Effect of musical experience on Mandarin tone and vowel discrimination and imitation

Terry Gottfried and Yangqing Xu

Lawrence University, P O Box 599, Dept. of Psychology, Appleton, WI 54912-0599, USA
gottfrit@lawrence.edu

.

Recent studies have demonstrated a relation between musical experience and accuracy in speaking and perceiving unfamiliar speech contrasts. For example, Gottfried [1] found that native speakers of American English with musical training perform better than non-musicians when discriminating and imitating the four lexical tones of Mandarin Chinese. In this study, four native speakers of Mandarin (Chengdu region) produced target words in a fixed carrier sentence. Their productions of high unrounded /i/ and high rounded /u/ were significantly affected by the initial consonant, especially /u/ (which is often transcribed as two different vowels in pinyin, *u* and *ü,* according to the initial consonant). Vowels that followed consonants with more anterior articulation (pinyin *j, q, x*) had lower F1 and higher F2 than vowels that followed more posterior consonants (pinyin *zh, ch, sh*). The F0 of the word immediately preceding the target words (*kan4*) was significantly affected by the tone of the target word. Current research tests whether musicians and non-musicians differ in their ability to discriminate and imitate these phonemic contrasts according to native speakers' phonemic categories, which allow contextual variation in acoustic information.

1   Introduction

Previous studies have found that musicians with conservatory training show greater accuracy than non-musicians in both discrimination and production of Mandarin Chinese tones [e.g., 1, 2, 3]. Although some studies [e.g., 4] have failed to show any significant relation between musical ability and second language (L2) proficiency, Slevc and Miyake [5], studying Japanese learners of English, found that musical ability (measured by tests of tone and chord perception and memory) significantly predicted L2 production and perception accuracy, but had no significant relation to lexical or syntactic knowledge in the L2.

Thus, musical experience or ability seems related specifically to phonetic learning in a second language, perhaps especially in the early stages of learning. The present study attempts to compare musicians and non-musicians on discrimination and production of unfamiliar Mandarin tone and vowel contrasts. The vowel contrasts were tested to see whether musicians' advantage was specific to tone perception—where they might approach tonal discrimination as a musical pitch task—or whether musicians have a more generalized advantage in discriminating important phonetic contrasts.

2   Acoustical Analysis

2.1   Method

Two female and two male native speakers (ages 20-21 years) of Standard Mandarin Chinese from Chengdu each recorded 14 syllables (Pinyin *zhi, chi, shi, ji, qi, xi, yi, zhu, chu, shu, ju, qu, xu, yu*) spoken with each of the four tones of Mandarin. Tone 1 is a level high tone; Tone 2 is a mid-rising tone; Tone 3 is a low-falling-rising tone; and Tone 4 is a high-falling tone. Each of these target words was recorded in the context of two carrier phrases *Qing3 shuo1* __ ("Please say __") and *Qing3 kan4* __ ("Please look at __"). These 56 (14 syllables × 4 tones) target words in two carrier phrase contexts were acoustically analyzed using the speech analysis computer program Praat [6].

2.2   Results

Measures of fundamental frequency (*F0*) contours of the Mandarin tones showed similar patterns to those reported by Howie [7] and Tseng [8]. Because the target words were in phrase-final position, Tone 3 consistently showed the "dipping" pattern (cf. [9, 10]). There were also significant tonal coarticulation effects (cf. [10, 11]). For example, the beginning *F0* of *kan4* (a word preceding the target words) was lower when the following target word was Tone 1 or 4; beginning *F0* was higher when the following target word was Tone 2 or 3.

The formant frequencies of the high unrounded /i/ and high rounded /u/ were affected by the initial consonant, especially /u/ (which is often transcribed as two different vowels in pinyin, *u* and *ü,* according to the initial consonant). Table 1 shows that vowels that followed consonants with more anterior articulation (pinyin *j, q, x*) had lower *F1* and higher *F2* than vowels that followed more posterior consonants (pinyin *zh, ch, sh*). The Euclidean distance (in the *F1-F2* space) between the allophones of /i/ following anterior and posterior consonants is smaller than the distance between allophones of /u/ following anterior and posterior consonants.

