

Reframing the ‘Strange’ as ‘Familiar’ – Teaching Mandarin Chinese Tones to Speakers of Non-Tonal Languages

Angela Cook

The University of Queensland, Australia

Yao Li*

The University of Queensland, Australia

Dongchen Yao

The University of Sydney, Australia

Correspondence

Email: yao.li5@uq.net.au

Abstract

A major challenge of learning Mandarin Chinese is mastering lexical tones. Conventional methods of teaching Mandarin tones assume that, for speakers of non-tonal languages, lexical tones are a ‘strange’ concept. Our small-scale pilot study drew on pedagogical theories of cross-linguistic transfer (CLT) to frame lexical tones as ‘familiar’ by modelling Mandarin pitch contours on English intonation patterns. We compared an experimental group with a control group, within a first-semester university Chinese language course. All participants completed a pre-test and a post-test (perception and production of tones), and a pre-survey and post-survey (attitudes towards learning). After 30-40 minutes of additional training, students in the experimental group demonstrated no significant improvement in perception or production of Mandarin tones vis-à-vis the control group but differentiated themselves in some interesting attitudinal ways. The results of the study suggest that it would be worth conducting further research into the new pedagogical approach on a larger scale.

ARTICLE HISTORY

Received: 15 March 2026

Revised: 06 September 2026

Accepted: 20 September 2026

KEYWORDS

CFL/CSL Teaching, Lexical Tones, Innovative Pedagogy, Non-Tonal Language Speakers, Tone Perception and Production

How to cite this article (APA 7th Edition):

Cook, A., Li, Y., & Yao, D. (2026). Reframing the ‘strange’ as ‘familiar’ – teaching Mandarin Chinese tones to speakers of non-tonal languages. *Language Teaching Research Quarterly*, 56, 108–131. <https://doi.org/10.32038/ltrq.2026.56.06>

Introduction

With nearly one billion native speakers around the globe, Mandarin Chinese is the most commonly spoken first language in the world, outranking its closest rival – Spanish – two to one (Eberhard et al., 2025). There has probably never been a time in the history of

humanity when such a large proportion of the world's population has shared a common mother tongue. This fact alone is sufficient to inspire many people from different language backgrounds with a desire to learn Mandarin Chinese. Add to this the rapid growth of the Chinese economy, the growing economic and trade ties between China and most OECD countries, and expanding cultural, scientific and technological exchanges at an international level, and it is no wonder that many people worldwide are actively seeking to acquire Mandarin Chinese cultural and linguistic capabilities¹.

A crucial element of effective communication is, of course, clear pronunciation. L2 learners who do not achieve at least an approximation of the conventions of pronunciation prevailing within the speech community of the language or dialect they are learning are likely to encounter frequent communication issues and misunderstandings. And since Mandarin Chinese is a tonal language, mastery of readily comprehensible pronunciation by L2 learners generally entails the reasonably standard realisation of Mandarin tones. At the same time, it is widely recognised that acquiring clear and comprehensible pronunciation can be a challenge for L2 learners of Mandarin Chinese (Gong et al., 2020; Li et al., 2021; Li et al., 2025; Yang, 2016). The fact that Mandarin Chinese learners have been found generally to prioritise the usefulness of listening (45%) and speaking (30%) over reading (15%) and writing (10%) (Renukadevi, 2014) underscores the importance of this aspect of Chinese as a foreign language (CFL) and Chinese as a second language (CSL) teaching.

Mandarin Chinese is one of the main foreign languages offered at universities and educational institutions in Australia, including at the University of Queensland (UQ), the site chosen for the current study. A range of conventional approaches has demonstrated pedagogical value in supporting learners' tone perception and production, including the audio-lingual method (Meng, 2021), gesture-based approaches (Baills et al., 2019), and visualisation-based approaches (Chun et al., 2013; Xu et al., 2019). However, evidence from both the literature and classroom experience suggests that these approaches may not fully address some persistent challenges faced by learners from non-tonal language backgrounds. These challenges include affective barriers, such as anxiety or apprehension about tone production, as well as difficulties with particular tonal categories, especially Tone 3. Therefore, in response to concerns voiced by students, and informed by personal Chinese teaching and learning experience, the authors developed a new method to facilitate the acquisition of Mandarin tones by students from non-tonal linguistic backgrounds, which they dubbed the 'YES' approach. Through a small-scale pilot project conducted with UQ students enrolled in their first semester of Mandarin Chinese, the current study aimed to test the effectiveness of the new approach, with the

¹ Reliable and transparent statistics on the numbers of people learning Chinese as a foreign language (CFL) worldwide are difficult to come by, with estimates ranging from two million to 100 million, depending in part on whether or not heritage language learners (who make up the bulk of CFL learners globally) are included.

long-term goal of providing an additional option to add to the repertoire of pedagogical strategies for teaching tones.

Literature Review

Mandarin Chinese is widely acknowledged to be one of the most difficult languages to learn as a second or foreign language, especially for L1 speakers of Indo-European languages such as English (US Foreign Service Institute, cited in Abdullah & Watson, 2021). While mastering Chinese orthography presents a major hurdle for learners from alphabetic language backgrounds wishing to improve their reading and writing skills, a key challenge within the macro skills of speaking and listening is believed to be the perception and production of lexical tones (Gong et al., 2020; Li et al., 2021; Xing, 2006; Zhang, 2019).

Not surprisingly, there is a large body of literature on different methods for teaching Mandarin Chinese tones, primarily focusing on gestures / kinaesthetic / embodied learning (e.g., Baills et al., 2019); audio-lingual methods (e.g., Meng, 2021); music-related and acoustic methods (e.g., Shi, 2018); computer-assisted / visualisation methods (e.g., Chun et al., 2013; Xu et al., 2019); and colour-coding (e.g., Godfroid et al., 2017). While these approaches have demonstrated varying degrees of pedagogical effectiveness, innovation in tone pedagogy has remained limited. Although AI development has expanded the possibilities for language learning, its benefits for Mandarin second language acquisition appear to remain modest, particularly in terms of tonal learning (Wang, 2026). These limitations point to the need for renewed pedagogical innovation that moves beyond established approaches and explores more adaptive ways of supporting L2 tone learning. What conventional methods have in common is that they are predicated on the assumption that, for speakers of non-tonal languages, lexical tones are a completely new and ‘strange’ concept. By contrast, the present study aims to propose an approach to transform conventional pedagogy and empower learners by surfacing their deliberate control of intonational pitch variation in their L1 (to express, e.g., mood, emotion, emphasis, illocutionary act) and supporting them to transfer these skills to the production of lexical tones in L2 Mandarin words (to express semantic meaning).

The literature review first defines lexical tone in the context of pitch variation in languages more generally, then presents a standard analysis of lexical tones in Mandarin Chinese. It goes on to describe conventional approaches to the teaching of Mandarin tones, as well as key observations and findings in this area, including a summary of the pros and cons of the various approaches.

Pitch Variation in Languages

It is a common feature of all human languages that they make use of pitch variation in their spoken forms. In this regard, broadly speaking the languages of the world can be

divided into tonal and non-tonal languages. In non-tonal languages such as English and other Indo-European languages, pitch variation has two main functions (Beckman, 1984). It is one component of the linguistic feature of stress at both the word and sentence level; this includes both the word-internal stress in polysyllabic words and the stress on particular words or phrases within a clause or sentence. Another important function of pitch variation is to encode pragmatic meaning – here it is known as intonation. Intonation provides information about the speaker’s attitudes and emotions, the function of the utterance (e.g., statement vs question), the relationship between interlocutors (Juhász & Bartos, 2024) etc. Importantly, intonation does not change the definition or semantic designation of a word; rather, it changes its pragmatic function and interpretation in context.

By contrast, in tonal languages such as Mandarin Chinese, pitch variation exists on three levels. As with non-tonal languages, pitch variation may be a component of stress and intonation applied at the phrase, clause or sentence level (Zhang, 2013). This type of pitch variation in tonal languages has similar functions to those described above for non-tonal languages; however, primarily owing to their interaction with lexical tone (see below), in practice there are some constraints on the manifestation of stress and intonation in tonal languages, which will not be discussed in this paper. The third function of pitch variation is the one which is most characteristic of tonal languages, namely lexical tone (Yip, 2006). This is applied at the level of the word or morpheme (which in tonal languages often means at the level of the syllable), and is quite different from the function of pitch variation in non-tonal languages in that it encodes semantic meaning; hence, pronouncing a given syllable with a different pitch contour will produce a different morpheme (Chao, 1965; Gussenhoven, 2004; Yip, 2006). It is lexical tone, or pitch variation at the morpheme level (which in Mandarin Chinese equates to the syllable level), that this paper is concerned with.

