

ARTICLE

The acquisition of lexical tones by Cantonese–English bilingual children

Peggy Pik Ki MOK^{1*} and Albert LEE^{2,1}

¹Department of Linguistics and Modern Languages, Leung Kau Kui Building, The Chinese University of Hong Kong, Shatin, Hong Kong and ²Department of English Language, University of Tokyo, Komaba 3-8-1, Tokyo 153-8902, Japan

*Corresponding author. E-mail: peggymok@cuhk.edu.hk

(Received 30 March 2017; revised 1 April 2018; accepted 30 May 2018;
first published online 1 August 2018)

Abstract

Previous studies on bilingual children found intact tonal development at the initial stages of interaction between Cantonese and English in successive bilingual children, whereas children exposed to both languages from birth have not been studied in this regard. We examined the production of Cantonese tones by five simultaneous bilingual children longitudinally at 2;0 and 2;6, and compared them with age-matched monolingual children using auditory analysis. Our results showed that some bilingual children had a delay at 2;0, compared to their monolingual peers. Some bilingual children also exhibited a ‘high–low’ template in their production, resembling the pitch pattern of English trochaic words. These findings suggest a possible early interaction of the Cantonese and English prosodic systems in which bilingual children adopted the English stress pattern in Cantonese production. The time-point along the trajectory of phonological development is important in modulating whether cross-linguistic transfer can be observed.

Keywords: sequential/successive bilingual; transfer; trochaic pattern

This study investigated longitudinally how Cantonese–English simultaneous bilingual children acquired lexical tones in Cantonese, and compared them with monolingual Cantonese children at two ages when prosody is developing rapidly: 2;0 and 2;6. Early research on bilingual phonological acquisition has often focused on segmental aspects, e.g., phonemic inventory and error patterns (e.g., Johnson & Wilson, 2002; Kehoe, 2002; Kehoe, Lleó, & Rakow, 2004). In recent years, more studies have examined the prosodic aspects of bilingual phonological acquisition, e.g., lexical stress (Paradis, 2001; Li & Mok, 2014) and speech rhythm (Bunta & Ingram, 2007; Mok, 2011, 2013). Nevertheless, to the best of our knowledge, there is still no published study on lexical tone development of bilingual children acquiring a tone language and a non-tone language simultaneously, although a few studies have investigated lexical tone development of successive bilingual children and reported no English influence on tone development (Holm & Dodd, 1999, 2006). Our study thus bridges an important gap in our understanding of bilingual interaction in early simultaneous

development of phonology by investigating this unexplored area with a language pair that differs typologically: Cantonese and English.

Acquisition of Cantonese lexical tones

The use of lexical tone (T) is a salient phonological characteristic of Cantonese. Each syllable (usually corresponding to a morpheme) carries a tone. Cantonese has a complex tone system. There are six distinct lexical tones based on pitch contrast alone: T1 [55] high-level, T2 [25] high-rising, T3 [33] mid-level, T4 [21] low-falling, T5 [23] low-rising, and T6 [22] low-level (Fok-Chan, 1974; Bauer & Benedict, 1997). The numbers in [] represent the relative starting and ending pitch height of each tone, with 5 being the highest and 1 being the lowest pitch height of a speaker's normal pitch range (Chao, 1930, 1947). The six lexical tones can be divided into two registers, T1, T2, and T3 in high register and T4, T5, and T6 in low register (Yip, 2002). The six tones appear in open syllables or syllables with nasal endings [-m, -n, -ŋ]. There are three allotones which are traditionally called the 'entering tones' in Chinese phonology. They only appear in syllables ending with unreleased stops [-p, -t, -k]: T7 [5] high-stopped, T8 [3] mid-stopped, and T9 [2] low-stopped. They are much shorter in duration and are considered allotones of the three corresponding unstopped level tones T1, T3, and T6, respectively (Chao, 1947; Bauer & Benedict, 1997).

Several studies based on auditory analysis have shown that Cantonese monolingual children have acquired all the six tones very early, by 2;0 (Tse, 1978; So & Dodd, 1995) or 2;6 (To, Cheung, & McLeod, 2013). The longitudinal conversational data of one child in Tse (1978) between 1;3 and 2;6, and those of four children aged 1;2 to 2;0 in So and Dodd (1995) showed that they had acquired all Cantonese tones by 2;0. Tse (1978) divided the acquisition of tone production in three stages: in Stage 1 (1;2–1;4) T1 [55] and T4 [21] were acquired; in Stage 2 (1;5–1;8) T3 [33], T2 [25], and the three allotones were acquired; in Stage 3 (1;9) T5 [23] and T6 [22] were acquired. The acquisition of the first to last tones covered a period of only eight months. The four children in So and Dodd (1995) had a similar pattern of order and rate of acquisition. They reported that the children acquired T1 [55] and T3 [33] first, then T2 [25] and the three allotones. Two children acquired T6 [22] before T4 [21] and T5 [23], while one child showed the opposite pattern. Another child acquired these three tones simultaneously. Their data showed that all four children had acquired the Cantonese tones by 2;0, although the specific order of acquisition might differ.

Cross-sectional data of many more children with elicited production present a similar picture. So and Dodd (1995) tested 268 Cantonese-speaking children aged 2;0 to 6;0. They found that only two children made tone errors, one four-year-old made two errors and a five-year-old made three errors. They concluded that by 2;0 most children had mastered the tonal contrasts in Cantonese. The large-scale study by To *et al.* (2013) tested 1,726 children aged 2;4 to 12;4 and also echoes their findings. To *et al.* found that for the youngest age group (2;4–2;9, 104 children), the averaged accuracy was already at ceiling (mean 98.02%, SD 5.19%). As there were no data before 2;4 in their study, they concluded that tone acquisition is complete by age 2;6. Both So and Dodd (1995) and To *et al.* (2013) have found that Cantonese-speaking children had finished acquiring tones well before consonants and vowels.

If the complex Cantonese tone system is acquired so early by monolingual children, how about the tonal development in children who acquire Cantonese and a non-tone

language (English) bilingually? So far, there is no published study on tone acquisition by simultaneous Cantonese–English bilingual children, so it is still an unknown. Nevertheless, a number of studies of successive Cantonese–English bilinguals using auditory analysis can give us some insights into this important yet unexplored question.

Holm and Dodd (1999) documented the speech development of two successive Cantonese–English bilingual children who were exposed to an English immersion environment (childcare centres in Australia) at 2;0 and 2;6, respectively. Since they were in an exclusively Cantonese-speaking environment before immersion, and since monolingual children have acquired the Cantonese tones by 2;0 or 2;6, as discussed above, it is not at all surprising that the Cantonese tones of these two successive bilingual children were found to be intact during the assessment periods of 2;3 to 3;1 and 2;9 to 3;5, respectively. Holm and Dodd (1999, p. 355) simply said that tone accuracy was monitored, but that errors were infrequent. Additionally, Holm and Dodd reported cross-sectional data from 40 Cantonese–English successive bilingual children aged 2;2–5;7 in Australia. They found that generally there was no difference between monolingual Cantonese and bilingual Cantonese, and there were only five atypical tonal errors (defined as errors used by less than 10% of the monolingual population). Both longitudinal and cross-sectional data point to the conclusion that late exposure to English does not affect Cantonese tones of successive bilingual children adversely, at least in the first years of English exposure.

