

Does Second Language Experience Modulate Perception of Tones in a Third Language?

Language and Speech

1–21

© The Author(s) 2015

Reprints and permissions:

sagepub.co.uk/journalsPermissions.nav

DOI: 10.1177/0023830915590191

las.sagepub.com**Zhen Qin and Allard Jongman**

Department of Linguistics, University of Kansas, USA

Abstract

It is unclear what roles native language (L1) and second language (L2) play in the perception of lexical tones in a third language (L3). In tone perception, listeners with different language backgrounds use different fundamental frequency (F0). While English listeners use F0 height, Mandarin listeners rely more on F0 direction. The present study addresses whether knowledge of Mandarin, particularly as an L2, results in speakers' reliance on F0 direction in their perception of L3 (Cantonese) tones. Fifteen English-speaking L2 learners of Mandarin constituted the target group, and 15 English monolinguals and 15 native Mandarin speakers, with no background in other tonal languages, were included as control groups. All groups had to discriminate Cantonese tones either by distinguishing a contour tone from a level tone (F0 direction pair) or a level tone from another level tone (F0 height pair). The results showed that L2 learners patterned differently from both control groups by using F0 direction as well as F0 height under the influence of L1 and L2 experience. The acoustics of the tones also affected all listeners' discrimination. When L2 and L3 are similar in terms of the presence of lexical tone, L2 experience modulates the perception of L3 tones.

Keywords

Speech perception, third language acquisition, Chinese tones

Introduction

The present study addresses the role of linguistic experience, in addition to psychoacoustic factors, in the perception of tones from a new language. Specifically, the study examines how English speakers who are learning Mandarin as a second language perceive Cantonese tones when they are exposed to the target language for the first time. Similar to previous studies (e.g., Rast, 2010; Sanz, Park, & Lado, 2015), this study defines the new language, Cantonese, as a

Corresponding author:

Zhen Qin, Department of Linguistics, University of Kansas, 1541 Lilac Lane, Blake Hall, Room 427, Lawrence, KS 66045-3129, USA.

Email: qinzhentientin2@ku.edu

third language (L3) for learners at the very initial stage of acquisition. For example, French participants in Rast (2010) performed a task in Polish without any prior exposure to or instruction in Polish. Their data were used to investigate language processing “at the level of absolute first contact” (Rast, 2010, p. 166) with the L3.

Linguistic experience is defined as previously acquired languages (i.e., L1 and L2) in this study. While only the L1 could be a source of potential transfer in the case of L2 acquisition, either L1 or L2, or both, could be sources of transfer in the very beginning stage of L3 acquisition (Gass & Selinker, 1983; Leung, 2007; Odlin, 1989; Ojima, Nakata, & Kakigi, 2005; Rothman, 2010, 2011; Steinhauer, White, & Drury, 2009). Several models have been proposed in the L3 acquisition field to account for the transfer source. For instance, the Cumulative Enhancement Model by Flynn, Foley and Vinnitskaya (2004) assumes that both L1 and L2 can facilitate L3 acquisition. According to the model, language acquisition is cumulative, meaning that any prior language either helps or does not affect subsequent acquisition at all. The L2 Status Factor Model proposed by Bardel and Falk (Bardel & Falk, 2007, 2012; Falk & Bardel, 2010, 2011) explicitly states that L2 plays a privileged role in L3 transfer compared with L1 due to the similar cognitive status of the two non-native languages. Besides the role of language status (i.e., L1 vs. L2) in L3 acquisition, the relative proximity of the language systems may influence how languages are transferred in L3 acquisition. Recent studies found that the transfer in L3 acquisition comes from the properties of L2 when L2 and L3 are structurally similar (Leung, 2007; Montrul, Dias, & Santos, 2011). The Typological Proximity Model (TPM) advanced by Rothman and colleagues (Rothman, 2010, 2011, 2013, 2015; Rothman & Cabrelli Amaro, 2010) proposes that the influence of source language (i.e., L1 and L2) for this transfer depends on the typological proximity between the source and target languages. Besides the influence of previously acquired languages, psychoacoustic factors can influence perception of L3 sounds. Basically, the larger the acoustic difference between two sounds, the easier their discrimination will be, regardless which languages have been previously acquired.

The present study is unique in that it addresses the question of what role L1 and L2 play in L3 acquisition by investigating lexical tones. The L3 models listed above were based on studies examining lexicon and morpho-syntax. Research on phonology and phonetics in L3 acquisition has been far from sufficient and often limited in scope (García Mayo, 2012); this is changing, though, and while several recent L3 studies focus on the acquisition of segments (e.g., Cabrelli Amaro, 2012; Wrembel, 2012; Wrembel, Ulrike, & Grit (2010); see Special Issue of *International Journal of Multilingualism* 2010 for more), very few L3 studies have examined the perception of tones. Lexical tone is an informative linguistic structure for testing the influence of L1 and L2 in L3 acquisition because L2 is probably the source of transfer to L3 if lexical tones are present in L2, but not L1 (Rothman, 2010, 2011, 2013, 2015). The present study recruited English-speaking learners with Mandarin as L2 and introduced them to Cantonese as L3. Given the greater structural similarity between L2 and L3 compared with L1 and L3 in terms of lexical tones, it is revealing to examine whether L2 experience can override L1 influence and modulate learners’ use of cues in their perception of L3 tones.

Speakers of tonal languages use F0 cues to distinguish lexical meanings whereas speakers of non-tonal languages typically use F0 cues to convey pragmatic meaning in intonation (Ladd, 2008). Previous studies found that English listeners attend to F0 height as a cue and process lexical tones as non-linguistic units because there is no lexical tone in English. In contrast, Mandarin listeners are sensitive to F0 direction as a cue and process lexical tones as linguistic units, because Mandarin is a tone language (Burnham & Mattock, 2007; Gandour, 1983). In addition to the contour tone contrasts found in Mandarin, Cantonese has level tone contrasts, so listeners need to attend to both F0 height and F0 direction to identify Cantonese tones. Therefore, the

present study uses Cantonese tones as stimuli to examine whether English-speaking L2 learners of Mandarin perceive Cantonese tones using F0 height like English listeners or F0 direction like Mandarin listeners.

The paper is organized as follows: First, the findings of previous studies investigating L2 perception of lexical tones are discussed; second, relevant facts are presented about the target L1–L2–L3 pairing (English–Mandarin–Cantonese) used in this study (based on previous findings, research questions and predictions are then formulated and tested with a tone discrimination experiment); finally, the paper concludes with a discussion of the results and their implications for theories of L2/L3 speech perception.

1.1 L2 perception of lexical tones

First of all, psychoacoustic factors affect all listeners in a similar way in the perception of tones. For example, some cues such as F0 onset, are acoustically more salient than other cues. Lee, Tao, and Bond (2008) found that when processing Mandarin tones with only F0 onset (i.e., the first six F0 periods from the beginning of voicing) Mandarin listeners were much more accurate than chance, which shows that the F0 onset is salient enough to provide some of the tone identity information. Burnham and Francis (1997) reported that the onset F0 height of Thai tones was acoustically more salient than the offset F0 height. Both native and non-native listeners relied on this cue to distinguish Thai tones.

Besides psychoacoustics, linguistic experience is generally considered to influence perception of non-native sounds. For instance, L2 speech perception models propose that phonetic similarities/dissimilarities determine whether two L2 sounds are assimilated into one sound or two L1 sounds (e.g., Best, 1995; Flege, 1995). Specifically for lexical tone, many studies explored the effects of L1 influence on non-native perception of tones in terms of perceptual cues (Francis, Ciocca, Ma, & Fenn, 2008; Gandour, 1983), as well as category assimilation (So & Best, 2010; Wu, Munro, & Wang, 2014), and the effects of L2 learning experience or phonetic training on adult acquisition (Guion & Pederson, 2007; Hao, 2012; Wang, Spence, Jongman, & Sereno, 1999).

Both native speakers of tone languages (i.e., Mandarin) and non-tone languages (i.e., English) are included in the present study. First, the phonemic status of F0 in L1, whether or not F0 is lexically decoded, plays an important role in determining tone perception performance. Second, L1 experience will influence which perceptual cues are used.

Most previous studies showed that speakers of tone languages performed better than speakers of non-tone languages in perceiving and processing lexical tones (e.g., Lee, Vakoch, & Wurm, 1996; Wang, Behne, Jongman, & Sereno, 2004; Wang, Jongman, & Sereno, 2001; Wayland & Guion, 2004), though some studies did not (e.g., Hao, 2012; So & Best, 2010). Burnham and Mattock (2007) suggested that speakers of non-tone languages (e.g., English listeners) perceived lexical tones primarily in an acoustic mode, relying mainly on the psychoacoustic similarity of tonal stimuli. Meanwhile, speakers of tone languages relied more on their native tonal system and perceived tones primarily in a linguistic mode.

