Cantonese tone production in pre-school Urdu-Cantonese bilingual minority children

Yao Yao[1,2,3], Angel Chan[1,2,3,4], Roxana Fung[1,3], Wing Li Wu[1], Natalie Leung[1], Sarah Lee[1], & Jin Luo[1]

[1]Department of Chinese and Bilingual Studies, The Hong Kong Polytechnic University, Hong Kong

[2]Speech Therapy Unit, The Hong Kong Polytechnic University, Hong Kong

[3]The Hong Kong Polytechnic University - Peking University Research Centre on Chinese Linguistics, Hong Kong

[4] Research Centre for Language, Cognition, and Neuroscience

Addresses for correspondence:

Angel Chan, Speech Therapy Unit, Department of Chinese and Bilingual Studies, The Hong Kong Polytechnic University, Hong Kong. Email: angel.ws.chan@polyu.edu.hk. Phone: +852 2766-7439.

Yao Yao, Department of Chinese and Bilingual Studies, The Hong Kong Polytechnic University, Hong Kong. Email: y.yao@polyu.edu.hk. Phone: +852 2766-7478.

Cantonese tone production in pre-school Urdu-Cantonese bilingual minority children

Abstract

Aim: In this study, we examine the production of Cantonese tones by preschool Urdu-Cantonese children living in Hong Kong.

Methodology: 21 L1-Urdu L2-Cantonese children (ages 4-6) and 20 age-matched L1-Cantonese children participated in a picture-naming experiment with 86 words (109 syllables in total).

Data and Analysis: Acoustic analysis was carried out for perceptually correct and incorrect tone productions of each tone. Comparisons were also made across speaker groups regarding accuracy rates and error patterns.

Findings: Overall, L1-Urdu participants had lower accuracy and greater tone confusion than L1-Cantonese participants. The pattern is attributable to influence from Urdu prosody, ongoing Cantonese tone mergers, and general sensitivity to phonetic information.

Originality: This is the first empirical study on the acquisition of Cantonese tones by children who are heritage speakers of a non-tone language.

Significance: This study extends the literature of early bilingual phonology by furthering our understanding of an under-studied bilingual population, i.e. heritage children of a non-tone language acquiring a tone language as the majority language. The findings of this study also produce implications for the practice of language educators and speech therapy professionals working with bilingual children.

Keywords

Hong Kong Cantonese, lexical tone, Urdu prosody, early bilingual acquisition, South Asian heritage speakers in Hong Kong

Introduction

The literature of bilingual phonology has produced influential theoretical models such as the Perceptual Assimilation Model (PAM; Best, 1995; Best & Tyler, 2007) and the Speech Learning Model (SLM; Flege, 1996). However, our understanding in this area is heavily skewed toward adult speakers and segmental acquisition. Research on bilingual children's acquisition of suprasegmentals is particularly lacking, but a notable exception is a vibrant cluster of studies documenting Cantonese tones acquired by overseas Cantonese heritage children (in Australia: Holm & Dodd, 1999, 2006; in the US: Kan &

Schmid, 2019) and bilingual children in Hong Kong (Cantonese-Mandarin: Law & So, 2006; Cantonese-English: Mok & Lee, 2018). This study extends this line of research by examining a distinctive child population, i.e. children who are heritage speakers of a non-tone language, Urdu, who acquire Cantonese as their second language (L2) and societal language. These children are born to and raised by families with Pakistani heritage in Hong Kong. This study thus also contributes to the scholarly works on South Asian heritage speakers in Hong Kong.

Background literature

L2 tone acquisition

A tone language (e.g. Cantonese) systematically uses pitch pattern to distinguish word meaning. The literature on L2 tone acquisition—focusing mostly on late learners—places much emphasis on the influence of the first language's (L1) prosody (e.g., Chang, Yao, & Huang, 2017; Francis, Ciocca, Ma, & Fenn, 2008; Hallé, Chang, & Best, 2004; Hao, 2012; Wang, Jongman, & Sereno, 2003; Wayland & Guion, 2004). For example, Hallé et al. (2004) reported that naïve L1-French listeners were able to perceive the acoustic differences among Mandarin tones, but did not show categorical perception like native Mandarin listeners did. This finding suggested that pitch perception occurred at the psychophysical level—but not the linguistic level—for

French listeners, whose non-tone L1 has sentence-level intonation but no lexical-level prosodic constraint.

A later study by So and Best (2010) further examined how specific L1 prosodic features would modulate the perception of L2 tones and theorized the influence of L1 prosody in an extension of the PAM framework, the PAM for Suprasegmentals (PAM-S) model. They conducted a Mandarin tone identification task with three groups of Mandarin-naïve listeners, who had lexical-tone L1 (Cantonese), pitch-accent L1 (Japanese), or non-tone L1 (English), respectively. All three groups had more difficulty with Mandarin tone pairs that had similar phonetic properties than those with dissimilar properties, indicating the presence of a universal sensitivity to phonetic information. The authors further showed how between-group differences were accounted for by PAM-S. Specifically, PAM-S predicts that if L1 also has tones, L2 tones may be assimilated into L1 tones, with assimilatory routes subject to the acoustic-phonetic mapping between the two systems; if L1 is a non-tone language (e.g. English, French), L2 tones can be assimilated either to L1 phrasal/sentential intonation categories, or as uncategorized intonation patterns, or as non-speech melodic units, and in any case, linguistic processing of tones may not be available, which is compatible with Hallé et al.'s findings of French listeners. Since the initial proposal, PAM-S has been successfully applied to a number of subsequent studies (Kan & Schmid, 2019; So & Best, 2011, 2014). In the current study, we adopt PAM-S as the main theoretical framework.

As mentioned above, there is a general lack of literature of early bilinguals' acquisition of tone, but the topic is getting increasing attention. In particular, a group of studies on Cantonese tone acquisition by bilingual children have emerged in recent years. Here we discuss three studies in detail, all involving Cantonese-learning bilingual children whose other language is a non-tone language: Holm & Dodd (1999, 2006; hereafter "H&D"), Kan & Schmid (2019; hereafter "K&S"), and Mok & Lee (2018; hereafter "M&L").

H&D examined the phonological development of L1-Cantonese L2-English children in Australia (ages 2;0 to 5;7), all of whom were Cantonese-heritage children who used Cantonese predominantly at home. H&D found high similarity between bilingual children and their Cantonese monolingual peers in Hong Kong in terms of both tone production accuracy and error patterns.

However, pronounced differences between bilingual and monolingual children were reported in later studies, for both perception and production. K&S tested a similar population in a different English-speaking country, the U.S., at a later stage (ages 5-11). K&S focused on the perception of two tone pairs: a "similar" pair with the high-rising

Tone 2 and the low-rising Tone 5, and a “dissimilar” pair with the high-level Tone 1 and the low-falling Tone 4. Bilingual children overall scored lower than monolingual controls in Hong Kong, especially in the “dissimilar” tone pair. K&S explained the between-pair differences in the PAM-S model: While the tones in the “dissimilar” pair are mapped to distinctive English intonational categories (i.e. high-flat Tone 1 to flat pitch in English; low-falling Tone 4 to statement intonation in English), the rising tones in the “similar” pair are mapped to the same category (i.e. both to the question intonation in English), thus leading to worsened distinction of the “similar” pair.

M&L examined a different group, which consisted of Cantonese-English simultaneous bilinguals (ages 2;0 to 2;6) born to mixed race parents in Hong Kong. Through the analysis of spontaneous speech recorded in a longitudinal corpus, M&L found that bilinguals showed a delay in tone development at 2;0, but caught up later at 2;6 as a group, although considerable individual differences persisted. M&L also noted the use of a “high-low” template by some bilinguals, presumably influenced by the trochaic stress pattern in English.

