Bidirectional Transfer Between Language and Musical Experience: A Study on Tonal Category Perception in Native Mandarin-Speaking Musicians

Authors: Yang Mingchuan, Li Xianzhuo, Liang Dandan, Li Xianzhuo

Date: 2024-06-02T00:00:00+00:00

Abstract

This study constructed a Mandarin tone T1-T2 continuum and employed tone categorical perception identification and discrimination tasks to investigate: (1) whether musical experience can influence the phonological abilities of Mandarin-native musicians; (2) whether the tone categorical perception patterns of Mandarin-native musicians can transfer to refined musical processing. The results revealed: (1) Mandarin-native musicians exhibit a higher degree of tone categorical perception, as reflected in metrics including identification curve steepness, category boundary width, between-category discrimination rate, and discrimination peakness; (2) Mandarin-native musicians demonstrated categorical perception patterns on both identification and discrimination curves for musical stimuli. The study demonstrates that musical experience can enhance the phonological abilities of Mandarin-native musicians across domains, and that tone categorical perception patterns can transfer across domains to musical perception. The research findings demonstrate bidirectional transfer between language and musical experience in Mandarin-native musicians, supporting the "training transfer effect" from the domain-specific perspective of tone categorical perception.

Full Text

Preamble

Self-Check Report for *Acta Psychologica Sinica*

Please complete the following items and paste them on the first page of your manuscript.

1. Please list up to three innovative contributions of this study in the form of "Research Highlights," with a total word count not exceeding 200.

Acta Psychologica Sinica aims to publish cutting-edge psychological research that is "both scientifically excellent and of particularly broad interest and significance." If your study only makes minor incremental contributions, does not attempt to open new areas of inquiry or propose unique and innovative perspectives, or particularly if it is purely an algorithm or technical work without clear psychological questions, such research has a low chance of being accepted by this journal and we recommend submitting elsewhere.

a. In the direction from music to language processing, this study is the first to demonstrate that musical pitch processing experience can comprehensively enhance the phonological abilities of native Mandarin-speaking musicians.

b. This research extends the "shared resources hypothesis" of music and language processing: at the suprasegmental level, it proves that music and language experience can transfer not only at the general acoustic level but also that domain-specific processing experience in language—tonal categorical perception—mutually influences musical pitch processing experience.

2. Have you published or submitted any articles using the same data as this study? If yes, please attach them for review. (We do not encourage authors to publish multiple articles using the same data with identical variables, nor do we support splitting a series of related studies into multiple publications.)

Answer: No.

3. For non-experimental, non-intervention studies in management, clinical, personality, and social psychology that rely solely on self-report (questionnaire methods), it is necessary to check for common method bias. What methods and measures have you used to control for or demonstrate that such bias does not affect the validity of your conclusions? (For literature on common method bias, see: http://journal.psych.ac.cn/xlkxjz/CN/abstract/abstract894.shtml) Studies based on cross-sectional data with only self-reports, measured merely in convenience samples, are easy to conduct but typically lack innovative value and have low acceptance chances at this journal.

Answer: This study is an experimental research with no questionnaire design.

4. Have you reported and analyzed effect sizes (e.g., Cohen's d for t-tests, ^{2} for ANOVA)? (Many studies mechanically report effect sizes without necessary analysis or explanation, such as whether the effect size is small, medium, or large, or what theoretical or applied significance it has.) (Searching "effect size calculator" on Google yields many convenient apps. For explanations of effect sizes in Chinese, see: http://journal.psych.ac.cn/xlkxjz/CN/abstract/abstract1150.shtml; in English, see: http://www.uccs.edu/lbecker/effect-size.html) Have you reported 95% CIs for statistical analyses? (e.g., 95% CI for differences, 95% CI for correlation/regression coefficients) For calculations and graphing of confidence intervals, see https://thenewstatistics.com/itns/esci/

Answer: Effect sizes and 95% CIs for statistical analyses have been reported.

5. Please state the planned sample size and actual sample size. If they differ, please explain why. The problem of low statistical power due to insufficient sample size is widespread in previous psychological research. We recommend explaining in the Methods section the basis for your sample size calculation and determination. Sample size should be determined based on a justified effect size and desired power, with the calculation software or program reported. For rationale and practices regarding sample size planning, see https://osf.io/5awp4/

Answer: Based on the parameters set in this study (reported in the article), the required sample size was 34. Referring to previous studies in the field, the actual sample size was set at 60.

6. For p-values, this journal requires reporting exact values (e.g., report $p = 0.034$ rather than $p < 0.05$), except for $p < 0.001$ (report as $p < 0.001$). Does your paper meet this requirement? If using Bayes factors, have you reported its sensitivity to prior distribution assumptions?

Answer: Yes, this requirement is met.

7. To ensure completeness of data reporting, if some data were excluded in statistical analyses, is this reported in the text? What were the reasons? How would the statistical results change if these data were included? How were missing data handled in statistical analyses? When using scales, were any individual items deleted? Why? How would the statistical results change if these items were included? Are there any measured items or variables not reported? Why? Please indicate where in the paper this information appears.

Answer: This study meets the requirements for data reporting completeness.

8. For experimental materials, scales, or questionnaires that have not undergone peer review and validation, are they attached at the end of the file for review? If not, please explain why. If this article is published, are you willing to share these materials with other researchers?

Answer: This study does not involve experimental materials, scales, or questionnaires that have not undergone peer review and validation.

9. This journal requires authors to provide raw data. Please choose one option below by checking √:

a) Send data to editorial office email after submission

b) Data can be obtained from the following link

c) Raw data and programs have been shared on Science Data Bank (https://www.scidb.cn/surl/xlxb)

d) If data cannot be provided, please explain the reason or provide relevant proof. (√)

10. Is your study a clinical intervention or laboratory experiment? Yes ( ) No (√)

If yes, please provide pre-registration number. If no, please explain why.
Note: For clinical interventions or laboratory experiments, pre-registration is recommended before data collection. Other experimental studies are also encouraged to pre-register. Pre-registration requires stating all research hypotheses and their support, as well as detailed experimental/intervention procedures. This journal's pre-registration website is https://os.psych.ac.cn/preregister (see "Download Center" on the journal website for instructions) or https://osf.io/ or https://aspredicted.org/. If your study is pre-registered, it will significantly increase acceptance chances. For the importance of pre-registration, see https://osf.io/5awp4/

11. If your study used human or animal subjects, was it approved by your institution's ethics committee? If yes, please send a scanned copy to the editorial office email. If no, please explain why.