Categorical perception of tones is also determined by the phonemic status of F0 in L1. Prior studies found that speakers of tone languages perceive tones in a categorical manner (e.g., Hallé, Chang, & Best, 2004; Peng et al., 2010; Wang, 1976; Xu, Gandour, & Francis, 2006). In contrast to speakers of tone languages, speakers of non-tone languages do not hear tones categorically. Wang (1976) reported that Mandarin listeners, but not English listeners, showed a typical pattern of categorical perception in perceiving a dynamic tonal continuum varying from a level tone to a rising tone. While Mandarin listeners exhibited a linguistic boundary in perceiving F0 contours,

English listeners appeared to make judgments on the basis of psychoacoustic properties of the stimuli. Although speakers of non-tone languages do not hear tones categorically, they may have a psychoacoustic advantage over speakers of tone languages. While Mandarin listeners are not sensitive to minor F0 changes in their tonal categorization, English listeners are able to detect subtle F0 variations, that is, within-category differences for Mandarin listeners (Leather, 1987; Stagray & Downs, 1993).

In short, most studies found that speakers of tone languages differ from those of non-tone languages in tone perception. Previous studies also showed that they differed in which cues they primarily use in the perception of tones.

Although other acoustic properties (e.g., duration, intensity and F0 turning point) can be perceptual cues to lexical tone, F0 is the primary cue in the perception of Mandarin tones (Howie, 1976; Lin & Wang, 1984; Moore & Jongman, 1997) and Cantonese tones (Fok-Chan, 1974; Khouw & Ciocca, 2007).

However, while F0 is the primary cue in tone perception, speakers of tone languages and non-tone languages often attend to different dimensions of F0. Previous studies agreed that English listeners attended to F0 height rather than F0 direction, whereas Mandarin listeners were more sensitive to F0 direction than F0 height (Francis et al., 2008; Gandour, 1983). Different perceptual cues were found not only between speakers of tone languages and non-tone languages, but also among speakers of different tone languages. In Gandour (1983), differences regarding F0 dimensions were found between Taiwan Mandarin and Cantonese listeners. While Taiwan Mandarin listeners assigned more weight to F0 direction than F0 height, Cantonese listeners were sensitive to both F0 direction and F0 height.

Previous studies found that L2 experience can facilitate learners' perception of L2 tones. For instance, Guion and Pederson (2007) found that late advanced learners of Mandarin began to behave like Mandarin listeners rather than English listeners by attending to both F0 slope and average F0 when judging synthesized tones.

Training studies in a laboratory setting consistently show that non-native listeners without exposure to tone languages improved both their perception and production of lexical tones after a short period of perceptual tone training (Wang, Jongman, & Sereno, 2003; Wang et al., 1999). For instance, in a recent study, Francis et al. (2008) included two groups of listeners, one with native speakers of a tone language (i.e., Mandarin Chinese) and one with native speakers of a non-tone language (i.e., English) and trained them to recognize Cantonese lexical tones (three contour tones and three level tones). Both groups showed significant improvement in identifying Cantonese tones after perceptual training. Interestingly, while the Mandarin listeners began to behave like native Cantonese listeners by showing almost equal weighting of F0 direction and height in the post test, the English listeners only attended to F0 height after training. In short, speakers' perceptual system demonstrates plasticity, and it can be modified by increasing exposure to L2 tones.

Most of the studies examining the effects of L2 experience are limited to perceptual training contexts (e.g., Francis et al., 2008; Wang et al., 1999, 2003), and studies exploring the effects of L2 experience in an adult acquisition context (e.g., Guion & Pederson, 2007; Hao, 2012). More importantly, few studies have investigated how L1 and L2 experience interact in shaping the perception of L3 tones. The present study fills these gaps by examining how Cantonese tones are perceived by English-speaking learners with Mandarin as L2 and Cantonese as L3.

1.2 The use of F0 in Cantonese, Mandarin and English

1.2.1 Cantonese tones. Cantonese uses F0 to distinguish lexical meanings. As illustrated in Figure 1, there are six lexical tones in Cantonese (Bauer & Benedict, 1997; Chao, 1947). Following Bauer

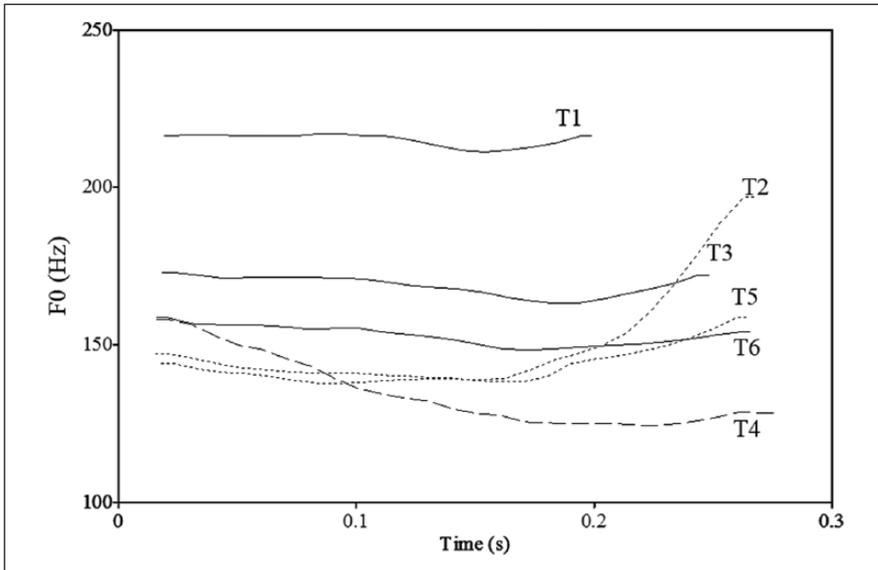


Figure 1. Tone contours of the six Cantonese tones as produced by a female native speaker of Hong Kong Cantonese used in the present study.

and Benedict (1997), the six tones are labeled as follows: Tone1 (T1), a High Level (55) tone; T2, a High Rising (25) tone; T3, a Mid-Level (33) tone; T4, a Low Falling (21) tone; T5, a Low Rising (23) tone; and T6, a Low Level (22) tone. The six Cantonese tones have similar duration, and duration is not a cue in tone perception for Cantonese listeners (Fok-Chan, 1974; Khouw & Ciocca, 2007). Glottalization may be a cue to perceive the low falling tone in Cantonese, but it does not function as a consistent cue for Cantonese listeners (Vance, 1976).

F0 is the primary cue to distinguish the six lexical tones. There are three level tones (T1, T3 and T6), which differ only in F0 height, and three contour-level tones (T2, T4 and T5), which differ primarily in F0 direction. Due to acoustic similarities among some of the Cantonese tones, some contour-level tones are harder to discriminate than others. For instance, the two rising tones (i.e., T2 and T5), which only differ in the magnitude of the rise in F0, were found to be difficult even for native listeners to distinguish and appear to be merging (Fok-Chan, 1974; Mok, Zuo, & Wong, 2013). Similarly, as T4, T5, and T6 share a similar starting point and diverge only at the offset, the three tones are found to be confused by native listeners (Fok-Chan, 1974).

Mandarin Tones Mandarin distinguishes four lexical tones: T1, a High Level (55) tone; T2, a Rising (35) tone; T3, a Dipping (214) tone; and T4, a Falling (51) tone, as illustrated in Figure 2 (Chao, 1968; Li & Thompson, 1989).

While the Mandarin tones are different from the Cantonese tones, the tonal systems in the two languages have some tones in common. On the one hand, the tone system in Mandarin appears to be simpler than that in Cantonese. There is no contrast between level tones, as there is only one level tone in Mandarin. On the other hand, the high level tone, T1, and the high rising tone, T2, in Mandarin are similar to T1 and T2 in Cantonese in terms of F0 height and shape.

Given the differences and similarities between the tonal systems in the two languages, it is predicted that Mandarin listeners will not have difficulty perceiving Cantonese T1 and T2, but they will have great difficulty distinguishing the three Cantonese level tones. Qin and Mok (2013) observed that although Mandarin listeners distinguished the Cantonese contour-level tones well,

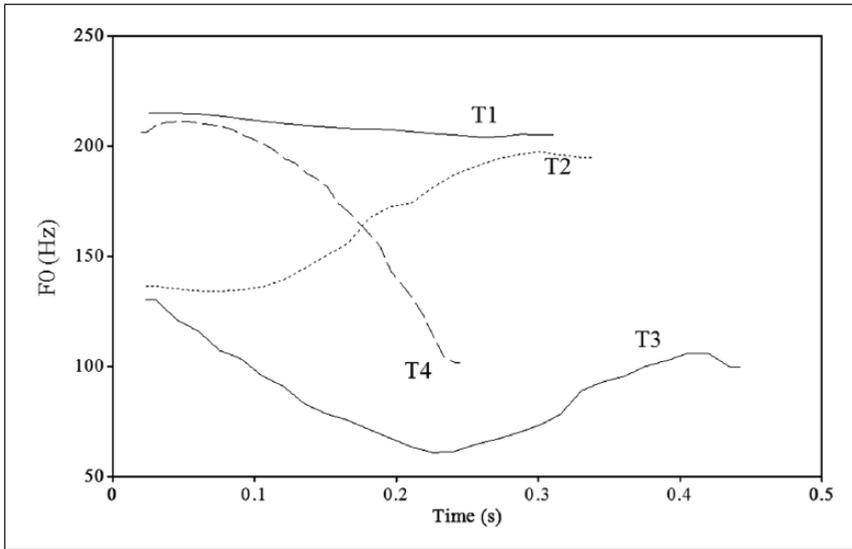


Figure 2. Tone contours of the four Mandarin tones as produced by a male native speaker of Beijing Mandarin.

they often confused the three level tones even more than English listeners. However, this observation was limited to a small sample. It is therefore of interest to examine whether L2 learners of Mandarin (initial-state L3 learners of Cantonese) would use the Mandarin tones they learned in L2 to perceive tones from an unfamiliar language. These questions motivate the present study.