Lexical Tone in Mandarin Chinese

Mandarin Chinese is generally analysed as having four tones, in addition to a neutral tone (Chao, 1965; Li & Thompson, 1981; Norman, 1988). The pitch contours of the tones were notably studied by one of the pioneers of modern Chinese linguistics and pedagogy, Yuen Ren Chao, who proposed an approach using a 5-scale system to illustrate the pitch of the four tones in Mandarin Chinese, as shown in Table 1 (Chao, 1965). Later research has proposed alternative or modified representations, either for scholarly or pedagogical purposes. These include characterisation through opposing features, such as [$\pm$ high], [$\pm$ rise], [$\pm$ fall] (Duanmu, 2000) and “tone targets,” whereby each tone is conceptualised as having an ideal “target pitch,” a theoretical approximation whose precise realisation in actual speech contexts depends on a range of factors such as ambient pitch, vocal cord control, linking effects etc. (Duanmu, 2000). In view of significant regional and dialectal variations across Chinese-speaking regions (see Lin & Wang, 1992), we focus on the

standard *Putonghua* ('common tongue') or spoken Modern Standard Chinese, as defined and stipulated by the National Language Commission of China.

Table 1

Five-Scale Illustration of Pitch Contours of Mandarin Tones (following Chao, 1965)

Tones	Pinyin Diacritic	Pitch	Scale	Example
1	ˉ	high level	55	mā 妈 'mother'
2	ˊ	middle rising	35	má 麻 'hemp'
3	ˇ	low falling rising	214	mǎ 马 'horse'
4	ˋ	high falling	51	mà 骂 'to curse'

Teaching Mandarin Chinese tones to CFL/CSL learners

The following section presents an overview of the existing research into second language acquisition of Mandarin Chinese tones; it starts by discussing common barriers and challenges for L2 learners before summarising relevant pedagogical approaches and theories of learning.

Commonly Highlighted Challenges

The literature has highlighted a number of common challenges and areas of confusion for L2 learners in improving their perception and production of Mandarin Chinese lexical tones. Studies have repeatedly found that the single most influential factor in determining successful tone acquisition is the tonality of a learner's L1, with L1 speakers of other tonal language varieties such as Cantonese or Vietnamese learning to perceive and produce Mandarin Chinese tones more readily than L1 speakers of non-tonal language varieties (e.g., Hao, 2012). Another key factor has been shown to be prior musical training (Meng, 2021), although the duration and intensity of training required to make a significant difference is yet to be determined. Overall, T3 is usually identified as the most difficult tone to acquire (e.g., Guo & Tao, 2008) and is commonly confused with T2 in both perception and production tests (H. Zhang, 2018). The other Mandarin Chinese tone pair that is frequently confused is T1-T4 (Yang, 2016). A third common challenge highlighted in the literature is mastering the rules of tone sandhi (Zhang, 2019), the most common of which relate to T3: one is the realisation of T3 as T2 when immediately followed by another T3; another is the truncation of T3 in non-phrase-final position to the so-called half form (Yang & Yang, 2019). In fact, this latter is so common that T3 occurs with statistically higher frequency in its half form than in its full form, leading to some debate within the literature as to which form should be treated as the base form and hence taught first to L2 learners (Zajdler, 2021). Apart from specific difficulties learners have in tone production, research in second language acquisition has identified pitch acuity as a predictor of L2 tonal learning success. Zhou and Veríssimo (2026) further propose a perception–production asymmetry in L2 tonal acquisition, suggesting that learners' ability to perceive tonal contrasts may not necessarily correspond to their ability to produce them accurately.

Conventional Pedagogical Approaches

A few pedagogical strategies have emerged over the years as quite popular in helping L2 learners deal with these challenges. Probably the most common is the audio-lingual method, whereby the teacher models the tone pronunciation for the students to repeat. For example, Meng (2021) found that practising words in varied contexts was more effective for correcting tone errors than practising tones in isolation or original contexts. Another popular strategy for both teachers and learners is the use of physical movement – either head movement or, more commonly, hand gestures – to indicate the pitch contour of each tone. In the gesture approach, no significant advantage of gesture production over gesture observation has been established (Baills et al., 2019). Visual methods such as musical notation to represent pitch contours are sometimes used, as is colour coding with four different colours for each of the tones T1-T4 and a fifth for the neutral tone (Godfroid et al., 2017); this is frequently visualised in videos and realised in apps. In addition, there are a number of computer-assisted language learning strategies, including the option for students to record and play back their own voice; this is sometimes combined with a visual representation of the student's own pitch modulation in comparison with the 'standard' or 'target' pitch contour shown on an open-source platform such as Praat (e.g., Chun et al., 2013; Chun et al., 2015). Furthermore, the potential of AI in second language learning has been noted by recent research, particularly in English. While AI has also demonstrated its capacity to facilitate Mandarin learning, including increasing vocabulary range and reducing affective participation, its contribution to tonal accuracy remains limited (Wang, 2026). In summary, of the approaches to teaching Mandarin Chinese tones outlined here, the visual and gestural approaches (Shi, 2018) and computer-assisted approaches (Li et al., 2021) have been demonstrated most frequently and consistently to have beneficial effects on learning outcomes.

Interestingly, these visual and kinaesthetic forms work to build connections between an unfamiliar and strange object (in this case, lexical tones) to something that students are familiar with and can readily use and relate to (readable line graphs and body movements). As such, they are based on a similar rationale to the proposed 'YES' approach, which is designed to bridge the gap between learners' innate L1 phonological knowledge and the target sound system by deliberately highlighting similar pitch contours in the learners' L1 and L2, hence making these salient. This will be discussed in more detail in the next section, with reference to cross-linguistic transfer theory.

Cross-Linguistic Transfer and 'Noticing'

The conventional pedagogical approaches described above are predicated on the assumption that tones are a 'strange' or 'alien' concept for speakers of non-tonal languages. However, an alternative approach founded on a well-established notion is emerging – to make the strange familiar. This principle, widely accepted in many branches of second or foreign language pedagogy outside lexical tone teaching, can be

considered an example of conscious (as opposed to subconscious) cross-linguistic transfer (CLT) (Kellerman, 1979). Learners' first language provides a perceptual and conceptual reference point against which features of the target language are 'noticed.' In this case, 'noticing,' an essential element in second language learning (Schmidt, 1990), mediates whether cross-linguistic transfer is facilitated, inhibited or avoided. Research has shown that relating features of the L2 back to features of the L1 can help learners to grasp linguistic concepts more easily and acquire them more readily; this is known as positive interlanguage transfer (Ringbom, 2006). The effectiveness of positive interlanguage transfer in enhancing L1-L2 phonological awareness, including the contextual application of lexical tone for learners of Mandarin Chinese, has been the topic of a few experimental projects and corroborated by a meta-analysis (Yang et al., 2017).

If the focus is on tonality, then Chinese and English are normally considered two languages without overlapping features, since one is classified as a tonal language and the other as a non-tonal language. If, however, we shift our focus to pitch variation, then commonalities can be observed across both languages; moreover, it is feasible to draw on these in pedagogy. That is, a bridge can be built for learners to 'notice' the similarities and allow them to use their L1 to understand Chinese tones from the perspective of pitch control.

One study, which examined whether English can be reframed in terms of tones for L1 Chinese-speaking learners to improve their pitch processing in rising and falling English intonations (analogous to Mandarin T2 and T4), found that explicitly mapping connections between tonality and atonality helped improve the sensitivity in both Mandarin and English pitch perception (Ortega-Llebaria & Wu, 2021). Similarly, Juhász and Bartos (2024) compared rising and falling pitch contours in Hungarian with their closest Mandarin Chinese equivalents T2 and T4 in a study exploring the possibility of utilising intonation patterns in Chinese tones teaching; they found that although some differences were observed between the pitch contours in the two languages, there were sufficient similarities for these to usefully inform pedagogy. In research investigating what degree of plasticity adults exhibit in learning new contrasts in L2, in this case Chinese Mandarin tones, results showed that all adult L2 learners, even absolute beginners, were able to acquire perceptual discrimination of tonemes at the lexical level (Wang et al., 2020). It has also been empirically demonstrated that there is a link between perception and production in Mandarin Chinese tone learning, suggesting that the former can benefit the latter (Leung et al., 2025), even if the linguistic transfer between perception and production may be subject to time lags and limitations (Nagle, 2018). This indicates that with sufficient explicit instruction and guidance, learners can enhance their 'noticing' skills and further promote their tone perception and production. Tapping into Chinese L2 learners' implicit knowledge of features of their own L1 to explain similar-but-different features of their chosen L2 thus seems to be an opportunity for applying the principles of deliberate cross-linguistic transfer and 'noticing' in language pedagogy.