Answer: Yes, ethics review documents have been sent.

12. Have you written a 400-500 word English abstract following the "English Abstract Writing Guidelines" published on the editorial office website? Has the English title and abstract been reviewed by a proficient English speaker or professionally edited by an SCI/SSCI paper editing company?

Answer: Yes, it has been polished.

13. If the first author is a student, please have the supervisor send a separate email to the editorial office (xuebao@psych.ac.cn) stating that they have read the paper and carefully reviewed it. Have you reminded your supervisor to send this email? (The editorial office will only consider the manuscript for processing after receiving the supervisor's email.)

Answer: The first author is a student, and the supervisor has been reminded to send the email.

14. Please go to the "Download Center" on the right side of the editorial office website homepage to download and complete the "Manuscript Non-Confidentiality Certificate," stamp it with the confidentiality office seal of the corresponding author's institution, and send the scanned copy to the editorial office email (xuebao@psych.ac.cn). If there is no confidentiality office seal, please use the institution's official seal. Has the email been sent?

Answer: Yes.

The Bi-Directional Transfer Between Language and Music Experience: A Study Based on the Tonal Categorical Perception of Native Mandarin-Speaking Musicians

Abstract

This study constructed a Mandarin tonal continuum from T1 to T2 and employed tonal categorical perception identification and discrimination tasks to investigate: (1) whether musical experience can influence the phonological abilities of native Mandarin-speaking musicians; and (2) whether the tonal categorical perception patterns of native Mandarin-speaking musicians can transfer to refined musical processing. The results revealed: (1) Native Mandarin-speaking musicians exhibited higher degrees of tonal categorical perception, as reflected in metrics including identification curve steepness, categorical boundary width, between-category discrimination rate, and discrimination peakedness; (2) Native Mandarin-speaking musicians demonstrated categorical perception patterns on both identification and discrimination curves for musical stimuli. The findings indicate that musical experience can comprehensively enhance the phonological abilities of native Mandarin-speaking musicians across domains, and that tonal categorical perception patterns can transfer cross-domain to musical perception. The results demonstrate bi-directional transfer between language and music experience in native Mandarin-speaking musicians, supporting the "transfer of training effects" from the perspective of domain-specific tonal categorical perception.

Keywords: language, music, tonal categorical perception, transfer of training effects
Classification Code: B842

Speech and music are two human products used for communication that share close connections and involve complex cognitive processes. Both domains utilize pitch information (Nan, 2017; Nan & Friederici, 2013; Patel, 2008). Pitch refers to the perceived "highness" or "lowness" of sound, determined by fundamental frequency (F0). In music, unique arrangements of pitch constitute melody (Bakan, 2012). In language, pitch can distinguish semantic meaning, typically achieved through changes in pitch contour patterns on monosyllables. Languages with this feature are called tonal languages (Wang, 1976).

1.1 Transfer of Training Effects and Tonal Categorical Perception

Based on the similarity of resource invocation in language and music processing, researchers have proposed the "shared resources hypothesis," suggesting that language and music processing share common neurophysiological foundations and cognitive and neurophysiological mechanisms for sound category learning (Patel, 2003, 2008; Friederici, 2017). Besson et al. (2011) further noted that language and music share neural processing mechanisms, and that experience in the two domains can transfer not only at the domain-general acoustic level but also at the domain-specific level, reflecting "transfer of training effects."

At the domain-general level, relevant studies have confirmed bi-directional transfer of pitch processing experience between language and music (Cooper & Wang, 2012; Giuliano et al., 2011; Marie et al., 2011; Zhao & Kuhl, 2015). For example, Marie et al. (2011) found that French-speaking musicians discriminated Mandarin lexical tones significantly better than non-musicians, demonstrating transfer from music to language. Giuliano et al. (2011) found that native Mandarin speakers (tonal language speakers) showed significantly improved precision in representing musical pitch compared to non-tonal language speakers, demonstrating transfer from language to music. However, beyond domain-general acoustic resources, language and music processing also require domain-specific resources.

Language processing requires phonological resources specific to the language domain. Phonology refers to the sound system of a particular language (or dialect), with its components being the minimal sound units that can distinguish meaning in that language—phonemes. Processing at the phonological level involves identifying phonemes within a language; stronger phonological ability enables more effective discrimination of minimal phonological units that distinguish meaning (Liu & Sun, 2009). Research shows that due to long-term tonal language experience, tonal language speakers establish reliable phonological representation templates in long-term phonological memory, forming categorical perception patterns for native contour tones (Francis et al., 2003; Liberman et al., 1957; Wang, 1976; Xu et al., 2006; Zhang, 2010). Unlike the continuous perception of musical pitch, in the categorical perception mode of language, continuously varying pitch in a tonal continuum is perceived as discrete, limited phonological representations. Tonal language speakers must ignore acoustic differences within phonological categories while remaining sensitive to acoustic differences across category boundaries (Liberman et al., 1957)—this is the manifestation of phonological ability.

1.2 Effects of Musical Experience on Tonal Categorical Perception

Current research has confirmed the influence of musical experience on phonological ability at the segmental level, supporting the "transfer of training effects." For example, Marie et al. (2011) found that musicians' ability to discriminate syllable differences was significantly stronger than non-musicians'; Chobert et al. (2011) found that musician children were more sensitive to subtle differences in voice onset time (VOT); Bidelman et al. (2014) found that musicians perceived vowel changes more accurately and rapidly. These studies confirm that musicians also have advantages in processing phonological differences at the segmental level, suggesting that enhanced abilities at low-level sensory processing may also influence higher-level cognitive constructs such as phonological representations in language (Chen & Peng, 2020; Besson et al., 2011).