1.2.2 The use of F0 in English. Different from Cantonese and Mandarin, English uses F0 variation to convey pragmatic meanings in intonation and to contrast stress (Ladd, 2008). However, it does not use F0 to distinguish lexical meaning. Since English does not have lexical tones, it is not possible for English listeners to transfer L1 tones to perceive non-native tones. Previous studies showed that English listeners perceive tones acoustically as non-linguistic units, and maintain some sensitivity to minor F0 changes within a tonal category (Burnham & Mattock, 2007; Leather, 1987; Stagray & Downs, 1993). It is predicted that English listeners, although not as sensitive to tones as Mandarin listeners are, may show greater sensitivity to the subtle differences among level tones than Mandarin listeners.

2 The present study

Since few studies have investigated how L1 and L2 experience modulates the perception of L3 tones, the present study fills this gap by examining how Cantonese tones are perceived by English-speaking learners of Mandarin, for whom Cantonese is an unfamiliar L3. Based on the prosodic systems of Cantonese, Mandarin and English, the research questions that this study intends to address are formulated as follows:

First, broadly speaking, is perception of L3 tones affected by linguistic experience and/or psychoacoustic factors?

Second, given that lexical tones are shared between L2 and L3 does L2 experience modulate listeners' use of cues in terms of F0 height versus F0 direction?

For the first question, it is predicted that both linguistic experience and psychoacoustic factors influence listeners' perception of Cantonese tones. In terms of linguistic experience, Mandarin listeners are sensitive to F0 direction, so they should discriminate contour–level tone pairs better than level–level tone pairs. In contrast, English does not have lexical tones, and listeners attend to F0 height rather than F0 direction. Therefore, English listeners should be more sensitive to level tones than the Mandarin listeners. In terms of psychoacoustic factors, acoustic properties of tonal stimuli may influence listeners' sensitivity to tones, regardless of listeners' language backgrounds. It is predicted that acoustically easy tone pairs with a larger acoustic distance will be more accurately distinguished than acoustically hard tone pairs with a smaller acoustic distance. Regarding the second question, it is predicted that English-speaking learners of Mandarin will use F0 direction to perceive lexical tones in L3 Cantonese due to the influence of their Mandarin learning experience.

3 Experiment

To explore whether L2 experience modulates learners' use of cues (i.e., F0 height and F0 direction) in perceiving L3 tones, an AX forced-choice tone discrimination task using Cantonese tones was conducted for the three groups of listeners.

3.1 Participants

Fifteen English-speaking learners of Mandarin were included as this study's target group. All of them were college students studying Chinese at the University of Kansas. Their average age of acquisition was 21.5 years (*SD*: 3.7) and they had received an average of 2.17 years (*SD*: 0.8) of Chinese instruction¹. Eleven of them self-reported their proficiency in Chinese as intermediate. To test the possible L1 and L2 influence on the perception of L3 tones, fifteen English native speakers with no exposure to any tone languages (i.e., English controls) and fifteen Mandarin native speakers with no background in other tone languages (i.e., Mandarin controls) were recruited as control groups.

A questionnaire was used to collect information about the participants' language backgrounds. None of the participants had been exposed to Cantonese before. They reported normal hearing and no history of speech or language disorders. Information about their musical experience (e.g., years of musical training, playing an instrument) was collected. While the Mandarin controls had less musical experience than the other two groups, the English-speaking L2 learners of Mandarin and the English controls were similar in their musical experience. Participants were each paid US\$10 for their participation in the experiment.

3.2 Stimuli

Level and contour Cantonese tones carried by two syllables were used as stimuli, allowing for a test of both F0 height and F0 direction. Two syllables, [jəu] and [se] (/jau/ and /se/ in Jyutping, a Cantonese Romanization system used in Hong Kong) were used as test stimuli for three reasons: First, the four target tones with the two syllables all form real Cantonese words; second, these stimuli were used in previous studies (e.g., Francis et al., 2008); and third, the two syllables were used to counterbalance the voicing of initial consonants, which may cause F0 variations.

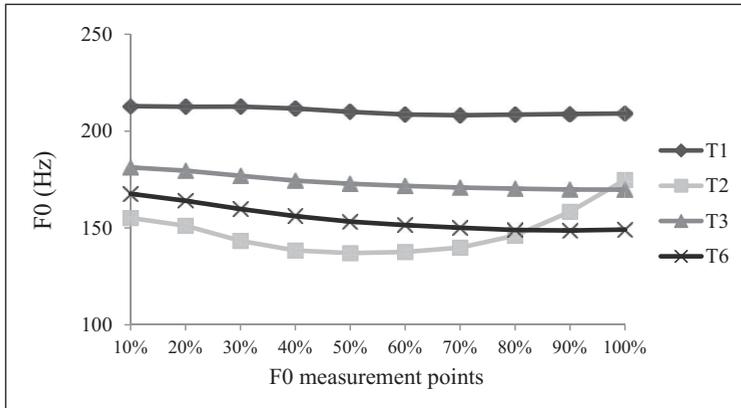


Figure 3. F0 measurements of our stimuli collapsed across tokens and syllables.

One female native speaker of Hong Kong Cantonese was recorded producing all the stimuli in a quiet room. The speaker read a randomized list of words embedded in a carrier phrase *ngo5duk6_zi6* “I read the word___” three times at a normal speech rate. The carrier sentence was selected to avoid the final lengthening effect on the tone stimuli. Tone sandhi does not apply in this sentence. The recording was done directly on disk using a sampling rate of 44.1 kHz. The three tokens (i.e., repetitions) of four target tones carried by each syllable were excised, which resulted in 24 (4 tones \times 2 syllables \times 3 repetitions) stimuli in total. Intensity of all the stimuli was normalized at 70 dB. The duration of each syllable was normalized by shortening or lengthening each token in Praat at an average value of /jau/ (345 ms) and /se/ (488 ms).

Among the six lexical tones of Cantonese, four tones (T1, T2, T3, and T6) were chosen as target tones. The other two tones, T4 (a low falling tone) and T5 (a low rising tone) were excluded from our study, because they are often confused with other tones (e.g., T6 and T2) even by native listeners (Fok-Chan, 1974; Mok et al., 2013). Ten temporally equidistant F0 measurements were taken for each stimulus using ProsodyPro (Xu, 2013) and averaged across tokens. The resulting F0 contours are shown in Figure 3. Consistent with previous studies (e.g., Mok et al., 2013), this study’s F0 measurements show that T2 is a rising tone, and T1, T3 and T6 are high-level, mid-level and low-level tones, respectively.

The four tones carried by the two syllables are plotted separately for each syllable in Figure 4. They have a tonal distribution which is consistent with previous studies (e.g., Mok et al., 2013), but some phonetic variations of tone contours are observed across the two syllables. Firstly, the tones for /se/ ($SD=28.6\text{ Hz}^2$) are more separated from each other than those for /jau/ ($SD=25.8\text{ Hz}$). For instance, T3 and T6 are much closer for /jau/ than for /se/. The tonal variations may result in a difference in discrimination accuracy between the two syllables.

Moreover, the tones carried by /se/ show a falling F0 at the initial part and have a slightly higher F0 onset than those carried by /jau/. This is because F0 onset often varies with syllable-initial consonant voicing (Hombert, 1978). /se/ has a voiceless initial obstruent, which raises the F0 onset of the following vowel and yields a falling F0 onset pattern.