Preceding consonant	/i/		/u/	
	F1	*F2*	*F1*	*F2*
Anterior	324.7	2277.7	338.1	2050.7
Posterior	399.1	1907.4	419.0	993.3

Table 1 Mean midpoint *F1* and *F2* values (in Hz) for /i/ and /u/ following anterior (*j, q, x*) and posterior (*zh, ch, sh*) consonants

3   Perceptual Discrimination

3.1   Method

Participants included 25 students at Lawrence University, all native speakers of American English and non-speakers of Mandarin; they were categorized as musicians and non-

musicians according to their self-rating of musicianship on a 8-point scale (14 musicians; 11 non-musicians). We performed a median split, using 4 as our cut-off for creating musician (scores 4-7) and non-musician (scores 0-3) groups. Five native speakers of Mandarin were also tested, all students at Lawrence from various regions in China where Mandarin Chinese is the dominant language.

Tone discrimination. The *shu* syllables were excised from the sentences spoken by the four native speakers (described above) and presented in pairs with either same or different tones. All possible tone comparisons were presented (1-1, 1-2, 1-3, 1-4, 2-2, 2-3, 2-4, 3-3, 3-4, 4-4) with order of the tones and order of the speakers counterbalanced. An equal number of same and different trials were presented. After a brief orientation to Mandarin tones, participants performed the same-different discrimination test of 192 randomized trials with a 500-ms interstimulus interval between the pair of words to be compared and a 3-s intertrial interval. Participants listened over headphones and indicated whether the tone was the "same" or "different" by pressing keys on a computer keyboard.

Vowel discrimination. Tone 4 syllables were excised from the sentences spoken by the native speakers and presented in pairs with either the same or different vowels. As noted above, the front and back versions of /u/ are in complementary distribution in Mandarin, even though they are often transcribed as different phonemes. The front and centralized /i/, also in complementary distribution, are usually transcribed with one phonetic symbol. In this test, only comparisons of the allophonic variations of /u/ and of /i/ were presented: *qi* vs. *chi*, *xi* vs. *shi*, *ji* vs. *zhi*, *qu* vs. *chu*, *xu* vs. *shu*, and *ju* vs. *zhu*.

3.2   Results

Figure 1 shows the mean correct tone discrimination by non-native musicians and non-musicians and by native listeners. Statistical analysis revealed an effect of musicianship, such that over all tone comparisons, musicians performed significantly more accurately ($M = 82\%$) than non-musicians ($M = 66\%$), $F (1, 23) = 23.14$, $p < .001$. However, these musicians did not perform as accurately as native listeners ($M = 97\%$). The Tone 2-3 comparison was the most difficult for all listeners ($M = 66\%$), including native listeners, compared to the other comparisons (mean percent correct ranging from 80% to 87%).
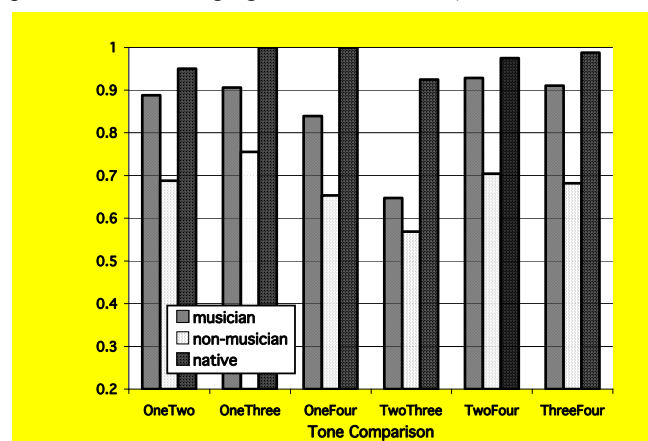

Fig. 1. Mean proportion of correct tone discrimination for each tone comparison by non-native musicians, non-musicians, and native listeners.

Likewise on the vowel discrimination test (Figure 2), musicians performed significantly better ($M = 85\%$) than non-musicians ($M = 75\%$), $F (1, 23) = 8.98$, $p = .006$. However, native listeners actually performed with significantly *lower* accuracy ($M = 74\%$) than musicians and about equal to non-musicians. Native listeners were significantly below chance performance on their discrimination of the /i/ allophones; although significantly above chance on their discrimination of the /u/ allophones, they were significantly less accurate on these vowel contrasts than non-native musicians.