However, at the suprasegmental level, research remains insufficient. In studies on transfer of training effects, Tang et al. (2016) found that Mandarin-speaking musicians showed larger mismatch negativity when perceiving linguistic tonal changes. However, musicians' advantages in tonal perception may stem from both acoustic and phonological aspects, and this study did not effectively separate acoustic from phonological processing. Classic tonal categorical perception experiments employ identification and discrimination tasks that require participants to identify or discriminate stimuli from different categories. Through calculations, acoustic and phonological processing levels can be measured separately, exploring the transfer of musical experience to language processing from these two dimensions (Wu et al., 2015; Xu et al., 2006). Previous studies examining the effects of musical experience in adults on native categorical perception through tonal language speakers number only three (Chen et al., 2020; Wu et al., 2015; Zhu et al., 2021), and none found advantages for musicians in phonological ability, only transfer of musical experience at the general acoustic level. These studies showed insufficient understanding of the linguistic properties of tonal continua and failed to provide appropriate experimental materials, which may explain why musicians' phonological advantages were not observed. Zhu et al. (2021) used a T2-T4 continuum with overly large differences in tonal contour and inter-stimulus intervals, likely causing ceiling effects in continuum perception. Wu et al. (2015), Zhu et al. (2021), and Chen et al. (2020) used overly short stimulus durations that could not elicit normal categorical perception patterns under natural speech perception.

1.3 Effects of Tonal Categorical Perception Patterns on Musical Perception

The categorical perception paradigm can also be used to observe the transfer of language experience to musical perception (Chang et al., 2016; Li et al., 2022). When perceiving tonal language, the categorical perception mode requires native speakers to ignore within-category acoustic differences and perceive pitch as discrete phonological representations. When perceiving music, however, it requires processing pitch information as precisely as possible, accurately perceiving note pitches through continuous perception. The requirements for pitch precision processing are opposite between the two domains. Combined with the "shared resources hypothesis," this suggests that tonal categorical perception experience may hinder refined musical pitch processing (Li et al., 2022; Patel, 2014). Currently, evidence supports this hypothesis. Chang et al. (2016) found that Mandarin-speaking non-musicians performed worse than English-speaking musicians and non-musicians on musical melody discrimination tasks, suggesting that linguistic tonal categorical perception experience may transfer to musical perception, weakening sensitivity to musical pitch changes. Li et al. (2022) found that tonal language speakers' categorical perception patterns could transfer cross-domain to music, forming similar categorical perception patterns in musical perception that hinder refined musical processing. Both studies examined Mandarin-speaking non-musicians. Mandarin-speaking musicians, however, have long-term experience in both tonal language and music domains, possessing both refined acoustic perception and phonological-level categorical perception patterns. Based on the "shared resources hypothesis" and the above evidence, there may be conflicts between processing experiences in the two domains for tonal language musicians, or more complex transfer patterns that require further investigation.

1.4 Research Questions and Hypotheses

In summary, this study aims to investigate native Mandarin-speaking musicians to examine: (1) whether musical experience can influence the phonological abilities of native Mandarin-speaking musicians, thereby modulating their tonal categorical perception patterns; and (2) how native Mandarin-speaking musicians' tonal categorical perception patterns transfer to refined musical processing.

To address these research questions, this study used a Mandarin T1-T2 continuum to construct corresponding linguistic and musical stimuli, comparing categorical perception between musicians and non-musicians for linguistic and musical stimuli. We predicted that musicians would show greater categorical perception than non-musicians, demonstrating transfer from music to language in the domain-specific processing of language. Simultaneously, we expected musicians to exhibit categorical perception patterns similar to language when perceiving musical stimuli, verifying transfer effects from language-specific categorical perception to musical processing. By examining both directions, this study comprehensively investigates bi-directional transfer effects between language and music experience in native Mandarin-speaking musicians, validating the "transfer of training effects" from the perspective of domain-specific phonological processing.

2.1 Participants

G*Power 3.1.9.7 (Faul et al., 2007) was used to estimate the required sample size. With parameters set at effect size f = 0.25, α err prob = 0.05, and power ($1 - \beta$ err prob) = 0.80, the calculated required sample size was 34. Referring to previous studies in the field (Li et al., 2022; Wu et al., 2015; Zhu et al., 2021), the experiment ultimately recruited 60 participants (N = 60). Participants were divided into a musician group (n = 30) and a non-musician group (n = 30). The musician group consisted of 16 males and 14 females with a mean age of 19.30 years (SD = 1.49). All had received at least 7 years of professional musical training, started training no later than age 12, and majored in music-related fields (following criteria from Wong et al., 2007; Wayland et al., 2010). The non-musician group consisted of 16 males and 14 females with a mean age of 22.10 years (SD = 2.14). All non-musicians had received no musical training within the past 5 years; if they had training more than 5 years ago, the total duration did not exceed 2 years (following criteria from Wong et al., 2007; Wayland et al., 2010). The Montreal Battery of Evaluation of Amusia (MBEA; Peretz et al., 2003) was used to assess amusia, and no participants in either group showed signs of amusia. All participants were right-handed, had no hearing, language, or neurological impairments, had normal IQ, and were native Mandarin or northern dialect speakers. All participants were recruited through online posters, signed informed consent before participating, and received compensation after completing the experiment.

2.2 Materials

The materials for this study were based on the syllable [i], constructing a T1-T2 continuum from [i] (T1, Mandarin high-level tone) to [i] (T2, Mandarin rising tone) with equal acoustic intervals. Compared to other continua, the T1-T2 continuum has several advantages. First, the T1-T2 continuum has a relatively small span (compared to the T2-T4 continuum), making identification and discrimination of stimuli within the continuum somewhat difficult, thus having the potential to observe enhanced categorical perception abilities in musicians. Second, changes in stimuli within the T1-T2 continuum are purely F0 changes; the quality of identification and discrimination depends directly on phonological and acoustic abilities with minimal interference from other factors. For other continua, such as T2-T3, discrimination often relates to the timing of tonal turning points (Shen & Lin, 1991) and can be interfered with by Mandarin connected speech patterns (Hao, 2012).

