

The effect of second-language learning experience on Korean listeners' use of pitch cues in the perception of Cantonese tones

Second Language Research

2026, Vol. 42(3) 529–545

© The Author(s) 2024



Article reuse guidelines:

sagepub.com/journals-permissions

DOI: 10.1177/02676583241244604

journals.sagepub.com/home/slr

**Quentin Zhen Qin** 

The Hong Kong University of Science and Technology, Hong Kong

Sang-Im Lee-Kim

Hanyang University, South Korea

Haifeng Qi

Shanghai International Studies University, China

Abstract

Past studies have found that the linguistic experience of previously-acquired languages, such as one's native-language (L1) and second-language (L2) learning experience, modulates the perception of novel sounds from a third language (L3). Lexical tone in L3 is a good case for testing the influence of L1 or L2, as listeners with varying language backgrounds may use different pitch cues (pitch contour or height) in tone perception. The present study focuses on L2 learners of Mandarin whose L1 variety is either Seoul Korean (SK), a non-tonal stressless language, or Gyeongsang Korean (GK), a tonal pitch-accent language. Intermediate-to-advanced SK-speaking and GK-speaking L2 learners of Mandarin were recruited as target groups, and naive listeners of respective L1 varieties were recruited as control groups. The participants completed an AX forced-choice tone discrimination task. Four Cantonese tones, one rising tone and three level tones, were used. Contour–level and level–level tonal contrasts were target tone pairs, allowing for testing the primary use of pitch contour and pitch height, respectively. The results showed that the two groups of naive listeners had greater accuracy in discriminating level–level than contour–level tonal contrasts. In contrast, L2 learners, independent of their L1 varieties, showed higher accuracy in discriminating contour–level than level–level tonal contrasts. The L2 learners'

Corresponding author:

Quentin Zhen Qin, Division of Humanities, The Hong Kong University of Science and Technology, Room 3368, Academic Building, Division of Humanities, Clear Water Bay, Kowloon, Hong Kong

Email: hmqin@ust.hk

perceptual pattern is consistent with Mandarin listeners, as reported in previous work. Taken together, the findings provide evidence for a possible developmental change in which Korean-speaking L2 learners might have a perceptual cue shift from pitch height to pitch contour through their L2 experience in Mandarin. The findings about the role of L2 proficiency in Mandarin further supported the effect of L2 experience on learners' increased use of pitch contour.

Keywords

Cantonese tones, pitch cues, L2 proficiency, native-language (L1) variety, second-language (L2)

1 Introduction

Previous studies have found that the linguistic experience of previously-acquired languages – for instance, one's native-language (L1) and second-language (L2) learning experience – modulates the perception of novel sounds in an unfamiliar language (i.e. a third language, L3). It remains unclear whether L1 or L2 (or both) is the source of transfer at the early stage of L3 acquisition (Falk and Bardel, 2011; Rothman, 2011; Westergaard, 2021). Past studies of L3 acquisition focused mainly on lexicon and morphosyntax, but a growing body of literature has investigated the influence of linguistic experience (i.e. L1 or L2) on the acquisition of L3 segments (Amengual, 2021; Chen and Han, 2019; Luo et al., 2020; Wrembel et al., 2019) and L3 suprasegmentals (lexical tone: Chan and Chang, 2019; Niu and Mok, 2022; pitch accent: Wiener and Goss, 2019).

Lexical tone, with its contrasts cued by different pitch dimensions, is a good case for testing the influence of linguistic experiences at the early stage of L3 tone acquisition (e.g. Chan and Chang, 2019). Pitch is a primary cue to contrast tones, but different pitch dimensions are used to contrast different types of tones, and ultimately, both need to be learned: pitch contour (different tone shapes; rising vs. falling tones) and pitch height (difference in height; high vs. low tones). For instance, Cantonese has a rich set of tonal categories, including both level and contour tones (/si 55¹/ 'silk' (Tone 1 (T1)), /si 25/ 'history' (T2), /si 33/ 'to try' (T3), /si 21/ 'time' (T4), /si 23/ 'city' (T5), and /si 22/ 'matter' (T6)) (Matthews and Yip, 2011), and learners of Cantonese would have to develop perceptual sensitivity to both types of pitch cues.

The literature on non-native speech perception finds that acoustic cues are perceptually weighted as a function of their informativeness for cueing L1 sound contrasts (e.g. stops). For instance, English listeners attended primarily to voice onset time (VOT), which is considered the most salient cue to the English stop distinction, in their perception of non-native Korean stops that are cued by F0, VOT, and others (Francis and Nusbaum, 2002; Holt and Lotto, 2006). In the suprasegmental domain, listeners with different language backgrounds are often attuned to different pitch cues to perceive lexical tones (Gandour, 1983). For instance, Mandarin listeners who encode pitch variations lexically in their native contour-tone system (/pa 55/ 'eight' T1, /pa 35/ 'to pull out' T2, /pa 214/ 'to hold' T3, /pa 51/ 'father' T4) mainly rely on pitch contour differences (Yip, 2002). In contrast, English listeners with no prior tone experience are tuned in to subtle pitch height variations and may use pitch height more than pitch contour in perceiving non-native tones (Francis et al., 2008; Gandour, 1981; Jongman et al., 2017). Beyond the

limitations posed by one's L1, L2 learners may demonstrate changes in using pitch cues through their experience of learning an L2 (e.g. Dmitrieva, 2019). These findings give rise to an idea that could be called a 'cue-weighting theory' that suggests a possibility of cue shifting in the L2 or L3 perception of non-native sounds (Francis and Nusbaum, 2002; Holt and Lotto, 2006).

To specifically test the influence of L1 and L2 on listeners' use of pitch cues, Qin and Jongman (2016) compared listeners with differing linguistic experiences for the perception of Cantonese tone pairs that have contrasting pitch contour and pitch height. Consistent with previous findings, Mandarin listeners performed well in distinguishing Cantonese contour tones from level tones (e.g. T2 /25/ – T6 /22/) in an AX discrimination task, but they struggled in discriminating level–level tonal contrasts (e.g. T3 /33/ – T6 /22/). The results suggested that they use pitch contour more than pitch height (for further evidence, see Jongman et al., 2017). Despite overall low accuracy, English listeners who were naive to lexical tones did not show greater difficulty discriminating level–level tonal contrasts than contour–level contrasts, a pattern opposite to Mandarin listeners. Mandarin listeners' reduced sensitivity to pitch height differences was attributed to the lack of level tonal contrasts in their native contour–tone system. That is, differences in pitch height are not used contrastively in their native tonal inventory. Crucially, the English-speaking L2 learners of Mandarin showed a pattern similar to Mandarin listeners, namely a greater sensitivity to pitch contour than pitch height. The findings suggest the *predominant influence of the L2* on English listeners' use of pitch cues in their L3 tone perception.