Informed by the ‘noticing’ framework, this study is designed to investigate the possibility and practicality of facilitating the construction of salient connections or pathways of conscious linguistic transfer between learners’ existing knowledge and skills regarding pitch modulation (English intonation patterns, in this study) and Mandarin Chinese tonality, with the intention of easing fear and reluctance and thus supporting the second language acquisition of lexical tone.

Research Questions

Based on previous research findings and the authors’ observations of certain similarities between Mandarin Chinese tonal pitch contours and the intonation contours commonly found in English monosyllabic utterances, the following research questions are proposed:

RQ1: Does the use of L1 intonation patterns as a model for L2 lexical tone contours have a beneficial effect on learners’ attitudes towards and confidence in learning lexical tones?

RQ2: Does the use of L1 intonation patterns as a model for L2 lexical tone contours enable learners to improve their perception and production of Mandarin Chinese tones?

Methodology

Study Design

The current pilot study was conducted at UQ in Semester 1 of 2025 in an introductory Chinese language course, CHIN1100 - Spoken Chinese IA, using one tutorial group as the control group and the other as the experimental group. The study ran for eight weeks and was designed in five stages: (1) information session and pre-survey; (2) pre-test (perception and production); (3) training sessions and exercises; (4) post-test (perception and production); and (5) post-survey. The novel training was dubbed the ‘YES’ approach. Each 10- to 15-minute training module was incorporated into the lesson plan for the relevant semester week, with the main tutor responsible for running the tutorial and the guest instructor assisting with targeted tones training. Participants were able to access the training materials after class via the course Blackboard site. The structure of training modules and the design of in-class activities are described in the section on Training; the tests and surveys are described in the section on Data collection instruments.

Participants

After obtaining ethics approval from the University of Queensland (no. 2025/HE000159), participants were recruited from the first-year course during a short in-class information session; the Consent Form and Participant Information Sheet were distributed via Blackboard announcement, email and QR code. There were 48 students enrolled in the course in Semester 1, 2025, of whom 21 signed the Participant Consent Form. Data from the Student Management System revealed that students in the course were enrolled in different degree programs and different majors across the university, with some students doing CHIN1100 as a foundational course in their Chinese language major and others completing it as an elective.

Participants were divided into a control group and an experimental group, such that only the students in the experimental group received additional tones training using the new method, while both groups received tones training using the more conventional audio-lingual and gestural methods. For practical considerations, one of the two course tutorial groups was chosen at random as the control group and the other as the experimental group; this meant that the additional tones training could be provided during scheduled tutorial times. In total, 21 students (eleven male and ten female, aged between 18 and 30) completed the pre-survey; 15 and 13 completed perception and production pre-tests, respectively; and six completed both post-tests (perception and production) as well as the post-survey. Of the six students who completed all tests and both surveys, three were in the control group and three in the experimental group, with two males and one female in each.

Training

The experimental group received three short 10- to 15-minute training sessions, scheduled in semester Weeks 8, 9 and 11. The in-class information session was given in Week 6, and the pre-survey and pre-tests were available for participants to complete until the end of Week 7. The post-survey and post-tests closed at the end of W14, approximately three weeks after completion of the training.

Training Session 1

The aim of Session 1 was to surface participants' implicit knowledge and control of pitch variation in day-to-day communication in English (the L1 of most learners) and to explain the linguistic differences between intonation and lexical tone. The training module was run as a mini-workshop, with students guided to complete the following activities in small groups of three to five: (1) uttering common single-word utterances such as 'hello' and 'yes' with a range of intonation contours; (2) drawing these contours; (3) providing a gloss of the emotional or pragmatic meanings associated with the different intonation patterns; and (4) reflecting on the significance of pitch modulation in Mandarin Chinese and English.

Figure 1

Selected activities from Training Session 1

Examples of "hello" with pitch contours

- ▶ Can you express these different meanings?
- ▶ Can you think of other ways to say "hello"?

Examples of "yes" with pitch contours

- Yes. (I'm not really convinced)
- Yes! (I know, and isn't it scandalous!)
- Yes. (I'm coming – for the umpteenth time.)
- Yes! (Of course! Whaddya think?)
- Yes. (Don't remind me. Please.)

Training Session 2

The aim of Session 2 was to make students aware that they use intonation patterns that mimic the pitch contours of Mandarin tones in their everyday English conversations. Training activities were structured around the word ‘yes,’ as an example of a very common monosyllabic English word: (1) the instructor modelled the word ‘yes’ in different interactional contexts with four common English intonation contours that are very similar to Mandarin Chinese tones while eliciting from the students the pragmatic nuances conveyed; and (2) students practised uttering the word ‘yes’ in contextually appropriate communication with the four lexical tones of Mandarin Chinese superimposed (Table 2).

Table 2

Mapping Mandarin Chinese Tones onto English Meanings

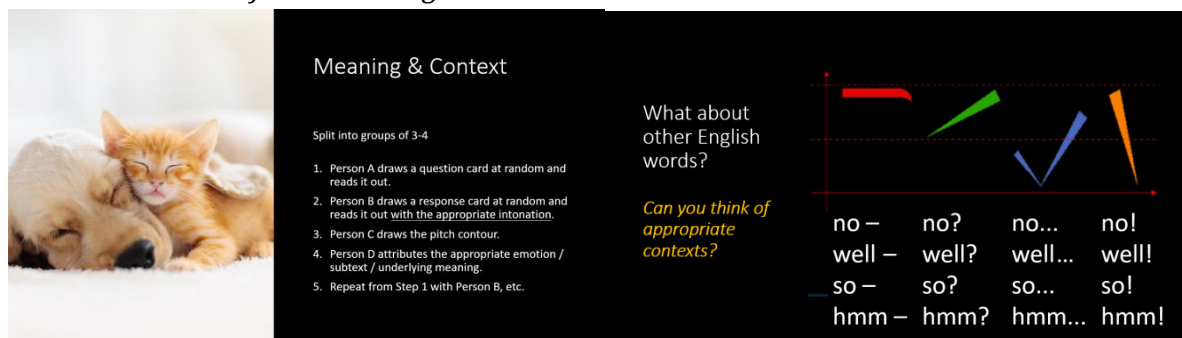
Intonation Pattern	Word	Context	Interpretation
Tone 1	‘yēs’	A: <i>(knocks on B’s office door)</i> B: ‘Yēs.’	yēs = ‘I’m busy, don’t interrupt me.’
Tone 2	‘yés’	A: <i>(knocks on B’s office door)</i> B: ‘Yés.’	yés = ‘Please come in.’
Tone 3	‘yěs’	A: ‘Mum, you know how you said I could borrow the car on the weekend?’ B: ‘Yěs.’	yěs = indicates uncertainty and apprehension as to what could come next
Tone 4	‘yès’	A: <i>(scrolling through a catalogue of items)</i> ‘No, no, no ... Yès!’	yès = ‘I’ve finally found what I was looking for!’

Training Session 3

The aim of Session 3 was to consolidate learners’ association of Mandarin Chinese lexical pitch contours with particular situations and pragmatic meanings in English, so that they were able to recall and reproduce them as required in L1 communicative contexts and gradually learn to apply them to L2 utterances. Group activities were as follows: (1) reminding participants of the four Mandarin-similar intonation contours with English ‘yes’ (Table 2); (2) experimenting with similar intonation patterns on other common monosyllabic words such as ‘no,’ ‘so’ and ‘well’; and (3) matching the different contours to contexts in which they might occur (Figure 2).

Figure 2

Selected activities from Training Session 3



Meaning & Context

Split into groups of 3-4

1. Person A draws a question card at random and reads it out.
2. Person B draws a response card at random and reads it out with the appropriate intonation.
3. Person C draws the pitch contour.
4. Person D attributes the appropriate emotion / subtext / underlying meaning.
5. Repeat from Step 1 with Person B, etc.

What about other English words?

Can you think of appropriate contexts?

no –	no?	no...	no!
well –	well?	well...	well!
so –	so?	so...	so!
hmm –	hmm?	hmm...	hmm!

Data Collection Instruments

The pre-survey collected data on the participants' age, gender, self-rated language proficiency and previous Chinese language learning experience. It included three Likert-scale rated questions on attitudes towards learning Mandarin Chinese tones, with five response options ranging from 'strongly agree' to 'strongly disagree':

- It will be useful for me to learn Mandarin tones.
- It will be easy for me to learn Mandarin tones.
- It will be enjoyable for me to learn Mandarin tones.

The post-survey asked similar Likert-scale rated questions to the pre-survey, to ascertain if there had been any change in the students' attitudes towards learning Mandarin tones. Participants in the control group were asked to rate the following statements:

- It is USEFUL for me to learn Mandarin tones.
- I find it EASY to learn Mandarin tones.
- I find it ENJOYABLE to learn Mandarin tones.
- I am SATISFIED with how well I have mastered Mandarin tones.

