8-2019

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

Danqing Zhou

UNIVERSITY OF NORTHERN COLORADO

Greeley, Colorado

The Graduate School

THE EFFECTS OF PITCH SINGING TRAINING ON RECOGNIZING AND ENUNCIATING CHINESE TONES

A Dissertation Submitted in Partial Fulfillment
of the Requirements for the Degree of
Doctor of Arts

Danqing Zhou

College of Performing and Visual Arts
School of Music
Music Education

August 2019

This Dissertation by: Danqing Zhou

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

has been approved as meeting the requirement for the Degree of Doctor of Arts in the College of Performing and Visual Arts in the School of Music, Program of Music Education.

Accepted by the Doctoral Committee

_____
Mark Montemayor, Ph.D., Research Advisor

_____
Nancy Glen, D.A., Committee Member

_____
Khalil Shafie Holighi, Ph.D., Committee Member

_____
Jingzi Huang, Ph.D., Faculty Representative

Date of Dissertation Defense _____

Accepted by the Graduate School

_____
Linda L. Black, Ed.D.
Associate Provost and Dean
Graduate School and International Admissions
Research and Sponsored Projects

ABSTRACT

Zhou, Danqing. *The effects of pitch singing training on recognizing and enunciating Chinese tones.* Published Doctor of Arts dissertation, University of Northern Colorado, 2019.

It is a challenge for non-tonal language speakers to enunciate the four Chinese tones accurately. In the present 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. The participants, 60 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 a long-term retention.

ACKNOWLEDGMENTS

Many people have contributed to this dissertation; without them I would not reach my goal. My special thanks to my research advisor Dr. Mark Montemayor, for believing, encouraging, and supporting me to complete my study and the dissertation. Dr. Montemayor has helped me in all kinds of ways — not only in my studies, but in academic writing and other aspects in life when I needed help.

I also need to give my sincere thanks to all the participants who donated their time and effort to learn Chinese with me, and to those who helped me to recruit, namely Dr. Milian Madeline, Sumiko Gibson, Dr. I-hui Lee, Dr. Shafie Khalil, Dr. Erin Noelliste, and Dr. Michelle Low. Without their help, I could not have reached the number of the participants that I needed.

Moreover, I should give my gratitude to my dissertation committee members Dr. Mark Montemayor, Dr. Shafie Khalil, Dr. Nancy Glen, and Dr. Jingzi Huang. Many thanks for the insightful suggestions and comments for my dissertation.

Last but not least, I need to dedicate my biggest gratitude to my family. My parents' unconditional love and support made me who I am today; a thousand thanks are not enough to express my appreciation (感谢爸爸妈妈，谢谢你们多年的无私奉献让我获得了今天的成绩). Many thanks to my husband, Matteo Sabattini, for giving up his job and accompanying me for four years at University of Northern Colorado.

TABLE OF CONTENTS

LIST OF TABLES

TABLE

LIST OF FIGURES

FIGURE

CHAPTER I

INTRODUCTION

Mandarin Chinese is the official language that is used in schools, newspapers, media, and governments in mainland China. However, there are 56 ethnic groups in the country. Each ethnic group speaks one or multiple languages such as Mongolian, Tibetan, and Uighur and one or multiple dialects, such as Cantonese, Shanghainese, Minnan, and Taiwanese. Chinese is not a language that uses the Roman alphabet like English, Italian, or French, in which the sound of each character or word can be reflected from the pronunciation of the letter or combination of letters. Instead, Chinese uses logograms to express meaning, like these: 我 (I or me), 你 (you), 她 (she).

Although Chinese use the same Chinese characters for their written language, people who speak different dialects apply different pronunciations to the same characters. By looking at these written texts, non-Chinese speakers would not know how to pronounce them unless someone teaches the pronunciation of written words. Thus, it was not easy to make Mandarin an official language across China before 1958, because there was not a standard pronunciation system that could apply to all the Chinese characters.

Around 1950, the Chinese government called a group of linguists to find a solution to help Chinese people be literate and be able to speak Mandarin in addition to their dialects (Luo, 2018). In 1958, Zhou You Guang (周有光), who led a group of Chinese linguists, finalized and established the standard system that uses the Roman

alphabet with a diacritical mark notation system for indicating Chinese characters' pronunciation, pīn yīn (Luo, 2018).

In the pīn yīn system, the four tones of Chinese characters are presented by four kinds of tone marks. Each tone carries different pitches and uses a unique tone mark to help people learn the intonation, for example: *mā*-mother-妈, *má*-hemp-麻, *mǎ*-horse-马, *mà*-curse-骂 (Xing, 2006). The four Chinese tones are categorized into five pitch levels, with the first tone being high-flat level pitch, the second tone low to high-rising pitch, the third tone low-dipping pitch, and the fourth tone high-falling pitch (Chao, 1948). To learn Chinese, all students often start from learning pīn yīn, because the Chinese characters themselves do not present the pronunciation. The sound of each character is learned separately from the pīn yīn system. Secondly, once students master the pīn yīn system, they can use this tool to help them learn the pronunciation of Chinese characters on their own. In the Chinese dictionaries, each character is labeled with one pīn yīn. By reading pīn yīn, students can pronounce words on their own. Consequently, students do not need to depend heavily on teachers to learn how to read Chinese aloud.

However, Chinese learners need to pay attention to the intonation of the four Chinese tones because these tones change the meaning of characters. Many Chinese words share the same letter spelling in pīn yīn but have different tone markings; for example, "wēn" means warm and "wèn" means ask. In general, the context of dialogue provides speakers the background of the conversation's meaning, even if the Chinese tone is not enunciated quite accurately. Under the guidance of situational context, listeners can guess and construct meaning to certain extent, but the more precisely speakers can enunciate the Chinese tones, the more efficient and clear the

communications can be delivered. It will also cause less misunderstanding if speakers pronounce the tones accurately. For example, if someone wants to ask a question, he or she would say “I would like to ‘wèn’ (ask) you,” but if the tone of asking is not pronounced correctly, it could sound like “I would like to ‘wěn’ (kiss) you” instead. Also, some locations’ names in China are very close to each other and their difference is from the tone. For example, one province is called shǎn xī, the other province is shān xī. Both provinces use the same alphabet letters in pīn yīn to indicate the sound. The only difference is on the first character’s tone mark. Using different tones on the words “shan xi” indicates various meanings. Therefore, recognizing and enunciating Chinese tones accurately are important for Chinese language learners.

Although the Roman alphabet is familiar to many people, the intonation of each tone mark that is used in pīn yīn is new to people whose native languages are not tonal. Studies on Chinese tones recognition revealed that tone confusion for Chinese language learners occurs because Chinese tones have a wider pitch range than English (Chen, 1997; Repp & Lin, 1990), and some pitch ranges do not appear in non-tonal languages (White, 1981). It is not usual to ask non-tonal language speakers to pronounce the words as if singing high and low pitches. According to Yang (2015), non-tonal English speakers have difficulty in pronouncing the Chinese tones accurately because they are not familiar with the intonations. In addition, some students feel awkward and strange when saying pīn yīn with high and low pitches (Tsai, 2011).

In accordance with the pitch range of the four Chinese tones, to pronounce Chinese tones accurately one almost needs to sing these tones. Language teachers often use songs to help students to learn languages. English songs are widely used in English

learning, and studies show the benefits of how singing improves listening and speaking skills, words' memorization, and motivation of learning English (Rahbar & Khodabakhsh, 2013; Schön et al., 2008; Shen, 2009). However, there are not sufficient studies to establish the effects of music training on tonal language pronunciation and sound recognition (Shi, 2018).

In teaching Chinese to American college students, I have found that students had difficulties in enunciating the four Chinese tones accurately, especially the first and fourth tones. both of which start from a high pitch. American students are not used to producing a sound in high pitch in speaking mode; they tend to pronounce pīn yīn close to their native English language without emphasizing the four distinct tones' intonations. Even though the students have learned the different pitch ranges of these four tones, they do not pronounce them dramatically enough to match the intonations.

In a language class, Chinese teachers are familiar with the content. Thus, they can guess most of the meaning that the students want to express, even with inaccurate tones. However, I have found if my students try to have improvised conversations with me, in which the content of the conversation is not arranged in advance, sometimes I have difficulties understanding their words due to inaccurate tone enunciation. Unlike non-tonal languages, each tone in Chinese carries a specific meaning. Therefore, it is important to recognize and enunciate the four Chinese tones accurately.

The traditional teaching method used to improve students' pronunciation and recognition of Chinese tones is the "audio-lingual method" (Xing, 2006). In this teaching method, students imitate teachers' pronunciation or practice speaking with recordings. Since Chinese tones are categorized by pitch levels, pitch singing activities may help

Chinese language learners to recognize and enunciate the four Chinese tones more accurately. In a pitch singing training approach, students can practice enunciating the four Chinese tones first by singing four groups of intervals. For example, the four tones can be interpreted as the first tone is sung note G4 to G4 (a unison interval), the second tone is sung middle E4 to G4 (an ascending major or minor third interval), the third tone is sung middle C4 (middle C), descending by half step to B, then rising to D4, and the fourth tone is sung G4 to F4 (a descending major second interval). The graphic notation of four groups of intervals is presented in Figure 1.

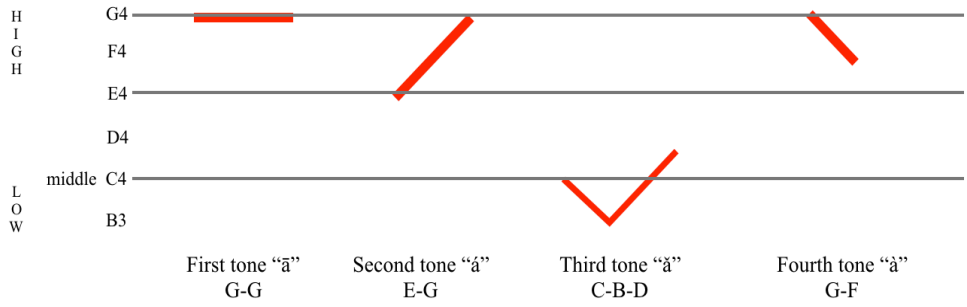

Figure 1. Graphic notation of four groups of intervals for the Chinese tones.

Once students have the perception of the four tones' pitch ranges and intonations, they could be asked to pronounce these tones without singing the intervals. Practicing singing various pitch groups according to each Chinese tone's intonation may help non-tonal language speakers to build schema of dramatic intonations of the Chinese tones, to extend Chinese language learner's voice range of the speaking mode, and to associate this new tonal concept with a familiar singing experience.

The current study is inspired by my foreign language learning and teaching experiences. I have been studying and teaching music and Chinese in the United States

for 10 years, and had taught music and Chinese in Italy for five years. My native language is Mandarin Chinese. Among the four aspects of foreign languages learning (speaking, listening, reading, and writing), I felt less challenged by listening and speaking, especially imitating the pronunciation. Language teachers whom I know and my friends often conclude the reason for having few difficulties in listening and speaking is because I am a musician. I was told that my "trained ears" can differentiate and recognize the sounds better than those without musical training. Also, I have known many musicians who speak two or more languages in addition to their native language. Is it true that music training benefits language learning? This theory inspired me to develop some research questions and to form a point of view regarding how to teach Chinese pronunciation effectively during the years I have studied and worked abroad. Reflecting on my own language learning experiences, I wondered if there was a music training activity that could help Chinese language learners to recognize and enunciate the four tones more accurately. With the question firmly in my mind, I started to explore studies of language learning, the relationship between music and language, and effective teaching methods in foreign language learning and Chinese tones learning.

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 audio-lingual method.

CHAPTER II

REVIEW OF LITERATURE

Chinese is a tonal language in which the four tones are visually presented in the pīn yīn system by four different tone marks. The meaning of Chinese words is expressed by the combination of pronunciation and tones. The same pronunciation enunciated with different tones may convey different meaning. To make communication more understandable and efficient, recognizing and enunciating the Chinese tones precisely is important to Chinese speakers. Chinese language learners often face the challenge of recognizing and enunciating the four Chinese tones because the intonations are different from non-tonal languages; for example, the pitch range of the Chinese tones is wider than English pronunciation (Chen, 1997; Repp & Lin, 1990). Also, some pitch ranges do not appear in non-tonal languages (White, 1981). Moreover, non-tonal English speakers have difficulty in pronouncing the Chinese tones accurately because the four Chinese tones are a new concept to them and are unfamiliar to non-tonal speakers (Yang, 2015). Lastly, because of the wide pitch range of the Chinese tones, some students feel awkward and strange when enunciating pīn yīn with high and low pitches (Tsai, 2011).

