

ARTICLE

Age-related changes in lexical tones and intonation in Cantonese infant-directed speech: A longitudinal study

Luchang Wang¹  and Patrick C. M. Wong^{2,3}

¹Department of Applied Linguistics, Xi'an Jiaotong-Liverpool University, Suzhou, China; ²Department of Linguistics and Modern Languages, The Chinese University of Hong Kong, Hong Kong, China; and ³Brain and Mind Institute, The Chinese University of Hong Kong, Hong Kong, China.

Corresponding authors: Luchang Wang and Patrick C. M. Wong; Emails: Luchang.Wang@xjtlu.edu.cn; p.wong@cuhk.edu.hk

(Received 16 May 2023; revised 11 July 2024; accepted 04 August 2024; first published online 27 September 2024)

Abstract

This longitudinal study investigated modifications in lexical tones and intonation in Cantonese infant-directed speech (IDS) to children aged 15 and 23 months. The results showed that to children at both ages, mothers increased intonational pitch height and pitch variability across utterances, and produced lexical tones with generally larger tonal space and greater intra-talker tone variation within categories in IDS compared to adult-directed speech. No significant changes were found in either lexical tones or intonation in IDS between the two ages. In addition, positive correlations were found between the degree of age-related changes in tonal space and intonational exaggerations in IDS as children grow older. The findings were discussed with a focus on the co-occurrence of an increase in tone variation along with tonal space expansion, the age-related changes in lexical tones and intonation, and the associations between the lexical and prosodic pitch modifications.

Keywords: infant-directed speech; lexical tones; intonation; age-related changes; Cantonese

Introduction

Infant-directed speech (IDS), the type of speech directed towards infants and children, usually constitutes the majority of the language input that these young learners receive (Soderstrom, 2007). IDS is largely characterized by its acoustic-phonetic features. Compared to adult-directed speech (ADS), IDS typically features higher overall pitch and larger pitch variability across and within utterances (Fernald et al., 1989; Jacobson et al., 1983; Stern et al., 1983). At the phonemic level, studies across languages have shown that IDS involves an expansion of the acoustic vowel space, which is usually measured as the area of the vowel triangle formed by plotting the first and second formants (F1 and F2) of the three most peripheral vowels /a/, /i/, and /u/ in two-dimensional space (Burnham et al., 2002; Kuhl et al., 1997; Marklund & Gustavsson, 2020; but see Benders, 2013; Englund & Behne, 2005). This

© The Author(s), 2024. Published by Cambridge University Press. This is an Open Access article, distributed under the terms of the Creative Commons Attribution licence (<http://creativecommons.org/licenses/by/4.0>), which permits unrestricted re-use, distribution and reproduction, provided the original article is properly cited.

has been considered evidence for a didactic linguistic function of IDS to facilitate child language learning (Burnham et al., 2002; Kuhl et al., 1997). On the other hand, an increase of intra-talker within-category vowel variation has also been observed, which may counteract the effects of vowel space expansion on exaggerating vowel contrasts, thus casting doubt on the linguistic function of phonemic modifications in IDS (Cristia & Seidl, 2014; McMurray et al., 2013; Miyazawa et al., 2017; Rosslund et al., 2022). In addition, longitudinal studies have demonstrated the constancy of vowel space expansion over the first two years of children's lives (Hartman et al., 2017; Kalashnikova & Burnham, 2018; see Cox et al., 2023 for a meta-analysis investigation).

Despite the extensive studies on vowels, there is a notable insufficiency in research on lexical tones, the phonemic categories used in tone languages such as Cantonese and Mandarin Chinese. While a similar expansion of tonal space has generally been reported in IDS compared to ADS (Tang et al., 2017; Xu Rattanasone et al., 2013), very limited knowledge has been obtained regarding the change in intra-talker within-category tone variation (Wang et al., 2021). Only a small number of studies on lexical tones have engaged in longitudinal observations (Han et al., 2018; Liu et al., 2009; Xu Rattanasone et al., 2013), gaining a restricted understanding of how lexical tones in IDS may change as children develop.

Furthermore, lexical tones are unique in that they are instantiated by changes in pitch (or acoustically f_0 , i.e., fundamental frequency), which is also the primary acoustic cue for intonation. Therefore, in IDS, pitch realization at the phonemic level (lexical tones) must be associated with prosodic pitch movement at the utterance level (intonation). However, to date, few IDS studies have concurrently investigated modifications on lexical tones and intonation, not to mention their relationships (Kitamura et al., 2002; Wang et al., 2021).