3.3 Tone pairs

In order to test listeners’ sensitivity to F0 height and F0 direction, the four tones (T1, T2, T3, and T6) were then paired into four tonal contrasts; two pairings were level–level (T1–T6 and T3–T6)

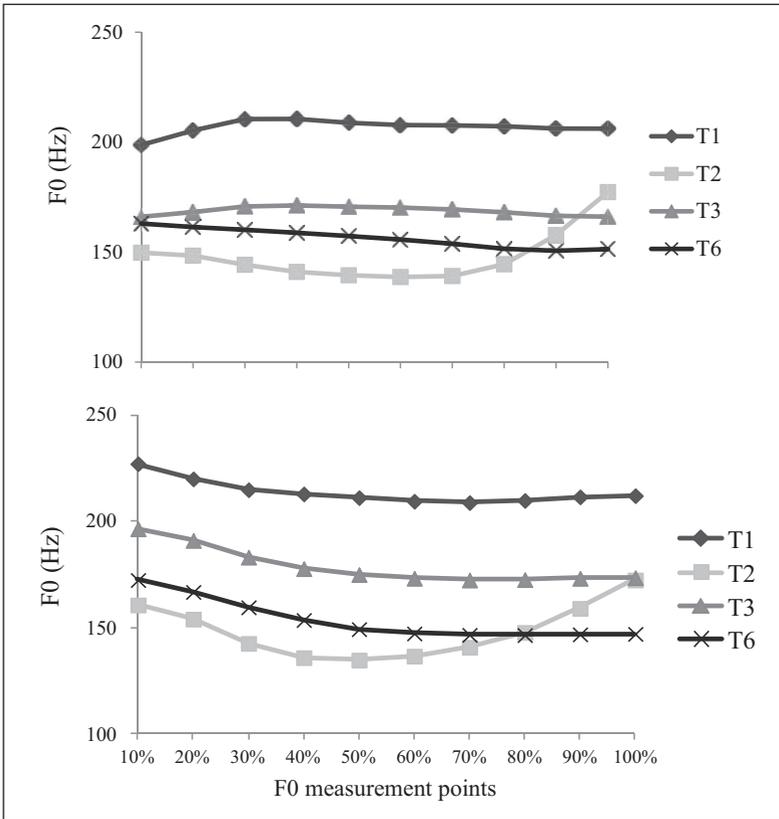


Figure 4. F0 measurements of stimuli carried by /jau/ (top panel) and /se/ (bottom panel) collapsed across tokens.

Table 1. The four target tone pairs included in the present experiment.

Types	Primary cues	Tone pairs
Level-level tone pairs	F0 height	T1-T6 (larger acoustic difference) T3-T6 (smaller acoustic difference)
Contour-level tone pairs	F0 direction	T2-T1 (larger acoustic difference) T2-T6 (smaller acoustic difference)

and the other two were contour-level (T2-T1 and T2-T6), as illustrated in Table 1. The level-level tone pairs in which two tones in each pair differ in F0 height were used to test listeners’ sensitivity to F0 height. While T1-T6 has a larger difference (50.5 Hz³ for /jau/; 60.2 Hz for /se/) in F0 height, T3-T6 has a smaller difference (12.4 Hz for /jau/; 25.2 Hz for /se/) in F0 height, as illustrated in Figures 3 and 4. Hence, T1-T6 and T3-T6 are acoustically easy and hard level-level tone pairs, respectively.

The contour-level tone pairs in which two tones in each pair differ primarily in F0 direction were used to assess listeners’ use of F0 direction. While T2-T1 has a larger acoustic difference (58.8 Hz for /jau/; 65.4 Hz for /se/), T2-T6 has a smaller acoustic difference (8.3 Hz for /jau/; 5.1

Hz for /se/). Hence, T2–T1 and T2–T6 are acoustically easy and hard contour–level tone pairs, respectively.

3.4 Procedure

An AX forced-choice discrimination task was conducted. Two types of tone pairs were used as stimuli, AA pairs (pairs with the same tone) as fillers and AB pairs (pairs with different tones) as target stimuli. 144 AB pairs (4 pairs \times 2 orders \times 2 syllables \times 3 tokens⁴ \times 3 tokens⁵) and 144 AA pairs (4 pairs \times 2 syllables \times 3 tokens \times 3 tokens \times 2 repetitions) were created to counterbalance the two types of tone pairs. The presentation order of each AB pair was also counter-balanced. The stimuli were presented in random order to the participants.

The stimuli were presented to the participants via stereo headphones with the volume adjusted to a comfortable level. The participants were told that they would hear pairs of sounds from a language they were not familiar with. Their task was to judge whether the two tones in each pair were “same” or “different” as quickly and accurately as possible by pressing one of two buttons which were labeled “same” or “different”, respectively, on the keyboard.

The task was conducted using Paradigm software (Perception Research Systems, 2007): The inter-stimulus interval (ISI) was 500 ms⁶; the time-out time was 6 seconds and invalid responses were excluded from analysis; and instructions were given orally as well as displayed visually on a computer screen before the test. No feedback was given. Before the actual experiment, the participants familiarized themselves with 12 practice trials randomly selected from the test stimuli. Accuracy percentage was collected only for different pairs.

3.5 Data analysis

Logit mixed-effects models were performed on the participants’ accuracy data (1 = correct, 0 = incorrect). The models were fitted in R, using the `lmer()` function from the `lme4` package for mixed-effects models (for discussion, see Baayen, 2008).

A first model analyzed the accuracy of each of the three groups in perceiving contour–level versus level–level tone pairs, respectively, with Tone pair (contour–level vs. level–level), Group (L2 Learners vs. English controls and Mandarin controls), and the interaction between the two as fixed variables. In this model, the contour–level tone pairs and the L2 learners group were used as baselines. A second model analyzed only the two monolingual groups’ accuracy, with Group (English vs. Mandarin) as a fixed variable. The baseline for Group was English.

Four more models were used in the same way for the tones carried by /jau/ (models three and four) and /se/ (models five and six), respectively.

A seventh model analyzed the accuracy of each of the three groups in perceiving tones carried by /jau/ versus /se/, with Syllable (/jau/ vs. /se/), Group (L2 Learners vs. English controls and Mandarin controls), and the interaction between the two as fixed variables. In this model, the /jau/ stimuli and the L2 learners group were used as baselines. An eighth model analyzed only the two monolingual groups’ accuracy, with Group (English vs. Mandarin) as a fixed variable. The baseline for Group was English.

A ninth model analyzed the accuracy of each of the three groups in perceiving the two different contour–level tone pairs, with Tone pair (T2–T1 vs. T2–T6), Group (L2 Learners vs. English controls and Mandarin controls), and the interaction between the two as fixed variables. In this model, T2–T1 and the L2 learners group were used as baselines. A tenth model analyzed only the two monolingual groups’ accuracy, with Group (English vs. Mandarin) as a fixed variable. The baseline for Group was English.

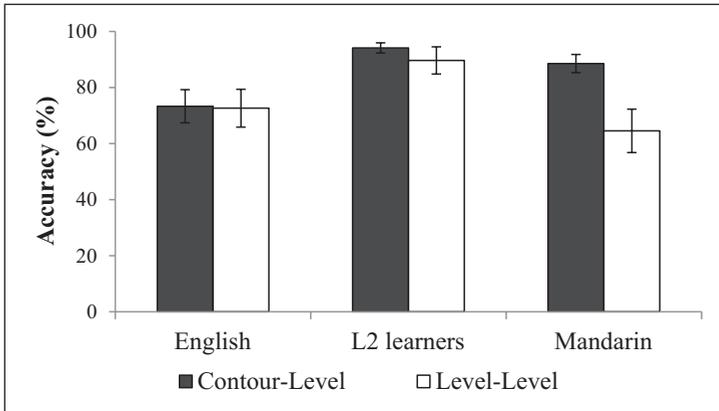


Figure 5. Discrimination accuracy (and standard errors) of contour-level versus level-level tone pairs in L3 Cantonese by English controls, L1 English/L2 Mandarin subjects and Mandarin controls.

An eleventh model analyzed the accuracy of each of the three groups in perceiving the two different level-level tone pairs, with Tone pair (T1–T6 vs. T3–T6), Group (L2 Learners vs. English controls and Mandarin controls), and the interaction between the two as fixed variables. In this model, T1–T3 and the L2 learners group were used as baselines. A twelfth model analyzed only the two monolingual groups' accuracy, with Group (English vs. Mandarin) as a fixed variable. The baseline for Group was English.

In all the models, participant and item were crossed random variables.

4 Results

4.1 The effect of Tone Pair Type

Discrimination accuracy of AA pairs was 95.9%, 98.5%, and 87.8% for the English controls, L2 learners and Mandarin controls, respectively. Moreover, accuracy of the three groups in discriminating AB pairs was 72.9%, 92% and 73.7%, respectively. Accuracy of both AA and AB pairs for each group was much higher than chance level (i.e., 50%). No bias to “same” or “different” responses was found in discrimination performance for any group. In addition to accuracy data, the *d*-prime data (MacMillan & Creelman, 1991) will also be reported. *d*-prime serves to tease apart any response bias from sensitivity and is commonly applied to discrimination of tones (e.g., Francis & Ciocca, 2003; Wayland & Guion, 2004).

The following analysis focuses on accuracy for AB pairs alone. Accuracy of the three groups in discriminating contour-level and level-level pairs is illustrated in Figure 5.

A first logit mixed-effects model, focusing on the L2 learners, was performed on the participants' accuracy. The results, as illustrated in Table 2, showed a significant Tone Pair Type x Group interaction for both the English group and the Mandarin group. This indicates that the L2 learners differed from the English and the Mandarin controls in their discrimination of contour-level and level-level tone pairs.