This literature review begins with discussing broad considerations of language learning, such as how people acquire languages, the differences between first and second language acquisition, and the effects of one language's background on the acquisition of other languages. Then, the characteristics and learning challenges of Chinese language and Chinese tones are explored. I proceed by examining the relationship between music

and language and the effects of music training on language learning. The chapter concludes by discussing the reasons for applying a singing training method to language learning.

General Perspective of Language Learning

Language is a tool that human beings use for expressing emotions, understanding others, and communicating with each other. Languages that are acquired since birth to the early childhood are considered as first languages or native languages (Saville-Troike, 2012), while those that are learned later, after the first language, are considered as second languages (Lightbown & Spada, 1993). The following aspects are going to be discussed in the proceeding section: (1) how people master a language, (2) whether humans are born with the ability of using languages or languages are acquired by learning, and (3) what the differences are between first language (L1) and second language (L2) learning. Linguists and researchers have provided different views on how humans learn languages.

To acquire first languages, infant children often start with their care providers. Parents often say words repeatedly to their babies and are proud and excited when they repeat the words. Babies receive positive feedback from their parents and are often very happy to continue repeating and imitating the sound that they have heard. Traditional behaviorists believed that language is learned by imitating sounds and forming habits from this feedback (Skinner, 1957). Language learners, especially babies, start with imitating the sounds and patterns that they have heard around them. By receiving positive feedback, language learners will continue imitating the sounds and form speaking routines (Lightbown & Spada, 1993). The following scenario illustrates how humans learn a language from imitating sounds:

Parents: say mama, mama

Babies: ma, ma

Parents: mama, mama (point at mom)

Babies: mama, mama

Parents: yeah! You said mama, mama! (Then mom hugs or kisses babies)

Babies: mama, mama

Mom: yes. (Mom will often give positive feedback to acknowledge babies by verbal or physical confirmation)

From the behaviorists' view, language learning happens when imitating and practicing, but people do not learn all the words, sentences, and structures from imitating. Babies can create their own stories and manipulate words to suit their needs. According to Chomsky (1995, 2013), language is an innate ability that humans are born with. He named this innate ability as the language acquisition device (LAD). LAD can be defined as

> an imaginary "black box" which exists somewhere in the brain. This device contains all and only the principles which are universal to all human languages […]. For the LAD to work, the child needs access only to samples of the natural language. These language samples serve as a trigger to activate the device (Lightbown & Spada, 1993, p. 8).

When children get exposed to language material, they can create and construct the meaning on their own. Children not only learn language through imitating and building habits, but they also have the innate ability to construct the language information they encounter and master the language by themselves.

While behaviorists believe that languages are learned through imitating, Chomsky (2013) states that language is a device that human beings possess from birth, and humans can activate the language acquiring device on their own as long as they have the appropriate environment. In addition to the imitation and natural language acquiring theories, interactionists claim that language is gained through interaction (Long, 1985). From the interactionists' view, language is learned through complex interplay between the learner and the environment in which the learner lives (Lightbown & Spada, 1993). Unlike Chomsky's innate theory, "the interactionists claim that language which is modified to suit the capability of the learner is a crucial element in the language acquisition process" (Lightbown & Spada, 1993, p. 14). The following conversation is an example of how an adult modified the conversation information to scaffold a child's language acquisition.

Child: Car red, fast, fast.

Mom: You mean the red car goes fast.

Child: Yes, that red car goes fast.

From the modified interaction, language learners receive the appropriate amount of information they need for building their language system (Gass & Selinker, 1994; White, 1987).

Like how first languages are acquired, second language can also be learned through sound imitation, using LAD to construct communications repeatedly and through interactions with native speakers or language teachers. L2 students often learn foreign languages in classrooms where they imitate the sound that the teachers provide. In a language classroom, students not only imitate and repeat the sound of their teachers, but

they also use the learned second language to construct meanings to make communication. In addition, students receive modified input from formal language classrooms, schools, and the environments around them. Long (1985) states that language learners receiving modified and comprehensible input through interaction with native speakers is a necessary process in acquiring languages.

Other linguists and psychologists like Stephen Krashen and Lev Vygotsky also share a similar learning perspective with interactionists. According to Krashen's (1982) theory of language acquisition, students will acquire language when the input is understandable and accessible. Comprehensible input will lead students to learn the elements of a language, such as pronunciations, grammar, and sentence structures. The manageable level of input information needs to be a little higher than the student's ability. Krashen (1982) names this i+1 (input plus one). If the input information is far beyond the student's ability, then language acquisition will not happen (Freeman & Freeman, 2001).

Similar to Krashen's input hypothesis, Vygotsky's (1978) Zone of Proximal Development theory also states that acquisition happens when a learner's existing developmental state is extended to his or her potential development by scaffolding. It can be analogous to stairs or ladders that teachers provide to students in learning. Teachers should provide students the suitable size or height of the stair or ladder to help them understand and absorb new knowledge. According to Vygotsky (1978, 1986), students do not develop from one point or level directly to the next and higher level. The developmental process requires connecting new knowledge to existing knowledge, then learners absorb and understand new knowledge by reflecting on their past experiences.

In the developmental process, students may need an outsider to help them to make connections with and to reflect from the past experiences. This progression is “not in a circle but in a spiral form, passing through the same point at each new revolution while advancing to a higher level” (Vygotsky, 1978, p. 56). Thus, it is important for a language teacher to offer students appropriate and comprehensible levels of input, so that language acquisition can be facilitated. Correspondingly, pronouncing the four Chinese tones is close to singing (Shi, 2018; Zhang, 2006). Using singing activities may help students to better understand the intonations of four Chinese tones. Students can hear and feel the pitch range and length of each tone through singing. Thus, singing activities are more comprehensible than directly imitating the tones in speaking mode for students to understand the pitch range and intonations. For language learners, whose native language is not tonal, perhaps teachers can use some musical activities to scaffold or provide comprehensible input for students to understand the intonations of the four Chinese tones. Then, students will be able to recognize and enunciate them more accurately and clearly.

According to Krashen’s (1982) acquisition and learning hypothesis, languages should be obtained in a more subconscious process (immersing in the language environment and using languages for other purposes), rather than in conscious learning (talking about languages and learning grammar, pronunciation, and structure of languages). He believes a subconscious process helps humans acquire language, while learning process does not make language acquisition happen. Krashen (1982) states that acquisition allows humans to speak, understand, read and write the language, while learning allows students to pass exams on a language or talk about language.

Freeman and Freeman (2001) state that students who learn a second language through a learning mode often display limited communication skills. They can talk about the language and they know the grammar and words or have enough knowledge to pass the language exams, but students have difficulties in speaking and carrying conversations. Thus, if a language course is offered in a traditional way, students will have fewer opportunities to practice speaking skills and will have difficulty in producing a language orally. Other researchers also indicate the importance of output in second language acquisition. Johnson (1995) states the process of producing or performing a language (production-based) is part of successful second language acquisition. Swain (1985) indicates that language learners need the opportunity for output.

Often in second language learning programs, students do not have many opportunities to immerse themselves into a foreign language learning environment. According to American Council on the Teaching of Foreign Languages (ACTFL), foreign language immersion education means to use the foreign language exclusively (use the foreign language 90% of the classroom time) and to use the foreign language to teach subjects to students (ACTFL, 2019). However, students often do not have opportunity to learn foreign languages in an immersed environment. They learn the grammar, vocabulary and other elements of a foreign language, but rarely have time to use the language. For instance, students cannot find appropriate people with whom to have conversations in the foreign language they are learning. Thus, students need more opportunities for the output of a foreign language, and more practice in using foreign languages. In a typical language classroom setting, the time is usually divided into speaking, listening, writing and reading. Once students leave the class, they will not have

the environment or audience to practice speaking and listening, but they can practice writing and reading on their own. Language students, especially adult learners, are often afraid of making mistakes in public. It will be even harder for them to seek opportunities to practice speaking outside classrooms, because they are embarrassed or afraid of inaccurate pronunciation and inappropriate usage of foreign languages (Freeman & Freeman, 2001). If there is an effective method to help students to improve their enunciation of foreign languages, so that their speaking is more clear and understandable, students may be more motivated to learn and use foreign languages.

Selinker (1972) proposed the term interlanguage for describing L2 learners' language learning system. Interlanguage means that L2 learners construct a learning system that is between their native languages and second languages. If the native languages and second languages share similar structure, then L2 learners may use L1 as background or preparation to comprehend new concepts. For example, Italian and Spanish are two languages that share many similarities. Thus, it is not difficult for Italian or Spanish speakers to learn the other language. The grammar structures in these two languages are similar, as is the pronunciation of many words. Therefore, it is quite easy for Italians to speak Spanish, and vice versa.

According to the contrastive analysis hypothesis (Brown, 2014), differences between the structures of the native language and the second language create difficulties in learning. Chinese and English are from two different systems. From the speaking aspect, Chinese is a tonal language, while English is non-tonal. For English speakers to learn Chinese, they do not have much information they could use from their L1 language. Thus, in learning Chinese pīn yīn and the four tones, students need to have an appropriate

preparation or previous knowledge that they can associate with or connect to the new concept correctly. Otherwise, wrong concepts or misunderstandings that apply to the new knowledge will cause errors. For example, in English the tones do not change words' meaning, but convey emotions and attitudes, while in Chinese, tones not only convey feelings and manners, but also more importantly carry different meanings. In teaching Chinese to American students, I find students often do not pay attention to the tone when they pronounce Chinese words because tones are not as important in English as in Chinese. Students often treat the intonations of four tones as if they are not marked. It is challenging for English speakers to pronounce Chinese words in accurate tones. Therefore, providing Chinese learners appropriate and comprehensible input of Chinese tones is crucial for helping them to recognize and enunciate Chinese tones accurately.

Language acquiring is one kind of learning and its process follows general learning principles. Many language concepts are abstract. To better understand and absorb abstract concepts, Pestalozzi, Piaget, and Gordon proposed similar ideas on how to internalize them. Pestalozzi believed that the foundation of all human learning was based upon sense experiencing (as cited in Silber, 1973). He stated:

> Actual sensory contact with things is only the source of correct ideas in so far as I am thereby brought into relation with unchanging qualities as distinguished from the accidental properties of individual objects or particular circumstances. Conversely, misconception only arises when it is these accidents of circumstance, which play the chief part in my sensory experience. The more varied our sensory experience of an object, the more accurate our knowledge of it is

> likely to be. If we put objects side by side which belong to the same class, we shall get a broader and clearer insight into their essential qualities. (pp.103–104)

According to Pestalozzi, humans' innate faculties should evolve as humans learn how to think, proceeding gradually from observation to comprehension, to the formation of clear ideas, and to abstract concepts (as cited in Silber, 1973). Thus, a learning process should be founded on sensory experience. Before learning abstract concepts, students should use various sensory functions to experience the concept. For example, in teaching the register of a sound, such as high and low, one can provide students with high and low sound listening samples. Then, the teacher can give students some percussion instruments that will make high and low sounds. Students will try the instruments and then sort out instruments according to the register. Then students and the teacher will discuss the difference between the two registers. Last, students will listen to some sounds and define whether they are in a high register or low register. Thus, students learn the abstract concept, high and low register, by listening and making sounds first, then gradually processing the input to the final stage of understanding the abstract concept.

Similar to Pestalozzi's belief that learning abstract concepts should start with observation and sensory experience, Piaget (1973) divided learning into stages, including cognitive developmental stages and the formal operation stage. Cognitive developmental stages includes sensory-motor intelligence, preoperational representation, and concrete operations, which are aspects of perceptual learning whereby students learn and

memorize a dominant aspect of an object. Then, the formal operation stage is the conceptual learning process where one's thinking is at a higher level and more complicated. The cognitive developmental stages provide the foundation for a higher level of mental operation, and by the formal operation stage the young person can use abstract concepts and logic to solve problems independently and creatively.

In music learning, Gordon (1980) also suggested that learners should experience concepts before understanding them. To him, the primarily goal of teaching music is to help students to understand music. One of the fundamental elements of music is sound. To understand sound, one should start by hearing and connecting with music aural elements, which Gordon named audiation. Basic audiation provides the foundation of intelligent listening to music and it "interacts with the rote performance and original performance of music, because basic audiation and performance are dependent upon each other" (p. 3). Later, when students have sufficient experiences with music, they are ready to learn the theoretical portion of music and use metaphors to define and conclude music theory.