Although the tones of the successive Cantonese–English bilingual children appeared not to be affected by their exposure to English after 2;0, their segmental development in the two languages was not perfect. Both the longitudinal data in Holm and Dodd (1999) and the cross-sectional data in Holm and Dodd (2006) suggested that these children exhibited segmental error patterns which were considered atypical for monolingual children acquiring the same two languages, even reminiscent of language disorders. They suggested that these children underwent a developmental period characterized by underspecified phonological realization rules when the two languages interact initially. The process of acquiring two phonological systems bilingually is different to the process of acquiring each system monolingually.

Holm and Dodd's findings of atypical segmental development at the initial stages of the interaction between Cantonese and English in successive bilingual children are highly relevant to our study on simultaneous bilingual acquisition of tone. The disparate patterns of intact tone production versus atypical segmental errors following the introduction of English can be attributed to the different timings of phonological development of tones and segments in monolingual children. As discussed above, both So and Dodd (1995) and To *et al.* (2013) confirmed that tones were acquired well before segments in Cantonese monolingual children. So when English was introduced after the completion of tone acquisition for Cantonese–English successive bilingual children, their tones were not adversely affected. However, the introduction of English coincided with the ongoing development of segments in both languages for these children, i.e., their segments were still incompletely acquired and were in a state of flux. This explains why their segments were affected and resulted in atypical error patterns.

Such findings raise the question of what would happen if English was introduced during a period when Cantonese tones were still developing in successive bilingual children. Following the above argument, we would expect to see atypical tone error patterns. This is exactly what was found. Light (1977) reported a case study of his Cantonese-speaking daughter Claire who moved with him to the US at 16 months.

She was in a predominantly Cantonese environment before 16 months. Observations of her Cantonese production from 19 months onward showed “disintegration of her tonal system” (Light’s wording). Light’s examples included changing [25] to [55]; ‘flattening out’ of the high–rising tone [25] in some words; alternation between the high and low registers for some tones, etc. He said that such tone errors abounded, but by the age of four, most traces of her tonal disintegration were gone and her tones were again intact. Even at around six years old, when she resisted speaking Cantonese, her tones still sounded reasonably accurate. Light (1977, p. 265) attributed her tonal disintegration to the strong influence of English phonology. He cited examples showing that many of the incorrect tonal usages actually reflected a pitch-contour approximation of the English equivalent items (their stress/intonation patterns). He suggested that these approximations of English intonation follow the rules of English loanwords in Cantonese very well. He believed that the later rectification of her tonal disintegration by around age four indicated that for a time her two languages had been confused in her performance as a result of the influx of newly learned English words.

Light’s (1977) observations, despite being a case study, are insightful for our study in illustrating that Cantonese tones can be influenced by English prosody in successive bilingual children whose English exposure coincided with tonal development. As a result, we can naturally expect to see similar interaction in simultaneous bilingual children who were exposed to both languages from birth, because English influence is already present at the incipient stage of tonal development. Nevertheless, so far, there is still no published study investigating this phenomenon. Our study is designed to fill this important gap.

The present study

We examined the accuracy of tone production of five Cantonese–English simultaneous bilingual children longitudinally at 2;0 and 2;6, the period when Cantonese monolingual children have been reported to have completed the acquisition of tone (Tse, 1978; So & Dodd, 1995; To *et al.*, 2013). We compared them with two groups of Cantonese monolingual children, three children longitudinally at 2;0 and 2;6; and ten cross-sectionally at 2;6.

Previous studies showing early acquisition of Cantonese tones by monolingual children discussed above all used transcription data by a native transcriber. However, in a recent paper, Wong, Fu, and Cheung (2017), by adopting more rigorous methods of transcription with low-pass filtered materials by multiple judges and acoustic analysis, demonstrated that three-year-old Cantonese children still had not fully acquired the Cantonese tones in production in that their production accuracy and acoustic patterns were still not adult-like. Similar findings were obtained for Mandarin tone acquisition as well: simple transcription data showed very early acquisition by 2;0 (Zhu & Dodd, 2006) while Wong and colleagues (Wong, Schwartz, & Jenkins, 2005; Wong, 2013) demonstrated a much more protracted acquisition process using their methods.

Wong *et al.*’s findings are important because they revealed a very different picture of the tone acquisition process by monolingual children. It is not surprising that more rigorous methods should reveal slower tone development compared to simple transcription. Unfortunately, we could not adopt their methods in our study for several reasons. First and foremost, the tone data were of a different nature and they

were not comparable. Wong *et al.* collected tone production in an experimental setting using a picture-naming task, i.e., the same set of monosyllabic words produced in isolation in the same session by all children, which rendered them suitable for low-pass filtering and using acoustic analysis, while our data were all from natural conversations between children and different interlocutors in multiple sessions in corpora. Tonal coarticulation occurred frequently in connected speech (Xu, 1997). Furthermore, a diverse set of words with various segmental and intonational contexts were produced by the children in the corpora. The resultant F0 patterns in these words would be quite deviant from the expected canonical patterns, although the tone production sounds natural and appropriate in the original contexts. The recordings we used were made in homes or kindergartens with much background noise. These conditions rendered our data unsuitable for acoustic analysis and transcription with filtering. However, auditory analysis would be less affected by these adverse conditions because human speech perception is very robust even in noisy environments (Lecumberri & Cooke, 2006; Cutler, Lecumberri, & Cooke, 2008; Alwan, Jiang, & Chen, 2011). Also, natural conversational data, although noisy and methodologically problematic in some ways, is also informative in a way that elicited speech is not.

Given the above reasons, and the fact that our study is the first study on bilingual tone acquisition, we considered that it would be best to follow the practice of previous studies on tone acquisition using auditory analysis instead of Wong *et al.*'s methods as a first step for comparability's sake. Nevertheless, previous studies on Cantonese tone acquisition used transcription data by a native judge to assess the accuracy of tone production, but with only a small portion of data cross-checked by another transcriber. For example, the data of 27 out of 268 children in So and Dodd (1995) and 130 out of 1,726 children in To *et al.* (2013) were cross-checked (i.e., ~10%). Only five Cantonese samples in Holm and Dodd (1999) were cross-checked (they did not give details of the total number of samples in their study). Although they all reported high inter-rater agreement (over 90%), there could still be quite a portion of data without agreement if we consider the size of the entire dataset. As a result, we decided to have two independent raters transcribe all tone production by the bilingual and monolingual children in context in our study. This allows us to assess their tone accuracy with two criteria: a lenient criterion in which the tone production was considered accurate by at least one rater; and a more stringent criterion in which correct production was confirmed by both raters. We believe that such a procedure is essential because of the nature of the recordings: corpus conversational data in various contexts spoken by young children with noisy background. Raters may be easily influenced by contextual cues (e.g., emotional and expressive utterances, background noise) which may bias their judgement, especially for those tone productions that are ambiguous (more details will be given in the 'Methods' section).