A second logit mixed-effects model, focusing on the two monolingual groups, also showed a significant Tone Pair Type x Group interaction for the Mandarin controls. This indicates that the Mandarin controls differed from the English controls in their discrimination of contour-level and level-level tone pairs.

Table 2. Logit mixed-effects models on all participants' overall accuracy.

Variable	Est.	SE	z	p
L2 learners' data (df = 6125)				
Group (E)	-1.9	0.15	-12.6	<0.001
Group (M)	-0.8	0.17	-4.7	<0.001
Tone Pair Type	-0.6	0.17	-3.9	<0.001
Group (E) × Tone Pair Type	0.6	0.20	3.1	0.002
Group (M) × Tone Pair Type	-1.0	0.21	-4.6	<0.001
English versus Mandarin (df = 3977)				
Variable	Est.	SE	z	p
Group	1.2	0.12	9.2	<0.001
Tone Pair Type	-0.1	0.10	-0.4	0.67
Group × Tone Pair Type	-1.6	0.16	-9.9	<0.001

Given the interaction effects found in the first two models, three *post-hoc* analyses were done to test the effect of Tone Pair Type for each group. The results revealed that while the English controls $z(1841) < |1|$, did not show a significant effect of Tone Pair Type, the L2 learners, $z(1845) = -4.2$, $p < 0.001$, and the Mandarin controls, $z(1833) = -13.7$, $p < 0.001$, both showed a significantly higher accuracy for the contour-level tone pairs than for the level-level tone pairs. This suggests that the L2 learners and Mandarin controls showed a similar advantage in discriminating the contour-level tone pairs versus the level-level tone pairs.

Moreover, six *post-hoc* analyses were done to test the group effect for each type of tone pair. Three *post-hoc* analyses on contour-level tone pairs showed that the L2 learners were better than the English controls, $z(2918) = 12.4$, $p < 0.001$, and the Mandarin controls, $z(2918) = 4.6$, $p < 0.001$, at using F0 direction to perceive the contour-level tone pairs. The Mandarin controls were better, $z(1844) = 9.2$, $p < 0.001$, than the English controls at using F0 direction to distinguish contour-level tone pairs. Three *post-hoc* analyses on level-level tone pairs showed the L2 learners did better than the English controls, $z(2904) = 10.4$, $p < 0.001$, and the Mandarin controls, $z(2904) = 13.9$, $p < 0.001$, at using F0 height to perceive the level-level tone pairs. The English controls were better, $z(1830) = 4.3$, $p < 0.001$, than the Mandarin controls at using F0 height to distinguish level-level tone pairs. This suggests that, different from the Mandarin controls that were good at using F0 direction and the English controls that were good at using F0 height, the L2 learners were sensitive to both F0 direction and F0 height compared with the other two groups.

In addition to accuracy data, to assess the participants' discrimination ability, a *d*-prime score for each participant was derived based on the "hit" rate (number of times the "different" button was pressed for AB pairs) and the "false alarm" rate (number of times the "different" button was pressed for AA pairs). *d*-prime score is the difference between "hit" rate and the "false alarm" rate when they are *z*-transformed (MacMillan & Creelman, 1991). Overall, as illustrated in Figure 6, the *d*-prime results show a very similar pattern to the accuracy data.

A repeated measures two-way ANOVA was conducted on the *d*-prime scores with Tone Pair Type as within-subjects factor and Group as between-subjects factor. The results showed main effects of Tone Pair Type, $F(1,42) = 57$, $p < 0.001$, Group, $F(2,42) = 9.9$, $p < 0.001$, and a significant interaction between the two, $F(2,42) = 15.4$, $p < 0.001$.

Three paired *t*-tests were conducted on the effect of Tone Pair Type for each group. While the English controls, $t(14) = 1.5$, $p = 0.14$, did not show a significant effect of Tone Pair Type, the

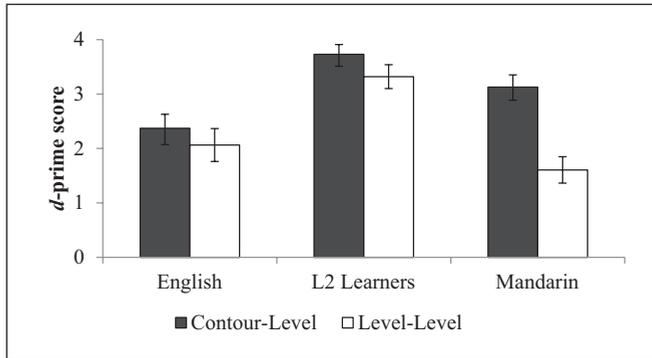


Figure 6. *d*-prime scores (and standard errors) of contour-level versus level-level tone pairs in L3 Cantonese by English controls, L1 English/L2 Mandarin subjects and Mandarin controls.

learner group, $t(14) = 2.9, p = 0.012$ and the Mandarin controls, $t(14) = 8.9, p < 0.001$, both showed a significant effect of Tone Pair Type. Consistent with accuracy data, this suggests that the L2 learners and native speakers of Mandarin showed a similar advantage in discriminating the contour-level tone pairs versus the level-level tone pairs.

Moreover, six independent sample *t*-tests were conducted on the effect of Group for each tone pair. Three *t*-tests on contour-level tone pairs showed that the L2 learners were better than the English controls, $t(28) = 4.6, p < 0.001$, and the Mandarin controls, $t(28) = 2.2, p = 0.035$, at using F0 direction to perceive the contour-level tone pairs. The Mandarin controls were better, $t(28) = 2.1, p = 0.049$, than the English controls at using F0 direction to distinguish contour-level tone pairs. Three *t*-tests on level-level tone pairs showed the L2 learners did better than the English controls, $t(28) = 3.3, p = 0.003$, and the Mandarin controls, $t(28) = 5.1, p < 0.001$, at using F0 height to perceive the level-level tone pairs. But there was no significant difference between the English and the Mandarin controls on level-level tone pairs ($p = 0.24$). Consistent with the accuracy data, this suggests the L2 learners were sensitive to both F0 direction and F0 height compared with the other two groups.

4.2 The effect of syllable

Since some phonetic variation in tone contours was observed across the two syllables, the contour-level and level-level tone pairs were analyzed separately for /jau/ and /se/, as illustrated in Figure 7, to investigate the potential difference in listeners' accuracy for the two syllables.

A third logit mixed-effects model, focusing on the L2 learners, was used for the participants' accuracy on the syllable /jau/. The results, as illustrated in Table 3, showed a significant Tone Pair Type \times Group interaction for the Mandarin controls, and a marginally significant interaction for the English controls, $z(2913) = 1.8, p = 0.07$. This indicates that the L2 learners differed from the Mandarin controls, and might also differ from the English controls given the marginal significance, in their discrimination of contour-level and level-level tone pairs.

A fourth logit mixed-effects model, focusing on the two monolingual groups, also showed a significant Tone Pair Type \times Group interaction for the Mandarin controls. This indicates that the Mandarin controls differed from the English controls in their discrimination of contour-level and level-level tone pairs.

Three *post-hoc* analyses of Tone Pair Type revealed that while the English group, $z(776) < |1|$, did not show a significant effect of Tone Pair Type, the L2 learners, $z(773) = -2.4, p = 0.016$, and

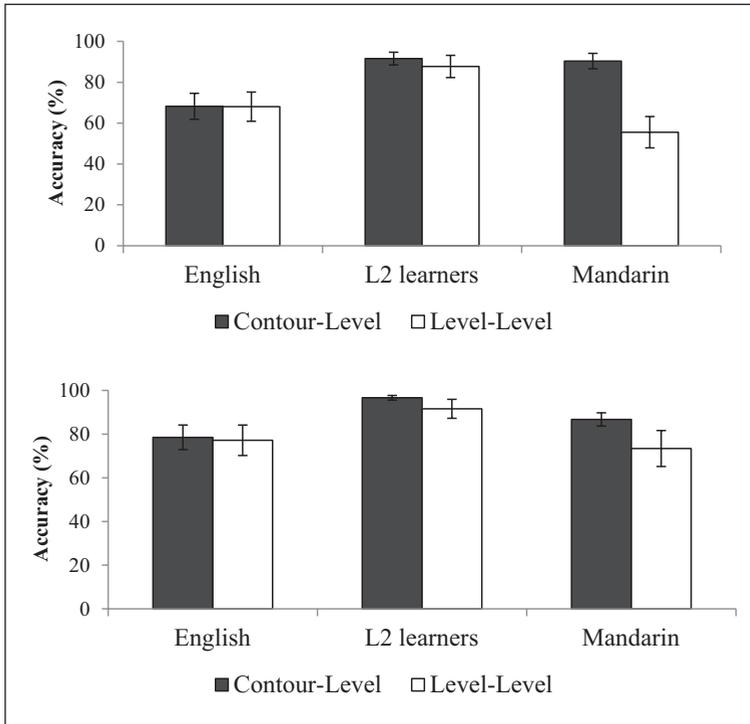


Figure 7. Discrimination accuracy (and standard error) of contour–level versus level–level tone pairs carried by /jau/ (top panel) and /se/ (bottom panel) in L3 Cantonese for English controls, L1 English/L2 Mandarin subjects and Mandarin controls.

