Prosodic interaction in Cantonese-English bilingual children's speech production

Jonathan Him Nok Lee[1], Regine Yee King Lai[2], Stephen Matthews[3] and Virginia Yip[2]
[1] University of Pennsylvania | [2] The Chinese University of Hong Kong | [3] The University of Hong Kong

This corpus-based study investigates intonation patterns in the production of Cantonese by Cantonese-English bilingual children. We examine the intonation patterns in eight simultaneous bilingual children acquiring a tonal (Cantonese) and an intonational language (English) from 2;0 to 3;0. Two intonation patterns are observed in all the bilingual children studied: high pitch followed by a fall (including H_H*L% and H_L*L%) and low pitch followed by a rise (including L_H*H% and L_L*H%), in which English-like intonation is applied to Cantonese and code-mixed utterances. They illustrate cross-linguistic influence in prosody from English in the bilingual children's early phonological development. Language dominance, use of sentence-final particles, and the children's grammatical complexity are found to be significant predictors for the production of bilingual intonation. First, the more dominant the child is in Cantonese, the less bilingual intonation is produced in Cantonese and code-mixed utterances. Second, bilingual intonation is significantly more likely to be produced in utterances with sentence-final particles than without. Third, the greater the child's grammatical complexity, the lower the predicted probability of producing bilingual intonation.

Keywords: bilingual intonation, prosody, sentence-final particles, cross-linguistic influence, language acquisition

1. Introduction

This corpus-based study investigates intonation patterns and prosodic interaction in utterances with Cantonese sentence-final particles (SFPs) produced by Cantonese-English bilingual children.

https://doi.org/10.1075/lab.21049.lee | Published online: 11 April 2023
Linguistic Approaches to Bilingualism 14:5 (2024), pp. 668–707. ISSN 1879-9264 | E-ISSN 1879-9272

1.1 Background

Early research showed that children regularly exposed to two languages from birth, or soon after, acquire their language systems as largely independent systems (Meisel, 1989). Despite this independence, recent studies have questioned the extent to which the two language systems of bilingual children are autonomous. In bilingual first language acquisition, interaction (transfer, interference, convergence, etc.) between the two languages of a bilingual child is collectively known as *cross-linguistic influence* (CLI), which refers to instances where there is evidence for the effect of one language on the other (Serratrice, 2013). Concerning the conditions under which CLI is likely to occur, there have been at least two hypotheses: the Overlap Hypothesis and Interface Hypothesis. Some overlap of the two systems at the surface level (Hulk & Müller, 2000), or even in their syntactic derivation (Schmitz et al., 2012) is necessary for CLI to occur. Additionally, it has been proposed that such influence takes place at the interfaces between different linguistic modules (e.g., syntactic, phonological, and semantic) (Sorace & Filiaci, 2006) and their external interfaces with discourse/pragmatics (Tsimpli & Sorace, 2006). However, many studies have reported phenomena of CLI which cannot be attributed to overlap or interfaces (e.g., Pérez-Leroux et al., 2011; Strik, 2012). Some scholars have argued that language dominance is a major determinant of the directionality of transfer (Yip & Matthews, 2000), while others have found no clear role for language dominance (e.g., Nicoladis, 2002).

Hypotheses concerning conditions for CLI have been predominantly based on morphosyntactic evidence (see Serratrice, 2013, for a review). The discussion has involved a wide range of morphosyntactic phenomena: word order (Döpke, 2000), subject realization (Paradis & Navarro, 2003), object omission (Hulk & Müller, 2000), wh-questions (Yip & Matthews, 2007), etc. While there have been limited studies investigating CLI in semantics, studies suggest that bilingual children's construal of semantic categories is guided by both non-linguistic cognitive factors and language-specific properties (Engemann, 2021). Regarding lexicon, scholars have focused on the use of "translation equivalents" by simultaneous bilinguals (Deuchar & Quay, 2000), which implies the children develop two distinct lexical systems.

Previous studies in early bilingual phonological acquisition have focused on segmental aspects of production. Recent studies have addressed intonation in bilinguals but mostly with non-tonal language pairs, e.g., German-English (Gut, 2000), Turkish-German (Queen, 2001, 2012), English-Hindi (Puri, 2013, 2018), Greek-English (Kontaxi & Chaida, 2015). In a longitudinal study of three German-English children from 2;1 to 5;6, Gut (2000) argued that the phonological functions of nucleus placement and pitch are mastered by the children as early as 2;1, while mastery of the phonetic production of the intonational systems is achieved much later, and acquisition is not yet complete at 5;6. Individual differences and clear separation of both language systems were apparent in all children studied, with a considerable time lag in the acquisition of the weaker language. Queen (2001) argued that the rises L*HH% and L%H%[1] used by Turkish-German bilingual children are non-normative for both Turkish and German and involve two-way fusion of the intonation of the two languages. Queen (2012) examined two phonetically, phonologically, and pragmatically distinct rises produced by second- and third-generation Turkish-German bilinguals living in Germany, which differ from what is typically reported for German monolinguals. She argued that the two rises are derived from the two languages and form an intonation system along with other edge phenomena. Comparing acoustic features in the speech of simultaneous and late Indian English-Hindi adult bilinguals, Puri (2013, 2018) argued that simultaneous bilinguals have a merged system of intonation because they acquired a nativized variety of English, but subtle acoustic features (e.g., the use of H*/H*L pitch accent) distinguish them from late bilinguals. Kontaxi and Chaida (2015) studied the influence of English on the tonal characteristics of Greek polar and wh-questions produced by Greek-English simultaneous bilinguals. They found that English tonal patterns did not radically affect Greek polar questions but wh-questions, which showed a different boundary tone. They argued that simultaneous bilinguals may use tonal features of both languages, which may differ across question types.

Only two studies have investigated the bilingual acquisition of intonation of a tonal and a non-tonal language. Examining the intonation of Yami-Mandarin adult bilinguals, Lai (2018) and Lai and Gooden (2018) reported transfer of Mandarin-like intonation to Yami neutral questions, and fusion between a non-Yami-native question type and pre-existing Yami intonation. These studies combined different types of adult bilinguals and adopted the perspective of language contact. The present study focuses specifically on bilingual interactions in intonation in simultaneous bilingual children acquiring a prosodically and typologically orthogonal language pair: Cantonese (tonal language) and English (intonational language). This study can bridge an important gap in our understanding of CLI in prosody in early bilingual development.

1. Based on Tones and Break Indices (ToBI) conventions, H* and L* annotate pitch accents of high and low tones respectively. H- and L- annotate phase accent, and H% and L% annotate boundary tones.

1.2 The prosodic systems of Cantonese and English

Every Cantonese syllable carries a tone to differentiate word meanings. There are six lexical tones (T1-T6) based on pitch contrast alone (Qian et al., 2007). Table 1 summarizes the six tones using Chao's 5-level annotation system. The numbers in the "relative pitch" column in Table 1 represent the relative starting and ending pitch of each tone, with 1 being the lowest and 5 the highest pitch of a speaker's normal range (Chao, 1930).

Table 1. The six contrastive Cantonese lexical tones based on pitch contrast alone

Tones	Descriptive pitch contours	Relative pitch
1	High level	55
2	High rising	25
3	Mid level	33
4	Low falling	21
5	Low rising	23
6	Low level	22

English has three levels of lexical stress: primary stress, secondary stress, and unstressed. English intonation is distinguished by global falling and rising of pitch (Grice et al., 2020), whereas Cantonese uses boundary tones on utterance-final syllables to signal intonation (Xu & Mok, 2011). It also possesses over 40 sentence-final particles (SFPs) with diverse meanings and functions including speech acts, evidentiality, affective and emotional coloring (Matthews & Yip, 2011). They are complex linguistic elements involving multiple interfaces between syntax, phonology, semantics/pragmatics, and discourse (Cheng & Tang, 2022; Chor, 2018; Lee, 2021; Lee et al., 2022; Wakefield, 2020). SFPs typically attach to the end of an utterance, while the part of the utterance excluding the SFPs is called the utterance body (Fox et al., 2008). For example, both (1) and (2) have the same utterance body *lei5 heoi3 sik6 faan6* ("you go to eat").[2] The use of *aa4* in (1) transforms the statement into a question, and the use of *le4* in (2) expresses a suggestion.

2. Cantonese examples are transcribed in *Jyutping* romanization (The Linguistic Society of Hong Kong, 2002). The onset and rhyme of each syllable are transcribed alphabetically, followed by a number (1–6) representing the lexical tones (T1-T6).

(1) Lei5 heoi3 sik6 faan6 aa4?
2SG go eat rice SFP
"Are you going to eat?"

(2) Lei5 heoi3 sik6 faan6 le4.
2SG go eat rice SFP
"I suggest that you should go to eat."

Similar to other lexical items, the tones on Cantonese SFPs can differentiate distinct meanings between particles having the same segmental form. For instance, *aa1* is a particle that signals "lively" suggestions or advice; *aa3* is an emotion softener; *aa4* is a question particle. It has been argued that tones on certain SFPs are intonational (Sybesma & Li, 2007). Wakefield (2020) even argued that Cantonese SFPs are the equivalent of English intonation. Ladd (1996/2008) was one of the first to propose that intonation be redefined to include particles if they have similar functions. It is argued that all languages use a combination of intonation and particles for expressing connotative meaning, with languages that use more intonation using fewer particles, and vice versa (Wakefield, 2020). For example, languages that use a large number of particles, such as Cantonese, German, Japanese, and Mandarin, have smaller repertoires of intonation. In contrast, languages with larger inventories of intonation, such as English and French, have fewer discourse particles (e.g., sentence-final *huh?* and *man!* in English) but use suprasegmental intonation to encode similar semantic/pragmatic functions (Ladd, 2008; Lee et al., 2022; Wakefield, 2020).

Regarding CLI, both the Overlap and Interface hypotheses will predict transfer in the production of Cantonese SFPs by Cantonese-English bilingual children. First, the two languages possess the overlapping feature of particles in utterance-final positions. The similarity in functions between Cantonese SFPs and English intonation may lead to transfer in the speech of bilingual children. Second, pitch serves as a prosodic cue in both Cantonese (intonation and tones) and English (intonation and stress). Third, Cantonese SFPs are situated at multiple language-internal and external interfaces. Both hypotheses therefore predict transfer of prosody in the presence of SFPs.

To our knowledge, there are no previous studies specifically investigating the prosody of SFPs in Cantonese-English bilingual children. Lai (2006), however, noticed a non-target usage of the SFP *aa4* by Charlotte, a Cantonese-English bilingual child in the same corpus used by the current study. Charlotte often showed a falling intonation, affected by English prosody, at the end of Cantonese and code-mixed utterances. Regarding the acquisition of Cantonese tones, previous studies showed that monolingual Cantonese-speaking children have acquired their tones by 2;0 or 2;6 (So & Dodd, 1995; To et al., 2013; Tse, 2008). Mok and

Lee (2018) investigated the acquisition of Cantonese tones by five simultaneous Cantonese-English bilingual children in the corpus used by the current study and reported U-shaped patterns of tonal accuracy in the children's development. Specifically, Alicia, Sophie, and Timmy were on a par with their monolingual counterparts in their Cantonese tone production at 2;0, whereas Charlotte and Llywelyn were behind in their Cantonese development until 2;6. Excluding SFPs from the analysis, they argued that T1 is the "default tone" in the errors in Charlotte and Llywelyn's productions.

1.3 Current study

This study investigates prosodic interaction between Cantonese SFPs and intonation in the phonological development of simultaneous bilingual children acquiring Cantonese and English. Examining the intonation produced by the bilingual children longitudinally from 2;0 to 3;0, we discuss how the use of pitch for intonation interacts with the pitch patterns of lexical tones and SFPs in Cantonese and code-mixed utterances, focusing on "bilingual intonation" which we define as English-like intonation applied to Cantonese and code-mixed utterances. We address the following questions: (i) is there CLI in prosody in the phonological development of Cantonese-English bilingual children? If so, (ii) are any intonation patterns exclusive to the bilingual children? If so, (iii) what factors affect the production of bilingual intonation?