However, another interpretation of the results would be that the functional weight of pitch cues for signaling lexical distinctions (i.e. stress contrasts; *OFFset* vs. *offSET*) is more limited than vowel quality cues in English (Cutler, 1986), so the L1 influence regarding the use of pitch cues (height differences) was not clearly demonstrated. The present study replicates Qin and Jongman (2016) to disambiguate the accounts of (limited) L1 functional weight of pitch and (predominant) L2 influence by testing speakers of a language variety that fully employs pitch cues for lexical contrasts. To control typological differences in word prosody, we chose non-tonal and tonal varieties of the same language: Seoul Korean (SK) vs. Gyeongsang Korean (GK). SK is neither tonal nor stressed and does not use pitch to mark lexical prosody (Jun, 2010). In contrast, GK is a lexical pitch accent language and uses pitch differences to realize lexically contrastive words (Jun et al., 2006; Lee and Jongman, 2015; Lee et al., 2016; Ramsey, 1975). Recent studies provided preliminary evidence supporting GK listeners' overall greater sensitivity to pitch cues (over the duration and intensity cues) than SK listeners in their L1 pitch accents (e.g. /u^H.li^L/ 'cage' vs. /u^L.li^H/ 'us') and L2 lexical stress processing (Kim and Tremblay, 2021; Lee et al., 2019). If an advantage of encoding pitch lexically in an L1 variety is borne out, we would expect GK listeners, regardless of L2 learning experience, to be more sensitive to pitch cues in discriminating non-native tonal contrasts than their SK counterparts.

This study primarily compares two groups of Korean learners (i.e. SK- or GK-speaking L2 learners of Mandarin), along with their naive counterparts, to test the relative importance of L1 variety (non-tonal or tonal Korean) vs. L2 (Mandarin) on their use of pitch cues in the perception of L3 (Cantonese) tonal contrasts. If the influence of the L1 variety

is predominant, the two groups of Korean-speaking learners would be expected to show different performances, with GK-speaking-L2 learners patterning more like Mandarin listeners by virtue of the contrastive pitch cues in their L1 variety. If the L2 learning experience is more integral in the way L3 tones are processed, both groups would be expected to show greater sensitivity to pitch contour than to pitch height.

In addition to the broad effect of L2 experience, we also explore the potential effect of L2 proficiency on L3 tone perception, a topic that is as yet understudied in the literature (see Niu and Mok, 2022). For instance, Qin and Jongman (2016) did not include objective measures of learners' L2 proficiency in Mandarin and thus did not test the potential (gradient) effect of L2 proficiency in further support of the L2 influence. Based on previous findings of the positive proficiency effect on L2 tone perception (Han and Tsukada, 2020; Tsukada and Han, 2019), it would be expected that Korean-speaking L2 learners' use of pitch cues will pattern more closely with Mandarin listeners over the course of L2 development. That is, they will show greater use of pitch contour than pitch height as a function of their increasing proficiency in Mandarin.

II Methods

I Participants

The target groups include 20 SK-speaking L2 learners of Mandarin (5 males) in their twenties (mean=24.7 years; SD=4.5) and 15 GK-speaking L2 learners of Mandarin (6 males) in their twenties (mean=24.1; SD=2.6). The control groups include 15 SK-speaking participants (9 males; mean of age=23.8 years, SD=3.3) and 15 GK-speaking participants (9 males; mean of age=22.8, SD=3.5) with no prior experience with Mandarin or any tone languages (i.e. naive participants). The SK speakers were born and raised in Seoul with parents who spoke SK, and the GK speakers were born and raised in the Gyeongsang region of South Korea with parents who spoke GK. The SK- and GK-speaking L2 learners of Mandarin were recruited and tested at a university in Shanghai, China,² and the SK- and GK-speaking naive participants were recruited and tested at a university in Seoul, South Korea.³ None of the participants reported a history of speech, hearing, or other language impairments. Each participant received monetary compensation for participation.

The SK- and GK-speaking learners were intermediate-to-advanced L2 learners of Mandarin, indicated by gaining a pass in HSK-4 (Hanyu Shuiping Kaoshi Level 4)⁴ and had been immersed in Chinese-speaking countries (see LOR in Table 1). In addition, the Mandarin proficiency levels of the SK- and GK-speaking L2 learners were matched based on the information obtained in a language background questionnaire and their proficiency scores. The proficiency tests include a Mandarin lexical decision task adapted from LexTALE (Chan and Chang, 2018; Lemhöfer and Broersma, 2012) and a Mandarin cloze (i.e. fill-in-the-blank) test (Yuan, 2010) to provide a comprehensive characterization of learners' global proficiency in L2 Chinese (Qin et al., 2019). The Mandarin LexTALE included 120 items, 80 of which were words. The cloze test included a total of 40 missing words. The L2 learners' Mandarin learning experience and proficiency information (mean scores converted into percentages) are summarized in Table 1. The output

Table 1. Two groups of second language (L2) learners' Mandarin learning experience and proficiency information.

First language (L1) group	AOE (year)	Mandarin instruction (year)	LOR (year)	Lexical decision test (%)	Cloze Test (%)
Seoul Korean (n=20)	20.8 (3.3)	2.6 (1.6)	2.0 (1.7)	72.3 (6.0)	72.4 (10.7)
Gyeongsang Korean (n=15)	19.1 (3.7)	3.0 (2.0)	3.0 (2.0)	73.4 (9.9)	69.0 (14.5)
BF_{10}	0.73	0.41	0.86	0.34	

Notes. Mean (standard deviation). AOE=age of first exposure to Mandarin. LOR=length of residence in a Chinese-speaking country. BF_{10} is Bayes factor calculated using the BayesFactor package in R, and the BF_{10} for proficiency tests was yielded for the average proficiency score across the two tests.

of Bayes factor tests (i.e. $BF_{10} < 1$) provided relative evidence supporting that there was no difference between the two L2 groups for each variable reported in Table 1 (Wagenmakers et al., 2018).