Given that acquiring two phonological systems simultaneously is different from acquiring each one monolingually, that Cantonese monolingual children have already acquired all tones by 2;6, and that there is no lexical tone in English, we asked the following research questions: (i) Would bilingual children show a delay in their tone acquisition compared to their monolingual counterparts? If so, (ii) when would they catch up? In addition, given that interaction between Cantonese tones and English prosody was reported in one successive Cantonese–English bilingual child resulting

Table 1. Background Information of the Five Bilingual Children

| Child | Native language of mother | Native language of father | Sex | Dominant language | Age range of data used |
|-----------|---------------------------|---------------------------|-----|-------------------|------------------------------------|
| Timmy | Cantonese | British English | M | Cantonese | 2;1.22 ~ 2;1.29 2;5.12 ~ 2;6.19 |
| Sophie | Cantonese | British English | F | Cantonese | 1;11.8 ~ 2;1.20 2;5.2 ~ 2;7.25 |
| Alicia | Cantonese | British English | F | Cantonese | 1;11.5 ~ 2;1.15 2;5.18 ~ 2;7.28 |
| Llywelyn | Cantonese | British English | M | Cantonese | 2;0.12 ~ 2;1.13 2;5.10 ~ 2;7.18 |
| Charlotte | Cantonese | British English | F | English | 1;11.5 ~ 2;1.22 2;5.19 ~ 2;7.23 |

in ‘tonal disintegration’, (iii) would similar English influence on Cantonese tones be observed in simultaneous bilingual children?

Methods

Participants

The five Cantonese–English bilingual children (two boys and three girls) are featured in the Hong Kong Bilingual Child Language Corpus, which is available through the YipMatthews corpus in CHILDES (<https://childes.talkbank.org/access/Biling/YipMatthews.html>). Yip and Matthews (2007) give detailed background for these children. They were children of mixed marriages who were exposed to Cantonese and English from birth, and grew up in a ‘one parent one language’ environment. Four children were Cantonese-dominant and one was English-dominant. Their language dominance was determined objectively by calculating MLU differentials between the two languages, and their language preferences and patterns of code-mixing. Please refer to Yip and Matthews (2007, §3.1) for details of the language background of these bilingual children. They were recorded longitudinally at weekly or bi-weekly intervals in two unstructured play situations on the same day, one for Cantonese and one for English, although some language mixing can be found in the recordings, especially the earlier ones. Table 1 shows the language background of these five bilingual children.

Data of the monolingual children came from two corpora. Ten children at around 2;6 are featured in the HKU-Cantonese-70 corpus available in CHILDES (<<https://childes.talkbank.org/access/Chinese/Cantonese/HKU.html>> Fletcher, Leung, Stokes, & Weizman, 2000). The corpus contains cross-sectional conversational data recorded in kindergartens in Hong Kong (i.e., preschools). Each recording is about 30 minutes long. The age range of the ten monolingual children we used covers one month (2;5.1–2;6.1). Since there is no data at 2;0 in the HKU-Cantonese-70 corpus, we also used longitudinal conversational data of three monolingual children (CGK, CKT, MHZ) at 2;0 and 2;6 in the Hong Kong Cantonese Child Language Corpus (CANCORP; Lee *et al.*, 1996; Lee & Wong, 1998; available at <<http://www.arts.cuhk.edu.hk/~lal/corpora.html#CANCORP>>) to supplement our comparisons. The original corpus contained recordings of eight children (4 female), each of whom was

observed for one year from the time when they were between one and a half to two years old. In our analysis of the CANCELP data, the time-point 2;0 covered recordings from 1;11.13 to 2;0.16; and the 2;6 time-point was from 2;5.0 to 2;6.18. Each recording in CANCELP is about one hour long. As these corpora were inherently different in nature (e.g., duration, number of speakers), their results (production accuracy) would inevitably be different, though comparable since all were natural conversational data. We saw the value of directly comparing them because suitable longitudinal data were otherwise non-existent.

Procedure

We first extracted the online transcripts of the recordings. A list of all syllables produced by the target child in the sound file in question was compiled. For each syllable, line number (approximate location in the original corpus annotation), the corresponding Chinese character, and the citation tone (extracted from Jyutping transliteration) were specified. Each sound file was listened to independently by two native speakers of Hong Kong Cantonese with phonetic training. They distinguished all six tones clearly in both production and perception, i.e., they did not merge any tones (Mok, Zuo, & Wong, 2013). Rater 1 listened to all three sets of recordings. Another rater listened to the bilingual and HKU-70 sets, and a third one only the CANCELP set; these two raters will be collectively referred to as Rater 2 below. The raters took turns to listen to the sound files in context, wearing circumaural headphones, and identified the tonal category they perceived for each syllable, without reference to the judgements of the other rater.

Two criteria were used to determine the accuracy of tone production. A lenient criterion accepts a tone to be correctly produced if either of the two raters judged it to be the intended tone as indicated in the transcript. A more stringent criterion only considers a tone to be correct if both of the raters perceived it as the intended tone.

One type of words required special attention in our calculation of tone accuracy. Sentence-final particles abound in Cantonese (Matthews & Yip, 2011), and their phonetic realizations are variable and subject to influences such as communicative functions (Wu, 2009). Some sentence-final particles can have multiple possible tones depending on discourse function, but in the original transcripts such variations were not fully specified. This means that what sounded natural to the raters could be deemed erroneous if we simply compared it against the citation tone listed in the transcript. It was difficult for the raters to decide whether it was used correctly for various communicative functions by listening to the recordings only. Therefore, all sentence-final particles in the recordings were excluded from subsequent analysis. This treatment affected only 1.6% ($N = 393$ syllables) of the bilingual data, thus its effect of increasing the tone accuracy rates should be minimal; also since the same treatment applies to all children, monolingual and bilingual alike, it should not bias the results unfairly in any direction.

Results

Inter-rater reliability

Table 2 shows the inter-rater reliability of different sets of data. They are all over 90% (Cronbach's alpha) and are comparable to previous studies on Cantonese tone

Table 2. Inter-rater Reliability

| Time-point | Dataset | Agreement (total no. of syllables) | Cronbach's alpha |
|------------|-----------|------------------------------------|------------------|
| 2;0 | Bilingual | 96% (N = 8,141) | 0.972 |
| | CANCORP | 91% (N = 4,242) | 0.961 |
| 2;6 | Bilingual | 96% (N = 15,581) | 0.978 |
| | HKU-70 | 96% (N = 3,704) | 0.976 |
| | CANCORP | 93% (N = 9,907) | 0.964 |

Table 3. Average Tone Production Accuracy under Two Criteria

| Time-point | Criteria | Bilingual | HKU-70 | CANCORP |
|------------|-----------|-----------|--------|---------|
| 2;0 | lenient | 96.78% | – | 95.66% |
| | stringent | 92.88% | – | 88.02% |
| 2;6 | lenient | 97.65% | 98.52% | 98.34% |
| | stringent | 93.53% | 95.03% | 91.69% |

acquisition using transcription data. Nevertheless, in our study, the inter-rater reliability is calculated based on all data, while in previous studies, only a small subset of their data (~10%) were cross-checked by another transcriber.

Overall production accuracy

Table 3 shows the average accuracy of tone production under both lenient and stringent criteria at the two time-points. We can see that the bilingual children were on a par with monolingual children at both ages, with similarly high production accuracy under both criteria. The slightly lower accuracy of the CANCORP data compared to the bilingual and HKU-70 data under the stringent criterion can be explained by its poorer recording quality. This is also supported by the relatively lower inter-rater reliability for CANCORP in **Table 2**. In addition, there were only three children in the CANCORP data, while there were five bilingual children and ten children in the HKU-70 data. Individual variations had a larger impact on the overall accuracy in the CANCORP data than the other two sets of data. Nonetheless, these differences among corpora do not seem to affect the relative production accuracy of individual lexical tones, as will be discussed below.