1.4 Hypotheses

Based on the studies reviewed, we formulate the following hypotheses and predictions for three factors to be investigated in the current study: language dominance, use of SFPs, and grammatical complexity.

First, the hypothesis that language dominance is a determinant of the directionality of transfer predicts that if the Cantonese-English bilingual child is more Cantonese-dominant, there will be less CLI from English, and thus less frequent production of bilingual intonation. Conversely, if language dominance plays no role in CLI, there should be no significant difference in the frequency of the production of bilingual intonation between children with different language dominance profiles.

Second, the Overlap and Interface hypotheses predict cross-linguistic transfer of prosody to occur predominantly in the presence of SFPs.

Third, as the bilingual children's grammar grows more complex, they should produce more SFPs. Grammatical complexity (measured by mean length of utterance (MLU)) is associated with age (Miller & Chapman, 1981). As the bilingual children develop, they should learn that bilingual intonation is a non-target intonation. We hypothesize that the production of bilingual intonation will decrease as the child's grammatical complexity increases.

2. Methods

2.1 Participants

We investigate the production of eight simultaneous Cantonese-English bilingual children (four female) from ages 2;0 to 3;0 using the longitudinal Hong Kong Bilingual Child Language Corpus (Yip & Matthews, 2007) in CHILDES (MacWhinney, 2000). The same corpus was previously the subject of prosodic analyses of rhythm by Mok (2011, 2013) and acquisition of lexical tones by Mok and Lee (2018). Six participants were children of mixed marriages, and two (Darren and Kasen) were children of native speakers of Cantonese who also spoke English fluently as an L2. The eight children came from similar socio-economic backgrounds: at least one of their parents was a high-school or university teacher. All children received care from Filipino domestic helpers who spoke English. They were thus exposed to both Cantonese and English from birth. Five were Cantonese-dominant, one was English-dominant, one showed balanced development in both languages, and one was originally balanced but became English-dominant after 2;10. Language dominance was determined based on MLU differentials between the two languages, language preferences, and patterns of code-mixing (Yip & Matthews, 2006, 2007).

The children were recorded longitudinally in naturalistic (typically play) situations at their homes by investigators within Cantonese and English contexts. Cantonese context refers to sessions where the child was addressed in Cantonese in the recording. Since our study is concerned with the usage of Cantonese SFPs, only recordings in the Cantonese context were considered. The corpus data were analyzed by sampling children's speech at 3-month intervals at ages 2;0, 2;3, 2;6, 2;9, and 3;0. One recording in the Cantonese context was sampled for each participant at each time-point. Table 2 summarizes the background information of the eight bilingual children.

To verify whether the intonation patterns observed are exclusive to bilingual children, we also investigated the production of eight monolingual Cantonese-speaking children in the HKU-70 corpus ($N=70$, cross-sectional study, naturalistic data, Fletcher et al., 2000) in CHILDES each at ages 2;6 and 3;0. Table 3 summarizes their background information.

Table 2. Background information of the eight bilingual children

Child	Sex	Dominant language	Age of data used	No. of utterances	No. of morphemes	MLU
Alicia	F	Cantonese	2;00.13	68	190	2.79
			2;03.02	148	395	2.67
			2;06.02	172	540	3.14
			2;09.15	187	462	2.47
			3;00.10	234	798	3.41
Charlotte	F	English	2;00.25	131	303	2.31
			2;03.17	152	380	2.50
			2;06.16	161	398	2.47
			2;09.04	35	68	1.94
			3;00.03	123	372	3.02
Darren	M	Balanced	2;00.10	152	327	2.15
			2;03.08	148	363	2.45
			2;06.11	107	215	2.01
			2;09.03	242	669	2.76
			3;00.07	210	463	2.21
Janet	F	Cantonese	2;10.16	142	545	3.84
			3;00.11	369	1433	3.88
Kasen	M	Balanced	2;06.16	256	752	2.94
			2;09.03	369	1221	3.31
		English	3;00.03	184	545	2.96
Llywelyn	M	Cantonese	2;00.12	149	215	1.44
			2;03.14	154	320	2.08
			2;06.20	300	980	3.27
			2;09.07	293	986	3.37
			3;00.27	292	865	2.96
Sophie	F	Cantonese	2;00.06	330	683	2.07
			2;03.01	237	675	2.85
			2;06.00	213	719	3.38
			2;09.06	361	1241	3.44
			3;00.03	226	682	3.02
Timmy	M	Cantonese	2;03.17	325	1116	3.43
			2;06.09	312	1110	3.56
			2;09.08	217	829	3.82
			3;00.09	413	1459	3.53

Table 3. Background information of the eight monolingual Cantonese-speaking children

Time-point	Age of data used	Sex	No. of utterances	No. of morphemes	MLU
2;6	2;05.23	F	145	451	3.11
	2;05.27	M	94	210	2.23
	2;05.30	M	133	398	2.99
	2;06.01	F	261	984	3.77
3;0	2;11.28	M	123	368	2.99
	2;11.29	M	167	674	4.04
	2;11.30	M	154	369	2.40
	3;00.29	F	128	470	3.67

2.2 Procedure

Data listed in Table 2 were divided into three datasets. Dataset 1 comprises all sampled recordings of Alicia and Janet. Dataset 2 comprises all sampled recordings of Charlotte and Darren, and those of Llywelyn at 2;9 and 3;0. Dataset 3 comprises all sampled recordings of Kasen, Sophie, and Timmy, and those of Llywelyn at 2;0, 2;3, and 2;6. A native adult speaker of Cantonese who spoke English as an L2 listened to all three datasets, judging the intonation pattern of each utterance. Two bilingual intonation patterns were observed: high pitch followed by a fall (HF) and low pitch followed by a rise (LR). These patterns are not recognized as being present in Cantonese, which does not signal intonation at the global utterance level but uses boundary tones on utterance-final syllables (Xu & Mok, 2011). Another three native adult speakers of Cantonese who also spoke English as an L2 were trained to identify the perceptual features of the two bilingual intonation patterns by being presented with more than ten representative samples of each pattern. They then each listened to one recording in a dataset and were asked to judge the intonation of all utterances in the recording blindly and independently, with or without the intonation patterns as judged by the first rater. The raters were asked to judge whether the utterance presented exemplified either HF or LR; and if so, which. The judgements of the raters achieved more than 95% similarity to the judgements of the first rater at this training phase. Next, the raters judged the remaining recordings in their dataset without referencing the judgements of the first rater (i.e., each dataset was listened to by two raters in total). Inter-rater reliability (Cohen's kappa) was calculated.

Of the eight previous studies of bilingual acquisition of intonation listed in Section 1.1, Puri (2013, 2018) and Kontaxi and Chaida (2015) were experimental studies which did not involve cross-checking of judgements. Among the remaining five corpus-based studies, Gut (2000) carried out an inter-rater reliability test of 200 random utterances (~1.43% of her data), while the other four studies did not mention any cross-checking of judgements. The present study calculated inter-rater reliability based on all data (no. of utterances = 7412). Of these utterances, 47.28% were produced with SFPs (N=3430) and 53.72% were not (N=3982). The total working hours were estimated to be around 500 hours.

Similarly, two native speakers of Cantonese listened to all data listed in Table 3, judging whether any bilingual intonation patterns were attested in the speech of monolingual Cantonese-speaking children.

3. Results

3.1 Inter-rater reliability

Table 4 shows inter-rater reliability (Cohen's kappa) for the three bilingual datasets listed in Table 2. There is substantial to almost perfect agreement between the judgements of raters (Cohen, 1960).

Table 4. Inter-rater reliability (bilingual data)

Dataset	Agreement (no. of utterances)	κ	*SE*	95% CI
1	99.985% (N=1320)	.672	.069	.537-.807
2	99.976% (N=2046)	.849	.021	.808-.890
3	95.587% (N=4046)	.651	.024	.604-.698

Regarding the monolingual data, the two raters agreed that no bilingual intonation patterns were attested in the utterances produced by the monolingual Cantonese-speaking children listed in Table 3 (i.e., agreement=100%). All utterances sounded native-like to both raters.

3.2 Bilingual intonation

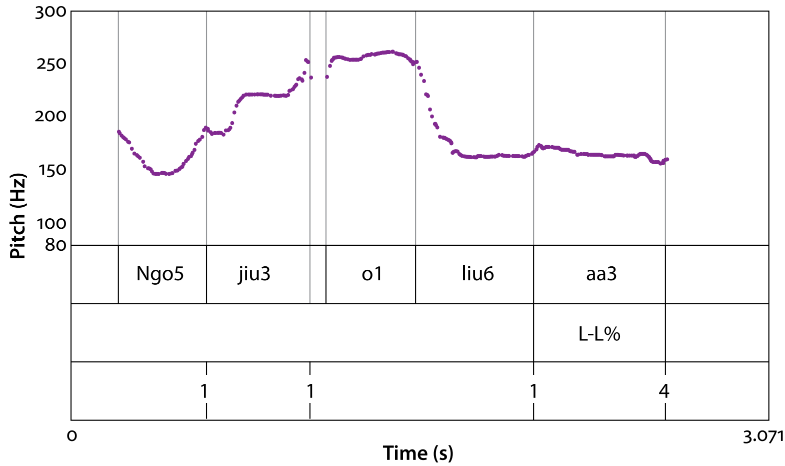
Two bilingual intonation patterns are observed in all the bilingual children studied: HF and LR. There are three features of the two patterns: utterance body, nuclear accent, and sentence-final position. ToBI equivalents for the pattern HF include H_H*L% and H_L*L%, and those for LR include L_H*H% and L_L*H%.

To illustrate the differences between the two bilingual intonation patterns and the target/native intonation, an utterance of example (3a) (= 3b) produced by a native adult speaker of Cantonese with native intonation was recorded. Its Fo measured by Praat (Boersma & Weenink, 2019) and ProsodyPro (Xu, 2013) is shown in Figure 1(A).[3] Figure 1(B)–(D) shows the F0 of examples (3a)–(3c). The pattern HF in (3a) can be attributed to superimposition of an English intonation with high pitch followed by a fall at the utterance-final position (Figure 1(B)). The intonation pattern LR in (3b) can be attributed to superimposition of an English intonation with low pitch followed by a rise at the utterance-final position (Figure 1(C)). In (3a) the high pitch at the utterance body overrides the lexical tones of the syllables, while the fall at the utterance-final position overrides the mid-level tone of the SFP *aa3*. In (3b) the low pitch at the utterance body overrides the high-level tone of *o1* "pee", while the final rise overrides the mid-level tone of the SFP *aa3*.

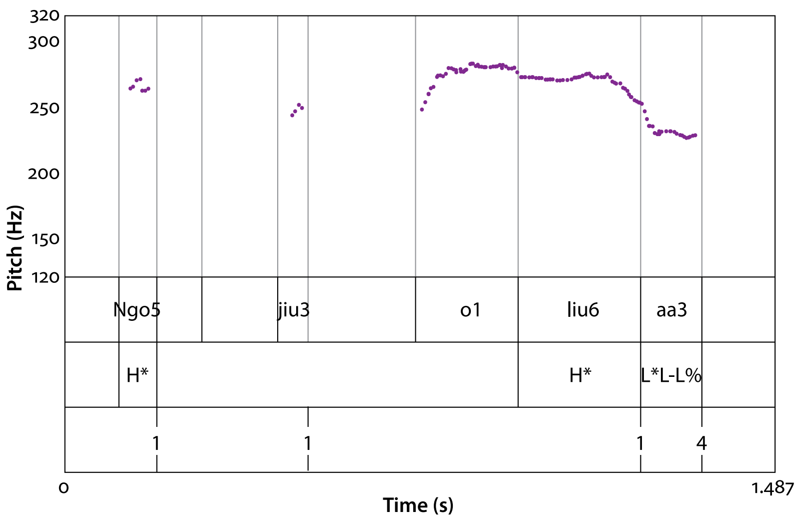
(3) a. Ngo5 jiu3 o1liu6 aa3. (Llywelyn 2;03.14)
1SG have.to pee SFP
"I have to pee."
b. Ngo5 jiu3 o1liu6 aa3. (Llywelyn 2;03.14)
1SG have.to pee SFP
"I have to pee."
c. O1liu6. (Llywelyn 2;03.14)
Pee
"Pee."