2 Stimuli and materials

The perception experiment adopted a tone discrimination task. As illustrated in Figure 1, the tone stimuli from Qin and Jongman (2016), a contour tone (i.e. T2 /25/, a rising tone) and three level tones (i.e. T1 /55/ high-level, T3 /33/ mid-level, T6 /22/ low-level) in Cantonese, were (re-)used for the current task. The other two tones, T4 /21/ (a low falling tone) and T5 /23/ (a low rising tone), were excluded from our study primarily because they are confused with other tones (e.g. T6 and T2, respectively) even by native listeners (e.g. Mok et al., 2013). In addition, T4 is often cued by the non-pitch information of creaky phonation (Yu and Lam, 2014; Zhang and Kirby, 2020) and may not be ideal for examining listeners' use of pitch cues in tone perception.

The tone stimuli from our previous study, produced by a female native speaker of Hong Kong Cantonese, were reused for replication purposes. Each tone was carried by two syllables, [jɐu] and [se], to counterbalance the voicing of initial consonants and increase stimuli variability. The four tones with the two syllables all form real Cantonese words, but they are not perceived as real Mandarin words given the phonotactic constraints in Mandarin (e.g. [se] is illegal in Mandarin). The three tokens of each tone carried by each syllable were chosen, giving rise to 24 tokens in total (4 tones × 2 syllables × 3 repetitions). The duration of tone stimuli was normalized at a mean value of length (i.e. 345 ms for [jɐu] and 488 ms for [se]). The intensity was normalized at 70 dB. None of the tokens was signaled by phonation cues such as creaky or breathy voice.

The four tones were paired into four tone pairs in a discrimination task: two contour–level tone pairs (i.e. T2 /25/ – T1 /55/, T2 /25/ – T6 /22/⁵) primarily contrasting in pitch contour, and two level–level tone pairs (i.e. T1 /55/ – T6 /22/, T3 /33/ – T6 /22/⁶) primarily contrasting in pitch height. The contour–level tone pairs and the level–level tone pairs allowed us to assess listeners' relative use of the pitch contour and height cues, respectively.

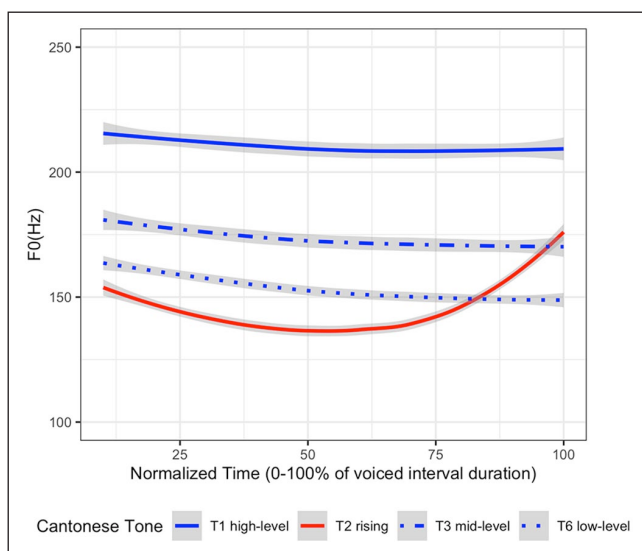


Figure 1. Time-normalized pitch values of a contour–rising tone (T2; e.g. /se 25/) in red, and three level tones (T1, T3, T6; e.g. /se 55/, /se 33/, /se 22/) in blue.

Notes. The pitch curve for each tone represents the smoothing function fit to acoustic data for the tone (i.e. method = 'loess' in R). Shaded areas represent 95% confidence intervals.

3 Procedure

The participants completed an AX forced-choice tone discrimination task. During the task, they were told they would hear pairs of sounds from a language with which they were unfamiliar. They were instructed (in written English, a neutral language understandable to all the participants) to press one of two buttons labeled 'same' and 'different' on the keyboard to assess whether the two tones in each pair were the 'same' or 'different' as accurately as possible. The inter-stimulus interval (ISI) was 500 ms.⁷ The time-out time was 3,000 ms. No feedback was given. The trials consisted of 144 AB pairs (4 pairs $\times$ 2 orders $\times$ 2 syllables $\times$ 3 tokens^{1st tone} $\times$ 3 tokens^{2nd tone}) and 144 AA pairs (4 pairs $\times$ 2 syllables $\times$ 3 tokens $\times$ 3 tokens $\times$ 2 repetitions) to counterbalance the two types of tone pairs. The trials were presented randomly in a block, with the presentation order of each AB pair counterbalanced throughout the experiment. Before the experiment, the participants familiarized themselves with 12 practice trials. The task was conducted using Paradigm software (Perception Research Systems; retrieved from <http://www.paradigmexperiments.com>). It took approximately 25–30 minutes to complete this task.

4 Statistical analysis

Data files, along with analysis scripts, are made publicly available at OSF (<https://osf.io/e2sdx/>). All the participants completed the perception task. Thirty-five missing responses were excluded from the analysis, which resulted in a loss of 0.18% of the data. The

proportion of correct responses (i.e. accuracy) of AA pairs was on the boundary across the four groups (mean=0.97; a range from 0.96 to 0.98), indicating a possible ceiling effect. Therefore, only the accuracy of AB pairs (i.e. 9,335 observations) was submitted for statistical analysis.⁸

Mixed-effects logistic regression models were performed on the participants' response accuracy (1=correct, 0=incorrect) of tonal contrasts. The models were fitted in R (R Core Team, 2022) using the lme4 package (Bates et al., 2015). The model diagnosis was conducted for each model with the DHARMA package (Hartig, 2020) by observing the residual QQ plots. We also followed a maximum Variance Inflation Factor (VIF) threshold of 5 (O'Brien, 2007) when building the models. None of the models reported in the study violated distributional assumptions and showed multicollinearity.

Two sets of models, with categorical predictors deviation coded to test for the main effect, were built to focus on the sensitivity to tonal contrasts by all the listeners and only L2 learners, respectively. In the first set of models, Learning experience (2 levels: naive vs. learners; coding: -0.5, 0.5), L1 variety (2 levels: SK vs. GK; coding: -0.5, 0.5), and Tonal contrast (2 levels: contour vs. height; coding: -0.5, 0.5) were entered as fixed effects. The interaction effects between fixed factors were also included to address the question of the combined effect of L1 varieties and L2 experience on the way different tonal contrasts are processed. Analyses yielding significant interactions between tonal contrasts and listener groups (L1 variety or L2 experience) were followed up by subsequent modeling with the same random structure conducted separately on each group. The models included random intercepts for participants and trials and the random slope of Tonal contrast for the participants.

