Using Online Applications to Improve Tone Perception among L2 Learners of Chinese
(网络应用对中文二语学习者声调辨识的有效性研究)

Xu, Hongying
(徐红英)
University of Wisconsin-La Crosse
(威斯康星大学拉克罗校区)
hxu@uwlax.edu

Li, Yan
(李艳)
University of Kansas
(堪萨斯大学)
liyan@ku.edu

Li, Yingjie
(李颖颉)
University of Colorado-Boulder
(科罗拉多大学博尔得分校)
Yingjie.Li@colorado.edu

Abstract: This study investigated the effectiveness of an online application in helping beginning-level Chinese learners improve their perception of the tones in Mandarin Chinese. Two groups—one experimental and one traditional—of beginning Chinese learners from two universities in the Midwest participated in this study. The experimental group (n=20) used the online application to practice tones for four 15-minute sessions in class. The traditional group (n=11) participated in traditional instructor-led practice in class in lieu of the online practice. Both groups completed a pre-test, an immediately administered post-test, and a delayed post-test designed to assess their perception of the tones of monosyllabic and disyllabic words. No statistically significant difference has been found between the two groups in their tone perception performance in the post-test and in the delayed post-test. However, the experimental group showed a positive trend in improving their perception on those tones which posed more difficulty than others. Their experience with this online application and the pronunciation learning strategies of participants in the experimental group were also examined through a survey. Based on the findings, it is proposed that the use of online tone practice is worthwhile in a Chinese language class, but might fit better into the curriculum as external assignments.

摘要：本研究调查了网络应用在帮助初级中文二语学习者提高声调辨识的准确度上的有效性。实验组和对照组的参与者均来自于美国中西部两所四年制大学的初级中文班。实验组的学生（n=20）参与了四次每次十五分钟的基于网络应用的课堂练习，对照组的学生（n=11）则在实验组使用网络应用进行声调练习的时间内进行传统的教师带领的声调练习。两组学生都参加了测试单音节词、双音节词中的声调辨识准确度的前测、后测和延时后测。实验组与对照组在前测、后测及延后测试中声调辨识的准确度上没有显著差异。然而，实验组在某些比较难辨识的声调上显示出较好的改进倾向。实验组的学生还完成了一份检测学习者使用网络应用进行声调练习的经验和学习发音

的策略的问卷。根据调查结果，本文建议在正式的教学中使用网络应用进行声调练习，但这种练习可能更适合作为作业让学生在课外完成。

Keywords: Online application, Chinese tone perception, learning experience

关键词：网络应用、中文声调辨识、学习体验

1. Introduction

As a tonal language, Mandarin Chinese employs different pitch heights and pitch contours to distinguish meaning. Research has shown that, regardless of their first language (L1), second language (L2) learners of Chinese have difficulty mastering Mandarin tones (Miracle, 1989; Shen, 1989; Shen & Lin, 1991; Sun, 1998; Jongman, Wang, Moore, & Sereno, 2006; Lee, Tao, & Bond, 2010; Chun, Jiang, Meyr, & Yang, 2015; Hao, 2012).

Despite the fact that Chinese language instructors and Chinese linguists recognize the critical role that the mastery of tones plays in developing communicative competence, the teaching of tones in the classroom lags far behind the teaching of other linguistic aspects and skills (Xing, 2006; Orton, 2013).

A key to channeling the genuine need and practice is finding an efficient way to improve learners' perception of tones in a short period of time. Incorporating online tone training practice into Chinese language class might solve the dilemma. Some studies carried out in lab settings show that intensive tone training can improve Chinese learners' perception accuracy on tones (Wang, Spence, Jongman, & Sereno, 1999) which can, in turn, be transferred to tone production (Rochet 1995; Wang, Jongman, & Sereno, 2003). Would similar effects be observed among Chinese learners who receive online tone training in a classroom setting? How would students of Chinese feel about online tone training? These questions remain unanswered because hardly any empirical studies that test the effectiveness of online tone training sessions on learners' perception of tones have been carried out in authentic classroom situations. In addition, very few studies have investigated the training effects of tone perception using true beginners during the time of their initial exposure to tones. Studies of the training effect in regular classrooms with true L2 Chinese beginners are thus needed.

The present study addresses this gap by exploring the effectiveness of computer-assisted online tone practice incorporated into classroom instruction in improving true beginning L2 Chinese learners' perception of tones. In addition, this study also investigates L2 Chinese learners' experience in their use and perception of online tone practice.

This article is organized as follows: Section 2 discusses research on L2 Chinese learners' tonal perception and the literature on computer-assisted tone training; Section 3 reports on the current empirical study, including research questions, experimental design, results and discussion; Section 4 comprises the conclusion drawn from the experiment and

the pedagogical implications, and; Section 5 discusses the limitations of the current study and suggests directions for future research.

2. Overview of L2 Chinese Learners' Tonal Perception and Current Literature on Computer-Assisted Tone Training

2.1 L2 Chinese Learners' Tonal Perception

Tones in Chinese are lexical in that the application of different tones to the same syllable results in different meanings. Linguists generally agree that there are four tones plus a neutral tone in Mandarin Chinese. These tones are different in pitch height and contour as can be seen in Fig. 1 below, in which a scale of 1 to 5 is used to describe pitch height and the critical points in the change of pitch height, i.e. the beginning point and endpoint, are depicted to represent different tones.
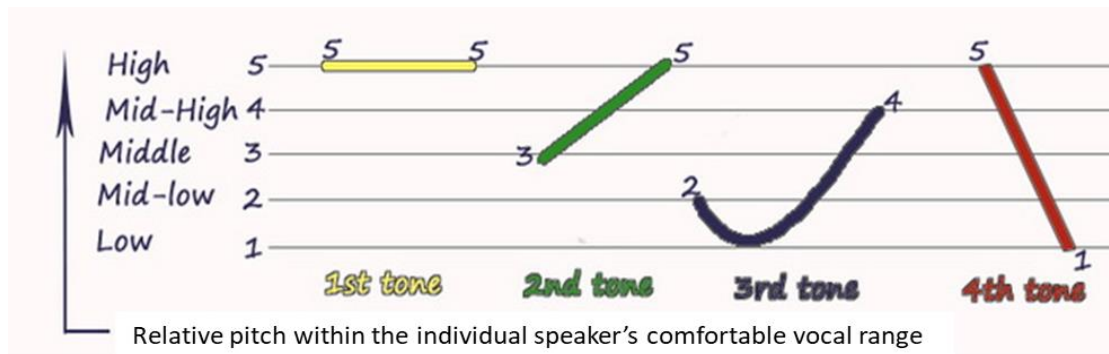

Figure 1 Tone Pitch and Contour Chart

The same Chinese syllable can express several unrelated meanings when pronounced using different tones. For example, if the syllable *ma* is pronounced using the first tone, a high level tone represented by 55, it means "mom"; when it is pronounced using the second tone, 35, it means "hemp"; when it is pronounced using the third tone, 214, it means "horse"; and it means "to scold" when pronounced using the fourth tone, 51. The neutral tone differs from the four tones in that its pitch height is decided by the tone of the preceding syllable and its pronunciation is light and short. The syllable *ma*, when pronounced using the neutral tone, functions as a question particle.

A failure to perceive Mandarin tones and establish the tonal categories can "have direct and drastic consequences" (Rohr, 2014) on both listening comprehension and communicative competence. It has been widely reported that L2 Chinese learners whose L1 has no tonal element have difficulty perceiving tones in Mandarin (Miracle, 1989; Shen, 1989; Shen & Lin, 1991; Sun, 1998; Jongman et al., 2006; Lee, et al., 2010). On monosyllabic words, L2 Chinese learners seem to acquire the rising tone (Tone 2) or dipping tone (Tone 3) later than the level tone (Tone 1) or the falling tone (Tone 4). A few studies that examined learners' tonal perception on disyllabic words found that learners' tonal performance on disyllabic words is generally worse than on monosyllabic words due to the complex relationship between adjacent tones and the great tonal variability in disyllabic words.

2.2 Research on Computer-assisted Tone Training

In today's Chinese language classroom in the United States, the teaching of pronunciation, including tones, lags far behind the teaching of other linguistic aspects and skills (Levis, 2007). Tones are introduced to learners mainly in isolation within a short period at the very beginning of the target language instruction: no priority is assigned to instruction of tones and less time is allotted to practice (Xing, 2006; Orton, 2013).