Figure 1 shows the production accuracy of individual tones by the bilingual and monolingual children at the two time-points under the two judgement criteria. We can see that at 2;0, the development of T2 and T5 was behind the other four tones for both bilingual and monolingual children. At 2;6, all the six tones were well developed using the lenient criterion, but problems with T2 and T5 still remained using the stringent criterion for the bilingual and the monolingual children in CANCORP. If we compare bilingual and monolingual children at 2;0, under the lenient criterion, the bilingual children were not producing T1, T3, T4, and T6 as accurately as the monolingual children, and yet they were slightly better than the

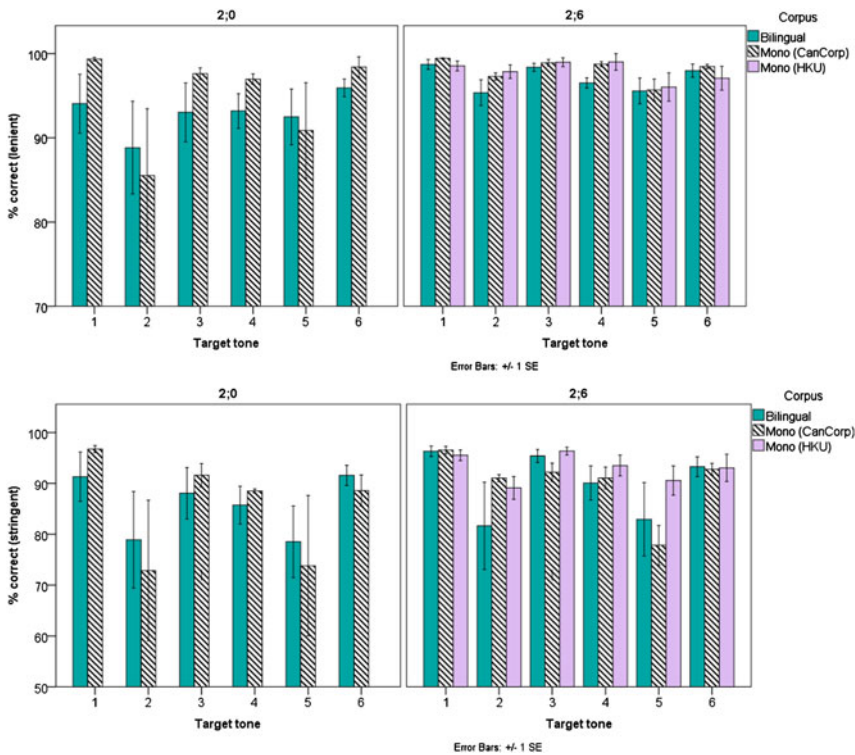


Figure 1. Production accuracy of individual tones by the bilingual and monolingual children at two time-points under two judgement criteria: lenient (upper panel) and stringent (lower panel).

monolingual children for T2 and T5. This explains why the overall accuracy of the bilingual children and the monolingual children were very similar. The same patterns persisted under the stringent criterion, except for T6. The generally longer error bars of the bilingual children indicate that they had more individual variation than the monolingual children.

Individual patterns

Table 4 shows the individual accuracy of the bilingual children at the two time-points. At 2;0, three children, Alicia, Sophie, and Timmy, already had very few errors under both lenient and stringent criteria, while Charlotte and Llywelyn were not as talkative as and made many more errors than the other three children. At 2;6, all five children appeared to have mastered the tones well using the lenient criterion, but Charlotte and Llywelyn still had more than 10% errors under the stringent criterion. It should be noted that while the English-dominant child Charlotte was still not speaking much Cantonese, Llywelyn was already as talkative as the other three Cantonese-dominant children, evidenced by the total number of syllables they produced.

Table 4. Individual Tone Error Patterns of the Bilingual Children under Lenient (Top) and Stringent (Bottom) Criteria

| | 2;0 | | | | | 2;6 | | | | |
|------------------------------|-------------------------|--------------------------|---------------------------|-------------------------|------------------------|-------------------------|--------------------------|---------------------------|-------------------------|------------------------|
| | Alicia | Charlotte | Llywelyn | Sophie | Timmy | Alicia | Charlotte | Llywelyn | Sophie | Timmy |
| Total # of syllables uttered | 2059 | 402 | 393 | 3930 | 1357 | 2741 | 755 | 3531 | 5459 | 3095 |
| Total # of wrong tone | 40 (1.9%) 117 (5.7%) | 55 (13.7%) 84 (20.9%) | 66 (16.8%) 112 (28.5%) | 76 (1.8%) 202 (5.1%) | 22 (1.6%) 66 (4.9%) | 60 (2.2%) 210 (7.7%) | 31 (4.1%) 124 (16.4%) | 132 (3.7%) 418 (11.8%) | 92 (1.7%) 159 (3.0%) | 51 (1.6%) 96 (3.1%) |
| Total | | | 259 (3.2%) 581 (7.1%) | | | | | 366 (2%) 1008 (6%) | | |

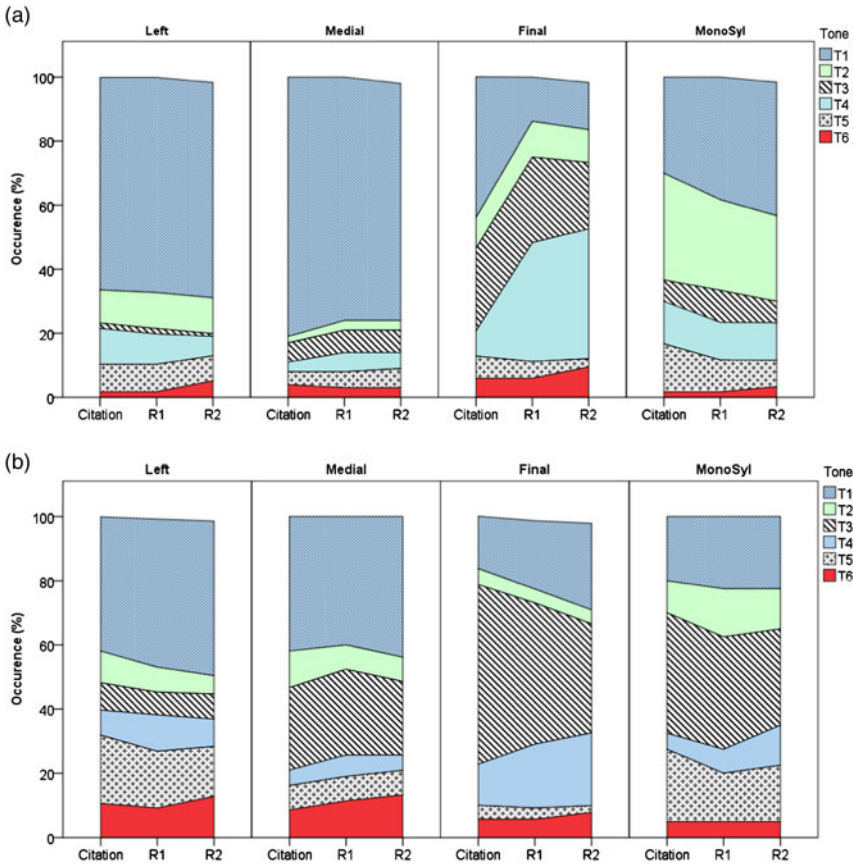


Figure 2. The occurrence of the six tones in different positions spoken by (a) Llywelyn and (b) Charlotte, and the judgements by the two raters.