The second set of models only included data from L2 learners to test the effect of L2 proficiency. A composite proficiency score (i.e. average scores of the two proficiency tests to avoid the collinearity issue), together with the interaction with other fixed effects (i.e. L1 variety and Tonal contrast), was entered as a continuous variable (and was centered prior to the analysis) to test whether Korean listeners' discrimination of tonal contrasts changes as their proficiency in Mandarin increases. Analyses yielding significant interactions among tonal contrasts, learner groups, and proficiency scores were followed-up by subsequent modeling with the same random structure conducted separately on each learner group. The models included random intercepts for participants and trials, and the random slope of Proficiency scores for the participants. All models followed the backwards stepping procedure to determine a maximal random effects structure that was best justified by the data (Matuschek et al., 2017). Only the results of the best-fit model after the backfitting procedure are reported, with *p* values calculated using the lmerTest package in R (Kuznetsova et al., 2018).

III Results

I Overall results

The first set of models focuses on the overall accuracy of different tonal contrasts (i.e. Contrast: contour and height) by four groups of participants with different L2 learning experience (i.e. Experience: naive participants and learners) and L1 varieties (i.e. L1: SK and GK).

Table 2. Results of the mixed-effects logistic regression model on all the listeners’ trial-level accuracy in the tone discrimination task.

Effects	Est.	SE	z	p
(Intercept)	2.06	0.15	13.72	< .001
Experience	0.16	0.24	0.68	.49
First language (L1)	−0.07	0.30	−0.24	.81
Contrast	−0.08	0.12	−0.68	.50
Experience × L1	−0.76	0.47	−1.61	.11
Experience × Contrast	−0.73	0.20	−3.62	< .001
L1 × Contrast	−0.14	0.21	−0.63	.53
Experience × L1 × Contrast	−0.10	0.40	−0.24	.81

Notes. A = .05; significant results are in bold; lmer model: accuracy ~ experience × L1 × contrast + (1 + contrast|participant) + (1|trial).

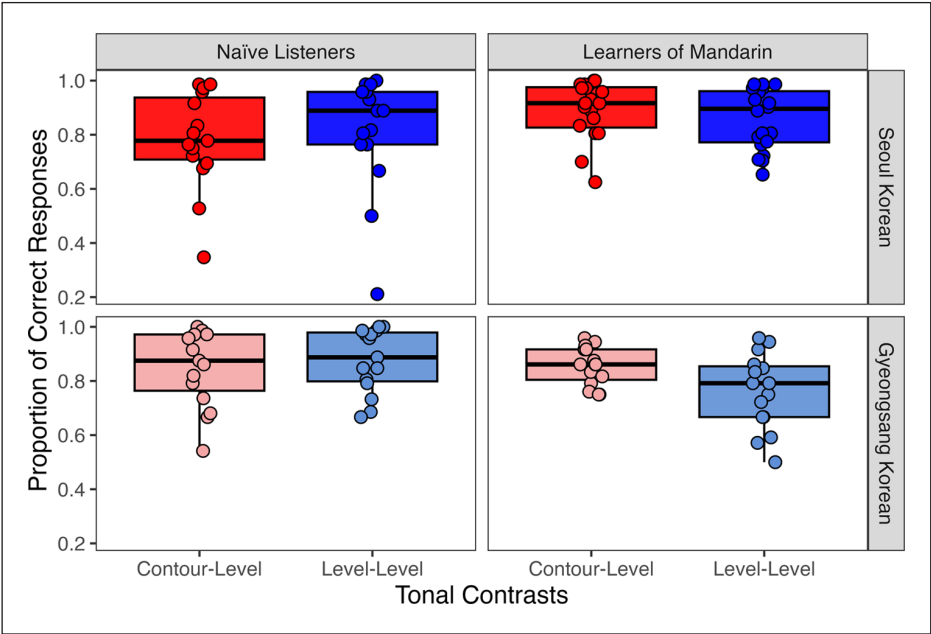


Figure 2. Discrimination accuracy of Cantonese contour-level tonal contrasts (red) and level-level tonal contrasts (blue) by Seoul Korean (SK) speaking (top) and Gyeongsang Korean (GK) speaking (bottom) naive listeners (left) and second language (L2) learners of Mandarin (right).

The results of the model are summarized in Table 2. The model only revealed a significant interaction between Learning experience and Tonal contrast.⁹ This is reflected in Figure 2 wherein the relative accuracy of contour-level and level-level tonal contrasts was reversed between the naive listeners and L2 learners of Mandarin.

Table 3. Results of the mixed-effects logistic regression model on the trial-level accuracy of L2 learners with different proficiency scores in the tone discrimination task.

Effects	Est.	SE	z	p
(Intercept)	1.93	0.17	11.02	<.001
Proficiency	1.84	1.42	1.30	.19
First language (L1)	-0.72	0.32	-2.20	.03
Contrast	-0.51	0.08	-6.19	<.001
Proficiency × L1	-3.43	2.49	-1.38	.17
Proficiency × Contrast	-0.30	0.95	-0.31	.75
L1 × Contrast	-0.34	0.17	-2.06	.04
Proficiency × L1 × Contrast	-4.46	1.90	-2.36	.02

Notes. A=.05; significant results are in bold; lmer model: accuracy~ proficiency × L1 × contrast + (1 + proficiency|participant) + (1|trial).

To better understand the source of the two-way interaction, follow-up models were run on the response accuracy of naive listeners and L2 learners, respectively. The model of naive participants revealed a significant main effect of Tonal contrast ($\beta=0.37$, $SE=0.19$, $z=1.96$, $p=.05$), with higher accuracy for the level–level tonal contrasts than for the contour–level ones. The model of L2 learners also revealed a significant main effect of Tonal contrast ($\beta=-0.47$, $SE=0.15$, $z=-3.21$, $p=.001$), but in a reversed manner, namely higher accuracy for the contour–level tonal contrasts than for the level–level ones. The findings implied that naive listeners, regardless of their L1 varieties (i.e. a lack of prosodic transfer from L1), might have an advantage of discriminating level–level tones over contour–level tones. In contrast, L2 learners, independent of their L1 varieties, might increase their use of pitch contour to distinguish a rising tone from level tones. The L2 learners’ discrimination of tonal contrasts is consistent with Mandarin listeners’ patterns, as reported in previous studies (e.g. Francis et al., 2008).