Lexical tones as phonemes are crucial in determining lexical meaning. For young learners of tone languages, mastery of lexical tones is an important step in building their vocabulary and language skills (W. Ma et al., 2017; Wang et al., 2022). Hence, it is necessary to enhance understanding of lexical-tone modifications in IDS. The current study collected longitudinal data on Cantonese IDS addressed to children in the early to late second year of life. With the aim to shed light on the unresolved questions related to lexical tones, this study 1) investigated age-related changes in tone modifications in IDS, 2) delved into not only tonal space but also within-category tone variation, 3) and concurrently analyzed lexical tones and intonation, and the relationship between their changes over time.

Lexical tones in IDS

In general, research on lexical tones in IDS has shown an expansion of the overall tonal space or enhancements of the acoustic differences between tone pairs in various tone languages, such as Mandarin (Han et al., 2018; Liu et al., 2007; Tang et al., 2017), Cantonese (Li, 2022; Wang et al., 2021; Xu Rattanasone et al., 2013), and Hakka (Cheng & Chang, 2014). These observations were made in IDS directed to children from as young as 3 months to around 2 years old.

Age-related changes

A small number of longitudinal studies have traced the developmental trajectory of lexical-tone modifications in IDS as children grow older. Han et al. (2018) reported a decrease in maximum and minimum pitch of lexical tones in Mandarin IDS to

24-month-olds compared to 18-month-olds, with only the younger age group showing significantly higher maximum and minimum pitch than in ADS. On the other hand, no significant difference in pitch range of lexical tones was observed between the two ages. Liu et al. (2009) compared Mandarin IDS to infants aged 7–12 months with IDS to children aged 5 years, and revealed a reduction in the mean pitch differences between the high and low tones over this time period. The findings on Cantonese IDS mainly came from Xu Rattanasone et al. (2013) on infants within the first year of life. Xu Rattanasone and colleagues measured tonal space by calculating the area of the tone triangle, which was established based on the average onset and offset f_0 of the three peripheral lexical tones of Cantonese – namely, Tone1, Tone2, and Tone4. They examined tonal space in IDS to infants at 3, 6, 9, and 12 months, and observed a decreasing trend from 3 to 12 months of age. Another study on Cantonese (P. Wong & Ng, 2018) gathered IDS data from mothers with their children at 1, 2, 3, and 5 years of age, and found that the pitch height measurements of most lexical tones had negative correlations with the age of the children.

With the limited number of longitudinal studies, our current understanding of age-related changes in lexical-tone modifications in IDS remains inadequate. Two of these studies (Han et al., 2018; P. Wong & Ng, 2018) focused on the pitch cues of individual tones without delving into the distinctions between different tone categories due to their research scope. The other two studies that measured the acoustic contrasts among tone categories (Liu et al., 2009; Xu Rattanasone et al., 2013) generally demonstrated a declining trend in tone contrast enhancements as children's age increased. This trend was specifically observed in Cantonese IDS within the first year of life and in Mandarin IDS from preverbal to preschool ages. The present study sought to enrich our understanding of the developmental trajectory of lexical-tone modifications in Cantonese IDS by testing an understudied age range, i.e., from the early to late second year of children's lives.

Intra-talker within-category tone variation

The findings of expanded tonal and vowel space in IDS have led to the proposal of the hyperarticulation hypothesis that phonemic contrasts are hyperarticulated in IDS to serve a didactic function to facilitate child language learning (Kuhl et al., 1997; Werker et al., 2007; Xu Rattanasone et al., 2013). This is based on the assumption that larger acoustic space in IDS indicates more distinguishable phonemic categories. However, this assumption has been challenged by the findings that IDS exhibits greater intra-talker variation within vowel categories than ADS (Cristia & Seidl, 2014; McMurray et al., 2013; Miyazawa et al., 2017; Rosslund et al., 2022). When controlling for other acoustic variables, greater within-category variation results in more overlap among phonemic categories, leading to less distinct contrasts. When taking into account both the overall acoustic space and within-category variation, studies have reported no evidence of enhanced vowel distinctions (Cristia & Seidl, 2014; McMurray et al., 2013; Miyazawa et al., 2017).