Taken together, previous evidence suggests that early bilinguals’ tone development is likely delayed and influenced by the (non-tone) prosody of their other language. In this study, we extend this line of research by examining a novel group of bilingual children, who are heritage speakers of a non-tone language (Urdu) living in Hong Kong. This group can be viewed as the mirror image of the Cantonese-heritage children in H&D and K&S, while also distinct from the simultaneous bilinguals in M&L. To the best of our knowledge, there has not been any formal investigation into the tone acquisition of bilinguals with such background.

Heritage speakers and their language acquisition

Before proceeding, it is worth noting the distinctive features of heritage speakers, i.e. individuals "raised in a home where one language is spoken who subsequently switch to another dominant language" (Polinsky & Kagan, 2007:368). Compared to the simultaneous bilinguals documented in M&L, heritage speakers are different in terms of both the order of acquisition and the nature of bilingual input. Most importantly, heritage speakers' exposure to the majority language (L2) typically only begins when they start schools, but increases rapidly with schooling, work, and socialization, and eventually becomes the speaker's dominant language in most cases.

A mounting literature has been devoted to heritage speakers' phonology, focusing predominantly on the heritage language (L1) phonology (as in H&D and K&S, see also Au et al., 2002; Chang & Yao, 2016, 2019; Chang et al., 2011; Oh et al., 2010; Polinsky, 2018; Rao, 2015). What is largely unknown, however, is the acquisition of the majority language. Some recent studies (e.g. Lloyd-Smith, Einfeldt, & Kupisch, in press) suggest that adult heritage speakers' accent in the majority language can be indistinguishable from that of native monolingual speakers, but this doesn't mean that heritage speakers follow the same developmental pathways as monolingual speakers. A main goal of this study is to open this line of inquiry by comparing heritage Urdu children' acquisition of Cantonese (majority language) tones with that of native Cantonese children. In doing so, we also contribute to the growing literature on heritage speakers' prosodic acquisition.

In the following, we provide a brief background of Cantonese tones, Urdu prosody, and the South Asian community in Hong Kong.

Cantonese lexical tones

Cantonese has six phonological tones: Tone 1 (high level [55]), Tone 2 (high rising [25]), Tone 3 (mid level [33]), Tone 4 (low falling [21]), Tone 5 (low rising [23]) and Tone 6 (low level [22]) (Mok, Zuo, & Wong, 2013). The tones are differentiated mainly in pitch height (high, mid, low), pitch contour (level, rising, falling), and the magnitude of pitch shift (Khouw & Ciocca, 2007). In this paper, we refer to these tones by shorthand names ("T1"-"T6") in text and by tone numbers (1-6) in IPA sequences (e.g. [sy4]). Figure 1 shows the F0 contours of all six tones based on a male adult speaker's production.

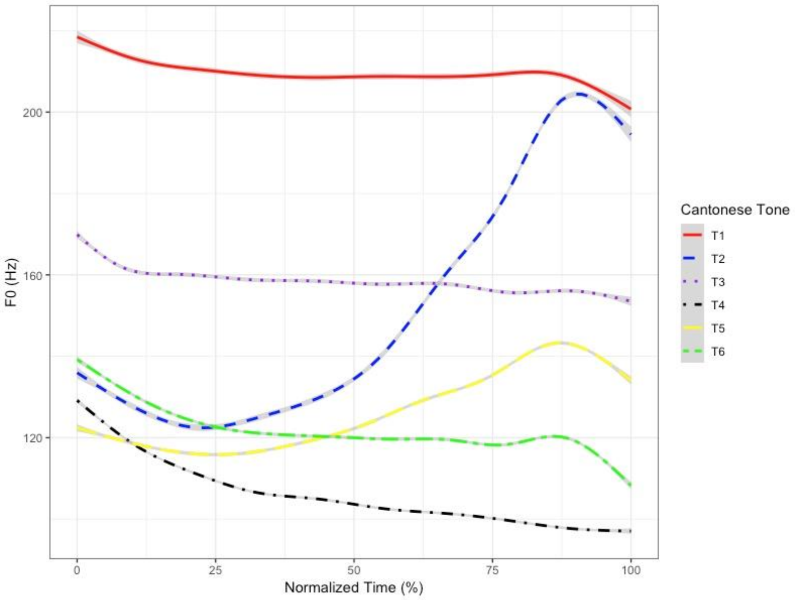


Figure 1. F0 traces of the six lexical tones averaged across four syllables ([fu], [ji], [sɛ], and [si]) produced by a male speaker

While tone acquisition is an important developmental milestone, it is not entirely clear when native Cantonese children complete the process. A number of studies found evidence of early acquisition before the age of 3 (So & Dodd, 1995; To, Cheung, & Mcleod, 2013). For example, To et al. (2013) reported near-perfect production (mean accuracy rate > 98%) even in the youngest age group (2;4-2;9) they investigated. Meanwhile, recent studies by Wong & Leung (2018) and Mok, Fung & Li (2019) reported incomplete acquisition at as late as 5 or 6, echoing Wong’s (2013) finding of protracted tone development in Mandarin children. It is unclear whether the discrepancy across studies regarding age of acquisition of tones can be completely explained away by differences in research methods. In this study, we infer between-group differences by comparing bilingual and monolingual children's performance in the same experimental condition.

Another factor that complicates the acquisition of Cantonese tones is the ongoing mergers of several tone pairs (T2-T5, T4-T6, and T3-T6; see Mok et al., 2013 and references therein). The merging tones are associated with heightened perceptual difficulty and weaker acquisition by monolingual children (Lee, Chan, Lam, Van Hasselt, & Tong, 2015), as a result of reduced perceptual distinction and greater variation in the input. Recent literature suggests that bilingual children may be susceptible to ongoing sound change in the input, too (see K&S and Chang & Yao, 2016).

Urdu prosody and Urdu speakers in Hong Kong

Urdu is a Hindustani (Indo-Aryan) language, most widely spoken in Pakistan and well known for its resemblance to Hindi. While the literature on Hindi-Urdu prosody is overall scarce (e.g. Hussain, 1997; Jabeen, 2010; Ohala, 1986), it is safe to conclude that both languages have lexical stress and phrase-/sentence-level intonation but no lexical tones. A salient feature of Hindi-Urdu prosody is the frequent use of a rising (LH) contour, especially on non-final content words (Hussain, 1997; Jabeen, 2010).

This feature can also be transferred to an L2, as shown by recent studies of Indian English (Puri, 2013).

The number of Urdu speakers in Hong Kong has been growing steadily in the past few decades, as part of the continuing wave of South Asian new immigrants. In particular, Pakistani-heritage children under 15 almost doubled in number from 2001 to 2011 (Census and Statistics Department of HKSAR, 2011), and contributed the largest group of ethnic minority admittees in local kindergartens (Hong Kong Unison, 2012). Despite early immersion in the largely Cantonese-speaking society, minority children showed great variation in Cantonese proficiency (Hong Kong Unison, 2012), allegedly due to the varying degrees of Cantonese knowledge of their parents. The rise of the South Asian-heritage population has stimulated scholarly works on language pedagogy and education policy (e.g. Li & Chuk, 2015; Shum et al., 2011), but a rigorous linguistic description of ethnic minority children's language acquisition is yet to be obtained.

Current research

We report on the first empirical study of Cantonese tone production by Urdu heritage children in Hong Kong, compared to age-matched Cantonese-dominant children. For simplicity, we refer to these two groups by their first language, i.e. L1-Urdu or L1-Cantonese, in this paper.