2 The results of L2 proficiency

To examine whether L2 learners’ discrimination of tonal contrasts changes as their proficiency in Mandarin increases, the second set of models was created using the accuracy of L2 learners, with Proficiency scores (i.e. Proficiency; mean proficiency score by averaging the last two columns in Table 1), L1 variety, and Tonal contrast as fixed effects.

The results of the model are summarized in Table 3. The model fit revealed the main effects of Tonal contrast and L1 variety, and a significant interaction between these two factors. The significant effect of Tonal contrast, again, suggests the L2 learners’ higher accuracy for the contour–level tonal contrasts for the level–level ones. While the effect of L1 variety indicates the SK-speaking L2 learners’ higher sensitivity to tonal contrasts, the significant interaction between L1 variety and Tonal contrast suggests that the two groups differed in their relative sensitivity to the types of tonal contrasts. Follow-up models (with random intercepts for participants and trials, and the random slope of Proficiency scores for the participants) conducted separately on each tonal

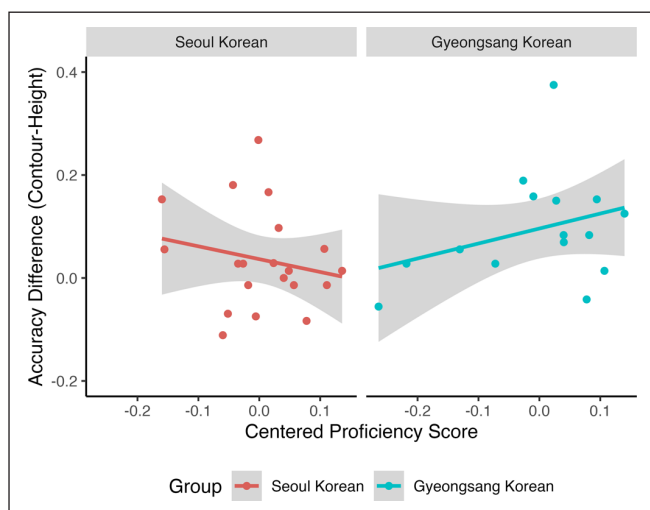


Figure 3. The accuracy (i.e. the proportion of correct responses to AB pairs) difference between contour-level and level-level tonal contrasts (i.e. Contour – Height; contrast coded: $-0.5, 0.5$ in the model) among individual learners with different proficiency scores (centered proportion of correct responses in two proficiency tests) from the Seoul Korean (SK) speaking second language (L2) group (left; red) and the Gyeongsang Korean (GK) speaking L2 group (right; green).

Notes. Shaded areas indicate 95% confidence intervals.

contrast confirmed that while the SK-speaking L2 learners had higher accuracy in discriminating the level-level tonal contrasts than the GK-speaking L2 learners ($\beta = -0.77$, $SE = 0.33$, $z = -2.32$, $p = .02$), the two L2 groups did not significantly differ in their accuracy in discriminating the contour-level ones ($\beta = -0.62$, $SE = 0.34$, $z = -1.82$, $p = .07$).

Importantly, the model revealed a three-way interaction between Proficiency scores, Tonal contrast, and L1 variety. This three-way interaction is reflected in Figure 3 in which a positive relationship between proficiency scores and accuracy difference (i.e. response accuracy of contour-level contrasts minus level-level contrasts to AB pairs) is observed for the GK-speaking L2 learners, but not for the SK-speaking L2 learners. To better understand the three-way interaction, separate models were run on the response accuracy data of SK-speaking L2 learners and GK-speaking L2 learners, respectively. The model of SK-speaking L2 learners did not show a significant interaction between Proficiency scores and Tonal contrast ($\beta = 1.95$, $SE = 1.56$, $z = 1.25$, $p = .21$). In contrast, the model of GK-speaking L2 learners showed a significant interaction between the two ($\beta = -2.45$, $SE = 1.05$, $z = -2.34$, $p = .02$). These results suggest that the accuracy difference between contour-level tonal contrasts (primarily cued by pitch contour) and level-level tonal contrasts (cued by pitch height), as illustrated in Figure 3, became larger as proficiency in Mandarin increased among the GK-speaking L2 learners, but not among the SK-speaking L2 learners.

IV Discussion and conclusions

By replicating the tone discrimination experiment from Qin and Jongman (2016), the present study tested the effect of linguistic experience (L1 and L2) on the use of pitch contour and pitch height cues by naive listeners and L2 learners of Mandarin, whose L1 is either a non-tonal variety (SK) or a tonal variety (GK). The results showed that naive Korean listeners had a greater sensitivity to pitch height than to pitch contour. In contrast, Korean-speaking L2 learners of Mandarin were comparable to native Mandarin listeners, showing greater sensitivity to pitch contour than to pitch height, under the influence of their L2 learning experience. The results of the GK-speaking L2 learners further suggested a significant role of L2 proficiency in the perception of L3 tones. The GK-speaking learners were fine-tuned more toward pitch contour (and less toward pitch height) as their proficiency in Mandarin increased. These results have theoretical implications for L3 tone perception.

First, the overall results provide preliminary evidence for a potential developmental change in which Korean-speaking L2 learners might have a *perceptual (pitch) cue shift* from pitch height to pitch contour through their L2 experience in Mandarin. This finding is consistent with the pattern observed for English-speaking L2 learners in Qin and Jongman (2016). As proposed by the so-called ‘cue-weighting theory’ of speech perception (Francis and Nusbaum, 2002; Holt and Lotto, 2006), listeners with various L1 backgrounds, such as (naive) Korean listeners in this study and Mandarin listeners in the previous study, were found to perceive Cantonese tonal contrasts differently, potentially due to their relative weighting of the pitch cues in L1. Our findings suggest that, like Mandarin listeners (and unlike naive Korean listeners), Korean-speaking L2 learners of Mandarin might have assigned greater weight to pitch contour in perceiving L2-Mandarin tonal contrasts. On the other hand, their sensitivity might have been reduced for pitch height that did not signal L2-Mandarin tonal contrasts (see similar findings of segment learning in Dmitrieva, 2019). Since there is no level tonal contrast in Mandarin, subtle differences in pitch height might be taken as secondary cues for L2 learners, as in the case of Mandarin listeners, resulting in their reduced sensitivity to Cantonese level tones (Francis et al., 2008; Jongman et al., 2017). The possibility of perceptual cue shift was corroborated across the two L2 groups with differing L1 varieties, suggesting the impact of the L2 learning experience on learners’ use of pitch cues in the perception of L3 tones.