Various factors contribute to this issue, some of which are quite justifiable. Students exhibit wide-ranging differences in pronunciation needs and challenges: some need help with retroflex fricatives; some need to work on the palatals; some want more practice with third tones, etc. It is almost impossible to cater to all individual needs in class. Some language instructors, without foundational training in phonetics and phonology, have a hard time effectively explaining the articulation position and manner of each sound or each tone. Thus, when their students make mistakes, it is hard for them to give immediate and effective corrective feedback (Levis, 2007). Without targeted corrective feedback, simple mimicking and repetition do not work that well in class (Xu, 2017). However, the first few weeks of exposure to the Chinese language is critical in L2 Chinese learners' ability to produce correct tones partly because adult production and perceptual systems demonstrate only a certain degree of plasticity (Wang et al. 1999; Wang et al., 2003). L2 Chinese learners need to repeatedly practice listening to and producing tones since "good pronunciation of a language is a matter of motor skills, coupled with ear training" (Hockett, 1951). Therefore, it would be beneficial to find a tool that provides many opportunities for practice, that caters to each learner's need, and that is attractive enough that learners will actually use it.

Computer-assisted language learning (CALL) might meet these requirements. It has been an effective pedagogical approach since it was first integrated into American foreign language pronunciation teaching in the 1980s. CALL offers many advantages to the modern foreign language classes, including increasing learning efficiency and effectiveness, and providing easy access, great convenience, strong motivation, and institutional efficiency (Hubbard, 2009).

Short-term training using CALL, or CAPT (computer-assisted pronunciation teaching), has proven to be effective in assisting learners in the acquisition of new phonetic contrasts that do not exist in the phonological systems of their native language, be it English, Chinese, German, Cantonese, Spanish, or any other languages (Logan, Lively, & Pisoni, 1991; Wang et al., 1999; Wang et al., 2003; Kingston, 2003; Francis, Ciocca, Ma, & Fenn, 2008; Herd, Jongman, & Sereno, 2013; Li, 2016; Godfroid, Lin, & Ryu, 2017; Xu, 2017). In their studies, Wang et al. (1999) and Wang et al. (2003) investigated the effect of short-term auditory training on tonal performance in monosyllabic Chinese words while Li (2016) examined improvement in L2 learners' tonal perception in both mono- and disyllabic words after two weeks of perceptual training. After only eight sessions of auditory training, Wang et al. (1999; 2003) found that English-speaking learners of Chinese had improved significantly both in their tonal perception (by 21%) and tonal production (by 18%) on monosyllabic words. Li (2016), using the same test stimuli as in Wang et al. (1999, 2003) conducted a two-week auditory training on L2 learners' tonal perception of both

monosyllabic and disyllabic words. She found an overall 12% significant improvement after the auditory training. Moreover, learners showed greater improvement on disyllabic stimuli (12%) over monosyllabic stimuli (8%). All these studies indicated a significant gain in Chinese language learners' perception of both monosyllabic and disyllabic tones. Godfroid, Lin, and Ryu (2017) compared the effectiveness of using three single-cue methods and two dual-cue methods in helping L1 English L2 Chinese learners improve their perception of Chinese tones. Participants improved an average of 16% and 17% on trained syllables and 12% and 10% on untrained syllables. Only monosyllables (CV or CVC-structured) were used in their training and on tests.

However, those training sessions were all conducted in a clinical-lab environment that required an extra time commitment from L2 learners, facility support from the institute, and more management from the language instructor. Many scholars have recommended more research in natural classroom settings (Chun, 2015; Chun, 2017; Derwing & Munro, 2015). Chun (2017) pointed out that, although experimental studies are still needed to examine the effectiveness of application, it is as important to conduct studies "in authentic learning situations, for example, as part of regular scheduled language courses." But such studies are scarce.

Previous studies that investigated the acquisition order of the four tones by L2 Chinese learners have reached, more or less, a consensus that L2 learners acquire the level tones (e.g. T1) earlier than the contour tones (T2, T3, and T4). As for the contour tones, learners acquired the falling contour tone (T4) earlier than the rising contour tone (T2), which is, in turn, acquired earlier than the dipping tone (T3) (Zhang, 2018). Do training effects on tone perception vary among individual tones? Previous studies have shown mixed results. Wang et al. (1999) found no significant difference among the four tones on either the pre-test or the post-test, although, numerically, trainees' tones improved in the following order T4 > T2 > T3 > T1. So (2006) found that trainees needed more training sessions to distinguish between T1/T4 and T2/T4 than they did on T1/T2 or T1/T3. Would similar patterns of training effects be observed if the training is self-paced in a natural classroom setting? Would the online application facilitate L2 Chinese learners' acquisition of the harder to acquired tones, i.e. the contour tones?

Xu (2017) incorporated an online tone training application available at http://www.pinyinpractice.com into her beginning Chinese class. After two 20-minute sessions of in-class practice, her L2 Chinese learners' tonal perception improved significantly, as shown on the immediately administered post-test. Xu also took a survey of the learners' user experience of this online application and found that it was well received by a majority of the students. However, because Xu did not include a control group for comparison, it remains unclear whether the significant improvement was due to natural learning or was the benefit of the online training. Moreover, Xu only tested the perception of tones in monosyllabic words, which leaves the question of whether such training might be as helpful to students in perceiving tones in disyllabic words unanswered. For these reasons, it was necessary to re-investigate the effectiveness of online tone training practice through a new study that included a comparison group and tested the training's effectiveness on both monosyllabic words and disyllabic words. The acquisition order of

tones, training effects on different tones and learners' attitude towards such online tone training session was investigated as well.

3. The Current Empirical Study

3.1 Purpose of This Study

The present study explored the effectiveness of computer-assisted online tone practice in improving beginning L2 Chinese learners' perception of tones on monosyllabic and disyllabic words. It also explored whether the effect varied among different tones. The goal of this study was to identify an effective alternative, one that Chinese language instructors could easily incorporate into the curriculum. A group of students who received the traditional instructor-led instruction was used as a baseline for comparison. The following research questions were addressed:

1. Do L2 Chinese learners who receive online tone training in class show equal acuity in tone perception compared to students who receive traditional instructor-led tone training in class?
2. Do L2 Chinese learners who receive online tone training in class improve differently in tone perception on monosyllabic words or on disyllabic words than those who receive traditional instructor-led tone training in class?
3. Do L2 Chinese learners who receive online training in class improve differently on their perception of different tones on monosyllabic words or on disyllabic words than those who receive traditional instructor-led tone training in class?
4. How do students feel about online tone training?

Questions 1 through 3 were investigated by comparing the performance of two groups of students, one group receiving traditional instruction and the other using online tone-training practice, on the three tone perception tests (i.e. pre-test, post-test, and delayed post-test). A delayed post-test was used in order to measure the long-term effects that online tone training practice may have on Chinese learners' tonal perception. Neither the experimental group nor the control group was aware that they would be given a delayed post-test after they completed the post-test. Question number 4 was answered by analyzing answers to the questionnaire given to the students in the experimental group after they completed the delayed post-test.

The hypotheses regarding each research question that require statistical tests are listed as follows:

1. Research Question 1. Null hypothesis: there is no difference between the two groups in their total acuity of tone perception across tests. Alternative hypothesis: there is a difference between the two groups in their total acuity of tone perception across tests.
2. Research Question 2. Null hypothesis: there is no difference between the two groups in the acuity of their tone perception in the monosyllabic or the disyllabic settings across tests. Alternative hypothesis: there is a difference between the two

groups in their acuity of their tone perception in the monosyllabic or the disyllabic settings across tests.

3. Research Question 3. Null hypothesis: there is no difference between the two groups in the acuity of their perception of each tone in the monosyllabic or the disyllabic settings across tests. Alternative hypothesis: there is a difference between the two groups in the acuity of their perception of each tone in the monosyllabic or the disyllabic settings across tests.

3.2 Method and Design

3.2.1 Participants

Participants in this study were students enrolled in elementary Chinese classes at two universities in the American Midwest. The textbooks used by the two programs were *Integrated Chinese* and *Chinese Link* respectively. Despite the different textbooks used by the two programs, the grammar and vocabulary for first-year Chinese language learning were not much different and, in fact, are not much different across many textbooks (Cai & Sciban, 2008). We used natural class sessions and randomly decided on the division of the participants into the experimental group and the traditional group. Initially, 49 students participated in the study, but 14 of those students were ultimately excluded from the analysis due to previous Chinese language learning experience or exposure to other tonal languages. In addition, 4 more students were excluded from the analysis because they missed either the pre-test or the post-test. As a result, the data from 31 native English-speaking learners of Chinese were analyzed in this study, with 20 in the experimental group and 11 in the traditional group. These participants represented a variety of majors including business, finance, and political science. None had any history of hearing, speech, or language difficulties.

