華語文教學研究 18.4: 67-110, 2021

The Effects of Pitch Singing Training on Recognizing and Enunciating Chinese Tones

Danqing ZHOU
Department of Education
Shanghai Normal University

Abstract

It is a challenge for non-tonal language speakers to enunciate the four Chinese tones accurately. In this study, I examined whether musical training benefits tonal language learning, evaluated whether the pitch singing training influence recognizing and enunciating the four Chinese tones, and compared the effects of a pitch singing training method with the traditional audio-lingual method. Sixty American college students who had not taken a Chinese course and who did not speak and write Chinese, were recruited in this study. The participants were divided evenly and randomly into two groups: the pitch singing training group and the traditional audio-lingual training group. They participated in a pretest/training/posttest program over the course of eight training sessions. Results revealed the pitch singing training method had a greater effect on both recognizing and enunciating the four Chinese tones than did the traditional audio-lingual method. The pitch singing training method can be used as an alternative and effective way to improve non-tonal language speakers' recognition and enunciation of the four Chinese tones, because it provides a sensory experience to the learners, builds associations to the concept of abstract Chinese tones, and elaborates on learners' memory of the Chinese tones to store it for long-term retention. Consequently, Chinese language learners can make conversations more efficient and be more motivated in learning Chinese.

Keywords: enunciation mandarin tones, learning tones with music, pitch singing exercise, recognition mandarin tone

1. Introduction

Chinese is a tonal language. Non-tonal language learners often face the challenge of recognizing and enunciating Chinese tones because the intonations are different from their native languages; for example, the pitch range of the Chinese tones is wider than English pronunciation (Repp and Lin 1990; Chen 1997). Also, some pitch ranges do not appear in non-tonal languages (White 1981). Moreover, non-tonal English speakers have difficulty pronouncing the Chinese tones accurately because these tones are a new concept and are unfamiliar to non-tonal speakers (Yang 2015). Lastly, because of the wide pitch range of the Chinese tones, some Chinese language learners feel awkward and strange when enunciating the tones with high and low pitches (Tsai 2011).

Singing is a common activity that is applied in language learning due to its various benefits, such as fostering motivation (Schön et al. 2008), strengthening memorization (Mora 2000), and promoting conversations (Wray and Perkins 2000). In accordance with different pitch characteristics of Chinese tones, to pronounce them accurately one almost needs to sing these tones. An adapted pitch singing exercise may help Chinese language learners to recognize and enunciate Chinese tones more accurately. The purposes of this study were to examine whether musical training benefits tonal language learning, to evaluate the effect of pitch singing training on recognizing and enunciating the four Chinese tones, and to compare the effects of a pitch singing training method with the traditional non-music teaching method.

2. Literature Review

2.1 The Characteristics of the Four Chinese Tones

In Mandarin Chinese (the standard Chinese pronunciation), there are four Chinese tones, which could be categorized into five-pitch levels, with the first tone (T1) being high-level pitch, the second tone(T2) high-rising pitch, the third

tone (T3) low-dipping pitch, and the fourth tone (T4) high-falling pitch (Chao 1948). The four Chinese tones are presented in Figure 1 (Chao 1948:85).

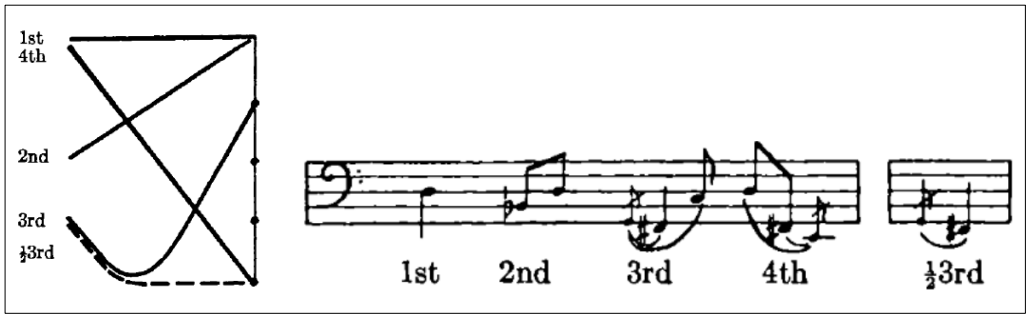

Figure 1: The Pitch Level of the Four Chinese Tones

To make the pitches of the four Chinese tones easier to understand and read, I converted the four tones' pitches into the first five notes (1-2-3-4-5/do-re-mi-fa-sol) of a major scale. It is based on Chao's (1948) description of Chinese tones' pitches in the bass clef. For phonetics, Chinese use a distinctive tool, "*pīn yīn*", to represent the sounds and tones of Chinese characters and to facilitate the pronunciation of words. The pitch level of the four Chinese tones in the form of "*pīn yīn*" is presented in Figure 2.

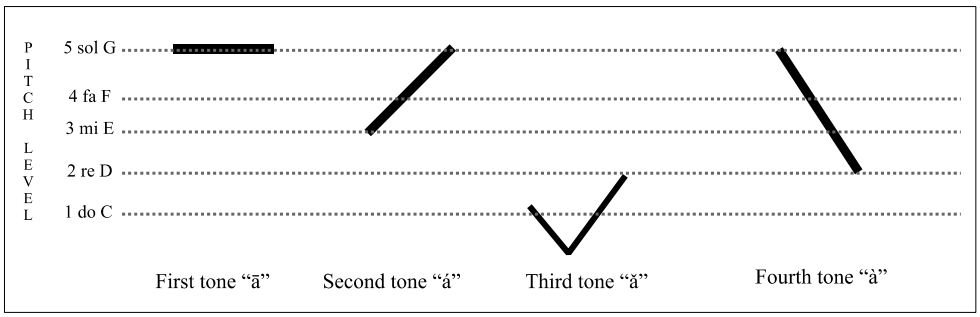

Figure 2: Pitch Level of the Four Chinese Tones in the Form of "*pīn yīn*"

Each of the four Chinese tones has its own characteristics. The duration of the tones differs from each other, with the fourth tone having the shortest duration and the third tone having the longest duration (Lin 1965). In addition, the amplitudes of the four tones are not the same either. The third tone has the lowest amplitude and the fourth one has the highest (Chuang et al. 1972).

2.2 Challenges of Recognizing and Enunciating the Four Chinese Tones

It is challenging for English speakers to recognize and enunciate the four Chinese tones accurately (Kiriloff 1969; Bluhme and Burr 1971; Shen 1989). One of the reasons is that the phonetic system of English and Mandarin is different from each other in pitch patterns, distributions, and functions (White 1981). In Chinese oral production, one syllable can be related to different tone marks to produce different morphemes, for example, [wǒ] means I, while [wō] means nest.

Contrastingly, in English, each syllable corresponds to either one morpheme, or several syllables are combined to equal one morpheme. For example, “happy” is a two-syllable word [hap-py] but has one morpheme. One morpheme carries one meaning. No matter how the tone of “happy” is changed, the meaning of the word “happy” does not change to others. Tones in English do not change a word’s meaning, while Chinese tones do.

Another reason for the difficulty in acquiring accurate pronunciation of a second language is the interaction of a learner’s native language and the second language system (Selinker 1972; Flege 1995; Best and McRoberts 2003; Brown 2014). Often English speakers use their native language background to perceive and produce Chinese. In a study investigating the different perceptions of Chinese tones between tonal and non-tonal language speakers, Gandour (1984) indicated that Mandarin speakers pay more attention to the tone’s contour, while English speakers focus more on the tone’s height and less on the direction. In addition, in a study comparing English speakers’ responses to English and Cantonese language stimuli, functional magnetic resonance imaging (fMRI) revealed comprehension-based differences between the two languages (Morrison et al. 2003). Wang and Honig (2010) pointed out that one of the difficulties that children experience in learning to read and write Chinese is lacking awareness of the meanings that Chinese tones change. Therefore, Chinese learners need to set up an accurate perception of Chinese tones and avoid the influence of their native language background.

2.3 Music and Language, and Their Shared Learning Mechanisms

Music and language are two domains, which both depend on a mental framework of learned sound categories (Handel 1989). Positron emission tomography (PET) brain studies show many musical pitch perception tasks activate the right hemisphere circuits of the brain, whereas many linguistic phonemic tasks show a greater activation in the left hemisphere (Zatorre et al. 1996; Stewart et al. 2006).

Some studies indicate music and language share a few mechanisms, while other studies state each domain has its individual system (Patel 2008). Strong and Plitnik (1992) state that tonal languages require engaging the right hemisphere of the brain for processing, while non-tonal languages mostly use the left hemisphere. Interestingly, Hsieh et al. (2001) found when a pitch was used in cuing lexical difference in a tonal language, the left hemisphere was activated. On the other hand, if a pitch is not used in recognizing lexical difference, then the right hemisphere is engaged (Zatorre and Gandour 2008). A tone perception comparing study of Chinese syllables also indicated that when pitch perception or pitch differences are processed linguistically (as a cue for differentiating lexical difference), the left hemisphere is engaged (Klein et al. 2001). On the other hand, when the speakers do not use pitch differences as a linguistic function, the right hemisphere region is activated (Klein et al. 2001). These studies suggested that non-tonal language speakers use the right hemisphere region where music is usually processed to differentiate tones (pitch levels) of tonal languages when they do not have much knowledge about languages.

Moreover, in a study examining adult English native speakers' cortical activation during a tone identification task (Wang et al. 2003), increased activation in the left posterior superior temporal gyrus (STG) and adjacent regions and engagement of the right inferior frontal gyrus (IFG) were found before and after a two-week training procedure. This study further confirmed that for non-tonal language speakers to acquire a new tonal language, they need to activate both the right and left hemisphere regions. The right hemisphere region for identifying lexical tones of a tonal language is used specifically at the beginning of learning a tonal language for non-tonal language students. Later, the

left hemisphere region will be engaged more as non-tonal language students have learned more of a tonal language. Nan and Friederici (2013) also found “shared neural resources engaged in pitch processing for music and tone language” (Nan and Friederici 2013:2053). These studies confirmed the positive transfer effects between the music pitch processing domain and the linguistic tone processing domain (Schön et al. 2004; Magne et al. 2006; Wong et al. 2007; Moreno et al. 2009). One domain is strengthened, and the benefits could be transferred or shared to the other domain.