[Figure 1: see original paper]

The experimental materials included two stimulus types: linguistic stimuli and musical stimuli. The original syllable for continuum construction was the monosyllable [i], spoken by a male Mandarin speaker with 二级甲等 (Level 2, Grade A) proficiency without contextual cues. Using Praat software (Boersma & van Heuven, 2001), the F0 at both the onset and offset of the original monosyllable sample was first adjusted to 225 Hz as stimulus 1 (T1). Then, while keeping the offset F0 unchanged, the onset F0 was lowered to 175 Hz to form stimulus 9 (T2). Next, stimuli 1 and 9 served as boundaries, and the remaining 7 stimuli with equal F0 intervals were constructed within these boundaries, ultimately forming a T1-T2 continuum containing 9 equally spaced stimuli (as shown in Figure 1; detailed pitch parameters for the continuum are in Table 1). The duration of individual stimuli was 400 ms, approximating the natural duration of monosyllabic high-level and rising tones in Mandarin (Guo, 1993). Using Praat's pitch extraction function, the pitch contours of the 9 linguistic stimuli in the continuum were extracted and synthesized into sine tones to simulate musical stimuli, creating a musical stimulus continuum (following Xu et al., 2006). These sine waves had the same pitch contours, amplitude, and duration as their corresponding linguistic stimuli, with the main acoustic difference being spectral content. For example, stimulus 9, perceived as Mandarin rising tone in language, approximates a glissando from F3 to A3 in music (using scientific pitch notation). Figure 2 [Figure 2: see original paper] presents the waveform and spectrogram for stimulus 9 in both the linguistic and musical continua.

[Figure 2: see original paper]

2.3 Procedure

The experiment was presented to participants via E-prime 3.0 on a computer. Participants sat in a quiet, enclosed environment, wore headphones, and responded by pressing keys according to instructions. The headphones played binaural sounds with equal volume in both channels, set at 80 dB.

In the identification task, a fixation cross appeared at the center of the screen for 500 ms, followed by a linguistic or musical stimulus. After stimulus presentation, a response screen with instructions appeared, and participants judged whether the stimulus was Mandarin T1 or T2 by pressing a key. Participants had up to 2500 ms to respond; the response screen disappeared after key press, followed by a 500 ms blank screen before the next trial began (as shown in Figure 3 [Figure 3: see original paper]). Musical and linguistic stimuli were presented in separate blocks. In each block, the 9 stimuli in the continuum were each repeated 10 times, with 90 stimuli per block. Stimulus order within blocks was random. The presentation order of linguistic and musical stimulus blocks was counterbalanced, with a brief rest period between blocks.

[Figure 3: see original paper]

In the discrimination task, two stimuli from the continuum were presented as a pair with a 500 ms inter-stimulus interval, making the total duration for a complete linguistic or musical stimulus pair 1300 ms. After a 500 ms fixation cross, the stimulus pair was played. Following presentation, a response screen appeared, and participants judged whether the two stimuli in the pair were the same or different by pressing a key. Participants had up to 2500 ms to respond; the response screen disappeared after key press, followed by a 500 ms blank screen before the next trial (as shown in Figure 4 [Figure 4: see original paper]). Musical and linguistic stimuli were presented in separate blocks. In each block, "different" stimulus pairs (pairs with two different stimuli) were second-order differences, with both forward and backward order combinations. For example, stimuli 1 and 3 combined to form pair 1-3 (stimulus 1 followed by stimulus 3); reversing the order formed pair 3-1. Following this rule, each block contained fourteen "different" stimulus pairs: 1-3, 3-1, 2-4, 4-2, 3-5, 5-3, 4-6, 6-4, 5-7, 7-5, 6-8, 8-6, 7-9, and 9-7. Additionally, each block contained nine "same" stimulus pairs: 1-1, 2-2, 3-3, 4-4, 5-5, 6-6, 7-7, 8-8, and 9-9. Thus, each block contained 23 types of "same" and "different" stimulus pairs, with each pair presented ten times, for a total of 230 stimulus pairs per block in random order. The presentation order of linguistic and musical stimulus blocks was counterbalanced, with a brief rest period between blocks.

[Figure 4: see original paper]

2.4 Data Analysis

The experiment calculated six basic metrics related to categorical perception: identification curve steepness, categorical boundary position, categorical boundary width, between-category discrimination rate, within-category discrimination rate, and discrimination peakedness (Jiang et al., 2012; Peng et al., 2010; Wu et al., 2015; Xu et al., 2006).

Identification task results were fitted using the logistic regression equation proposed by Xu et al. (2006) to obtain identification curves (Figure 5 [Figure 5: see original paper]):

$$P(T1) = 1 / (1 + e^{(-b(x - a))})$$

where P(T1) represents the probability of participants judging a stimulus as T1 in the identification task, and x represents stimulus number. We used the coefficient b (i.e., slope) in the regression equation as an indicator of identification curve steepness; larger absolute values of b indicate steeper identification curves and higher degrees of categorization. Additionally, setting $P(T1) = 0.5$ yields the corresponding x value as the categorical boundary position, giving $x = a$. Following Peng et al.'s (2010) algorithm, we calculated the absolute x values when P(T1) took values at 25% and 75% to obtain categorical boundary width $= 2 \ln 3 / |b|$; narrower categorical boundary width indicates stronger categorical perception.