3.2.2 Training Materials and Test Instruments

All participants received a language learning background information sheet (see Appendix A) on the first day of class. The purpose of this information sheet was to ascertain whether participants had studied Chinese before or had had previous exposure to Chinese or any other tonal language, which might influence their performance on the tone perception task.

We chose http://www.pinyinpractice.com/ as the online tone training tool for the experimental group based on the selection criteria for CAPT tools proposed in Levis (2007). According to Levis (2007), a CAPT tool should provide a substantial amount of input, immediate and useful feedback, and motivate the learners to practice. The website www.pinyinpractice.com provides ample randomized tone exercises accompanied by audio files. Users can either choose the tones from the screen or enter the tones using numbers. It also provides instant feedback by indicating "correct" or "incorrect" in Chinese characters when learners pick or supply a tone for the syllable they hear in the audio. The interface is so well designed that learners are motivated to use it on their phone at any time that is convenient. Students in Xu's (2017) study expressed favorable feelings about the website. Moreover, the website is free and thus adds no financial burden to the students or

the instructors. This makes it more likely to be adopted in Chinese language courses if proven to be effective.

To measure the effectiveness of this online application, a perception task with 64 monosyllabic stimuli and 32 disyllabic stimuli adapted from Li (2016) was designed. The monosyllabic stimuli included all possible combinations of various initial consonants and final vowels, and different syllabic structures in Mandarin Chinese (i.e. V, CV, CVNasal, VN, CGlideV, and CGVN). In contrast, each disyllabic stimulus was composed of a combination of two random individual syllables from the monosyllabic stimuli. In such a design, the identical tones and vowel consonant combinations in both monosyllabic and disyllabic words ensured that students' perception of tones in both settings would be comparable. All stimuli were produced and recorded in a language lab by a well-trained female Chinese language instructor who had not been apprised of the purpose of the study and who had no contact with any of the students in either of the two Chinese programs. Pre-test, immediate post-test and delayed post-test were identical across both programs. The same sets of audio files were randomized on the pre-test, immediately administered post-test, and delayed post-test. For each version of the test, there was a corresponding test sheet composed of the syllables without tones (See Appendix B-1 for the test stimuli). For each test, students received a test sheet (Appendix B-2[1]) and were asked to circle the correct tone(s) for each stimulus according to what they heard on the audio recording. In each setting, monosyllabic or disyllabic, there are four tone marks provided after each syllable. Therefore, participants had to focus on the tones in order to choose the correct tone mark.

A survey was designed and conducted along with the delayed post-test with the students in the experimental group to investigate Chinese learners' attitudes toward tone learning and the use of online applications in learning tones. The survey consisted of self-assessment questions targeting the difficulties learners have in perceiving the tones, strategies that students believe can help improve their tone perception, and their experience with the use of the online application. The self-assessment questions were included because a good self-awareness of learning difficulties may help learners do more targeted practice if they are to use the online applications on their own in the future.

3.2.3 Procedure

All participants in both the experimental group and the traditional group completed a language learning background information sheet on the first day of Chinese class. A tone perception pre-test was given to the participants on the second day of class immediately following a short introduction to the four tones in Mandarin Chinese. The purpose of this test was to provide a baseline against which to compare participants' improvement. One group of participants received traditional tone practice while the other did online tone training practice in class which consisted of four 15-minute sessions spread over a two-week period. The online tone practice included both monosyllabic and disyllabic words

---

[1] Appendix B-2 shows the test sheet that is used for the pre-test. The test sheets used for the immediate post-test and the delayed post-test were different only in the order of test items, which are not included in the appendix to save space.

found on the website www.pinyinpractice.com, with the first two sessions focusing on monosyllable tone practice and the last two sessions on disyllable tone practice.

A tone perception post-test was administered one day after the last training session. Four weeks later, a delayed tone perception post-test was administered to both groups. Learners in the experimental group were also asked to complete a questionnaire that addressed their experience with and their attitudes towards the online application.

During the time that the experimental group practiced tones in class using www.pinyinpractice.com, the traditional group practiced tones in class with the instructors in the traditional way, namely, through repetition after the instructor and reading aloud, both followed by immediate corrective-feedback from the instructor. For instance, the instructor usually pronounced a syllable with the correct tones after which the students repeated the syllable in groups or individually. If a student made an error, whether it was on the initial, final or the tone, the instructor would correct the student by pronouncing the syllable correctly and asking him/her to repeat.

3.2.4 Data Coding

All data from the language learning background information sheets, the three tone perception tests, and the tone learning experience questionnaire were input into Excel files. Data were cross-checked by two of the authors to ensure accuracy.

The points for the 64 monosyllabic stimuli and the 32 disyllabic stimuli were collected. One point was given for each correct tone in each section of the tone perception tests. One token from the monosyllable section and two from the disyllable section were removed from the final analysis due to some ambiguity. As a result, the total number of points available from the monosyllable section was 63, and the total number from the disyllable section was 60. The total number of points received by each participant from each test was converted into an accuracy rate to make the results comparable across different sections and different tests.

The coding of the semi-open-ended questionnaire was done using a mixed method. For those questions with multiple choice answers, the percentage of each choice was calculated by compiling all the answers to each question and then calculating the percentage of each choice. For the open-ended questions, all answers were compiled and read through to identify patterns and categorize them.

3.3 Results

In this section, the results of the tone perception tasks from the pre-test, immediately administered post-test, and delayed post-test and the results from the questionnaire will be reported.

3.3.1 Results of Tone Perception Tests

Tone perception accuracy rates were used as the measurement on all tone perception tests. An overall score from both the monosyllabic words section and disyllabic words section was calculated for each participant.

3.3.1.1 Overall Results of Tone Perception Tests

The overall accuracy rates of both groups on the pre-test, immediately administered post-test, and delayed post-test are shown in Figure 2 and Table 1. These results answer the first research question. Both groups showed improvement in their perception of tones from pre-test to the immediately administered post-tests. The tone perception accuracy rate of the experimental group increased from 60.9% on the pre-test to 78.8% on the immediately administered post-test and 85.9% on the delayed post-test. Similarly, the tone perception accuracy rate of the traditional group increased from 53.8% to 76% on the immediately administered post-test and 78.6% on the delayed post-test. The experimental group performed better on perceiving tones than did the students in the traditional group in absolute value. In order to see whether the differences between the experimental group and the traditional group within each test were statistically significant, a one-way repeated measures ANOVA was conducted with three levels: pre-test, immediately administered post-test, and delayed post-test. The results of homogeneity of variances tests were not significant at any level, which indicates that the two groups were comparable at any level. However, the null hypothesis of sphericity in Mauchly's test was rejected (p=.006), therefore, the Greenhouse-Geisser adjustment was used. The results indicated that there was no significant interaction between the time and group: $F(1.53, 44.34) = .872$, $p = .399$. Therefore, the null hypothesis was accepted for research question 1. In other words, there was no difference between the two groups in their total acuity of tone perception across tests.
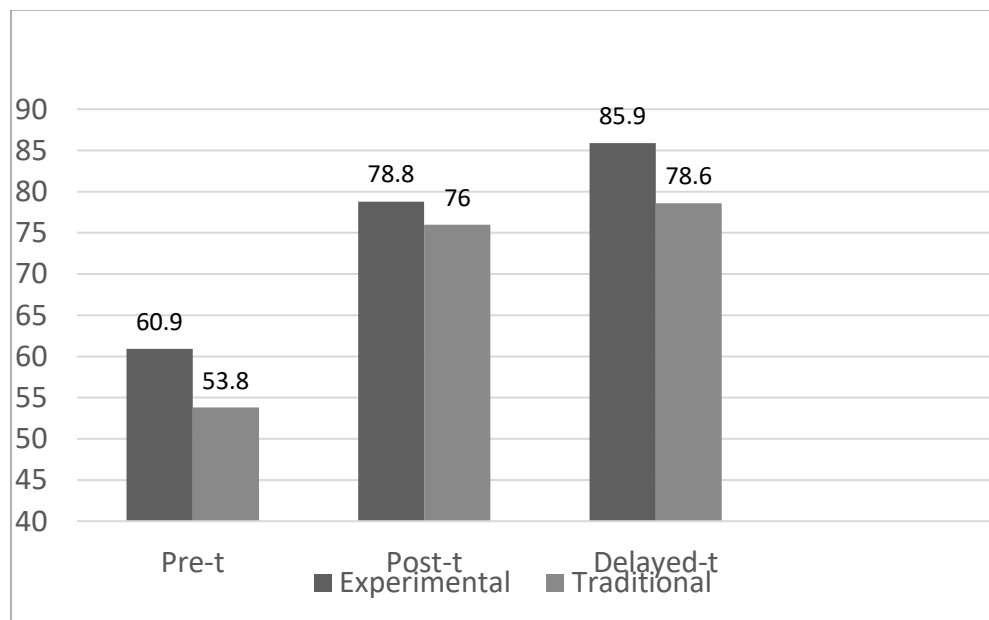