For both non-native listener groups, overall vowel discrimination ($M = 80\%$) was significantly more accurate than tone discrimination ($M = 74\%$), $F (1, 23) = 8.91$, $p = .007$. The reverse was the case for native listeners: Vowel discrimination was less accurate than tone discrimination.
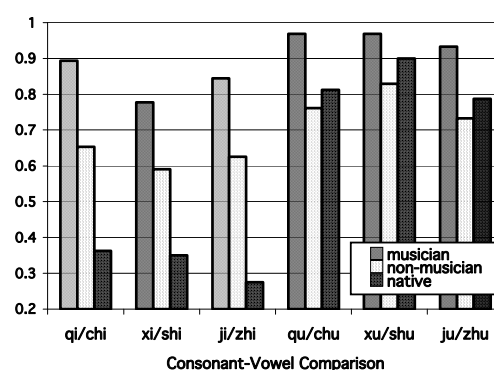

Fig. 2. Mean proportion of correct vowel discrimination for each consonant-vowel comparison by non-native musicians, non-musicians, and native listeners.

4   Imitation

4.1   Method

Spoken imitations were also collected from 10 musicians and 7 non-musicians who had also participated in the perceptual tests. For tone imitation, participants heard the syllable /shu/ spoken with the four tones by the four native speakers (same speakers as used in the perceptual tests). These 16 syllables were randomized and presented over headphones. For vowel imitation, participants heard /shi/, /shu/, /xi/, and /xu/, all with Tone 4, spoken by the four native speakers, in random order. Participants were told to repeat each syllable exactly as they heard it into a microphone. These imitations were recorded and then randomly presented to three native speakers of Mandarin who rated the quality of the tone or vowel imitation on a 10-point scale (0=very poor; 9=native).

4.2   Results

There was no significant effect of musicianship on imitation accuracy for tones or vowels, $p = .07$. However, there were significant interactions of musicianship and the particular syllable to be imitated. Musicians were significantly *more* accurate in their imitation of Tone 2 ($M = 7.02$) than non-musicians ($M = 5.58$), $p = .014$; musicians were

significantly *less* accurate than non-musicians in imitating /xi/ (*M* = 6.03) than non-musicians (*M* = 7.11), *p* = .003. This inconsistent effect of musicianship on imitation accuracy might be explained, in part, by other participant variables, such as experience with many languages other than English. None of the participants had experience with tonal languages, but increased facility with other phonetic systems might have led them to attend to the vowel and tone contrasts differently than participants with less linguistic experience. Although participants were asked about their fluency in non-native languages, a multi-dimensional assessment may be necessary to determine the extent to which these previous linguistic experiences affected performance on the imitation or the perceptual tests.

5   Conclusion

Because musicians are trained to be sensitive to F0 variations in their perception of music, we predicted a positive relationship between musicianship and Mandarin tone perception. The current study confirmed this prediction, which is also consistent with previous studies [1, 2, 3, 12]. Although the musicians were not as accurate in tone discrimination as native listeners, they were significantly better than non-musicians. The present study has also demonstrated a positive relationship between musicianship and Mandarin vowel perception, suggesting that previous musical training not only enhances people's sensitivity to F0 contours but also sensitivity to other phonetically relevant acoustic information (cf. [5]). Consistent with other studies [10, 13, 14, 15], the discrimination of Tone 2 and Tone 3 was the hardest among all the tone comparisons for both musician and non-musician non-native listeners, but also for the native listeners. One possible explanation for the difficulty in discriminating Tone 2 and Tone 3 is their acoustic similarity: F0 at the starting points of these two tones are relatively close to each other, and these two tones are also the only two tones that have a full or partial rising contour.

Regardless of their level of musicianship, non-native speakers in general found it easier to distinguish Mandarin vowels than tones. This difference in level of difficulty of the two tasks was not too surprising considering that native speakers of American English, a non-tonal language, are more familiar with distinguishing words by formant information and duration than by F0 contour. Tasks that focus on discrimination of tones are therefore less familiar and more difficult to American English speakers. Results also revealed that discrimination between the front /i/ and the central /i/ was harder than discrimination between /ü/ and /u/. The acoustical analysis showed that the allophonic variation of /i/ was less (as measured by distance in the F1-F2 "vowel space") than the variation of /u/ and /ü/. This difference in acoustical similarity may lead to the differences in discrimination difficulty for both non-native and native listeners

It may seem odd that native listeners were less accurate than non-native musicians in their discrimination of the contextually conditioned variations of /i/ and /u/, but this may have been a function of the particular instructional set given to the listeners. Indeed, some of the native participants reported after the test that they hesitated on those

front and central /i/ comparison trials because they could clearly hear the difference between the two vowels, but because they had learned those syllables share the same vowel, they decided to choose "same" as the response. This experiment might have generated different results if native participants were instructed to respond whether they heard a phonetic difference and to ignore what they may have learned from the pinyin system. On the other hand, it is noteworthy that even though native speakers *were* taught that /u/ and /ü/ are different phonemes in pinyin, their performance on this discrimination was still less accurate than that of the non-native musicians. This may due to the fact these vowels are in a complementary distribution with each other following anterior (where /ü/ occurs) and posterior (where /u/ occurs) consonants: For native listeners, then, there is no phonemic contrast. Non-native listeners would not know this, and they presumably relied on the acoustic differences in these allophones to respond "different."