Next, discrimination curves were plotted based on discrimination accuracy for seven second-order stimulus units in the discrimination task (Figure 7 [Figure 7: see original paper]). To calculate accuracy in the discrimination task, we divided the 230 stimulus pairs in each block into seven second-order stimulus units. Each A-B stimulus unit (where A and B are stimuli spanning two steps) contained four stimulus pairs: A-B, B-A, A-A, and B-B. For example, the 1-3 stimulus unit contained pairs 1-3, 3-1, 1-1, and 3-3. Each stimulus unit contained 40 stimulus pairs. Adjacent stimulus units shared AA or BB pairs (e.g., the 1-3 and 3-5 units shared pair 3-3). Discrimination accuracy for each stimulus unit was calculated using the following formula (Xu et al., 2006):

$$\text{Accuracy} = P(\text{“same”}|\text{same}) \times P(\text{same}) + P(\text{“different”}|\text{different}) \times P(\text{different})$$

where P("same"|same) is the probability of correctly judging "same" pairs (number of correct judgments for AA and BB pairs divided by total AA and BB pairs), and P("different"|different) is the probability of correctly judging "different" pairs (number of correct judgments for AB and BA pairs divided by total AB and BA pairs). P(same) and P(different) represent the probabilities of "same" (AA and BB) and "different" (AB and BA) stimulus pairs within each unit, respectively. Using this method, seven data points were obtained for each participant in both linguistic and musical blocks.

Combining the previously calculated categorical boundary position value a, we calculated the average accuracy of stimulus units crossing the categorical boundary as the between-category discrimination rate and the average accuracy of remaining units as the within-category discrimination rate (Jiang et al., 2012). For example, if the categorical boundary $a = 5.5$, the average discrimination accuracy of the 4-6 and 5-7 stimulus units was recorded as the between-category discrimination rate, while the average accuracy of all other units was recorded as the within-category discrimination rate. Considering that within-category discrimination rate reflects acoustic-level processing ability while between-category discrimination rate depends on both acoustic and phonological processing abilities, we further purified the metrics to extract an indicator of phonological ability by subtracting within-category from between-category discrimination rate to obtain discrimination peakedness, which reflects phonological ability (Wu et al., 2015; Xu et al., 2006).

3.1 Identification Task

Figure 5 presents identification curves for each group. Table 2 presents descriptive statistics for metrics in the identification task. Figure 6 [Figure 6: see original paper] presents bar charts of these metrics.

[Figure 5: see original paper]

Both linguistic and musical stimulus identification curves (Figure 5) showed that the difference in identification rates between two adjacent stimuli across category boundaries was much larger than between adjacent stimuli on the same side of the boundary, indicating that both groups exhibited typical categorical perception patterns for linguistic and musical stimuli (Liberman et al., 1957; Xu et al., 2006). A 2 (group: musicians vs. non-musicians) × 2 (stimulus type: musical vs. linguistic) repeated measures ANOVA was conducted on identification task metrics (descriptive statistics in Table 2).

Results showed significant main effects of group on identification curve steepness, $F(1, 58) = 22.51$, $p < 0.001$, $p^2 = 0.28$, 95% CI = [-0.47, -0.19], with musicians showing significantly smaller identification curve slopes (i.e., steeper curves) than non-musicians. The stimulus type main effect was also significant, $F(1, 58) = 14.69$, $p < 0.001$, $p^2 = 0.20$, 95% CI = [-0.32, -0.10], with musical stimuli showing significantly smaller slopes (steeper curves) than linguistic stimuli. For categorical boundary position, the group main effect was significant, $F(1, 58) = 5.36$, $p = 0.024$, $p^2 = 0.09$, 95% CI = [-0.62, -0.05], with musicians' categorical boundaries significantly closer to the T1 end than non-musicians'. The stimulus type main effect was significant, $F(1, 58) = 66.69$, $p < 0.001$, $p^2 = 0.54$, 95% CI = [-1.53, -0.93], with musical stimuli boundaries significantly closer to the T1 end than linguistic stimuli. For categorical boundary width, the group main effect was significant, $F(1, 58) = 7.75$, $p = 0.007$, $p^2 = 0.12$, 95% CI = [-0.61, -0.10], with musicians showing significantly narrower boundaries than non-musicians. The stimulus type main effect was not significant, $F(1, 58) = 1.49$, $p = 0.227$, $p^2 = 0.03$, 95% CI = [-0.33, 0.08] (Figure 6 [Figure 6: see original paper]). No interaction effects were found for any metrics ($ps > 0.05$).

3.2 Discrimination Task

Figure 7 presents discrimination curves for each group. Table 3 presents descriptive statistics for metrics in the discrimination task. Figure 6 [Figure 6: see original paper] presents bar charts of these metrics.

[Figure 7: see original paper]

Discrimination curves (Figure 7) also showed that both groups exhibited categorical perception for music and language, specifically demonstrated by higher discrimination rates for cross-boundary stimulus units than for within-category units (Liberman et al., 1957; Xu et al., 2006). A 2 (group: musicians vs. non-musicians) × 2 (stimulus type: musical vs. linguistic) repeated measures ANOVA was conducted on discrimination task metrics (descriptive statistics in Table 3).

Results showed significant main effects of group on between-category discrimination rate, $F(1, 58) = 14.02$, $p < 0.001$, $p^2 = 0.20$, 95% CI = [0.03, 0.10], with musicians showing significantly higher between-category discrimination rates than non-musicians. For within-category discrimination rate, the group main effect was significant, $F(1, 58) = 4.95$, $p = 0.030$, $p^2 = 0.08$, 95% CI = [0.00, 0.05], with musicians showing significantly higher within-category discrimination rates than non-musicians. For discrimination peakedness, the group main effect was significant, $F(1, 58) = 4.95$, $p = 0.041$, $p^2 = 0.07$, 95% CI = [0.00, 0.08], with musicians showing significantly greater discrimination peakedness than non-musicians (Figure 6 [Figure 6: see original paper]). No stimulus type effects or interaction effects were significant (ps > 0.05).

4.1 Musical Experience Enhances Phonological Ability in Native Mandarin-Speaking Musicians

Steeper identification curves and narrower categorical boundaries indicate enhanced categorical perception (Jiang et al., 2012; Li et al., 2022; Peng et al., 2010; Xu et al., 2006; Zhang, 2010). The present results show that musicians exhibited higher degrees of tonal categorical perception than non-musicians, indicating stronger phonological abilities. In discrimination curves, within-category discrimination rate generally reflects acoustic information processing ability, while between-category discrimination rate depends on both acoustic and phonological processing abilities (Wu et al., 2015; Xi et al., 2010; Yu et al., 2014; Zhang et al., 2011; Zhu et al., 2021). Discrimination peakedness, obtained by subtracting within-category from between-category discrimination rate, further extracts the phonological ability indicator (Wu et al., 2015; Xu et al., 2006). Results showed that musicians' between-category discrimination rate, within-category discrimination rate, and discrimination peakedness were all significantly higher than non-musicians', indicating that musicians possess stronger phonological and acoustic abilities than non-musicians.