Those in the experimental group were asked to think back on the additional training sessions when rating the following slightly different statements:

- The training has been USEFUL in helping me to learn Mandarin tones.
- The training has made it EASY for me to learn Mandarin tones.
- The training has made it ENJOYABLE for me to learn Mandarin tones.
- I am SATISFIED with how well I have mastered Mandarin tones.

The perception pre-test and post-test were designed as tests on the course *Blackboard* site and consisted of two parts. Part A, which tested students' perception of Mandarin Chinese monosyllables, contained ten multiple-choice questions, with each prompt drawn at random from a pool of eight syllables with the same phonetic value but four different tones (e.g., *mā, má, mǎ, mà*) read by two different L1 Mandarin speakers, one male and one female. The monosyllables were deliberately chosen to include only consonants and vowels commonly found in English (the L1 of the majority of participants) to avoid potential confusion caused by unfamiliar phonemes.

Part B, which tested participants' perception of Mandarin Chinese disyllabic compounds, included ten pools of two questions each: one disyllabic compound read by the same two L1 Mandarin speakers who recorded the questions in Part A. The syllables were again composed of phonemes that could be easily pronounced and aurally distinguished by the predominantly L1 English participants. Syllables were deliberately combined in such a way that they did not produce genuine Mandarin Chinese words, so as to avoid the possible influence of prior vocabulary knowledge. The compounds were created with approximately equal numbers of each of the four tones (five x T1, four x T2, six x T3, five x T4). Since previous literature on L2 learners has highlighted frequent confusion between T1 and T4 and between T2 and T3, participants were asked to select the correct representation from four different options, with each correct tone alternating with the tone it is most commonly confused with. For example, the disyllabic combination for Question 11 was *mǎ-dīng*, or T3-T1. Based on the assumption that students were most likely to mishear T3 as T2 and T1 as T4, the four answer options provided were T3-T1, T3-T4, T2-T1 and T2-T4, or *mǎ-dīng*, *mǎ-dìng*, *má-dīng* and *má-dìng*.

Like the perception tests, the production pre-test and post-test were set up in *Blackboard* and consisted of two parts. Participants used *Vocaroo*² (<https://vocaroo.com>) to record their own voice and pasted the link to the recording into the answer box. Part A tested students' production of Mandarin Chinese monosyllables with ten multiple-choice questions, each one drawn at random from a pool of four written pinyin prompts with the same syllabic spelling but different diacritic tone markings. As with the perception tests, Mandarin syllables composed of familiar, English-friendly phonemes were chosen.

Part B, which tested participants' production of disyllabic combinations, included ten fixed prompts, being the same disyllabic combinations used in the perception test. As the purpose was specifically to test students' production of tones rather than their recall of vocabulary or their pronunciation of new phonemes, syllables were carefully selected and combined to avoid both genuine Mandarin Chinese words and sounds that were considered difficult for L2 learners to articulate.

Data Analysis

Responses to the perception pre- and post-tests were collected in *Blackboard* and collated and analysed in *Excel*. The students' recorded responses to the production pre- and post-tests were independently evaluated by L1 Mandarin-speaking Authors 2 and 3, with the emphasis on the adequacy of tone production for communication purposes. Each individual tone was rated on a scale of one to three (in increasing acceptability of performance) across three separate criteria:

² Vocaroo is a free online voice recording tool without sign-up, enabling recordings to be temporarily stored online for later review and download. This was chosen for the production test for its efficiency and ease of use.

- Accuracy – the extent to which the student is able to correctly produce the pitch contour of each Mandarin tone at the monosyllabic level, in alignment with L1 speaker norms;
- Tone differentiation – the ability of the student to consistently distinguish among the four tones by producing sufficiently distinct pitch patterns, so that minimal pairs differing only in tone are perceptually separable; and
- Understandability – the degree to which the student's tone production can be readily identified and interpreted by listeners as the intended tone, regardless of minor deviations from L1-like accuracy.

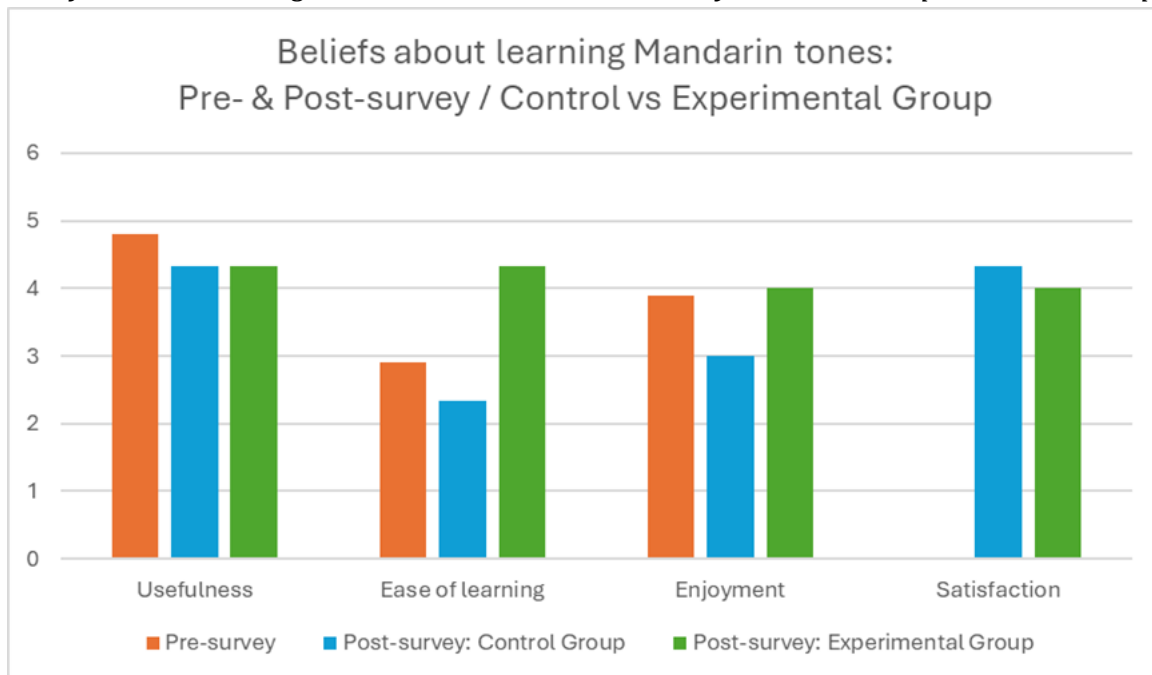
This score was averaged across the two assessors and (for the disyllabic utterances) across the two syllables, resulting in a maximum of nine points for each of the twenty recorded responses for each participant. Where tone issues were observed in individual syllables within the disyllabic utterances, these were noted separately.

Results

Surveys

Of the 21 students who completed the pre-survey, all spoke a non-tonal language as their L1, with English the L1 of the majority (17) of participants and the others speaking Indonesian, Malay and Norwegian. About half the participants spoke another language in addition to English and Mandarin Chinese and approximately one third had prior (formal or informal) Mandarin learning experience, mostly at secondary school.

Responses to the Likert-scale rated attitudinal questions in the 21 pre-surveys showed moderate levels of optimism and confidence before commencement of the research project. In response to the three prompts "It will be ___ for me to learn Mandarin tones," the descriptor "useful" was rated 4.8 (strongly agree) on average, "enjoyable" 3.9 (agree) and "easy" 2.9 (neither agree nor disagree). In other words, participants reported quite a positive attitude towards learning Mandarin Chinese tones; believed that tones were very important; but were diffident about their ability to acquire tones. The attitudinal differences between students in the control and experimental groups following tones training are illustrated in Figure 3.

Figure 3*Beliefs About Learning Mandarin Tones in Post-Survey: Control vs Experimental Group*

After several weeks of training using only conventional audio-lingual and gestural methods, the three students in the control group who completed the post-survey showed slightly negative to moderately positive attitudes, averaging 4.3 (agree) for *usefulness*, 2.3 (disagree) for *ease of learning*, 3.0 (neither agree nor disagree) for *enjoyment* and 4.3 (agree) for *satisfaction*. By contrast, the three students in the experimental group, who received training with the new method in addition to (slightly reduced) training via conventional methods, developed more positive attitudes and greater confidence over the same period, averaging 4.3 (agree) for *usefulness*, 4.3 (agree) for *ease of learning*, 4.0 (agree) for *enjoyment* and 4.0 (agree) for *satisfaction*. In other words, although both groups were satisfied with their overall progress, the experimental group found learning tones both **more easy** than and **just as enjoyable** as anticipated, while the control group found tones **less easy** and **less enjoyable** than anticipated.