To fully examine the discriminability of tone categories in IDS, research has to be conducted not only on tonal space, but also on within-category tone variation. To our knowledge, only one study (Wang et al., 2021) probed into this dimension of tone modifications in IDS. It reported an increase in within-category tone variation in individual Cantonese-speaking caregivers' IDS to their 15-month-old children. By collecting data at

two age points (15 and 23 months), the current study aimed to expand our knowledge on the modification concerning tone variation in IDS to older children and its developmental trajectory during the second year of children's lives.

Intonation in IDS

As summarized at the beginning, intonational exaggerations in IDS typically involve an increase in overall pitch height, and pitch variability within and across utterances compared to ADS. These exaggerations on pitch at the utterance level are believed to mainly fulfill a communicative role by capturing and regulating children's attention, and conveying communicative intentions such as positive emotions (Cooper et al., 1997; Fernald, 1989). While the rise in pitch height is a robust feature seen across various languages (Fernald et al., 1989), including Cantonese (Wang et al., 2021; N. Xu & Burnham, 2010), pitch variability within utterances, often measured as the range between the highest and lowest pitch in individual utterances, does not consistently exhibit an increase. Studies on tone languages have reported either an expanded pitch range (Grieser & Kuhl, 1988 in Mandarin), or no significant changes in pitch range (Kitamura et al., 2002 in Thai; Papoušek & Hwang, 1991 in Mandarin), or even a reduced pitch range (N. Xu & Burnham, 2010 in Cantonese) in IDS compared to ADS. A more recent study (Wang et al., 2021) employed a different metric, which involved calculating the standard deviation of pitch at multiple points within each utterance, and found no significant difference in pitch variability within utterances between IDS and ADS in Cantonese. This study also confirmed an increase in pitch variability across utterances in Cantonese IDS, which was quantified using the standard deviation of mean pitch across utterances.

Age-related changes

Longitudinal studies on intonational exaggerations in IDS have yielded mixed results. A majority of the studies have reported a reduction of intonational exaggerations in IDS with increasing age of children (see Cox et al., 2023 for a meta-analysis investigation). As a representative instance, Stern et al. (1983) found that intonational exaggerations on pitch height and range in English IDS peaked when infants aged 4 months and decreased as they reached 1 and 2 years old. However, there are also studies failing to detect age effects on intonation in IDS. Narayan and McDermott (2016) measured pitch height and range in IDS of Tamil, Tagalog, and Korean addressed to children aged 4 to 15 months, and reported no significant age effects on these measurements. Similarly, Kalashnikova and Burnham (2018) observed no significant differences in pitch height in English IDS to children aged 7, 9, 11, 15, and 19 months. Moreover, one study on IDS of Dutch (Benders, 2013) even reported larger pitch height and greater pitch range when speaking to 15-month-olds than to 11-month-olds.

While most previous research on this topic focused on non-tone languages, some studies included investigations into tone languages. Kitamura et al. (2002) revealed a decrease in pitch height in Thai IDS from 9 to 12 months of age. Han et al. (2020) reported a decrease in intonational exaggerations in IDS of Mandarin Chinese from 18 to 24 months of age; pitch height and range were only significantly exaggerated in IDS to 18-month-olds, but not 24-month-olds, compared to ADS. Despite evidence of a decreasing trend in intonational exaggerations in IDS of Thai and Mandarin, there is still a gap in knowledge regarding Cantonese IDS specifically. Moreover, previous

longitudinal studies have mainly focused on the overall pitch height and pitch range within utterances, paying little attention to the pitch variability across utterances in IDS. The present study provided longitudinal observations of intonational exaggerations, including pitch variability across utterances, in Cantonese IDS to children in the early to late second year of life, which will contribute to a more comprehensive understanding of the developmental trajectory of intonation in IDS.

Lexical tones and intonation

As noted above, lexical tones share the prosodic cue, pitch, with intonation. Hence, acoustic realization of lexical tones needs to accommodate – and is inevitably influenced by – intonational pitch modulation. For example, lexical tones carried by syllables at the utterance-final position are often compromised under the effects of large intonational modulation at this position (J. K.-Y. Ma et al., 2006). Accordingly, in IDS, how lexical tones are modified may also be related to intonational pitch exaggerations for communicative purposes. However, limited research on IDS has looked into the associations between lexical tones and intonation.