Given the current literature, we put forward two main hypotheses:

(1) Hypothesis I: L1-Urdu children have lower tone accuracy than L1-Cantonese controls, as a result of less and later-onset exposure.
(2) Hypothesis II: L1-Urdu children's acquisition of Cantonese tones is modulated by (a) influence from Urdu prosody, (b) ongoing tone changes in Cantonese, (c) general sensitivity to phonetic information.

Specifically, regarding Hypothesis IIa, we predict that both rising tones (T2 and T5) will be assimilated to the Urdu LH contour according to PAM-S, undergoing either a *Single Category* (when both tones assimilate equally well or poorly) or a *Category Goodness* (when one tone assimilates better than the other) assimilation. In either case, some degree of perceptual difficulty is predicted, more strongly in *Single Category* routine than *Category Goodness* assimilation (see K&S, So & Best, 2010). The other four Cantonese tones, presumably with no corresponding intonational categories in Urdu, may be assimilated as either uncategorized pitch patterns (i.e. *Uncategorized*) or non-speech melodic units (i.e. *Non-Assimilable*). Furthermore, greater difficulty is predicted for tones undergoing the merging process than those that are not (Hypothesis IIb), and for tones that are perceptually similar than those that are perceptually distinctive (Hypothesis IIc). While IIb and IIc may also affect L1-Cantonese children, IIa is unique to L1-Urdu children.

Methods

Participants

The participants are 21 L1-Urdu children (13M, 8F; ages 4;5 to 6;6) and 20 L1-Cantonese children (10M, 10F; ages 4;6 to 6;3). Born and raised in Hong Kong, all the participants were attending local kindergartens at the time of testing, where Cantonese was used as the medium of instruction. The L1-Urdu participants grew up in Urdu-speaking households, with both parents being native Urdu speakers. Due to the high degree of bi(multi)lingualism in the society, both groups may have some exposure to other languages (English, Mandarin, etc), but the exposure should be minimal at their age. All the participants are typically developing, with no known history of developmental delay or speech, language, hearing or cognitive disorder.

Given the general lack of knowledge of L1-Urdu children's Cantonese development, we conducted a Cantonese vocabulary assessment for the L1-Urdu participants, consisting of two receptive vocabulary tests—CRVT (the Cantonese Receptive Vocabulary Test; Cheung, Lee, & Lee, 1997) and PPVT-4 (the Peabody Picture Vocabulary Test, 4^{th} edition; Dunn & Dunn, 2007)—and a parental questionnaire.

All but one L1-Urdu participants took the two vocabulary tests (one participant was absent from school that day). Both tests had similar procedure: in each trial, the participant was asked to point to the picture depicting the target word they heard in lieu of three distractor pictures. CRVT is commonly used in Hong Kong and has gathered extensive L1-Cantonese developmental norms, whereas PPVT-4 has been adapted into Cantonese and Urdu to assess children's vocabulary knowledge in these two languages. Our results showed that all the L1-Urdu participants' CRVT scores were in the bottom 10^{th} percentile for their age group in CRVT norms, indicating a significantly smaller Cantonese vocabulary than average L1-Cantonese peers'. Their PPVT-4 scores were similar for Cantonese (M = 43.6, SD =8.6) and Urdu (M = 38.4, SD =10.5; $t(19)=1.54$, $p > .1$, paired sample t-test, two-tailed).

The majority of L1-Urdu participants (N = 17) returned a completed parental questionnaire, where parents rated how often their child spoke Cantonese/Urdu at home on a 5-point scale (1 = *never*, 5 = *all the time*). Overall Urdu tends to be more often used (M = 3.59, SD = 1.18) than Cantonese (M = 2.76, SD = 1.15), but the difference did not reach significance ($t(16)= -1.69$, $p > .1$, paired sample t-test, two-tailed).

Taken together, the L1-Urdu participants clearly lagged behind their L1-Cantonese peers in Cantonese vocabulary development, likely due to the reduced amount of Cantonese input.

Production experiment

All the participants completed a picture naming experiment with two tasks in a quiet room of their school or community centre: a standard picture naming task (hereafter the "naming" task) where the participant saw a picture and provided an oral response to the experimenter's question (e.g. [mɛ1 lɐi4 ka3] 'What is this?'), followed by a modified naming task where the participant saw a picture and named after the experimenter did (hereafter the "repetition" task). The stimuli of the naming task consisted of 41 monosyllabic items (e.g. [fa1] 'flower') and 23 disyllabic items (e.g. [tin6 wa2] 'telephone'). The repetition task used 22 monosyllabic items in the form of 11 minimal pairs that contrasted in tone (e.g. [tɐŋ1] 'lamp' and [tɐŋ3] 'stool').

Altogether the experiment used 86 (=41+23+22) word stimuli—with 109 syllable tokens and 84 unique syllables types (segments + tone), covering all the tones (T1-T6) and vowels in Cantonese (11 monophthongs: [i], [y], [ɪ], [ɛ], [œ], [a], [ɵ], [ɐ], [u], [ʊ], [ɔ];11 diphthongs: [ai], [ei], [ɐi], [ui], [ɔi], [au], [ɐu], [iu], [ou], [ɵy], [ɛu]). Each tone appeared in 14 to 23 stimuli (list of stimuli in Supplementary Materials). All the target words were familiar to the participants, and no instance of unknown words was reported.

The participants' productions in the picture naming experiment were recorded using a Tascam DR-44WL audio recorder connected to an AKG SE300B condenser microphone.

Perceptual judgment of accuracy

Participants' productions of the target words were perceptually evaluated for production accuracy by two native Cantonese speakers independently, both with extensive training in Cantonese phonetics (interrater agreement > 95%). For perceptually incorrect productions, the raters further annotated error type (segmental, tonal, or both) and the closest resembling tone in case of tonal errors. It should be noted that perceptual judgment was carried out with two judges using unfiltered speech productions, i.e. the unfiltered-two-judges condition recommended in Mok et al. (2019). Only productions judged to be correct by both raters were coded as correct in the analysis.

Acoustic analysis

For each syllable token, F0 measurements of the voiced interval (manually annotated as the first and last complete and regular vocal cycles) were extracted from Praat (Boersma

& Weenink, 2017), smoothed by the trimming algorithm in ProsodyPro (Xu, 2013), and normalized in both time and pitch domains in order to facilitate cross-speaker comparison. The normalization procedure resulted in 10 measurements of T (see formula (3); Shi, 1986) per token, equidistant from 5% to 95% of the interval. The T measure is in the range of [0, 5], comparable to the conventional 5-point tone description.

(3) $T_x = \frac{5\times(\log(F0_x)-\log(F0_{min}))}{\log(F0_{max})-\log(F0_{min})}$, where $F0_{min}$ and $F0_{max}$ are talker-specific minimal and maximal F0 measurements in the data

Six acoustic features were derived for each syllable token to characterize the pitch pattern in the voiced interval:

- **MinT:** Minimal pitch (in T)
- **MaxT:** Maximal pitch (in T)
- **MeanT:** Mean pitch (in T)
- **ExcursionT:** The difference between MinT and MaxT. If MinT occurs before MaxT, ExcursionT is positive (i.e. upward pitch contour); otherwise ExcursionT is negative (i.e. downward pitch contour).
- **SlopeT:** ExcursionT divided by the lapse of time (in number of time points) between MinT and MaxT. The sign of SlopeT follows the sign of the corresponding ExcursionT.

- **Dur:** log-transformed duration (in *s*) of the voiced interval.