The low accuracy of Charlotte and Llywelyn at 2;0 prompted us to further examine their tone production in detail. Figure 2 shows the occurrence of the six tones in different positions spoken by the two children, and the judgements by the two independent raters. For illustration, the four panels give the patterns of tones produced as the first syllable of a word, in a word-medial position, as the final syllable of a word, and as stand-alone monosyllables. The three items on the horizontal axis show the percentage of tone occurrence according to the transcripts (citation), and the judgements by the two raters (R1, R2). The total occurrence judged by the two raters falls slightly below 100% because some of the tokens were inaudible due to noisy background or the children speaking too softly. All the inaudible tokens were classified as wrong productions and not assigned a perceived lexical tone. For Llywelyn (Figure 2a), the occurrences of tones as the first syllable and in word-medial position, and to a lesser extent also when the tones were uttered as monosyllables, were quite similar in citation and in the raters' judgements. However, a striking pattern is observed when the tones were produced as the last syllable of a word. There are obvious differences between the occurrence of T1 and

T4 in citation and that judged by the two raters. Specifically, there was a sharp decrease of T1 (with a shrunk area from left to right) and a sharp increase of T4 (with an expanded area from left to right) in actual realization as compared to citation.

The recordings of these special error tokens were then culled and checked. It was found that many of these tokens by Llywelyn had a stable pattern of T1–T4 sequence (i.e., high–low). Most of these errors (N = 29) occurred in reduplicated words (e.g. 波波 bo1bo1 ‘ball ball’ and 車車 ce1ce1 ‘car car’, T1T1 becoming T1T4, transliteration in Jyutping), but they were also observed in non-reduplicated forms (e.g., 香蕉 hoeng1ziu1 ‘banana’ and 單車 daan1ce1 ‘bicycle’, T1T1 becoming T1T4), albeit to a much lesser extent (N = 5).

Charlotte (Figure 2b) had a slightly different error pattern. In word-final position (third panel), the T3 area shrinks towards the right, while the area of T4 expands. This is attributed to the highly recurring error (N = 10) of the reduplicated word 靚靚 leng3leng3 ‘beautiful beautiful’, T3T3 becoming T1T4.

The base patterns of the recurring tone errors of these two bilingual children are different: T1T1 for Llywelyn and T3T3 for Charlotte. Nevertheless, they converge on the same T1T4 ‘high–low’ template, which resembles the stress pattern of trochaic words (stressed–unstressed) in English. Although Cantonese is a tonal language, whereas English has lexical stress with multiple acoustic cues, it has been repeatedly demonstrated that pitch is an important, if not the most important, cue in English stress perception. A higher pitch syllable is perceived as stressed (e.g., Fry, 1967; Cooper, Eady, & Mueller, 1985; Chrabaszcz, Winn, Lin, & Idsardi, 2014). The weight of pitch cues in English thus strengthens this hypothesized connection between the T1T4 ‘high–low’ template and English trochaic stress. It should be pointed out that the ‘high–low’ template was used alongside the correct production of the same word. For instance, Llywelyn produced the correct forms of both reduplicated words (e.g., 車車 ce1ce1 ‘car car’ T1T1) and non-reduplicated words (e.g., 單車 daan1ce1 ‘bicycle’ T1T1) in the same recording, with the forms using the T1T4 ‘high–low’ template. Charlotte also produced one correct form of 靚靚 leng3leng3 (‘beautiful beautiful’ T3T3) in the same recording, together with the many tokens having the ‘high–low’ template. The parallel use of the correct and templatic forms demonstrates the variability in child production.

This is an interesting finding, given the difference in language dominance of the two children: Charlotte was dominant in English while Llywelyn was dominant in Cantonese. Their language dominance was determined objectively using various methods in Yip and Matthews (2007), who specifically commented on Charlotte’s Cantonese production: “Charlotte’s Cantonese shows strong English influence such as producing Cantonese words with non-target tones and sentences with English prosody, sounding very much like a non-native speaker of Cantonese”, and Llywelyn’s English: “Llywelyn’s English shows some of the same features observed in the Cantonese-dominant siblings, such as *wh*-in-situ questions and null objects” (p. 66). Nevertheless, one common characteristic of the two children is that the templatic errors occurred when neither of them was very productive in Cantonese (producing much fewer syllables and making more tone errors than the other three Cantonese-dominant bilingual children; see Table 4).

In addition to the ‘high–low’ template, both bilingual children seemed to use T1 as a ‘default’ option for tone errors as well. In Charlotte’s production, regardless of syllable position, the T1 areas expand rightwards. This is because for many non-recurring (N < 2) errors, T1 was the tone perceived by the raters. Figure 2b shows that in all syllable

Table 5. Tone Error Patterns of Four Bilingual Children at 1;9 under Lenient (Top) and Stringent (Bottom) Criteria

| | Alicia | Charlotte | Sophie | Timmy |
|------------------------------|-------------------------|--------------------------|------------------------|-----------------------|
| Total # of syllables uttered | 344 | 130 | 722 | 574 |
| Total # of wrong tone | 28 (8.1%) 71 (20.6%) | 15 (11.5%) 29 (22.3%) | 16 (2.2%) 49 (6.8%) | 27 (4.7%) 69 (12%) |
| Total | | 86 (4.9%) 218 (12.3%) | | |

positions Charlotte produced more T1 than any other tone. Of the 52 valid counts of errors, 20 (38%) were perceived as T1. Possibly Charlotte might have produced a different tone, but the native raters heard them as T1 nonetheless. For Llywelyn, although Figure 2a seems to suggest that T1 remained stable in non-final positions, it was actually the most frequently perceived tone in all errors, according to the raters' impression. Of the 21 errors in non-final position, 7 (33.3%) were perceived as T1. The number of perceived cases of T1, however, was cancelled out by 9 counts of T1 targets being heard as other tones. Although these 9 errors on T1 constitute only 5.1% of all T1 targets produced in these conditions, graphically they suffice to mask the dominant status of T1 as a 'default' tone in errors.

The observations of these systematic error patterns led us to expand our investigation to the other three bilingual children as well. Although their tone production was judged to be very accurate even at 2;0, we hypothesized that it might be possible to find similar templatic errors in earlier recordings, when they were not as articulate as at 2;0. We examined their tone production at 1;9 following the same procedure we used for the data at 2;0 and 2;6. As there was no recording of Llywelyn at 1;9, we could only examine the recordings of Alicia, Charlotte, Sophie, and Timmy, covering the period between 1;9.10 and 1;10.02.

The inter-rater reliability for the judgements of the 1;9 recordings by the two raters was 92.3% (N = 1487). Table 5 shows the individual tone error patterns of the four children. It can be seen that in addition to Charlotte, Alicia also had quite a lot of tone errors at 1;9. No tone error with the 'high-low' template was found in Charlotte's recording, given the small number of syllables she produced. Instead, we found templatic errors in Alicia's reduplication. Of 25 奶 naai5 'milk' syllables produced by her (12 counts of the reduplicated form 奶奶 naai1naai1 T1T1), 9 exhibited the 'high-low' template, with 8 being perceived as T1T4 and 1 perceived as T1T3. Another example of the 'high-low' template was the word 嗶嗶 bi4bi1 'baby' T4T1 becoming T1T4, which occurred only once. There was no other error in her recording that occurred more than once.