Participants' responses to the open-ended questions acknowledged a number of challenges and barriers to mastering Mandarin tones, including recognising individual tones; differentiating tones in multisyllable words and sentences; and confusing tone pairs (T1-T4 and T2-T3). Their suggestions included clearer initial guidance to foster tone awareness; personalised, one-to-one feedback, particularly to prevent fossilisation of errors; authentic learning materials; and distinct tone learning strategies for more advanced language acquisition levels.

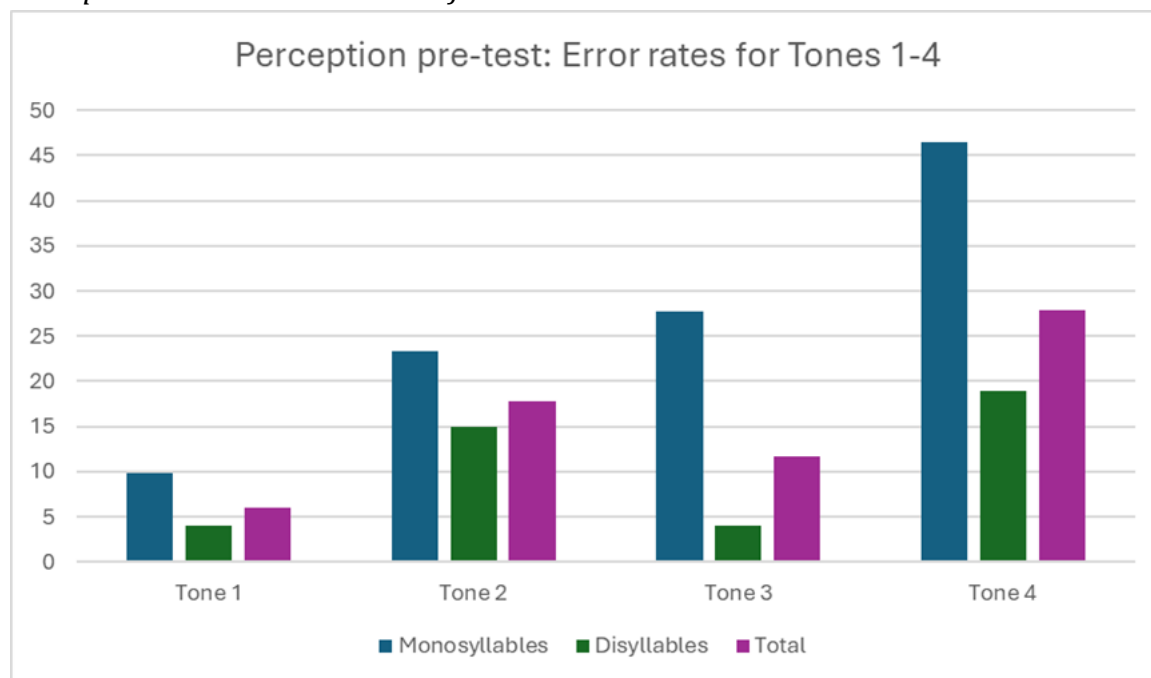
Perception Tests

Figure 4 shows the results for the perception pre-test, completed by fifteen participants. Results were very similar for participants in the experimental and control groups, both

in overall patterns and in individual error rates. Somewhat surprisingly, students scored higher on the correct identification of tones in disyllabic compounds than in monosyllables. Overall, students were most adept at recognising T1 (error rate 6%), followed by T3 (11%), T2 (18%) and T4 (28%). Closer analysis revealed that students often mistakenly identified T2 as T3 and vice versa (cf. H. Zhang, 2018), and T1 as T4 and vice versa (cf. Yang, 2016), both in monosyllables and in disyllables.

Figure 4

Perception Pre-Test: Error Rates for Tones 1-4



A comparison of the perception post-tests with the pre-tests submitted by the same six participants reveals that learners in both the control and the experimental groups benefited from the training received (see Figures 5 and 6), with males performing better overall than females. Once again, tone perception was better on Part 2 of the test (disyllables) than on Part 1 (monosyllables) for both groups (see Figure 6). Overall, T4 remained the most challenging tone to identify correctly (see Figure 5), followed by T2 and T3, with T1 showing the lowest rates of misidentification, despite the somewhat anomalous jump in T1 perception errors for the control group in the post-test. The control group achieved a greater overall reduction in perception errors from 27% to 17%, compared to a decrease in the experimental group from 13% to 11% (Figure 5). Both groups improved their test scores by a similar amount (Figure 6).

Figure 5

Error Rates in Perception of Tones 1-4: Control vs Experimental Group

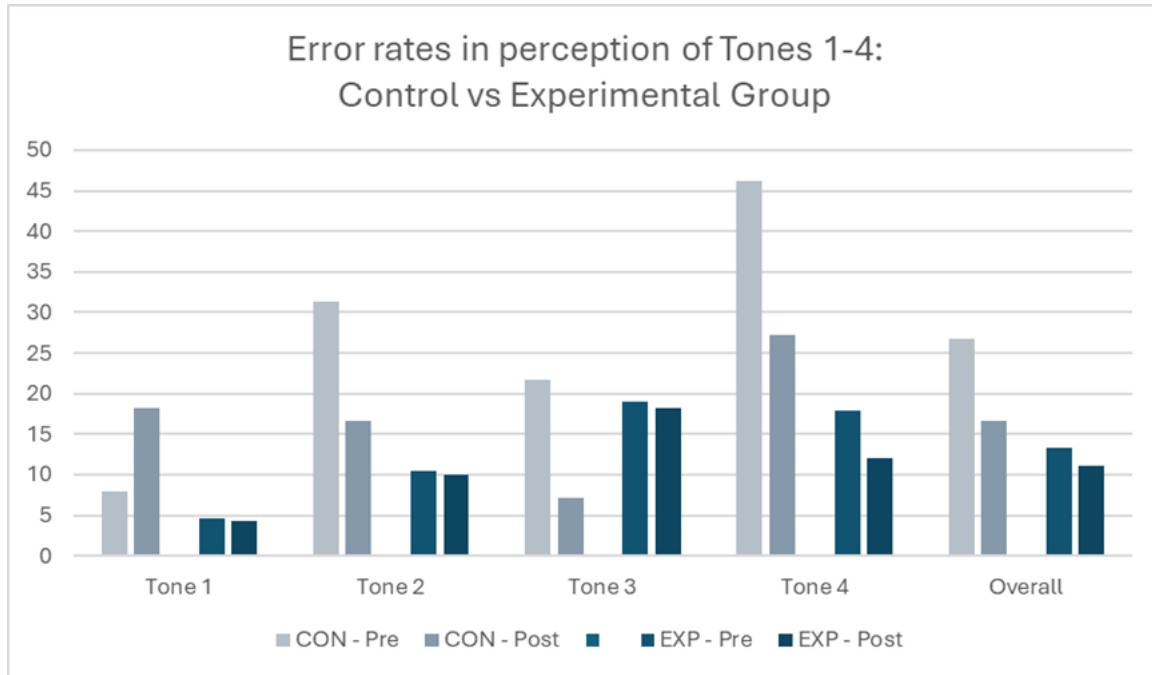
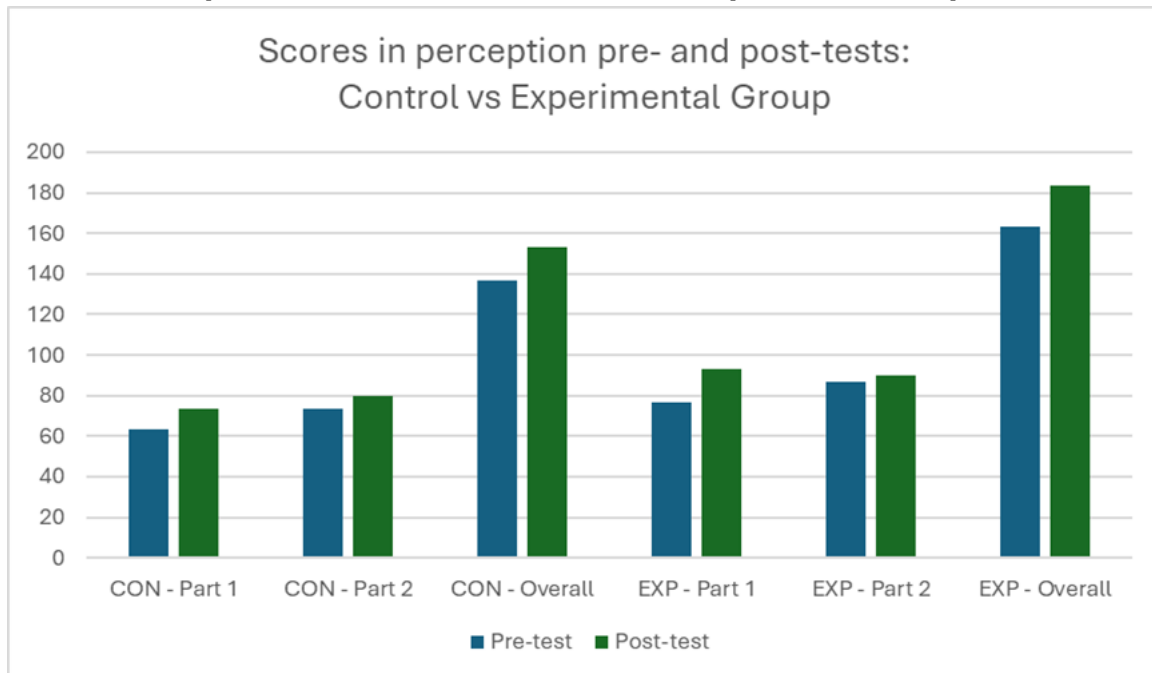