The role of L2 proficiency observed from the data of L2 learners are consistent with previous findings of L2 tone perception (Han and Tsukada, 2020; Tsukada and Han, 2019), and may provide further support for the effect of L2 learning experience. Specifically, the greater sensitivity to contour-level tonal contrasts than to level-level tonal contrasts emerged more clearly as proficiency in Mandarin increased among the GK-speaking L2 learners (see similar findings in Niu and Mok, 2022). GK-speaking L2 learners with higher proficiency presented more like Mandarin listeners than those with lower proficiency. However, this proficiency-dependent sensitivity to novel L3 tones was absent for the SK-speaking L2 learner. One possible explanation for these results is that a small range of proficiency scores for the SK-speaking L2 learners did not allow the proficiency effect to emerge with a limited sample size (see the lesser degree of proficiency variability in Figure 3). Although the Mandarin proficiency levels of the SK- and

GK-speaking L2 learners were roughly matched, the SK-speaking L2 learners exhibited smaller variances in Mandarin learning experience and proficiency scores than the GK-speaking L2 learners (see Section C of supplemental material for exploratory analyses regarding variables of learning experience).

This study did not find conclusive evidence supporting an effect of L1 varieties on L2 learners' performance, except for the GK-speaking L2 learners being greatly desensitized to pitch height.¹⁰ The lack of the L1 variety effect in this study may appear to stand at odds with the findings of recent studies reporting L1 prosodic transfer. In Kim and Tremblay (2021), SK-speaking and GK-speaking L2 learners of English completed sequence-recall tasks for GK pitch accent patterns and English stress patterns. The results showed that the GK-speaking learners outperformed the SK-speaking learners in the use of English stress patterns in conjunction with their native pitch accent pattern, indicating an L1 transfer to L2 prosodic processing. One possibility is that learners were able to transfer L1 pitch cues to L2 prosodic processing when L1 (GK) pitch accents and L2 (English) stress contrasts share a prosodic domain (i.e. relative pitch differences that spread over multiple syllables) as in Kim and Tremblay (2021). However, pitch-cue transfer of L1 pitch accents may be limited in the case of L3 tone perception, in which pitch is systematically contrasted on a monosyllable (Best, 2019; Schaefer and Darcy, 2014). This could be a potential reason why the effect of L1 influence, the advantage for GK listeners over SK listeners, did not emerge in L3 tone perception. The difference in the experimental paradigms is an alternative explanation. The present study adopted an AX forced-choice discrimination task to examine lexical tone perception. In contrast to the sequence-recall task, which imposes significant memory demands on listeners and allows them to tap into listeners' category encoding (e.g. lexical stress) at a phonological/lexical level (Dupoux et al., 2001), the discrimination task with a short ISI (500 ms) often taps into listeners' sensitivity to pitch cues at an acoustic/phonetic level (Chen et al., 2023; Wayland and Li, 2008). Since SK and GK varieties differ in whether pitch could signal lexically contrastive words (e.g. Lee and Jongman, 2015; Lee et al., 2016), future studies are warranted to apply the sequence recall task, for example, to further examine the effect of the two Korean varieties on encoding tones at the phonological/lexical level (see Gussenhoven et al., 2022).

In closing, we note some limitations of the present study in assessing the L1 and L2 influence on L3 tone perception. While this study provided the initial evidence supporting the influence of L2 learning experience by examining participants' L3 tone processing, future studies are needed to further examine the participants' perception of L2-Mandarin tone perception (e.g. Qin and Jongman, 2016) and L1-Korean pitch accent perception (e.g. Kim and Tremblay, 2021) to confirm the important role of L2 in the perception of L3 tones. Furthermore, the use of naturally occurring tones from Cantonese did not allow us to draw a robust conclusion about listeners' general sensitivity to pitch contour and pitch height cues (Jongman et al., 2017), and assess the universal tendencies in pitch contour and height perception (Burnham et al., 2014). Future studies may use synthesized stimuli from a pseudo-language and manipulate pitch contour and height differences systematically to address the potential difference in the processing of dynamic pitch variations (e.g. falling–rising tones) and static pitch information (e.g. higher–lower tones) (Laméris, 2022; Laméris and Post, 2023).

Acknowledgements

The authors would like to thank Drs Allard Jongman and Peggy Mok for their involvement in the earlier stages of this project and Yuqi Wang for her assistance in data analysis. They would also like to thank the anonymous reviewers and editors for their comments and suggestions, which helped improve the manuscript.

Declaration of Conflicting Interests

The authors declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The authors disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This work was supported by a Publication Capacity Fund from the Division of Humanities, the Hong Kong University of Science and Technology, awarded to the first author.

ORCID iD

Quentin Zhen Qin  <https://orcid.org/0000-0002-6117-7809>

Supplemental material

Supplemental material for this article is available online.

Notes

1. Chao numbers denote relative pitch heights and are used to describe tone contours; 1 is the lowest pitch, and 5 is the highest pitch.
2. Two GK-speaking L2 learners were recruited and tested in Seoul. The two participants moved to Seoul to pursue higher education but have lived in Seoul for less than three years. They had both lived in China for at least 12 months. Additional analyses included the length of residence and other variables of learning experiences (i.e. AOE and years of instruction), and the effects of learning experiences did not further contribute to the models on top of proficiency scores (see Section A of supplemental material).
3. All the GK-speaking naive participants moved to Seoul to pursue higher education but have lived in Seoul for less than three years. They were short-term transplants of Seoul-Korean, using measures from Kim and Jongman (2022).
4. HSK (Hanyu Shuiping Kaoshi) is a standard Chinese language proficiency test. There are six levels, and level-6 is the highest level of proficiency for the 2010–21 version of HSK. This study used this version to screen Korean-speaking L2 learners of Mandarin (Tsukada and Han, 2019).
5. The two pairs, T2–T1 with a larger acoustic difference (dissimilar tone pairs) and T2–T6 with a smaller acoustic difference (similar tone pairs), were included to test the effect of acoustic similarity in Qin and Jongman (2016; see also Chan and Chang, 2019). Given the word limit, the analyses incorporating the difference of acoustic similarity (i.e. tone pairs) were reported in Section B of supplemental material. The tone pairs were grouped into the contour–level tonal contrast in the analyses reported in the manuscript to focus on the effect of linguistic experience (L1 variety and L2 learning experience) on sensitivity to tonal contrasts.