2.4 Singing Training Effects in Foreign Language Learning

Singing is often used in language classes. Studies have reported various benefits that singing has on foreign language learning (Shen 2009; Rahbar and Khodabakhsh 2013). Singing has a great effect on learning motivation and acquiring language structures (Schön et al. 2008). Singing songs can also help students to retrieve the lexical patterns stored in long-term musical memory easily, with mental rehearsal, memorization, and conversation (Mora 2000).

In learning English, studies have reported the effectiveness of using English lyrics to improve the listening comprehension and pronunciation ability of both adult English learners (Rahbar and Khodabakhsh 2013) and young students (Ghanbari and Hashemian 2014). Furthermore, another study presented the positive effects of using children’s songs in learning English vocabulary and improving pronunciation (Shehadeh and Farrah 2016). Thus, these results suggest that singing songs brings more positive effects on learning English than without using songs. Also, these findings shed light on the benefits of using songs or singing in foreign language teaching, especially in speaking and listening.

2.5 Pitch Singing in Teaching Chinese Tones

Previous studies support that pitch singing training may have a great effect on tonal language learning, especially on tone differentiation and pronunciation (Shen 1989; Schön et al. 2004; Shi 2018). However, after searching through the current and available Chinese tones teaching methods in the United States (Ning

2001; Luo 2010), I did not find one that uses pitch singing training methods to help learners in recognizing and enunciating the four Chinese tones.

There are plenty of studies that focus on Chinese tonal phonology and tonal error analysis, but not much research focuses on practical teaching methods of Chinese tone acquisition (Shi 2018). Some teachers have tried to apply musical elements in teaching the Chinese tones, but the methods were shown as not practical in classroom teaching (Shi 2018).

Shi (2018) conducted a study for examining the effectiveness of using high, middle, and low sounds in teaching the four Chinese tones. Participants were K-12 Chinese language students and adult learners. Their Chinese proficiency level was beginning and intermediate. Shi (2018) used the "opera voice" for the high pitch and tones in the high register, a normal speaking voice for the middle pitch and tones in the middle register, and a hesitate "ugh" sound for the low sound and the tones in the low register. In addition, a visual tone map (similar to Figure 3, presented later) was used to present the four Chinese tones. There was a total of five groups that received the same training. The training session was 30 to 50 minutes long according to students' age. Thirty minutes is for elementary school students, 45 minutes for middle school students, and 50 minutes for high school students and up. The training session includes two parts, an oral and a written production practice. In the oral part, students read Chinese sentences that written with "*pin yin*", tone marks, and Chinese characters. The researcher corrected students' tone pronunciation by referring to opera pitch, middle register, and low sound. In the written part, students were asked to mark the tones on the Chinese words that they had just learned in the oral part. Two recordings (before and after the training) of students' reading Chinese sentences were collected and compared. Results indicated that all participants learned the four Chinese tones effectively. However, Shi (2018) did not compare the results between common teaching methods and a music related-teaching method. Thus, the results cannot suggest that pitch singing training methods perform better than other methods.

Therefore, in this study, I sought to investigate the effect of pitch singing training on two important language outputs: recognizing and enunciating the four Chinese tones, and to compare the effects of a pitch singing training method with

the traditional audio-lingual method. The audio-lingual teaching method is based on sound imitation. Students learn pronunciations from imitating teachers' speaking or sound from audio recordings. Specific research questions were:

Q1: Do pitch singing training and a traditional audio-lingual teaching method differ in their effect on the accuracy of Chinese tones' recognition for non-tonal Chinese language adult learners?

Q2: Do pitch singing training and a traditional audio-lingual teaching method differ in their effect on the accuracy of Chinese tones' enunciation for non-tonal Chinese language adult learners?

3. Method

3.1 Overview

In order to examine the possible benefits of musical training on learning the four Chinese tones and to compare the different effects of pitch-singing training and a traditional audio-lingual teaching method on non-tonal language speakers' Chinese tone recognition and enunciation skills, an experimental study was conducted at a medium-sized American public doctoral research university in the southwestern United States with the approval of the university institutional review board.

3.2 Participants and Setting

Before recruiting participants, I calculated the necessary sample size for ensuring statistical power of 0.8 (Cohen 1988) by using G*Power 3.1.9.2 software (Faul et al. 2007). A repeated-measures (ANOVA), a significance level of α = 0.05, and an effect size of 0.25 were used in calculating the sample size. The minimum sample size was 34. In total, 66 college students (both undergraduate and graduate) who neither spoke Chinese nor had studied Chinese before this study were recruited. Their majors were not in the music area. None of the participant self-reported being a musician before taking the study. Two participants were over 50 years old; the rest of the participants' ages ranged from 18 to 26.

Among the 66 participants, there were three Japanese students and the rest were American students. Some American participants were bilingual speakers who could speak Spanish in addition to English, but none of the participants had studied Chinese before participating in this study.

American Council on the Teaching of Foreign Languages (ACTFL 2019) divided language learning into four skills: speaking, listening, reading, and writing. Each skill is divided into 10 levels from novice-low to superior (ACTFL 2012). Based on field studies of student proficiency levels in immersion programs and traditional language classes, it takes at least one to two school years for new students to arrive at novice-low to novice mid-level (Bai et al. 2016). Since none of the 66 participants had learned Chinese before, they had minimum knowledge about Chinese and its pronunciation; therefore, they were all qualified to participate in this study. All participants self-reported having normal hearing, speech, and the ability for learning languages.

The participants were randomly assigned to one of two training methods: one was the pitch singing training method, and the other one was the traditional audio-lingual method. I divided participants evenly into each training method.

3.3 Procedure

The procedure consisted of an introduction session, a pretest, eight group-training sessions, and a posttest. The procedure was modeled after the study of Wong and Perrachione (2007). They investigated whether non-tonal language (English) speakers can learn the use of pitch patterns (suprasegmentals) for lexical identification. In their study, they used English pseudowords with pitch patterns to present Mandarin Tones 1, 2, and 4 as the training stimuli. Each word was heard four times. According to their study, the number of training sessions for achieving 95% success in using pitch patterns to identify English words learned ranged from 7.22 sessions of a group of successful subjects to 9.38 sessions of a group of less successful subjects. In their study, each group received three to four sessions per week and no more than one session per day. Thus, for the current study, the median number of eight sessions was chosen. The eight total training sessions were given over three to four weeks with each week having

two to three sessions.

To complete the whole process, the participants and I met nine times in total. In the first meeting, all the participants were given an introductory session (15-20 min) of the Chinese language and pronunciations of "*pīn yīn*" since none of the subjects had any previous knowledge of the Chinese language. The introduction of "*pīn yīn*" and tones was given in a normal speaking mode with visual aids for all the participants, and no pitch singing was involved. The visual aid, the tone marks chart is presented in Figure 3.

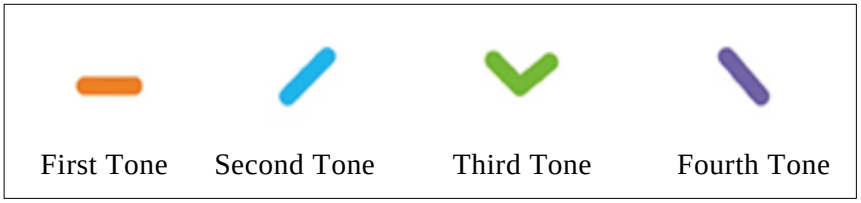

Figure 3: Tone Marks Chart

After the introductory session, the students took the pretest immediately. The pretest was designed to test each of the participants' accuracy of recognizing and enunciating the four Chinese tones without training. The introductory session and pretest were completed in one session. Afterward, all of the participants took eight training sessions with the method to which they were assigned and completed in three to four weeks' period of time. A posttest was scheduled at the end of the eighth training session, which was the last meeting. The posttest was designed to reexamine the participants' accuracy of recognizing and enunciating the four Chinese tones after they experienced different training methods.

3.4 Materials

3.4.1 Pitch Singing Materials

I used Praat software to record and analyze the four Chinese tones that I enunciated. The results were congruent with Chao's (1948) definition of different pitch levels of Chinese tones. In addition, the pitch intervals for the four Chinese tones are very close to the ones that Chao (1948) pitched in the bass clef in Figure 1. To indicate the music notes' pitch and intervals, other than writing on the music staff, one common way is to use seven letters to present the pitch and

use numbers to indicate which octave the notes belong to (see Figure 4 music notes range).

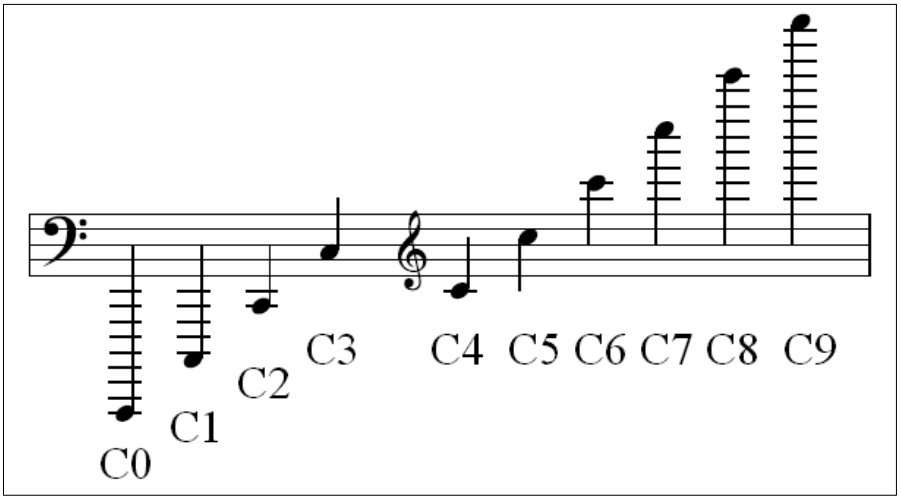

Figure 4: Music Notes Range

Smaller numbers correspond to lower octaves, while larger numbers correspond to higher octaves. The middle C octave uses number 4 for each note in that range, such as C4 (216.63 Herz), D4 (293.66 Herz), E4 (329.63 Herz), F4 (349.23 Herz), G4 (392 Herz), A4 (440 Herz), B4 (493.88 Herz).