Table 3. Logit mixed-effects models on all participants' accuracy on /jau/.

Variable	Est.	SE	z	p
L2 learners' data (df = 2913)				
Group (E)	-1.8	0.19	-9.5	<0.001
Group (M)	-0.2	0.22	-0.8	0.44
Tone Pair Type	-0.5	0.21	-2.2	0.03
Group (E) x Tone Pair Type	0.5	0.25	1.8	0.07
Group (M) x Tone Pair Type	-1.8	0.28	-6.4	<0.001
English versus Mandarin (df = 1839)				
Group	1.6	0.18	9.0	<0.001
Tone Pair Type	-0.1	0.14	-0.1	0.9
Group x Tone Pair Type	-0.6	0.23	-9.8	<0.001

the Mandarin group, $z(766) = -12.3, p < 0.001$, both showed a significantly higher accuracy on the contour–level tone pairs than the level–level tone pairs. This suggests that the L2 learners and native speakers of Mandarin showed a similar advantage in discriminating the contour–level tone pairs versus the level–level tone pairs carried by /jau/.

Table 4. Logit mixed-effects models on all participants' accuracy on /se/.

Variable	Est.	SE	z	p
L2 learners' data (df = 2909)				
Group (E)	-2.3	0.27	-8.3	<0.001
Group (M)	-1.6	0.28	-5.6	<0.001
Tone Pair Type	-1.0	0.30	-3.5	<0.001
Group (E) x Tone Pair Type	1.0	0.34	2.8	0.005
Group (M) x Tone Pair Type	0.1	0.34	0.1	0.9
English versus Mandarin (df = 1835)				
Group	0.7	0.18	3.8	<0.001
Tone Pair Type	-0.1	0.16	-0.6	0.5
Group x Tone Pair Type	-0.9	0.23	-3.9	<0.001

A fifth logit mixed-effects model, focusing on the L2 learners, was used for the participants' accuracy for the syllable /se/. The results, as illustrated in Table 4, showed a significant Tone Pair Type x Group interaction for the English group, but not for the Mandarin group. This indicates that the L2 learners differed from the English controls, but not from the Mandarin controls, in their discrimination of contour-level and level-level tone pairs.

A sixth logit mixed-effects model, focusing on the two monolingual groups, showed a significant Tone Pair Type x Group interaction for the Mandarin controls. This indicates that the Mandarin controls differed from the English controls in their discrimination of contour-level and level-level tone pairs.

Three *post-hoc* analyses of tone pair type revealed that while the English group, $z(770) < |1|$, did not show a significant effect of Tone Pair Type, the L2 learners, $z(774) = -3.6$, $p < 0.001$, and the Mandarin group, $z(771) = -6.0$, $p < 0.001$, both showed a significantly higher accuracy on the contour-level tone pairs than the level-level tone pairs. This suggests that, similar to the pattern of /jau/, the L2 learners and native speakers of Mandarin showed the same direction in discriminating the contour-level tone pairs versus the level-level tone pairs carried by /se/.

To further investigate the interaction between syllable and group, a seventh logit mixed-effects model, focusing on the L2 learners, was used for the participants' accuracy with Syllable and Group as fixed variables. The results showed a main effect of Syllable, $z(6125) = 3.9$, $p < 0.001$, but did not show a significant Syllable x Group interaction for either the English group ($z < |1|$) or the Mandarin group, $z(6125) = -1.1$, $p = 0.3$. This indicates that listeners found the tones carried by /jau/ harder to discriminate than those carried by /se/, and the L2 learners did not differ from the other two groups in distinguishing tones across the two syllables.

An eighth logit mixed-effects model, focusing on the two monolingual groups, was used for the participants' accuracy with Syllable and Group as fixed variables. The results showed a main effect of Syllable, $z(3977) = 5.4$, $p < 0.001$, but did not show a significant Syllable x Group interaction for the Mandarin group ($z < |1|$). This indicates that listeners found the tones carried by /jau/ harder to discriminate than those carried by /se/, and the Mandarin controls did not differ from the English controls in distinguishing tones across the two syllables.

In summary, the comparison of the listeners' performance on the two syllables indicates that the three groups shared a similar pattern between the two syllables. The only difference was that the L2 learners differed from the Mandarin controls on /jau/ but not on /se/ in terms of contour-level and level-level tone pairs. All groups found the tones carried by /jau/ harder to discriminate than those carried by /se/, and the three groups behaved similarly across the two syllables in their overall performance.

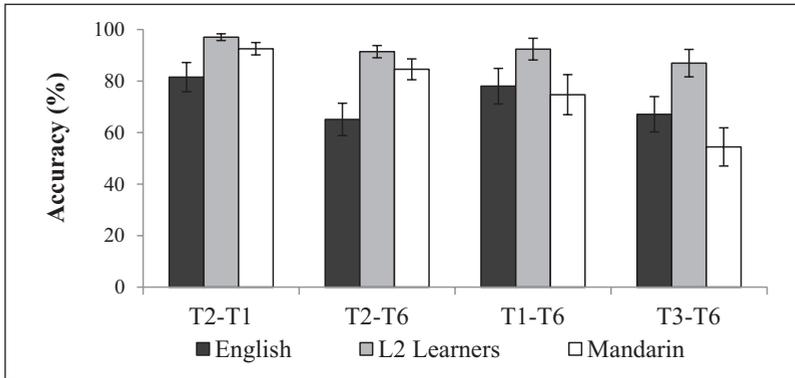


Figure 8. Discrimination accuracy (and standard errors) of the four tone pairs in L3 Cantonese for English controls, L1 English/L2 Mandarin subjects and Mandarin controls.

4.3 The effect of acoustics

Since some tone pairs (i.e., T3–6 and T2–6) may be more difficult to discriminate than others (i.e., T1–6 and T1–2), this study analyzed each tone pair separately for each group, as illustrated in Figure 8.

First, accuracy of the three groups in perceiving the two different contour–level tone pairs (T2–T1 and T2–T6) was analyzed. A ninth logit mixed-effects model was used for the participants' accuracy on the two contour–level tone pairs. The results showed a significant effect of Tone pair, $z(3018) = -3.9, p < 0.001$, but no significant Tone pair \times Group interaction was found. Similarly, a tenth logit mixed-effects model, focusing on the two monolingual groups, showed a significant effect of Tone pair, $z(1844) = -6.5, p < 0.001$, but no significant Tone pair \times Group interaction was found. This indicates that T2–T1 was perceived more accurately than T2–T6 by all three groups, and the groups did not differ in their discrimination of the two contour–level tone pairs.

Second, accuracy of the three groups in perceiving the two different level–level tone pairs (T1–T6 and T3–T6) was analyzed. An eleventh logit mixed-effects model was used for the participants' accuracy for the two level–level tone pairs. The results showed a significant effect of Tone pair, $z(2904) = -3.0, p = 0.003$, but no significant Tone pair \times Group interaction was found. A twelfth logit mixed-effects model, focusing on the two monolingual groups, showed a significant effect of Tone pair, $z(1830) = -4.3, p < 0.001$, as well as a significant Tone pair \times Group interaction, $z(1830) = -2.1, p = 0.04$. The results showed that L2 learners of Mandarin did not differ from the other two groups in their discrimination of the two level–level tone pairs. Although the English and Mandarin controls differed regarding the two level–level tone pairs, a *post-hoc* analysis showed that both the English group, $z(771) = -4.8, p < 0.001$, and the Mandarin group, $z(765) = -8.3, p < 0.001$, showed a significant effect of Tone pair. This indicates that T1–T6 was perceived more accurately than T3–T6 by all the groups.

5 Discussion

This study investigated how linguistic experience and acoustic factors affect perception of L3 tones. Overall, our results showed that both factors had an impact on the perception of L3 tones.

In terms of linguistic experience, this study found that L2 experience did modulate listeners' use of F0 cues in the perception of L3 tones. The results of this study showed that the L2 learners had

a pattern which was different from either that of the English controls or Mandarin controls in their discrimination of Cantonese tones. The L2 learners, as well as the Mandarin controls, were significantly better at discriminating the contour–level tone pairs than the level–level tones. This suggests that L2 experience increased L2 learners' sensitivity to F0 direction in the perception of L3 tones. Additionally, *post-hoc* tests showed that the L2 learners, as well as the English controls, were significantly better than the Mandarin controls at using F0 height to discriminate the level–level tone pairs. This suggests that L1 experience still influenced how L2 learners of Mandarin perceived Cantonese tones. Thus, L2 influence augments, but does not override, L1 influence in the perception of L3 tones.