The findings from two studies have suggested the possibility that lexical-tone modifications in IDS may be driven by intonational effects. Tang et al. (2017) found that tonal space expansion only occurred at the utterance-final position in Mandarin IDS, matching the pattern of pitch exaggerations at this position in happy speech where intonation would be exaggerated to convey the positive emotion. P. Wong and Ng (2018) explored the errors made by native adult listeners when perceiving lexical tones in IDS. They discovered that these error patterns were similar to those found in the perception of lexical tones in the utterance-final position, where pitch is heavily influenced by intonational modulation. These findings have lent support to the prosodic hypothesis that IDS mainly serves a communicative function through prosodic exaggerations, and modifications of phonemes are unintended byproducts of the prosodic modulations.

This hypothesis has been further corroborated by Wang et al. (2021), a more recent study examining lexical tones and intonation in tandem. It reported significant positive correlations between Cantonese caregivers' intonational exaggerations (particularly the increase of pitch variability) and lexical-tone modifications (including the expansion of tonal space and increase of within-category tone variation) in IDS. Caregivers who exaggerated prosody more at the utterance level tended to have more expanded tonal space and produce more varied tone tokens. This previous study collected IDS data from only one age point of children. As children develop, caregivers may adjust the degree of intonation exaggerations in their IDS (Cox et al., 2023; Han et al., 2020). According to the prosodic hypothesis, the degree of lexical-tone modifications in IDS, as byproducts of the intonational modulations, would change correspondingly. The present study with longitudinal data would enable an investigation into whether the age-related changes in lexical tones and intonation in IDS positively correlate, as predicted by the prosodic hypothesis.

The current study

This study collected recordings of IDS from Cantonese-speaking mothers with their children when the children were around 15 and 23 months old. The second year of life was targeted considering that children learn words rapidly showing a spurt of vocabulary

during this developmental period (Goldfield & Reznick, 1990). Lexical tones as phonemic categories differentiate lexical meanings. Hence, it is crucial to understand how lexical tones are modified in speech input to children throughout this critical period of lexical learning. To date, longitudinal research into this time period has been scarce for lexical tones as well as intonation, especially in the context of Cantonese. The onset of our observation was set at 15 months of age to replicate the findings on lexical tones (especially the understudied within-category tone variation) in IDS to children at this age reported by Wang et al. (2021), before proceeding to the unexplored later developmental stage.

In addition to IDS, ADS was also recorded from the mothers as the baseline for comparison. For IDS at each age point as well as ADS, lexical tones were examined from two aspects - tonal space and intra-talker within-category tone variation; intonation was also analyzed by measuring the overall pitch height, and pitch variability within and across utterances. Three specific research questions were addressed based on the measurements: 1) compared to when speaking to adults, how Cantonese-speaking mothers modify lexical tones and intonation when speaking to their children at the two age points; 2) how the lexical-tone and intonational modifications in IDS differ between the two age points; 3) whether the age-related changes in lexical tones in IDS correlate with the age-related changes in intonation.

The study aimed to enrich our knowledge about the lexical and prosodic pitch modifications in IDS of a tone language mainly in three aspects: 1) to draw the developmental trajectories for lexical-tone and intonational modifications in IDS during the second year of children's lives when they undergo the vocabulary spurt; 2) to uncover whether an increase of tone variation consistently co-occurs with the expansion of tonal space in IDS showing a similar developmental pattern; 3) to reveal whether lexical tones and intonation in IDS co-vary with increasing age of children, as predicted by the prosodic hypothesis.

Method

Participants

Nineteen Cantonese-speaking mothers and their Cantonese-learning monolingual children (8 males and 11 females) were examined when the children were around 15 months ($M = 1;02.20$; $SD = 11.1$ days) and 23 months ($M = 1;10.23$; $SD = 38.54$ days) of age. They lived in Hong Kong and were recruited through social media. All the mothers and children had no known mental, sensory, or language impairments. Prior to participating, the mothers provided written informed consent, which was approved by The Joint Chinese University of Hong Kong - New Territories East Cluster Clinical Research Ethics Committee.