Comparing the results of my four tones' pitch intervals to Chao's version, the differences are in tone 2, mine is one semitone smaller than Chao's tone. In tone 3, the rising part of Chao's interval is 3 semitones wider than mine. In tone 4, Chao's interval is 6 semitones wider than mine. Using small pitch intervals for the four Chinese tones is consistent with Patel's (2008) finding that both musical melodies and speech melodies are dominated by small intervals. Thus, using small pitch intervals to present the four Chinese tones is appropriate. Furthermore, the music intervals for each tone can be transposed to higher or lower pitches to accommodate participants' singing ranges because the numbers of the semitone in each interval are the same. This is similar to singing one song in different keys. Two versions of music intervals for the four Chinese tones are presented in Table 1.

Table 1: Music Intervals for the Four Chinese Tones

Music Intervals	Chao (1948)	My Version
Tone 1	a sustained note D3, as a unison interval, 0 semitone	a sustained note G4, as a unison interval, 0 semitone
Tone 2	notes Bb2 to D3, as a major third interval, 4 semitones	E4 to G4, as a minor third interval, 3 semitones
Tone 3	G2 to F#2, then go up to C3 as a minor second interval (1 semitone) plus a tritone (6 semitones),	C4 to B3, then go up to D4 as a minor second interval (1 semitone) plus a minor third interval (3 semitones)
Tone 4	D3 to F#2, as a minor sixth interval, 8 semitones	G4 to F4, as a major second interval, 2 semitones

In Moore's (1991) study, the participants' (who are teachers) singing ranges are from E3 to D5. Moore (1991) also states that the female adults' mean singing ranges are from E3 to F5, while male adults' mean singing ranges are almost an octave lower than females' ranges, from F2 to F4. Since the participants in my study are all adults, using notes from B2 to G4 (B2-C3-D3-E3-F3-G3 for males, and B3-C4-D4-E4-F4-G4 for females) is a comfortable singing range for them. This is congruent with Chao's (1948) indication of singing four Chinese tones with the music notes, which females should sing an octave higher than the pitches he wrote in the bass clef.

Because of the difference between males' and females' singing ranges, I interpreted the four Chinese tones in the small pitch intervals and transposed the intervals depending on the singing range of the participants. To make transposing the intervals convenient and clear, I composed the four Chinese tones' intervals in a major scale system. The musical intervals that were used for the four Chinese tones for female and male participants are presented in Figure 5.

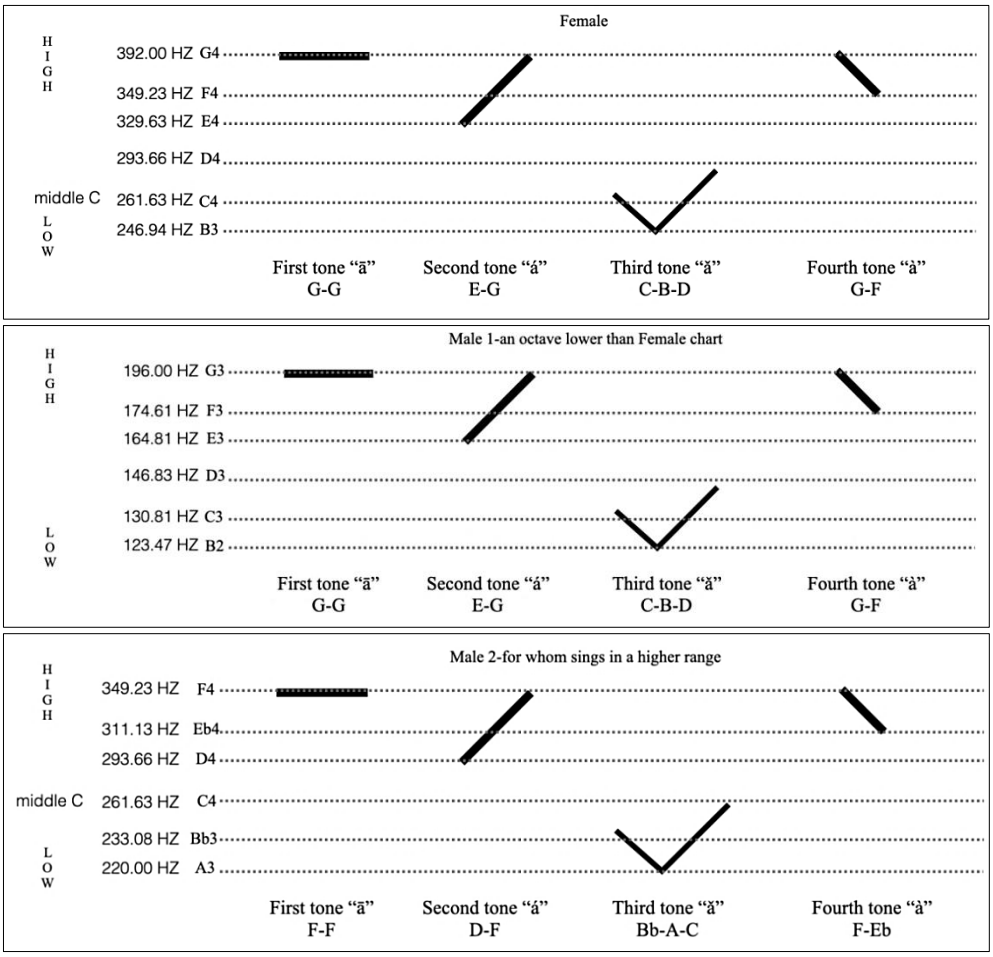

Figure 5: Music Intervals for the Four Tones for Female and Male Learners

3.4.2 Material for an Introduction Session

In the introductory session, I first explained the purpose of learning the "*pīn yīn*" system. Then, the components of *pīn yīn* were presented in a normal speaking mode, in which each "*pīn yīn*" is composed of three parts: consonant(s), vowel(s), and a tone mark. The tone marks chart was then presented in front of the groups. I pronounced the four tones' intonation with each vowel along with the tone marks chart (see Figure 3, presented previously). In addition, I used a finger to trace the tone mark shape when enunciating the tone. The students imitated the sounds that I enunciated and learned four tone marks' intonation on

each vowel, for example, [ā], [á], [ǎ], [à]. Furthermore, I taught the pronunciation of 21 consonants, then, demonstrated some combinations of vowels and consonants to the participants. The material of the introduction session is presented below.

Introduction Material:

1. Vowel

   ā á ǎ à, ō ó ǒ ò, ē é ě è, ī í ǐ ì, ū ú ǔ ù, ǖ ǘ ǚ ǜ

2. Consonant

   b p m f, d t n l, g k h, j q x, zh, ch, sh, z, c, s, r

3. Some Combinations of Vowel and Consonant

   Example: *bāng, hóng, děng, lìng, pái, fèi, mù, yü, jiā, zhōng, guó, cǐ, rì*

3.4.3 Pretest and Posttest Materials

The pretest contained three parts, recognition, enunciation, and singing test. The posttest contained two parts, recognition and enunciation. In both pretest and posttest, the structure and the form of recognition and enunciation section were the same.

The material for the recognition section of both the pretest and the posttest was recorded by me, a native Chinese speaker, on a Mac laptop with the recording software Praat (version 6.0.14). Before giving the tests to the participants, the recordings for the sections were assessed by one male and one female native Chinese speaker for validation. For both listeners, identification accuracy was 100% for the recognition sections. The pretest and posttest are presented in Appendix A.

The recognition section asked participants to recognize the tone of 20 "*pīn yīn*". Each "*pīn yīn*" was presented three times. All participants were asked to select the appropriate tone number 1 to 4, or write four kinds of tone marks in the answer choices according to the recording. The pretest recognition part was delivered in groups.

The posttest was in the same form as the pretest, but half of the tested Chinese "*pīn yīn*" were from the training sessions and the other half were new to the participants. The new "*pīn yīn*" were added to avoid possible test familiarity

effects and to examine if the participants can apply what they have learned in the training sessions to the new "*pīn yīn*". There were 20 "*pīn yīn*" in the recognition section with ten learned "*pīn yīn*" from the training sessions and ten new "*pīn yīn*". The same portion was used in the enunciation part. The recognition materials of pretest and posttest are presented in Table 2. The enunciation portion of the pretest and posttest are presented in Table 3.

Table 2: Pretest and Posttest Recognition (Listening) Materials

Pretest					
First tone	*tā*	*bō*	*hē*	*xī*	*wū*
Second tone	*bá*	*tú*	*rén*	*lí*	*fó*
Third tone	*mǎ*	*wǒ*	*kě*	*nǐ*	*shǔ*
Fourth tone	*là*	*mò*	*wù*	*zì*	*dòu*
Posttest					
First tone	*mā*	*zhuān*	*bō*	*kē*	*tī*
Second tone	*míng*	*guó*	*rén*	*dú*	*fá*
Third tone	*hěn*	*wǒ*	*tǔ*	*jǐ*	*gǔ*
Fourth tone	*jiào*	*lù*	*è*	*shì*	*dòng*

Table 3: Pretest and Posttest Enunciation Materials

Pretest				
1. *mā*	2. *bó*	3. *kě*	4. *dì*	5. *tú*
6. *dǐ*	7. *tù*	8. *má*	9. *kē*	10. *bǒ*
11. *kè*	12. *dī*	13. *bò*	14. *tǔ*	15. *mǎ*
16. *bō*	17. *mà*	18. *dí*	19. *tū*	20. *ké*
Posttest				
1. *bā*	2. *fàn*	3. *bú*	4. *wǒ*	5. *sì*
6. *hěn*	7. *zhōng*	8. *mén*	9. *fù*	10. *yī*
11. *mò*	12. *hǎo*	13. *jí*	14. *wǔ*	15. *kū*
16. *bò*	17. *lǐ*	18. *péng*	19. *shàng*	20. *tū*

A singing accuracy test (the third part of the pretest) was given to all the participants for examining the possible influences that the singing-in-tune skill may apply to the performances of Chinese tones. In the singing test, participants were asked to sing the first verse of "Mary Had a Little Lamb." The reasons for choosing this song were because this song is known by many people and because it includes all of the intervals that are used in the four Chinese tones.

The second enunciation portion and the third singing part were completed individually, and each participant's enunciation of the tones and the singing part was recorded on their computers or cell phones and then sent to the researcher by email for later analysis.