The effect of L1 experience was not only found for L2 learners of Mandarin, but also for the Mandarin controls and the English controls. Interestingly, the Mandarin controls performed much better on the contour–level tone pairs than the level–level tone pairs. Since F0 direction is used to distinguish contour and level tones, Mandarin controls, that pay most attention to F0 direction, did well in distinguishing the contour–level tone pairs. Additionally, Cantonese T2, a rising tone, and Cantonese T1, a level tone, may be assimilated to the rising T2 and the level T1 in Mandarin, respectively. This could explain why the Mandarin controls had no difficulty discriminating contour tones from level tones. However, the Mandarin controls had significantly lower discrimination accuracy than the other two groups for level–level tone pairs. Therefore, the native tonal system does not always facilitate perception of non-native tones, which is consistent with previous findings (Hao, 2012; So & Best, 2010). Since there is no level tone contrast in Mandarin, subtle differences between level tones might become within-categorical differences for Mandarin controls (i.e., they assimilate the three level Cantonese tones to their single native level tone), resulting in reduced sensitivity to the different level tones in Cantonese. These findings are consistent with L2 speech perception models claiming that perception of non-native tones may be based on the phonetic similarities and the phonological relationships between L1 and L2 tones (Best, 1995; Flege, 1995).

In contrast to the Mandarin controls, the English controls did not distinguish the contour–level tone pairs significantly better than the level–level tone pairs, and they showed a higher accuracy on the level–level tone pairs than the Mandarin controls. Since lexical tone is absent in English, English controls are not very sensitive to F0 direction, which is a linguistic cue to lexical tone. However, they maintain a better sensitivity to the subtle difference between level tones than Mandarin controls (Burnham & Mattock, 2007; Leather, 1987; Stagray & Downs, 1993).

In terms of acoustics, this study found that all listeners showed higher accuracy on acoustically easy pairs than acoustically hard pairs within the level–level as well as contour–level tone pairs. This suggests that acoustic properties of the tonal stimuli affected listeners' perception regardless of their language backgrounds. Because the acoustically easy versus hard pair within each comparison in this study is based on F0 height differences, it would be interesting to include an easy versus hard pair involving contour differences (e.g., big rise–level vs. small rise–level) in a further study.

Moreover, some difference was found between tones carried by /jau/ and /se/. First, it was harder for listeners to discriminate tones carried by /jau/ than by /se/. The accuracy difference between the two syllables may be due to the phonetic variations in tone contours across the two syllables, as illustrated in Figure 4. The level tones carried by /jau/ are closer to each other than those carried by /se/. Second, the L2 learners differed from the Mandarin controls on /jau/ but not on /se/ in terms of contour–level and level–level tone pairs. It might be easier to tease apart groups in terms of linguistic experience when the tones are closer to each other, like those carried by /jau/, than when tones are more separate from each other, like those carried by /se/.

Overall, knowledge of lexical tones in L2 was found to influence the perception of lexical tones in L3. Our findings about the role of L2 in perception of L3 Cantonese tones are in agreement with

studies of L3 morpho-syntax acquisition (e.g., Bardel & Falk, 2007, 2012; Flynn et al., 2004; Leung, 2007). On the one hand, the findings of the L2 influence seem to support all three models of L3 acquisition. While the Cumulative Enhancement Model and the L2 Status Factor Model both predict that L2 will influence perception of Cantonese tones, the Typological Primacy Model also predicts that L2 Mandarin comes into play in this setting, as Mandarin is more typologically similar to Cantonese than English in our design. On the other hand, L3 learners of Cantonese were found to remain sensitive to F0 height to perceive level–level tone pairs under the influence of L1 English besides perceiving L3 contour–level tone pairs under the influence of L2 Mandarin. The findings seem to support the Cumulative Enhancement Model, which assumes that both L1 and L2 can facilitate but not inhibit L3 acquisition. Since the present study did not specifically intend to tease apart different L3 models, it is premature to conclude that the Cumulative Enhancement Model is better than the Typological Proximity Model and the L2 Status Factor Model in accounting for the acquisition and perception of L3 Cantonese tones. To test the extent to which different L3 models can account for L3 acquisition and perception, future studies will need to tease apart several factors (e.g., language status and typology) by including a number of different language pairings.

6 Conclusion

Both linguistic experience and acoustic factors were found to affect the perception of L3 tones in this study. First, the acoustics of the tones affected all listeners' discrimination. Second, in the perception of Cantonese tones, the L2 learners of Mandarin showed a pattern different from either that of the English controls or the Mandarin controls in their discrimination of contour–level and level–level tone pairs. Hence, L2 experience as well as L1 experience was found to modulate the listeners' use of cues in their perception of L3 tones. More importantly, when L2 and L3 are similar in terms of the presence of lexical tone, L2 experience modulates the perception of L3 tones. Therefore, this study contributes to the fields of both L2 and L3 acquisition by providing evidence supporting the influence of L2 on the acquisition of L3 in the speech domain.

Acknowledgements

We thank Dr Peggy Mok for her contribution to an early version of this research. For their insightful comments on this research, we are very grateful to: Drs Robert Fiorentino, Alison Gabriele, Joan Sereno, Annie Tremblay, and Jie Zhang, as well as the student members of LING 850 at the University of Kansas.

Funding

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

Notes

1. L2 learners were recruited from different levels of Chinese classes (first-year: 1 participant; second-year: 9 participants; and third-year: 5 participants). However, an identification test using Mandarin tones demonstrated that all L2 learners showed a high mean identification accuracy (90%; *SD*: 6.1), and identification accuracy was not significantly correlated with level of Chinese class ($p > .1$). Thus, the L2 learners were treated as a single group.
2. The standard deviation of F0 values for each syllable was used to quantify the distribution of the tones carried by each syllable.
3. The difference of average F0 values between two tones was used to quantify the acoustic distance between tones carried by each syllable.

4. Three tokens of the first sound in AB pairs.
5. Three tokens of the second sound in AB pairs.
6. 500 ms was chosen as ISI, because it is considered to encourage a phonetic mode of processing and has been used in previous studies. Moreover, pilot testing indicated that participants found it more comfortable to do the task using this ISI than a longer one (i.e., 1500 ms).

References

- Baayen, R. H. (2008). *Analyzing linguistic data: A practical introduction to statistics*. Cambridge, UK: Cambridge University Press.
- Bardel, C., & Falk, Y. (2007). The role of the second language in third language acquisition: The case of Germanic syntax. *Second Language Research*, 23(4), 459–484.
- Bardel, C., & Falk, Y. (2012). *The L2 status factor and the declarative/procedural distinction*. In J. Cabrelli Amaro, S. Flynn, & J. Rothman (Eds.), *Third language acquisition in adulthood* (pp. 61–78). Amsterdam, The Netherlands: John Benjamins.
- Bauer, R. S., & Benedict, P. K. (1997). *Modern Cantonese phonology*. New York, NY: Mouton de Gruyter.
- Best, C. T. (1995). *A direct realistic view of cross-language speech perception*. In W. Strange (Ed.), *Speech perception and linguistic experience: Issues in cross-language research* (pp. 171–204). Baltimore, MD: York Press.
- Burnham, D., & Francis, E. (1997). *The role of linguistic experience in the perception of Thai tones*. In A. Abramson (Ed.), *Southeast Asian linguistic studies in honour of Vichin Panupong* (pp. 29–47). Bangkok, Thailand: Chulalongkorn University Press.
- Burnham, D., & Mattock, K. (2007). *The perception of tones and phones*. In O.-S. Bohn & M. J. Munro (Eds.), *Language experience in second language speech learning: In honor of James Emil Flege* (pp. 259–280). Amsterdam, The Netherlands: John Benjamins.
- Cabrelli Amaro, J. (2012). *L3 Phonology: An understudied domain*. In J. Cabrelli Amaro, S. Flynn, & J. Rothman (Eds.), *Third language acquisition in adulthood* (pp. 33–60). Amsterdam, The Netherlands: John Benjamins.
- Chao, Y. R. (1947). *Cantonese primer*. Cambridge, UK: Cambridge University Press.
- Chao, Y. R. (1968). *A grammar of spoken Chinese*. Berkeley, CA: University of California Press.
- Falk, Y., & Bardel, C. (2010). The study of the role of the background languages in third language acquisition: The state of the art. *International Review of Applied Linguistics in Language Teaching*, 48(2–3), 185–219.
- Falk, Y., & Bardel, C. (2011). Object pronouns in German L3 syntax: Evidence for the L2 status factor. *Second Language Research*, 27(1), 59–82.
- Flege, J. E. (1995). *Second-language speech learning: Theory, findings, and problems*. In W. Strange (Ed.), *Speech perception and linguistic experience: Issues in cross-language research* (pp. 233–273). Baltimore, MD: York Press.
- Flynn, S., Foley, C., & Vinnitskaya, I. (2004). The cumulative enhancement model for language acquisition: Comparing adults and children's patterns of development in first, second and third language acquisition of relative clauses. *International Journal of Multilingualism*, 1(1), 3–16.
- Fok-Chan, Y. Y. (1974). *A perceptual study of tones in Cantonese*. Hong Kong: Centre of Asian Studies, University of Hong Kong.
- Francis, A. L., & Ciocca, V. (2003). Stimulus presentation order and the perception of lexical tones in Cantonese. *Journal of the Acoustical Society of America*, 114(3), 1611–1621.
- Francis, A. L., Ciocca, V., Ma, L., & Fenn, K. (2008). Perceptual learning of Cantonese lexical tones by tone and non-tone language listeners. *Journal of Phonetics*, 36(2), 268–294.
- Gandour, J. T. (1983). Tone perception in far Eastern languages. *Journal of Phonetics*, 11(2), 149–175.
- Garcia Mayo, M. P. (2012). Cognitive approaches to L3 acquisition. *International Journal of English Studies*, 12(1), 1–20.
- Gass, S. M., & Selinker, L. (Ed.) (1983). *Language transfer in language learning*. Rowley, MA: Newbury House.