Stimuli and materials

Six target words (see Table 1) were elicited in both IDS and ADS, with each word corresponding to one of the six Cantonese lexical tones illustrated in Figure 1. Cantonese tones consist of three level tones (Tone1: high level; Tone3: mid level; Tone6: low level), two rising tones (Tone2: high rising; Tone5: low rising), and one falling tone (Tone4: low falling). To minimize the influence of phonetic environment on the lexical tones, the target words were chosen to share the same vowel and have a voiceless plosive or fricative

Table 1. The target words and the toys used to elicit them.

Target words in Chinese	Tone	IPA	Gloss	Toys to elicit the target words
飛	Tone1	[fei]	Fly	Miniature planes; toy birds
畀(俾)	Tone2	[pei]	Give	Furry toys that can be given from one person to another
四	Tone3	[sei]	Four	Toy packs containing four animals/people
旗	Tone4	[k ^h ei]	Flag	Miniature national flags; learning cards depicting national flags; triangular flag bunting
被	Tone5	[p ^h ei]	Quilt	A miniature doll quilt
鼻	Tone6	[pei]	Nose	Animal nose matching boards; an elephant toy

Notes. IPA: International Phonetic Alphabet.

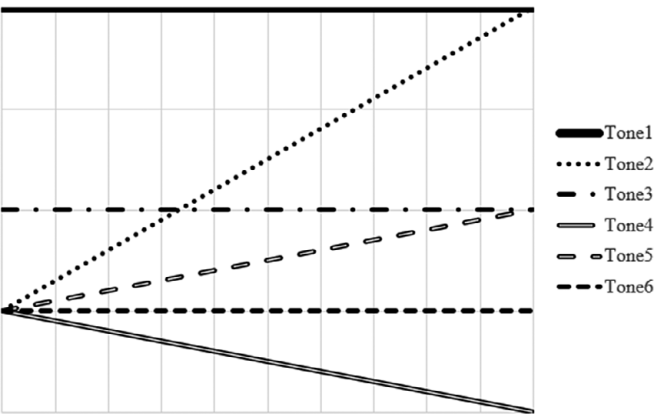


Figure 1. The six Cantonese lexical tones.

as the initial consonant; their semantic appropriateness for mothers to use with young children was also considered. To elicit the target words, the mothers were provided with six sets of toys labeled with the corresponding Chinese characters (see Table 1) during mother-child interaction.

Recording procedure

The IDS recording took place when the children were around 15 and 23 months old. During the recording, the mothers and children were left alone in a sound-attenuated booth where they played and interacted face-to-face with the toys. The mothers were asked to speak naturally and use the target words when appropriate during the interaction, as they would at home. The six sets of toys were presented one at a time in a random order. To record the audio, the mothers wore a head-mounted microphone (Audio-Technica BP894) that was connected to a laptop (MacBook Air) via an audio interface (Roland Quad-Capture). The recordings were made with a sampling rate of 44100Hz and a precision of 16 bits per sample.

The ADS recording was carried out only once at the initial age point since we did not anticipate any changes in ADS associated with the increase of children's age. Following the IDS recording, the mothers were interviewed in Cantonese by an adult native Cantonese speaker. To elicit the target words in ADS, the mothers were asked questions about the words and toys, such as "In what situations would you use the word ____ when speaking to your child in daily life?" and "Do you think your child likes the toy ____?". In the interview, the mothers sometimes imitated how they spoke to their children; these utterances were excluded from ADS analysis. The same recording equipment and configuration were utilized for ADS as for IDS.

To ensure adequate tokens for lexical-tone analysis, all the recordings were monitored by an experimenter outside the booth using headphones attached to the audio interface. We would not proceed to a second target word until we had collected a minimum of 10 and 8 utterances containing the current target word for IDS and ADS respectively. A lower criterion was applied to ADS, given the greater difficulty in eliciting the target words in the interview. The total duration of the recording was kept within 30 minutes.

