COMPARISON OF VOCALISTS AND INSTRUMENTALISTS ON LEXICAL TONE PERCEPTION AND PRODUCTION TASKS

Joe Kirkham, Shuang Lu, Ratree Wayland & Edith Kaan

Department of Linguistics, University of Florida, USA
LBDL003@ufl.edu

ABSTRACT

Musical experience has been shown to positively affect non-native lexical tone perception and production. The current study divided musicians into vocalists and instrumentalists to see if either background was particularly advantageous. The results showed no difference between the musician groups on either task. Additional analyses of individual tones yielded results consistent with previous research.

Keywords: music, lexical tone, vocalists, perception, production

1. INTRODUCTION

A number of behavioral and brain-imaging studies support the view of shared neural and cognitive systems between the music and language domains. Research has shown that musical experience affects particular facets of linguistic ability [1, 5, 12], and experience with a tonal language, like Mandarin Chinese, facilitates the perception and production of musical pitch [8]. Such findings could only occur if some neural system(s) were to have a functional role in and could be influenced by both music and language experience.

In tonal languages, pitch differences are used to make semantic distinctions at the word level. In Mandarin Chinese, single CV-combinations may have different meanings when used with each of the four tones, which can be described as a high-level tone (T1), a rising tone (T2), a falling-rising tone (T3), and a falling tone (T4), respectively. Musical experience has been shown to positively affect the ability to perceive and produce lexical tones [1, 5, 11, 12]. Concerning lexical tone production, Gottfried and Ouyang [3] found that English musicians and non-musicians showed a difference in production accuracy only for T4.

Assuming a model in which speech is perceived by identifying the intended articulatory gestures [6], we would expect that vocal musicians, who have extensive neuromuscular experience with producing sounds of various pitches, should outperform instrumental musicians in both the perception and production of lexical tones. To our knowledge, this prediction has not been tested. Using non-linguistic pitch stimuli, Nikjeh, et al. [7] found that the relation between production and perception measures was different for vocalists and instrumentalists, suggesting that the specific type of musical experience differentially affects non-linguistic pitch perception and production.

Unrelated to the effect of musical experience, some studies [5, 11] observed that the Mandarin Chinese T2-T3 pair-type was more difficult to distinguish than any other tone pair-type. Gottfried and Suiter [4] attributed the T2-T3 confusion to subjects primarily focusing their attention on initial pitch heights, which are similar for T2 and T3. With regard to lexical tone production, Shen [9] found that native English speakers who were beginning learners of Chinese produced T4 the least accurately and T2 the most accurately, with T1 and T3 in between. Shen [9] attributed the difficulty of T4 production to L1 interference.

Taking previous research into account, our predictions were as follows: musicians who were native speakers of English would be more accurate at both perception and production of lexical tones than non-musicians; vocalists would perform better than instrumentalists on production of lexical tones, but not necessarily on perception of lexical tones; the T2-T3 pair-type would be the most difficult for all groups to discriminate; native English speakers would have the most difficulty producing T4; and vocalists and instrumentalists would show some differences in production of individual tones.

2. METHODS

2.1. Participants

There were 28 total participants in four distinct groups: 7 native Mandarin Chinese-speaking non-musicians, 7 native English-speaking non-musicians, 7 native English-speaking instrumentalists, and 7 native English-speaking vocalists. The instrumentalists had diverse instrumental backgrounds. The vocalists all had

significant vocal experience in addition to instrumental experience, primarily on piano. All of the native English speakers had no prior experience with tonal languages. In order to qualify for either musician group, participants must have had at least four years of formal training on their instrument and had to have been currently still playing their instrument or singing.

2.2. Stimuli

Six monosyllabic syllables [t^hi], [li], [mi], [t^ho], [lo], [mo] produced with four Mandarin tones, were used in both the perception and production tasks. These particular syllables were selected because (1) American English has these segments as well, thus, English speakers would focus their attention on the tones; (2) all stimuli were real words in Mandarin Chinese. The stimuli were recorded by two female native speakers. The target syllables were produced in the phrase "Qing shuo TARGET zhege zi" (please say TARGET this word), in order to mimic natural speech, and subsequently excised from the recording. Five native Mandarin speakers listened to and transcribed the isolated syllables; only recordings that had a 95% correct transcription rate were used as stimuli.