Figure 2 Overall Tone Perception Accuracy Across Groups

Table 1 Group Perception Accuracy Across Time: Total Scores

Time	Group	Pre-test		Post-test		Delayed-post-test	
		M	SD	M	SD	M	SD
Total	Experimental	60.9	13.0	78.8	9.2	85.9	5.8
	Traditional	53.8	14.1	76.0	9.6	78.6	7.4

3.3.1.2. Tone Perception on Monosyllabic Words and Disyllabic Words

The second research question focused on participants' perception accuracy in two different settings. Figures 3 and 4 show the tone perception accuracy rates on monosyllabic words and disyllabic words respectively, and detailed descriptive statistics are summarized in Table 2. As can be seen in Figure 3, both groups of learners performed at an accuracy rate of over 68.3%, even on the pre-test on monosyllabic words after only a brief introduction to tones. The accuracy rate increased to over 92.4% on the immediately administered post-test and to over 94.4% on the delayed post-test.

Figure 4 shows participants' accuracy rates on tones of disyllabic words. Compared to learners' performance in a monosyllabic setting, their performance in the disyllabic setting was not ideal. On the pre-test, the accuracy rate was 38.6% for the traditional group, and 45.8% for the experimental group. On the immediately administered post-test, the accuracy rate for the traditional group was 58.8% and 63.5 % for the experimental group, and the accuracy rate increased only slightly on the delayed post-test (62.1% for traditional group and 73.2% for the experimental group).
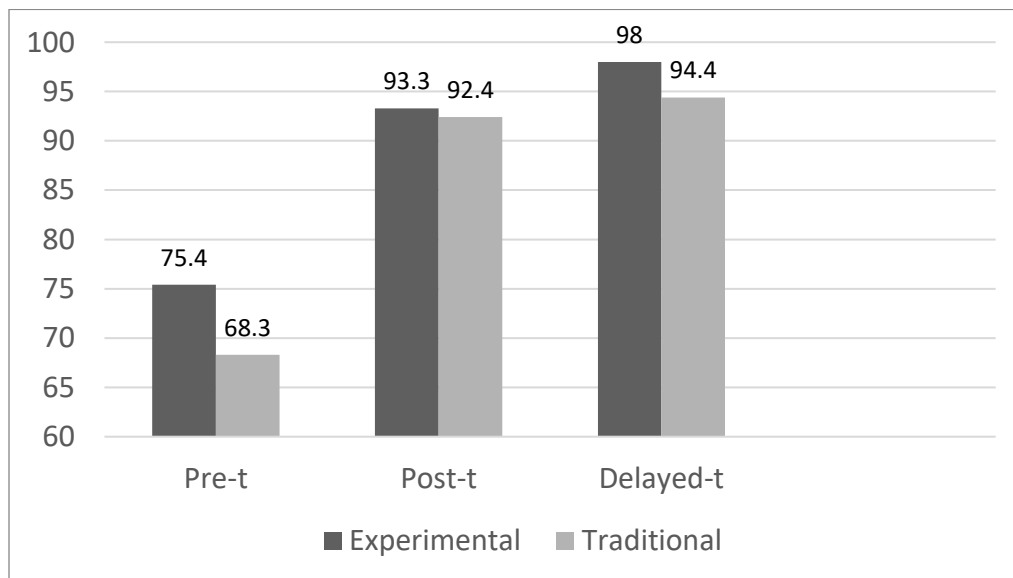

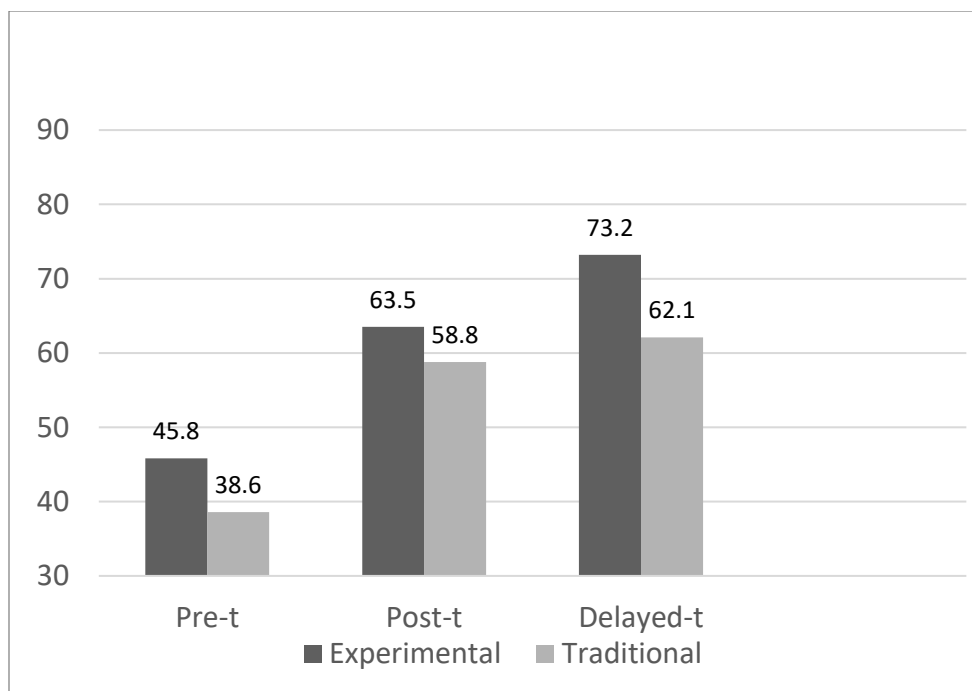**Figure 3 Perception Accuracy on Monosyllabic Words Across Groups**

Figure 4 Perception Accuracy on Disyllabic Words Across Groups

Table 2 Group Perception Accuracy Across Time in Different Settings

Time	Group	Pre-test		Post-test		Delayed-post-test	
		M	SD	M	SD	M	SD
Monosyllabic words	Experimental	75.4	13.6	93.3	8.3	98.0	2.2
	Traditional	68.3	19.6	92.3	6.7	94.4	6.0
Disyllabic words	Experimental	45.8	14.3	63.5	12.2	73.2	11.2
	Traditional	38.6	9.6	58.8	13.9	62.1	9.5

In order to test if there were any significant differences between the two groups' performances on monosyllables and disyllables across tests, a doubly multivariate repeated measures ANOVA was conducted. There were repeated measures over time (pre-test, post-test, and delayed post-test), the independent variable was the group, and the dependent variables were their accuracy scores for monosyllabic and disyllabic words. The results indicated that overall there was no significant difference between the two groups in the acuity of their tone perception in the monosyllabic or disyllabic settings across tests: Wilks' Lambha = .935, $F(4, 114) = .977$, $p = .423$. When broken down into different settings, there was no significant difference between the groups either. In the monosyllabic setting, Mulchly's test rejected sphericity ($p < .001$), so the Greenhouse-Geisser-adjusted results were used. There was no interaction effect between the time and group: $F(1.33, 38.65) = .654$, $p = .466$. In the disyllabic setting, Mulchly's test of sphericity was not significant ($p = .186$). There was no interaction effect between the time and group on tones for disyllabic words: $F(2, 58) = 1.248$, $p = .295$. Therefore, the second null hypothesis was accepted. In other words, the results indicated that both groups showed a similar pattern in their performance in the two different settings across the tests.

3.3.1.3. Tone Perception Accuracy on Different Tones

The results of research question 3 are reported in this section. To discover whether online tone training has an equal effect as the traditional training on different tones, participants' tone perception accuracy in the monosyllabic setting and disyllabic setting was further broken down by tone. Table 3 summarizes the groups' tone perception accuracy on different tones across the tests in the monosyllabic setting. On the pre-test, the experimental group performed best on perceiving the 1st tone (93.8%) and the traditional group performed best on perceiving the 3rd tone (79%). On the immediately administered post-test, both groups performed best on perceiving the 1st tone: the experimental group achieving a 98.1% accuracy rate and the traditional group 94.9%. Both groups recorded their worst performance in perceiving the 2nd tone with the experimental group achieving an 85.7% accuracy rate and the traditional group 85.5%. The same pattern was displayed on the delayed post-test: both groups performed the best in perceiving the 1st tone (the experimental group 99.7%; the traditional group 98.3%) and the worst in perceiving the 2nd tone (the experimental group 96.3%; the traditional group 87.9%).