3.4.4 Eight Training Sessions and Materials

After the introduction session and the pretest, all participants took eight training sessions with either taking pitch singing training or audio-lingual training. In each session, 8 to 15 new Chinese words were taught to the participants with each tone appearing in at least one "*pīn yīn*", such as tone 1, [tā] means he or she; tone 2, [rén] means people; tone 3, [wǒ] means I; and tone 4, [shì] means verb to be. The pronunciation and the tone of each word were practiced four times. Thus, both groups received the same amount of training time on each "*pīn yīn*". All the training sessions were given by me. Thus, the teaching demonstrations and approaches for different groups were consistent. The study was examining the effects of two training methods on the four tones. To encourage all the participants to complete all the sessions and to increase their learning interests, I presented this study as a mini Chinese introducing course. I added a conversation and cultural facts demonstration part at the end of each session. Thus, each training session is divided into three parts: learning tones and pronunciation (main focus), reading a dialogue (dialogue was composed by the newly learned words from each session), and presenting sentence structure or cultural facts.

For the pitch singing group (experimental group), I firstly delivered the pronunciation of the four tones in pitch interval singing and played the pitch intervals on the keyboard application on my cellphone at the same time. Then, the

participants repeated the tones in pitch interval singing. I transposed the intervals to lower or higher pitches to accommodate participants' singing range. Females sang the pitch intervals an octave higher than male participants. The singing materials are presented in Figure 5, the pitch intervals for the four Chinese tones. The specific procedure is presented below.

Pitch Singing Method of Each Session:

1. The instructor uses pitches to sing each character's tone from the vocabulary section (twice for each character). For example, the first tone uses the note G, the second tone uses notes E to G, the third tone uses notes C to B (lower) to D, and the fourth tone uses notes G to F.
   For example, teaching *pīn yīn* [nǐ-you]. The instructor sings the vowel [ǐ] with the intervals (C-B-D) first, then sings the whole *pīn yīn* [nǐ] with the intervals (C-B-D).
   The students repeat the sound of each character as the instructor pronounces it (twice).
2. The instructor pronounces each character in normal speaking mode, without the pitch tones (twice for each character).
   The students repeat each character as to how the instructor pronounces it (twice for each character).
3. The instructor calls each student to enunciate the pronunciation of each character independently in pitch singing mode first then in normal speaking mode. Then, the instructor corrects students' pronunciation individually.
4. The instructor reads the dialogue sentence by sentence with the normal speaking mode. The students repeat what the instructor says.
5. Before finishing one session, the instructor explains some grammar points and cultural facts that are related to the dialogue.

For the other "audio-lingual" group, I taught the pronunciation of the four tones in a normal speaking mode without any singing music intervals involved. All the participants in the control group first listened to my pronunciation, then repeated the "*pīn yīn*" in a normal speaking mode. The specific procedure is

presented below.

Audio-Lingual Method (Non-Music Training) of Each Session[1]:

1. The instructor pronounces each character from one session (twice for each character):
   For example, teaching *pīn yīn*: [nǐ-you].
   The instructor says the vowel [ǐ] with the normal speaking mode first, then says the whole *pīn yīn* [nǐ] with the normal speaking mode.
   The students repeat each character as the instructor pronounces it (twice for each character).
2. The instructor pronounces each character again in a normal speaking mode (twice for each character).
   The students repeat each character as the instructor pronounces it (twice for each character).
3. The instructor calls each student to pronounce each character in normal speaking mode twice independently. Then, the instructor corrects students' pronunciation individually.

4 and 5 steps are the same as the pitch singing method.

Complete lesson materials of training sessions for control and experimental groups are listed in Appendix B.

3.5 Scoring

The pretest and posttest scores were calculated as follows: There were a total of 20 questions in each recognition and enunciation section with each question worth 1 point for a correct answer. The lowest possible score could be 0 points (if none of the questions were answered correctly), and the highest possible score could be 20 points (if all the questions were answered correctly).

The section of enunciating the four Chinese tones in the pretest and posttests was constructed by asking participants to enunciate the "*pīn yīn*"

[1] The differences between the two methods are from the first to the third step. The fourth and fifth steps are the same for both groups.

tones that they had read. The scores were calculated only based on the accuracy of tones' enunciation, and the segmental errors were excluded. For example, if the participant pronounced [qí] as [kí], the point was still given if the tone was enunciated as the second tone. The participants' enunciation recordings were graded by two native (one female and one male) Chinese speakers from the university to minimize personal bias. The recordings of all the participants were numbered when given to the Chinese graders. Therefore, the two Chinese native speakers could not recognize the participants, which might have affected their judgments on the enunciation performance.

In the pretest, participants were asked to sing the first verse of "Mary Had a Little Lamb." Then, their recorded singing was graded based on the following points: (1) the stepwise notes are in tune or not (E-D, C-D, 2 points), (2) the repeated notes are sung the same or not (E, E, E and D, D, D, 2 points), (3) the skipped notes are in tune or not (E-G, 1 point). Each point was calculated as 0 points for not singing in tune, 1 point for singing slightly off, and 2 points for accurate singing. Thus, the total singing score was ranged from zero to ten. The participants' singing recordings were graded by two professional musicians from the School of Music at the university. The recordings of all the participants' singing were numbered when given to the musicians. Therefore, the two musicians could not recognize the participants, which (again) might have affected their judgments on the singing skills.

4. Results

In this study, data were collected from only those participants who completed all training sessions. Six participants' data were not used because they did not complete all sessions. Also, three of the six participants were Japanese. The Japanese language has pitch accents. To avoid bias in analyzing the results, the data of three Japanese participants were not used as well. Thus, at the end of the training sessions, sixty valid participants' data were used.

There were two sections where the scores were evaluated by independent graders, the singing-in-tune test and the enunciation of the four Chinese tones. To check reliability, Cronbach's alpha was used to examine the consistency

between the two graders' scores for each section. The coefficient result between the two independent graders of the singing-in-tune test indicated very strong agreement (Cronbach's $\alpha = .97$). The coefficient results between the two independent graders of enunciating the four Chinese tones were also strong (Cronbach's $\alpha = .96$ for the pretest and Cronbach's $\alpha = .95$ for the posttest). Therefore, I used the average score between the two graders for each section to produce a single score for each of the 60 participants in the subsequent data analysis. For the singing-in-tune test, the mean score for the pitch singing group was 7.83 (SD = 1.68), and the mean score for the audio-lingual group was 8.23 (SD = 1.68).

The two dependent variables in this study were recognizing scores and enunciating scores. The two different training methods were two levels of one independent variable, including the pitch singing training group participants (n = 30) and the audio-lingual instruction participants (n = 30). I took the difference between the pretest and posttest scores (gain score) on each of those and analyzed the differences. Means (and Standard Deviations) of pretest and posttest scores for recognition and enunciation are presented in Table 4. The descriptive data (minimum, median, maximum, mean, and range) of gain scores for recognition and enunciation is presented in Table 5. The distribution of gain scores of the pitch singing group and audio-lingual group are presented in Figure 6.

Table 4: Means (and Standard Deviations) of Pretest and Posttest Scores for Recognition and Enunciation

	Pretest	Posttest	Gain
Pitch Singing Group			
Recognition	13.30 *(4.62)*	15.26 *(4.63)*	1.97 *(3.41)*
Enunciation	12.93 *(3.48)*	16.27 *(2.42)*	3.33 *(2.78)*
Audio-lingual Group			
Recognition	15.23 *(3.78)*	15.30 *(3.94)*	0.07 *(2.20)*
Enunciation	14.20 *(2.72)*	15.53 *(2.61)*	1.33 *(2.44)*

Table 5: The Descriptive Data of Gain Scores for Recognition and Enunciation

	Gain S. Recognition Audio	Gain S. Recognition Pitch	Gain S. Enunciation Audio	Gain S. Enunciation Pitch
Minimum	-4	-5	-3	0
Median	0	1	1	3
Maximum	7	10	8	9
Mean	0.07	1.97	1.33	3.33
Range	11	15	11	9

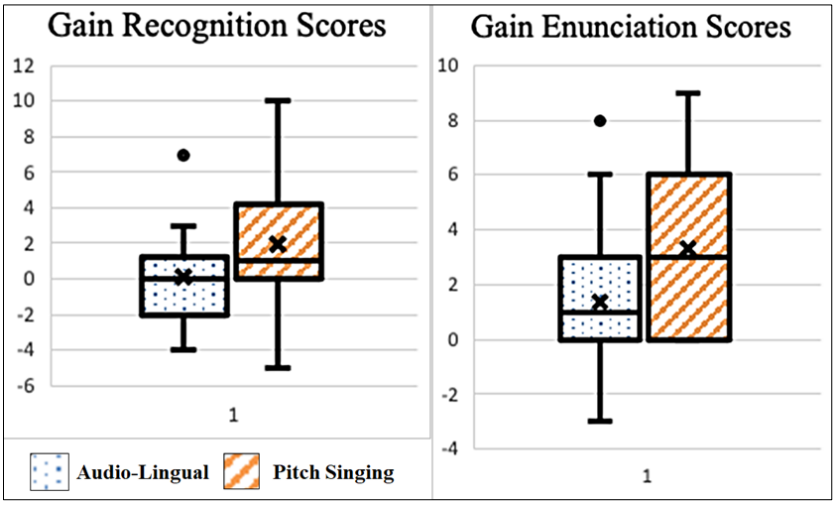

Figure 6: Distribution of Gain Scores for Recognition and Enunciation

In order to account for possible influences of preexisting musical abilities of non-tonal language speakers on the accuracy of recognition and enunciation of the four Chinese tones, I first conducted a multivariate analysis of covariance (MANCOVA), with the singing-in-tune scores used as a covariate.

According to the study of Rankin and Tracy (1965:226), "Residual gain does not require that initial and final measures be expressed in equal interval scales. Residual gain always correlates zero with initial status by definition. Initial differences on the pre-test do not contaminate the measure of gain." Thus, I used the residuals of the gaining scores (posttest – pretest = gain score) of recognizing and enunciating Chinese tones in the MANCOVA test to avoid possible effects

from initial differences of the two groups on the pre-test. The descriptive data of the residuals on recognition and enunciation are reported in Table 6.

Table 6: Descriptive Data of the Residuals of Gain Scores on Recognition and Enunciation

	N	Mean	*SD*	Median	Min	Max	Kurtosis
Recognition	60	0.00	2.84	-0.07	-6.97	8.03	1.120
Enunciation	60	0.00	2.60	-0.33	-4.33	6.67	-0.337

Before conducting the MANCOVA test, the multivariate normality assumption was checked. The residuals of gain scores on recognition and enunciation sections have bivariate normal distribution: Mardia Skewness test's *p*-value was 0.320 and Mardia Kurtosis test's *p*-value was 0.791, which both *p*-value are greater than α = 0.05. Therefore, the data met the multivariate normality assumption. Then, I used Box's *M* test (Box *M* = 5.881, *p* = .129) to test the assumption of homogeneity of covariance matrices, where the *p*-value is greater than α = 0.05. Thus, the data met the assumption of homogeneity of covariance matrices.