2.3. Procedure

Participants completed a perceptual discrimination task, and a production task. The order of the production and perception tasks was counterbalanced between subjects.

2.3.1. Perception task methods

The perceptual discrimination task consisted of 288 experimental pairs of Mandarin syllables. The inter-stimulus interval within a pair was set at 500 milliseconds. The interval between trials was set at 1500 milliseconds. The two syllables in each pair consisted of the same CV combination. Participants decided whether each pair of syllables presented to them was comprised of the same or different tones. All possible combinations of tone and CV-combination pairings were presented.

2.3.2. Production task methods

For tone production, the design was a self-paced imitation task, in which the participants first heard one syllable and then were asked to imitate it. All participants were tape recorded in a sound-attenuating booth. A total of 96 trials were presented in two blocks, with at least 30 seconds in between the blocks, in order to prevent fatigue.

3. RESULTS

3.1. Perception task results

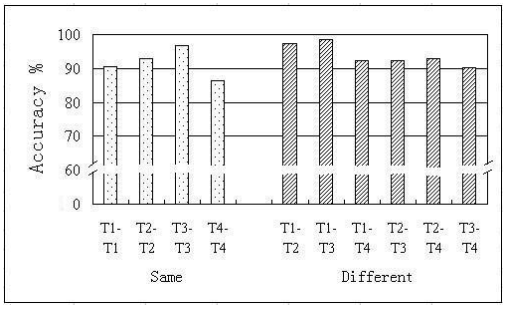
3.1.1. Effects of group

In the discrimination task, the Chinese group had a mean accuracy score of 98.6% (SD = 0.9%), the non-musicians had a mean score of 86.5% (SD = 5.6%), instrumentalists had a mean score of 91.7% (SD = 3.2%), and vocalists had a mean score of 94.4% (SD = 1.8%). These scores were submitted to a one-way ANOVA, which showed a main effect of group [$F(3,27)=15.65$, $p < .001$]. The Chinese group outperformed the non-musicians ($p < .001$), the instrumentalists ($p < .001$), and the vocalists ($p < .03$). Neither musician group outperformed the other ($p = .14$), though both the vocalists and instrumentalists outperformed the non-musicians (instrumentalists $p < .01$, vocalists $p < .001$; p-values here and below are Bonferroni-corrected when applicable).

3.1.2. Effects of tone pair-type

A repeated-measures ANOVA was conducted on the perception data with tone pair-type (collapsed over order) as the within-participant factor and group as the between-participant factor. The overall means are given in Figure 1. Analysis restricted to the "same" tone pairs showed a main effect of tone pair-type [$F(3,72)=16.69$, $p < .001$]. More errors were made on the T1-T1 pair-type than on the T3-T3 pair-type ($p < .05$). The T2-T2 pair-type was responded to less accurately than the T3-T3 pair-type ($p < .05$), and the T4-T4 pair-type had more errors than the T2-T2 and T3-T3 pair-types (both ps $< .001$). There was no significant interaction of group and tone pair-type [$F(9,72)=1.80$, $p = .11$].

Figure 1: The overall mean discrimination accuracy for the 'same' and 'different' pair-types.



For the "different" tone pairs analysis, there was a main effect of tone pair-type [$F(5,120)=8.01$, $p < .001$]. Relative to the T1-T2 and T1-T3 pairs, more errors were made on all pair-types involving the fourth tone: T1-T4 (both ps < .05), T2-T4 (both ps < .05), and T3-T4 ($p < .005$, $p < .001$, respectively). There was a significant effect of group and tone pair-type [$F(15,120)=2.31$, $p < .05$], so further analyses on the effect of tone pair-type were conducted for each group separately. The Chinese group showed no main effect of "different" tone pair-type [$F(5,30)=1.35$, $p = .29$]. The non-musician group did show a main effect [$F(5,30)=6.21$, $p < .05$], where the T3-T4 pair-type was more difficult to discriminate than the T1-T3 pair-type ($p < .05$). The instrumentalist group showed no main effect of tone pair-type [$F(5,30)=1.31$, $p = .306$]. However, the vocalists did show a main effect of tone pair-type [$F(5,30)=4.15$, $p < .05$], though no specific tone pair was statistically more difficult than any other tone pair.