All raters agreed that the penultimate syllables of both examples sounded metrically prominent and stressed, like the nuclear accent in English. The pre-final *liu6* "pee" in (3a) was produced with a high pitch (mean *F0* = 270.51 Hz) similar to the preceding syllable *o1* "to pee" with high-level tone (mean *F0* = 275.80 Hz). In addition, its mean intensity (61.00 dB) is 3.35 times that of the preceding syllable *o1* (59.02 dB). Similarly in (3b), the penultimate syllable *liu6* should have a lower target tone than the preceding syllable *o1*, yet its mean F0 (335.31 Hz) is 41.78 Hz higher than that of *o1* (293.53 Hz). Its syllable duration (285 ms) is 26.7% longer than that of *o1* (225 ms). These acoustic data match the native intuition of the penultimate syllables of the two instances of *liu6* being metrically prominent and stressed. However, it is not that Llywelyn had not yet acquired the tone of *liu6*, because he produced the target tones of the phrase (3c) *o1liu6* "to pee" (Figure 1(D)) elsewhere in the same recording when there was no SFP.

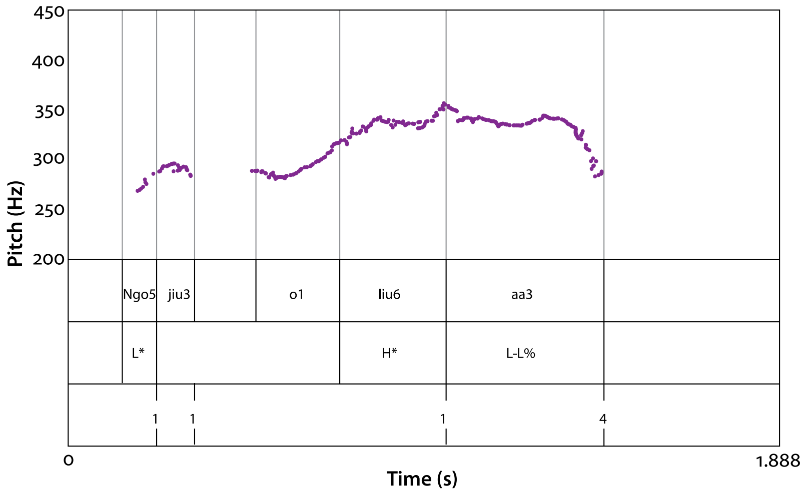
3. In each pitch-tracking diagram, the first tier shows the F0 and the second tier the transcription of the utterance. The third and fourth tiers show tone and break-index tiers using ToBI conventions. The numbers in the fourth tier are the break-index values: 1 = phrase-internal word end; 3 = intermediate phrase end; and 4 = intonational phrase end.


300
250
200
150
100
80
Pitch (Hz)
Ngo5
jiu3
o1
liu6
aa3
L-L%
1
1
1
4
0
3.071
Time (s)

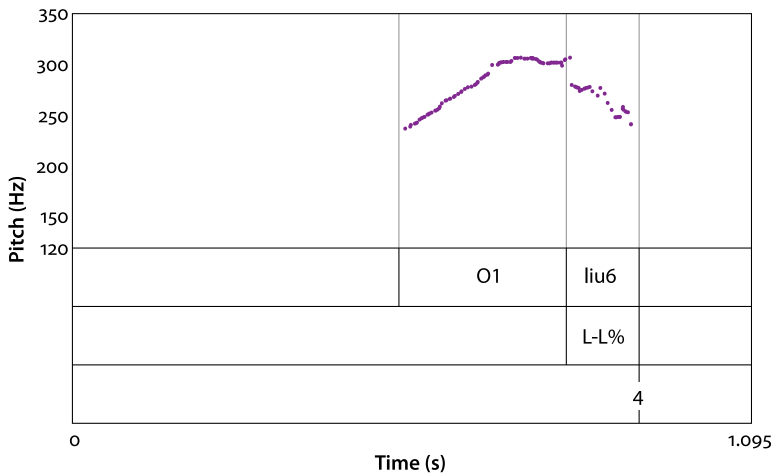
A.


320
300
250
200
150
120
Pitch (Hz)
Ngo5
jiu3
o1
liu6
aa3
H*
H*
L*L-L%
1
1
1
4
0
1.487
Time (s)

B.



C.



D.

Figure 1. (A) F0 of (3a/3b) produced by a native adult speaker of Cantonese; (B) F0 of (3a); (C) F0 of (3b); (D) F0 of (3c)

The two bilingual intonation patterns occur in both Cantonese utterances ((3a) and (3b)) and code-mixed utterances ((4a) and (5a)). Figure 2(B)–(C) shows the F0 of examples (4a) and (5a). Figure 2(A) shows the F0 of (4a) produced by a native adult speaker of Cantonese who also spoke English as an L2. The utterance body of (4a) is superimposed with an intonation showing high pitch, followed by a fall at the SFP *aa3*. In addition to the primary stress at the first syllable of the word *battery*, Llywelyn also stressed the unstressed syllable -ry of *battery* (the penultimate syllable of the utterance), as judged by two native adult speakers of English. In contrast, the initial low pitch in (5a) overrides the prominence of the stressed syllable Tim-, followed by a high boundary tone at the SFP *le1*. The stressed syllable *Tim-* (mean *F0* = 488.03 Hz) does not show prominence relative to the unstressed syllable *-my* in pitch (mean *F0* = 493.27 Hz) or in volume (mean intensity = 66.04 dB for *Tim-* vs. 66.62 dB for *-my*). These examples are not cases where the bilingual children had yet to acquire the prosody of the words, because we find ample evidence for them producing the target prosody elsewhere, especially in utterances without SFPs. For example, Llywelyn had already acquired the target prosody (the appropriate English stress patterns) of *battery* no later than 2;01.13 (4b); Alicia had acquired that of *Timmy* by 2;00.15 (5b).

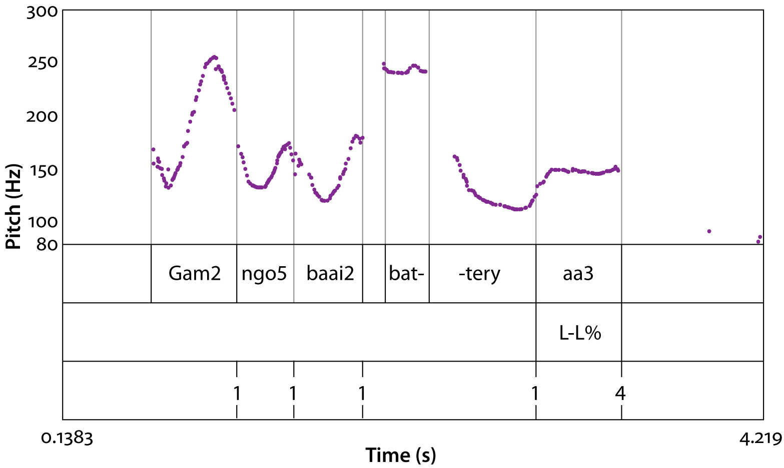
(4) a. Gam2 ngo5 baai2 battery aa3. (Llywelyn 3;00.27)
then 1SG put battery SFP
"Then I'll put batteries in."
b. Jau5 mou5 battery? (Llywelyn 2;01.13)
have not.have battery
"Do you have batteries?"

(5) a. Timmy le1? (Alicia 2;10.29)
Timmy SFP
"What about Timmy?"
b. Timmy. (Alicia 2;00.15)

We find no predictive relationship between the two bilingual intonation patterns and sentence types. Although the pattern LR resembles interrogative intonation in both English and Cantonese yes-no questions, it appears in both questions (5a) and statements (3b). Similarly, although the pattern HF resembles declarative intonation in English, it is found in both statements ((3a) and (4a)) and questions (6). Figure 3 shows the F0 of (6), a Cantonese A-not-A question superimposed with an English intonation with high pitch overriding the low-falling tone of *m4* "not", followed by a fall at the SFP *gaa3*. Despite the superimposition of the bilingual intonation, we can still hear traits of trochaic stress (pitch differences with high-low tone patterns) in *li1go3* "this CL" (CL = classifier), approximating the target tones in Cantonese.

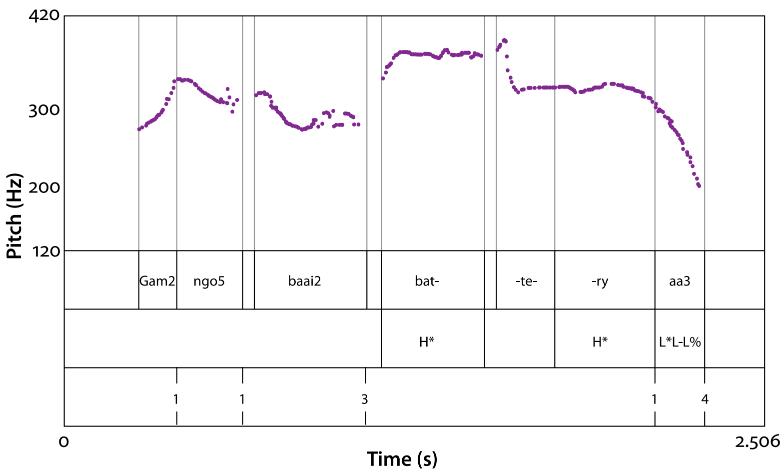
(6) Li1go3 li1go3 dak1-m4-dak1 gaa3? (Charlotte 2;03.17)
this.CL this.CL can-not-can SFP
"This, this, is it okay?"

The two English-like intonation patterns were also attested in the bilingual children's English. For instance, (7) carries an English intonation with high pitch,



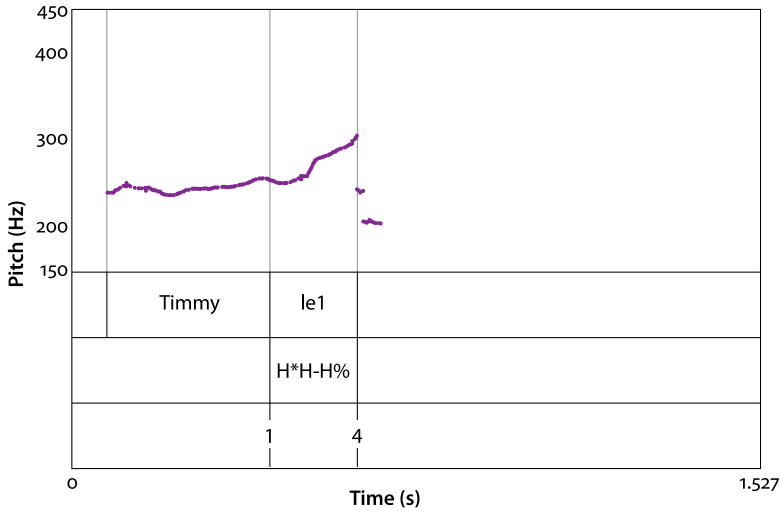
300
250
200
150
100
80
Pitch (Hz)
Gam2
ngo5
baai2
bat-
-tery
aa3
L-L%
1
1
1
1
4
0.1383
4.219
Time (s)

A.



420
300
200
120
Pitch (Hz)
Gam2
ngo5
baai2
bat-
-te-
-ry
aa3
H*
H*
L*L-L%
1
1
3
1
4
0
2.506
Time (s)

B.



C.