Gordon (1980) stated that music learning starts from discrimination (aural, verbal association, partial synthesis, symbolic association, composite synthesis) to inference (generalization, creativity, theoretical understanding). Discrimination and inference functions operate concurrently. Inference learning assimilates discrimination learning. Inference learning is more complex and at a higher stage than discrimination, where conceptualization occurs. "Inference learning subsumes all corresponding levels of discrimination learning, such as aural/oral in discrimination learning-generalization aural/oral" (p.12). Thus, according to Pestalozzi, Piaget, and Gordon's theories of

learning, one should gain actual sensory experience, then move on to the stage of understanding and utilizing abstract concepts. Chinese tones carry different intonations, which are similar to pitch intonations. Using pitch singing to experience the intonations of the four Chinese tones may provide a foundation for better understanding of this concept.

Challenges of Recognizing and Enunciating the Four Chinese Tones

Chinese is a tonal language, which is different from non-tonal languages such as English, Greek, or German. The difference is due to the tones that can change meanings of words; while non-tonal languages can use different tones to enunciate a pronunciation, the meaning does not change. “Tones are word length pitch variations that affect word meaning. Tone languages are categorized in terms of register tones, those tones in which contrasts are made among two or more constant pitch levels, or contour tones” (Strong & Plitnik, 1992, p. 229). For Chinese phonetics, Chinese people use a distinctive tool, pīn yīn, to represent the sounds and tones of Chinese characters and to facilitate the pronunciation of Chinese words. Pīn yīn means “spelling and sound” in Chinese and it uses Roman alphabets and tone marks to indicate the pronunciation of each Chinese character (Xing, 2006). Pīn yīn is not a part of Chinese written language. It is a tool to teach people how to read and recognize the sound of Chinese characters. In written versions, such as newspapers and books, pīn yīn does not appear in text, unless the materials are designed for teaching Chinese characters.

Each Chinese character uses one pīn yīn, each composed of three parts: consonant(s), vowel(s) and a tone mark. A tone mark is a kind of diacritics that is added

on vowels. There are six vowels (finals) and 21 consonants (initial) that are used in the pīn yīn system (Chao, 1968; Cheng, 1973).

Vowels: a, o, e, i, u, ü

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

For indicating the tones of each Chinese character, the pīn yīn system uses one tone mark on one of the main vowels (a, o, e, i, u, ü), for example, mián, hōng, bǔ, lèi. The sequence of which vowel in one pīn yīn gets the tone mark is following this order: a-o-e-i-u-ü. If there is vowel "a" in one pīn yīn, then the tone mark is on "a." If there is no "a," but vowel "o" in one pīn yīn, then the tone mark is on "o." For example, if the vowel section is "ia," then the tone mark is added on the main vowel "a." If the vowel section is "ou," then the tone mark is added on the main vowel "o."

In the pīn yīn system, the four Chinese tones are categorized into five pitch levels, with the first tone being high-level pitch, the second tone high-rising pitch, the third tone low-dipping pitch, and the fourth tone high-falling pitch (Chao, 1948). The first tone uses a bar on top of the vowel (mā) to indicate a flat tone; the second tone uses an acute accent mark to indicate the raising tone (má); the third tone uses a V shape mark to indicate the tone goes down first and then up (mǎ); the fourth tone uses a downward accent mark to indicate the falling tone (mà) (see Figure 2).

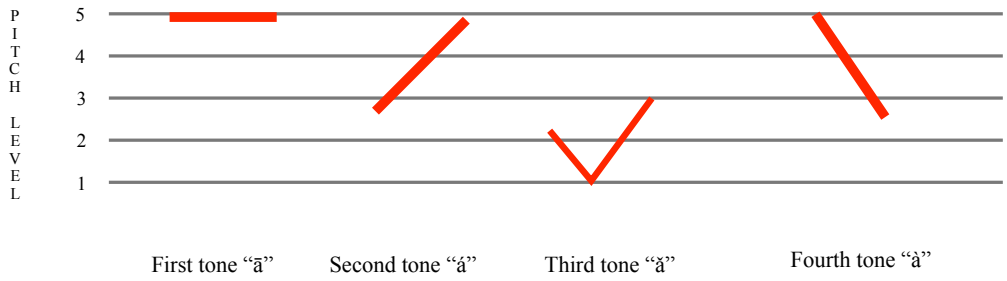

Figure 2. Pitch level of the four Chinese tones (Chao, 1948).

However, not all the individual pīn yīn has a tone mark on the vowel. Some Chinese teaching method books and Chinese linguists address this situation as a neutral tone (Shi, 2018). For example, the word “mother” is composed by two identical characters “妈妈.” The pīn yīn for the two characters are mā ma. The second “ma” does not have a tone mark on the vowel. This “ma” has a neutral tone that does not have a defined pitch. Thus, some scholars and linguists state that there are five Chinese tones in which first four tones have pitch and the fifth one is a neutral tone (Chao, 1948). Since the pronunciation of the neutral tone does not have a pitch, most of the alphabet letters that are used in pīn yīn are pronounced the same as in the non-tonal language. From my Chinese teaching experiences, Chinese learners will not have much difficulty in recognizing and enunciating the neutral tone compared to the four Chinese tones. Therefore, this study focuses on how to help Chinese learners to recognize and enunciate the four Chinese tones more accurately and clearly.

It is challenging for English speakers to recognize and enunciate the four Chinese tones accurately (Bluhme & Burr, 1971; Kiriloff, 1969; 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). Each pīn yīn, a Chinese morpheme, is the smallest meaningful unit in a language that includes both meanings and pronunciations (Wang & Honig, 2010). The basic unit of spoken Chinese is one pīn yīn without a tone mark or syllable, but each one can be related to several different tones to produce different morphemes (Ho & Bryant, 1997), for example, "wǒ – I, wō – nest, wò – lie down." Contrastingly, in English, each syllable corresponds to either one morpheme or several syllables are combined together to equal one morpheme. For example, "happy" is a two-syllable word (hap-py) but has one morpheme.

Each of the four Chinese tones has its own characteristic. 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 fourth one has the highest (Chuang, Hiki, Sone, & Nimura, 1972).

Another reason for the difficulty in acquiring accurate pronunciation of a second language is the interaction of learners' L1 and L2 system (Best & McRoberts, 2003; Brown, 2014; Flege, 1995; Selinker, 1972). Often English speakers use their L1 background to perceive and produce Chinese. In a study investigating the different perceptions of Chinese tones between tonal language speakers 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, Demorest,

Aylward, Cramer, & Maravilla, 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. Tones in English do not change a word's meaning, while Chinese tones do. Therefore, it is important for Chinese learners to set up an accurate perception of Chinese tones and avoid the influence of their L1 background.

Music and Language, and Their Shared Learning Mechanisms

Before a young child acquires the vocabulary and grammar of a language, the mother uses singing speech to communicate with her child. Tafuri (2016) states that the vocal communications between parents and children are more like musical conversations in early childhood. This kind of communication involves all the sound elements, such as dramatic pitch contour, timbre, dynamic, pattern, and sequence. Later, this vocal communication will lead young children to acquire the words and grammar for speech (Tafuri, 2016; Patel, 2008). In addition, singing activities are widely used around the world for parents and early childhood teachers to amuse and comfort babies (Custodero, 2006; Ilari, Moura, & Bourscheidt, 2011), because their dramatic elements convey the meaning and emotions of spoken words, facilitate language and musical development, and strengthen social binding (Trehub, 2001). People who speak different languages often can be affected with a common feeling from listening to a music piece. In a study investigating infants' preferences to song or spoken stimuli (Costa-Giomi & Ilari, 2014), babies were attracted to both singing a foreign language song and speaking the song's lyrics. The study showed that singing had the same power as speaking for attracting

infants' attention. Thus, music and language have been connected traditionally, as sound is a fundamental element for both domains.

Music and language are two domains based on sound. Handel (1989) states that both systems are acquired by humans and therefore they are part of cognition. Music and language both depend on a mental framework of learned sound categories. 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 (Stewart, von Kregstein, Warren, & Griffiths, 2006; Zatorre, Meyer, Gjedde, & Evans, 1996). Some studies indicate music and language share a few mechanisms, while other studies state each domain has its individual system (Patel, 2008). However, these studies only present the different regions of activation; they cannot prove that the two domains do not share some mechanisms. The different activation regions can be analogous to an auto factory that produces both motors and cars. The factory has two different areas for building these two kinds of machines, but it does not mean the two building areas do not share the process of making them (Patel, 2008).

Strong and Plitnk (1992) state that tonal languages require engaging the right hemisphere of the brain for processing, while non-tonal languages mostly use the left hemisphere. In fact, Hsieh and colleagues (2001) found when pitch was used in cuing lexical difference in a tonal language, the left hemisphere was activated. On the other hand, if pitch is not used in recognizing lexical difference, then the right hemisphere is engaged (Zatorre & Gandour, 2008). A tone perception comparing study of Chinese syllables indicated that the left hemisphere is activated for tonal speakers (Klein, Zatorre,

Milner, & Zhao, 2001). Twelve native Chinese speakers and 12 native non-tonal language speakers participated in this study. All participants were asked to compare pairs of Mandarin words and identify these words' tones and meanings as the same or different, as quickly and accurately as possible. While the participants were making comparisons, the PET scan was conducted simultaneously. The PET images show that only Mandarin speakers' frontal, parietal, and parieto occipital regions of the left hemisphere were activated. In contrast, the right inferior frontal cortex of the English group was activated, which was not observed in the Chinese group. This finding confirmed the results of another study comparing pitch processing of Thai syllables by Thai, Chinese and English speakers (Gandour et al., 2000). In that study, the left frontal region was engaged only by the Thai group, while both Chinese and English speakers showed the significant activation in the anterior insula region. Such results were not seen in the Thai group. Thus, these studies demonstrated that when pitch perception or pitch differences are processed linguistically (as a cue for differentiating lexical difference), the left hemisphere is engaged. On the other hand, when the speakers do not use pitch differences as a linguistic function, the right hemisphere region is activated. These studies also indicated 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, Sereno, Jongman, & Hirsch, 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 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. In a more recent study, Nan and Friederici (2013) found "shared neural resources engaged in pitch processing for music and tone language" (p. 2053). In this study, 18 female Chinese musicians (adults) were recruited from the China Conservatory of Music. Eighty quadrisyllabic Chinese phrases and 80 four-note musical phrases were used as stimulus materials. In the Chinese sentence stimuli, half of the sentences (n = 40) were incongruous, which means the sentences were either semantically incorrect or syntactically incorrect or both semantically and syntactically incorrect. The other 40 Chinese sentences were congruous. Similar to the Chinese sentence stimuli, half of the music phrases (n = 40) were composed incongruous, which means the interval between two notes was dissonant. The 18 participants were asked to listen to all the stimuli and identify each stimulus as congruous or incongruous. While the participants were taking the test, functional magnetic resonance imaging (fMRI) was conducted on each participant. Results indicated that pitch processing in music and tonal language share neural circuits in part of the right superior temporal gyrus (STG) and the pars triangularis of Broca's area. Thus, this study confirmed the positive transfer effects between music pitch processing domain and linguistic tone processing domain (Magne, Schön, & Besson, 2006; Moreno et al., 2009; Schön, Magne, & Besson, 2004; Wong, Skoe, Russo, Dees, & Kraus, 2007). This study

also explains the positive impacts of music training on language learning. Because both music pitch and language tone processing share the same neural network, one domain is strengthened, and the benefits could be transferred or shared to the other domain.

Chinese tones share many musical features with the musical element pitch: “Pitch is defined as having a particular quality of a sound that fixes its position in a scale” (Haynes & Cooke, 2017, p. 1). Pitch is presented by combining a frequency value and a note name (Haynes & Cooke, 2017). Compared to the musical features of Chinese tones, pitch can carry similar characteristics, such as high and low sounds, duration, and volume. Thus, the characteristics of Chinese tones exhibit a close relationship with musical pitch. A few studies examined the influences of musical ability on identifying Chinese tones and pronunciation. The results all indicated positive effects of musical ability on tone recognition and differentiation (Cooper & Wang, 2012; Delogu, Lampis, & Belardinelli, 2006; Lee & Hung, 2008; Marie, Delogu, Lampis, Belardinelli, & Besson, 2011).