In the direction of transfer from musical experience to language processing, previous research has focused on transfer effects in domain-general acoustic processing, demonstrating that linguistic and musical pitch processing share general acoustic processing resources (Cooper & Wang, 2012; Marie et al., 2011; Marques et al., 2007). The present study approaches from the domain-specific phonological processing level and, for the first time, finds that musical experience can cross-domain enhance suprasegmental phonological processing ability. This discovery validates Besson et al.'s (2011) proposed transfer of training effects, showing that experience in music and language domains can transfer at the domain-specific level. Specifically, enhanced domain-general abilities from musical experience can also influence domain-specific language abilities. Yao and Chen's (2020) longitudinal study found that preschool children showed significantly reduced categorical boundary width after one year of musical training, confirming that musical training can improve children's tonal categorical perception. The present study demonstrates that this benefit also exists in adults and, through further optimized materials, reveals more comprehensive promotion effects across more metrics. We offer two plausible explanations: First, at the brain region level, primary auditory cortex and nearby temporal plane and temporal pole regions are responsible for early phoneme identification in speech perception (Diesch et al., 1996; Levy & Wilson, 2020; Poeppel et al., 1997). Such processing systems are not language-specific but must process various types of auditory input (Friederici, 2017). We speculate that musical training can enhance processing ability in the temporal plane region responsible for separating and matching time-frequency patterns, thereby improving efficiency in analyzing sound signal time-frequency features (Griffiths & Warren, 2002), and transfer this benefit to language phoneme perception, enhancing musicians' phonological abilities. Bidelman et al. (2014) found that musicians' primary auditory processing structures could more effectively extract spectral cues related to speech information (e.g., F0, F1), ultimately making sound signal representations clearer, providing evidence for our speculation. Second, musicians' enhanced categorical perception ability may relate to strengthened sound-meaning mapping ability. Bidelman et al. (2014) noted that musicians have more robust and selective internalized representations of speech signals in the auditory pathway; these strengthened representations provide more reliable phonological templates for decision processes controlling speech recognition, enabling extracted acoustic signals to access phonological information stored in the lexicon more accurately and rapidly (Friederici, 2017).

However, this finding contradicts conclusions from Wu et al. (2015), Zhu et al. (2021), and Chen et al. (2020), none of which found effects of musical experience on phonological ability. We speculate this may relate to continuum selection, as tonal continuum design affects categorical perception patterns: different continuum designs produce different categorical perception results (Wu & Wang, 2018). The main problems with continua selected in the above three studies include: First, the continuum types made cross-category perception too easy. Zhu et al. (2021) used a T2-T4 continuum that actually spanned two category boundaries (essentially a T2-T1-T4 continuum), with overly large internal tonal contour differences involving perception differences between contour and level tones, likely causing ceiling effects in continuum identification and discrimination. Second, the F0 span of continua was too large. Zhu et al. (2021) used a continuum spanning over 100 Hz, making discrimination and identification too coarse; overly large spans may mask musicians' phonological perception advantages, also producing ceiling effects. Third, stimulus duration was too short. According to Guo (1993), the minimum necessary durations for maintaining naturalness of high-level and rising tones are 210 ms and 230 ms, respectively; durations below these limits cause tones to lose their phonemic characteristics. Wu et al. (2015), Zhu et al. (2021), and Chen et al. (2020) used 200 ms stimulus durations; overly short perception times may have recruited more acoustic processing resources and could not elicit normal categorical perception patterns under natural speech perception. Our T1-T2 continuum used smaller pitch steps, making identification and discrimination more difficult and requiring more refined processing. The 400 ms stimulus duration approximates natural monosyllabic tone duration in Mandarin, allowing musicians' advantages in phonological processing to emerge significantly. How different continuum selections and parameter settings affect experimental results requires further discussion in future research.

4.2 Transfer of Native Mandarin-Speaking Musicians' Tonal Categorical Perception Patterns to Musical Perception

Both identification and discrimination curves showed that, for musicians and non-musicians alike, identification curves for linguistic and musical stimuli showed high steepness near categorical boundaries, and discrimination curves showed clear discrimination peaks. These results indicate that both groups exhibited categorical perception patterns in musical stimuli sharing the same F0 range as linguistic stimuli. Notably, categorical perception is a domain-specific ability in language, suggesting that native phonological perception patterns stored in long-term memory can transfer cross-domain to musical perception.

According to the OPERA-e hypothesis (Patel, 2014), music requires higher precision than language in domain-shared acoustic processing. For pitch processing, continuous perception in music requires higher processing precision to identify continuously changing pitch cues, representing fine-grained processing; in language, tonal changes in pitch are less refined than in music, and tonal perception does not require discriminating within-category pitch differences, representing coarse-grained processing. Previous research on Mandarin-speaking non-musicians found that coarse-grained linguistic pitch categorical processing transfers to fine-grained musical pitch processing, thereby hindering refined musical pitch processing (Chang et al., 2016; Li et al., 2022). Mandarin-speaking musicians possess both long-term tonal categorical memory and more refined acoustic processing, suggesting that while linguistic categorical perception may hinder refined musical pitch processing, their musical experience simultaneously demands fine-grained processing of musical pitch. Our results show that when linguistic categorical perception transfers to musical pitch processing, musicians' fine-grained pitch processing does not weaken their categorical perception degree, indicating that native Mandarin-speaking musicians exhibit transfer from language experience to musical perception, and that domain-specific categorical processing experience in language can influence general acoustic processing. Previous research by Wu et al. (2015) also found that musicians activated language-like categorical perception patterns when processing nonlinguistic stimuli. These results demonstrate that the influence of linguistic pitch categorical processing experience on acoustic processing is robust, and this domain-specific linguistic perception pattern can transfer to auditory processing in other domains. Related neuroscience research found that brain regions processing phonological information are located in deeper structures of the auditory processing system, such as the superior temporal sulcus and middle temporal gyrus (Zhang et al., 2011). Based on this, we infer that transfer of linguistic categorical perception patterns to musical processing may occur because native phonological processing experience begins at birth, with phonological memory existing long-term and stably in deep cognitive resources and influencing lower-level auditory input (Wu et al., 2015; Xu et al., 2006).