Figure 2. F0 of (4a) produced by a native adult speaker of Cantonese; (B) F0 of (4a); (C) F0 of (5a)

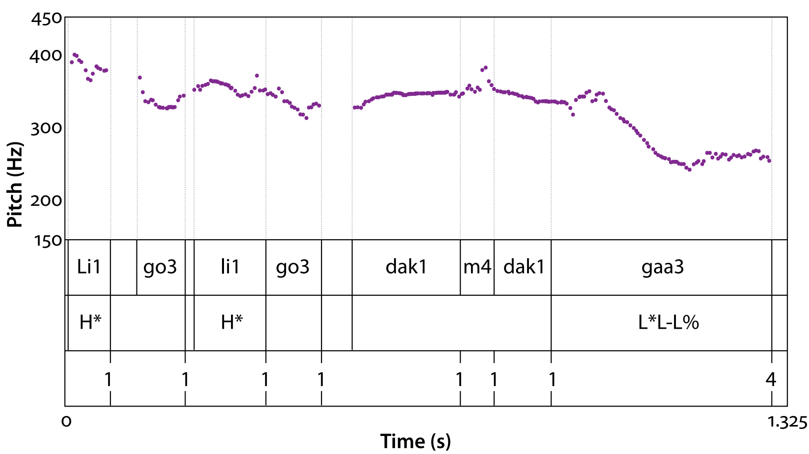
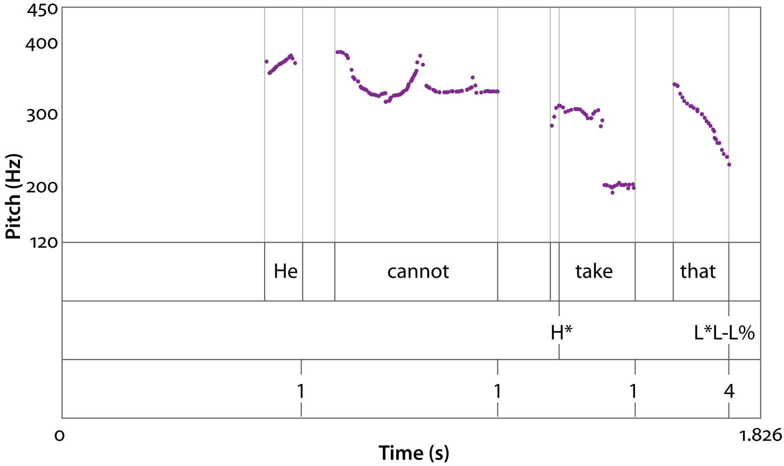


Figure 3. F0 of (6)

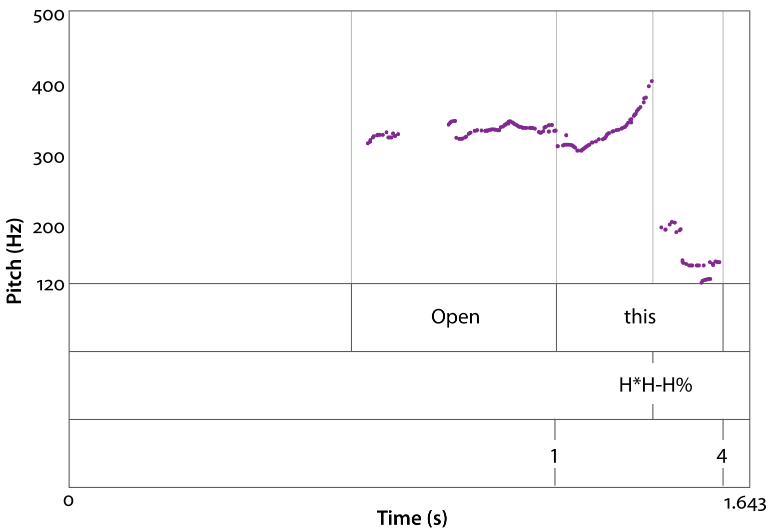
followed by a fall on "that". (8) carries an English intonation with low pitch, followed by a rise on "this". Figure 4 shows their F0.

(7) He cannot take that. (Llywelyn 2;06.20)

(8) Open this? (Llywelyn 2;06.20)



A.



B.

Figure 4. (A) F0 of (7); (B) F0 of (8)

Figure 5 shows the production of intonation patterns by the eight bilingual children at five time-points. We consider an utterance as belonging to either of the two bilingual intonation patterns only when both raters perceived it as the same intonation pattern. If only one rater perceived it as an instance of bilingual intonation while the other did not, that utterance was considered as "other". The category "other" also included target intonation, unknown intonation patterns, and unintelligible utterances. Both intonation patterns were attested in the speech of all eight participants, the pattern HF more frequently than LR.

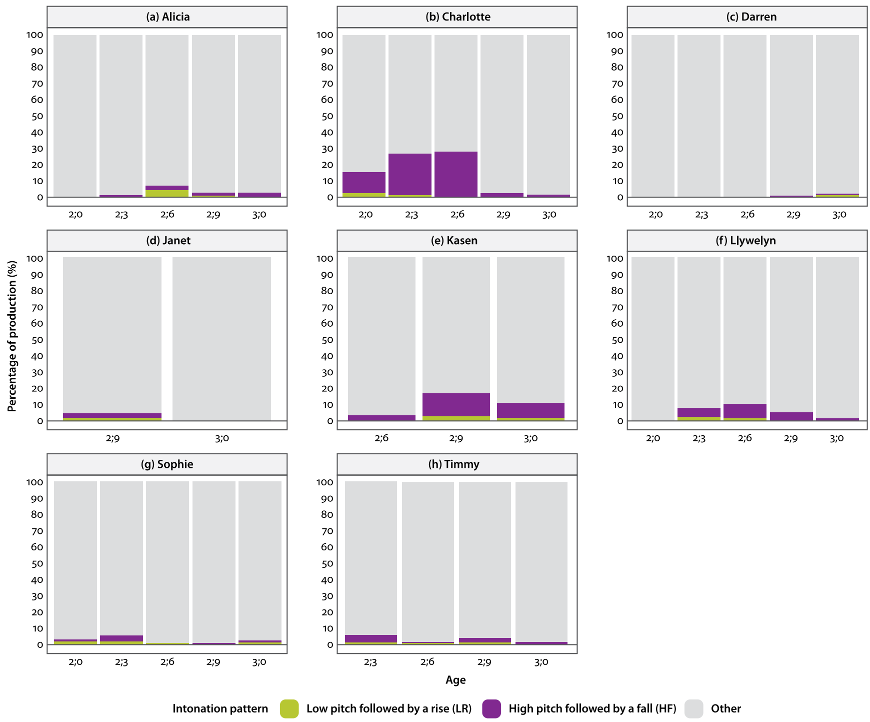


Figure 5. Intonation patterns produced by eight Cantonese-English bilingual children; (a) Alicia; (b) Charlotte; (c) Darren; (d) Janet; (e) Kasen; (f) Llywelyn; (g) Sophie; (h) Timmy

Bilingual intonation was attested most frequently in the English-dominant child Charlotte, who produced it in over 25% of her utterances in the sampled recording at 2;6. In contrast, bilingual intonation was attested in the speech of the Cantonese-dominant children (Alicia, Janet, Llywelyn, Sophie, and Timmy) at much lower frequencies (generally < 10% of utterances).

Regarding the two balanced bilingual children, bilingual intonation was attested frequently in Kasen's speech (~15% of his speech in a recording at 2;9) but rarely in Darren's. This can be attributed to Darren's unusually sparse production of Cantonese SFPs: he produced zero SFPs in the sample recordings at 2;0 and 2;6 and very few tokens at other time-points. We observe an idiosyncratic intonation pattern in Darren's speech – constant high pitch, as in (9) where all syllables have a high pitch (Figure 6). The absence of SFPs seems to preclude a fall at the utterance-final position. Despite the superimposition of English intonation, there are still traits of trochaic stress (high-low pitch pattern) in *jan1wai6* "because", approximating the pitch pattern of the Cantonese lexical tones.

(9) Jan1wai6 mou5 wun6 din6. (Darren 2;09.03)
because not.have change battery
"Because [we] did not change the battery."

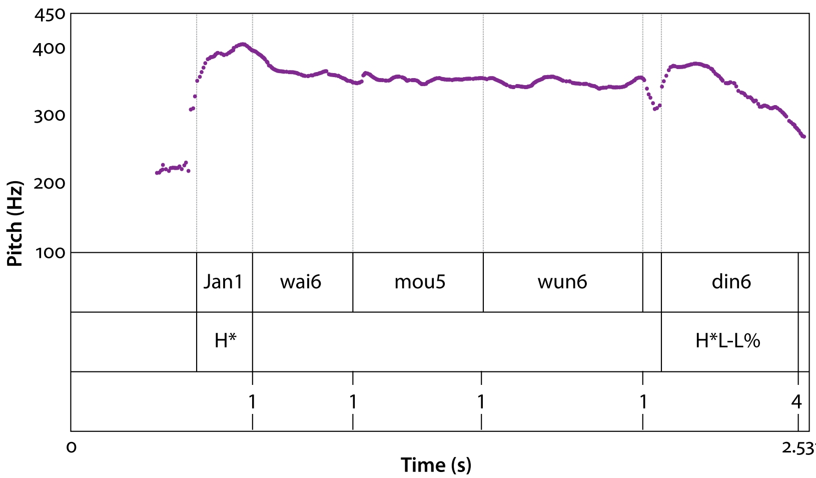


Figure 6. F0 of (9)

Apart from Darren, we observe an inverted U-shaped pattern in the production of bilingual intonation by the other seven children. This pattern is especially conspicuous when the data for the child are available at all five time-points. Alicia, Charlotte, and Llywelyn produced more and more bilingual intonation from ages 2;0 to 2;6, with a decline after 2;9. Although we do not have data for Janet, Kasen, and Timmy at all five time-points, the data available fit the inverted U-shaped pattern. Kasen produced bilingual intonation more frequently at 2;9 than at 2;6 and 3;0. Janet and Timmy showed a gradual decrease in the production of bilingual intonation which could represent regression.

Although bilingual intonation was attested in all children studied, it was produced only a small proportion of the time (average ~5.67% of total utterances). Most of the time, the children produced target-like intonation, which constituted the majority of "other" cases (grey bars) in Figure 5. Apart from target intonation, "other" also includes a small number of unknown intonation patterns and unintelligible utterances.

No bilingual intonation patterns were attested in the monolingual control data listed in Table 3, which implies that the bilingual intonation patterns are exclusive to bilingual children and reflect CLI from English prosody.

3.3 Case illustration – Kasen's m4 hai6 aa3 "no"

Since the bilingual corpus data involve naturalistic speech, there are formidable challenges in reducing the utterances with bilingual intonation to simple acoustic plots (Liberman, 2019). Despite the high variability, we illustrate the two bilingual intonation patterns and compare them with the target intonation using a case study of Kasen's production of the phrase *m4 hai6 aa3* "no".

In the recording at 2;09.01, just before Kasen turned English-dominant, he manifested the temperament of the "terrible twos" (Hughes et al., 2020), frequently saying "no". He often did not differentiate the negative phrases in Cantonese and typically used the phrase *m4 hai6 aa3* (negator be SFP; "no").

Kasen produced 96 tokens of *m4 hai6 aa3* "no" in the recording of 2;09.01. These tokens are phonetically comparable as they were produced by the same speaker on the same day. Four native speakers of Cantonese independently judged the intonation patterns as either HF, LR, "target intonation", or "others". "Target intonation" refers to intonation which is native-like (i.e., native speakers can perceive the tones of the three syllables as the target sequence T4-T6-T3). Fleiss' kappa showed substantial agreement between the raters' judgements, $\kappa = .765$ (95% CI, .716 to .815), $p < .001$. Adopting a stringent criterion, we only consider an utterance as instantiating either of the two bilingual intonation patterns or the target intonation when at least three raters perceived it as the same intonation. After discarding tokens lacking raters' consensus and of poor recording quality, each syllable of all utterances ($n = 65$; usage of data = 67.71%) was segmented. F0 was sampled at ten equal time-points in all syllables. Smoothing Spline ANOVA (SSANOVA) was conducted on the data using *gss* package (Gu, 2014) in R (R Core Team, 2020).