Thus, our hypothesis that templatic errors might be found in earlier recordings of the other three bilingual children was confirmed. Specifically, templatic errors were found in the child Alicia, who was not as talkative as and made more errors than the other two Cantonese-dominant children, i.e., her Cantonese was not as well developed as the other two. We did not extend the search to an even earlier age because only two bilingual children, Alicia and Sophie, had earlier recordings available, but, understandably, they were not very productive in Cantonese in those recordings.

Discussion

Our study is the first to investigate the development of lexical tones in bilingual children acquiring Cantonese and English simultaneously from birth. Previous studies on successive Cantonese–English bilingual children showed that their tone development was not affected by the introduction of English after age two, when Cantonese monolingual children were shown to have acquired all of their tones. Our data on five simultaneous bilingual children illustrated that the development of their Cantonese tones was indeed influenced by their English. First, although the overall accuracy (Table 3) gives the impression that bilingual tones were on a par with monolingual tones, bilingual children made more errors than monolingual children for some specific tones at 2;0 (Figure 1). As a group, they appeared to have caught up with the monolingual children by 2;6. When we consider the bilingual children individually, they did not form a uniform pattern. While three bilingual children were already quite accurate in their tone production at 2;0, two children, Charlotte and Llywelyn, had over 20% of errors. They were obviously behind in their Cantonese development, evidenced not only by the number of tone errors they made, but also by the fewer Cantonese syllables uttered by them in the recordings. Detailed analysis of their tone production revealed that there were systematic errors in both children, which corresponded well to the English ‘high–low’ trochaic stress pattern. These templatic errors appeared most often in reduplicated words, but they were also found in non-reduplicated forms in Llywelyn’s speech. Further exploration of earlier recordings at 1;9 of four bilingual children also found similar templatic errors in Alicia’s speech, whose Cantonese development was behind the other Cantonese-dominant children at the same time-point as well.

We have found clear evidence that the lexical tone development of the bilingual children was influenced by their simultaneous exposure to English. This stands in stark contrast to the findings on successive Cantonese–English bilingual children (Holm & Dodd, 1999, 2006), whose tones were found to be intact, but echoes well the independent findings in Light’s (1977) case study of his daughter, whose sudden exposure to English during a period when her Cantonese tones were still developing (1;6) resulted in ‘tonal disintegration’. Light specifically pointed out that many of her tone errors resembled the stress or intonation patterns of the English equivalent items. The ‘high–low’ errors found in our study demonstrate a similar influence. It should be noted that the three bilingual children producing ‘high–low’ errors in our study were from different families. The convergence of their ‘high–low’ template was likely due to the influence of English exposure. The trochaic pattern constitutes about 90% of disyllabic words in English (Cutler & Carter, 1987). It is unsurprising that such a dominant pitch pattern was adopted by the bilingual children as a template.

Not all bilingual children in our study exhibited the ‘high–low’ template, however. Even for Charlotte, who produced such a template at 2;0, there was no example at 1;9. This seems a bit counter-intuitive, given our argument that such template would be found when their Cantonese tones were not well developed and thus more easily affected by the English stress pattern. We would expect to see more such examples in earlier recordings. For studies using corpus data, one important consideration is the amount of language use captured by the corpus. Tomasello and Stahl (2004) estimated that a weekly or bi-weekly one-hour recording sampling frequency only constituted about 1–1.5% of actual language use by the children. Given that the templatic errors were only found when the children were not very fluent in

Cantonese, that is, the base of production was already quite small, it is not surprising at all that we did not find examples from all children, or at all time-points. In fact, we considered ourselves lucky to have found some examples from three bilingual children!

Notwithstanding the sparsity of data sampled in a corpus, our data indicate that the window of occurrence of this 'high-low' template is probably quite short (only a few months). It is particularly worth noting that the base forms of those templatic errors were different, but they all converged on the same 'high-low' pitch pattern. We believe that they represented a genuine and non-idiosyncratic influence from English trochaic stress pattern.

The 'high-low' template found in our study resembles, but is not the same as, the 'phonological templates' proposed by Vihman (2010, 2014a, 2014b, 2016) for first-language phonological acquisition. Vihman (2014b) defined phonological templates as "idiosyncratic child production patterns typically developed in the period of single-word use and often maintained or further developed through the first months of word combination, after which they fade out of use" (p. 466). They refer to certain segmental combinations, although the prosodic and rhythmic structures of the adult languages can also influence the shapes of the templates. She also suggested that, on the one hand, the use of templates and the timing of template use cannot be readily predicted. On the other hand, the templates are similar both within and across languages. The 'high-low' template used by the bilingual children in our study was not segmental in nature, and the children were well over the one-word stage when they produced such a template. In addition, the phonological templates proposed by Vihman are stepping-stones used by children to approximate adult phonology, so they were preliminary attempts and were not very accurate. In contrast, the 'high-low' template in our study was used alongside the correct production of the same word, so they were not stepping-stones to the ultimate forms. Instead, they were alternative forms which co-existed with the correct forms. The 'high-low' template resulted from the bilingual interaction between Cantonese and English prosody, while phonological templates can be found in both monolingual and bilingual children. Nevertheless, there are also some common properties between the 'high-low' template and the phonological templates. They are similarly shaped by the prosodic patterns of the adult language, and are abstract in nature. Neither type of template is necessarily found in all children, and yet they can be similar across children. The 'high-low' template can be viewed as a specific type of phonological template for bilingual children.

The co-occurrence of the correct forms and the templatic forms is intriguing. Why would the bilingual children produce the wrong 'high-low' form if they already knew and could produce the correct forms? The same question applies to Light's (1977) study. Her daughter could produce the correct tones before intensive exposure to English which resulted in 'tonal disintegration' for about two years. Her tone production was good again at around age four (i.e., demonstrating a U-shaped development pattern). We believe that these instances demonstrate the dynamic interaction of the two prosodic systems, and the creativity and flexibility of the bilingual children. Although phonological development is generally viewed as a gradual process, it does not mean that the path is linear. In fact, many studies reported a U-shaped phonological development in which regressions are not uncommon, even for monolingual children (e.g., Vihman & Velleman, 1989; Bleile & Tomblin, 1991; Werker, Hall, & Fais, 2004). Regressions are temporary loss of phonetic accuracy in a later time relative to an earlier time. Previous studies usually

documented segmental regressions. Our data illustrate that regressions can occur at the suprasegmental level as well. In addition, Werker *et al.* (2004) argued that the U-shaped development, which is often found in various aspects of infant development, represents a reorganization, rather than actual loss, of the relevant ability/system. This idea fits our data well, as the bilingual children were able to produce the correct forms. These regressions or errors demonstrate the fluidity of their prosodic systems, with interaction between the two languages.

That the 'high-low' template occurs predominantly in reduplicated forms is interesting. Why was it observed mainly in reduplicated words, when presumably any disyllabic words could host such a template? Reduplication in Chinese is a morphological means to express a diminutive connotation, used often in child speech or child-directed speech (Matthews & Yip, 2011). In this sense, the cases of 'high-low' template could also be seen as the children applying a morphological template, in which they also added a tonal pattern, reminiscent of them using diminutive forms in English (e.g., *piggy*, *kitty*, *poo-poo*, *wee-wee*) which most often involved a trochaic pattern. In addition, the lexical meaning of the reduplicated forms is less affected if the tones are not produced accurately, as both syllables are the same. Nevertheless, it remains unclear why reduplicated words were produced differently – a more extensive survey of the acquisition of tones by bilingual children will be useful to shed light on this.