Two additional features regarding the timing of Min(Max)T were derived for tokens of rising tones (T2 and T5), which typically have more complex pitch contours as shown in Figure 1.

- **MinTPos:** The time point where MinT occurs
- **MaxTPos:** The time point where MaxT occurs

Statistical analysis

The statistical analysis was carried out in R (version 3.5.1; R Core, 2018), using mainly the lme4 package (version 1.1-18-1; Bates, Maechler, Bolker, & Walker, 2015).

Results

The complete dataset consists of 4469 syllable tokens (i.e. 109 syllable tokens × 41 participants). After excluding tokens that were judged by at least one rater as only containing segmental errors but no tonal errors (N = 106), or as overall wrong but no identified segmental or tonal error (N = 3), the final dataset has 4360 syllable tokens (L1-Urdu: N =2187; L1-Cantonese: N = 2173). Accuracy was coded as 1 (Correct) if both raters considered the token correct in tone, or 0 (Incorrect) otherwise.

L1-Urdu participants achieved an average tone accuracy of 82.6%, significantly lower than L1-Cantonese participants' 95.9% ($X^2(1, N= 4360) = 198.1$, $p < .001$). Only 5 L1-Urdu participants achieved accuracy over 90%, while all L1-Cantonese participants did so. Significant variation in accuracy is also observed across tones ($X^2(5, N= 4360) = 72.2$, $p < .001$), especially for L1-Urdu participants, who had more difficulty with T3 and T4 than other tones (see Figure 2).

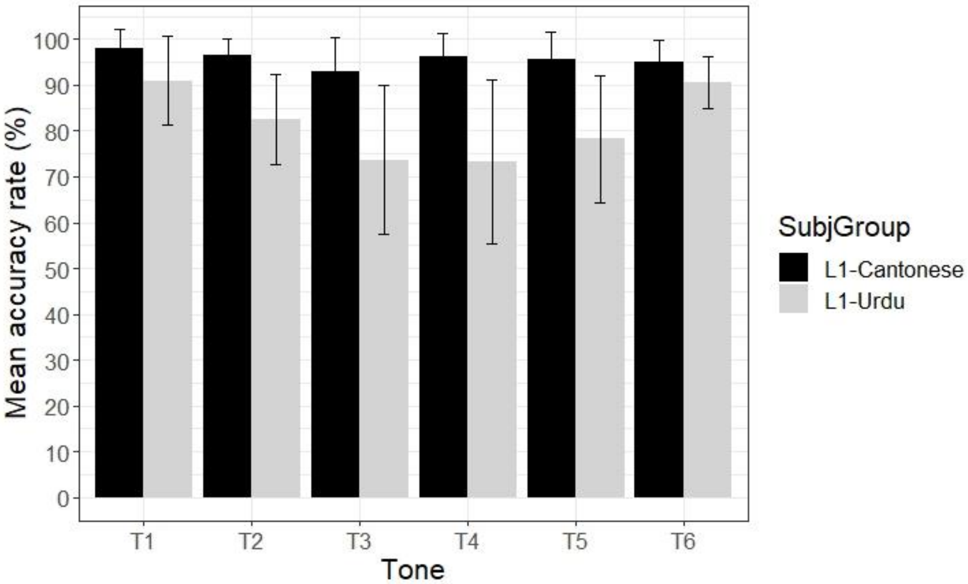


Figure 2. Barplot of production accuracy by tone by speaker group. Error bars indicate the standard deviation of by-speaker accuracy rate.

Acoustic comparison of perceptually correct and incorrect tone productions

To determine the acoustic differences between perceptually correct and incorrect tonal productions, we built generalized mixed-effects models for each tone. All the models started with a common structure, with Accuracy as the dependent variable, all the acoustic measures (centered) and speaker/item properties (SubjGroup = [L1-Cantonese, L1-Urdu], SubjGender = [F, M], WordLen (i.e. number of syllables in the embedding word) = [1, 2], Task = [Naming, Repetition]) as fixed-effects predictors, and random intercepts for speaker and item.

Each initial model went through a process of backward elimination: in each step, the weakest fixed-effects predictor (i.e. with the highest p value) was tested for significance by comparing the model fit with and without this predictor and removed if nonsignificant, until all the remaining predictors were significant. The resulting models were checked for collinearity: if the model contains inherently correlated acoustic measures (i.e., MinT, MaxT, and MeanT; SlopeT and ExcursionT), only the one with the highest significance would be retained (final model summaries available in Supplementary Materials).

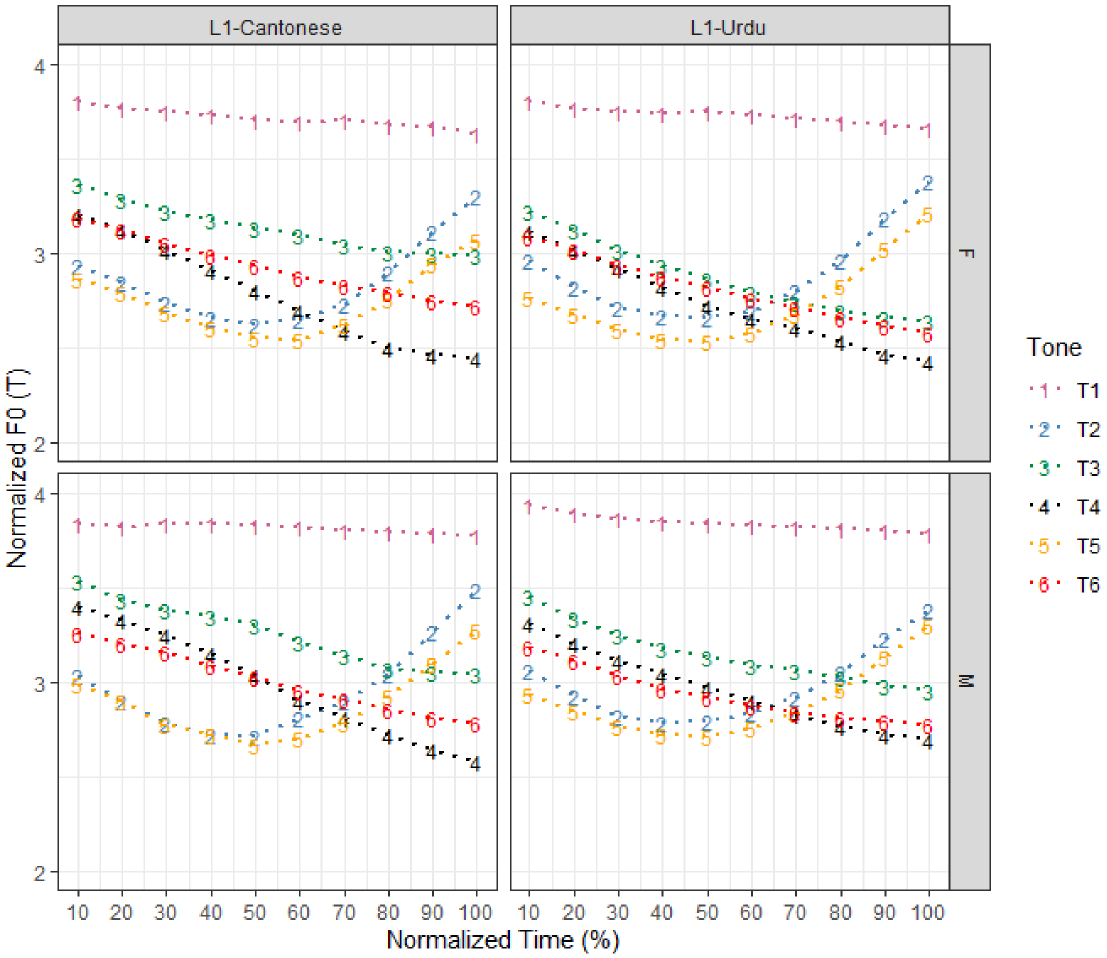
The models revealed significant acoustic differences between correct and incorrect productions. The accuracy of T1 and T3 (both level tones) crucially relies on higher MeanT (T1: $\beta = 1.49$, $p < .001$; T3: $\beta = 0.46$, $p = .03$); T2 (high-rising), on greater MaxTPos ($\beta = 0.14$, $p < .001$) and MaxT ($\beta = 0.32$, $p = .04$); T4 (low-falling), on lower MinT ($\beta = -0.62$, $p = .01$), lower ExcursionT ($\beta = -0.65$, $p = .003$), and shorter Dur ($\beta =$

