Perception of Mandarin Tones by Native Tibetan Speakers

Wenfu Bao1, Hui Feng1,2,3, Jianwu Dang3,4, Zhilei Liu3, Yang Yu1, Siyu Wang3

1 Research Center for Linguistic Sciences, Tianjin University, China
2 School of Liberal Arts and Law, Tianjin University, China
3 Tianjin Key Laboratory of Cognitive Computing and Application, China
4 School of Information Science, Japan Advanced Institute of Science and Technology, Japan

Abstract

Previous studies have demonstrated that the second language (L2) learners’ linguistic backgrounds and L2 proficiency have an effect on their perception of L2 sounds. This paper attempts to investigate the assimilation patterns of Mandarin and Tibetan tones, and the influences of first language (L1) backgrounds and Mandarin proficiency on the perception of Mandarin tones. A total of 46 Tibetan participants, including 14 Khams and 32 U-Tsang speakers, were instructed to assimilate the Mandarin tones they’ve heard to their most similar native tones. Results suggest that the four-tone system U-Tsang speakers match M155 to T155, M452 to T452, with great disparities in mapping M235 and M314 to T212 and T312, while the two-tone-system Khams speakers tend to assimilate M155 and M452 to the high tone, and M235 and M314 to the low tone. Mandarin Chinese proficiency does show the progressive tone-mapping patterns, that is, the higher the learners’ proficiency is, the more possible he or she will have the tone mapping which could be predicted by the tone values. As proficiency level increases, standard deviation of learners’ mapping tends to get smaller, especially among Khams speakers.

Index Terms: speech perception, Tibetan-Mandarin tone mapping, L1 backgrounds, L2 proficiency, second language acquisition

1. Introduction

It is well-attested that linguistic experience affects the perception of non-native tones. In the case of non-tone language speakers, for instance, [1] found that the discrimination of Mandarin tone contrasts by Japanese listeners is affected by their language backgrounds. [2] discovered that English listeners’ lack of tone experience may cause more rising responses than their Mandarin counterparts when identifying non-speech flat and rising pitch contours. Similarly, for tone language speakers, Cantonese subjects in [3] would probably ignore the inflection point of Mandarin Tone 3, due to the lack of similar experience in their native language. The perceptual experiment in [4] asked Tibetan speakers to identify the four Mandarin tones, and results showed distinct perception performance in different dialect groups and people with varying Mandarin proficiency. However, few studies focus on the influence of their L1 experience on L2 learning by asking them to give the mapping between L1 and L2 tone systems, the results of which may cast some light on the causes of poor production of Mandarin Chinese tones by some speakers whose L1 is not Mandarin.

Most Sino-Tibetan languages are tonal [5]. Depending on the pitch contours of the tones, the four citation tones in Mandarin could be categorized as either high level (M155), low rising (M235), low dipping (M314), or high falling (M452) [6]. Of the three Tibetan dialects, U-Tsang and Khams are tonal, though there is no agreement on the tone patterns and tone values of the two dialects, for example, in [6], [7] and [8]. An integrated tone system for Lhasa Tibetan is proposed, which is composed of four basic tones, namely, high level (T155), high falling (T452), low rising (T212) and low dipping (T312) [9]. A simpler tone system with only two tones, high tone (T1) and low tone (T1i), is agreed on for Khams Tibetan [7]. Table 1 shows the correspondence between the two Tibetan tone systems, and between Tibetan and Mandarin tones.

### Table 1. Comparison of Mandarin and Tibetan tone systems. The subscript number represents its tone value.

<table>
<thead>
<tr>
<th></th>
<th>Mandarin</th>
<th>U-Tsang</th>
<th>Khamsp</th>
</tr>
</thead>
<tbody>
<tr>
<td>First Tone (high level)</td>
<td>M155</td>
<td>T155</td>
<td>High tone (Tn)</td>
</tr>
<tr>
<td>Fourth Tone (high falling)</td>
<td>M452</td>
<td>T452</td>
<td></td>
</tr>
<tr>
<td>Second Tone (low rising)</td>
<td>M235</td>
<td>T212</td>
<td>Low tone (T1i)</td>
</tr>
<tr>
<td>Third Tone (low dipping)</td>
<td>M314</td>
<td>T312</td>
<td></td>
</tr>
</tbody>
</table>