In a study investigating the influence of musical expertise on segmental and tonal processing in Mandarin, non-tonal language (French) musicians and non-musicians were recruited (Marie et al., 2011). None of the participants had prior knowledge of Chinese. Participants were provided with one warm-up training session for understanding the tonal feature of Chinese, its four Chinese tones, and the Chinese consonants and vowels. Then, the musician group and non-musician group were each presented with a sequence of four words orally and were asked to differentiate the segmental variations in Chinese words and tones. Results indicated the musician group performed better on detecting tonal differences and segmental variations than the non-musician group.

In another study (Cooper & Wang, 2012), the effects of musical ability and tonal language (Cantonese) tone identification were compared. The study recruited native tonal language (Thai) and non-tonal language (English) speakers. Participants were divided into four groups: musician (Thai), non-musician (Thai), musician (English), and non-musician (English). These participants did not speak Cantonese, and they were given a training session for understanding the five distinguished tones of Cantonese words. The researchers pointed out that the musical ability was found to have a greater effect than tonal language background on tone identification. Also, test results indicated that the musical ability was significantly correlated with tone identification for English speakers, but not for Thai speakers. These results indicate that musical expertise affected how non-tonal language speakers process the tonal language as well as how they categorize tonal linguistic features. These studies present positive transfer effects from music training to tonal languages and open new perspectives for language teachers in teaching tonal languages.

Singing Training Effects in Foreign Language Learning

Many studies have reported various benefits that singing has on foreign language learning, especially in the first language learning stage (Rahbar & Khodabakhsh, 2013; Shen, 2009). Also, singing has a great effect on learning motivation and acquiring language structures because students need to segment new words (Schön et al., 2008). Mora (2000) states that songs can help students to retrieve the lexical patterns stored in long-term musical memory easily with mental rehearsal, with memorization, and with conversation. Research suggests that pop songs are useful for not only remembering and recognizing vocabulary, but also for forming phrases, producing longer sentences and

promoting conversations (Wray & Perkins, 2000). Most speech is the repetition and variations of memorized language patterns. Therefore, most songs match this feature and can facilitate the acquisition of a language.

In a study investigating the effectiveness of using English lyrics to improve the listening comprehension ability of adult English learners in Iran (Rahbar & Khodabakhsh, 2013), results indicated that the mean score of the experimental group, which was trained by singing English songs, was much higher than the control group, which was trained without singing English songs. Also, similar effects were found in an examination of young students' listening comprehension and pronunciation skills in learning English in Iran (Ghanbari & Hashemian, 2014).

Furthermore, another study conducted in Jerusalem, Palestine, examined the effects of using children's songs in learning English vocabulary and improving pronunciation (Shehadeh & Farrah, 2016). In this study, there were 123 fourth grade EFL (English as a Foreign Language) students; 72 students were female and 51 were male. Their native language was Arabic. The participants were divided into four groups: two groups of girls and two groups of boys. One group of boys and one group of girls were the experimental groups who practiced the English vocabulary and phonemes with English children's songs. The other two groups learned English without any English children songs. The results of vocabulary and pronunciation posttests showed a significant difference between the experimental groups and the control groups. In both tests, the experimental groups' mean scores were much higher than the mean of the control groups. In conclusion, these studies suggested that singing songs has more positive effects on learning English than without using songs.

Furthermore, in examining the relationship between musical aptitude and second language pronunciation skills of 40 students from ages 10 to 11, Milovanov, Huotilainen, Välimäki, Esquef, and Tervaniemi (2008) showed that children with good linguistic skills performed better on musical aptitude tests than children who had less accurate linguistic skills. In this study, 40 children were recruited from four elementary schools in the Turku district in southwestern Finland, and these students had little knowledge of English. Before taking an English pronunciation skill test, the children received an eight-week English phonemic discrimination training. After the training, students took a pronunciation skill test, and they were divided into two groups. One group had an English pronunciation accuracy rate over 60%, while the other group had an accuracy rate under 60%. After the English test, both groups took the Seashore Musical Aptitude Exam. Results confirmed that the group who had the higher English pronunciation accuracy rate also got a higher score in musical aptitude test. Thus, these results suggest that musical skills have great effects on language pronunciation skills. Also, these findings shed light on the benefits of using songs or singing in foreign language teaching, especially in speaking and listening.

What are the common teaching methods that Chinese language teachers currently use in America? Since Chinese tones have a close relationship with pitch, and because previous studies have indicated the benefits of applying singing training in foreign language learning, do Chinese language teachers use singing training activities to help students to learn Chinese tones?

Pitch Singing in Teaching Chinese Tones

Previous studies indicate the close relationship between music and language learning and the shared mechanism of music and language in human brains. Also, many studies show that singing has a great effect on various aspects of learning English as a foreign language, such as pronunciation and the retention of vocabulary. Moreover, musicians performed better on Chinese tone recognition and pronunciation tasks than non-musicians. These studies indicate that pitch singing training may have a great effect on tonal language learning, especially on tone differentiation and pronunciation. After searching through the current and available Chinese tones teaching methods in the US and Italy market, I did not find one that uses pitch singing training methods to help learners in recognizing and enunciating the four Chinese tones. Chinese language teachers often use audio-lingual methods, visual aids, hand movements and a few computer-based programs that help students to learn the four Chinese tones.

The audio-lingual teaching method is based on sound imitation. Students learn pronunciations from imitating teachers' speaking or sound from audio recordings. However, as indicated previously in the literature review, non-tonal language speakers have difficulties understanding the dramatic Chinese tones' intonations because they are different than those in non-tonal languages. Also, non-tonal language speakers' background negatively influences on recognizing and enunciating the Chinese tones. Learning the four Chinese tones by sound imitation cannot help students to establish the accurate Chinese tone concept in their minds, nor can it clear the negative influence that non-tonal language speakers' background brought into their learning of Chinese tones.

The visual aids that are often used in Chinese classes are images and symbols that can illustrate the contour and intonations of the tones. The diacritical marks along with the five pitch levels' tone chart that Chao (1948) developed well represented each Chinese tone's sound shape. Chinese teachers often use these two visual aids to explain to students the intonations of Chinses tones. In China, when teaching Chinese tones to students, teachers often ask students to use their hands or fingers to trace the four tone marks. For example, when students enunciate the first flat tone, they will use their hands or fingers to draw a flat line while they are enunciating the tone. I have also found benefits in using the visual aids. However, understanding the concept is a part of a learning process; to recognize and produce the Chinese tones accurately is another procedure. Students may understand the intonation clearly, but to perform the Chinese tones, vocal exercises are needed in order to get non-tonal speakers' vocal muscles used to producing these sounds.

Computer-based teaching programs often mix the previous three teaching methods together to provide students various learning approaches. However, these programs are often expensive. The current online Chinese learning program (Integrated Chinese level 1 & 2 by Cheng & Tsui) that is used at the university where I teach costs $170 USD for a one-year subscription. The popular language learning computer program Rosetta Stone is $180 USD for a one-year subscription. These programs are not affordable to many language learners. Moreover, these two teaching programs do not provide the necessary exercises to extend the vocal range that Chinese language learners need, and the method of mastering the Chinese tones' pronunciation is founded on the appearance or production of the sound. Students learn Chinese tones by imitating and

correcting the pronunciation of tones when they receive feedback. However, the core concept of enunciating the Chinese tones accurately is using dramatic pitch intonation, which is not clearly presented and established in computer-based programs. Therefore, in addition to the common teaching methods (audio-lingual, visual aids, movements), I wonder if there are any other effective and economic training methods for mastering the Chinese tones. Since Chinese tones have a close relationship with the musical element pitch, it is not clear why the pitch singing training method is not applied in learning Chinese tones.

One possible reason based on previous studies on Chinese tones could be that insufficient empirical studies were found to compare the effects of music training on the learning of Chinese tones to other training methods. Although there are many studies focused on the effect of songs or singing on language learning, most the studies focused on non-tonal languages, such as English. In addition, 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).

There are some teachers who have tried to apply musical elements in teaching the Chinese tones, but the methods were showed as not practical in classroom teaching (Shi, 2018). Since the first Chinese linguist Chao (1948) had defined the four Chinese tones into five pitch levels, a few linguists and teachers have developed different music training methods in teaching the four Chinese tones. Zhao (1987) developed a vocal-cord training method to help students to differentiate the various registers of the Chinese tones. Zhao explained how to produce the Chinese tones by using different levels of vocal cord tension. The first tone is produced by keeping vocal cords intense. The second tone is

produced by not keeping the vocal cords too tense nor too relaxed. The third tone is produced by making vocal cords intense-relaxed-then tense. The fourth tone is produced by making the vocal cords suddenly intense and then to relax gradually. However, this training method is quite abstract to learners who are not singers, nor who have good understanding and control over their vocal cords (cited in Zhang, 2006, p. 54). Shen (1989) proposed an idea of using high, low and middle pitches to train Chinese language learners to understand and master the four Chinese tones. Shen (1989) suggested that the first tone can use a sustained (equal to two music notes' time) high key to present the pitch contour; the second tone can use a mid-key, then rise to the high key to show the contour; the third tone can start from a mid-key then goes down to a low-key and comes back to the mid-key to present the contour; and the fourth tone can start from a high-key then drop to a low-key. However, Shen did not apply the idea into practice, and did not examine the effects of this idea.

Among previous studies and trials, no study has been found to apply a pitch singing training method in teaching Chinese tones, to examine the effects of pitch singing training on recognizing and enunciating the four Chinese tones, or to compare the effects between pitch singing training and other common teaching methods. The most recent study that Shi (2018) conducted for examining the effectiveness of using musical elements in teaching the four Chinese tones did not compare the results between common teaching methods and a music related-teaching method. In the study, Shi compared the pre-test result with the post-test result from the participants. However, all the participants took the same training method that involved musical elements. The "opera voice" was used to refer the high pitch and tones in the high register, a normal speaking voice was

used to refer the middle pitch and tones in the middle register, and a hesitate “ugh” sound was used to refer the low sound and the tones in the low register. In addition, a visual tone map (similar to Figure 2, presented earlier) was used to present the four Chinese tones. Results indicated that all participants learned the four Chinese tones effectively, but there was no data comparing different teaching methods. It was not clear if this teaching method had more positive effects than common training methods or if her training method had the same effect as other training methods. Therefore, in this study, I am interested in examining the effects of pitch singing training on the learning of Chinese tones. If the results of this study indicate positive effects, then the pitch singing training method could be an alternative way to help learners to master the four Chinese tones.

Summary

Acquiring a foreign language successfully depends on many aspects, such as learners’ L1 background, learners’ L1 and L2 language differences, and learners’ learning environment and methods. According to the Foreign Service Institute’s School of Language Studies (2018), Chinese is considered a difficult foreign language to learn:

> The Foreign Service Institute has classified various languages into four groups according to their level of difficulty for native speakers of English. Romance languages such as Spanish, French, and Italian fall into Group I, the least difficult for native English speakers. Languages that are character-based and those whose structures are quite different from English, such as Chinese, Japanese, Korean and Arabic, fall into Group IV. Students learning these languages will require approximate 2200 class hours to meet the reading and writing goals, and they will

also face a greater challenge as they become culturally competent in the societies where these languages are spoken.

Thus, mastering Chinese requires more effort in every respect. The intonation of Chinese tones is a new concept to non-tonal language speakers. Therefore, it is challenging for non-tonal language learners to recognize and enunciate them accurately. Non-tonal language learners need to first build a network or background to connect the new concept to their existing knowledge. The four distinct tones of Chinese share many characteristics with musical pitch, such as pitch range and duration. Previous studies of how humans process music and language revealed that these two domains share neural networks. Music training has positive effects on language learning, especially in tone recognition. Moreover, singing has been widely used in language learning, and studies have indicated many benefits that singing can bring to language learning, especially in speaking. Thus, pitch singing training may help non-tonal language speakers to master Chinese tones more effectively.

In doing this study, I sought to examine the benefits of musical training on learning Chinese tones, to investigate the effect of pitch singing training on two important language outputs (namely, recognizing and enunciating the four Chinese tones), and to compare the effects of pitch singing training method with the traditional audio lingual method. 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 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 learners?

Q3 Does the pitch singing training have a greater effect on the accuracy of Chinese tones' recognition and enunciation than the traditional audio-lingual method?

CHAPTER III

METHOD

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 an American university with selected American college students.