-1.33, p = .004); T5 (low-rising), on greater ExcursionT (β = 0.28, p = .02) and longer Dur (β = 0.68, p = .04); and T6 (low-level), on higher MinT (β = 0.32, p = .02). All the acoustic effects were expected, except for the positive effect of MinT on T6, which is probably due to the mispronunciation of T6 as the low-falling T4 (see *Analysis of confusable tone pairs*).

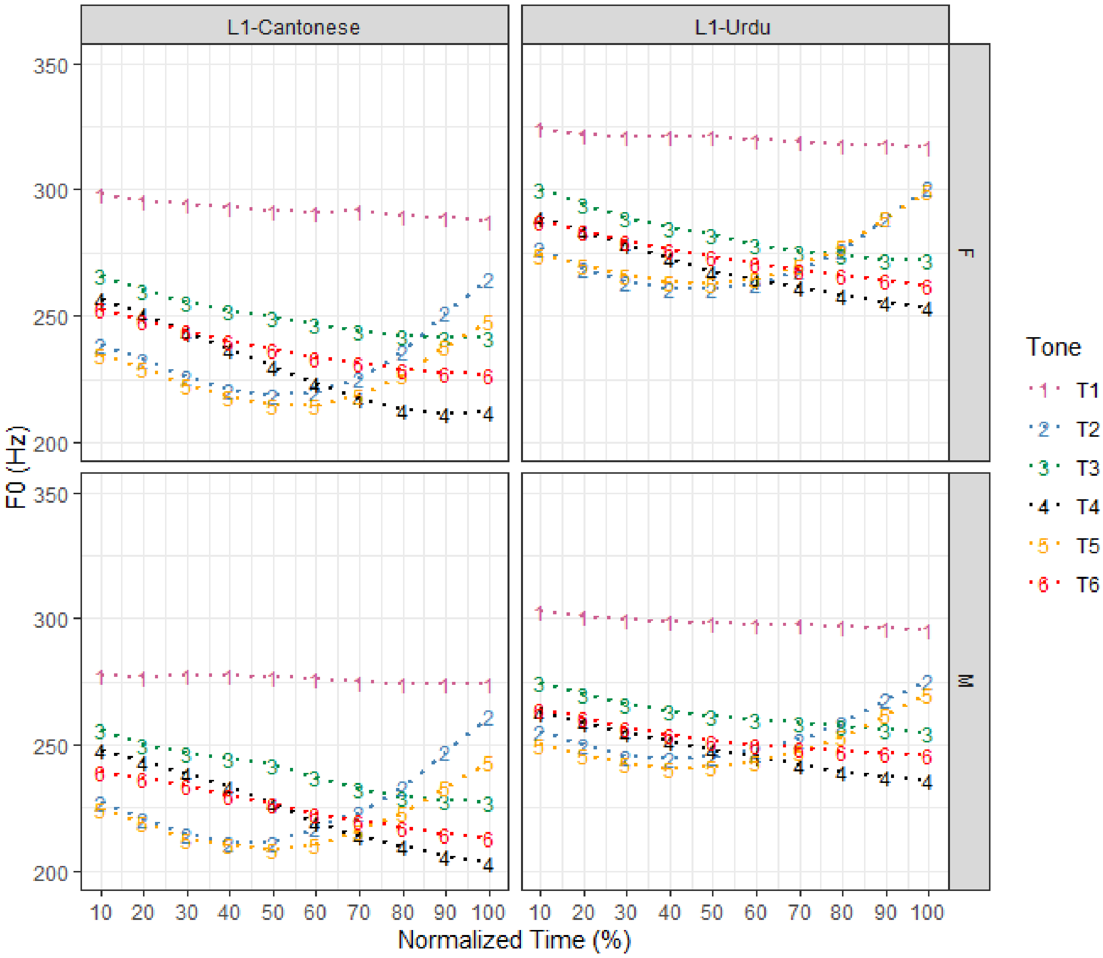
The models also revealed a strong effect of speaker group, with L1-Urdu productions being less accurate than L1-Cantonese ones in all the models (all $|z| \geq 2.5$, $p \leq .01$), consistent with the overall accuracy difference between the two speaker groups. Only models of T3 and T5 showed significant effects of WordLen (T3, T5) and Task (T3), with tokens from longer (disyllabic) embedding words and the repetition task being more accurate than those from shorter (monosyllabic) words and the naming task, respectively.

Tone spaces of L1-Urdu and L1-Cantonese speakers

Figure 3 plots the average tonal contours—with and without pitch normalization—of each tone by L1-Urdu and L1-Cantonese speakers, based on perceptually correct productions only. One may notice that the plotted tonal contours were in a compressed range of T, roughly between 2.5 and 4. This is probably because of a small number of unusually high- or low-pitched tokens, which would greatly shift the upper/lower bound of the pitch range used for T normalization although their effect on average contours was minimal. Overall L1-Urdu participants showed greater overlap among non-high tones (T2-T6) than L1-Cantonese counterparts. Another interesting observation from Figure 3b is that L1-Urdu participants had higher pitch, ~20-30Hz higher on average than L1-Cantonese counterparts (see Table 1; significant difference for each tone, with all $|t| > 2.45$, all $p < .019$). We return to this in the *Discussion* section.



(a) Tone space in normalized F0 (T).



(b) Tone space in raw F0.

Figure 3. Tonal space of L1-Cantonese and L1-Urdu children in (a) normalized F0 and (b) raw F0, based on correct productions only.

	T1	T2	T3	T4	T5	T6
L1-Can	248.12 (23.23)	241.50 (21.16)	240.22 (25.14)	239.08 (25.19)	244.85 (23.81)	238.19 (20.56)
L1-Urdu	278.24 (32.36)	272.81 (32.91)	267.50 (34.21)	265.68 (35.15)	268.33 (37.73)	261.58 (37.55)

Table 1. Mean F0 in Hz (SD) by tone by speaker group.

Analysis of confusable tone pairs

To better understand the error patterns in the production data, we conducted a confusability analysis, using the raters' transcription of the closest resembling tone for erroneous productions. Table 2 shows the mean rates of mispronouncing a target tone as another, averaged across the two raters' transcriptions.

		Mispronounced as											
		By L1-Cantonese participants						By L1-Urdu participants					
		T1	T2	T3	T4	T5	T6	T1	T2	T3	T4	T5	T6
Target tone	T1	-	0	0.9	0	0	0.5	-	0	3.9	1.0	0	2.3
	T2	0	-	0	0	2.3	0.5	1.0	-	0.4	3.5	7.5	1.0
	T3	0.9	0.5	-	1.1	0.7	2.5	1.6	1.2	-	8.3	2.6	9.3

	T4	0	0	0.7	-	0	1.5	0.7	0.3	6.1	-	4.0	10.3
	T5	0	2.5	0.1	0.7	-	0	0.3	8.1	1.9	8.1	-	2.4
	T6	0	0.1	0.3	2.9	0.4	-	0.3	1.0	2.4	3.2	0.9	-

Table 2. Average percentages of target tone productions that were mispronounced as another tone, separated by tone by speaker group.