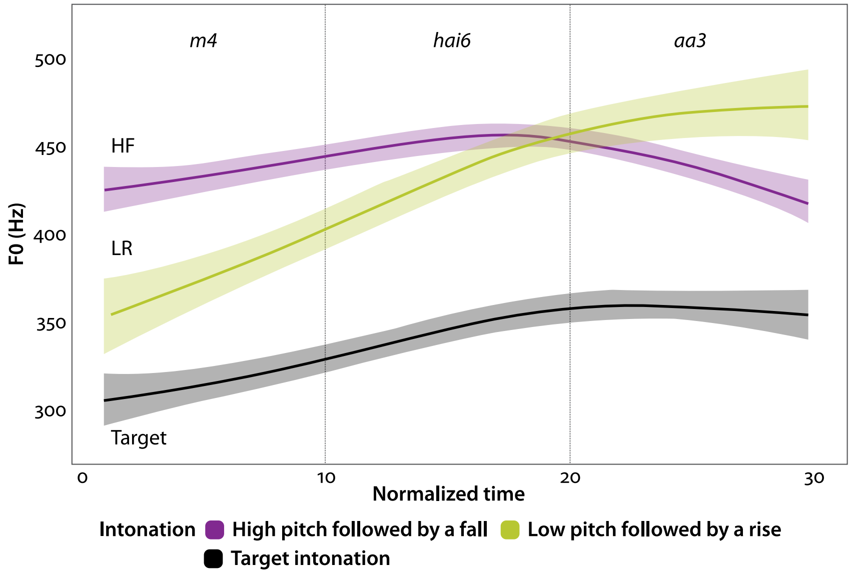
Figure 7(A) shows the SSANOVA results for the F0 of the three intonation patterns attested in Kasen's (2;09.01) production of *m4 hai6 aa3*: HF ($n = 36$), LR ($n = 8$), and "target intonation" ($n = 21$). First, the F0 of the utterance body of the pattern HF is significantly higher than that of the target intonation. The high pitch of the utterance body overrides the low tones of *m4* and *hai6*, which were perceived by the raters as high-level T1. Although the absolute F0 of the SFP *aa3* is higher than that of *aa3* in target intonation, the relative fall in pitch at the SFP was perceived by the raters as low-falling T4. Conversely, the F0 of *m4* in LR is significantly higher than in target intonation. It was not low enough to be perceived as T4 but was perceived as low-level T6 by native Cantonese speakers. The F0 of the SFP *aa3* is also significantly higher than in target intonation, perceived as high-level T1 by the raters.

Since both the canonical contours of LR and the pitches of the lexical tones (T4-T6-T3) in "target intonation" are rising, one may argue that Kasen's wide pitch range covers both the green and black curves in Figure 7(A). To examine this possibility, we compared Kasen's prosody with that produced by monolingual Cantonese-speaking children and an older Cantonese-English bilingual child. The older bilingual participant, Sophie (8;11.23) was recorded incidentally while visiting Kasen's home and her speech contained six tokens of *m4 hai6 aa3* "no".

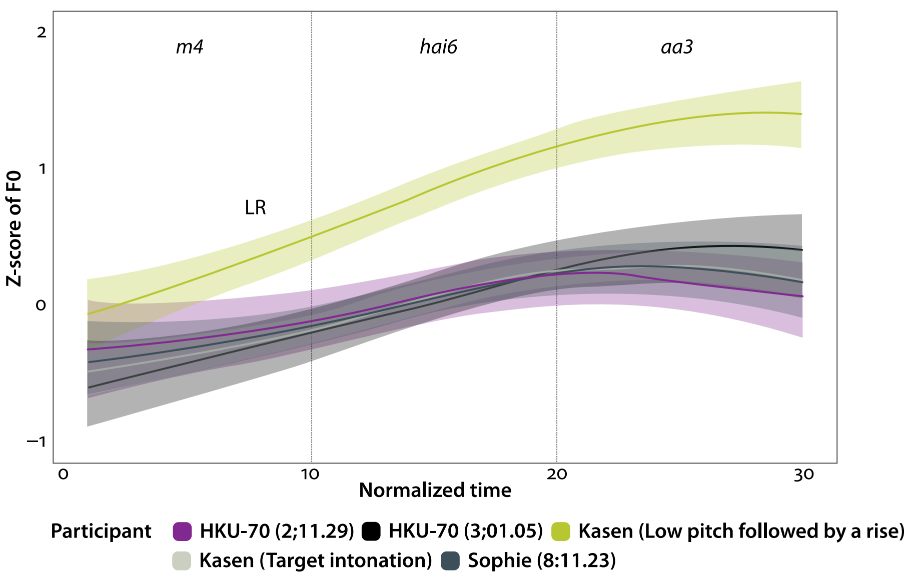
We compared the tones of the phrase *m4 hai6 aa3* produced by the above participants: Kasen's (2;09.01) LR and target intonation, the two monolingual Cantonese-speaking children (2;11.29 and 3;01.05) in the HKU-70 corpus (Fletcher et al., 2000), and the older bilingual child Sophie (8;11.23). Similar to the earlier procedures, all syllables were segmented and time-normalized. Owing to speaker variation, the F0 of each participant was standardized using z-scores. SSANOVA was conducted and the results are reported in Figure 7(B).

Kasen's "target intonation" does not differ significantly from that of the two monolingual children and the older bilingual child (Figure 7(B)). This demonstrates Kasen's ability to produce target prosody and supports the raters' perception of the "target intonation" as target-like. Moreover, the pattern LR (green curve) differs significantly from all other curves in Figure 7(B), supporting the perception of this pattern as distinct from the target. The pitch of *m4* in LR was not low enough to be perceived as T4 by native speakers of Cantonese, while the pitch of *aa3* was too high and was perceived as T1.

Not only does the pattern LR have a rising canonical contour at the utterance body, but the pattern HF also has a slightly rising contour at the utterance body. Raters perceived *hai6* as having a higher pitch than *m4*, even though both syllables were perceived as T1. This can be attributed to the ceiling effects of this highest tone (T1) in Cantonese phonology. The rising contours may indicate an interaction of the HF intonation with the rising sequence of pitch patterns in Cantonese (T4-T6) which reflects Kasen's knowledge of the target tones.



A.



B.

Figure 7. SSNOVA results. (A) F0 of the three intonation patterns attested in Kasen's (2;09.01) *m4 hai6 aa3*; (B) Comparison of F0 of Kasen's bilingual intonation pattern LR with his own target intonation, two monolingual Cantonese-speaking children, and older bilingual child Sophie in the phrase *m4 hai6 aa3*

3.4 Mixed-effects logistic regression

3.4.1 *Predictors*

Based on our hypotheses in Section 1.4, there are three predictors to be compared: language dominance (three levels: Cantonese-dominant, balanced, and English-dominant), use of SFPs (binary variable: yes/no), and grammatical complexity (continuous variable). Since our participants are bilingual children with differing language dominance, their grammatical complexity is measured by the mean MLU for both Cantonese and English contexts[4] to average the grammatical complexities of the child's two languages.

3.4.2 *Model comparison*

Mixed-effects logistic regression was performed using the R *lme4* package (Bates et al., 2015) with bilingual intonation as a binary outcome (yes/no). All models included random intercepts for the variability of utterances, participants, and participants' sex.

Table 5 summarizes all model comparisons between the selected model (M1) and competing additive models (M2-M7), null model (M8), and multiplicative models (M9-M11). The selected model M1 is an additive model including fixed effects of language dominance (LD), grammatical complexity (GC), and use of SFPs (SFP).[5] It is obtained by backward elimination from the global model M1, which is also the best attained model based on the Akaike Information Criterion (AIC), the Bayesian Information Criterion (BIC), deviance, and statistical significance of the change in deviance against the Chi-squared distribution, by performing ANOVA on the competing models (M2-M8). The subsequent stepwise introduction of interactions (M9-M11) is found to be non-significant.

3.4.3 *Mixed-effects logistic regression results*

Table 6 reports mixed-effects logistic regression results of the selected model (M1). Post hoc analyses using estimated marginal means (EMMs) were performed using R *emmeans* package (Lenth, 2021). Using odds ratio (OR), average marginal predicted probabilities of all predictors were calculated using *effects* package (Fox & Weisberg, 2019). Figure 8 shows the post hoc test results and predicted probabilities of all three predictors.

---

4. English MLU was calculated based on words. Cantonese MLU was calculated based on words as represented in the transcript, where the corpus was transcribed by native speakers of Cantonese.

5. R code: glmer(Bilingual intonation~ LD + GC + SFP + (1|Utterance) + (1|Participant) + (1|Sex)

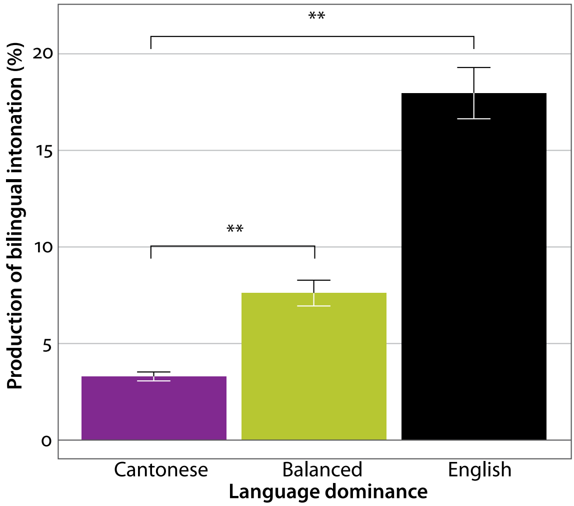
Table 5. Model comparison

Predictor	AIC	BIC	Deviance	χ^2	*df*	*p*
M1. LD + GC + SFP	**2227.146**	**2282.432**	**2211.146**			
M2. GC + SFP	2230.308	2271.773	2218.308	7.163	2	.028
M3. LD + GC	2866.168	2914.544	2852.168	641.022	1	<.001
M4. LD + SFP	2243.456	2291.832	2229.456	18.311	1	<.001
M5. GC	2866.476	2901.031	2856.476	645.331	3	<.001
M6. SFP	2248.343	2282.897	2238.343	27.197	3	<.001
M7. LD	2869.305	2910.770	2857.305	646.159	3	<.001
M8. (Null model)	2869.828	2897.471	2861.828	650.682	4	<.001
M9. LD + GC + SFP + LD*SFP	2230.077	2299.186	2210.077	1.068	2	.586
M10. LD + GC + SFP + LD*GC	2226.415	2295.524	2206.415	4.730	2	.094
M11. LD + GC + SFP + SFP*GC	2228.094	2290.292	2210.094	1.051	1	.305

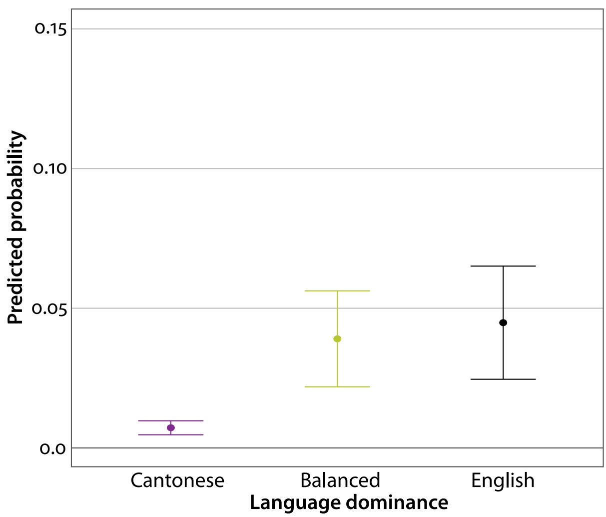
Table 6. Regression results of the selected model (M1)

Factor	β	*SE*	*Z*	*p*	OR (95% CI)
(Intercept)	−2.652	.705	−3.760	<.001	.070 (.018, .281)
LD (Cantonese)	−1.821	.560	−3.251	.001	.162 (.054, .485)
LD (English)	.142	.276	.515	.607	1.153 (.671, 1.980)
SFP (Yes)	4.147	.275	15.071	<.001	63.211 (36.863, 108.389)
GC	−.849	.196	−4.325	<.001	.428 (.291, .629)

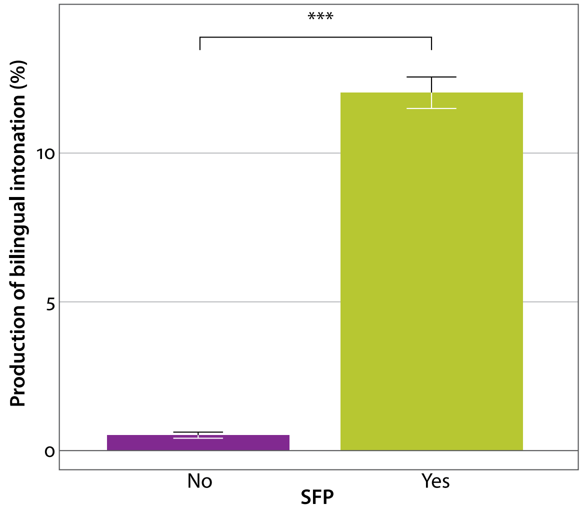
Bilingual intonation was attested in 17.94% of utterances produced by English-dominant children, but in only 7.55% and 3.25% of utterances produced by balanced and Cantonese-dominant children respectively. Post hoc analyses (Figure 8(A)) show that production of bilingual intonation by Cantonese-dominant children is significantly lower than that of balanced ($p = .003$) and English-dominant children ($p = .002$), but no significant difference holds between balanced and English-dominant children ($p = .864$). The predicted probabilities of balanced (.0387) and English-dominant children (.0443) are about 6 and 6.8 times that of Cantonese-dominant children (.0065) respectively (Figure 8(B)). This supports the hypothesis of Yip and Matthews (2000): if the child is more Cantonese-dominant, we predict less CLI from English, and thus less frequent production of bilingual intonation.