Table 3 Group Perception Accuracy on Different Tones in Monosyllabic Words across Tests

Time	Group	Pre-test		Post-test		Delayed-post-test	
		M	SD	M	SD	M	SD
Tone 1	Experimental	93.8	8.1	98.1	7.1	99.7	1.4
	Traditional	73.9	25.0	94.9	6.1	98.3	4.0
Tone 2	Experimental	51.3	26.5	85.7	20.2	96.3	5.1
	Traditional	49.7	19.6	85.5	12.6	87.9	19.0
Tone 3	Experimental	84.7	17.4	95.3	7.6	97.5	4.3
	Traditional	79.0	20.4	93.2	8.6	93.8	8.4
Tone 4	Experimental	70.3	25.6	94.4	8.3	98.4	4.9
	Traditional	68.2	28.2	92.0	10.5	97.2	4.3

Participants' average accuracy in perception of different tones in the disyllabic setting is shown in Table 4. On the pre-test, both groups performed the best on perceiving the 1st tone with the experimental group achieving an accuracy rate of 55% and the traditional group 42.6%. Both groups performed the worst on the 2nd tone with the experimental group achieving an accuracy rate of 38.8% and the traditional group 31.8%. On the immediately administered post-test, both groups still performed the worst on the 2nd tone (experimental group 52.5%; traditional group 51.7%), and the experimental group still performed the best on perceiving the 1st tone (73.8%), while the traditional group performed the best on the 4th tone (65.9%). On the delayed post-test, both groups performed the best on the 4th tone (the experimental group 81.9%; the traditional group 69.9%). The experimental group performed the worst on the 3rd tone (60.4%) while the traditional group performed the worst on 2nd tone (51.7%).

Table 4 Group Perception Accuracy on Different Tones in Disyllabic Words Across Tests

Time	Group	Pre-test		Post-test		Delayed-post-test	
		M	SD	M	SD	M	SD
Tone 1	Experimental	55.0	18.1	73.8	15.5	78.4	18.4
	Traditional	42.6	12.4	64.2	19.8	67.0	14.0
Tone 2	Experimental	38.8	17.2	52.5	20.6	68.4	13.1
	Traditional	31.8	8.6	51.7	24.5	51.7	22.0
Tone 3	Experimental	45.4	22.0	54.6	17.2	60.4	16.0
	Traditional	34.8	18.2	52.3	18.3	53.8	9.4
Tone 4	Experimental	43.8	18.8	71.6	16.7	81.9	15.3
	Traditional	43.8	21.8	65.9	21.7	69.9	20.7

In order to test if the two groups performed differently among the tones in the monosyllabic and the disyllabic settings, a doubly multivariate repeated measures ANOVA was conducted. There were repeated measures for the tones (T1, T2, T3, and T4) and the time (pre-test, post-test, and delayed post-test), the independent variable was the group, and the dependent variable was their accuracy scores for both settings (monosyllabic vs. disyllabic).

The multivariate test results indicated that overall there was no significant difference between the two groups in the acuity of their perception of different tones in the monosyllabic and disyllabic settings across tests, because there was no significant interaction effect between tone, time, and group: Wilks' Lambha = .910, $F(12, 346) = 1.394$, $p = .166$. The reason being that when broken down into different settings, there was no significant interaction between tone, time, and groups for the univariate tests. In the monosyllabic setting, Mulchly's test of sphericity was significant ($p < .001$), therefore the Greenhouse-Geisser adjustment was used. There was no interaction effect among tone, time, and group: $F(3.63, 105.21) = 1.905$, $p = .383$. In the disyllabic setting, Mulchly's test of sphericity was not significant ($p = .538$) and there was no interaction effect between tone, time, and group either: $F(6, 174) = 1.069$, $p = .383$. Therefore, the third null hypothesis was accepted and the alternative hypothesis was rejected. In summary, the results indicated that both groups showed a similar pattern in their performance on different tones in the two different settings across the tests.

Although no significant interaction effect was found between the groups across the tones, an interesting trend was noticed regarding T2 in both monosyllabic and disyllabic settings. After we singled out the accuracy rates of T2 and compared the change of each group's acuity across the tests, the following graph (Figure 5) was created:
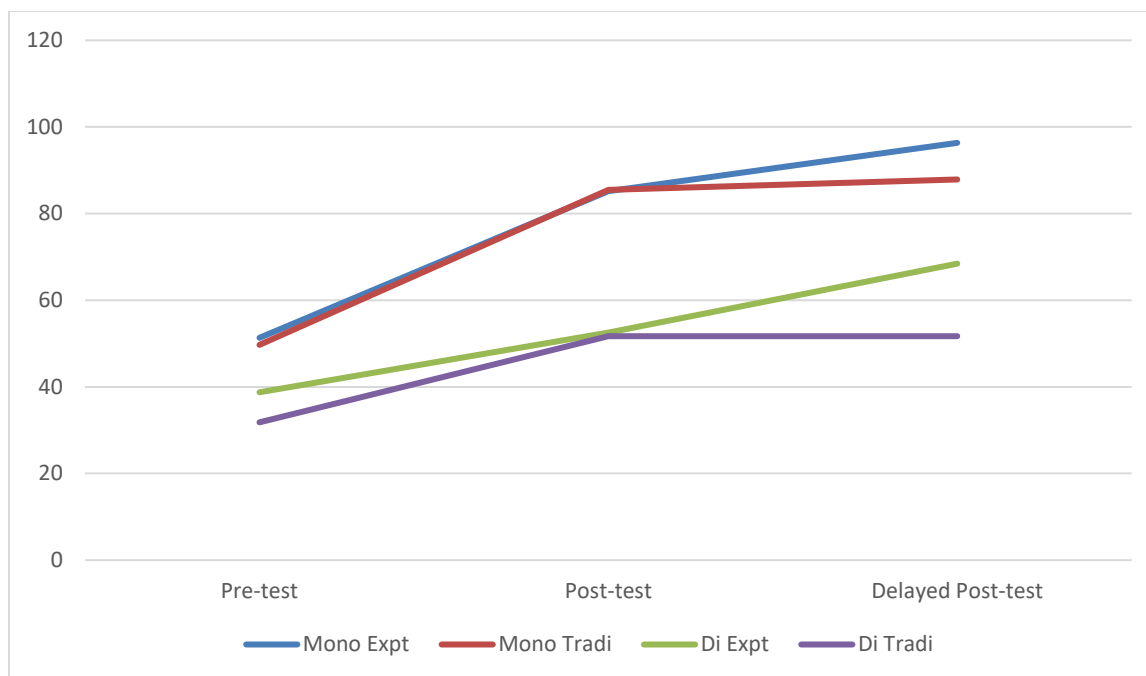
Figure 5 Perception Accuracy of T2 in Different Settings Across Time

Both groups started around the same accuracy, and they reached similar accuracy at the end of the training sessions. However, the experimental group's accuracy kept improving, whereas the traditional group's accuracy stopped increasing.

3.3.2 Questionnaire Results

The questionnaire focused on Chinese learners' self-assessment of the difficulties they have in perceiving tones, their tone learning strategies, their experience with the online tone training practice, and their preference regarding using online or traditional tone training practices in the classroom. Only learners in the experimental group completed the questionnaire. Because the questionnaire is qualitative in nature, 24 responses[2] were included in the final analysis.

3.3.2.1 Students' Perception of Different Tones

Part I of the questionnaire elicited students' perception of different tones: 62.5% of the participants rated the 1st tone as the easiest one to identify while 25% of the participants rated the 3rd tone and 12.5% rated the 4th tone as the easiest. Nobody rated the 2nd tone as easy. As for the most difficult tone to identify: 58% of the participants rated the 2nd tone and 25% of participants rated the neutral tone as being most difficult to accurately perceive while 8.3% of participants rated the 3rd and 4th tones as being the hardest to identify. This

---

[2] Four of these subjects' tone perception data were not included in the analysis of tone perception tests because they either missed the pre-test or the post-test. We included their questionnaire responses here because we believe the questionnaire is qualitative in nature and these subjects' experiences using the online tone training websites are as valuable as the other Chinese learners in the experimental group.

was consistent with students' performance on the tone perception tasks: their accuracy was highest on the 1st tone and lowest on the 2nd tone. These results are in line with the results of other experimental studies that found the 2nd tone and the 3rd tone to be harder than the 1st or 4th tone for students to perceive after training (Wang et al., 1999; Wang et al., 2003; Li, 2016; Sun, 1998; He & Wayland, 2013). Our participants also mentioned that the neutral tone was difficult to perceive. However, the neutral tone was not included in our study.