As shown in Table 2, L1-Urdu participants had higher error rates and more diverse errors than L1-Cantonese participants. We further extracted major error patterns for each speaker group, which were defined as mispronunciation patterns that accounted for more than 20% of the errors regarding the target tone in both raters' transcription (20% is the chance level, assuming a tone is mispronounced randomly as one of the other five tones). As shown in Table 3, T1, T2 and T5 have identical major mispronunciation patterns across speaker groups: T1 tends to be mispronounced as the other two level tones (T3 and T6); the two rising tones, T2 and T5, are mutually confusable, while T5 also tends to be mispronounced as low-falling T4. These patterns concur with the acoustic analysis presented above. The mispronunciation patterns regarding T3, T4, and T6 reveal greater distinction across speaker groups: L1-Urdu participants showed confusion in all possible directions (i.e. T3 → T4, T3→T6, T4→T3, T4→T6, T6→T3, T6→T4), while L1-Cantonese participants' errors were more focused (i.e. T3→T6, T4→T6, T6→T4) and patterned closely with the documented tone mergers (T3/T6, T4/T6; see Mok et al., 2013). We return to this in the *Discussion* section.

	L1-Cantonese	L1-Urdu
Major mispronunciation patterns (→ means "is mispronounced as")	T1 → T3, T1 → T6 T2 → T5 T3 → T6 T4 → T6 T5 → T2, T5 → T4 T6 → T4	T1 → T3, T1→ T6 T2 → T5 T3 → T4, T3 → T6 T4 → T3, T4 → T6 T5 → T2, T5 → T4 T6 → T3, T6 → T4

Table 3. Major mispronunciation patterns by speaker group.

Discussion

In this study, we collected single word productions with a picture naming experiment from two groups of preschool children in Hong Kong: L1-Urdu children who acquire Cantonese as an L2 and societal language, and age-matched L1-Cantonese children as the control group. All the productions were perceptually evaluated for tone accuracy.

The results show that both L1-Urdu and L1-Cantonese children at ages 4-6 are prone to erroneous tone productions. The specific acoustic features that distinguish correct and incorrect productions vary across tones, but overall, the errors are attributable to inaccurate pitch height (min, max and mean pitch) or pitch contour (direction and magnitude of pitch change), both of which are critical for Cantonese tone perception (see Khouw & Ciocca, 2007), suggesting a link between perception and production in tone acquisition (see Mok et al., 2019).

Our first hypothesis of L1-Urdu children having lower tone production accuracy is borne out by multiple between-group comparisons. In all six tones, L1-Urdu participants are significantly less accurate than L1-Cantonese counterparts, consistent with previous findings of early developmental delay in bilinguals (K&S, M&L). The finding of high tone accuracy in L1-Cantonese children (>90%) patterns with previous studies supporting early tone acquisition (So & Dodd, 1995; To et al., 2013), but not with those suggesting a more protracted tonal development (Mok et al., 2019; Wong & Leung, 2018).

Our second hypothesis regarding possible sources of influence for L1-Urdu children's tone development is largely supported by the current data, although the evidence is compatible with multiple accounts. The great confusion between the two rising tones, T2 and T5, is compatible with the prediction that both tones are assimilated to the Urdu LH intonation category—more likely via *Single Category* assimilation, which is associated with greater perceptual difficulty. Meanwhile, the confusion can also be explained by the ongoing T2-T5 merger and the accompanying reduced distinction from the input.

Regarding the other four tones (T1, T3, T4, T6), the PAM-S predicts them to be *Uncategorized* or *Non-Assimilable* into Urdu prosody, depending on whether the tone in question is acquired as a linguistic (*Uncategorized*) or non-linguistic (*Non-Assimilable*) unit. Neither route facilitates the formation of linguistic categories, but the acquisition of high-level T1 was greatly aided by its perceptual distinctiveness as the only tone in the high pitch range. Unsurprisingly, L1-Urdu participants' production of T1 was on par with that of L1-Cantonese controls; by contrast, L1-Urdu acquisition of T3, T4 and T6—all of which are in the mid-low pitch range and have similar contours—showed both low accuracy and a high degree of confusion, probably under concomitant influence from being *Uncategorized/Non-Assimilable*, lack of salient acoustic distinction, and ongoing tone mergers. It should be noted that the prediction of the *Uncategorized/Non-Assimilable* status of these four tones is based on our current understanding of Urdu prosody, and thus should be revisited if the knowledge is updated.

To be fair, L1-Cantonese children's acquisition is affected by tone mergers and reduced perceptual distinction as well, as shown by their relatively weak acquisition of T2/T5 and T3/T4/T6 (see also Lee et al., 2015). However, compared with L1-Urdu participants, L1-Cantonese participants' error rates are consistently lower, and their error patterns more closely trace the ongoing tone mergers. One may ask if it is possible that L1-Urdu children are somehow more affected by ongoing mergers, maybe because of the nature of their Cantonese input. This speculation is hardly substantiated by the current evidence. Although L1-Urdu children have less Cantonese input, as reflected by their smaller vocabulary size, the input should nonetheless be more homogeneous in nature, as it is predominantly from kindergarten teachers. If so, L1-Urdu children's input likely has less variability and more hyperarticulation, both of which should alleviate the influence of tone mergers.

On a related note, we also observed that L1-Urdu children had significantly higher F0 when producing Cantonese tones than L1-Cantonese children. In fact, both raters commented voluntarily on the unusually high pitch—which almost sounded unnatural—of some L1-Urdu participants. We offer three possible accounts for the observation. First, as mentioned above, L1-Urdu children's Cantonese input mainly comes from kindergarten teachers, most of whom are female. Second, L1-Urdu children may use pitch range to separate their two languages, using high pitch for Cantonese and lower pitch for Urdu. Relatedly, they may be in a mode of “performing”—as opposed to speaking—when producing the nonnative language, and hence exaggerate tones with elevated pitch.

To sum up, both L1-Urdu and L1-Cantonese children’s tone developmental pathways are influenced by the general phonetic properties of the Cantonese tones and ongoing tone mergers. L1-Urdu children face additional difficulty because Urdu prosody lacks correspondence with Cantonese tone categories.

The current results provide the first empirical evidence for reduced production accuracy in the majority language by preschool-age heritage speakers, compared with L1-majority language counterparts. Follow-up studies tapping into perception and the perception-production link in these children are much needed, because the findings allow a collective understanding of how children develop abstract phonological categories in early development when influence from the heritage(non-tone) language is likely prominent. While the current production results seem to pattern with similar findings of proficiency deficiency in other early bilingual acquisition contexts, heritage speakers’ acquisition of the majority language has a unique trajectory, with the speaker (more likely) achieving native-like proficiency in the end, which is often not the case for other types of L2 acquisition. Thus, we may ask: how do heritage speakers catch up with L1 speakers in the acquisition of the majority language? Does development of the majority language interact with decline (or retainment) of the heritage language? These questions should be addressed in future research to further our understanding of heritage speakers' language development.