Participants and Setting

This study was conducted at the University of Northern Colorado (UNC) main campus in Greeley, with the approval of the UNC Institutional Review Board (Appendix A). I first presented my study to the language classes (German, French, and Japanese classes), elementary education classes, and statistics classes at UNC. Also, I posted flyers to advertise this study in the School of Music building. In total, 66 UNC college students (both undergraduate and graduate) who neither speak Chinese nor had studied Chinese for more than one semester before this study were recruited. Two participants were over 50 years old; the rest of the participants' ages ranged from 18 to 26.

Among the 66 participants, there was one Japanese student, four Thai 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 for more than one semester or learned Chinese before participating in this study. There were also six music major students who participated in this study.

According to the Asia Society Center for Global Education and the American Council on the Teaching of Foreign Languages (ACTFL), Chinese is classified in the Level IV difficult foreign language learning group. ACTFL divided language learning in four skills: speaking, listening, reading and writing. Each skill is divided into 10 levels from novice-low to superior. 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, Lien & Spring, 2016). Moreover, according to ACTFL's foreign language speaking proficiency standards, the novice-low level in speaking is defined as "…no real functional ability because of their pronunciation…," while the novice-mid level in speaking is defined as "…communicate minimally by using a number of isolated words and memorized phrases limited by the particular context in which the language has been learned" (ACTFL Proficiency Guideline, 2012). Since none of the 66 participants had studied Chinese for more than one semester or 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 ability for learning languages.

The participants were randomly assigned to one of two training methods: one is the pitch singing training method, and the other one is the traditional audio-lingual mode. Based on the students' schedules and their availability, they were grouped in small sizes with each group having one to four people. I divided participants evenly to each training

method. The training sessions were scheduled mostly either before or after students' classes where I went to present the study. Some groups were scheduled in evenings. The participants needed to attend all eight training sessions. If the participants missed a training session, they were asked to schedule a make-up session with the instructor before the next training session started. Data were collected from only those participants who completed all eight training sessions. Thus, at the end of the training sessions, 60 valid participants' data were used.

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). According to their study, the number of training sessions for achieving 95% success in using pitch pattern 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. Also, 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 two to three weeks with each week having three to four 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 Chinese language and pronunciations of pīn yīn since none of the subjects had any previous knowledge of Chinese language. The introduction of pīn yīn and tones were given in a normal speaking mode with visual aids for all the participants, and no pitch singing exercise was involved. The visual aid is presented in Figure 3.

First Tone Second Tone Third Tone Fourth Tone

Figure 3. Tone marks chart.

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

Materials

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. To indicate the music notes' pitch and range 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 are in (see Figure 4). The lowest octave notes use number 0 along with the seven letters (C0, D0, E0, F0, G0, A0, B0) to indicate the

register. Larger numbers correspond to higher octaves. The middle C octave uses number 4 for each note in that range, such as C4, D4, E4, F4, G4, A4, B4.

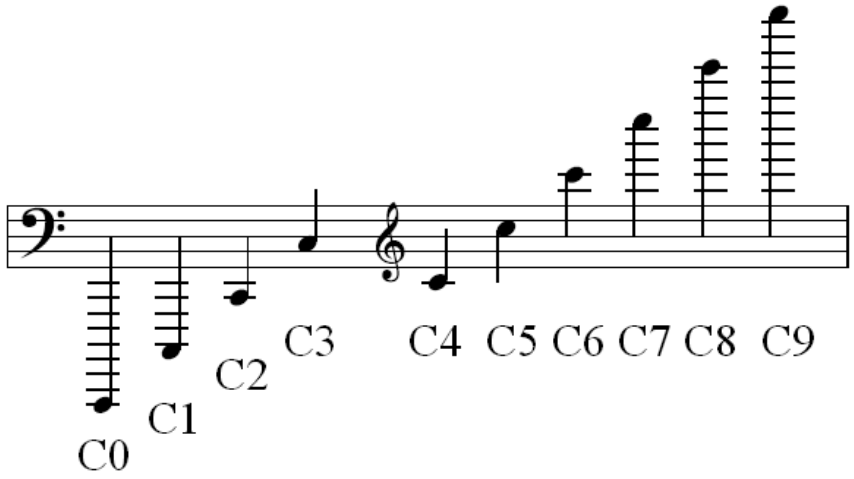

Figure 4. Music notes range.

After examining the pitch and register of the four Chinese tones that I spoke, the results show that the first tone is close to singing a sustained note G4 (as a unison interval), the second middle rising tone is similar to singing notes from E4 to G4 or G4 sharp (as a third interval), the third low falling, then rising tone is close to singing notes from C4 to lower B3 then go up to D4 (as a second interval plus a third interval), and the fourth falling tone is close to singing notes from G4 to lower F4 (as a second interval).

Using these 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. An interval depicts the distance between two pitches. Patel (2008) stated that 2-semitones intervals, for example note C to D, are commonly used in musical melodies. In speech melodies, 1-semitone intervals, for example note C to C sharp are

commonly used, and this usage of small intervals applies to languages generally. Thus, using small pitch intervals to present the four Chinese tones is appropriate.

Because each person's voice range is different, I interpreted the four Chinese tones in the following intervals and transposed the intervals depending on the voice range of the Chinese learners. It can be compared to singing a song in different keys. Depending on the voice range of a singer, the singer will choose an appropriate key to sing a song. To make transposing the intervals convenient and clear, I composed the four Chinese tones' intervals in a major scale system. The first tone is a unison interval that starts from the fifth degree of a major scale. For example, if the scale starts from the note D (D-E-$F^{\#}$-G-A), then the first tone is sung from note A to A (the same note). The second tone is a third interval that starts from the third degree of a scale and finishes at the fifth degree of a scale. For example, if the scale starts from the note A (A-B-$C^{\#}$-D-E), then the second tone is sung from note C sharp to E. The third tone is composed of two intervals: one is a downward second interval and an upward third interval. The first part of downward second interval starts from the first degree of a scale and goes one note down. The second part of upward third interval starts from the previous note then goes to the second degree of a scale. For example, if the scale starts from C (B-C-D-E-F-G), the third tone starts from C, then goes down to the lower B, and ends with D. The fourth tone is a downward second interval that starts from the fifth degree of a scale and ends with one note down. For example, if the scale starts from E (E-$F^{\#}$-$G^{\#}$-A-B), the fourth tone is sung from note the B then goes down to A.

Based on Moore's (1991) study, the participants' (who are teachers) vocal ranges are from E3 to D5. Similar to his finding, Geringer and his colleagues (1980) found the

undergraduates' vocal ranges are from E3 to D5 too. Thus, from these studies, one can conclude that adults' singing ranges usually span two octaves and the range is from E3 to D5. Moore (1991) also states that the female adults' mean vocal ranges are from E3 to F5, while male adults' mean vocal ranges are lower, from F2 to F4. Since the participants in this study are all adults, using notes from A3 to G4 (A3-B3-C4-D4-E4-F4-G4) is a comfortable vocal range for them; therefore, I used the following pitch interval singing patterns for the four Chinese tones in the experimental group. For female participants, the first tone used note G4 to G4, the second middle rising tone used notes from E4 to G4, the third low falling rising tone used notes from C4 to lower B3 then go up to D4, and the fourth falling tone used note from G4 to lower F4. For male participants, each tone was sung one note lower, the first tone was F4 to F4, the second tone was D4 to F4, the third tone was B3 flat-A3-C4, and the fourth tone was F4 to E4 flat. The musical intervals that were used for four Chinese tones for female and male participants are presented in Figure 5.

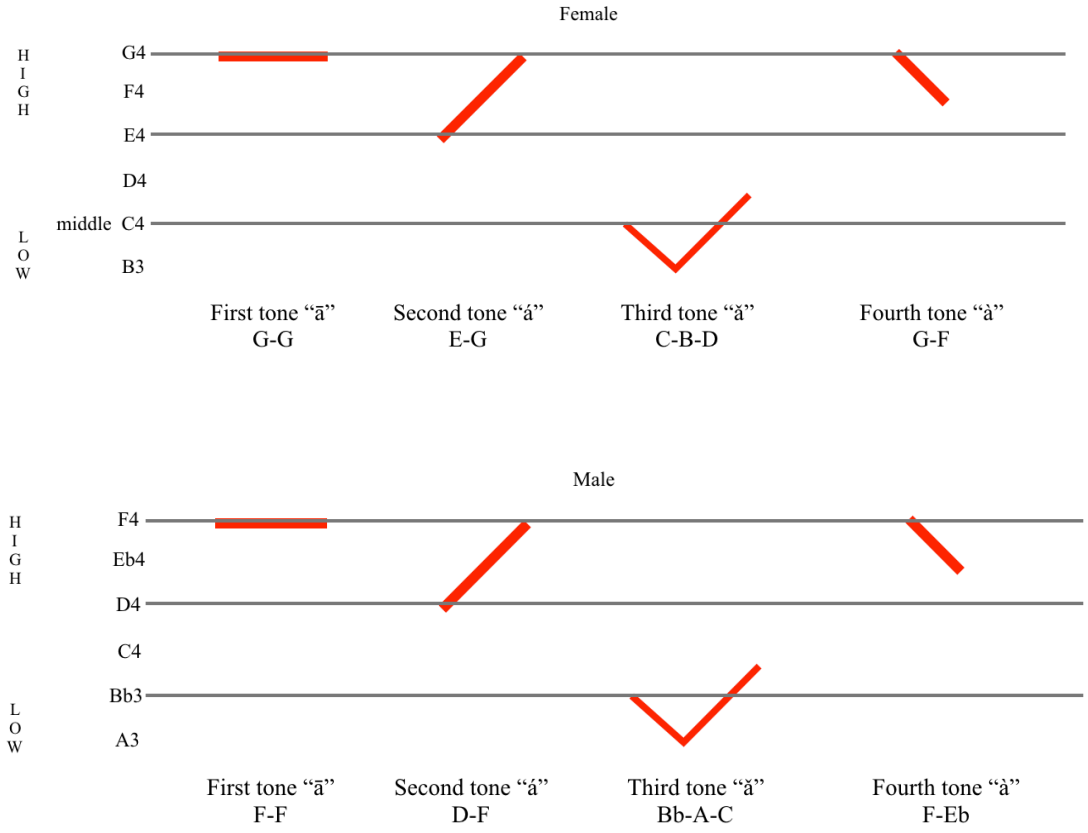

Figure 5. Music intervals for the four tones for female and male learners.

Material for Introduction Session

In the introduction session, I first explained the purpose of learning pīn yīn system. Then, the components of pīn yīn was presented, which each pīn yīn is composed of three parts: consonant(s), vowel(s) and a tone mark. A tone mark is a kind of diacritics that is added on vowels. There are six vowels (finals) and 21 consonants (initial) that are used in the pīn yīn system.

Vowels: a, o, e, i, u, ü

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

Then, the tone marks chart was presented in front of the groups. I pronounced four tones' intonation with each vowel along with the pīn yīn 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; ō, ó, ǒ, ò; ē, é, ě, è; ī, í, ǐ, ì. Then, I taught the pronunciation of 21 consonants to all the participants in normal speaking mode as well. Last, combinations of vowels and consonants were taught to the participants.

Pretest and Posttest Materials

The pretest contained three parts, namely, recognition, enunciation, and singing test. The recognition section was listening to 20 Chinese characters' sounds with a three-second pause between each pronunciation. Each tone was presented with five different vowels in pīn yīn (see Table 1).

Table 1

Pretest Listening and Reading Materials

Listening part					
First tone	tā	bō	hē	xī	wū
Second tone	bá	guó	rén	lí	yú
Third tone	mǎ	wǒ	kě	nǐ	dǔ
Fourth tone	là	mò	è	zì	fù
Reading part					
First tone	mā	dī	bō	kē	tū
Second tone	má	bó	ké	dí	tú
Third tone	mǎ	bǒ	kě	dǐ	tǔ
Fourth tone	mà	bò	kè	dì	tù

The material for the recognition sections of both the pretest and the posttest were 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 materials are presented in Appendix B.

Each pīn yīn was presented three times. The order of these 20 pīn yīn was generated randomly by using the Excel program. 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's recognition part was delivered in groups, and it lasted about five minutes. The second enunciation portion was asking participants to read 20 Chinese pīn yīn. Each tone was presented with five different pīn yīn. The order of these 20 pīn yīn was generated randomly by using the Excel program as well. The second enunciation portion and the third singing part were completed individually, and each participant's enunciation of tones was recorded on their computers or cell phones and then sent to the researcher by email for later analysis.