6. Likewise, T1–T6 (dissimilar tone pair) and T3–T6 (similar tone pair) were grouped into the level–level tonal contrast in the analyses reported in the manuscript. See the effect of acoustic similarity in Section B of supplemental material.
7. 500 ms was chosen as ISI, as this ISI is known to tap into a phonetic mode of processing (Soo and Monahan, 2023). Pilot testing indicated that participants with non-tonal language backgrounds found it easier to do the task using the shorter ISI than a longer one (i.e. 1,500 ms).
8. The *d*-prime scores, which serve to tease apart any response bias from sensitivity (Macmillan and Creelman, 2004), were calculated for different tonal contrasts and submitted to a mixed-effects ANOVA. The analysis of *d*-prime scores yielded similar results (i.e. a significant interaction between the Learning experience and Tonal Contrast; $F(1, 61) = 10.9$, $p < .01$) as the accuracy-based analysis. Since the measures of *d*-prime scores did not allow for modeling trial-level accuracy and there is a word limit, the results of *d*-primes were not reported in the manuscript.
9. The results of the additional analyses (see Section B of supplemental material) examining the effect of acoustic similarity (i.e. tone pairs) showed a main effect of acoustic similarity (T2–T1/T1–T6 vs. T2–T6/T1–T3) with dissimilar tone pairs having higher accuracy than similar tone pairs. However, the acoustic similarity did not modulate the significant interaction effect between Learning Experience and Tonal Contrast.
10. It has been well documented that younger SK speakers use pitch rather than voice onset time (VOT) as the primary cue to distinguish native lenis and aspirated stops (e.g. Lee and Jongman, 2019) and may use the cue to perceive L2 segmental contrasts (e.g. Schertz et al., 2015). An alternative explanation of the L1 variety difference is that SK-speaking participants' great reliance on pitch cues in L1 segments may have masked or even reversed the SK-speaking L2 learners' disadvantage of using the pitch cue (pitch height) if it can be transferred from the L1 segmental domain to the L2/L3 suprasegmental domain (see the findings of Kim and Tremblay, 2022).

References

- Amengual M (2021) The acoustic realization of language-specific phonological categories despite dynamic cross-linguistic influence in bilingual and trilingual speech. *The Journal of the Acoustical Society of America* 149: 1271–84.
- Bates DM, Maechler M, Bolker B, and Walker S (2015) lme4: Linear mixed-effects models using Eigen and Eigen++. *Journal of Statistical Software* 67: 1–48.
- Best CT (2019) The diversity of tone languages and the roles of pitch variation in non-tone languages: Considerations for tone perception research. *Frontiers in Psychology* 10: 364.
- Burnham D, Kasisopa B, Reid A, et al. (2014) Universality and language-specific experience in the perception of lexical tone and pitch. *Applied Psycholinguistics* 36: 1459–91.
- Chan IL and Chang CB (2018) LEXTALE_CH: A quick, character-based proficiency test for Mandarin Chinese. In: Bertolini AB and Kaplan BMJ (eds) *Proceedings of the 42nd Annual Boston University Conference on Language Development*. Somerville, MA: Cascadia Press, pp. 114–30.
- Chan IL and Chang CB (2019) Perception of nonnative tonal contrasts by Mandarin–English and English–Mandarin sequential bilinguals. *The Journal of the Acoustical Society of America* 146: 956–72.
- Chen HC and Han QW (2019) L3 phonology: Contributions of L1 and L2 to L3 pronunciation learning by Hong Kong speakers. *International Journal of Multilingualism* 16: 492–512.
- Chen J, Antoniou M, and Best CT (2023) Phonological and phonetic contributions to perception of non-native lexical tones by tone language listeners: Effects of memory load and stimulus variability. *Journal of Phonetics* 96: 101199.

- Cutler A (1986) Forbear is a homophone: Lexical prosody does not constrain lexical access. *Language and Speech* 29: 201–20.
- Hartig F (2020) *DHARMa: Residual diagnostics for hierarchical regression models*. Available at: <https://cran.r-project.org/web/packages/DHARMa/vignettes/DHARMa.html> (accessed March 2024).
- Dmitrieva O (2019) Transferring perceptual cue-weighting from second language into first language: Cues to voicing in Russian speakers of English. *Journal of Phonetics* 73: 128–43.
- Dupoux E, Peperkamp S, and Sebastián-Gallés N (2001) A robust method to study stress ‘deafness’. *The Journal of the Acoustical Society of America* 110: 1606–18.
- Falk Y and Bardel C (2011) Object pronouns in German I3 syntax: Evidence for the I2 status factor. *Second Language Research* 27: 59–82.
- Francis AL, Ciocca V, Ma L, and Fenn K (2008) Perceptual learning of Cantonese lexical tones by tone and non-tone language speakers. *Journal of Phonetics* 36: 268–94.
- Francis AL and Nusbaum HC (2002) Selective attention and the acquisition of new phonetic categories. *Journal of Experimental Psychology: Human Perception and Performance* 28: 349–66.
- Gandour J (1981) Perceptual dimensions of tone: Evidence from cantonese. *Journal of Chinese Linguistics* 9: 20–36.
- Gandour JT (1983) Tone perception in Far Eastern languages. *Journal of Phonetics* 11: 149–75.
- Gussenhoven C, Lu Y-A, Lee-Kim S-I, et al. (2022) The sequence recall task and Lexicality of tone: Exploring tone ‘deafness’. *Frontiers in Psychology* 13: 902569.
- Han J-I and Tsukada K (2020) Lexical representation of Mandarin tones by non-tonal second-language learners. *The Journal of the Acoustical Society of America* 148: EL46–50.
- Holt LL and Lotto AJ (2006) Cue weighting in auditory categorization: Implications for first and second language acquisition. *The Journal of the Acoustical Society of America* 119: 3059–71.
- Jongman A, Qin Z, Zhang J, and Sereno JA. (2017) Just noticeable differences for pitch direction, height, and slope for Mandarin and English listeners. *The Journal of the Acoustical Society of America* 142: EL163–69.
- Jun J, Kim J, Lee H, and Jun S-A (2006) The prosodic structure and pitch accent of Northern Kyongsang Korean. *Journal of East Asian Linguistics* 15: 289–317.
- Jun SA (2010) Korean intonational phonology and prosodic transcription. In: Jun SA (ed.) *Prosodic typology: The phonology of intonation and phrasing*. Oxford: Oxford University Press, pp. 201–29.
- Kim H and Jongman A (2022) The influence of inter-dialect contact on the Korean three-way laryngeal distinction: An acoustic comparison among Seoul Korean speakers and Gyeongsang speakers with limited and extended residence in Seoul. *Language and Speech* 65: 531–53.
- Kim H and Tremblay A (2021) Korean listeners’ processing of suprasegmental lexical contrasts in Korean and English: A cue-based transfer approach. *Journal of Phonetics* 87: 101059.
- Kim H and Tremblay A (2022) Intonational cues to segmental contrasts in the native language facilitate the processing of intonational cues to lexical stress in the second language. *Frontiers in Communication* 7: 845430.
- Kuznetsova A, Brockhoff PB, and Christensen RHB (2018) *lmerTest: Tests in linear mixed effects models: R package version 2.0-20*. Available at: <http://CRAN.R-project.org/package=lmerTest> (accessed March 2024).
- Laméris TJ (2022) The effect of L1 pitch status and extralinguistic factors on L2 tone learning. In: *Proceedings of the International Conference on Speech Prosody*. ISCA, pp. 708–12.
- Laméris TJ and Post B (2023) The combined effects of L1-specific and extralinguistic factors on individual performance in a tone categorization and word identification task by English-L1 and Mandarin-L1 speakers. *Second Language Research* 39: 833–71.