Combined with the discussion in Section 4.1, our results show that fine-grained pitch processing and categorical perception patterns coexist in native Mandarin-speaking musicians and are shared cross-domain (language, music). At the level of categorical perception—a domain-specific linguistic processing ability—mutual transfer between linguistic and musical experience occurs simultaneously in Mandarin-speaking musicians, confirming that brain plasticity changes brought by language and music processing experience are bi-directional (Bidelman et al., 2013).

5 Conclusion

Using the Mandarin tonal categorical perception paradigm and examining categorical perception performance, this study investigated bi-directional transfer effects between music and language experience in native Mandarin-speaking musicians. The findings reveal that musical experience can transfer to domain-specific linguistic perception, enhancing phonological processing ability, and that linguistic phonological experience can transfer to musical perception, causing native Mandarin-speaking musicians to exhibit categorical patterns in musical perception. The results demonstrate bi-directional transfer between language and music experience in native Mandarin-speaking musicians, supporting the "transfer of training effects" (Besson et al., 2011) from the perspective of domain-specific tonal categorical perception. The study supports the "shared resources hypothesis" of language and music processing (Giuliano et al., 2011; Patel, 2003, 2008; Xu et al., 2006).

References

Bakan, M. B. (2012). *World music: traditions and transformations* (2nd ed.). New York: McGraw-Hill.

Besson, M., Chobert, J., & Marie, C. (2011). Transfer of training between music and speech: common processing, attention, and memory. *Frontiers in Psychology*, 2, 94.

Bidelman, G. M., Hutka, S., & Moreno, S. (2013). Tone language speakers and musicians share enhanced perceptual and cognitive abilities for musical pitch: evidence for bidirectionality between the domains of language and music. *PLoS ONE*, 8(4), e60676.

Bidelman, G. M., Weiss, M. W., Moreno, S., & Alain, C. (2014). Coordinated plasticity in brainstem and auditory cortex contributes to enhanced categorical speech perception in musicians. *The European Journal of Neuroscience*, 40(4), 2662–2673.

Boersma, P., & van Heuven, V. (2001). Speak and unSpeak with Praat. *Glot International*, 5(9–10), 341–347.

Chang, D., Hedberg, N., & Wang, Y. (2016). Effects of musical and linguistic experience on categorization of lexical and melodic tones. *Journal of the Acoustical Society of America*, 139(5), 2432–2447.

Chen, F., & Peng, G. (2020). Reduced Sensitivity to Between-Category Information but Preserved Categorical Perception of Lexical Tones in Tone Language Speakers With Congenital Amusia. *Frontiers in psychology*, 11,

Chen, S., Zhu, Y., Wayland, R., & Yang, Y. (2020). How musical experience affects tone perception efficiency by musicians of tonal and non-tonal speakers?. *PloS one*, 15(5), e0232514.

Chobert, J., Marie, C., François, C., Schön, D., & Besson, M. (2011). Enhanced passive and active processing of syllables in musician children. *Journal of cognitive neuroscience*, 23(12), 3874–3887.

Cooper, A., & Wang, Y. (2012). The influence of linguistic and musical experience on Cantonese word learning. *The Journal of the Acoustical Society of America*, 131(6), 4756–4769.

Diesch, E., Eulitz, C., Hampson, S., & Ross, B. (1996). The neurotopography of vowels as mirrored by evoked magnetic field measurements. *Brain and Language*, 53(2), 143–168.

Faul, F., Erdfelder, E., Lang, A. G., & Buchner, A. (2007). G*Power 3: a flexible statistical power analysis program for the social, behavioral, and biomedical sciences.* Behavior research methods*, 39(2), 175–191.

Francis, A. L., Ciocca, V., & Ng, B. K. C. (2003). On the (non)categorical perception of lexical tones. *Perception and Psychophysics*, 65(7), 1029–1044.

Friederici, A. D. (2017). *Language in our brain: the origins of a uniquely human capacity.* Cambridge, MA: The MIT Press.

Giuliano, R. J., Pfordresher, P. Q., Stanley, E. M., Narayana, S., & Wicha, N. Y. (2011). Native experience with a tone language enhances pitch discrimination and the timing of neural responses to pitch change. *Frontiers in psychology*, 2, 146.

Greenwood, D. D. (1961). Critical bandwidth and the frequency coordinates of the basilar membrane. *Journal of the Acoustical Society of America*, 33(10), 1344–1356.

Griffiths, T. D., & Warren, J. D. (2002). The planum temporale as a computational hub. *Trends in neurosciences*, 25(7), 348–353.

Guo, J. (1993). *An Explication and Investigation of Chinese Tones and Intonation.* Beijing: Beijing Language Institute Press. [郭锦桴. (1993). 汉语声调语调阐要与探索. 北京: 北京语言学院出版社.]

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.

Jiang, C., Hamm, J. P., Lim, V. K., Kirk, I. J., & Yang, Y. (2012). Impaired categorical perception of lexical tones in Mandarin-speaking congenital amusics. *Memory and Cognition*, 40(7), 1109–1121.

Levy, D. F., & Wilson, S. M. (2020). Categorical encoding of vowels in primary auditory cortex. *Cerebral Cortex*, 30(2), 618–627.

Li, X., Xiao, R., & Liang, D. (2022). The cross-domain influence of tonal categorical perception and tonal complexity on musical pitch perception. *Acta Psychologica Sinica*, 54, 1021–1030. [李贤卓, 肖容, 梁丹丹. (2022). 声调范畴感知和声调复杂度对音乐音高感知的跨领域影响. 心理学报, 54, 1021–1030.]