According to the previous studies, musicians can identify the Chinese tones more accurately than non-musicians. Therefore, a singing accuracy test (the third part of the pretest) was given to all the participants for examining the possible influences that the singing skills 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 to choose this song to test whether singing skills have effects on recognizing and

enunciating the Chinese tones are because this song is known by many people and it includes all the tonal features that are used in the four Chinese tones.

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. There were 20 pīn yīn in total with ten learned and ten new ones.

Eight Training Sessions 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, 10 to 15 new Chinese words were taught to the participants with each tone appearing in at least two pīn yīn, such as tone 1, tā-he/she, zhōng–middle; tone 2, rén–people, guó–country; tone 3, wǒ–I, nǐ–you; and tone 4, shì–verb be, jiào–call. All the training sessions were given by me, a native Chinese speaker who has taught the Chinese 101 and 102 beginner level courses at UNC for three years. I have also received the world language teaching certification from UNC and have two master's degrees, one in music performance and the other one in music education. Thus, I am confident to present the four Chinese tones in accurate pitch levels. For the pitch singing group (experimental group), I delivered the pronunciation of four tones in pitch interval singing first, and then the participants repeated the tones in pitch interval singing. This sound imitating exercise

was repeated twice. Then, I pronounced the pīn yīn in a normal speaking mode, and the participants repeated it in the same way. This exercise was repeated twice.

On the other hand, for the audio-lingual group, I taught the pronunciation of four tones in normal speaking mode without any pitch intonation involved. All the participants in the control group first listened to my pronunciation, then repeated the pīn yīn in a normal speaking mode. This exercise was repeated for four times. Thus, both groups received the same amount of training time on each pīn yīn. After the pīn yīn instruction, a quick assessment of tone enunciating was given to all the participants to evaluate their learning results. The participants were asked to read aloud each pīn yīn and I corrected participants' pronunciations either in pitch singing mode or normal speaking mode depending on their assigned training method group. Last, all the new pīn yīn in one session were grouped together to make sentences and short conversations. The individual training structures of both kinds of instruction are presented in Appendix C. Lesson materials of training sessions for control and experimental groups are also listed in Appendix D.

Scoring

The pretest and posttest scores were calculated as follows: The recognition section was calculated as correct (1 point) or incorrect (0 points) for each pīn yīn's tone identification, and the enunciation section was calculated as accurate (1 point) or inaccurate (0 points) for each tone's articulation. The result of the enunciation section was graded by two native (one female and one male) Chinese speakers from UNC to minimize personal bias.

The section for recognizing the four Chinese tones in the pre- and posttests was designed as a multiple-choice questionnaire. Participants chose a number from one to four to indicate the tone they had heard. There were total 20 questions in this 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 pre- and posttests was constructed as asking participants to enunciate the pīn yīn tones that they had read. The results were recorded for later data analysis. There were total 20 questions in this section as well, with each question worth 1 point for a correct enunciation. 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). It is impossible to ask the participants to only pronounce the tone without saying the letter. Thus, the scores were calculated only based on the accuracy of tones' enunciation, and mispronunciation of the letters were not counted. 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 Chinese speakers from UNC. 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 singing recordings 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 skip 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 UNC. 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.

After collecting all the data, I sought to examine the possible benefits of musical training on learning Chinese tones, to investigate the effect of pitch singing training on two important language outputs, namely, recognizing and enunciating the four Chinese tones, and to compare the effects of a pitch singing training method with the traditional audio-lingual method through my analysis of the results.

CHAPTER IV

RESULTS

In this study, 66 American college students who had not taken Chinese course for more than one semester or did not speak and write Chinese at all before this study were recruited for participation. Six participants' data were not used in the data analysis because they did not complete all the sessions. Thus, at the end, 60 valid data were used in examination. I divided the participants evenly and randomly into two groups. One (control group) was taking the traditional audio-lingual training method, and the other group (experimental group) was taking the pitch singing training method. Both groups were first given an introduction session of the Chinese pīn yīn system and the sounds of the four Chinese tones. Then, all the participants took a pretest immediately after the introduction session. The pretest was comprised of three sections: recognizing the four Chinese tones, enunciating the four Chinese tones, and a singing-in-tune test. After completing the introduction session and the pretest, all the participants took eight training sessions with the method to which they were assigned, and completed a posttest at the end of the eighth training session. The posttest was comprised of two sections: recognizing and enunciating the four Chinese tones.

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. Two independent graders gave a score for each section. None of the graders personally knew any of the participants whom they heard in the recordings. To check reliability, Cronbach's alpha

was used to examine the consistency between the two graders' scores for each section. In this study, the coefficient result between the two independent graders of singing-in-tune test indicated very strong agreement ($\alpha = .97$). The coefficient results between the two independent graders of enunciating the four Chinese tones was also strong ($\alpha = .99$ for the pretest and $\alpha = .99$ 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 purpose of this study was to examine the effects of musical training on Chinese tones learning and to compare the effects of two training methods on the accuracy of recognizing and enunciating the four Chinese tones. By comparing the gain scores between pre- and posttests, I investigated which method had a greater effect on learning the four Chinese tones. 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 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 learners?

Q3 Does the pitch singing training have a greater effect on the accuracy of Chinese tones' recognition and enunciation than the traditional audio-lingual method?

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. 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 on each of those and analyzed the differences.

Results for pretest and posttest scores are presented in Table 2. For the pitch singing method group, the mean difference of the recognition test score was 1.97 and the standard deviation was 3.41. The mean difference of their enunciation test score was 3.33 and the standard deviation was 2.78. For the audio-lingual group, the mean difference for recognition was 0.07 and the standard deviation was 2.20. Their mean was 1.33 and the standard deviation was 2.44 for enunciation.

Table 2

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)*

To check the MANCOVA test's multivariate normality assumption, I conducted multivariate normality distribution tests on the residuals of the gaining scores of recognizing and enunciating Chinese tones. The descriptive data of the residuals on recognition and enunciation are reported in Table 3. 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 results are greater than $\alpha = 0.05$.

Table 3

Descriptive Data of the Residuals of Gain Scores on Recognition and Enunciation

	N	Mean	*SD*	Median	Min	Max	Skew	Kurtosis
Recognition	60	-1.201	2.838	-0.115	-7.142	8.144	0.371	0.784
Enunciation	60	7.037	2.590	-0.289	-4.164	6.836	0.438	-0.429

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. 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 results indicated that the singing-in-tune score had no effect on the recognizing and the enunciating scores in the presence of the factor. The MANCOVA results also 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 singing intonation skills, $F(2, 56) = 6.158$, $p = .004$, Wilks' $\Lambda = .819$, partial $\eta^2 = .180$. (Similar results were found using the other multivariate tests, with Pillai's Trace = .180, and Hotelling's Trace and Roy's Largest Root both = .220.)

Subsequent ANCOVA testing results also 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 6.

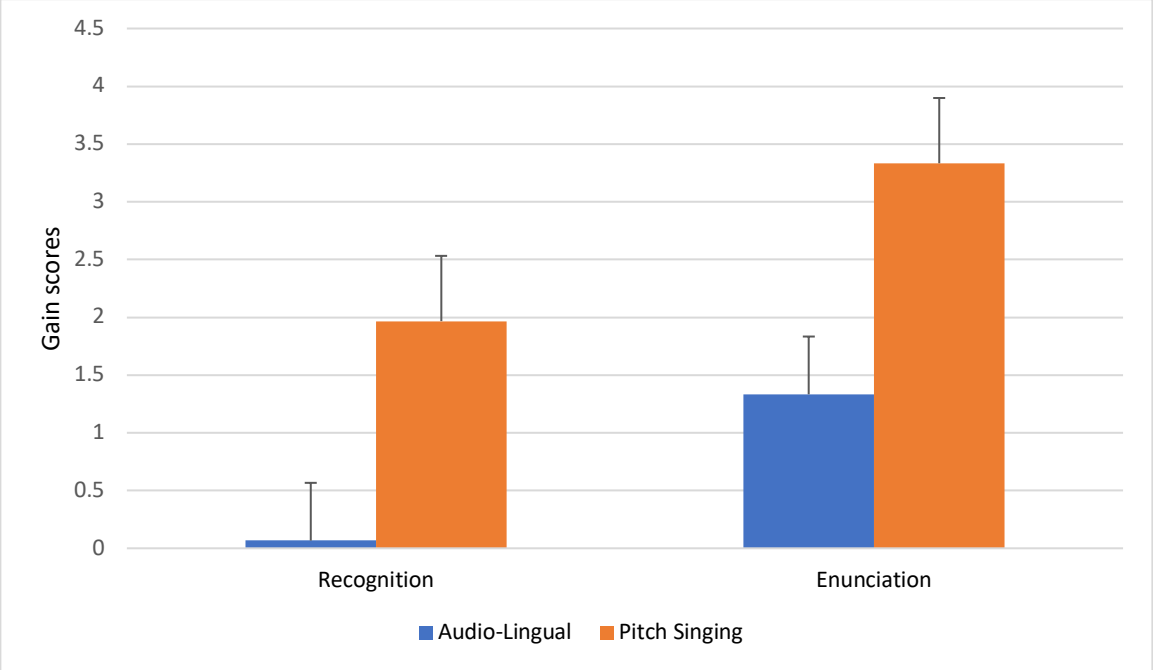

Figure 6. Gain scores for recognition and enunciation tests for audio-lingual and pitch singing groups. Error bars indicate standard error.

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 $t(49.536) = -2.566$, $p = 0.013$ (2-tailed). Also, the pitch singing training group performed significantly better on the enunciation than the traditional audio-lingual group $t(58) = -2.959$, $p = 0.004$ (2-tailed).

CHAPTER V

DISCUSSION

Learning Chinese has become more popular in recent years. According to the Modern Language Association's (MLA) survey in 2007, Chinese language course enrollment rate for American college students increased 20% from 1998 to 2002, and from 2002 to 2006 it increased by 51% (Wang & Honig, 2010). Despite the popularity of learning Chinese, Chinese is not an easy language to learn, especially for non-tonal language speakers. Chinese language's orthographic system is totally different from English.

In Chinese, the characters are not presented in letters, but rather in logograms. The sound of each character is not expressed with the logogram, and the pronunciation of words is learned separately. Chinese people learn the sound of the characters from the sound and spelling system that is called pīn yīn. A salient difference compared to English is that Chinese tones change meanings. Each Chinese character's pronunciation carries one tone. There are four different tones and each one presents a specific meaning of a Chinese word.

In learning Chinese, the most challenging tasks are learning to write characters and accurately speaking the tones (Wang & Honig, 2010; Huang, 2000). In Wang and Honigs' (2010) study, the greatest difficulty for non-Chinese children to speak Chinese was enunciating the Chinese characters with accurate intonations. Students often speak Chinese in a toneless way. Wang and Honigs' study confirmed others' findings that

American Chinese language learners are often not aware of the important meaning that the Chinese tones carry (Shi, 2018; Huang, 2000). They tend to ignore and speak Chinese without the tones. 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.

In the previous chapter, the data were analyzed according to the three research questions. In this chapter, results will also be discussed in detail.

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 conducted a multivariate analysis of covariance (MANCOVA), with the singing-in-tune scores used as a covariate. Results indicated that singing-in-tune did not affect the performance 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. However, many studies indicated that musicians who had refined ears could distinguish the difference in sound and intonations (Schütze, 2017), and they could also recognize the four different Chinese tones better than non-musicians (Cooper & Wang, 2012; Delogu et al., 2006; Lee & Hung, 2008; Marie et al., 2011). However, in this study, the results showed the skill of singing in tune does not influence the performances of recognizing and enunciating the four Chinese tones.

Music skill is a broad concept, and it includes various specific skills such as listening, singing, and reading musical elements. The indication of singing-in-tune ability having no influence on the performances of the Chinese tones could be affected because participants did not associate pitch and the Chinese tones' intonations. Since all the participants either did not speak Chinese or had not taken Chinese more than one semester 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, the influence of singing in tune skill was not strongly presented in the posttest.

Also, in this study, the participants were not grouped as musicians versus non-musicians. There was not enough data to compare the different performances between such groups. There were only six music major students who participated in this study. This result was limited in suggesting that music skills do not have effects on recognizing and enunciating Chinese tones because each music skill needs a specific and appropriate test to evaluate. However, even with such tests, one cannot truly give an evaluation of a music skill because the evaluation cannot be fair to everyone and cover every aspect of music skill; it can only provide a reference. For example, the singing-in-tune test that was used in this study had participants sing "Mary Had a Little Lamb." The scores were given based on how accurate the specific music intervals were sung by the participants.