Figure 6

Scores in Perception Pre- and Post-Tests: Control vs Experimental Group



Production Tests

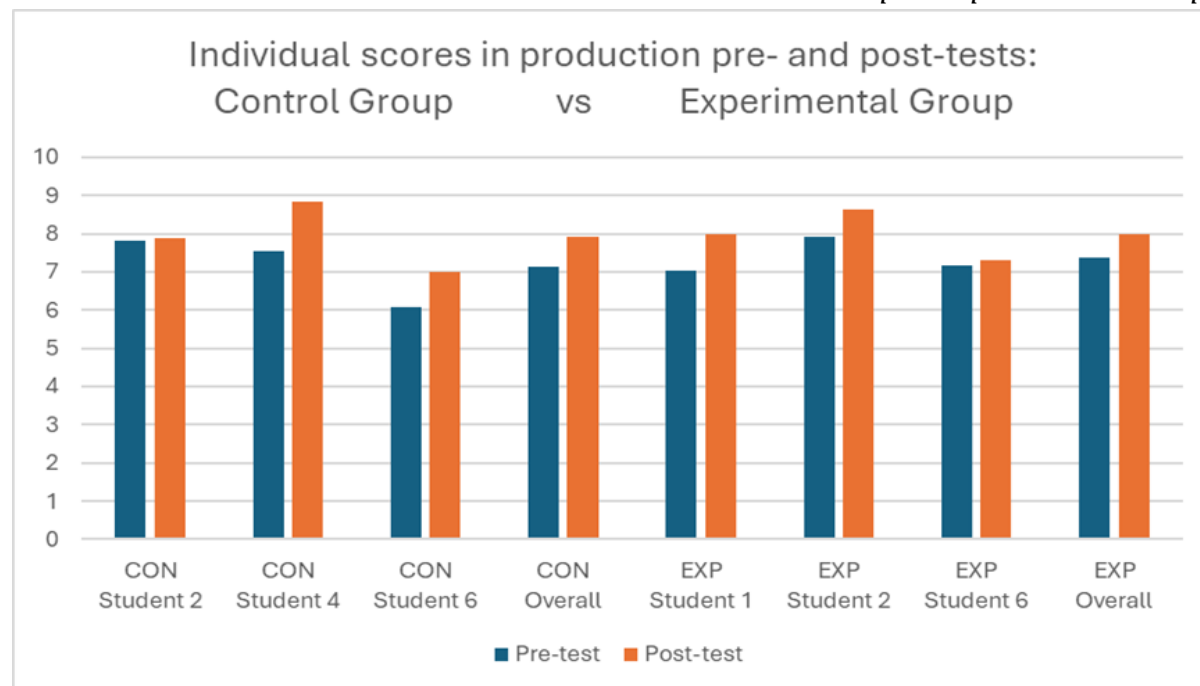
The production pre-test was completed by thirteen participants. Due to technical glitches a few recordings were lost; statistics were calculated after excluding the missing data. The Pearson correlation coefficient for inter-rater agreement was 0.978, indicating high agreement. Detailed qualitative analysis of the pre-tests revealed, as predicted by

previous research, firstly that there was common confusion between tone pairs T1-T4 and T2-T3 (cf. Yang, 2016; H. Zhang, 2018) and secondly that production of T3 presented the greatest challenge (cf. Guo & Tao, 2008). The most common production error was to mispronounce T3 as T2; this mirrors the confusion between T2 and T3 observed in the perception tests. Other common issues with T3 arose from unstable or insufficiently controlled pitch contours (e.g., ambiguous pitch turn; more than one pitch turn; rising too high at the end). The second most challenging tone for production was T4, with some participants pronouncing T4 as T1 or abbreviated T1 (similar to Mandarin neutral tone) or occasionally as T2. The second easiest tone for production was T2, which was generally pronounced quite clearly, although it was sometimes hard to distinguish from T3 in students' recordings. Participants' pronunciation of T1 was deemed by the assessors to be most adequate; however, in some cases other tones were realised as T1. Overall, it is clear from the production pre-tests that while some students had already attained a reasonably adequate level of tone production, others had not yet developed a clear, systematic cognition of Mandarin tones. This was evident in the fact that for some students, tone production was somewhat 'hit and miss': all four tones could be pronounced correctly or incorrectly seemingly at random.

A comparison of the production post-tests with the pre-tests submitted by the same six participants reveals that students in both the control and the experimental groups improved during the period in question (see Figure 7). Despite individual variation amongst participants, both groups achieved very similar averages in both pre- and post-tests. Once again, male participants outperformed females.

Figure 7

Individual Scores in Production Pre- and Post-Tests: Control Group vs Experimental Group



Considering only the six final participants, a detailed analysis of tone production in the pre- and post-tests revealed that tone pair confusions (T1-T4 and T2-T3) persisted, as did the relative difficulty of the four tones, with T1 the easiest and T3 the hardest to acquire. T3 improved most markedly in both groups but remained the most significant challenge. For example, T3 sometimes started with a too-obvious dropping pitch over an unnecessarily large range or ended on an overly high pitch and/or a fluctuating pitch contour. The authors surmise that some students in the experimental group may have overgeneralised their training techniques in exaggerating the T3 pitch turn or entirely copying English intonation patterns onto Chinese syllables.

Discussion

The current research set out to ascertain whether the use of common English intonation patterns on monosyllabic English utterances as a model for Mandarin Chinese tones might show beneficial effects either in the efficacy of tones teaching or in students' attitudes to learning. The findings presented here indicate that while there was no measurable advantage in this small-scale pilot study in improving the perception or production of tones vis-à-vis conventional methods, participants in the experimental group did develop more positive attitudes towards learning than those in the control group, reporting greater enjoyment and ease of acquisition. Of course, training methods that promote effective learning are very important; however, learner attitudes are known to be similarly important. Learner satisfaction and enjoyment have been associated with greater persistence and engagement in learning, both of which are important for learner retention (Guo, 2021; Schreiner & Nelson, 2013).

In relation to RQ1 (effects on students' attitudes towards and confidence in learning lexical tones), the findings suggest that the novel approach piloted in the current research project does indeed help alleviate learners' negative preconceptions and reinforce positive attitudes. In particular, the L1 speakers of non-tonal languages who attended the 'YES' training sessions were shown that they actually already have the necessary skills (pitch variation, control of pitch production, discernment of pitch contours) to learn a tonal language; consequently, they developed more confidence and enjoyed the tones training more. This aligns with cross-linguistic transfer theory: drawing learners' attention to similarities between L1 and L2 facilitates 'noticing,' reduces cognitive load, and promotes positive transfer (Ringbom, 2006). By contrast, pedagogies that 'foreignise' the target language and emphasise its 'otherness' risk increasing learner anxiety and inefficiency.

These findings have clear implications for CFL teaching. The study, although small in scale, produced results that were strongly suggestive of improved learner confidence and enjoyment with the new training method. CFL/CSL instructors can thus consider incorporating the 'YES' method into tones teaching, not to replace but rather to augment existing tones teaching methods. Since links between learner attitudes and continued

study motivation are well established, the new 'YES' method could well have some positive flow-on effects on long-term student retention in tertiary level Mandarin Chinese courses. The findings also suggest that CFL/CSL pedagogy should not be restricted to existing methods of tones training but should continue to develop and experiment with novel approaches.

The training sessions described in this research project were designed to help students learn Chinese tones by presenting them as a familiar concept in a slightly different guise, enabling learners to draw on the intonation patterns of their non-tonal L1 when acquiring a tonal language. In the pre-survey, participants pointed out some barriers and challenges they had encountered in Chinese tones learning in the first six weeks of their Mandarin Chinese university course, indicating that they were aware of differences between English and Mandarin. The training sessions tapped into this kind of 'noticing' to help them focus more on similarities than on differences; this allowed participants to consciously employ cross-linguistic transfer, hence building their confidence and gaining more motivation in the learning process, eventually achieving a "similarity-easiness" learning outcome (J. Zhang, 2018). Recent national evidence indicates substantial attrition in language study across secondary schooling (AFMLTA, 2024; Jiang et al., 2024), while a 2026 parliamentary inquiry identified discontinuity in Chinese learning pathways as a particular barrier to sustained participation and proficiency (House of Representatives Standing Committee on Education, 2026).