Exactly the same combination of languages (Cantonese and English) and features (tones and stress) could result in vastly different patterns (intact Cantonese tones in successive bilingual children in previous studies vs. systematic templatic errors in simultaneous bilingual children in our study). Bilingual interaction would be most easily observed when the relevant features were still not fully developed, still in a state of flux, as it were. The notion of reorganization discussed above concurs well with this point. It is interesting to note that the time-point along the developmental trajectory appears to be even more important than language dominance in this respect. When their Cantonese was not so strong, the English-dominant child Charlotte, and the Cantonese-dominant children Llywelyn and Alicia, all exhibited the 'high-low' English template.

In addition to the 'high-low' template, our data also suggest that T1 seemed to be another 'default' tone used by the two bilingual children. T1 [55] was among the first tone acquired by Cantonese monolingual children (Tse, 1978; So & Dodd, 1995). It is perceptually very salient. T1 is also more frequent than other tones in Cantonese (Fok-Chan, 1974; Leung, Law, & Fung, 2004). All this may have contributed to the bias of adopting T1 as a 'default' tone.

Paradis and Genesee (1996) suggested that, if the two language systems in bilingual children are interdependent, there are three possible interaction effects: delay, acceleration, and transfer. There was no obvious instance of acceleration in our data, but three bilingual children, Timmy, Sophie, and Alicia, were on a par with monolingual children in their Cantonese tone production at 2;0, i.e., there was no difference. Charlotte and Llywelyn were behind in their Cantonese development at the same time-point, but they had caught up with the monolingual children at 2;6, using the lenient criterion. The delay appeared to be short-lived. The most interesting finding in our data is the effect of transfer or cross-linguistic influence, resulting in the 'high-low' template, but not all bilingual children exhibited such a template. Our findings clearly suggest that bilingual phonological interactions are variegated. Divergent patterns would emerge at different time-points according to

different children. Which factors contribute to the diverse patterns of bilingual phonological interactions warrants further investigation.

As an aside, the relative production accuracy of each target tone was different between monolingual and bilingual children. Figure 1 shows that, under the stringent criterion, the most accurately produced tones were in the order $1 > 3 > 4 + 6 > 2 + 5$ for monolinguals at 2;0, and $1 + 6 > 3 > 4 > 2 + 5$ for bilinguals at the same age. At 2;6, the least accurately produced tones were T2 and T5 for all groups. These results agree with previous studies, in which T1 was found to be the earliest to be acquired, followed by the rising tones (Tse, 1978; So & Dodd, 1995). However, these works also found T6 to be the last tone to be mastered, which is different from our data. That said, as noted in So and Dodd (1995), the timing of acquisition of T6 manifested substantial cross-speaker variability, and may thus be irrelevant to whether the child is bilingual or otherwise. More longitudinal studies with more children are needed in order to determine the order of acquisition for individual tones.

The major limitation of our study is that we could not perform an acoustic analysis of tone production given the nature of the data. We compensated for it by having two independent native judges transcribing all the data to increase reliability. In addition to relying on corpus data, further studies should elicit experimental data suitable for acoustic analysis to investigate bilingual tone acquisition more comprehensively.

A recent review by Singh and Fu (2016) pointed out a distinct course of first-language development of tone languages as compared to non-tone languages. They reviewed a large number of perception and production studies on tone development (mostly focusing on the first two years of life), and suggested theoretical advancement could be made by further research on tone acquisition. Our study complements their review by discussing bilingual acquisition of lexical tone and demonstrating that first-language tone development can be diverse as well. Given that we have found interesting cross-linguistic influence on early bilingual tone acquisition using Cantonese–English bilingual children, it would be useful to expand the investigation to other bilingual children also acquiring English and a tone language simultaneously, e.g., Mandarin–English bilinguals. If, as we contend, the ‘high–low’ template is a genuine cross-linguistic influence from the English trochaic stress pattern onto tone realization, it is quite possible that similar templatic errors could be found in Mandarin–English bilingual children, too. However, so far, data on early Mandarin–English bilingual children are still very scarce. Lin and Johnson (2010) reported phonological development patterns of successive Mandarin–English bilingual children at 5;0, but they did not comment on their tones; presumably their tones were accurate, just like the data reported in Holm and Dodd (1999, 2006) on successive Cantonese–English bilingual children, as similarly early acquisition of Mandarin tones by monolingual children (before 2;0) was reported (Zhu & Dodd, 2000; Zhu, 2002). It is likely that only dense longitudinal data during the first two years of life would reveal cross-linguistic influences on tonal development. We look forward to seeing similarly interesting patterns from various types of bilingual children to corroborate our findings.

Conclusions

Our study compared Cantonese tone production by simultaneous Cantonese–English bilingual children at 2;0 and 2;6 with their Cantonese monolingual peers. Previous studies showed that monolingual children have acquired their tones by 2;0 or 2;6,

and that the tone development of successive Cantonese–English bilingual children was not affected by the introduction of English after age two. Our results illustrated that while some simultaneous bilingual children were on a par with their monolingual peers, some had a delay at 2;0. Some bilingual children also exhibited a ‘high–low’ template in their Cantonese production, resembling the pitch pattern of English trochaic words. Cross-linguistic prosodic influence was evidenced when the bilingual children were not so fluent in Cantonese. Diverse patterns of first language tone development are demonstrated.

Acknowledgements. We thank Professor Thomas Lee for giving us access to the CANCORP recordings analyzed in this paper. We also would like to thank Meeko Mak for his involvement in the earlier stages of the project, and Mercy Wong for transcribing parts of the recordings.

References

- Alwan, A., Jiang, J., & Chen, W. (2011). Perception of place of articulation for plosives and fricatives in noise. *Speech Communication*, 53, 195–209.
- Bauer, R. S., & Benedict, P. K. (1997). *Modern Cantonese phonology*. Berlin/New York: Mouton de Gruyter.
- Bleile, K., & Tomblin, B. (1991). Regressions in the phonological development of two children. *Journal of Psychological Research*, 20, 483–99.
- Bunta, F., & Ingram, D. (2007). The acquisition of speech rhythm by bilingual Spanish- and English-speaking 4- and 5-year-old children. *Journal of Speech, Language, and Hearing Research*, 50(4), 999–1014.
- Chao, Y. R. (1930). A system of tone-letters. *Le Maître Phonétique*, 45, 24–7.
- Chao, Y. R. (1947). *Cantonese primer*. New York: Greenwood Press.
- Chrabaszcz, A. V., Winn, M. B., Lin, C. Y., & Idsardi, W. J. (2014). Acoustic cues to perception of word stress by English, Mandarin and Russian speakers. *Journal of Speech, Language, and Hearing Research*, 57, 1468–79.
- Cooper, W. E., Eady, S. J., & Mueller, P. R. (1985). Acoustical aspects of contrastive stress in question/answer contexts. *Journal of the Acoustical Society of America*, 77, 2142–56.
- Cutler, A., & Carter, D. M. (1987). The predominance of strong initial syllables in the English vocabulary. *Computer Speech and Language*, 2, 133–42.
- Cutler, A., Lecumberri, M. L. G., & Cooke, M. (2008). Consonant identification in noise by native and non-native listeners: effects of local context. *Journal of the Acoustical Society of America*, 124, 1264–8.
- Fletcher, P., Leung, S. C. S., Stokes, S. F., & Weizman, Z. O. (2000). *Cantonese preschool language development: a guide*. Hong Kong: Department of Speech and Hearing Sciences.
- Fok-Chan, Y. Y. (1974). *A perceptual study of tones in Cantonese*. Hong Kong University Press.
- Fry, D. B. (1967). Duration and intensity as physical correlates of linguistic stress. In I. Lehiste (Ed.), *Readings in acoustic phonetics* (pp. 155–8). Cambridge MA/London: MIT Press.
- Holm, A., & Dodd, B. (1999). A longitudinal study of phonological development of two Cantonese–English bilingual children. *Applied Psycholinguistics*, 20(3), 349–76.
- 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.
- Johnson, C. E., & Wilson, I. L. (2002). Phonetic evidence for early language differentiation: research issues and some preliminary data. *International Journal of Bilingualism*, 6, 271–89.
- Kehoe, M. (2002). Developing vowel systems as a window to bilingual phonology. *International Journal of Bilingualism*, 6, 315–34.
- Kehoe, M., Lleó, C., & Rakow, M. (2004). Voice onset time in bilingual German–Spanish children. *Bilingualism: Language and Cognition*, 7, 71–88.
- Lecumberri, M. L. G., & Cooke, M. (2006). Effect of masker type on native and non-native consonant perception in noise. *Journal of the Acoustical Society of America*, 119, 2445–54.