3.3.2.2 Student Strategies for Learning Tones

Part II of the questionnaire collected the strategies employed by students for learning tones. Not surprisingly, students employed a variety of strategies. Table 5 shows the percentage of students who used each strategy. Since students could check multiple items, the total number of strategies was tallied by adding up all items the participants had checked, then calculating the percentage for using a certain strategy.

The strategy used by most students (87%) was "listening to audio files" followed by "using online resources to practice" (83.3%). Almost 80% of the participants used "comparing what I hear with what I say" and 50% of the participants used the strategy "repeating after audio files." In addition, 29.2% of the participants employed a strategy of "practicing listening and reading pinyin with fellow students." Strategies other than those listed in the questionnaire were employed by 20.8% of participants.

Table 5 Student Strategies for Learning Tones

Tone learning strategies	Frequency percentage
Listening to audio files	87.5%
Using online resources to practice	83.3%
Comparing what I hear with what I say	79.2%
Repeating after audio files	50%
Practicing listening to and reading pinyin with fellow students	29.2%
Other	20.8%

3.3.2.3 Students' Computer-assisted Language Learning (CALL) Experience

Part III of the questionnaire focused on students' computer-assisted language learning (CALL) experience. All participants liked the website www.pinyinpractice.com, and thought it was both useful and helpful, especially when it came to improving their perception of tones in disyllabic words where tones are not as obvious in actual speech as they are when demonstrated by an instructor whose enunciation is more deliberate. Some participants felt frustrated using the website, especially at the beginning, but with greater familiarity came greater confidence. A large majority (83.3%) of the participants reported that they used the website outside of class for 15-30 minutes, two to three times a week.

Less than half (41.6%) of the participants also tried other features offered by www.pinyinpractice.com, although they had not been instructed to do so. A little over half

(54.2%) of the participants expressed an interest in trying other features in the future. However, 58.3% of the participants indicated that they would prefer to practice other features in class before using them outside of class.

In general, students' attitudes towards www.pinyinpractice.com were positive. They liked the set-up and the simple straightforward interface. They also appreciated other features such as instant feedback, the syllable replay option, and the variety of input methods (choosing or entering tones). At the same time, the participants also offered their thoughts on improvements that they would like to see on the website. These included an option for revealing the correct tones after a few attempts. The current interface only gives very general feedback such as "right" or "wrong," and does not reveal the correct answer until the learner gets it right. Nor does it give feedback on any specific error that students might have made. This vagueness sometimes caused anxiety and frustration for the students. Other suggestions included a way for the missed tones to be marked in the instant feedback, the ability to adjust the speed at which each syllable is read, and an option to enter tones in a way other than using numbers, such as to input the tone marks as diacritics.

When participants were asked whether they preferred traditional in-class instructor-led tone practice or self-paced on-line tone practice, 67.57% of the participants said they preferred traditional in-class tone practice while only 32.43% preferred online practice. Students who preferred in-class instructor-led tone practice appreciated the interactive nature of this method, pointing out that such practice allowed them to say the syllable aloud and get instant feedback from the instructors. Students who preferred the online tone practice liked the fact that they could practice at their own pace, and practice as many times as they wanted to without feeling intimidated by the perceived judgment of their instructor or classmates when they made mistakes.

3.4 Discussion

The first research question examined the overall performance of participants who received online training as compared to those who received traditional in-class tone training. The results of the tone perception tests indicate that L2 Chinese learners who received the online tone training in class showed equal acuity in tone perception compared to students who receive traditional instructor-led tone training in class. In terms of gains in acuity, the experimental group showed an increase of 18.1% on the immediate post-test and 25% on the delayed post-test respectively. This finding is similar to the training gains reported in Godfroid et al. (2017) of 10-20%. All participants in the experimental group and the control group practiced tones during the training period. The difference between the traditional group and the experimental group lay in the format of such tone training, i.e. instructor-led vs. self-paced online tone practice. Both groups showed improved tone perception after the training period was over as can be seen on the immediate post-test. Therefore, we can see that tone training, no matter what the format might be, has an immediate effect on tone perception, a finding which is consistent with those reported in tone training literature (Wang et al., 1999, 2003; Li, 2016; Godfroid, et al. 2017). The online self-paced tone practice helped the students as much as the instructor-led tone practice as far as their tone perception is concerned.

The second research question examined whether L2 Chinese learners who received the online tone training in class improve differently in tone perception on monosyllabic words or on disyllabic words than those who received traditional instructor-led tone training in class. The results indicated that the experimental group's acuity improved significantly over time in both settings and they showed a similar pattern of improvement as the traditional group. The two groups did not differ significantly from each other in the acuity of their tone perception in the monosyllabic or the disyllabic settings across tests. The results also indicated that both groups' tone perception accuracy on monosyllabic tokens was higher than their accuracy on disyllabic tokens. This finding is consistent with findings from previous studies, i.e. the tones of monosyllabic words are easier for Chinese learners to identify than those of disyllabic words (He & Wayland, 2013; Hao, 2012; Li, 2016). According to Xu (1997), the tones of monosyllabic words are more canonical and stable while tones of disyllabic words are always affected by adjacent tones making them more difficult to discern. The two groups of learners showed the same patterns in their performance in perceiving the tones of monosyllabic words and disyllabic words. This indicates that tone-training practice, no matter whether it is online or in-class instructor-led, is effective in helping L2 Chinese learners correctly perceive tones in both monosyllabic and disyllabic words.

The third research question addressed whether students who received online tone training showed a similar pattern as their peers who received traditional instructor-led tone training in their perception across the tones in both the monosyllabic and disyllabic settings. The tone perception tests' results indicate that improvement was not just dependent on time but also on specific tone: students performed very well in identifying the T1 on the pre-test even after just a short introduction to tones, but they did not do as well in perceiving T2, T3, and T4. Through training, they improved their perceptions on T2, T3, and T4 in the monosyllabic setting greatly, as well as in the disyllabic setting. However, the pattern of improvement of different tones across time looked similar between the online application experimental group and the traditional instruction group. What is noteworthy is that for T2, the experimental group' acuity continued to improve after the training ended in both the monosyllabic and disyllabic settings, whereas the traditional group' acuity stopped improving, even though the difference is not significant. Since T2 was the most difficult tone to perceive as indicated from both students' performance in the tests and their self-evaluation in the questionnaire, the results are very encouraging.

Why did the online application experimental group keep improving their perception of T2? Based on the information provided by the questionnaire, we offer this possible explanation: while the traditional group stopped practicing tones after the classroom tone-training period was over, some participants in the experimental group continued using the website to practice tones. A majority of the participants reported that they used the website outside of class for 15-30 minutes, two to three times a week, although the participants did not indicate whether they devoted extra time to practice outside of class during the experiment period or after it. Regardless of when they practiced, the significance of introducing the online tone practice website to the students in the classroom not only has immediate, short-term, tone training benefits, but also has extended, long-term benefits, particularly as a tool students can use outside the classroom to help them develop autonomy. Although the reported extra use of the online practice made the present study less

controlled in the sense that the experimental group spent more time practicing their perception of tones, the fact that these participants did, in fact, practice outside class without being told to is encouraging. Another explanation could be that the self-paced online training gave students greater exposure to the tones that they found more difficult, making their training more targeted and individualized and thus increasing retention.

The fourth research question investigated students' experience using the online training website. Overall, the students had a positive attitude toward the pinyin practice website. They acknowledged that the website was helpful and a majority (83.3%) used online-training in class as well as outside the classroom. However, according to their responses, more students preferred in-class instructor-led tone practice to online self-paced practice (67.57% vs. 32.43%). Several factors may contribute to this preference. One of the most obvious differences between the traditional classroom instructor-led practice and the online tone-training practice is that students interact with human beings in the former while interacting with computers in the latter. Secondly, the feedback they get from the instructor is different than the feedback they get from the computer. Although both can be counted as immediate feedback, that from the instructor is more specific, targeting the errors that the students made in production, whereas the latter focuses only on whether the tone the student perceived was correct or not; in other words, the traditional classroom instructor-led practice focuses on production while the online tone-training is only about perception. Students may feel that they are more engaged in the traditional classroom instructor-led practice than in the online tone-training activities.