The findings revealed a scale of decreasing ease of acquisition in tone production: T1, T2, T4, T3. Of the four tones, T1 was not only the easiest to acquire, it also seemed at times to be the default realisation for other tones. Since it is well known that T1 is usually the most accessible for CFL/CSL learners, this could be a case of overgeneralisation or oversimplification. The fact that T4 was occasionally realised as T2 could potentially be a result of mismatched transferral of rising and falling pitches in the learners' L1 (mostly English). The unstable and insufficiently controlled pitch contours observed in the production of T3 were not particularly surprising, in view of previous research highlighting this as the most difficult tone to pronounce adequately. Moreover, the mapping of L1 (English) 'apprehensive, dubious or uneasy' intonation onto Mandarin syllables by the students in the study could have exacerbated this manifestation.

In relation to RQ2 (effectiveness in helping students acquire lexical tones), comparison of the two groups' performance on pre- and post-tests indicated that both conventional and new methods of tones teaching benefited students' perception and production of Mandarin tones. Although the control group achieved a greater overall reduction in perception errors than the experimental group (Figure 5), this could be related to the fact that the control group started at a higher error rate (27% cf. 13% for the experimental group), leaving more capacity for them to improve their error rates.

There are several limitations of the current study. With only six participants completing all phases of the research project, the results of this study cannot claim to be representative of the larger student cohort. If treated with the caution generally accorded to the interpretation of pilot studies, the findings can nevertheless provide some useful pointers for further research and pedagogical practice. The brief training offered to participants – only 30-40 minutes in total – is likewise a limitation. A longer training duration could presumably produce more conclusive findings and perhaps reveal greater discrepancies between the control and the experimental groups. Another improvement to consider is the inclusion of delayed post-tests to measure learning retention and attitudinal change weeks or even months after completion of the training. Finally, it should be acknowledged that the ‘YES’ approach in its current form is heavily tailored to L1 English speakers. Although adaptation for other L1 learner cohorts would be relatively straightforward for some languages (e.g., German), significant modifications may be needed for others (e.g., Hindi). Thus, there is still much work to do in adapting the new approach and creating more teaching and learning materials to make it more broadly applicable.

In addition to conducting research on a larger scale and/or for a longer duration, some future directions for tones training were identified by trialling the module in this pilot study. Firstly, there are pitch-height distinctions between Chinese tones and English intonation contours. It was observed that some participants in the experimental group overgeneralised the English intonation patterns presented in the ‘YES’ training modules and consequently produced in the post-test some non-standard Mandarin lexical tones that were closer to pitch contours common in spoken English. This points to the success of the training in providing an L1 model that learners could consciously apply in acquiring L2 tones. At the same time, it highlights a potential weakness of the ‘YES’ training, namely that when pitch contours in L1 and L2 are not exactly the same, there may be a tendency for learners to opt for a ‘next best’ approximation drawn from more familiar L1 linguistic resources. In future training sessions, to reduce the processing load, it might be advisable not merely to copy intonation patterns from English, but also to give learners opportunities to ‘notice’ pitch contour differences once they have acquired a basic level of competence. A second area of focus is suggested by the somewhat tenacious tone pair confusion (T1-T4, T2-T3) that persisted within the experimental group in both perception and production post-tests, despite relating these tone contours back to L1 models. Future training could include a stronger emphasis on practising and comparing confusing pairs to help students distinguish more reliably between them. Another aspect for future monitoring and investigation is the potential (unintended) transference from L1 onto L2 of learners’ associations between a given pitch contour and its L1 modal or pragmatic affect. Students who have been trained to consciously associate certain situations and pragmatic connotations with particular pitch contours in their L1 in order to recall and reproduce these pitch contours at will may require specific guidance to dissociate the pragmatic value from the tone contour in their L2.

Conclusion

This paper argues for a re-evaluation in CFL teaching of the perceived differences between Mandarin Chinese and non-tonal languages. The authors' experience shows that L1 speakers of Indo-European and other non-tonal languages bring a range of linguistic skills to the CFL classroom, knowledge which is often overlooked by CFL instructors but which can be tapped into and built upon. Appropriate application of the pedagogical principles of cross-linguistic transfer and 'noticing' can be beneficial in assisting learners to master the macro skills of speaking and listening in a tonal language such as Mandarin, as well as in developing positive attitudes towards and a curiosity for learning. Rather than directly pointing out differences and highlighting barriers to learning, it may be more effective to transform differences into connections, awkwardness into familiarity, and fear into curiosity. Despite acknowledged limitations of the project, including a small sample size and short training duration, the current research underscores the continued relevance of a key pedagogical principle in language teaching, namely highlighting the similarities between learners' first and second languages and acknowledging differences without magnifying them.

It is pleasing to observe in the literature on the second language acquisition of lexical tone an incipient willingness to acknowledge some transferable linguistic knowledge between non-tonal and tonal languages, as in a previous study comparing Hungarian intonation contours with Mandarin tones (Juhász & Bartos, 2024). The findings of the present pilot study suggest that the approach trialled here should be adapted for learners from a wider range of L1 backgrounds and that researchers and language educators should continue to investigate novel, complementary approaches for teaching Mandarin Chinese tones to L1 speakers of non-tonal languages. In so doing, it is worth bearing in mind that supporting a healthy attitude towards learning may be just as important in the long term as improving measurable learning outcomes. One of the guiding principles in second language teaching, that of making the 'strange familiar,' can claim as much relevance today as when it was first proposed nearly half a century ago.

ORCID

 <https://orcid.org/0000-0002-3232-8508>

 <https://orcid.org/0000-0002-0775-7132>

 <https://orcid.org/0009-0001-6975-9037>

Publisher's Note

The claims, arguments, and counter-arguments made in this article are exclusively those of the contributing authors. Hence, they do not necessarily represent the viewpoints of the authors' affiliated institutions, or EUROKD as the publisher, the editors and the reviewers of the article.

Acknowledgements

Not applicable.

Funding

This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

CRedit Authorship Contribution Statement

Angela Cook: Conceptualization, Methodology, Project Administration, Writing – Original Draft Preparation, Writing – Review & Editing

Yao Li: Methodology, Investigation, Formal Analysis, Writing – Original Draft Preparation, Writing – Review & Editing

Dongchen Yao: Methodology, Investigation, Formal Analysis, Writing – Review & Editing

Generative AI Use Disclosure Statement

Generative AI tools were used only for language editing and formatting assistance. The authors take full responsibility for the content of the manuscript.

Ethics Declarations

This study was conducted in accordance with the ethical principles of the World Medical Association (WMA) Declaration of Helsinki for research involving human participants. Ethical approval was obtained from The University of Queensland (Approval No.: 2025/HE000159).

Consent for Publication

Not applicable. The manuscript does not contain any identifiable individual information.

Competing Interests

The authors declare that they have no competing interests.

Data Availability

The datasets generated and/or analysed during the current study are available from the corresponding author on reasonable request.

References

- Abdullah, K., & Watson, P. (2021, September 21). *Foreign Service Institute Language Difficulty Rankings*. Atlas and Boots. <https://www.atlasandboots.com/travel-blog/foreign-service-institute-language-difficulty/>
- Australian Federation of Modern Language Teachers Associations. (2024). *The state of languages education in Australia: Provision and participation: Towards a national plan and strategy for languages education in Australia*.
- Baills, F., Suarez-Gonzalez, N., Gonzalez-Fuente, S., & Prieto, P. (2019). Observing and producing pitch gestures facilitates the learning of Mandarin Chinese tones and words. *Studies in Second Language Acquisition*, 41(1), 33-58. <https://doi.org/10.1017/s0272263118000074>