Although “Mary Had a Little Lamb” is considered to be a well-known song for many people, one can claim that he or she can sing better in other songs. Moreover, some people may state that they are not good at singing, or do not have adequate experiences to exercise their singing skills. Thus, their singing intonation skills may not have been truly presented.

Lastly, singing in tune is only one aspect of music skill. In this study, the result can only suggest that singing intonation skill did not have an influence on the performance of recognizing and enunciating the four Chinese tones but cannot suggest that other music skills or musicianship in general do not have an impact. The purpose of using the singing-in-tune score as a covariate was to account for possible effects of this skill on the performance of recognizing and enunciating the four Chinese tones.

Since all three research questions are closely connected to each other, they are going to be discussed together. I used MANCOVA with singing-in-tune scores used as a covariate to examine if there is significant difference between the two groups. The two dependent variables in this study were recognizing scores and enunciating scores. I compared the difference between pretest and posttest scores on each of those. Variables for group comparisons were pitch singing training group participants (n = 30) and audio-lingual instruction participants (n = 30). According to the results, the effects of the two teaching methods are significantly different from each other. Then, to further confirm which training method had a greater effect on recognizing and enunciating the four Chinese tones, I administered a t-test to compare the gain scores of the recognition section between the two groups. Then, the same test was conducted on the gain scores of the enunciation section. These results indicated the pitch singing training method had a

greater effect on the accuracy of recognizing and enunciating the four Chinese tones than the traditional audio-lingual method. In addition, the Cohen results suggested the pitch singing training method had a large effect size on enunciating the four Chinese tones, while the effect size on recognizing the four Chinese tones was a little over medium.

Many language concepts are abstract. In order to construct and absorb abstract concepts and to build a complicated neuron network to comprehend the information, different strategies have been invented and applied in learning. In my study, the pitch singing training method had a greater effect on recognition and enunciation of Chinese tones than the traditional audio-lingual method. The findings indicate that using pitch singing training provides Chinese language learners with a sensory experience of the four tones and a foundation of better understanding the tone concept.

Moreover, using various strategies to build associations between learned knowledge and new knowledge can accelerate learning and enhance memory. Oxford (1990) generated a practical system of language learning, which involves six kinds of strategies: cognitive strategies, metacognitive strategies, memory strategies, affective strategies, social strategies, and compensation strategies. Vygotsky (1978) believed the important goals for educators is to teach and enable students to use the learning strategies independently and creatively and to help students combine the learning strategies so that they can reach and internalize higher psychological functions. Also, according to Vygotsky (1978), the process of internalization consists of a series of transformations: at the beginning, students learn through shared problem solving experiences with someone else, such as a parent, teacher, sibling or peer. Originally, the person interacting with the student undertakes the responsibility for providing guidance for problem solving, but

gradually this responsibility transfers to the student. Thus, it is important for Chinese language teachers to provide appropriate information for students to absorb, associate it with other knowledge, and understand it. The pitch singing training method gives Chinese language learners a more precise perception of Chine tones' contour and pitch registers.

Vygotsky (1978) also pointed out that humans use and invent strategies to help themselves remember and process information, such as drawing a map for direction, and using a birth date as a security number for a bank account so that it would not be forgotten. Human beings' memory is divided to short-term memory and long-term memory. For learning languages, memorizing words is a crucial aspect of the learning process (Schütze, 2017). Schütze (2017) stated the memory people have for a word lasts only for about half a second, which is stored in the short-term memory. To store and later recall a word from long-term memory, besides practicing and rehearsing the word often enough, one can add associations to the memory or elaborate on the memory (Schwartz, Tsang & Blair, 2016; Schütze, 2017). The pitch singing training can enrich the memory of the four Chinese tones and associate the intonations with singing high and low pitches.

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 vocal range in speaking and singing, but to accurately enunciate the four Chinese tones, one needs to use both low and high pitches. In addition, non-tonal language speakers do not have the awareness that the

Chinese tones are one key component in conveying meaning. Therefore, they tend to forget the tones when speaking Chinese and pronounce the Chinese words as toneless, which may result in inefficient communications. This non-complementary feedback could cause discouragement in learning, which would lead to motivation issues. Thus, in order to understand, memorize and use the four Chinese tones accurately, Chinese language learners need to first experience the intonations of the Chinese tones and build appropriate associations to them; then students can store the Chinese tones' intonations in memory for a longer period of time. A pitch singing training method provides the sensory experience for non-tonal language speakers to be familiar with the low and high abstract pitch intonation concept that is used in the four Chinese tones. Also, based on the four Chinese tones' intonation, singing pitch intervals exercises and extends the vocal range of Chinese language learners. Last, singing is an inevitable activity that humans do from birth and use in various occasions: festivals, memorial, worship, entertainment, etc. Therefore, singing groups of high and low pitches (intervals) can build appropriate associations for learning the four Chinese tones and to help non-tonal language learners to establish and store accurate input of the Chinese tones in their memory.

Summary, Recommendations and Conclusion

In this study, I compared two training methods' effects on recognizing and enunciating the four Chinese tones. American college students who did not take Chinese for more than one semester before this study or did not know Chinese at all were recruited for participating in this study. The participants were divided evenly and randomly into two groups. One group took the traditional audio-lingual training method, and the other group took the pitch singing training method. Both groups took one pretest

before the trainings and one posttest after the trainings. The differences between pre- and posttest scores were used to examine whether there were different effects between the two training methods and which training method had a greater effect on recognizing and enunciating the four Chinese tones. The score of singing-in-tune was used as a covariate. The results indicated the singing-in-tune skill did not influence the performance of recognizing and enunciating the four Chinese tones. Moreover, the Cohen results suggested the pitch singing training method had a large effect size on enunciating the four Chinese tones, while the effect size on recognizing the four Chinese tones was a little over medium. Lastly, the pitch singing training method had a greater effect than the audio-lingual training method on recognizing and enunciating the four Chinese tones. Thus, from the results of this study, 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.

Findings from this study support the notion that music training has positive effects on language learning (Bidelman, Hutka & Moreno, 2013). Results of my study are also congruent with the finding of Herrera, Lorenzo, Defior, Fernandez-Smith, and Costa-Giomi (2011) that musical training improves phonologic awareness, accelerates the speed of retrieving memory, and helps to store information into long-term memory. Thus, singing pitch intervals can help Chinese language learners to first experience the high and low pitch registers, to build connections with other existing knowledge, to accustom their vocal muscles to produce the dramatic intonations of the four Chinese tones, and to elaborate on the memory of the Chinese tones to store it for long-term retention.

In this study, I only focused on examining the effects of two training methods on individual Chinese character's pīn yīn pronunciation. It will be interesting to know if the pitch singing training method also has an effect on groups of characters' pronunciations, such as examining the enunciation of two and more characters together. For example, one can conduct a study on examining the accuracy of recognizing and enunciating the four Chinese tones in sentences, "wǒ jīn tiān qù nǐ jiā kàn diàn yǐng – Today I go to your home to watch movies." Also, it will 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 can examine the influence of music skills on recognizing and enunciating the Chinese tones. Furthermore, one can compare different effects between the pitch singing training method and other methods, such as computer-based training program, or learning the tones with hand movements. Finally, one can apply this study to a bigger 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 using this training method is organized as follows:

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".

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. Then the teacher can 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.

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”.

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

I used these steps in the pitch singing training group, and the results suggested that the pitch singing group performed better than the audio-lingual method group on recognizing and enunciating the four Chinese tones after the trainings.

From this study, one can conclude that the pitch singing training is an effective method that can build associations between the pitch intonation and the 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. Chinese language teachers do not need to be professional musicians to offer this pitch singing training method. The pitch singing exercise is similar to any vocal warm-up exercise that is used in singing classes in schools. If Chinese language teachers do not know how to play an instrument, they can pre-record these pitches and singing materials by using online digital software or applications. Then, they can play these recordings in the class and ask students to sing along. 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). https://www.actfl.org/publications/guidelines-and-manuals/actfl-proficiency-guidelines-2012

American Council on the Teaching of Foreign Languages. (2019). https://www.actfl.org/membership/special-interest-groups-sigs/immersion

Bai, J., Lien, L., & Spring, M.K. (2016). Mapping Chinese language learning outcomes in grades K-12. In S.C. Wang & J.K. Peyton (Eds.), *CELIN (Chinese Early Language and Immersion Network) Briefs Series*. New York, NY: Asia Society.

Best, C. T. & McRoberts, G. W. (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, G., 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. doi:10.1371/journal.pone.0060676

Brown, H. D. (2014). *Principles of language learning and teaching* (6th ed.). White Plains, NY: Pearson Longman.

Bluhme, H., & Burr, R. (1971). An audio-visual display of pitch for teaching Chinese tones. *Studies in Linguistics*, *22*, 51–57.

Chao, Y. R. (1948). *Mandarin primer: An intensive course in spoken Chinese*. Cambridge, MA: Harvard University Press.

Chao, Y. R. (1968). *A grammar of spoken Chinese*. Berkeley, CA: University of California Press.

Chen, Q. (1997). Toward a sequential approach for tonal error analysis. *Journal of the Chinese Language Teachers Association, 32*(1), 21–39. Retrieved from https://scholarsarchive.byu.edu/dlls/vol19/iss1/8/

Cheng, C. C. (1973). *A synchronic phonology of mandarin Chinese*. The Hague, Netherlands: Mouton.

Chomsky, N. (1995). *The minimalist program*. Cambridge, MA: MIT Press.

Chomsky, N. (2013). Lecture I: What is language? *Journal of Philosophy, 110*, 645–662. Retrieved from http://www.jstor.org/stable/43820808

Chuang, C. K., Hiki, S., Sone, T., & Nimura, T. (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. https://asa.scitation.org/doi/pdf/10.1121/1.1981919?class=pdf

Cooper, A., & Wang, Y. (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

Costa-Giomi, E., & Ilari, B. (2014). Infants' preferential attention to sung and spoken stimuli. *Journal of Research in Music Education, 62*, 188–194. doi:10.1177/0022429414530564

Custodero, L. A. (2006). Singing practices in ten families with young children. *Journal of Research in Music Education*, *54*, 37–56. doi:10.1177/002242940605400104

Delogu, F., Lampis, G., & Belardinelli, M. O. (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

Flege, J. E. (1995). Second language speech learning: theory, findings, and problems. In W. Strange (Eds.), *Speech perception and linguistic experience* (233–273). Baltimore, MD: York Press.

Foreign Service Institute's School of Language Studies (2018). FSI's Experience with Language Learning. Retrieved from https://www.state.gov/m/fsi/sls/c78549.htm

Freeman, D. E., & Freeman, Y. S. (2001). *Between worlds: Access to second language acquisition* (2nd ed.). Portsmouth, NH: Heinemann.

Gandour, J. (1984). Tone dissimilarity judgements by Chinese listeners. *Journal of Chinese Linguistics, 12*, 235–261. Retrieved from http://www.jstor.org.unco.idm.oclc.org/stable/23767002

Gandour, J., Wong, D., Hsieh, L., Weinzapfel, B., Van Lancker, D. & Hutchins, G. D. (2000). A PET cross-linguistic study of tone perception. *Journal of Cognitive Neuroscience, 12*, 207–222. doi:10.1162/089892900561841

Gass, S., & Selinker, L. (1994). *Second language acquisition*. Hillsdale, NJ: Lawrence Erlbaum.

Geringer, J. M., Nelson, J. K., & Kostka, M. C. (1980). Differential assessment of child and adult singing ranges. *Contributions to Music Education, 8*(8), 39–46. Retrieved from http://www.jstor.org/stable/24127432

Ghanbari, F., & Hashemian, M. (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.

Gordon, E. (1980). *Learning sequences in music: Skill, content, and patterns*. Chicago, IL: GIA Publications.

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

Haynes, B., & Cooke, P. (2017). "*Pitch*." Grove Music Online. Oxford Music Online. Oxford University Press. Retrieved from http://www.oxfordmusiconline.com/subscriber/article/grove/music/40883.