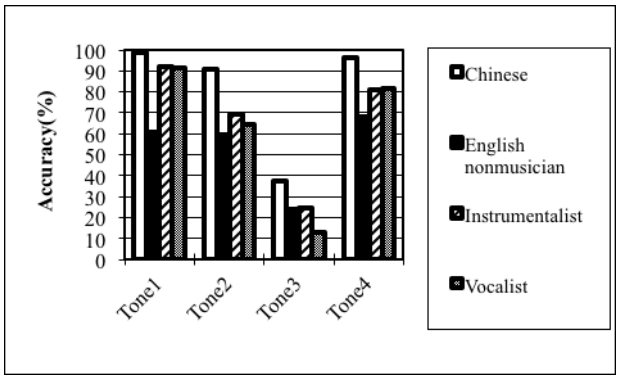
3.2. Production task results

The production responses were randomly presented to two native Mandarin judges. The percentage of correctly identified productions was used as the dependent measure. Figure 2 shows the mean percentage of correctly identified imitations of each tone. The Chinese group had a mean accuracy score of 84.4% (SD = 9.6%), the non-musicians had a mean accuracy of 50.9% (SD = 14.2%), instrumentalists had a mean accuracy of 67.9% (SD = 5.8%), and vocalists had a mean score of 65.2% (SD = 6.0%). A repeated-measures ANOVA was conducted on the production data with tone as the within-participant factor and group as the between-participant factor. There were significant main effects of group [$F(3,24)=9.60$, $p < .001$], and tone [$F(3,66)=122.49$, $p < .001$], and a significant interaction between tone and group [$F(8,66)=2.34$, $p < .05$]. The Chinese group outperformed the non-musicians ($p < .001$), the instrumentalists ($p < .001$), and the vocalists ($p < .001$). There was no significant difference between instrumentalists and vocalists, but both the vocalists and instrumentalists outperformed the non-musicians (both ps < .05).

Across groups, T1 was imitated more accurately than T2 ($p < .001$) and T3 ($p < .001$), with T3 being the least accurately imitated ($p < .001$). No significant difference was found between T1 and T4 ($p = .19$). For T1, the non-musician group was less accurate than the Chinese ($p < .001$), instrumentalists ($p < .001$), and vocalists ($p < .001$). For T2, Chinese participants outperformed non-musicians ($p < .001$), instrumentalists ($p < .05$), and vocalists ($p < .001$). For T3, the vocalist group was less accurate than the Chinese group ($p < .05$). . For T4, the Chinese outperformed the non-musicians ($p < .001$), instrumentalists ($p < .05$), and vocalists ($p < .05$). Non-musicians were less accurate relative to instrumentalists ($p < .05$) and vocalists ($p < .05$). No other differences between groups were significant.

In addition, the proportion of correctly identified imitations was positively correlated with the accuracy in the perception task ($r = 0.69$, $p < .001$).

Figure 2: Mean percentage of correctly identified productions for each tone.



4. DISCUSSION

4.1. Effects of group

The results confirmed the prediction that both musician groups outperform non-musicians on both the perception task and production task, which is highly consistent with previous research [1, 4, 11, 12]. This suggests a great degree of overlap between the processing systems associated with lexical tone processing and those associated with musical pitch processing. Furthermore, the instrumentalist and vocalist groups did not differ significantly on the lexical tone perception task. This suggests that vocal training in addition to instrumental training is not an advantage in terms of lexical tone perception ability. Unexpectedly, the vocalists and instrumentalists did not show a significant difference in tone production either, although both groups outperformed the non-musicians. This suggests that vocal expertise does

not benefit musicians in the production of lexical tones, as would be expected under some models of speech processing [6].