Perceptual Assimilation Model (PAM) compares L2 sound contrasts to L1 categories, and predicts the levels of difficulty in differentiating L2 sounds on the basis of how a pair of L2 segments is assimilated to L1 sounds [10] [11] [12]. PAM classifies the major categorization patterns for non-native contrasts along with the expected discrimination performance as Two Category (TC), Single Category (SC), Category Goodness (CG) or Categorized-Uncategorized (CU) contrast depending on the similarities and discrepancy between L2 contrasts and between L2 and L1 sounds. When two non-native sounds are assimilated to two different L1 categories (TC assimilation), discrimination is predicted to be very good, whereas discrimination is poor when they are mapped onto a single L1 category as equally good exemplars of it (SC assimilation). In cases where both members of the L2 contrast are mapped to the same L1 sound, but one member is a “better” exemplar of the native category than the other, discrimination is expected to vary depending on the degree of
the CG difference. Furthermore, L2 sounds can be
uncategorizable, that is, still heard as speech sounds but not
mapped onto any single L1 category, or heard as non-speech
[12].

Based on PAM, the similarity and difference among the
tone systems in Table 1 may lead to such prediction that U-
Tsang speakers possibly encounter more difficulty in
assimilating M2\textsubscript{5} and M3\textsubscript{314}, and Khams speakers may
outperform U-Tsang speakers in matching Mandarin tones to
their native tonal categories. Thus, this study aims to
investigate the cross-linguistic tonal assimilation pattern
within the Sino-Tibetan language family, and to explore the
role of L1 backgrounds and L2 proficiency play in the
perception of L2 tones.

2. Methods

2.1. Participants

The participants were 46 native Tibetan speakers,
including 14 Khams speakers and 32 U-Tsang speakers, with
an average age of 21.20 years (SD=1.59).

In order to discriminate the subjects into different groups
according to their L2 proficiency, all of them were required to
take a Putonghua test before the study. And two state-level
Putonghua assessors were invited to mark their performance in
reading monosyllabic words and polysyllabic words.

Table 2. Results of the perceptual assimilation of Mandarin-
Tibetan tones by U-Tsang speakers (%).

<table>
<thead>
<tr>
<th>T1\textsubscript{55}</th>
<th>T2\textsubscript{13}</th>
<th>T3\textsubscript{132}</th>
<th>T4\textsubscript{52}</th>
<th>None</th>
</tr>
</thead>
<tbody>
<tr>
<td>M1\textsubscript{55}</td>
<td>11.04</td>
<td>1.35</td>
<td>14.58</td>
<td>1.88</td>
</tr>
<tr>
<td>M2\textsubscript{35}</td>
<td>7.71</td>
<td>4.79</td>
<td>7.29</td>
<td>5.00</td>
</tr>
<tr>
<td>M3\textsubscript{214}</td>
<td>4.90</td>
<td>2.50</td>
<td>10.42</td>
<td>2.92</td>
</tr>
<tr>
<td>M4\textsubscript{51}</td>
<td>2.60</td>
<td>16.77</td>
<td>2.92</td>
<td>16.15</td>
</tr>
</tbody>
</table>

Despite the high portion of “None”, a choice meaning no
counterparts of Mandarin tones in Tibetan, a general tendency
still reveals itself in Table 2. M1\textsubscript{55} was most frequently
mapped onto T3\textsubscript{132} (14.58%), with a percentage slightly higher
than the expected M1\textsubscript{55}-T5\textsubscript{55} matching (11.04%). M2\textsubscript{35} was
mapped most often onto T1\textsubscript{55} (7.71%) and T3\textsubscript{132} (7.29%).
M3\textsubscript{214} was mapped most of the time onto T3\textsubscript{132} (10.42%).
However, it should be noted that though the largest percentage
of Tibetan speakers (79.27%) considered that there was no
perfect match in Tibetan for M3\textsubscript{214}, a majority of those who
thought otherwise agreed on the mapping between M3\textsubscript{214}-T3\textsubscript{132},
M4\textsubscript{51} was most frequently judged to be similar to T2\textsubscript{13}
(16.77%), which was the most readily perceived Mandarin
tone, with subtly higher frequency and similarity ratings than
T4\textsubscript{52} (16.15%).

As shown in Table 2, a vast majority of U-Tsang speakers
could not map the Mandarin tone they’ve heard onto its
corresponding Tibetan tonal category, so this case can be
interpreted as the uncategorized. And most of the Mandarin-
Tibetan tone matching by U-Tsang speakers is the TC
assimilation pattern. Additionally, their perception of M1\textsubscript{55}
and M3\textsubscript{314} as T3\textsubscript{132} was an example of the GC. The mapping
data also indicate that the most similar and expected tonal
categories in Mandarin and Tibetan languages are M4\textsubscript{51}-T4\textsubscript{52}
and M1\textsubscript{55}-T1\textsubscript{55}. For U-Tsang speakers, the most difficultly
perceived tone is M2\textsubscript{35}, followed by M3\textsubscript{314}. And the tone
which presents the largest gap in similarity rating is M4\textsubscript{51}.