**
**
Production of bilingual intonation (%)
20
15
10
5
0
Cantonese
Balanced
English
Language dominance

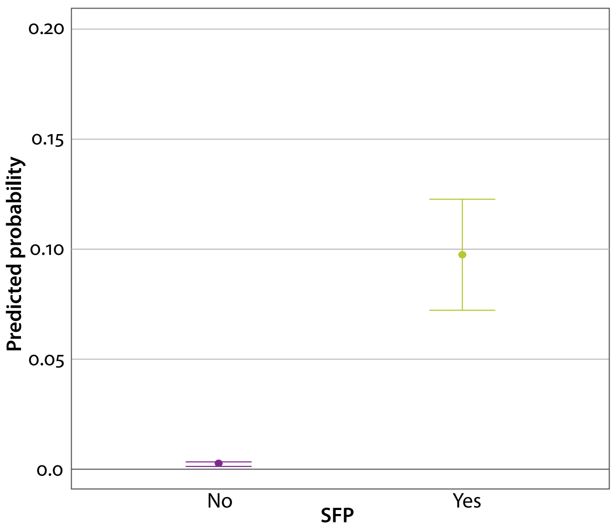
A.


Predicted probability
0.15
0.10
0.05
0.0
Cantonese
Balanced
English
Language dominance

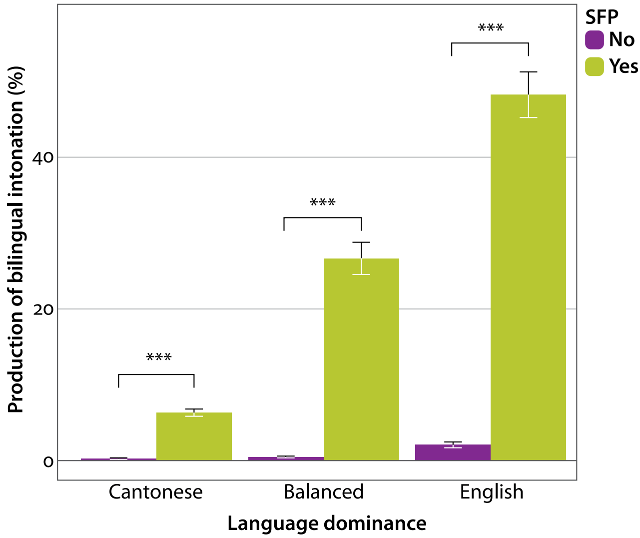
B.


***
Production of bilingual intonation (%)
10
5
0
No
Yes
SFP

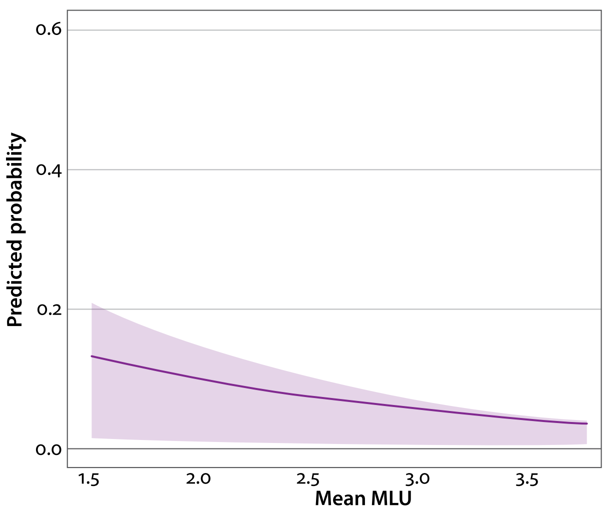
C.


0.20
0.15
0.10
0.05
0.0
Predicted probability
No
Yes
SFP

D.



E.



F.

Figure 8. (A) Post hoc test results of language dominance; (B) Predicted probability of three levels of language dominance; (C) Post hoc test results of SFPs; (D) Predicted probability of use of SFPs (No = absence of SFPs, Yes = presence of SFPs in the utterance); (E) Post hoc test results for use of SFPs at three levels of language dominance; (F) Predicted probability of grammatical complexity
$^{**}p < .01$, $^{***}p < .001$.

Post hoc analyses (Figure 8(C)) show that bilingual intonation was attested far more frequently (26.3 times) in sentences with SFPs (11.75%) than without (.43%) ($p<.001$). Moreover, in all three language dominance groups, bilingual intonation was attested in utterances with SFPs significantly more than those without ($p<.001$ for all three within-group comparisons) (Figure 8(E)). Figure 8(D) shows that the predicted probability of sentences with SFPs showing bilingual intonation is .097, 55.9 times higher than without SFPs (.002). Our results are consistent with the prediction of the Overlap and Interface hypotheses: cross-linguistic transfer of prosody predominantly occurs in the presence of SFPs.

Holding other factors constant, for each unit decrease in grammatical complexity, the odds of producing bilingual intonation increase by a multiple of 2.34 ($OR=.428$, $p<.001$) (Table 6). Figure 8(F) shows that as mean MLU increases, the predicted probability of grammatical complexity in producing bilingual intonation declines, which suggests the children learn that bilingual intonation is a non-target intonation as their grammar grows more complex.

3.5 Mixed-effects quadratic regression

3.5.1 *Procedures*

SFPs are known to be bound up with intonation in Cantonese (Lee, 2021; Sybesma & Li, 2007). We estimate the relationship between usage of SFPs and bilingual intonation by mixed-effects quadratic regression. Usage of SFPs was calculated by the percentage of SFPs used in the total number of utterances in the recordings listed in Table 2 ($n=34$). It was centered (SFP_C) by subtracting the mean (.3914) from each datum to alleviate multicollinearity. The production of bilingual intonation as the response variable (BI) was calculated by the percentage of utterances produced with bilingual intonation in a recording. Two outliers were detected (.1707 and .2698) using the 1.5×Interquartile Range Rule (Tukey, 1977). In comparison to the two next highest values (.1615 and .1527), the outlier .2698 was conspicuously larger and was winsorized to .1707 (highest value in the rest of the dataset). Our model included predictors of SFP_C² (quadratic) and SFP_C (linear) and random intercepts for participants and their sex.[6]

3.5.2 *Results*

There are significant effects of the quadratic term SFP_C² on BI, $\beta=-.350$, $SE=.137$, $t(25.205)=-2.548$, $p=.017$, and the linear term SFP_C on BI, $\beta=.142$, $SE=.036$, $t(27.811)=3.971$, $p<.001$. Overall, the model shows a significant qua-

6. R code: lmer(BI~ I(SFP_C^2) + SFP_C + (1|Participant) + (1|Sex)

dratic association between the SFP_C and BI. Marginal R^2 is .344, and conditional R^2 is .870, indicating that fixed factors alone and the total model can respectively account for 34.4% and 87.0% of variability in the production of bilingual intonation.

Figure 9 shows an inverted U-shaped relationship between SFP_C and BI. At first, as SFP_C increases, BI also increases. However, when SFP_C is around 0% (~39.14% uncentered usage of SFPs), the more frequent usage of SFPs leads to a fall in BI. If the variables SFP_C² and SFP_C in the equation equal 0, BI represents the estimated mean production of bilingual intonation (6.20% of total utterances), which is equal to the intercept ($\beta = .062$, $SE = .022$, $t(2.882) = 7.725$, $p = .021$).

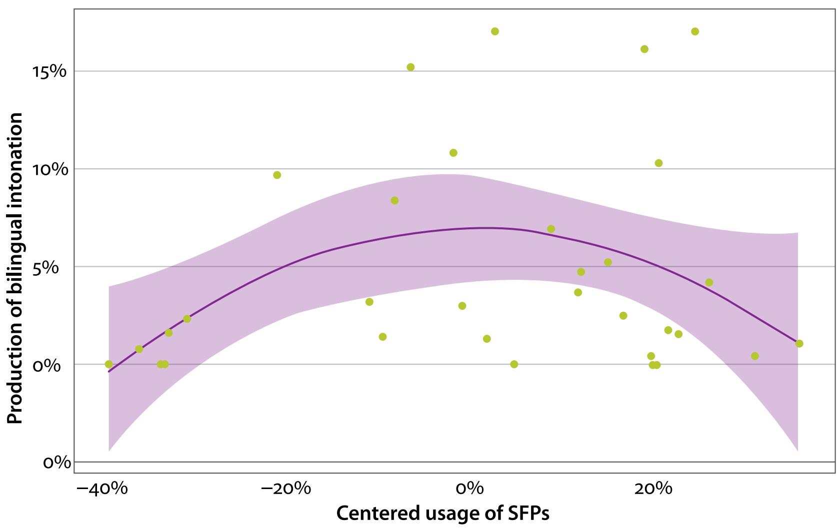


Figure 9. Quadratic regression predicting BI with SFP_C (on percentage scale)

Multiple factors result in this inverted U-shaped pattern which echoes the developmental pattern of the production of bilingual intonation reported in Section 3.2. As the children acquire more SFPs, increasing usage of SFPs results in increased production of bilingual intonation. However, at the apex, more frequent usage of SFPs leads to a fall in the production of bilingual intonation. This can be attributed to a more mature control of prosody as the children learn that bilingual intonation is a non-target prosody. They were also likely showing higher grammatical complexity (i.e., higher mean MLU) at the regressive period since more frequent usage of SFPs would lead to an increase in MLU. The mixed-effects logistic regression results suggest that the predicted probability of producing bilingual intonation decreases as grammatical complexity increases. Multicollinearity between the two factors may lead to a regressive relationship between the usage of SFPs and the production of bilingual intonation.

4. Discussion

This study is the first to investigate the development of intonation in simultaneous Cantonese-English bilingual children. We show that some instances of their intonation in Cantonese and code-mixed utterances were influenced by English. Alongside the target intonation, two bilingual intonation patterns HF and LR are observed in all children studied. Language dominance, use of SFPs, and grammatical complexity are significant predictors of the production of bilingual intonation.

The prominent pattern HF echoes observations reported in previous studies. The superimposition of intonation with high pitch at the utterance body echoes the argument of Mok and Lee (2018) that T1 is the "default tone" in errors in the speech of Charlotte and Llywelyn, while the fall at the utterance-final position leads to the non-target production of SFPs with T4 as observed by Lai (2006). Moreover, Mok and Lee (2018) argued that while Alicia, Sophie, and Timmy, were on a par with their monolingual counterparts in their Cantonese tone production at 2;0, Charlotte and Llywelyn were behind in their Cantonese development until 2;6. This echoes with our findings reported in Figure 5 that the three bilingual children with the most conspicuous production of bilingual intonation were Charlotte, Kasen (not studied by Mok and Lee), and Llywelyn. Such a match between tonal accuracy/acquisition of tones and production of bilingual intonation implies that bilingual intonation contributes to inaccurate production of tones.