- Lee H and Jongman A (2015) Acoustic evidence for diachronic sound change in Korean prosody: A comparative study of the Seoul and South Kyungsang dialects. *Journal of Phonetics* 50: 15–33.
- Lee H and Jongman A (2019) Effects of sound change on the weighting of acoustic cues to the three-way laryngeal stop contrast in Korean: Diachronic and dialectal comparisons. *Language and Speech* 62: 509–30.
- Lee H, Jongman A, and Zhang J (2016) Variation and change in the nominal pitch-accent system of South Kyungsang Korean. *Phonology* 33: 325–51.
- Lee H, Kong EJ, and Holliday JJ (2019) An investigation of the effect of L1 dialect on the L2 perception of lexical stress. In: Calhoun S, Escudero P, Tabain M, and Warren P (eds) *Proceedings of the 19th International Congress of Phonetic Sciences (ICPhS)*. Canberra: Australasian Speech Science and Technology Association, pp. 2281–85.
- Lemhöfer K and Broersma M (2012) Introducing LexTALE: A quick and valid Lexical Test for Advanced Learners of English. *Behavior Research Methods* 44: 325–343.
- Luo J, Li VG, and Mok PPK (2020) The perception of Cantonese vowel length contrast by Mandarin speakers. *Language and Speech* 63: 635–59.
- Macmillan NA and Creelman CD (2004) *Detection theory: A user's guide* (2nd ed.). Mahwah, NJ: Psychology.
- Matthews S and Yip V (2011) *Cantonese: A comprehensive grammar* (2nd ed.). London: Routledge.
- Matuschek H, Kliegl R, Vasishth S, Baayen H, and Bates D (2017) Balancing Type I error and power in linear mixed models. *Journal of Memory and Language* 94: 305–15.
- Mok P, Zuo D, and Wong PWY (2013) Production and perception of a sound change in progress: Tone merging in Hong Kong Cantonese. *Language Variation and Change* 25: 341–70.
- Niu Y and Mok PPK (2022) *The effect of L2 proficiency on pitch and phonation sensitivity in naïve tone perception: A study of Wenzhou dialect tone perception by English learners of Mandarin*. In: *Proceedings of First International Conference on Tone and Intonation (TAI)*, pp. 181–85.
- O'Brien RM (2007) A caution regarding rules of thumb for variance inflation factors. *Quality and Quantity* 41: 673–90.
- Qin Z and Jongman A (2016) Does second language experience modulate perception of Tones in a third language? *Language and Speech* 59: 318–38.
- Qin Z, Tremblay A, and Zhang J (2019) Influence of within-category tonal information in the recognition of Mandarin–Chinese words by native and non-native listeners: An eye-tracking study. *Journal of Phonetics* 73: 144–57.
- Ramsey SR (1975) *Accent and morphology in Korean dialects: A descriptive and historical study*. New Haven, CT: Yale University.
- R Core Team (2022). *R: A language and environment for statistical computing* [software]. Vienna: R Foundation for Statistical Computing. Available at: <http://www.R-project.org> (accessed March 2024).
- Rothman J (2011) L3 syntactic transfer selectivity and typological determinacy: The typological primacy model. *Second Language Research* 27: 107–27.
- Schaefer V and Darcy I (2014) Lexical function of pitch in the first language shapes cross-linguistic perception of Thai tones. *Laboratory Phonology* 5: 489–522.
- Schertz J, Cho T, Lotto A, and Warner N. (2015) Individual differences in phonetic cue use in production and perception of a non-native sound contrast. *Journal of Phonetics* 52: 183–204.
- Soo R and Monahan PJ (2023) Phonetic and lexical encoding of tone in Cantonese heritage speakers. *Language and Speech* 66: 652–77.
- Tsukada K and Han JI (2019) The perception of Mandarin lexical tones by native Korean speakers differing in their experience with Mandarin. *Second Language Research* 35: 305–18.

- Wagenmakers EJ, Marsman M, Jamil T, et al. (2018) Bayesian inference for psychology. Part I: Theoretical advantages and practical ramifications. *Psychonomic Bulletin and Review* 25: 35–57.
- Wayland RP and Li B (2008) Effects of two training procedures in cross-language perception of tones. *Journal of Phonetics* 36: 250–67.
- Westergaard M (2021) Microvariation in multilingual situations: The importance of property-by-property acquisition. *Second Language Research* 37: 379–407.
- Wiener S and Goss S (2019) Second and third language learners' sensitivity to Japanese pitch accent is additive: An information-based model of pitch perception. *Studies in Second Language Acquisition* 41: 897–910.
- Wrembel M, Marecka M, and Kopečková R (2019) Extending perceptual assimilation model to L3 phonological acquisition. *International Journal of Multilingualism* 16(4): 513–33.
- Yip M (2002) *Tone*. Cambridge: Cambridge University Press.
- Yu KM and Lam HW (2014) The role of creaky voice in Cantonese tonal perception. *The Journal of the Acoustical Society of America* 136(3): 1320–33.
- Yuan B (2010) Domain-wide or variable-dependent vulnerability of the semantics–syntax interface in L2 acquisition? Evidence from wh-words used as existential polarity words in L2 Chinese grammars. *Second Language Research* 26(2): 219–60.
- Zhang Y and Kirby J (2020) The role of F_0 and phonation cues in Cantonese low tone perception. *The Journal of the Acoustical Society of America* 148(1): EL40–45.