The results of MANCOVA test indicated that there was a statistically significant difference between the pitch singing and audio-lingual training method groups on the combined dependent variables after controlling for the singing-in-tune skill, $F(2, 56) = 6.158$, $p = .004$, Wilks' $\Lambda = .820$, partial $\eta^2 = .180$. Furthermore, the covariate *p*-value for the recognition part was 0.864 (*df* = 1, *MS* = .249, *F* = .030), and the covariate *p*-value for the enunciation part was 0.488 (*df* = 1, *MS* = 3.374, *F* = .488), both of which were greater than α = .05. Thus, the covariate, the singing-in-tune score, had no influences on the recognizing and the enunciating scores in the presence of the factor.

To understand whether there were differences between the pitch singing group and the audio-lingual group without singing-in-tune as a covariate, a subsequent ANCOVA test was conducted. Results revealed significant differences between groups for both recognition, $F(1, 57) = 6.144$, $p = .016$, partial $\eta^2 = .097$, and for enunciation, $F(1, 57) = 8.192$, $p = .006$, partial $\eta^2 = .126$. The pitch-singing group gained an average of 1.97 points (*SD* = 3.409) on

their recognition test, compared to .07 points (*SD* = 2.196) for the audio-lingual method group. For the enunciation test, the pitch-singing group gained an average of 3.33 points (*SD* = 2.783), and the audio-lingual method group gained an average of 1.33 points (*SD* = 2.440). These data are depicted in Figure 7.

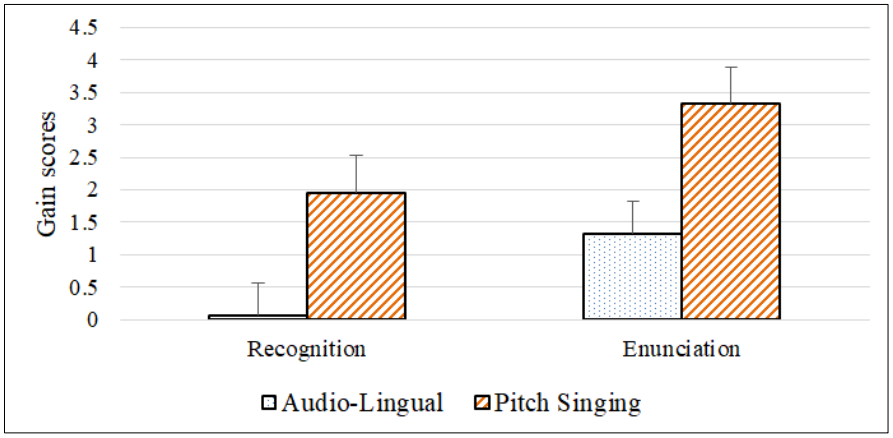

Figure 7: Gain Scores for Recognition and Enunciation Tests for Audio-lingual and Pitch Singing Groups.

Note: Error bars indicate standard error.

Then, I used Cohen's *d* (calculated as the mean difference between two groups divided by the pooled standard deviation) to examine the effect size of the two training method groups. Cohen believed that values of *d* around 0.2 to 0.3 were considered a small effect size, while those around 0.5 represented a medium effect size, and values from 0.8 and above indicated a large effect size (Russell 2018). Comparing the experimental group and control group's recognizing section, Cohen's *d* = (1.9667 – 0.0667) ⁄ 2.8672 = 0.6626 suggested a little over medium effect size, while the result of comparing the experimental group and control group's enunciating section, Cohen's *d* = (3.3333 – 1.3333) ⁄ 2.617364 = 0.764128 suggested a close-to-large effect size.

Lastly, to further determine which training method had a greater effect on recognizing and enunciating the four Chinese tones, an independent *t*-test was run on the data with a 95% confidence interval (CI) for the mean difference of gain scores on recognizing and enunciating the four Chinese tones. It was found that

the pitch singing training group performed significantly better on the recognition than the traditional audio-lingual group $t(49.536) = -2.566$, $p = .013$ (2-tailed). Also, the pitch singing training group performed significantly better on the enunciation than the traditional audio-lingual group $t(58) = -2.959$, $p = .004$ (2-tailed).

5. Discussion and Conclusion

Learning Chinese has become more popular in recent years. Despite the popularity of learning Chinese, one difficulty for non-tonal students to speak Chinese was enunciating the Chinese characters with accurate intonations (Wang and Honig 2010). Chinese tones are an abstract concept for non-tonal language speakers to comprehend. According to the previous studies, non-tonal language speakers are not used to the pitch range of the Chinese tones. The vocal range for enunciating the Chinese tones is much wider than the range for syllables in non-tonal languages. As indicated in Moore's (1991) study, adults often use their low portion of the vocal range in speaking and singing, but to accurately enunciate the four Chinese tones, one needs to use both low and high pitches.

Furthermore, some studies indicate that American Chinese language learners are often not aware of the important meaning that the Chinese tones carry (Huang 2000; Shi 2018). They tend to ignore and speak Chinese without the tones, which may result in inefficient communications. This non-complementary feedback could cause discouragement in learning, which would lead to motivation issues.

In the current study, I applied pitch singing training in learning the Chinese tones and compared this training method with the traditional audio-lingual method to explore and examine the different effects of the two training methods. By doing so, I hope to find an effective training method to help Chinese language learners to recognize and enunciate the four Chinese tones more accurately so that they are more motivated in learning and are proud to use Chinese more often.

5.1 Influences of the Singing-in-tune Skill

Previous studies reported that musicians who had refined ears could distinguish the difference in sound and intonations and they could also recognize

the four different Chinese tones better than non-musicians (Delogu et al. 2006; Lee and Hung 2008; Marie et al. 2011; Cooper and Wang 2012). In this study, the results showed the singing-in-tune ability did not influence the performances of recognizing and enunciating the four Chinese tones. This suggests that non-tonal language speakers who have better intonation skills in singing may not always perform better in recognizing and enunciating the four Chinese tones than those who cannot sing-in-tune.

The indication of singing-in-tune ability having no influence on the performances of the Chinese tones could be affected because participants had not to associate pitch differences and perceptions with the Chinese tones' intonations yet. Since all the participants either did not speak Chinese or had not taken Chinese courses before the study, they did not know the relationship between the intonation of pitch and the Chinese tones. Even participants who can sing-in-tune without practicing may have not used the intonation concept in distinguishing the four Chinese tones in the pretest. For those participants who took the pitch singing training method and already practiced pitch singing enough to build the connection between pitch intonation and the Chinese tones, the effects of their singing-in-tune skill already merged in with the effects of pitch singing skill. Thus, it is important to associate the Chinese tones' intonation with music pitch so that one could utilize the music skills to recognize and enunciate the Chinese tones more accurately.

5.2 Pitch Singing Training Effects

Chinese tones are an abstract concept that non-tonal language speakers often experience difficulty in associating with their native non-tonal languages. In order to construct and absorb the abstract Chinese tones' concept and to build a connection with learners' existing knowledge and skills, different strategies have been invented and applied to accelerate learning. Musical training improves phonologic awareness, accelerates the speed of retrieving memory, and helps to store information into long-term memory (Herrera et al. 2011).

Other studies indicated that for non-tonal language speakers to acquire a new tonal language, the right hemisphere region (where music is usually

processed) for identifying lexical tones of a tonal language is used specifically at the beginning of learning a tonal language. Later, the left hemisphere region will be engaged more as students have learned more of a tonal language (Gandour et al. 2000; Klein et al. 2001; Wang et al. 2003; Zatorre and Gandour 2008). In this study, the pitch singing training method helped the participants to build connections between the pitch intonation and the intonations of Chinese tones, as the positive effects transferred from the music pitch processing domain to the linguistic tone processing domain (Schön et al. 2004; Magne et al. 2006; Wong et al. 2007; Moreno et al. 2009). In the current study, the findings support the notion that music training has positive effects on language learning (Bidelman et al. 2013).

Furthermore, the results indicate that the pitch singing training method had a greater effect than the audio-lingual training method on recognizing and enunciating the four Chinese tones. Using pitch singing training provides Chinese language learners with a sensory experience of the four tones. The method stipulates learners a more precise perception of the four Chinese tones' contour and pitch registers. Nevertheless, the method accustoms learners' vocal muscles to produce the dramatic intonations of the tones and elaborates on the memory to store them for long-term retention. Therefore, Chinese language learners could understand the Chinese tones' characteristics better and enunciate and recognize them more accurately.

5.3 Limitation

In this study, I only focused on examining the effects of two training methods on individual Chinese character's tone pronunciation. I did not compare if there is any difference in training effects for different tones. It will be interesting to know that among the four Chinese tones, which one gained the most effects. Also, if the pitch singing training method has an effect on groups of characters' pronunciations instead of individual characters, such as examining the enunciation of two and more characters together. In addition, it would also be interesting to conduct the same study with one non-musician group and one musician group; each group would be divided into two sub-groups for taking two

training methods. Then, one could examine the influence of music skills on recognizing and enunciating the Chinese tones. Finally, one could apply this study to a larger sample size of participants in order to examine if the same results will be generated.

According to the results of this study, the pitch singing training method is effective in helping non-tonal language speakers to recognize and enunciate the four Chinese tones more accurately. To make the pitch singing training method applicable in classroom teaching, a possible and practical procedure for the first time (the first lesson) using this training method is presented in Appendix C.

5.4 Conclusion

From this study, one can conclude that pitch singing training is an effective method that can build associations between pitch intonation and Chinese tones and can set up an accurate concept of the Chinese tones in learners' minds. In addition, the pitch singing training made the participants aware of the difference between the Chinese tones and the tones they use in their non-tonal language, English. Moreover, this training method is practical for Chinese language teachers to implement in classroom teachings, as singing is a common activity that humans do for many occasions. Therefore, I believe the pitch singing training method can be an effective alternative exercise to help English speakers recognize and enunciate the four Chinese tones more accurately. Consequently, Chinese language learners can reduce misunderstandings because of inaccurate recognition and enunciation of tones in communication in Chinese, make conversations more efficient, and be more motivated in learning Chinese.