Herrera, L., Lorenzo, O., Defior, S., Fernandez-Smith, G., & Costa-Giomi, E. (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

Ho, C. S.-H., & Bryant, P. (1997). Development of phonological awareness of Chinese children in Hong Kong. *Journal of Psycholinguistic Research*, *26*(1), 109–126. Retrieved from https://link.springer.com/content/pdf/10.1023/A:1025016322316.pdf

Hsieh, L., Gandour, J., Wong, D., & Hutchins, G. D. (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, J. (2000). *Students' major difficulties in learning Mandarin Chinese as an additional language and their coping strategies.* Retrieved from ERIC database. (ED440537)

Ilari, B., Moura, A., Bourscheidt, L.A. (2011). Between interactions and commodities: Musical parenting of infants and toddlers in Brazil. *Music Education Research*, *13*, 51–67. doi:10.1080/14613808.2011.553277

Johnson, K. E. (1995). *Understanding communication in second language classrooms*. Cambridge, United Kingdom: Cambridge University Press.

Kiriloff, C. (1969). On the auditory discrimination of tones in Mandarin. *Phonetica*, *20*, 63–67.

Klein, D., Zatorre, R. J., Milner, B., & Zhao, V. (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

Krashen, S. (1982). *Principles and practice in second language acquisition*. Oxford, United Kingdom: Pergamon Press.

Lee, C.-Y., & Hung, T.-H. (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

Lightbown, P., & Spada, N. (1993). *How languages are learned*. Oxford, United Kingdom: Oxford University Press.

Lin, M. C. (1965). The pitch indicator and the pitch characteristics of tones in Standard Chinese. *Acta Acoustica (China)*, *2*, 8–15.

Long, M. H. (1985). A role for instruction in second language acquisition: Task-based language teaching. In K.Hyltenstam & M. Pienemann (Eds.), *Modelling and assessing second language acquisition* (pp. 77–99). San Diego, CA: College-Hill Press.

Luo, W. (2018). *Chinese alphabetization reform: Intellectuals and their public discourse, 1949–1958* (Order No. 13418856). Retrieved from ProQuest Dissertations & Theses Global. (2177028153)

Magne, C., Schön, D., & Besson, M. (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, 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*, 2701–2715. doi:10.1162/jocn.2010.21585

Milovanov, R., Huotilainen, M., Välimäki, V., Esquef, P. A., & Tervaniemi, M. (2008). Musical aptitude and second language pronunciation skills in school-aged children: neural and behavioral evidence. *Brain Research*, *1194*, 81–89. doi:10.1016/j.brainres.2007.11.042

Moore, R. 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. Retrieved from http://www.jstor.org/stable/40318417

Mora, C. F. (2000). Foreign language acquisition and melody singing. *ELT Journal*, *54*, 146–152. doi:10.1093/elt/54.2.146

Moreno, S., Marques, C., Santos, A., Santos, M., Castro, S. L., & Besson, M. (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, S. J., Demorest, S. M., Aylward, E. H., Cramer, S. C., & Maravilla, K. R. (2003). FMRI investigation of cross-cultural music comprehension. *Neuroimage*, *20*, 378–384. doi:10.1016/S1053-8119(03)00300-8

Nan, Y., & Friederici, A. D. (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

Oxford, R.L. (1990). *Language learning strategies: What every teacher should know*. Boston, MA: Heinle.

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

Piaget, J. (1973). *Main trends in psychology*. New York, NY: Harper & Row.

Rahbar, S., & Khodabakhsh, S. (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. Retrieved from http://www.journals.aiac.org.au/index.php/IJALEL/article/view/955

Repp, B., & Lin, H. (1990). Integration of segmental and tonal information in speech-perception: A cross-linguistic study. *Journal of Phonetics, 18*, 481–495. Retrieved from https://psycnet.apa.org/record/1991-29383-001

Russell, J. A. (2018). *Statistics in music education research: A reference for researchers, teachers, and students*. New York, NY: Oxford University Press.

Saville-Troike, M. (2012). *Introducing second language acquisition* (2nd ed.). Cambridge, United Kingdom: Cambridge University Press.

Schön, D., Boyer, M., Moreno, S., Besson, M., Peretz, I., & Kolinsky, R. (2008). Songs as an aid for language acquisition. *Cognition, 106*, 975–983. doi:10.1016/j.cognition.2007.03.005

Schön, D., Magne, C., & Besson M. (2004). The music of speech: Music training facilitates pitch processing in both music and language. *Psychophysiology, 41*, 341–349. doi:10.1111/1469-8986.00172.x

Schütze, U. (2017). *Language learning and the brain: Lexical processing in second language acquisition*. Cambridge, United Kingdom: Cambridge University Press.

Schwartz, D. L., Tsang, J. M., & Blair, K. P. (2016). *The ABCs of how we learn: 26 scientifically proven approaches, how they work, and when to use them* (1st ed.). New York, NY: W.W. Norton & Company, Inc.

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

Shehadeh, I. I., & Farrah, M. A. (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. Retrieved from 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, X. S. (1989). Toward a register approach in teaching Mandarin tones. *Journal of the Chinese Language Teachers Association, 24*(3), 27–47.

Shen, C. (2009). Using English songs: An enjoyable and effective approach to ELT. *English Language Teaching*, *2*, 88–94. Retrieved from ERIC database. (EJ1082242)

Shi, H. (2018). *A method of teaching English speaking learners to produce Mandarin-Chinese tones* (Order No. 10792055). Retrieved from ProQuest Dissertations & Theses Global. (2042860149)

Silber, K. (1973). *Pestalozzi: The man and his work*. (3rd ed.). New York: Schocken Books.

Skinner, B. F. (1957). *Verbal behavior*. New York, NY: Appleton-Century-Crofts.

Stewart, L., von Kregstein, K., Warren, J. D., & Griffiths, T. D. (2006). Music and the brain: Disorders of musical listening. *Brain*, *129*, 2533–2553. doi:10.1093/brain/asl171

Strong, W. J., & Plitnik, G. R. (1992). Music speech audio. Provo, UT: Soundprint.

Swain, M. (1985). Communicative competence: Some roles of comprehensible input and comprehensible output in its development. In S. Gass & C. Madden (Eds.), *Input in second language acquisition* (pp. 235–253). Rowley, MA: Newbury House.

Tafuri, J. (2016). *Infant musicality: New research for educators and parents* (G. Welch, Ed., & E. Hawkins, Trans.). New York, NY: Routledge.

Trehub, S. E. (2001). Musical predispositions in infancy. *Annals of the New York Academy of Sciences*, *930*, 1–16. doi:10.1111/j.1749-6632.2001.tb05721.x

Tsai, R. (2011). Teaching and learning the tones of Mandarin Chinese. *Scottish Languages Review, 24*. 43–50. Retrieved from https://www.scilt.org.uk/portals/24/library/slr/issues/24/24_5_tsai.pdf

Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological processes*. Cambridge, MA: Harvard University press.

Vygotsky, L. S. (1986). *Thought and language*. Cambridge, MA: MIT Press.

Wang, H. H., & Honig, A. S. (2010). What difficulties do children experience while learning to read and write Chinese? In Chen, J., Wang, C., & Cai, J., (Eds.), *Teaching and learning Chinese: Issues and perspectives* (pp. 139–154). Charlotte, NC: Information Age Publishing.

Wang, Y., Sereno, J. A., Jongman, A. & Hirsch, J. (2003). FMRI evidence for cortical modification during learning of Mandarin lexical tone. *Journal of Cognitive Neuroscience, 13*, 994–1005. doi:10.1162/089892903770007407

White, C. M. (1981). Tonal pronunciation errors and interference from English intonation. *Journal of the Chinese Language Teachers Association, 16*(2), 27–56.

White, L. (1987). Against comprehensible input: The input hypothesis and the development of second language competence. *Applied Linguistics, 8*, 95–110. doi:10.1093/applin/8.2.95

Wong, P. C., & Perrachione, T. K. (2007). Learning pitch patterns in lexical identification by native English-speaking adults. *Applied Psycholinguistics, 28*, 565–585. doi:10.1017/S0142716407070312

Wong, P. C., Skoe, E., Russo, N. M., Dees, T., & Kraus, N. (2007). Musical experience shapes human brainstem encoding of linguistic pitch patterns. *Nature Neuroscience*, *10*, 420–422. doi:10.1038/nn1872

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

Xing, J. Z. (2006). *Teaching and Learning Chinese as a Foreign Language: A Pedagogical Grammar*. Hong Kong, China: Hong Kong University Press.

Yang, J. (2015). *Teacher and pupil beliefs about beginning to learn Chinese language in English secondary schools* (Doctoral dissertation, University of Warwick). Retrieved from http://webcat.warwick.ac.uk/record=b2863329~S1

Zatorre, R. J., & Gandour, J. T. (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, R. J., Meyer, E., Gjedde, A., & Evans, A. C. (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

Zhang, F. Z. (2006). *The teaching of Mandarin prosody: A somatically-enhanced approach for second language learners* (Doctoral thesis, University of Canberra, Australia). Retrieved from https://www.researchgate.net/profile/Felicia_Zhang2/publication/262561996_TH

E_TEACHING_OF_MANDARIN_PROSODY_A_SOMATICALLY-ENHANCED_APPROACH_FOR_SECOND_LANGUAGE_LEARNERS/links/02e7e53802c7e64fd8000000.pdf

Zhao, J. M. (1987). 基础汉语语境教学的若干问题。[Some problems in the teaching of basic Chinese.]. *Proceedings of the First International Conference on Teaching Chinese*, Beijing, China.

APPENDIX A

INSTITUTIONAL REVIEW BOARD APPROVAL

UNIVERSITY OF
NORTHERN COLORADO

Institutional Review Board

DATE:	April 2, 2018
TO:	Danqing Zhou, doctoral program
FROM:	University of Northern Colorado (UNCO) IRB
PROJECT TITLE:	[1210468-1] The effects of pitch singing training on recognizing and pronouncing Chinese tones
SUBMISSION TYPE:	New Project
ACTION:	APPROVAL/VERIFICATION OF EXEMPT STATUS
DECISION DATE:	March 30, 2018
EXPIRATION DATE:	March 30, 2022

Thank you for your submission of New Project materials for this project. The University of Northern Colorado (UNCO) IRB approves this project and verifies its status as EXEMPT according to federal IRB regulations.

Danquing -

Thank you for your patience with the UNC IRB process. Your materials and protcols are clear, thorough and verified/approved exempt. You may begin participant recruitment and data collection.

Best wishes with this research.

Sincerely,

Dr. Megan Stellino, UNC IRB Co-Chair

We will retain a copy of this correspondence within our records for a duration of 4 years.

If you have any questions, please contact Sherry May at 970-351-1910 or Sherry.May@unco.edu. Please include your project title and reference number in all correspondence with this committee.

This letter has been electronically signed in accordance with all applicable regulations, and a copy is retained within University of Northern Colorado (UNCO) IRB's records.

APPENDIX B

PRETEST AND POSTTEST MATERIALS

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 C

PITCH TRAINING MODEL VS. AUDIO-LINGUAL MODEL

Each training session is divided into three parts learning: vocabulary learning, reading a dialogue and sentence structure or cultural facts learning.

Pitch Singing Training Procedure of Each Session:

1. The instructor uses pitches to pronounce each character from one session (twice for each character):

   the first tone using the note G, the second tone using notes E to G,

   the third tone using notes C to B (lower) to D, and the fourth tone using 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 each character as the instructor pronounces it (twice for each character).

2. The instructor pronounces each character in normal speaking mode, without the pitch tones (twice for each character).

   The students repeat each character as 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. Then students are paired in groups and read the dialogues. The instructor takes the notes of mispronounced characters and corrects the students at the end of each group's performance.
6. Before finishing one session, the instructor explains some grammar points and cultural facts that are related in the dialogue.

Audio-Lingual Model (Non-Music Training) of Each Session

(The differences between two models are from the first to the third step. The rest is the same for both models.)

Procedure:

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. The instructor reads the dialogue sentence by sentence with the normal speaking mode. The students repeat what the instructor says.
5. Then students are paired in groups and read the dialogues. The instructor takes the notes of mispronounced characters and corrects the students at the end of each group's performance.
6. Before finishing one session, the instructor explains some grammar points and cultural facts that are related in the dialogue.

APPENDIX D

MATERIALS OF EIGHT TRAINING SESSIONS

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

ü ü ü ü

Consonant

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

Session 2 — Lesson One: Greetings

- Vocabulary

wǒ 我 (I), 2. nǐ 你 (you), 3. hǎo 好 (good-adj.), 4. jiào 叫 (to be called-verb), 5. shénme 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

1 3 6 5- 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

Ex: 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.

Ex: 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

Ex: a. xī cān hái shì zhōng cān-western food or Chinese food

b. xué shēng hái shì lǎo shī- student or teacher