- Guion, S. G., & Pederson, E. (2007). *Investigating the role of attention in phonetic learning*. In O.-S. Bohn & M. Munro (Eds.), *Language experience in second language speech learning: In honor of James Emil Flege* (pp. 57–77). Amsterdam, The Netherlands: John Benjamins.
- Hallé, P. A., Chang, Y. C., & Best, C. T. (2004). Identification and discrimination of Mandarin Chinese tones by Mandarin Chinese vs. French listeners. *Journal of Phonetics*, 32(3), 395–421.
- 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.
- Hombert, J. M. (1978). *Consonant types, vowel quality, and tone*. In Fromkin, V. (Ed.), *Tone: A linguistic survey* (pp. 77–111). New York, NY: Academic Press.
- Howie, J. M. (1976). *Acoustical studies of Mandarin vowels and tones*. Cambridge, UK: Cambridge University Press.
- Khouw, E., & Ciocca, V. (2007). Perceptual correlates of Cantonese tones. *Journal of Phonetics*, 35(1), 104–117.
- Ladd, D. R. (2008). *Intonational phonology*. Cambridge, UK: Cambridge University Press.
- Leather, J. (1987). *F0 pattern inference in the perceptual acquisition of second language tone*. In A. James & J. Leather (Eds.), *Sound patterns in second language acquisition* (pp. 59–81). Dordrecht, The Netherlands: Foris Publications.
- Lee, C. Y., Tao, L., & Bond, Z. S. (2008). Identification of acoustically modified Mandarin tones by native listeners. *Journal of Phonetics*, 36(4), 537–563.
- Lee, Y. S., Vakoch, D. A., & Wurm, L. H. (1996). Tone perception in Cantonese and Mandarin: A cross-linguistic comparison. *Journal of Psycholinguistic Research*, 25(5), 527–542.
- Leung, Y.-K. I. (2007). Third language acquisition: Why it is interesting to generative linguists. *Second Language Research*, 23(1), 97–114.
- Li, C. N., & Thompson, S. A. (1989). *Mandarin Chinese: A functional reference grammar*. Berkeley, CA: University of California Press.
- Lin, T., & Wang, W. S.-Y. (1984). The perception of tones. *Journal of Chinese Linguistics*, 2(1), 59–69.
- MacMillan, N. A., & Creelman, C. D. (1991). *Detection theory: A user's guide*. New York: Cambridge University Press.
- Mok, P., Zuo, D., & Wong, P. (2013). Production and perception of a sound change in progress: Tone merging in Hong Kong Cantonese. *Language Variation and Change*, 25(3), 341–370.
- Montrul, S., Dias, R., & Santos, H. (2011). Clitics and object expression in the L3 acquisition of Brazilian Portuguese: Structural similarity matters for transfer. *Second Language Research*, 27(1), 21–58.
- Moore, C. B., & Jongman, A. (1997). Speaker normalization in the perception of Mandarin Chinese tones. *Journal of the Acoustical Society of America*, 102(3), 1864–1877.
- Odling, T. (1989). *Language transfer: Cross-linguistic influence in language learning*. Cambridge, UK: Cambridge University Press.
- Ojima, S., Nakata, H., & Kakigi, R. (2005). An ERP study of second language learning after childhood: Effects of proficiency. *Journal of Cognitive Neuroscience*, 17(8), 1212–1228.
- Peng, G., Zheng, H. Y., Gong, T., Yang, R. X., Kong, J. P., & Wang, W. S.-Y. (2010). The influence of language experience on categorical perception of F0 contour-levels. *Journal of Phonetics*, 38(4), 616–624.
- Perception Research Systems (2007). *Paradigm stimulus presentation*. Retrieved from <http://www.paradigm-experiments.com>
- Qin, Z., & Mok, P. (2013). Discrimination of Cantonese tones by speakers of tone and non-tone languages. *Kansas Working Papers in Linguistics*, 34, 1–25.
- Rast, R. (2010). The use of prior linguistic knowledge in the early stages of L3 acquisition. *International Review of Applied Linguistics in Language Teaching*, 48(2–3), 159–183.
- Rothman, J. (2010). On the typological economy of syntactic transfer: Word order and relative clause high/low attachment preference in L3 Brazilian Portuguese. *International Review of Applied Linguistics in Teaching (IRAL)*, 48(2–3), 245–273.
- Rothman, J. (2011). L3 syntactic transfer selectivity and typological determinacy: The typological primacy model. *Second Language Research*, 27(1), 107–27.

- Rothman, J. (2013). *Cognitive economy, non-redundancy and typological primacy in L3 Acquisition: Evidence from initial stages of L3 Romance*. In S. Baauw, F. Dirjkonigen, & M. Pinto (Eds.), *Romance languages and linguistic theory 2011* (pp. 217–247). Amsterdam, The Netherlands & Philadelphia, PA: John Benjamins.
- Rothman, J. (2015). Linguistic and cognitive motivations for the typological primacy model of third language (L3) transfer: Considering the role of timing of acquisition and proficiency in the previous languages. *Bilingualism: Language and Cognition*, 18(2), 179–190.
- Rothman, J., & Cabrelli Amaro, J. (2010). What variables condition syntactic transfer?: A look at the L3 initial state. *Second Language Research*, 26(2), 189–218.
- Sanz, C., Park, H. I., & Lado, B. (2015). A functional approach to cross-linguistic influence in ab initio L3 acquisition. *Bilingualism: Language & Cognition*, 18(2), 236–251.
- So, C. K., & Best, C. T. (2010). Cross-language perception of non-native tonal contrasts: Effects of native phonological and phonetic influences. *Language and Speech*, 53(2), 273–293.
- Stagray, J. R., & Downs, D. (1993). Differential sensitivity for frequency among listeners of a tone and a nontone language. *Journal of Chinese Linguistics*, 21(1), 143–163.
- Steinhauer, K., White, E. J., & Drury, J. (2009). Temporal dynamics of late second language acquisition: Evidence from event-related brain potentials. *Second Language Research*, 25(1), 13–41.
- Vance, T. J. (1976). An experimental investigation of tone and intonation in Cantonese. *Phonetica*, 33(5), 368–392.
- Wang, W. S.-Y. (1976). Language change. *Annals of the New York Academy of Sciences*, 28, 61–72.
- Wang, Y., Behne, D. M., Jongman, A., & Sereno, J. (2004). The role of linguistic experience in the hemispheric processing of lexical tones. *Applied Psycholinguistics*, 25(3), 449–466.
- Wang, Y., Jongman, A., & Sereno, J. (2001). Dichotic perception of linguistic tones by Mandarin and American listeners. *Brain and Language*, 78(3), 332–348.
- Wang, Y., Jongman, A., & Sereno, J. (2003). Acoustic and perceptual evaluation of Mandarin tone production before and after perceptual training. *Journal of the Acoustical Society of America*, 113(2), 1033–1044.
- Wang, Y., Spence, M., Jongman, A., & Sereno, J. (1999). Training American listeners to perceive Mandarin tones. *Journal of the Acoustical Society of America*, 106(6), 3649–3658.
- Wayland, R. P., & Guion, S. G. (2004). Training English and Chinese listeners to perceive Thai tones: A preliminary report. *Language Learning*, 54(4), 681–712.
- Wrembel, M. (2012). *Foreign accentedness in third language acquisition: The case of L3 English*. In J. Cabrelli Amaro, S. Flynn, & J. Rothman (Eds.), *Third language acquisition in adulthood* (pp. 281–309). Amsterdam, The Netherlands: John Benjamins.
- Wrembel, M., Ulrike, G., & Grit, M. (2010). Phonetics/phonology in third language acquisition—Introduction. *International Journal of Multilingualism (Special Issue on L3 phonology)*, 7(1), 1–3.
- Wu, X., Munro, J. M., & Wang, Y. (2014). Tone assimilation by Mandarin and Thai listeners with and without L2 experience. *Journal of Phonetics*, 46(1), 86–100.
- Xu, Y. (2013). *ProsodyPro—A tool for large-scale systematic prosody analysis*. In Proceedings of Tools and Resources for the Analysis of Speech Prosody, Aix-en-Provence, France, 30 August 2013 (pp. 7–10). Aix-en-Provence, France: Laboratoire Parole et Langage.
- Xu, Y. S., Gandour, J., & Francis, A. L. (2006). Effects of language experience and stimulus complexity on the categorical perception of F0 direction. *Journal of the Acoustical Society of America*, 120(2), 1063–1074.