between phonetic feature has been reported on bilingual’s segment production. For example, English-French bilinguals produce a merged category in which the VOT values of both their L1 and L2 /t/ were in-between English’s /t/ (long-lag VOT) and French’s /t/ (short-lag VOT) [26, 27]. However, this merged category in our observation is context-specific, suggesting the L1-L2 interaction at both phonetic and phonological levels.

When a contour tone is in a position with less tone-bearing capacity, such as a non-phrasal final position, it might undergo (a) no change, (b) partial contour reduction (flattening) or (c) complete contour reduction (becoming a level tone) depending on the ranking of the phonological constraints [28, 29]. Our finding shows that M13 exhibited (a) for Mandarin and (c) for TSM. While the other three subjects exhibited (b) for Mandarin and (c) for TSM. Their phonetic differences might result from the different rankings of constraints in their L2 phonology.

Studies have shown that Mandarin Tone 2 in Taiwan varies in its contour shapes [19, 20, 30]. At sentence-medial position, Mandarin Tone 2 was more likely to have a level contour. Furthermore, generally bilingual speakers are more likely to have a low-rising contour, while Mandarin monolinguals are more likely to have a high rising contour [31]. This study also provides a possible motivation of the flattened pitch contour by the bilingual speaker.

The L1-L2 phonetic influences are bidirectional [32]. Furthermore, multiple factors such as age of L2 acquisition, language use, and the length in L2 environment affect the L2 speech. More data controlling these factors are needed to further understand the tonal interactions among these bilinguals.

5. Conclusions

This preliminary analysis suggests that while L1 Southern Min speakers mapped L2 Mandarin T1, T2, and T4 onto TSM high level (T1), low-falling (T3) and high falling (T2) respectively due to their surface phonetic similarity, Mandarin rising tone (T2) is mapped to TSM rising tone (T5) only at the phrase-final position, but not at the non-XP-final position, due to L1 phonological constraints. Instead, the bilinguals tend to create a new phonetic category that is between rising and mid-level for Mandarin Tone 2 at this position. This study suggests that the L1 TSM sandhi rule might influence its L2 tonal production in a context-sensitive way. The preliminary results further our understanding in second language tone acquisition and dialect formation, and may also have potential implications on speech recognition or related fields.
6. References