4. Conclusion and Pedagogical Implications

Overall, the current study shows that using online applications helped students perceive tones as effectively as the traditional instructor-led training in class in terms of overall acuity, acuity in monosyllabic or disyllabic words, or acuity with individual tones. The introduction of online practice tools such as the one used in the current study is beneficial in that it gives students access to a platform on which they can —and are motivated to—continue working on their tone perception even after the formal pinyin training period has ended. It thus can have long-term positive effects in the training of students in correct tone perception. The fact that students voluntarily used the website outside of the classroom to practice tones is also encouraging. Although the introduction of the tone-training website in class is an effective alternative, such online tone training practice is, perhaps, better integrated into the curriculum as an assignment rather than as a classroom activity. In other words, instructor-led practice in the classroom is still necessary and well-received, while the online applications can be used as an effective alternative to help students improve their tone perception outside of class since there is not enough time to be spent on tone training in class.

Although it has not been empirically proven, it can be inferred from the current study that assigning tone training three times a week for four weeks after the completion of the initial pronunciation practice could be beneficial. To consolidate the effects, a tone perception test given in the fifth week might be desirable.

5. Limitation and Future Studies

Although the authors did their best to recruit as many participants as possible, the sampling in this study was not large enough to reach more solid conclusions, especially when the large variation within each group was taken into consideration. The training tool used in this study was not perfect either. For example, it only gave a general feedback such as "right" or "wrong," without targeting any specific error that students might have made. This vagueness made the training less efficient and raised students' anxiety and frustration in some cases, as was indicated in the findings of the questionnaire. In the future, if a more efficient and targeted online application can be used to explore this issue, a better result may be obtained. In addition, recent studies on Chinese tone training show that perceptual training has a clear, facilitative effect on both the perception and production of tones (Leather, 1990; Wang et al., 1999; Wang et al., 2003). More studies on the tone perception training effects on tonal production are hence clearly called for.

References

Cai, W., & Sciban, S.-N. (2008). How is Chinese grammar taught? A comparison of grammar materials in five Chinese textbooks. *Journal of the Chinese Language Teachers Association, (43)*1, 17-31.

Chun, D. M. (2015). Language and culture learning in higher education via telecollaboration. *Pedagogies: An International Journal*, *10*(1), 5-21.

Chun, D. M. (2017). Research methods for investigation technology for language and culture learning. In C. Chapelle & S. Sauro (Eds.), *The handbook of technology in second language teaching and learning* (pp. 393-408). Hoboken, NJ: Wiley-Blackwell.

Chun, D. M., Jiang, Y., Meyr, J., & Yang, R. (2015). Acquisition of L2 Mandarin Chinese tones with learner-created tone visualizations. *Journal of Second Language Pronunciation*, *1*, 86-114.

Derwing, T. M., & Munro, M. J. (2015). *Pronunciation fundamentals: Evidence-based perspectives for L2 teaching and research* (Vol. 42). Amsterdam: John Benjamins Publishing Company.

Francis, A. L., Ciocca, V., Ma, L., & Fenn, K. (2008). Perceptual learning of Cantonese lexical tones by tone and non-tone language speakers. *Journal of Phonetics*, *36*(2), 268-294.

Godfroid, A., Lin, C. H., & Ryu, C. (2017). Hearing and seeing tone through color: An efficacy study of web-based, multimodal Chinese tone perception training. *Language Learning, 67*(4), 819-857.

Hao, Y-C. (2012). Second language acquisition of Mandarin Chinese tones by tonal and non-tonal language speakers. *Journal of Phonetics*, *40*(2), 269-279.

He, Y., & Wayland, R. (2013). Identification of Mandarin co-articulated tones by inexperienced and experienced English learners of Mandarin. *Chinese as a Second Language Research*, *2*(1), 1-21.

Herd, W., Jongman, A., & Sereno, J. (2013). Perceptual and production training of intervocalic/d, ɾ, r/in American English learners of Spanish. *The Journal of the Acoustical Society of America*, *133*(6), 4247-4255.

Hockett, C. F. (1951). *Progressive exercises in Chinese pronunciation* (Vol. 2). Institute of Far Eastern Languages, New Haven: Yale University.

Hubbard, P. (Ed.). (2009). *Computer assisted language learning: Critical concepts in linguistics. Present trends and future directions in CALL.* New York: Routledge.

Jongman, A., Wang, Y., Moore, C. B., & Sereno, J. A. (2006). Perception and production of Mandarin Chinese tones. In P. Li, L. Tan, E. Bates, & O.J.L. Tzeng (Eds.), *The handbook of East Asian psycholinguistics: Volume 1, Chinese*, (pp. 209-217). Cambridge: Cambridge University Press.

Kingston, J. (2003). Learning foreign vowels. *Language and Speech, 46*(2-3), 295-348.

Leather, J. (1990). Perceptual and productive learning of Chinese lexical tone by Dutch and English speakers. *New sounds*, *90*, 72-97.

Lee, C. Y., Tao, L., & Bond, Z. S. (2010). Identification of multi-speaker Mandarin tones in noise by native and non-native listeners. *Speech Communication*, *52*(11), 900-910.

Levis, J. (2007). Computer technology in teaching and researching pronunciation. *Annual Review of Applied Linguistics, 27*, 184-202.

Li, Y. L. (2016). *Effects of high variability phonetic training on monosyllabic and disyllabic Mandarin Chinese tones for L2 Chinese learners*. Retrieved from ProQuest Dissertations and Theses Global. (1800269699).

Logan, J. S., Lively, S. E., & Pisoni, D. B. (1991). Training Japanese listeners to identify English/r/and/l: A first report. *The Journal of the Acoustical Society of America*, *89*(2), 874-886.

Miracle, W. C. (1989). Tone production of American students of Chinese: A preliminary acoustic study. *Journal of the Chinese Language Teachers Association*, *24*(3), 49-65.

Orton, J. (2013). Developing Chinese oral skills: a research base for practice. In I. Kecskés (Ed.), *Research in Chinese as a second languag*e (pp. 3-26). Berlin: De Gruyter Mouton.

Rochet, B. L. (1995). Perception and production of second-language speech sounds by adults. In W. Strange (Ed.), *Speech perception and linguistic experience: Issues in cross-language research* (pp. 379-410). Baltimore: York Press.

Rohr, J. (2014). Training naïve learners to identify Chinese tone: An inductive approach. In N. Jiang (Ed.), *Advances in Chinese as a Second Language: Acquisition and Processing*, (pp. 158-178). Newcastle: Cambridge Scholars Press.

Shen, X. S. (1989). Toward a register approach in teaching Mandarin tones. *Journal of the Chinese Language Teachers Association*, *24*(3), 27-47.

Shen, X. S., & Lin, M. C. (1991). A perceptual study of Mandarin tones 2 and 3. *Language and Speech, 34*, 145–156.

So, C. K. L. (2006). *Effects of L1 prosodic background and AV training on learning Mandarin tones by speakers of Cantonese, Japanese, and English*. Ph.D. dissertation, Simon Fraser University.

Sun, S. H. (1998). *The development of a lexical tone phonology in American adult learners of standard Mandarin Chinese* (No. 16). Honolulu: University of Hawaii Press.

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

Wang, Y., Jongman, A., & Sereno, J. (2003). Acoustic and perceptual evaluation of Mandarin tone production before and after training. *Journal of the Acoustical Society of America*, *113*, 1033-1043.

Xing, J. Z. (2006). *Teaching and learning Chinese as a foreign language: A pedagogical grammar*. Hong Kong: Hong Kong University Press.

Xu, H. (2017). *Self-paced use of online applications on Chinese tone perception among beginners in Chinese*. Paper presented at the 2017 UW Spring Conference on Teaching and Learning, La Crosse, WI.

Xu, Y. (1997). Contextual tonal variations in Mandarin. *Journal of Phonetics*, *25*(1), 61-83.

Zhang, H. (2018). Current trends in research of Chinese sound acquisition. In C. Ke (Ed.), *The Routledge handbook of Chinese second language acquisition*. London & New York: Routledge Taylor & Francis Group.

Appendix A: Language background information sheet

English Name: _____   Email: _____

Phone#: _____   School year: _____

Major: _____

Please answer the following questions.

1. Do you have any experience studying Chinese? If the answer is yes, please specify where and when. *Please remember to talk to your lecture teacher if you have any background speaking Chinese, whether or not you have taken placement test.

2. Have you ever been to China? If the answer is yes, please specify where and when.

3. Do you speak any Chinese dialects such as Cantonese etc.?

4. Have you learned any other foreign languages (excluding Mandarin Chinese)? If have, how long?

5. Do you have a Chinese name? If you do, what is it? Please write it down in Chinese characters or describe the meaning of the characters. If you do not have a Chinese name, do you have any preference?