4.2. Effects of tone

There were several peculiarities with regard to the results of the tone pair-type analyses. In the perception task, the T2-T3 pair-type was not significantly more difficult than any other tone pair overall or for any individual group, despite past studies [5, 11] showing the T2-T3 pair-type to be the most difficult. Some unexpected differences were found. All groups considered, all of the "different" tone pair-types that involved T4 (falling) were more difficult than the T1-T2 and T1-T3 pair-types. In some previous research using identification tasks [4, 10], T4 was found to be easily confusable with T1 due to similar initial pitch heights. Wang, et al. [11] found that T2-T4 was the second most confusing tone pair, following the T2-T3 pair for native English speakers. This could be probably attributed to the English interference: T4 was similar to the falling intonation in statement sentences, while T2 resembles rising intonation in interrogative sentences. These similarities make T2 and T4 less marked for native English speakers, especially when these tones were presented in isolation rather than in sentences. The confusability of T3-T4 is presumably because of our experimental design: when the T3 stimuli were segmented from the phrase and presented in isolation, the durations, which are a crucial cue for identifying T3, were reduced. Moreover, the rising segment of T3 was partially assimilated to the following word, making T3 sound more like T4 rather than T2. However, the Chinese participants could still differentiate T3 from T4 in the perception task because of the creaky quality of T3, which is another important cue for identifying T3 [2].

Diverging from previous studies [3, 9], the present study showed that T3 was the most difficult tone for both English musicians and non-musicians to produce. The production accuracy of T3 was low even for the Chinese speakers. Again, this was due to our experimental design. Participants were more likely to perceive T3 as T4, in turn resulting in incorrect production. Even Chinese participants could not perceive/produce T3 correctly because T3 was more difficult to identify in isolation than in comparison with other tones. Moreover, T4 was produced as accurately as T1. Both instrumentalists and vocalists outperformed non-musicians for T1 and T4, but not for T2 and T3. For T1, the musicians even showed native-like accuracy.

In sum, the current study's findings support that the same cognitive mechanisms are involved in musical tone processing and lexical tone processing. Vocal training in addition to instrumental training was found not to have any effect on either lexical tone perception or production. Finally, certain acoustic features of the tones made them more difficult to perceive and procedure for the different groups.

5. REFERENCES

[1] Alexander, J.A., Wong, P.C.M., Bradlow, A.R. 2005. Lexical tone perception in musicians and non-musicians. *Proceedings of Interspeech* Lisbon, Portugal.

[2] Belotel-Grenié, A., Grenié, M. 1994. Phonation types analysis in Standard Chinese. *Proceedings of ICSLP* Yokohama, Japan, 343-346.

[3] Gottfried, T.L., Ouyang, G.Y.H. 2005. Production of Mandarin tone contrasts by musicians and non-musicians. *Journal of the Acoustical Society of America* 118(3), 2025. (Abstract)

[4] Gottfried, T.L., Suiter, T.L. 1997. Effect of linguistic experience on the identification of Mandarin Chinese vowels and tones. *Journal of Phonetics* 25, 207-231.

[5] Gottfried, T.L., Xu, Y. 2008. Effect of musical experience on Mandarin tone and vowel discrimination and imitation. *Acoustics '08 Paris*. Paris, France. *http://intellagence.eu.com/acoustics2008/acoustics2008/cd1/data/articles/003043.pdf*

[6] Liberman, A.M., Mattingly, I.G. 1989. A specialization for speech perception. *Science* 243, 489-494.

[7] Nikjeh, D.A., Lister, J.J., Frisch, S.A. 2009. The relationship between pitch discrimination and vocal production: Comparison of vocal and instrumental musicians. *Journal of the Acoustical Society of America* 125(1), 328-338.

[8] Pfordresher, P.Q., Brown, S. 2009. Enhanced production and perception of musical pitch in tone language speakers. *Attention, Perception, & Psychophysics* 71(6), 1385-1398.

[9] Shen, X.S. 1989. Toward a register approach in teaching Mandarin tones. *Journal of Chinese Language Teachers Association* 24, 27-47.

[10] So, C.K., Best, C.T. 2010. Cross-language perception of non-native tonal contrasts: effects of native phonological and phonetic influences. *Language and Speech* 53(2), 273-293.

[11] Wang, Y., Spence, M.M., Jongman, A., Sereno, J.A. 1999. Training American listeners to perceive Mandarin tones. *Journal of the Acoustic Society of America* 106(6), 3649-3658.

[12] Wong, P., Skoe, E., Russo, N.M., Dees, T., Kraus, N. 2007. Musical experience shapes human brainstem encoding of linguistic pitch patterns. *Nature Neuroscience* 10, 420-422.