- Beckman, M. E. (1984). *Toward Phonetic Criteria for a Typology of Lexical Accent* (Publication Number 8407396) [Ph.D., Cornell University]. ProQuest Dissertations & Theses Global. New York. <https://www.proquest.com/dissertations-theses/toward-phonetic-criteria-typology-lexical-accent/docview/250816579/se-2?accountid=14723>
- Chao, Y. R. (1965). *A grammar of spoken Chinese*. University of California Press.
- Chun, D. M., Jiang, Y., & Natalia, A. (2013). Visualization of tone for learning Mandarin Chinese. *Pronunciation in Second Language Learning and Teaching Proceedings*, 4(1).
- 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(1), 86-114. <https://doi.org/10.1075/jslp.1.1.04chu>
- Duanmu, S. (2000). *The phonology of standard Chinese*. Oxford University Press.
- Eberhard, D. M., Simons, G. F., & Fennig, C. D. (2025). *What are the top 200 most spoken languages?* SIL International. <https://www.ethnologue.com/insights/ethnologue200/>
- 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. <https://doi.org/10.1111/lang.12246>
- Gong, Y., Gao, X., & Lyu, B. (2020). Teaching Chinese as a second or foreign language to non-Chinese learners in mainland China (2014–2018). *Language Teaching*, 53(1), 44-62. <https://doi.org/10.1017/S0261444819000387>
- Guo, L., & Tao, L. (2008). Tone production in Mandarin Chinese by American students: A case study. In M. K. M. Chan and H. Kang (Eds.), *Proceedings of the 20th North American Conference on Chinese Linguistics (NACCL-20)* (Vol. 1, pp. 123–138). The Ohio State University. <https://u.osu.edu/naccl/proceedings/naccl-20/>
- Guo, Y. (2021). Exploring the dynamic interplay between foreign language enjoyment and learner engagement with regard to EFL achievement and absenteeism: A sequential mixed methods study. *Frontiers in Psychology*, 12, 766058. <https://doi.org/10.3389/fpsyg.2021.766058>
- Gussenhoven, C. (2004). *The phonology of tone and intonation*. Cambridge University Press. <https://doi.org/10.1017/CBO9780511616983>
- 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. <https://doi.org/10.1016/j.wocn.2011.11.001>
- House of Representatives Standing Committee on Education. (2026). *Security and prosperity in Asia: Building Asia capability in Australia through the education system and beyond*. Parliament of Australia. https://www.aph.gov.au/About_Parliament/House_of_Representatives/About_the_House_News/Media_Releases/Security_and_Prosperty_in_Asia_Building_Australias_Asia_Capability
- Jiang, W., Cook, A., Shen, C., Mikhaylova, A., & Richards, H. (2024). Mission impossible for L2 learners: Teachers' perspectives on teaching and learning Chinese in Queensland secondary schools. In K. Hashimoto (Ed.), *Rethinking the Asian language learning paradigm in Australia* (pp. 73–97). Palgrave Macmillan. https://doi.org/10.1007/978-3-031-74149-4_4
- Juhász, K., & Bartos, H. (2024). Could L1 intonation patterns be applied in teaching Mandarin tones to atonal learners of Chinese?—An acoustic phonetic study. *Chinese as a Second Language Research*, 13(2), 157-182. <https://doi.org/10.1515/caslar-2024-2001>
- Kellerman, E. (1979). Transfer and non-transfer: Where we are now. *Studies in Second Language Acquisition*, 2(1), 37-57. <https://doi.org/10.1017/S0272263100000942>
- Leung, K. K. W., Lu, Y.-A., & Wang, Y. (2025). Examining speech perception–production relationships through tone perception and production learning among Indonesian learners of Mandarin. *Brain Sciences*, 15(7), 671. <https://doi.org/10.3390/brainsci15070671>
- Li, C. N., & Thompson, S. A. (1981). *Mandarin Chinese: A functional reference grammar*. University of California Press. <https://doi.org/10.1525/9780520352858>
- Li, L., Valcke, M., Dessein, B., Badan, L., & Anderl, C. (2025). Chinese as a foreign language (CFL) teachers' pedagogical content knowledge in teaching Chinese pronunciation. *Language Teaching Research*, 29(6), 2316-2347. <https://doi.org/10.1177/13621688221117605>
- Li, L., Valcke, M., Dessein, B., Badan, L., & Anderl, C. (2021). Pedagogical content knowledge: A systematic review of Chinese language pronunciation teaching in the CFL context. *Foreign Language Annals*, 54(2), 525–557. <https://doi.org/10.1111/flan.12558>
- Lin, T., & Wang, L. (1992). *Yuyinxue jiaocheng* [A course in phonetics]. Peking University Press.
- Meng, N. (2021). “Repeat after me”: Is there a better way to correct tone errors in teaching Mandarin Chinese as a second language? In C. Yang (Ed.), *The acquisition of Chinese as a second language pronunciation: Segments and prosody* (pp. 163-173). Springer. https://doi.org/10.1007/978-981-15-3809-4_7

- Nagle, C. L. (2018). Examining the temporal structure of the perception–production link in second language acquisition: A longitudinal study. *Language Learning*, 68(1), 234-270. <https://doi.org/10.1111/lang.12275>
- Norman, J. (1988). *Chinese*. Cambridge University Press.
- Ortega-Llebaria, M., & Wu, Z. (2021). Chinese-English speakers' perception of pitch in their non-tonal language: Reinterpreting English as a tonal-like language. *Language and Speech*, 64(2), 467-487. <https://doi.org/10.1177/0023830919894606>
- Renukadevi, D. (2014). The role of listening in language acquisition: The challenges and strategies in teaching listening. *International Journal of Education and Information Studies*, 4(1), 59-63.
- Ringbom, H. (2006). Different types of cross-linguistic similarities. In *Cross-linguistic similarity in foreign language learning* (pp. 5–9). Multilingual Matters. <https://doi.org/10.21832/9781853599361-003>
- Schmidt, R. W. (1990). The role of consciousness in second language learning. *Applied Linguistics*, 11(2), 129-158. <https://doi.org/10.1093/applin/11.2.129>
- Schreiner, L. A., & Nelson, D. D. (2013). The contribution of student satisfaction to persistence. *Journal of College Student Retention: Research, Theory & Practice*, 15(1), 73-111. <https://doi.org/10.2190/CS.15.1.f>
- Shi, H. (2018). *A method of teaching English-speaking learners to produce Mandarin-Chinese tones* [Doctoral dissertation, West Virginia University]. West Virginia University Research Repository. <https://doi.org/10.33915/etd.6633>
- Wang, T., Potter, C. E., & Saffran, J. R. (2020). Plasticity in second language learning: The case of Mandarin tones. *Language Learning and Development*, 16(3), 231-243. <https://doi.org/10.1080/15475441.2020.1737072>
- Wang, Y. (2026). Exploring student engagement with multimodal generative AI in task-based Chinese language learning. *Computer Assisted Language Learning*. Advance online publication. <https://doi.org/10.1080/09588221.2026.2645428>
- Xing, J. Z. (2006). *Teaching and learning Chinese as a foreign language: A pedagogical grammar*. Hong Kong University Press.
- Xu, H., Li, Y., & Li, Y. (2019). Using online applications to improve tone perception among L2 learners of Chinese. *Journal of Technology and Chinese Language Teaching*, 10(1), 26-56. <http://www.tclt.us/journal/2019v10n1/xuli.pdf>
- Yang, B., & Yang, N. (2019). Development of disyllabic tones in different learning contexts. *International Review of Applied Linguistics in Language Teaching*, 57(2), 205-233. <https://doi.org/10.1515/iral-2016-0004>
- Yang, C. (2016). *The acquisition of L2 Mandarin prosody: From experimental studies to pedagogical practice*. John Benjamins Publishing Company. <https://doi.org/10.1075/bpa.1>
- Yang, M., Cooc, N., & Sheng, L. (2017). An investigation of cross-linguistic transfer between Chinese and English: A meta-analysis. *Asian-Pacific Journal of Second and Foreign Language Education*, 2, Article 15. <https://doi.org/10.1186/s40862-017-0036-9>
- Yip, M. (2006). Tone: Phonology. In K. Brown (Ed.), *Encyclopedia of language & linguistics* (2nd ed., pp. 761–764). Elsevier. <https://doi.org/10.1016/B0-08-044854-2/00054-7>
- Zajdler, E. (2021). Between sound and voice: Teaching Chinese tones to non-tonal language speakers. *Roczniki Humanistyczne*, 69(9), 141-158. <https://doi.org/10.18290/rh21699-7>
- Zhang, H. (2018). *Second language acquisition of Mandarin Chinese tones* (Vol. 33). Brill Rodopi. <https://doi.org/10.1163/9789004364790>
- Zhang, J. (2013). *The effects of duration and sonority on contour tone distribution: A typological survey and formal analysis*. Routledge. <https://doi.org/10.4324/9781315024134>
- Zhang, J. (2018). Cross-linguistic influence on Chinese-L2 learners' acquisition of classifiers. *ANU Undergraduate Research Journal*, 9, 156-171. <https://studentjournals.anu.edu.au/index.php/aurj/article/view/138>
- Zhang, L. (2019). Tone features of Chinese and teaching methods for second language learners. *International Journal of Chinese Language Education*, 5, 45-66. https://ijcle.eduhk.hk/uploads/issues/376/IJCLE_No.5_Jun19_03.pdf
- Zhou, C., & Veríssimo, J. (2026). L2 difficulties in the perception of Mandarin tones: Phonological universals or domain-general aptitude? *Bilingualism: Language and Cognition*, 29(2), 405-419. <https://doi.org/10.1017/S1366728925100114>