Contrary to our prediction, we did not find a significant relationship between musicianship and Mandarin tone and vowel imitation. This finding is inconsistent with previous studies in which musical experience was positively related to Mandarin tone imitation [1]. Compared to the previous study, the current study used fewer participants. However, the number of syllables each participant was asked to imitate in the current study was more than those in the previous. Another factor that needs to be noted is that in the present study the imitation task always immediately followed the perception task. Given that the perception task was considerably long and intense (192 x 2 trials in 30-45 minutes), participants might have become tired or bored by the time when they were asked to perform the imitations, impairing their performance. Further studies may explore a different design and have the participants perform the perception and imitation tasks on two different days or two sessions on the same day but with sufficient break in between. However, better performance on a perceptual task may not necessarily guarantee better performance on an imitation task. Music major students with certain instrument trainings (e.g., keyboard or percussion) may be very sensitive to acoustic differences in F0 or formant frequencies, and thus have an advantage in tone and vowel discrimination, but may not be good at vocally producing these sounds. By contrast, students with musical training in vocal performance may be good at both perceiving and imitating unfamiliar phonetic contrasts (see [3]). It would be interesting to explore whether there is a relationship between specific types of musical training and success on imitation tasks by selecting a broader range of musicians with different types of musical training.

References

[1] Gottfried, T. L. "Music and language learning: Effect of musical training on learning L2 speech contrasts." In

O.-S. Bohn & M. J. Munro (Eds.), *Second-language speech learning: The role of language experience in speech perception and production.* Amsterdam: John Benjamins (2007).

[2] Alexander, J. A., Wong, P. C. M., & Bradlow, A. R. "Lexical tone perception in musicians and nonmusicians." *Interspeech 2005—Eurospeech—9th European Conference on Speech Communication and Technology*. Lisbon, 5 September (2005). [Abstract]

[3] Rø, M. H., Behne, D., & Wang, Y. "The effects of musical experience on linguistic pitch perception: A comparison of Norwegian professional singers and instrumentalists." *J. Acoust. Soc. Am., 120,* 3168 (2006). [Abstract]

[4] Thogmartin, C. "Age, individual differences in musical and verbal aptitude, and pronunciation achievement by elementary school children learning a foreign language." *IRAL, 20,* 65-72 (1982).

[5] Slevc, L. R., & Miyake, A. "Individual differences in second-language proficiency: Does musical ability matter?" *Psychological Science, 17,* 675-681 (2006).

[6] Boersma, P., & Weenink, D. *Praat: Doing phonetics by computer* (Version 4.4.24) [Computer program]. Retrieved 30 March 2007, from http://www.praat.org/ (2007).

[7] Howie, J. M. *Acoustical studies of Mandarin vowels and tones*. Cambridge Univ. Press. (1976)

[8] Tseng, C. Y. *An acoustic phonetic study of tones in Mandarin Chinese.* Taipei: Inst. of History and Phonology (1990).

[9] Chin, T. "Half-third first: On the nature of the third tone." *J. Chinese Lang. Teachers Assn. 22*(1), 87-101 (1987).

[10] Gottfried, T. L., & Suiter, T. L. "Effect of linguistic experience on the identification of Mandarin Chinese vowels and tones." *J. Phonet., 25,* 207-231 (1997).

[11] Xu, Y. "Production and perception of coarticulated tones." *J. Acoust. Soc. Am., 95,* 2240-2253 (1994).

[12] Gottfried, T. L., & Ouyang, G. Y.-H. "Training musicians and nonmusicians to discriminate Mandarin tones." *J. Acoust. Soc. Am., 120,* 3167. (2006). [Abstract]

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

[14] Shen, X. S., & Lin, M. "A perceptual study of Mandarin tones 2 and 3." *Lang. Speech, 34,* 145-156 (1991).

[15] Wang, Y., Spence, M. M., Jongman, A., & Sereno, J. A. "Training American listeners to perceive Mandarin tones." *J. Acoust. Soc. Am. 106,* 3649-3658 (1999).