References

American Council on the Teaching of Foreign Languages. 2012. ACTFL Proficiency Guidelines 2012. Accessed online, November 5, 2018. https://www.actfl.org/publications/guidelines-and-manuals/actfl-proficiency-guidelines-2012

American Council on the Teaching of Foreign Languages. 2019. Immersion. Accessed online, November 5, 2018. https://www.actfl.org/membership

/special-interest-groups-sigs/immersion

Bai, Jianhua, Luyi Lien, and Madeline K. Spring. 2016. Mapping Chinese language learning outcomes in grades K-12. *CELIN (Chinese Early Language and Immersion Network) Briefs Series*, eds. by Shuhan C. Wang, and Joy K. Peyton, 1-2. New York, NY: Asia Society.

Best, Catherine C., and Gerald W. McRoberts. 2003. Infant perception of non-native consonant contrasts that adults assimilate in different ways. *Language and Speech* 46: 183-216. doi:10.1177/00238309030460020701

Bidelman, Gavin M., Stefanie Hutka, and Sylvain Moreno. 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. doi:10.1371/journal.pone.0060676

Bluhme, Hermann, and Ronald Burr. 1971. An audio-visual display of pitch for teaching Chinese tones. *Studies in Linguistics* 22: 51-57.

Brown, H. Douglas. 2014. *Principles of Language Learning and Teaching* (6th edition). White Plains, NY: Pearson Longman.

Chao, Yuenren. 1948. Lesson A. The tones. *Mandarin Primer: An Intensive Course in Spoken Chinese*, 85-91. Cambridge, MA and London, England: Harvard University Press. https://doi.org/10.4159/harvard.9780674732889.c9

Chen, Qinghai. 1997. Toward a sequential approach for tonal error analysis. *Journal of the Chinese Language Teachers Association* 32.1: 21-39. Accessed online, November 5, 2018. https://scholarsarchive.byu.edu/dlls/vol19/iss1/8/

Chuang, ChiuKuang, and Shizuo Hiki. 1972. The acoustical features and perceptual cues of the four tones of standard colloquial Chinese. *Proceedings of the Seventh International Congress on Acoustics (Budapest)*, 297-300. Accessed online, November 5, 2018. https://asa.scitation.org/doi/pdf/10.1121/1.1981919?class=pdf

Cohen, Jacob. 1988. *Statistical Power Analysis for the Behavioral Sciences* (2nd edition). Hillsdale, NJ: Lawrence Erlbaum Associates.

Cooper, Angela, and Yue Wang. 2012. The influence of linguistic and musical experience on Cantonese word learning. *Journal of the Acoustical Society of America* 131: 4756-4769. doi:10.1121/1.4714355

Delogu, Franco, Giulia Lampis, and Marta O. Belardinelli. 2006. Music-to-language transfer effect: May melodic ability improve learning of tonal languages by native non-tonal speakers? *Cognitive Processing* 7: 203-207. doi:10.1007/s10339-006-0146-7

Faul, Franz, Edgar Erdfelder, Albert-Georg Lang, and Axel Buchner. 2007. G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods* 39: 175-191. doi: 10.3758/BF03193146

Flege, James E. 1995. Second language speech learning: Theory, findings, and problems. *Speech Perception and Linguistic Experience*, ed. by Winifred Strange, 233-273. Baltimore, MD: York Press.

Gandour, Jack. 1984. Tone dissimilarity judgements by Chinese listeners. *Journal of Chinese Linguistics* 12: 235-261. Accessed online, November 5, 2018. http://www.jstor.org.unco.idm.oclc.org/stable/23767002

Gandour, Jack, Donald Wong, Li Hsieh, Bret Weinzapfel, Diana Van Lancker, and Gary D. Hutchins. 2000. A PET cross-linguistic study of tone perception. *Journal of Cognitive Neuroscience* 12: 207-222. doi:10.1162/089892900561841

Ghanbari, Fariba, and Mahmood Hashemian. 2014. The effects of English songs on young learners' listening comprehension and pronunciation. *International Journal of Language Learning and Applied Linguistics World (IJLLALW)* 6: 367-375.

Handel, Stephen. 1989. *Listening: An Introduction to the Perception of Auditory Events*. Cambridge, MA: MIT Press.

Herrera, Lucia, Oswaldo Lorenzo, Sylvia Defior, Gerard Fernandez-Smith, and Eugenia Costa-Giomi. 2011. Effects of phonological and musical training on the reading readiness of native- and foreign-Spanish-speaking children. *Psychology of Music* 39: 68-81. doi:10.1177/0305735610361995

Hsieh, Li, Jack Gandour, Donald Wong, and Gary D. Hutchins. 2001.

Functional heterogeneity of inferior frontal gyrus is shaped by linguistic experience. *Brain and Language* 76: 227-252. doi:10.1006/brln.2000.2382

Huang, Jingzi. 2000. *Students' Major Difficulties in Learning Mandarin Chinese as an Additional Language and Their Coping Strategies*. Accessed online, November 5, 2018. https://eric.ed.gov/?id=ED440537

Kiriloff, C. 1969. On the auditory discrimination of tones in Mandarin. *Phonetica* 20: 63-67.

Klein, Denise, Robert J. Zatorre, Brenda Milner, and Viviane Zhao. 2001. A cross-linguistic PET study of tone perception in Mandarin Chinese and English speakers. *Neuroimage* 13: 646-653. doi:10.1006/nimg.2000.0738

Lee, Chao-Yang, and Tsun-Hui Hung. 2008. Identification of Mandarin tones by English-speaking musicians and non-musicians. *Journal of the Acoustical Society of America* 124: 3235-3248. doi:10.1121/1.2990713

Lin, MaoCan. 1965. The pitch indicator and the pitch characteristics of tones in standard Chinese. *Acta Acoustica (China)* 2: 8-15.

Luo, Chunying. 2010. General description on the present status of Chinese textbooks in USA. *Journal of Jiangxi Science & Technology Normal University* 5: 71-77+124.

Magne, Cyrille, Daniele Schön, and Mireille Besson. 2006. Musician children detect pitch violations in both music and language better than non-musician children: Behavioral and electrophysiological approaches. *Journal of Cognitive Neuroscience* 18: 199-211. doi:10.1162/jocn.2007.19.9.1453

Marie, Céline, Franco Delogu, Giulia Lampis, Marta O. Belardinelli, and Mireille Besson. 2011. Influence of musical expertise on segmental and tonal processing in Mandarin Chinese. *Journal of Cognitive Neuroscience* 23: 2701-2715. doi:10.1162/jocn.2010.21585

Moore, Randall S. 1991. Comparison of children's and adults' vocal ranges and preferred tessituras in singing familiar songs. *Bulletin of the Council for Research in Music Education* 107: 13-22. Accessed online, November 5, 2018. http://www.jstor.org/stable/40318417

Mora, Carmen F. 2000. Foreign language acquisition and melody singing. *ELT*

Journal 54: 146-152. doi:10.1093/elt/54.2.146

Moreno, Sylvain, Carlos Marques, Andreia Santos, Manuela Santos, São L. Castro, and Mireille Besson. 2009. Musical training influences linguistic abilities in 8-year-old children: More evidence for brain plasticity. *Cerebral Cortex* 19: 712-723. doi:10.1093/cercor/bhn120

Morrison, Seven J., Steven M. Demorest, Elizabeth H. Aylward, Steven C. Cramer, and Kenneth R. Maravilla. 2003. FMRI investigation of cross-cultural music comprehension. *Neuroimage* 20: 378-384. doi:10.1016/S1053-8119(03)00300-8

Nan, Yun, and Angela D. Friederici. 2013. Differential roles of right temporal cortex and Broca's area in pitch processing: Evidence from music and Mandarin: Shared pitch network across domains. *Human Brain Mapping* 34: 2045-2054. doi:10.1002/hbm.22046

Ning, Cynthia. 2001. Second-language studies and college-level Chinese-language textbooks in the United States. *China Review International* 8.1: 34-56. https://doi.org/10.1353/cri.2001.0038

Patel, Aniruddh D. 2008. *Music, Language, and the Brain*. Oxford, United Kingdom: Oxford University Press.

Rahbar, Samira, and Samaneh Khodabakhsh. 2013. English songs as an effective asset to improve listening comprehension ability; evidence from Iranian EFL learners. *International Journal of Applied Linguistics and English Literature* 2.6: 63-66. Accessed online, November 5, 2018. http://www.journals.aiac.org.au/index.php/IJALEL/article/view/955

Rankin, Earl F., and Robert J. Tracy. 1965. Residual gain as a measure of individual differences in reading improvement. *Journal of Reading* 8.4: 224-233.

Repp, Bruno H., and Hwei-bing Lin. 1990. Integration of segmental and tonal information in speech-perception: A cross-linguistic study. *Journal of Phonetics* 18: 481-495. Accessed online, November 5, 2018. https://psycnet.apa.org/record/1991-29383-001

Russell, Joshua A. 2018. *Statistics in Music Education Research: A Reference for Researchers, Teachers, and Students*. New York, NY: Oxford

University Press.

Schön, Daniele, Maud Boyer, Sylvain Moreno, Mireille Besson, Isabelle Peretz, and Régine Kolinsky. 2008. Songs as an aid for language acquisition. *Cognition* 106: 975-983. doi:10.1016/j.cognition.2007.03.005

Schön, Daniele, Cyrille Magne, and Mireille Besson. 2004. The music of speech: Music training facilitates pitch processing in both music and language. *Psychophysiology* 41: 341-349.

Selinker, Larry. 1972. Interlanguage. *International Review of Applied Linguistics in Language Teaching* 10: 209-232. doi:10.1515/iral.1972.10.1-4.209

Shehadeh, Iman I A., and Mohammed A. A. Farrah. 2016. The effectiveness of using children's songs in developing elementary graders' English vocabulary and pronunciation in Jerusalem, Palestine. *Journal of Creative Practices in Language Learning and Teaching* 4.2: 44-67. Accessed online, November 5, 2018. https://www.researchgate.net/publication/313983571_The_Effectiveness_of_Using_Children's_Songs_in_Developing_Elementary_Graders'_English_Vocabulary_and_Pronunciation_in_Jerusalem_Palestine

Shen, Xiaonan S. 1989. Toward a register approach in teaching Mandarin tones. *Journal of the Chinese Language Teachers Association* 24.3: 27-47.