3.2. Cross-linguistic tone confusion for Khams
speakers

Compared with U-Tsang speakers, the percentage of “None”
option selected by Khams listeners is much smaller. They
perceived M4\textsubscript{51} most similar to T5\textsubscript{10} (31.90%), followed by M1\textsubscript{55}.
Then M2\textsubscript{35} are most frequently perceived as T\textsubscript{10} (30.23%),
followed by M3\textsubscript{314}. Considering the disparity between T\textsubscript{510}
and T\textsubscript{10}, within each Mandarin tone, the discrepancy between the T\textsubscript{510}
and T\textsubscript{10} is much smaller.
and $T_L$ for M415 and M2214 (above 20%) is twice as much as that for M135 and M3214. Due to the characteristics of Khams dialect, Khams speakers tend to perform well in perceiving Mandarin tones as their similar native tonal categories.

From the perspective of categorization patterns, around 60% of the matching could be deemed as uncategorized. Different from the U-Tsang speakers, most of the Khams speakers present a CG assimilation pattern, with two L2 tones mapped to the same tone in Khams dialect but in varying goodness.

Table 3. Results of the perceptual assimilation of Mandarin-Tibetan tones by Khams speakers (%).

<table>
<thead>
<tr>
<th></th>
<th>$T_H$</th>
<th>$T_L$</th>
<th>None</th>
</tr>
</thead>
<tbody>
<tr>
<td>M135</td>
<td>21.19</td>
<td>15.00</td>
<td>63.81</td>
</tr>
<tr>
<td>M431</td>
<td>31.90</td>
<td>9.05</td>
<td>59.05</td>
</tr>
<tr>
<td>M235</td>
<td>10.00</td>
<td>30.23</td>
<td>59.76</td>
</tr>
<tr>
<td>M3214</td>
<td>12.86</td>
<td>22.86</td>
<td>64.29</td>
</tr>
</tbody>
</table>

3.3. Effects of L1 backgrounds

Both dialect groups show the predominance of “None” option in the Mandarin-Tibetan tonal matching, to the effect that most of the Tibetan participants thought Mandarin tones were uncategorizable. Nevertheless, the percentage of “None” selected by Khams group is much smaller than that by U-Tsang group, which suggests that Khams speakers are more sensitive to the assimilation of Mandarin-Tibetan tones than U-Tsang speakers.

Overall, the comparison between the perceptual results of U-Tsang and Khams speakers reveals that U-Tsang speakers’ tonal assimilation pattern is more complicated than that of Khams participants. Specifically, the four-tone system U-Tsang speakers match M135 to T135, M431 to T432, with great disparities in mapping M235 and M3214 to T213 and T3132, while the two-tone system Khams speakers tend to assimilate M135 and M431 to $T_L$, and M235 and M3214 to $T_L$. Consistent with previous findings in [1] and [2], the difference in these two groups’ perceptual results indicates the role of L2 learners’ L1 play in their perception of L2 sounds.

3.4. Effects of Mandarin proficiency

According to the test grades of the Tibetan participants, Khams and U-Tsang speakers were separately divided into three groups, as illustrated in Table 4 and Table 6. The number of subjects and the mean score in each grade internal are also given. And these scores were considered as a representation of their L2 proficiency.

Table 4. Distribution of Khams speakers with different Mandarin proficiency

<table>
<thead>
<tr>
<th>Group</th>
<th>Grades</th>
<th>Number of subjects</th>
<th>Mean score (SD)</th>
</tr>
</thead>
<tbody>
<tr>
<td>K1</td>
<td>47 &lt; score ≦ 67</td>
<td>2</td>
<td>51.00 (1.17)</td>
</tr>
<tr>
<td>K2</td>
<td>67 &lt; score ≦ 77</td>
<td>6</td>
<td>73.00 (1.04)</td>
</tr>
<tr>
<td>K3</td>
<td>77 &lt; score ≦ 100</td>
<td>6</td>
<td>82.00 (0.63)</td>
</tr>
</tbody>
</table>

Results of Mandarin-Tibetan tone matching for Khams speakers are shown in Table 4. As their test grades increase, the standard deviation of their mean scores declines, while from Table 5, the percentage of M1&M4-$T_H$ and M2&M3-$T_L$ displays a rising trend. And the gap between Group K1 and Group K2 is the most substantial. In addition, $T_H$ was assimilated more often than $T_L$ by Khams participants.

Table 5. Tonal assimilation patterns of Khams speakers with different Mandarin proficiency (%).