Appendix B-1: Test stimuli used in the pre-test, post-test, and delayed post-test (the syllables were randomized in each test)

Monosyllables:

1. ji ā	2.ji á	3.ji ǎ	4.ji à
5. w ān	6.w án	7.w ǎn	8.w àn
9. li ū	10.li ú	11.li ǔ	12.li ù
13.sh ēn	14.shén	15.shěn	16.shèn
17. xu ē	18. xué	19.xuě	20.xuè
21.tu ō	22.tuó	23. tuǒ	24. tu ò
25.ku ī	26.tu í	27.tuǐ	28.tuì
29.t āo	30.táo	31. tǎo 3	2.t ào
33. huān	34.huán	35. hu ǎ n	36. huàn
37.shī	38.shí	39. shǐ	40.shì
41.wā	42.wá	43.wǎ	44.w à
45.xī	46.xí	47.xǐ	48. xì
49.chōng	50.chó ng	51.chǒng.	52chòng
53.yōu	54.yóu	55.yǒu	56.yòu
57. pū	58.pú	59.pǔ	60.pù
61. yāng	62. yáng	63. yǎng	64. yàng

Disyllables

1. xīshǐ	2. huánliù	3.pùxuē	4.jiàkuǐ
5. chóngshí	6. tǎojiā	7. xuěyòu	8. tuōxí
9. wǎtuǒ	10.yóuliǔ	11.tuóxǐ	12.shěnhuān
13. liūhuǎn	14.púshī	15.wāhuàn	16.chǒngshì
17.yǒushén	18.jiǎwǎn	19.wānyōu	20.xìxué
21.yángxuè	22.wànwán	23.jiákuí	24.shèntuò

25.yāngwà   26.pūchōng   27.yǎngwá   28.kuìshēn

29. tàopǔ   30.kuīliú   31.chòngyàng 32.táotāo

Appendix B-2: Test sheet used in the study

Tone Perception

Name: _____ Lecture time (circle one): 9AM/10AM

You are going to hear some syllables. Please indicate the tone of each syllable you hear by circling the tone marker provided. For example, if you see the following syllable: ma (–   /   V   \ ) , and then you hear: mǎ. You circle the third tone: "v": ma (–   /   V   \ ) .

Part I: Monosyllables

1. jia (– / V \ )	24. liu (– / V \ )
2. shi (– / V \ )	25. kui (– / V \ )
3. shen (– / V \ )	26. tao (– / V \ )
4. kui (– / V \ )	27. xue (– / V \ )
5. xue (– / V \ )	28. huan (– / V \ )
6. shen (– / V \ )	29. tuo (– / V \ )
7. huan (– / V \ )	30. yang (– / V \ )
8. xi (– / V \ )	31. tuo (– / V \ )
9. xue (– / V \ )	32. xi (– / V \ )
10. shen (– / V \ )	33. pu (– / V \ )
11. wa (– / V \ )	34. you (– / V \ )
12. wan (– / V \ )	35. yang (– / V \ )
13. yang (– / V \ )	36. pu (– / V \ )
14. shi (– / V \ )	37. you (– / V \ )
15. chong (– / V \ )	38. wa (– / V \ )
16. yang (– / V \ )	39. chong (– / V \ )
17. liu (– / V \ )	40. pu (– / V \ )
18. shi (– / V \ )	41. kui (– / V \ )
19. xue (– / V \ )	42. wan (– / V \ )
20. huan (– / V \ )	43. liu (– / V \ )
21. jia (– / V \ )	44. you (– / V \ )
22. huan (– / V \ )	45. jia (– / V \ )
23. wa (– / V \ )	46. xi (– / V \ )

47.   tao (–   /   V   \ )
48.   pu (–   /   V   \ )
49.   chong (–   /   V   \ )
50.   tuo (–   /   V   \ )
51.   wan (–   /   V   \ )
52.   you (–   /   V   \ )
53.   shen (–   /   V   \ )
54.   tao (–   /   V   \ )
55.   tuo (–   /   V   \ )
56.   wa (–   /   V   \ )
57.   xi (–   /   V   \ )
58.   kui (–   /   V   \ )
59.   tao (–   /   V   \ )
60.   liu (–   /   V   \ )
61.   shi (–   /   V   \ )
62.   chong (–   /   V   \ )
63.   jia (–   /   V   \ )
64.   wan (–   /   V   \ )

Part II: Disyllables

1.   xi (–   /   V   \ )   shi (–   /   V   \ )

2.   huan (–   /   V   \ ) liu (–   /   V   \ )

3.   pu (–   /   V   \ )   xue (–   /   V   \ )

4.   jia (–   /   V   \ )   kui (–   /   V   \ )

5.   chong (–   /   V   \ )   shi (–   /   V   \ )

6.   tao (–   /   V   \ )   jia (–   /   V   \ )

7.   xue (–   /   V   \ )   you (–   /   V   \ )

8.   tuo (–   /   V   \ )   xi (–   /   V   \ )

9.   wa (–   /   V   \ )   tuo (–   /   V   \ )

10.   yang (–   /   V   \ ) wa (–   /   V   \ )

11.   tuo (–   /   V   \ )   xi (–   /   V   \ )

12.   shen (–   /   V   \ ) huan (–   /   V   \ )

13.   liu (–   /   V   \ )   huan (–   /   V   \ )

14.   pu (–   /   V   \ )   shi (–   /   V   \ )

15.   wa (–   /   V   \ )   huan (–   /   V   \ )

16.   chong (–   /   V   \ )   shi (–   /   V   \ )

17.   kui (–   /   V   \ )   shen (–   /   V   \ )

18.   jia (–   /   V   \ )   wan (–   /   V   \ )

19.   wan (–   /   V   \ ) you (–   /   V   \ )

20.   xi (–   /   V   \ )   xue (–   /   V   \ )

21.   yang (–   /   V   \ ) xue (–   /   V   \ )

22.   wan (–   /   V   \ ) wan (–   /   V   \ )

23.   jia (–   /   V   \ )   kui (–   /   V   \ )

24.   shen (–   /   V   \ )   tuo (–   /   V   \ )

25.   you (–   /   V   \ )   liu (–   /   V   \ )

26.   pu (–   /   V   \ )   chong (–   /   V   \ )

27.   yang (–   /   V   \ ) wa (–   /   V   \ )

28.   you (–   /   V   \ )   shen (–   /   V   \ )

29.   tao (–   /   V   \ )   pu (–   /   V   \ )

30.   kui (–   /   V   \ ) liu (–   /   V   \ )

31.   chong (– /   V   \ )   yang (–   /   V   \ )

32.   tao (–   /   V   \ ) tao (–   /   V   \ )

Appendix C: Tone learning experience questionnaire

Now that we have finished learning Pinyin, we would like you to share your learning experience with us. Your opinions will be invaluable for us to improve our curriculum and teaching in the future. Thank you very much!

1. According to your learning experience, you feel:

    a. The easiest tone to identify is (circle one): 1st, 2nd, 3rd, 4th, toneless

    b. The most difficult tone to identify is: (circle one): 1st, 2nd, 3rd, 4th, toneless

    c. Do you feel that it is harder to identify tones in disyllable words? (circle one): Yes
No.

2. The strategies that you used to help your tone learning (check all that apply)

a. Listen to the audio files of the pinyin exercises in the textbook

b. Repeat after the audio file

c. Compare what I hear with what I say

d. Practice listening and reading of pinyin with my fellow classmate(s) on a regular base, such as meeting one to two times a week.

e. Using available online resources to practice

f. Others: please specify_____

3. Computer-assisted learning experience

a. How do you like the interactive online practice (www.pinyinpractice.com) in helping you get the correct tones? How do you feel while you are doing the exercises?

b. How do you like the way the online exercises are set up? (Either to pick the tones or to supply the missing tones; to get instant feedback; etc.)

c. What improvement do you hope to see of the online exercises?

d. Which one do you prefer: Big group practice in class or on-line practice? Why?

e. Did you use the website, pinyinpractice.com, outside of class? Did you use similar applications to help you get the right tones? If yes, how often?

f. Have you shared the website (www.pinyinpractice.com) with any of your friend? Do you think you will continue to use online applications to help you learn Pinyin? If so, how often do you think you will use them?

g. Have you ever used any other online applications to help you learn a foreign language? If so, what kind of application?

 h. The website, Pinyinpractice.com, also has practices for initials and finals. Have you tried those practices?   Yes   No

If you haven't tried, do you plan to try? Yes No

Do you think we should use the website, Pinyinpractice.com, to practice initials and finals in class? Why or why not?