Conclusion

In line with previous studies (K&S, M&L), we show that bilingual children aged 4 to 6 have smaller vocabulary and are significantly less proficient in producing Cantonese tones than their L1-Cantonese peers. Specifically, the production patterns of bilingual children reflect multiple, concurrent sources of influence that are either specific to their (non-tone) L1 background or universally available for L2 and native speakers alike. Findings from this study contribute to the general understanding of the language development of Urdu heritage children in Hong Kong, and open a new line of inquiry into the developmental trajectory of heritage language speakers' acquisition of the majority language.

Declaration of Conflicting Interests

The authors declare no conflicting interests.

References

Au, T. K., Knightly, L. M., Jun, S., & Oh, J. S. (2002). Overhearing a language during childhood. *Psychological Science*, *13*(3), 238–243.

Bates, D., Maechler, M., Bolker, B., & Walker, S. (2015). Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software*, *67*(1), 1–48.

Best, C. T. (1995). A Direct realist view of cross-Language speech perception. In W. Strange (Ed.), *Speech Perception and Linguistic Experience: Issues in Cross-Language Research*.

Best, Catherine T, & Tyler, M. D. (2007). Nonnative and second-language speech perception: Commonalities and complementarities. In O. S. Bohn & M. Munro (Eds.), *Second-language Speech Learning: The Role of Language Experience in Speech Perception and Production. A Festschrift in Honour of James E. Flege* (pp. 13–34).

Boersma, P., & Weenink, D. (2017). *Praat: Doing phonetics by computer*. Retrieved from http://www.praat.org/

Census and Statistics Department of HKSAR. (2011). *Thematic Report: Ethnic Minorities*.

Chang, C. B., & Yao, Y. (2016). Toward an understanding of heritage prosody: Acoustic and perceptual properties of tone produced by heritage, native, and second language speakers of Mandarin. *Heritage Language Journal*, *13*(2), 134–

160.

Chang, C. B., & Yao, Y. (2019). Production of neutral tone in Mandarin by heritage, native and second language speakers. *Proceedings of the International Conference on Phonetic Siences (ICPhS) 2019*.

Chang, C. B., Yao, Y., Haynes, E. F., & Rhodes, R. (2011). Production of phonetic and phonological contrast by heritage speakers of Mandarin. *The Journal of the Acoustical Society of America*, *129*(6), 3964–3980.

Chang, Y. S., Yao, Y., & Huang, B. H. (2017). Effects of linguistic experience on the perception of high-variability non-native tones. *The Journal of the Acoustical Society of America*, *141*(2), EL120–EL126.

Cheung, P. S. P., Lee, K. Y. S., & Lee, L. W. T. (1997). The development of the 'Cantonese Receptive Vocabulary Test' for children aged 2–6 in Hong Kong. *International Journal of Language & Communication Disorders*, *32*(1), 127–138.

Dunn, L. M., & Dunn, D. M. (2007). *PPVT-4: Peabody Picture Vocabulary Test*. Boulder, CO: Pearson Assessments.

Flege, J. E. (1996). English vowel production by Dutch talkers: more evidence for the "similar" vs "new" distinction. In A. R. James & J. Leather (Eds.), *Second-language speech: Structure and process* (pp. 11–52).

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

268–294.
Hallé, P. A., Chang, Y., & Best, C. T. (2004). Identification and discrimination of Mandarin Chinese tones by Mandarin Chinese vs . French listeners. *Journal of Phonetics*, *32*, 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. https://doi.org/10.1016/j.wocn.2011.11.001
Holm, A., & Dodd, B. (1999). A longitudinal study of the phonological development of two Cantonese–English bilingual children. *Applied Psycholinguistics*, *20*(3), 349–376.
Holm, A., & Dodd, B. (2006). Phonological development and disorder of bilingual children acquiring Cantonese and English. In H. Zhu & B. Dodd (Eds.), *Phonological Development and Disorders in Children: A Multilingual Perspective* (pp. 286–325). Clevedon: Multilingual Matters.
Hong Kong Unison. (2012). *Report: Survey on Kindergarten Education for Ethnic Minority Students in Hong Kong*. (April).
Hussain, S. (1997). *Phonetic correlates of lexical stress in Urdu*. Northwestern University.
Jabeen, F. (2010). *An acoustic study of the influence of Urdu on the intonation patterns of English in Pakistan*. Government College University Faisalabad.

Kan, R. T. Y., & Schmid, M. S. (2019). Development of tonal discrimination in young heritage speakers of Cantonese. *Journal of Phonetics*, *73*, 40–54.

Khouw, E., & Ciocca, V. (2007). Perceptual correlates of Cantonese tones. *Journal of Phonetics*, *35*, 104–117.

Law, N. C. W., & So, L. K. H. (2006). The relationship of phonological development and language dominance in bilingual Cantonese-Putonghua children. *International Journal of Bilingualism*, *10*(4), 405–427.

Lee, K. Y. S., Chan, K. T. Y., Lam, J. H. S., Van Hasselt, C. A., & Tong, M. C. F. (2015). Lexical tone perception in native speakers of Cantonese. *International Journal of Speech-Language Pathology*, *17*(1), 53–62.

Li, D. C. S., & Chuk, J. Y. P. (2015). South Asian students' needs for Cantonese and written Chinese in Hong Kong: a linguistic study. *International Journal of Multilingualism*, *12*(2), 210–224.

Lloyd-Smith, A., Einfeldt, M., & Kupisch, T. (in press). Italian-German bilinguals: The effects of heritage language use on accent in early-acquired languages. *International Journal of Bilingualism*.

Mok, Peggy P. K., Zuo, D., & Wong, P. W. Y. (2013). Production and perception of a sound change in progress: Tone merging in Hong Kong Cantonese. *Language Variation and Change*, *25*(03), 341–370.

Mok, Peggy Pik Ki, Fung, H. S. H., & Li, V. G. (2019). Assessing the Link Between

Perception and Production in Cantonese Tone Acquisition. *Journal of Speech, Language and Hearing Research*, *62*(5), 1243–1257.

Mok, Peggy Pik Ki, & Lee, A. (2018). The acquisition of lexical tones by Cantonese-English bilingual children. *Journal of Child Language*, 1–20.

Oh, J. S., Au, T. K., & Jun, S. A. (2010). Early childhood language memory in the speech perception of international adoptees. *Journal of Child Language*, *37*, 1123–1132.

Ohala, M. (1986). A search for the phonetic correlates of Hindi stress. In B. Krishnamurti (Ed.), *South Asian Languages: structure, convergence, and diglossia* (pp. 81–92). Delhi: Motilal Banarsidass.

Polinsky, M. (2018). *Heritage languages and their speakers*. Cambridge University Press.

Polinsky, M., & Kagan, O. (2007). Heritage languages : In the 'wild' and in the classroom. *Language and Linguistics Compass*, *1*(5), 368–395.

Puri, V. (2013). *Intonation in Indian English and Hindi late and simultaneous bilinguals*. University of Illinois at Urbana-Champaign.

R Core, T. (2018). *R: A language and environment for statistical computing*. Retrieved from https://www.r-project.org/

Rao, R. (2015). Manifestations of /bdg/ in Heritage Speakers of Spanish. *Heritage Language Journal*, *12*(1), 48–74.

Shi, F. (1986). Tianjin fangyan shuangzizu shengdiao fenxi [An analysis of the bisyllabic tones in Tianjin dialect]. *Yuyan Yanjiu*, *1986.1*, 77–90.

Shum, M. S. K., Gao, F., Tsung, L., & Ki, W. W. (2011). South asian students' Chinese language learning in Hong Kong: Motivations and strategies. *Journal of Multilingual and Multicultural Development*, *32*(3), 285–297.

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.

So, C. K., & Best, C. T. (2011). Categorizing mandarin tones into listeners' native prosodic categories: The role of phonetic properties. *Poznan Studies in Contemporary Linguistics*, 133–145.