- Lee, T. H. T., & Wong, C. H. (1998). CANCORP: The Hong Kong Cantonese Child Language Corpus. *Cahiers de Linguistique Asie Orientale*, 27, 211–28.
- Lee, T. H. T., Wong, C. H., Leung, C. S., Man, P., Cheung, A., Szeto, K., & Wong, C. S. P. (1996). The development of grammatical competence in Cantonese-speaking children. *Report of RGC earmarked grant 1991–94*. Retrieved from <<https://childes.talkbank.org/access/Chinese/Cantonese/LeeWongLeung.html>>.
- Leung, M. T., Law, S. P., & Fung, S. Y. (2004). Type and token frequencies of phonological units in Hong Kong Cantonese. *Behavior Research Methods, Instruments & Computers*, 36(3), 500–5.
- Li, J., & Mok, P. (2014). The acquisition of English lexical stress by Cantonese–English bilingual children at 2;6 and 3;0. In *Proceedings of Speech Prosody 7* (pp. 688–92), Dublin. A version of the Proceedings can be retrieved from <https://www.researchgate.net/publication/315880083_Proceedings_of_the_7th_International_Conference_on_Speech_Prosoydy>.
- Light, T. (1977). CLAIRETALK: a Cantonese-speaking child's confrontation with bilingualism. *Journal of Chinese Linguistics*, 5, 261–75.
- Lin, L. C., & Johnson, C. J. (2010). Phonological patterns in Mandarin–English bilingual children. *Clinical Linguistics and Phonetics*, 24, 369–86.
- Matthews, S., & Yip, V. (2011). *Cantonese: a comprehensive grammar*, 2nd ed. New York: Routledge.
- Mok, P. (2011). The acquisition of speech rhythm by three-year-old bilingual and monolingual children: Cantonese and English. *Bilingualism: Language and Cognition*, 14(4), 458–72.
- Mok, P. (2013). Speech rhythm of monolingual and bilingual children at 2;06: Cantonese and English. *Bilingualism: Language and Cognition*, 16, 693–703.
- Mok, P., Zuo, D., & Wong, P. (2013). Production and perception of a sound change in progress: tone merging in Hong Kong Cantonese. *Language Variation and Change*, 25, 341–70.
- Paradis, J. (2001). Do bilingual two-year-olds have separate phonological systems? *International Journal of Bilingualism*, 5, 19–38.
- Paradis, J., & Genesee, F. (1996). Syntactic acquisition in bilingual children: Autonomous or independent? *Studies in Second Language Acquisition*, 18, 1–25.
- So, L. K. H., & Dodd, B. (1995). The acquisition of phonology by Cantonese-speaking children. *Journal of Child Language*, 22(3), 473–95.
- 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, 103–22.
- Tomasello, M., & Stahl, D. (2004). Sampling children's spontaneous speech: How much is enough? *Journal of Child Language*, 31, 101–21.
- Tse, J. K. P. (1978). Tone acquisition in Cantonese: a longitudinal case study. *Journal of Child Language*, 5, 191–204.
- Vihman, M. (2010). Phonological templates in early words: a cross-linguistic study. In C. Fougeron, B. Kühnert, M. D'Imperio, & N. Vallée (Eds.), *Laboratory phonology 10* (pp. 261–84). New York: Mouton de Gruyter.
- Vihman, M. (2014a). *Phonological development: the first two years*. Chichester: John Wiley and Sons.
- Vihman, M. (2014b). Phonological templates. In P. Brooks, V. Kempe, & J. G. Golson (Eds.), *Encyclopedia of language development* (pp. 466–7). Los Angeles: Sage.
- Vihman, M. (2016). Prosodic structures and templates in bilingual phonological development. *Bilingualism: Language and Cognition*, 19, 69–88.
- Vihman, M., & Velleman, S. (1989). Phonological re-organization: a case study. *Language and Speech*, 32, 149–70.
- Werker, J. F., Hall, D. G., & Fais, L. (2004). Reconstructing U-shaped functions. *Journal of Cognition and Development*, 5, 147–51.
- Wong, P. S. (2013). Perceptual evidence for protracted development in monosyllabic Mandarin lexical tone production in preschool children in Taiwan. *Journal of the Acoustical Society of America*, 133, 434–43.
- Wong, P. S., Fu, W. M., & Cheung, E. Y. L. (2017). Cantonese-speaking children do not acquire tone perception before tone production: a perceptual and acoustic study of three-year-olds' monosyllabic tones. *Frontiers in Psychology*, 8, 1450. Retrieved from <<https://www.frontiersin.org/articles/10.3389/fpsyg.2017.01450/full>>.

- Wong, P. S., Schwartz, R. G., & Jenkins, J. J.** (2005). Perception and production of lexical tones by 3-year-old Mandarin-speaking children. *Journal of Speech, Language, and Hearing Research*, *48*, 1065–79.
- Wu, W. L.** (2009). Sentence-final particles in Hong Kong Cantonese: Are they tonal or intonational? *Proceedings of the 10th Interspeech* (pp. 2291–4), Brighton, UK. Retrieved from <https://www.isca-speech.org/archive/interspeech_2009/i09_2291.html> .
- Xu, Y.** (1997). Contextual tonal variations in Mandarin. *Journal of Phonetics*, *25*(1), 61–83.
- Yip, M.** (2002). *Tone*. Cambridge University Press.
- Yip, V., & Matthews, S.** (2007). *The bilingual child: early development and language contact*. Cambridge University Press.
- Zhu, H.** (2002). *Phonological development in specific contexts: studies of Chinese-speaking children*. Clevedon: Multilingual Matters.
- Zhu, H., & Dodd, B.** (2000). The phonological acquisition of Putonghua (modern standard Chinese). *Journal of Child Language*, *27*, 3–24.
- Zhu, H., & Dodd, B.** (2006). A multilingual perspective on phonological development and disorders. In H. Zhu & B. Dodd (Eds.), *Phonological development and disorders in children: a multilingual perspective* (pp. 3–22). Clevedon: Multilingual Matters.