<table>
<thead>
<tr>
<th>Group</th>
<th>M1 &amp; M4 - $T_H$</th>
<th>M2 &amp; M3 - $T_L$</th>
</tr>
</thead>
<tbody>
<tr>
<td>K1</td>
<td>1.67</td>
<td>1.67</td>
</tr>
<tr>
<td>K2</td>
<td>26.11</td>
<td>25.55</td>
</tr>
<tr>
<td>K3</td>
<td>29.44</td>
<td>27.22</td>
</tr>
</tbody>
</table>

The similar rising trend can also be found in U-Tsang subjects. Table 6 shows that the difference among the three groups is much smaller than that of the Khams group. Moreover, as displayed in Table 7, the U-Tsang learners performed the best in matching M431-T432, followed by M135-T135 and M3214-T3132, and relatively undesirable in assimilating M235-T213. As a result, no matter what Mandarin proficiency they remain, U-Tsang speakers are supposed to pay more attention to the learning of M235.

Table 6. Distribution of U-Tsang speakers with different Mandarin proficiency.

<table>
<thead>
<tr>
<th>Group</th>
<th>Grades</th>
<th>Number of subjects</th>
<th>Mean score (SD)</th>
</tr>
</thead>
<tbody>
<tr>
<td>U1</td>
<td>67 &lt; score ≦ 73</td>
<td>5</td>
<td>70.20 (0.51)</td>
</tr>
<tr>
<td>U2</td>
<td>73 &lt; score ≦ 80</td>
<td>20</td>
<td>76.70 (0.47)</td>
</tr>
<tr>
<td>U3</td>
<td>80 &lt; score ≦ 100</td>
<td>7</td>
<td>84.50 (0.84)</td>
</tr>
</tbody>
</table>

Table 7. Tonal assimilation patterns of U-Tsang speakers with different Mandarin proficiency (%).

<table>
<thead>
<tr>
<th>Group</th>
<th>M1-T1</th>
<th>M2-T2</th>
<th>M3-T3</th>
<th>M4-T4</th>
</tr>
</thead>
<tbody>
<tr>
<td>U1</td>
<td>4.67</td>
<td>2.67</td>
<td>3.33</td>
<td>6.67</td>
</tr>
<tr>
<td>U2</td>
<td>10.67</td>
<td>3.33</td>
<td>10.00</td>
<td>15.17</td>
</tr>
<tr>
<td>U3</td>
<td>16.67</td>
<td>10.48</td>
<td>16.67</td>
<td>25.71</td>
</tr>
</tbody>
</table>

From the mapping data of participants with varying Mandarin proficiency, it can be noted that with the increase of their Mandarin test grades, the likelihood of their matching accuracy was also boosted. For example, the gap between K1 and K3 when assimilating M1&M4 to $T_H$ was about 28%. Therefore, L2 proficiency has a tremendous effect on the perception of L2 suprasegmental features.

4. Conclusion

To explore the cross-linguistic tonal matching within the same language family, and the effects of L1 backgrounds and L2 proficiency on the perception of L2 sounds, this study recruited 46 native Tibetan speakers as subjects to test their assimilation patterns of Mandarin-Tibetan tones. And according to their dialects, these Tibetan participants were...
divided into two groups to examine the role of their L1 backgrounds may play in the perception of L2 tones. Also their L2 proficiency was taken into consideration as a possible contributing factor.

There are three major findings in this study. First, although a majority of the Tibetan speakers could not map the Mandarin tones onto their native Tibetan tones, there is still a general tendency of matching Mandarin-Tibetan tones, suggesting that the cross-linguistic similarity between learners’ L1 and L2 sounds contributing to the perception of L2. Second, the two Tibetan dialect groups present different tonal assimilation patterns, which implies the effect of L1 backgrounds on L2 perception. For instance, U-Tsang speakers thought the most similar tonal categories were M15-T15 and M45-T45, and they had more difficulty in perceiving M25 and M35. Third, Mandarin proficiency shows the progressive tone-mapping patterns, that is, the higher the learners’ proficiency, the more possible he or she will have the tone mapping which could be predicted by the tone values. As proficiency level increases, standard deviation of learners’ mapping tends to get smaller, especially among Kham speakers.

For future research, a relevant production study is needed to explore the relation between perception and production of Mandarin tones, and the effects of L1 on the production of Mandarin tones by native Tibetan speakers. Also, contributing factors like gender difference, age of L2 acquisition etc. should be taken into account to further our understanding of L2 learning.

5. Acknowledgements

This work is partially supported by the National Program on Key Basic Research (No. 2013CB329301), the National Natural Science Foundation of China (No. 61233009) and Social Science Fund of Tianjin (No. TJWWW13-001).

6. References