Shen, Chunxuan. 2009. Using English songs: An enjoyable and effective approach to ELT. *English Language Teaching* 2: 88-94. doi:10.5539/elt.v2n1p88

Shi, Hui. 2018. *A Method of Teaching English Speaking Learners to Produce Mandarin-Chinese Tones* (Order No. 10792055). https://doi.org/10.33915/etd.6633

Stewart, Lauren, Katharina von Kregstein, Jason D. Warren, and Timothy D. Griffiths. 2006. Music and the brain: Disorders of musical listening. *Brain* 129: 2533-2553. doi:10.1093/brain/asl171

Strong, William J., and George R. Plitnik. 1992. *Music Speech Audio*. Provo, UT: Soundprint.

Tsai, Rachel. 2011. Teaching and learning the tones of Mandarin Chinese. *Scottish Languages Review* 24: 43-50. Accessed online, November 5, 2018.

https://www.scilt.org.uk/portals/24/library/slr/issues/24/24_5_tsai.pdf

Wang, Hui-Hua, and Alice S. Honig. 2010. What difficulties do children experience while learning to read and write Chinese? *Teaching and Learning Chinese: Issues and Perspectives*, eds. by Jianguo Chen, Chuang Wang, and Jinfa Cai, 139-154. Charlotte, NC: Information Age Publishing.

Wang, Yue, Joan A. Sereno, Allard Jongman, and Joy Hirsch. 2003. FMRI evidence for cortical modification during learning of Mandarin lexical tone. *Journal of Cognitive Neuroscience* 13: 994-1005. doi:10.1162/089892903770007407

White, Carolyn M. 1981. Tonal pronunciation errors and interference from English intonation. *Journal of the Chinese Language Teachers Association* 16.2: 27-56.

Wong, Patrick C., and Tyler K. Perrachione. 2007. Learning pitch patterns in lexical identification by native English-speaking adults. *Applied Psycholinguistics* 28: 565-585. doi:10.1017/S0142716407070312

Wong, Patrick C., Erika Skoe, Nicole M. Russo, Tasha Dees, and Nina Kraus. 2007. Musical experience shapes human brainstem encoding of linguistic pitch patterns. *Nature Neuroscience* 10: 420-422. doi:10.1038/nn1872

Wray, Alison, and Michael R. Perkins. 2000. Functions of formulaic language. *Language and Communication* 20: 1-28. doi:10.1016/S0271-5309(99)00015-4

Yang, Juan. 2015. *Teacher and Pupil Beliefs about Beginning to Learn Chinese Language in English Secondary Schools*. Coventry: University of Warwick Doctoral dissertation. Accessed online, November 5, 2018. http://webcat.warwick.ac.uk/record=b2863329~S1

Zatorre, Robert J., and Jackson T. Gandour. 2008. Neural specializations for speech and pitch: Moving beyond the dichotomies. *Philosophical Transactions of the Royal Society of London B: Biological Sciences* 363: 1087-1104. doi:10.1098/rstb.2007.2161

Zatorre, Robert J., Ernst Meyer, Albert Gjedde, and Alan C. Evans. 1996. PET studies of phonetic processing of speech: Review, replication, and reanalysis. *Cerebral Cortex* 6.1: 21-30. doi:10.1093/cercor/6.1.21

[審查：2021.9.30 修改：2021.11.2 接受：2021.12.22]

周丹卿
Danqing ZHOU
200234 上海市徐匯區桂林路 100 號　上海師範大學初等教育學系
Department of Education
Shanghai Normal University
No.100, Guilin Road, Xuhui district, Shanghai, 200234, China
danqingzhou@shnu.edu.cn, clara.danching@gmail.com

Appendix A: Pretest and Posttest

Pretest

Name Email

I. Listen the recording and choose the appropriate tone
Ex: mā

a. b. c.  d. 

The answer is a

a	b	c	d	answer	answer
				1	11
				2	12
				3	13
				4	14
				5	15
				6	16
				7	17
				8	18
				9	19
				10	20

Contact Email: danqing.zhou@unco.edu
II. Pronounce the Chinese tones

1. mā	2. bó	3. kě	4. dì	5. tú
6. dǐ	7. tù	8. má	9. kē	10. bǒ
11. kè	12. dī	13. bò	14. tǔ	15. mǎ
16. bō	17. mà	18. dí	19. tū	20. ké

III Sing Mary had a little lamb

Mary had a little lamb, little lamb, little lamb, Mary had a little lamb, la la la la la.

Posttest

Name Email

I. Listen the recording and choose the appropriate tone
Ex: mā

a. b. c.  d. 

The answer is a

a	b	c	d	answer	answer
				1	11
				2	12
				3	13
				4	14
				5	15
				6	16
				7	17
				8	18
				9	19
				10	20

II. Pronounce the Chinese tones: a, o, e, i, u

1. bā	2. fàn	3. bú	4. wǒ	5. sì
6. hěn	7. zhōng	8. mén	9. fù	10. yī
11. mò	12. hǎo	13. jí	14. wǔ	15. kū
16. bò	17. lǐ	18. péng	19. shàng	20. tū

Appendix B: Materials of Introduction & Eight Training Sessions

Introduction	
Vowel	*ā á ǎ à, ō ó ǒ ò, ē é ě è* *ī í ǐ ì, ū ú ǔ ù, ǖ ǘ ǚ ǜ*
Consonant	*b p m f, d t n l, g k h, j q x, zh, ch, sh, z, c, s, r*
Some Combinations of Vowel and Consonant	Example: *bāng, hóng, děng, lìng, pái, fèi, mù, yü, jiā, zhōng, guó, cǐ, rì*
Session 1 — Review: Pinyin and Tones	
Vowel	*ā á ǎ à, ō ó ǒ ò, ē é ě è* *ī í ǐ ì, ū ú ǔ ù, ǖ ǘ ǚ ǜ*
Consonant	*b p m f, d t n l, g k h, j q x, zh, ch, sh, z, c, s, r*
Some Combinations of Vowel and Consonant	Example: *tāng, nóng, fěng, dìng, bái, fèi, zhù, sī, qiā, zhōng, guó, zǐ, lì*
Session 2 — Lesson One: Greetings	
Vocabulary	1. *wǒ* 我 (I), 2. *nǐ* 你 (you), 3. *hǎo* 好 (good-adj.), 4. *jiào* 叫 (to be called-verb), 5. *Shén me* 什么 (what-qpr), 6. *míng zì* 名字 (name-noun), 7. *hěn* 很 (very-adv.), 8. *gāo xìng* 高兴 (happy-adj.), 9. *rèn shí* 认识 (know-verb), 10. *yě* 也 (also/too-adv.)
Dialogue	Person A-*nǐ hǎo*, Person B-*nǐ hǎo*, *wǒ jiào* (YOUR NAME). *nǐ jiào shénme míng zì*? Person A-*wǒ jiào* (…). *hěn gāo xìng rèn shí nǐ*. Person B-*wǒ yě hěn gāo xìng rèn shí nǐ*. Translation: -Hello! -Hello! I call myself (…). What is your name? -My name is (…). Nice to meet you. -Nice to meet you too.
Sentence Structure	a. Subject+Verb+Object *wǒ* (subject) *jiào* (verb) Mary (object). b. In Chinese, the question is composed without changing the position of a verb. In this sentence "what is your name?" the question word "*shénme*" is used for indicating this is a question. *nǐ* (subject) *jiào* (verb) *shénme* (what-qpr) *míng zì* (object)?

Session 3 — Lesson Two: Introduce Myself	
Vocabulary	1. *shì* 是 (be-verb), 2. *bú shì* 不是 (not-adv.), 3. *xué shēng* 学生 (student-noun), 4. *xué* 学 (learn-verb), 5. *ne* 呢 (question mark, no meaning), 6. *rén* 人 (people-noun), 7. *nǎ lǐ* 哪里 (where-qpr), 8. *ma* 吗 (question mark, no meaning)
Dialogue	Person A-*nǐ hǎo*! Person B-*nǐ hǎo*! *wǒ jiào* (YOUR NAME). *nǐ jiào shénme míng zì*? Person A-*wǒ jiào* (…). *wǒ shì xué shēng*. *nǐ shì xué shēng ma*? Person B-*wǒ shì xué shēng*. *wǒ xué* (major: education/music/engineering). *nǐ ne*? Person A-*wǒ xué* (the major). *wǒ shì* Denver *rén*. *nǐ shì nǎ lǐ rén*? Person B-*wǒ yě shì* Denver *rén*. *hěn gāo xìng rèn shí nǐ*. Person A-*wǒ yě hěn gāo xìng rèn shí nǐ*. Translation: -Hello! -Hello! I call myself (…). What is your name? -My name is (…). I am a student. Are you a student? -I am a student. I study (…) How about you? -I study… I am from Denver, and you? -I am from Denver too. Nice to meet you. -Nice to meet you too.
Sentence Structure	a. Subject+Verb+Object *wǒ* (subject) *shì* (verb) *xué shēng* (object). *wǒ* (subject) *xué* (verb) music (object). b. Question: add "*ne*" and "*ma*" at the end of the sentence to make a question. *nǐ ne*? - How about you/And you? *nǐ shì xué shēng ma*? - Are you a student?
Session 4 — Lesson Three: Making Friends	
Vocabulary	1. *zhè* 这 (this-pr.), 2. *de* 的 (a possessive or descriptive particle-p)/*wǒ de* (my), 3. *péng yǒu* 朋友 (friend-n.), 4. *zhuān yè* 专业 (major-n.), 5. *jǐ* 几 (question word for number), 6. *nián jí* 年级 (grade-n.), 7. *yī* 一 (one-n.), 8. *èr* 二 (two-n.), 9. *sān* 三 (three- n.), 10. *sì* 四 (four-n.), 11. *wǔ* 五 (five-n.)
Dialogue	Alex-*zhè shì wǒ de péng yǒu* Linda. *zhè shì wǒ de péng yǒu* Max. Linda-*nǐ hǎo*, Max. *hěn gāo xìng rèn shí nǐ*. Max-*nǐ hǎo*, Linda. *wǒ yě hěn gāo xìng rèn shí nǐ*. Linda-*nǐ xué shénme zhuān yè*? Max-*wǒ xué* music/education/language… *zhuān yè*, *nǐ ne*? Linda-*wǒ xué* (…) *zhuān yè*. *nǐ shì jǐ nián jí*?