We show that Cantonese SFPs are pertinent to the transfer of bilingual intonation. First, bilingual intonation was attested significantly (26.3 times) more in utterances with SFPs than without (Figure 8(C)). Such asymmetrical distribution indicates that bilingual intonation is transferred predominantly in the presence of SFPs. Second, the inflection points of the intonation contours (i.e., fall/rise in the two bilingual intonation patterns) occur at the utterance-final position, even though the span of the intonation contour is not limited to the utterance-final position, and in some cases extends over the whole utterance. The fact that SFPs serve as the domain for intonation realization (fall/rise at the utterance-final position) in bilingual children is consistent with the location of boundary tones signaling intonation in native Cantonese (Xu & Mok, 2011).

We argue that the two bilingual intonation patterns reflect transfer of English intonation to Cantonese and code-mixed utterances. First, the two patterns are included in the ToBI system of English intonation (Ladefoged & Johnson, 2015). Second, the patterns HF and LR resemble the English intonation in (10) and (11) respectively, where the two patterns occur in the parental input to the children. These two examples were produced by the father (FAT), a native speaker of

British English, of the three bilingual siblings (Alicia, Sophie, and Timmy) in the corpus.

(10) FAT: Put it down. (Alicia 2;03.16)

(11) FAT: The dog ran away, huh? [high-rising] (Alicia 2;01.15)

The overall pitch of "Put it" in (10) is high, followed by the particle "down" with a lower pitch. This resembles the pattern HF. In contrast, the overall pitch of "The dog ran away" in (11) is lower even though there is a rise on "away" as a global realization of English intonation, followed by a high-rising tag "huh". This resembles the pattern LR.

Overall, LR/HF is anchored at the phrase-final position, where boundary tones are associated at the right edge of the intonation phrase. This is similar to intonation in Japanese, where phrase-final intonation with various (paralinguistic) meanings is often realized with SFPs (Vance, 2008).[7] However, all raters agreed that the penultimate syllables of the two bilingual intonation patterns are often metrically prominent and stressed (Section 3.2). Specifically, we propose the following phonological analysis involving the interaction of nuclear stress (NS) with bilingual intonation.

Assuming Cantonese lacks NS (Cheung, 2010), no particular syllable should be conspicuously more prominent or "stressed". However, in the bilingual children's speech, we suggest that NS is assigned to the penultimate syllables of the utterances with bilingual intonation, while the SFPs represent post-nuclear unstressed tails.[8] Chomsky & Halle (1968) proposed the Nuclear Stress Rule (NSR) for English:

(12) Assign primary stress to the primary-stressed vowel of the rightmost lexical category in the context.

Since Cantonese is a SVO language, and SFPs follow the object, the words immediately preceding the SFPs are mostly lexical categories (content words). Following (12), the primary-stressed vowels of content words will be assigned NS. Given Cantonese is a highly monosyllabic language without word stress, the rightmost content words, which are typically monosyllables, will receive NS as the nucleus. This gives rise to the perception of the penultimate syllables of utterances with bilingual intonation as stressed.

7. We thank an anonymous reviewer for suggesting this parallel.

8. We thank Bob Ladd for suggesting such an analysis.

Unlike Cantonese, English has word stress and polysyllabic words. Code-mixed utterances with English polysyllabic words immediately preceding the SFPs in the bilingual intonation patterns have different metrical structures. For instance, the trisyllabic word *battery* immediately precedes the SFP *aa3* in (4a). Following (12), NS should be assigned to the primary-stressed vowel *ba-* of *battery* with the remaining syllables analyzed as tail. Intriguingly, Llywelyn stressed the unstressed syllable *-ry* of *battery* in addition to the primary stress on *ba-*. This is not compatible with NSR unless, as we suggest, Llywelyn treated the trisyllabic word *battery* as three monosyllabic words, in accordance with the pervasive monosyllabicity of Cantonese. The NSR assigns NS to the final syllable of the utterance body, similar to the aforementioned Cantonese utterances. The interaction between NS and bilingual intonation showcases CLI at the level of abstract phonological rules and prosodic features.

It has been suggested that some Cantonese SFPs are actually tags (Tang, 2015), similar to English tags (e.g., "huh" in (11)) and that the tones on some SFPs are intonational. This property is similar to English particles which can carry complex intonation. English particles often serve as post-nuclear tails in intonation, allowing multiple targets of phrasal and boundary tones to be realized. For instance, "huh" in (11) can be produced with both high-rising and low-rising tone, the latter encoding a sense of irony. As an intonational language, English has a copious repertoire of intonation. Bilingual children may exploit the close relationship of SFPs with prosody/intonation in transferring English intonation to Cantonese SFPs.

The transfer of English prosody while respecting prosodic features of Cantonese, with the fall/rise of bilingual intonation licensed at the utterance-final position (predominantly on SFPs), demonstrates fusion between the prosodic systems of the two languages, echoing the previously reported fusion of intonation in Turkish-German bilingual children (Queen, 2001), Yami-Mandarin adult bilinguals (Lai, 2018; Lai & Gooden, 2018), and English-Hindi adult bilinguals (Puri, 2013, 2018). Alternatively, the occurrence of bilingual intonation indicates that the bilingual children have not fully acquired the prosodic features of two languages and produce an intermediate form of intonation, extending the locus of intonation backward from its utterance-final realization in native Cantonese. Additionally, the fact that both bilingual intonation patterns can be applied to both questions and statements suggests that the bilingual children have not fully acquired the meanings and function of intonation, though the patterns may serve to signal prosodic phrasing at the end of an utterance.[9] What remains puzzling is that the patterns LR/HF are not consistently interrogative/declarative. They may

9. We thank an anonymous reviewer for suggesting this possibility.

be sourced from English interrogative and declarative intonation, but the bilingual children re-analyze/over-generalize them to other sentence types in Cantonese. Intriguingly, Charlotte's father is a native speaker of New Zealand English where rising intonation in statements is common, but the rising pattern LR (green bars in Figure 5(b)) was not attested frequently in her speech, while the falling pattern HF (purple bars in Figure 5(b)) constituted the majority of her bilingual intonation patterns. It is also noteworthy that even though the accents of English input from parents to the bilingual children are not homogeneous (Charlotte received New Zealand English input from her father, Darren and Kasen received L2 English from their parents), bilingual intonation was attested in all children studied, suggesting that the occurrence of bilingual intonation is due to the interplay between the intonation systems of English and Cantonese rather than to particular English accents.

Additionally, we do not find any specific tones associated with the production of either of the two bilingual intonation patterns. For instance, utterances (3a) and (3b) contain the same words (with identical tone sequences) but were produced with two different bilingual intonation patterns by Llywelyn. Another example is Kasen's *m4 hai6 aa3* "no", where exactly the same phrase was produced with different intonation patterns.

We have also explored how the use of pitch for intonation interacts with SFPs and the pitch patterns of lexical tones. We argue that, despite the superimposition of English intonation, the pitch patterns of Cantonese lexical tones can occasionally be reflected by traits of word stress. For instance, despite the superimposition of the intonation HF, we can hear traits of trochaic stress in *li1go3* "this Cl" in (6), approximating the pitch patterns of the target tones in Cantonese. Another example is (9) where the superimposition of English intonation is so robust that the whole utterance was produced with high pitch, yet there are still traits of trochaic stress in *jan1wai6* "because", approximating the pitch patterns of lexical tones in Cantonese. Both patterns LR and HF had rising canonical contours at the utterance body in Kasen's (2;09.01) *m4 hai6 aa3*. Although the latter pattern showed a generally falling intonation, there was still a slight rise in the utterance body where native adult raters could perceive *hai6* as having a higher pitch than *m4*. These examples demonstrate how the use of pitch for intonation interacts with lexical tones in the bilingual children's speech. Furthermore, there is interaction with English stress. For example, the non-target NS assigned to the unstressed final syllable of the word *battery* in (4a) did not preclude the target assignment of primary stress to the first syllable. The above examples reflect not only children's knowledge of the target Cantonese tones and English word stress, but also their knowledge of not assigning NS to SFPs (functional words).

The inverted U-shaped developmental pattern of the production of bilingual intonation in the current study corresponds to the U-shaped development of tonal accuracy reported in Mok and Lee (2018), because the production of non-target bilingual intonation naturally influences tonal accuracy. While their study excluded SFPs from the analysis, the present study shows an inverted U-shaped quadratic relationship between SFP usage and bilingual intonation. U-shaped development represents re-organization of the relevant ability/system (Werker et al., 2004). This implies fluidity in the prosodic systems of the bilingual children, with interaction between the two languages. Our data showcase how regression can also occur in complex (cross-linguistic) particle-prosody interaction. The production of bilingual intonation is a stage that bilingual, but not monolingual children go through. Bilingual children appear to have taken a different path from monolinguals toward the target (although we lack data to show whether bilingual intonation is expunged as the children grow older).

Besides, the occurrence of bilingual intonation is consistent with the Overlap and Interface hypotheses. Since there are overlapping structures with particles in utterance-final positions in both languages, and pitch is an overlapping prosodic cue, the transfer of English intonation to Cantonese and code-mixed utterances produced by bilingual children is consistent with the Overlap Hypothesis. The fact that bilingual intonation is transferred predominantly in the presence of SFPs is consistent with the Interface Hypothesis, because Cantonese SFPs are linguistic elements situated at multiple language-internal and external interfaces.

Furthermore, we have shown that the more dominant the bilingual child's Cantonese, the less bilingual intonation is produced in Cantonese and code-mixed utterances. Quantitatively speaking, this supports the role of language dominance in determining the magnitude of CLI in prosody; qualitatively speaking, however, unlike the unidirectional syntactic transfer from the dominant to less dominant language in Yip and Matthews (2000), bilingual intonation was transferred from a less dominant (English) to a more dominant language (Cantonese) in the speech of all Cantonese-dominant bilingual children in the present study.

5. Significance and limitations

This is the first study investigating bilingual interaction in intonation in simultaneous bilingual children acquiring a tonal language (Cantonese) and a stress/intonational language (English). It contributes to the understanding of early bilingual acquisition of prosody, particularly when the prosodic systems of the two languages investigated are typologically divergent. It also lays foundations for building models of CLI in prosody.

Furthermore, bilingual intonation, particularly the transfer of English prosody to Cantonese utterances, instantiates the problem of tune-text association raised by Liberman (1975), where English tune, but not Cantonese tune, is aligned with the Cantonese text in bilingual intonation. This may reflect a syntax-prosody integration problem in bilingual children where prosody must be integrated with grammar at the spell-out domain according to the Y-model (Chomsky, 1995). It remains unclear how bilingual children mentally encode the prosody of the two languages and how they integrate it with syntax. Future studies can investigate the psycholinguistic mechanisms of the production of prosody by bilingual children, and how monolingual and bilingual children differ in their encoding of SFPs and prosody. We acknowledge the production of bilingual intonation by the bilingual children may not be attributed solely to CLI from English, because it is unclear whether the monolingual Cantonese-speaking children were matched with the bilingual children in the corpora in terms of cognitive abilities, socioeconomic status, etc.[10] These factors can be better controlled and further examined in future studies.

Another limitation of the current study is that there are more Cantonese-dominant than English-dominant children in the corpus. Nevertheless, the transfer of bilingual intonation was attested in the speech of all Cantonese-dominant children, demonstrating prosodic transfer from a less dominant language to a dominant one. As predicted, bilingual intonation is more frequent in English-dominant children, supporting language dominance as a determinant of the magnitude of CLI. We predict that monolingual English-speaking children will have stronger grounds for producing the intonation patterns based on word stress and sentential functions of intonation.

Funding

This research has been funded by Childhood Bilingualism Research Centre, University of Cambridge-CUHK Joint Laboratory for Bilingualism, and the CUHK-Peking University Joint Research Centre for Language and Human Complexity.