So, C. K., & Best, C. T. (2014). Phonetic influences on english and french listeners' assimilation of mandarin tones to native prosodic categories. *Studies in Second Language Acquisition*, *36*(2), 195–221.

So, L. K. H., & Dodd, B. J. (1995). The acquisition of phonology by Cantonese-speaking children. *Journal of Child Language*, *22*(3), 473–495.

To, C. K. S., Cheung, P. S. P., & Mcleod, S. (2013). A population study of children's acquisition of Hong Kong Cantonese consonants, vowels, and tones. *Journal of Speech, Language and Hearing Research*, *56*, 103–122.

Wang, Y., Jongman, A., & Sereno, J. A. (2003). Acoustic and perceptual evaluation of

Mandarin tone productions before and after perceptual training. *The Journal of Acoustical Society of America*, *113*(2), 1033–1043.

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.

Wong, P. (2013). Perceptual evidence for protracted development in monosyllabic Mandarin lexical tone production in preschool children in Taiwan. *Journal of Acoustical Society of America*, *133*(1), 434–443.

Wong, P., & Leung, C. T.-T. (2018). Suprasegmental features are not acquired early: Perception and production of monosyllabic Cantonese lexical tones in 4- to 6-year-old preschool children. *Journal of Speech, Language, and Hearing Research*, *61*(5), 1070–1085.

Xu, Y. (2013). ProsodyPro—A tool for large-scale systematic prosody analysis. *Proceedings of Tools and Resources for the Analysis of Speech Prosody (TRASP 2013)*, 7–10. Aix-en-Provence, France.

Appendix A: List of stimuli for the picture naming experiment

Stimuli	Task	IPA	English gloss
電話	Naming	tin6 wa2	“telephone”
星	Naming	sɪŋ1	“star”
匙羹	Naming	ts^{h}i4 kɐŋ1	“spoon”
食嘢	Naming	sɪk6 jɛ5	“eat”
番茄	Naming	fan1 k^{h}ɛ2	“tomato”
樹	Naming	sy6	“tree”
馬	Naming	ma5	“horse”
兔	Naming	t^{h}ou3	“rabbit”
船	Naming	syn4	“ship”
鞋	Naming	hai4	“shoes”
石	Naming	sɛk6	“stone”
餅	Naming	pɛŋ2	“biscuit”
燈	Naming	tɐŋ1	“lamp”
杯	Naming	pui1	“cup”
書	Naming	sy1	“book”
龜	Naming	kwɐi1	“turtle”
信	Naming	sɵn3	“letter”
凳	Naming	tɐŋ3	“chair”
郵票	Naming	jɐu4 p^{h}iu3	“stamp”
錶	Naming	piu1	“watch”
繩	Naming	sɪŋ2	“rope”
月亮	Naming	jyt6 lœŋ6	“moon”
耳	Naming	ji5	“ear”
肚	Naming	t^{h}ou5	“belly”
唇	Naming	sɵn4	“lips”
豆	Naming	tɐu2	“bean”
鼓	Naming	ku2	“drum”
奶	Naming	nai5	“milk”
魚	Naming	jy2	“fish”
菜	Naming	ts^{h}ɔi3	“vegetable”
護士	Naming	wu6 si6	“nurse”

薯	Naming	sy4	“potato”
蟹	Naming	hai5	“crab”
碟	Naming	tɪp2	“plate”
貓	Naming	mau1	“cat”
瞓覺	Naming	fɐn3 kau3	“sleep”
喵	Naming	mɛu1	“meow”
尾	Naming	mei5	“tail”
彩虹	Naming	ts^{h}ɔi2 hʊŋ4	“rainbow”
綠色	Naming	lʊk6 sɪk1	“green”
大象	Naming	tai6 tsœŋ6	“elephant”
四	Naming	sei3	“four”
二	Naming	ji6	“two”
妹妹	Naming	mui4 mui2	“younger sister”
老虎	Naming	lou5 fu2	“tiger”
雨	Naming	jy5	“rain”
遮	Naming	tsɛ1	“umbrella”
女仔	Naming	nɵy5 tsɐi2	“girl”
唔係	Naming	m4 hɐi6	“no”
炒	Naming	ts^{h}au2	“fry”
錫	Naming	sɛk3	“kiss”
唱歌	Naming	ts^{h}œŋ3 kɔ1	“sing”
游水	Naming	jɐu4 sɵy2	“swim”
跳繩	Naming	t^{h}iu3 sɪŋ2	“skip”
吹	Naming	tsɵy1	“blow”
病	Naming	pɛŋ6	“sick”
損	Naming	syn2	“injured”
企	Naming	k^{h}ei5	“stand”
重	Naming	ts^{h}ʊŋ5	“heavy”
肚餓	Naming	t^{h}ou5 ŋɔ6	“hungry”
學校	Naming	hɔk6 hau6	“school”
早晨	Naming	tsou2 sɐn4	“good morning.”
唔該	Naming	m4 kɔi1	“thank you.”
出面	Naming	ts^{h}ɵt1 min6	“outside”

燈	Repetition	tɐŋ1	“lamp”
凳	Repetition	tɐŋ3	“chair”
書	Repetition	sy1	“book”
樹	Repetition	sy6	“tree”
錫	Repetition	sɛk3	“kiss”
石	Repetition	sɛk6	“stone”
餅	Repetition	pɛŋ2	“biscuit”
病	Repetition	pɛŋ6	“sick”
損	Repetition	syn2	“injured”
船	Repetition	syn4	“ship”
魚	Repetition	jy2	“fish”
雨	Repetition	jy5	“rain”
薯（片）	Repetition	sy4	“potato”
樹	Repetition	sy6	“tree”
鞋	Repetition	hai4	“shoes”
蟹	Repetition	hai5	“crab”
耳	Repetition	ji5	“ear”
二	Repetition	ji6	“two”
星	Repetition	sɪŋ1	“star”
繩	Repetition	sɪŋ2	“rope”
兔	Repetition	t^hou3	“rabbit”
肚	Repetition	t^hou5	“belly”

Appendix B: Summaries of the fixed effects in generalized mixed-effects models on accuracy.

Model on T1 accuracy

Predictor	*β*	z	*p*
MeanT	1.49	5.20	< .001
SubjGroup (=L1-Urdu)	-1.83	-2.82	.005

Model on T2 accuracy

Predictor	*β*	z	*p*
MaxT	0.32	2.09	0.04
MaxTPos	0.14	4.26	< .001
SubjGroup (=L1-Urdu)	-1.92	-6.20	< .001

Model on T3 accuracy

Predictor	*β*	z	*p*
MeanT	0.46	2.16	0.03
SubjGroup	-1.67	-4.45	< .001

(=L1-Urdu)			
WordLen	0.88	2.79	.005
Task (=Repetition)	1.04	2.68	.007

Model on T4 accuracy

Predictor	*β*	z	*p*
MinT	-0.62	-2.44	0.01
ExcursionT	-0.65	-2.89	.003
Dur	-1.33	-2.86	.004
SubjGroup (=L1-Urdu)	-2.88	-5.11	< .001

Model on T5 accuracy

Predictor	*β*	z	*p*
ExcursionT	0.28	2.26	.02
Dur	0.68	2.06	.04
SubjGroup (=L1-Urdu)	-1.89	-5.05	< .001

WordLen	0.92	2.45	.01

Model on T6 accuracy

Predictor	*β*	z	*p*
MinT	0.32	2.38	.02
SubjGroup (=L1-Urdu)	-0.73	-2.56	.01