Liu, F., & Sun, W. (2009). *General theory of linguistics.* Beijing: Beijing Language and Culture University Press. [刘富华, 孙炜. (2009). 语言学通论. 北京: 北京语言大学出版社.]

Liberman, A. M., Harris, K. S., Hoffman, H. S., & Griffith, B. C. (1957). The discrimination of speech sounds within and across phoneme boundaries. *Journal of experimental psychology*, 54(5), 358–368.

Marie, C., Delogu, F., Lampis, G., Belardinelli, M. O., & Besson, M. (2011). Influence of musical expertise on segmental and tonal processing in mandarin Chinese. *Journal of Cognitive Neuroscience*, 23(10), 2701-2715.

Marques, C., Moreno, S., Castro, S. L., & Besson, M. (2007). Musicians detect pitch violation in a foreign language better than nonmusicians: behavioral and electrophysiological evidence. *Journal of Cognitive Neuroscience*, 19(9), 1453–1463.

Nan, Y. (2017). The facilitation effect of music learning on speech processing. *Advances in Psychological Science*, 25(11), 1844-1853. [南云. (2017). 音乐学习对语言加工的促进作用. 心理科学进展, 25(11), 1844-1853.]

Nan, Y., & Friederici, A. D. (2013). Differential roles of right temporal cortex and Broca's area in pitch processing: evidence from music and Mandarin. *Hum Brain Mapp*, 34(9), 2045-2054.

Patel A. D. (2003). Language, music, syntax and the brain. *Nature neuroscience*, 6(7), 674–681.

Patel, A. D. (2008). *Music, language, and the brain.* New York: Oxford University Press.

Patel, A. D. (2014). Can nonlinguistic musical training change the way the brain processes speech? The expanded opera hypothesis. *Hearing Research*, 308, 98–108.

Peretz, I., Champod, A. S., & Hyde, K. (2003). Varieties of musical disorders——The Montreal battery of evaluation of amusia. *Annals of the New York Academy of Sciences*, 999, 58—75.

Peng, G., Zheng, H. Y., Gong, T., Yang, R. X., Kong, J. P., & Wang, W. S. Y. (2010). The influence of language experience on categorical perception of pitch contours. *Journal of Phonetics*, 38(4), 616–624.

Poeppel, D., Phillips, C., Yellin, E., Rowley, H. A., Roberts, T. P. L., & Marantz, A. (1997). Processing of vowels in supratemporal auditory cortex. *Neuroscience Letters*, 221, 145–148.

Shen, X. S., & Lin, M. (1991). A Perceptual study of Mandarin tones 2 and 3. *Language and Speech*, 34(2), 145–

Tang, W., Xiong, W., Zhang, Y. X., Dong, Q., & Nan, Y. (2016). Musical experience facilitates lexical tone processing among mandarin speakers: behavioral and neural evidence. *Neuropsychologia*, 91, 247–253.

Wang, W. S. (1976). Language change. *Annals of the New York Academy of Sciences*, 280(1), 61–72.

Wayland, R., Herrera, E., & Kaan, E. (2010). Effects of musical experience and training on pitch contour perception. *Journal of Phonetics*, 38(4), 654–662.

Wong, P. C. M., Skoe, E., Russo, N. M., Dees, T., & Kraus, N. (2007). Musical experience shapes human brainstem encoding of linguistic pitch patterns. *Nature Neuroscience*, 10(4), 420–422.

Wu, H., Ma, X., Zhang, L., Liu, Y., Zhang, Y., & Shu, H. (2015). Musical experience modulates categorical perception of lexical tones in native Chinese speakers. *Frontiers in Psychology*, 6, 436.

Wu, Q., & Wang, Y. (2018). Categorical perception of lexical tone and the neural mechanisms. *Advances in Psychological Science*, 26(01), 62–71. [吴倩, 王韫佳. (2018). 声调的范畴知觉及其神经机制. 心理科学进展, 26(01), 62–71.]

Xi, J., Zhang, L., Shu, H., Zhang, Y., & Li, P. (2010). Categorical perception of lexical tones in Chinese revealed by mismatch negativity. *Neuroscience*, 170(1), 223–231.

Xu, Y., Gandour, J. T., & Francis, A. L. (2006). Effects of language experience and stimulus complexity on the categorical perception of pitch direction.

Journal of the Acoustical Society of America, 120(2), 1063–1074.

Yao, Y., & Chen, X. (2020). The effects of music training on categorical perception of Mandarin tones in 4- to 5-year-old children. *Acta Psychologica Sinica*, 52(04), 456–468. [姚尧, & 陈晓湘. (2020). 音乐训练对 4~5 岁幼儿普通话声调范畴感知能力的影响. 心理学报, 52(04), 456–468.]

Yu, K., Wang, R., Li, L., & Li, P. (2014). Processing of acoustic and phonological information of lexical tones in Mandarin Chinese revealed by mismatch negativity. *Frontiers in Human Neuroscience*, 8, 729.

Zhang, L. (2010). On the influence of native language on the categorical perception of Chinese tones. *TCSOL Studies*, 38(2), 15–20. [张林军. (2010). 母语经验对留学生汉语声调范畴化知觉的影响. 华文教学与研究, 38(02), 15–20.]

Zhang, L., Xi, J., Xu, G., Shu, H., Wang, X., & Li, P. (2011). Cortical dynamics of acoustic and phonological processing in speech perception. *PLoS ONE*, 6(6), e20963.

Zhao, T. C., & Kuhl, P. K. (2015). Effect of musical experience on learning lexical tone categories. *Journal of the Acoustical Society of America*, 137(3), 1452–1463.

Zhu, J., Chen, X., & Yang, Y. (2021). Effects of amateur musical experience on categorical perception of lexical tones by native Chinese adults: an ERP study. *Frontiers in Psychology*, 12: 611189.

Note: Figure translations are in progress. See original paper for figures.