10. We thank an anonymous reviewer for pointing this out.

Acknowledgements

We thank the Editor, Margaret Kehoe, and the anonymous reviewers for their insightful comments and suggestions. We are indebted to Shengli Feng, Jianjing Kuang, Bob Ladd, Charles Lam, Chaak Ming Lau, Tan Lee, Mark Liberman, Peggy Mok, Sze-wing Tang, Sue-mei Wu, and Ka-Fai Yip for discussion and useful input. We recognize the assistance in statistics that Sihan Chen and Jie Yin have provided. We thank Tsun Hei Choi, Charles Lok, Herman Ng, Hailey Ngai, and Jason Tai for their help with the corpus data. Naturally, all errors remain our own.

References

doi Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, *67*(1), 1–48.

Boersma, P., & Weenink, D. (2019). *Praat: doing phonetics by computer* (6.0.47). http://www.praat.org/

Chao, Y. R. (1930). A system of tone-letters. *Le Maître Phonétique*, *45*, 24–27.

doi Cheng, S.-P., & Tang, S.-W. (2022). Syntax of sentence-final particles in Chinese. In C.-R. Huang, Y.-H. Lin, I.-H. Chen, & Y.-Y. Hsu (Eds.), *The Cambridge Handbook of Chinese Linguistics* (pp. 578–596). Cambridge University Press.

Cheung, L. Y. L. (2010). The syntactic manifestation of the nuclear stress rule in Cantonese. *Proceedings of 28th West Coast Conference on Formal Linguistics*. Los Angeles, California.

Chomsky, N. (1995). *The Minimalist Program*. MIT Press.

Chomsky, N., & Halle, M. (1968). *The Sound Pattern of English*. Harper & Row.

doi Chor, W. (2018). Sentence final particles as epistemic modulators in Cantonese conversations: A discourse-pragmatic perspective. *Journal of Pragmatics*, *129*, 34–47.

doi Cohen, J. (1960). A coefficient of agreement for nominal scales. *Educational and Psychological Measurement*, *20*(1), 37–46.

Deuchar, M., & Quay, S. (2000). *Bilingual Acquisition: Theoretical Implications of a Case Study*. Oxford University Press.

doi Döpke, S. (2000). *Cross-Linguistic Structures in Simultaneous Bilingualism*. John Benjamins.

Engemann, H. (2021). How (not) to cross a boundary: Crosslinguistic influence in simultaneous bilingual children's event construal. *Bilingualism: Language and Cognition*, 1–13.

Fletcher, P., Leung, S., Stokes, S., & Weizman, Z. (2000). *Cantonese pre-school language development: a guide*. (Report of a project funded by the Language Fund, 1996–1999).

Fox, A., Luke, K.-K., & Nancarrow, O. (2008). Aspects of intonation in Cantonese. *Journal of Chinese Linguistics*, *36*(2), 321–367.

Fox, J., & Weisberg, S. (2019). *An R Companion to Applied Regression* (3rd ed.). Sage.

Grice, M., German, J. S., & Warren, P. (2020). Intonation systems across varieties of English. In C. Gussenhoven & A. Chen (Eds.), *The Oxford Handbook of Language Prosody* (pp. 285–302). Oxford University Press.

doi Gu, C. (2014). Smoothing Spline ANOVA Models: R Package gss. *Journal of Statistical Software*, *58*(5), 1–25.

Gut, U. (2000). *Bilingual Acquisition of Intonation: A Study of Children Speaking German and English*. Niemeyer.

Hughes, C., Devine, R.T., Mesman, J., & Blair, C. (2020). Understanding the terrible twos: A longitudinal investigation of the impact of early executive function and parent–child interactions. *Developmental Science*, *23*(6), e12979.

Hulk, A., & Müller, N. (2000). Bilingual first language acquisition at the interface between syntax and pragmatics. *Bilingualism: Language and Cognition*, *3*(3), 227–244.

Kontaxi, M., & Chaida, A. (2015). Question intonation of Greek-American simultaneous bilinguals. *Proceedings of the 6th ISEL Conference ExLing 2015*, 33–36. Athens, Greece.

Ladd, D.R. (1996). *Intonational Phonology* (1st ed.). Cambridge University Press.

Ladd, D.R. (2008). *Intonational Phonology* (2nd ed.). Cambridge University Press.

Ladefoged, P., & Johnson, K. (2015). *A Course in Phonetics* (7th ed.). Cengage Learning.

Lai, L.-F. (2018). Intonation in contact: Prosodic transfer and innovation among Yami-Mandarin bilinguals. Ph.D. dissertation, University of Pittsburgh.

Lai, L.-F., & Gooden, S. (2018). Intonation in contact: Mandarin influence in Yami. *Proceedings of Speech Prosody 2018*, 952–956.

Lai, R.Y.K. (2006). Language mixing in an English-Cantonese bilingual child with uneven development. MPhil thesis, The University of Hong Kong.

Lee, J.H.N. (2021). Low boundary tone: Evidence from the acoustic differences between Cantonese sentence-final particles with low-falling tone. *1st International Conference on Tone and Intonation (TAI)*, 176–180.

Lee, J.H.N., Chui, H., Lee, T., Luk, S., Tao, D., & Lee, N.W.T. (2022). Formality in psychotherapy: How are therapists' and clients' use of discourse particles related to therapist empathy? *Frontiers in Psychiatry*, *13*, 1018170.

Lenth, R.V. (2021). *emmeans: Estimated marginal means, aka least-squares means* (1.5.4). https://cran.r-project.org/package=emmeans

Liberman, M. (1975). The Intonational System of English. Ph.D. dissertation, Massachusetts Institute of Technology.

Liberman, M. (2019). Corpus Phonetics. *Annual Review of Linguistics*, *5*, 91–107.

MacWhinney, B. (2000). *The CHILDES Project: Tools for Analyzing Talk* (3rd ed.). Lawrence Erlbaum.

Matthews, S., & Yip, V. (2011). *Cantonese: A Comprehensive Grammar* (2nd ed.). Routledge.

Meisel, J. (1989). Early differentiation of language in bilingual children. In K. Hyltenstam & L. Obler (Eds.), *Bilingualism across the lifespan: Aspects of acquisition, maturity and loss* (pp. 13–40). Cambridge University Press.

Miller, J., & Chapman, R. (1981). The Relation between Age and Mean Length of Utterance in Morphemes. *Journal of Speech, Language, and Hearing Research*, *24*(2), 154–161.

Mok, P.P.K. (2011). The acquisition of speech rhythm by three-year-old bilingual and monolingual children: Cantonese and English. *Bilingualism: Language and Cognition*, *14*(4), 458–472.

Mok, P.P.K. (2013). Speech rhythm of monolingual and bilingual children at age 2;6: Cantonese and English. *Bilingualism: Language and Cognition*, *16*(3), 693–703.

Mok, P.P.K., & Lee, A. (2018). The acquisition of lexical tones by Cantonese-English bilingual children. *Journal of Child Language*, *45*(6), 1357–1376.

Nicoladis, E. (2002). What's the difference between 'toilet paper' and 'paper toilet'? French-English bilingual children's crosslinguistic transfer in compound nouns. *Journal of Child Language*, *29*(4), 843–863.

Paradis, J., & Navarro, S. (2003). Subject realization and crosslinguistic interference in the bilingual acquisition of Spanish and English: what is the role of the input? *Journal of Child Language*, *30*(2), 371–393.

Pérez-Leroux, A. T., Cuza, A., & Thomas, D. (2011). Clitic placement in Spanish-English bilingual children. *Bilingualism: Language and Cognition*, *14*(2), 221–232.

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

Puri, V. (2018). Focus in Indian English and Hindi late and simultaneous bilinguals. *Linguistic Approaches to Bilingualism*, *8*(3), 343–371.

Qian, Y., Lee, T., & Soong, F. K. (2007). Tone recognition in continuous Cantonese speech using supratone models. *The Journal of the Acoustical Society of America*, *121*(5), 2936–2945.

Queen, R. (2001). Bilingual intonation patterns: Evidence of language change from Turkish-German bilingual children. *Language in Society*, *30*(1), 55–80.

Queen, R. (2012). Turkish-German bilinguals and their intonation: Triangulating evidence about contact-induced language change. *Language*, *88*(4), 791–816.

R Core Team. (2020). *R: A Language and environment for statistical computing*. https://www.r-project.org/

Schmitz, K., Patuto, M., & Müller, N. (2012). The null-subject parameter at the interface between syntax and pragmatics: Evidence from bilingual German-Italian, German-French and Italian-French children. *First Language*, *32*(1–2), 205–238.

Serratrice, L. (2013). Cross-linguistic influence in bilingual development. *Linguistic Approaches to Bilingualism*, *3*(1), 3–25.

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

Sorace, A., & Filiaci, F. (2006). Anaphora resolution in near-native speakers of Italian. *Second Language Research*, *22*(3), 339–368.

Strik, N. (2012). Wh-questions in child bilingual acquisition of French: Derivational complexity and cross-linguistic influence. *The Canadian Journal of Linguistics*, *57*(1), 133–151.

Sybesma, R., & Li, B. (2007). The dissection and structural mapping of Cantonese sentence final particles. *Lingua*, *117*(10), 1739–1783.

Tang, S. (2015). A generalized syntactic schema for utterance particles in Chinese. *Lingua Sinica*, *1*(3), 1–23.

The Linguistic Society of Hong Kong. (2002). *Hong Kong Jyut Ping Character Table* (2nd ed.). Hong Kong: The Linguistic Society of Hong Kong.

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*(1), 103–122.

Tse, J. K.-P. (2008). Tone acquisition in Cantonese: a longitudinal case study. *Journal of Child Language*, *5*(2), 191–204.

Tsimpli, I., & Sorace, A. (2006). Differentiating interfaces: L2 performance in syntax-semantics and syntax-discourse phenomena. In D. Bamman, T. Magnitskaia & C. Zaller (Eds.), *Proceedings of the 30th Boston University Conference on Language Development* (pp. 653–664). Cascadilla Press.

Tukey, J. W. (1977). *Exploratory Data Analysis*. Addison-Wesley Pub.

Vance, T. J. (2008). *The Sounds of Japanese*. Cambridge University Press.

Wakefield, J. C. (2020). *Intonational Morphology*. Springer.

Werker, J. F., Hall, D. G., & Fais, L. (2004). Reconstruing U-shaped functions. *Journal of Cognition and Development*, *5*, 147–151.

Xu, B. R., & Mok, P. P. K. (2011). Final rising and global raising in Cantonese intonation. *Proceedings of 17th International Congress of Phonetic Sciences*, 2173–2176.

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

Yip, V., & Matthews, S. (2000). Syntactic transfer in a Cantonese-English bilingual child. *Bilingualism: Language and Cognition*, *3*(3), 193–208.

Yip, V., & Matthews, S. (2006). Assessing language dominance in bilingual acquisition: A case for mean length utterance differentials. *Language Assessment Quarterly*, *3*(2), 97–116.

Yip, V., & Matthews, S. (2007). *The Bilingual Child: Early Development and Language Contact*. Cambridge University Press.

Address for correspondence

Jonathan Him Nok Lee
Department of Linguistics
University of Pennsylvania
3401-C Walnut Street, Suite 300, C Wing
Philadelphia, PA 19104
United States

jonhnlee@sas.upenn.edu
https://orcid.org/0000-0003-3990-8801

Co-author information

Regine Yee King Lai
Department of Linguistics and Modern Languages
The Chinese University of Hong Kong
ryklai@cuhk.edu.hk
https://orcid.org/0000-0003-3764-5892

Stephen Matthews
School of Humanities
The University of Hong Kong
matthews@hku.hk
https://orcid.org/0000-0001-7683-8051

Virginia Yip
Department of Linguistics and Modern Languages
The Chinese University of Hong Kong
vcymatthews@cuhk.edu.hk
https://orcid.org/0000-0002-0164-3484

Publication history

Date received: 13 August 2021
Date accepted: 16 February 2023
Published online: 11 April 2023