	Max-*wǒ shì yī nián jí, nǐ ne*? Linda-*wǒ shì* (*yī*/*èr*/*sān*/*sì*) *nián jí*. Translation: Alex-This is my friend, Linda. This is my friend, Max. Linda-Hello, Max. Nice to meet you. Max-Hello, Linda. Nice to meet you too. Linda-What is your major? Max-I study music/education/language…, and you? Linda-I study (…) Which grade are you in? Max-I am in the first grade, and you? Linda-I am in the grade (…).
Sentence Structure	This is +Topic *zhè shì* (this is) *wǒ de péng yǒu* Linda (topic).
Session 5 — Lesson Four: Time	
Vocabulary	1. *xiàn zài* 现在 (now-time), 2. *diǎn* 点 (measure word for time-one o'clock/lit. dot, point), 3. *le* 了 (a dynamic particle, no actual meaning), 4. *liù* 六 (six-n.), 5. *qī* 七 (seven-n.), 6. *bā* 八 (eight-n.), 7. *jiǔ* 九 (nine-n.), 8. *shí* 十 (ten-n.), 9. *qù* 去 (to go-verb), 10. *shàng kè* 上课 (take class), 11. *kè* 课 (class-noun), 12. *zài jiàn* 再见 (goodbye)
Dialogue	Person A-*xiàn zài jǐ diǎn le*? Person B-*nǐ hǎo, xiàn zài jiǔ diǎn le. nǐ qù nǎ lǐ*? Person A-*wǒ qù shàng kè. nǐ qù nǎ lǐ*? Person B-*wǒ yě qù shàng kè. nǐ qù shàng shénme kè*? Person A-*wǒ qù shàng* (history) *kè, nǐ ne*? Person B-*wǒ qù shàng* (…) *kè. zài jiàn.* Person A-*zài jiàn.* Translation: -Hello, what time is it now? -Hello, it is 9 o'clock. Where are you going? -I am going to a class. Where are you going? -I am also going to a class. What class are you going to? -I am going to the history class, and you? -I am going to the (…) class. Goodbye. -Goodbye.
Sentence Structure	What time is it now? *xiàn zài jǐ diǎn le*? In Chinese, it is literally saying: Now what time? Where are you going? *nǐ qù nǎ lǐ*? In Chinese, it is literally saying: You go where? Goodbye-*zài jiàn*

Session 6 — Lesson Five: Numbers and Counting	
Vocabulary	0. *líng* 0, 1. *yī* 一 (one-n.), 2. *èr* 二 (two-n.), 3. *sān* 三 (three-n.), 4. *sì* 四 (four-n.), 5. *wǔ* 五 (five-n.), 6. *liù* 六 (six-n.), 7. *qī* 七 (seven-n.), 8. *bā* 八 (eight-n.), 9. *Jiǔ* 九 (nine-n.), 10. *shí* 十 (ten-n.), 11. *shí yī* 十一 (11), 12. *èr shí yī* 二十一 (21), 13. *yī bǎi* 一百 (100), 14. *duì bù qǐ* 对不起 (I am sorry)
Culture Facts: Good and Bad Numbers	• Good (*hǎo*): *liù* 六 (six-n.), *bā* 八 (eight-n.), *jiǔ* 九 (nine-n.), *líng* 0 • Bad (*bù hǎo*): *sì* 四 (four), *shí sān* 十三 (thirteen), *shí sì* 十四 (fourteen) • Counting: *qiān* 千 (thousand), *bǎi* 百 (hundred), *shí* 十 (ten), *gè* 个 (single). 1365- *yī qiān sān bǎi liù shí wǔ*
Session 7 — Lesson Six: Students	
Vocabulary	1. *tā* 她 (she-pr.), 2. *měi guó* 美国 (America), 3. *zhōng guó* 中国 (China), 4. *è* 饿 (hungry-verb), 5. *nǐ mén* 你们 (you plural), 6. *wǒ mén* 我们 (we), 7. *chī* 吃 (eat-verb), 8. *fàn* 饭 (rice/meal-n.), 9. *qù chī fàn* 去吃饭 (go to eat) 10. *tā* 他 (he-pr.)
Dialogue	Alex-Max, *zhè shì wǒ de péng yǒu* Linda. *tā shì měi guó rèn*. *tā shì sān nián jí de xué shēng*. Max-*nǐ hǎo*, Linda. *wǒ shì zhōng guó rén*, *wǒ shì yī nián jí de xué shēng*. Linda-*hěn gāo xìng rèn shí nǐ*. Max-*wǒ yě hěn gāo xìng rèn shí nǐ*. Alex-*xiàn zài jǐ diǎn le*? Max-*xiàn zài shí èr diǎn*. Alex-*wǒ è le*, *nǐ mén è bú è*? Max & Linda-*wǒ yě è le*. Alex-*hǎo*, *wǒ mén qù chī fàn*. Translation: Alex-Max, this is my friend, Linda. She is American. She is a third-year student. Max-Hello, Linda. I am a Chinese. I am a first-year student. Linda-Nice to meet you. Max-Nice to meet you too. Alex-What time is it now? Max-It is 12 o'clock. Alex-I am hungry. Are you guys hungry? Max & Linda-I am hungry too. Alex-Ok, let's go to eat.

Sentence Structure	1. A not A. E.g.: a. *è bú è*-hungry or not hungry; b. *shì bú shì*- yes or no/isn't it, c. *hǎo bù hǎo*-good or bad *2. le*- to state a situation, something happened. *le* by itself doesn't have a meaning. E.g.: *wǒ è le* (我饿了) - I am hungry. *wǒ shàng kè le* (我上课了)- I am in the class. *wǒ chī le* (我吃了) - I ate. 3. Subject+verb+[…..*de*]+object. *wǒ shì* [*yī nián jí de] xué shēng.* I am a first year student.
Session 8 — Lesson Seven: Eat	
Vocabulary	1. *xiǎng* 想 (want to/think-verb), 2. *xī cān* 西餐 (western meal-noun), 3. *zhōng cān* 中餐 (Chinese meal-noun), 4. *hái shì* 还是 (or), 5. *xǐ huān* 喜欢 (like-verb), 8. *hàn bǎo bāo* 汉堡包 (burger-n.), 9. *wǎn fàn* 晚饭 (dinner), 10. *jiàn* 见 (meet-verb), 11. *jiā* 家(home-n.)
Dialogue	Max-*nǐ xiǎng chī shénme, xī cān hái shì zhōng cān*? Linda-*wǒ xiǎng chī xī cān, wǒ xǐ huān chī hàn bǎo bāo. nǐ ne*? Max-*wǒ xǐ huān chī zhōng cān, wǒ yě xǐ huān chī xī cān*. *wǒ mén jǐ diǎn chī wǎn fàn*? Linda-*wǔ diǎn, hǎo bù hǎo*? Max-*hǎo, wǒ mén nǎ lǐ jiàn*? Linda-*wǒ jiā jiàn*. Max-*nǐ jiā jiàn*. Linda-*zài jiàn*. Translation: -What do you want to eat, western meal or Chinese meal? -I want to eat western meal. I like to eat burgers. What about you? -I like to eat Chinese food, I also like western food. What time are we going to eat dinner? -Five o'clock, is it fine? (good or not good)? -Fine, where are we going to meet? -At my home. -Ok, see you at your home. -Goodbye.
Sentence Structure	A *hái shì* B- A or B E.g.: a. *xī cān hái shì zhōng cān*-western food or Chinese food b. *xué shēng hái shì lǎo shī*- student or teacher

Appendix C: Procedure for the Pitch Singing Training Method (First Lesson)

1. Step One

Vocal warm-up: sing scale “b, c, d, e, f, g” with six vowels and use both upward and downward scale.

For example: use vowel “a” sing the upward scale “b, c, d, e, f, g” and then sing the downward scale “g, f, e, d, c, b”.

2. Step Two

Use assigned intervals to sing four tones with six vowels.

The first flat and high pitched tone can use note G4 to G4, the second middle rising tone can use notes from E4 to G4 or G4 sharp, the third low falling rising tone can use notes from C4 to lower B3 then go up to D4, and the fourth falling tone can use note from G4 to lower F4.

Teachers can use a digital piano keyboard app or music instrument app that can play an octave of notes from a smart phone to play the intervals first. Teachers can also record these interval sound samples from an instrument before using them in the class. Teacher can then lead students to sing the interval with a vowel along with the digital musical instrument app or recordings.

For example: sing “*ā*” along with playing note G4-G4, sing “*á*” along with playing note E4 to G4, sing “*ǎ*” with playing note from C4 to lower B3 then go up to D4, and sing “*à*” with playing note G4 to lower note F4.

For students whose vocal ranges are lower, the pitch can be lowered by one or two notes. For example, the first tone can use note F4 to F4, the second tone can use notes D4 to F4, the third tone can use notes B3 flat-A3-C4, and the fourth tone can use notes F4 to E4 flat.

3. Step Three

Sing one *pīn yīn* with the intervals.

After singing six vowels with the pitch intervals, students can follow the teacher in singing one complete *pīn yīn* with a group of pitches by combining vowels and consonants together. For example, sing *pīn yīn* “*kě*” along with the pitch intervals “C4-B3-D4” and *pīn yīn* “*mā*” with the pitch interval “G4-G4”.

4. Step Four

Sing a *pīn yīn* in a faster speed for a few times, then transition to saying it without singing the pitches.

唱音高訓練對辨別和念讀漢語四聲調的作用研究

周丹卿
上海師範大學初等教育學系

摘要

對於母語是非音調語言的學習者來說，準確辨別和念讀漢語的四聲調是一個大挑戰。在這項研究中，調查員測試了唱音高訓練是否有益於帶音調語言的學習，評估了唱音高訓練是否對辨別和念讀拼音中的四種聲調有作用，並將唱音高訓練方法與傳統的聽－模仿跟讀模式（無音樂）的效果進行了比較。本研究招募了 60 名從沒有上過中文課程，也不會說和寫漢語的美國大學生。參與者被平均隨機分為兩組：唱音高訓練組和傳統聽－模仿跟讀訓練組。每一組的實驗模式為：一次前測／八次訓練活動／一次後測。結果表明，與傳統的聽－模仿跟讀訓練方法相比，唱音高訓練方法可以提高母語非音調語的學習者對漢語四聲的辨別和念讀的精準度。唱音高訓練法可以為母語非音調語的學習者提供一種感官體驗，建立對抽象漢語聲調概念的聯想，並加深學習者對漢語四聲調的記憶，以便存入長期記憶模式。因此，在漢語四聲調教學中可融入唱音高訓練法，從多維度和感性體驗上幫助學習者更快、更精準地掌握四聲調，從而增強學習動力和成功率。

關鍵詞：四聲調音樂學習法　念讀拼音四聲調　唱音高訓練
聽辨拼音四聲調