Data analysis

The analyses described below were conducted on every participant's IDS at each age point as well as their ADS. Using Praat (Boersma & Weenink, 2017), the recordings were segmented and labeled to identify the target words and utterances that contained them. Following the approach of earlier studies (Fernald et al., 1989; Kitamura et al., 2002), an utterance was defined as a speech segment that was preceded and followed by a pause or non-speech lasting at least 300 milliseconds. We attempted to extract up to 10 utterances for each target word, or the maximum number available in cases where there were not sufficient utterances, particularly in ADS recordings. Table 2 displays the average number of extracted utterances per target word and the number of target words contained in the utterances. The target words were analyzed for lexical tones, while intonation analysis utilized all the utterances combined. On average, each utterance contains 9.75 syllables ($SD = 4.68$) in IDS to children aged 15 months, 10.43 syllables ($SD = 5.14$) in IDS to children aged 23 months, and 12.2 syllables ($SD = 6.83$) in ADS. The criteria for inclusion of an utterance were as follows: 1) free of any noise, such as that produced by toys or children's crying; 2) not disrupted by children's vocalizations; 3) containing not only the

Table 2. The average number of utterances and target words analyzed for IDS at 15 months, 23 months, and ADS.

Target words	Tone	Number of utterances			Number of target words		
		IDS-15m	IDS-23m	ADS	IDS-15m	IDS-23m	ADS
飛	Tone1	10	9.95	9.11	13.63	13.74	10.79
界(俾)	Tone2	9.89	10	9.05	12.21	11.68	11.16
四	Tone3	10	10	9.21	11.16	11.79	10.58
旗	Tone4	9.95	9.89	7.47	11.21	11.16	8.26
被	Tone5	10	10	8.63	14.16	13.74	9.68
鼻	Tone6	10	9.95	8.42	15.47	16.63	10.37

target word. To ensure accurate pitch analysis, we carefully marked the onset and offset of the target words and utterances manually, following the criterion of previous studies (Liu et al., 2007) which involved identifying the zero crossing point of the first and last pulse when F1 and/or F2 were visible on the spectrogram. To prepare for lexical-tone and intonation measurements, f0 data were obtained from Praat with the help of a script 'prosodypro' (Y. Xu, 2013). The f0 values were calculated using the vocal cycle marks generated by Praat with a sampling rate of 100Hz. With the assistance of 'prosodypro', manual corrections were applied to add missing marks and remove redundant marks. All f0 data were converted to the Equivalent-rectangular-bandwidth-rate (ERB) scale to better reflect f0 variations from the auditory perspective (Hermes & van Gestel, 1991).

Lexical tones were assessed from two aspects, i.e., tonal space and within-category tone variation. For tonal space, following Xu Rattanasone et al. (2013), we first quantified the overall tonal space by the area of the tone triangle which was created by plotting the averaged onset and offset f0 of the three peripheral lexical tones in Cantonese, i.e., T1, T2, and T4, in two-dimensional space.

Furthermore, to explore the dispersion of each tone category from the center of the tonal space, we adopted a metric based on f0 data collected at 10 equally spaced time points along individual tone contours. Firstly, the overall central tone contour was calculated by averaging the f0 of all the tone contours at each of the 10 time points (see Figure 2A). Secondly, the central tone contour of each tone category was obtained by averaging the f0 of all the contours of that tone at each of the 10 time points (see Figure 2B). Then, for each tone category, the absolute difference between the f0 of its central tone contour and the overall central tone contour was calculated at each time point and then averaged to quantify the degree of its dispersion. This metric adapted from Zhao and Jurafsky (2009) also reflects the magnitude of tonal space: greater dispersion of individual tone categories from the center would result in a larger tonal space. In comparison to the tone triangle area denoting the overall tonal space, it enables assessment of the dispersion of individual tones. Besides, using data collected from multiple time points allows for a better consideration of the temporal information of lexical tones.

For within-category tone variation of each lexical tone, it was computed by first calculating the absolute difference between the f0 of each contour of that tone and its central tone contour at each time point, and then averaging across all the contours and time points.

Three aspects of intonation were measured for every participant, as listed and explained below:

- 1) Pitch height of utterances: the f0 data sampled along each utterance were averaged to determine its mean, which was then averaged across all utterances.
- 2) Pitch variability within utterances: the standard deviation (SD) of the f0 data sampled along each utterance was computed and then averaged across all utterances. This was employed instead of pitch range (= maximum f0 - minimum f0 in individual utterances) due to its advantage of considering more than just the highest and lowest f0 in utterances.
- 3) Pitch variability across utterances: the SD of the mean f0 of all the selected utterances.

Linear mixed-effects (LME) models were used to compare lexical tones and intonation among the three conditions – IDS collected at 15 months old (IDS-15m), IDS collected at 23 months old (IDS-23m), and ADS. The analyses were conducted in the R environment

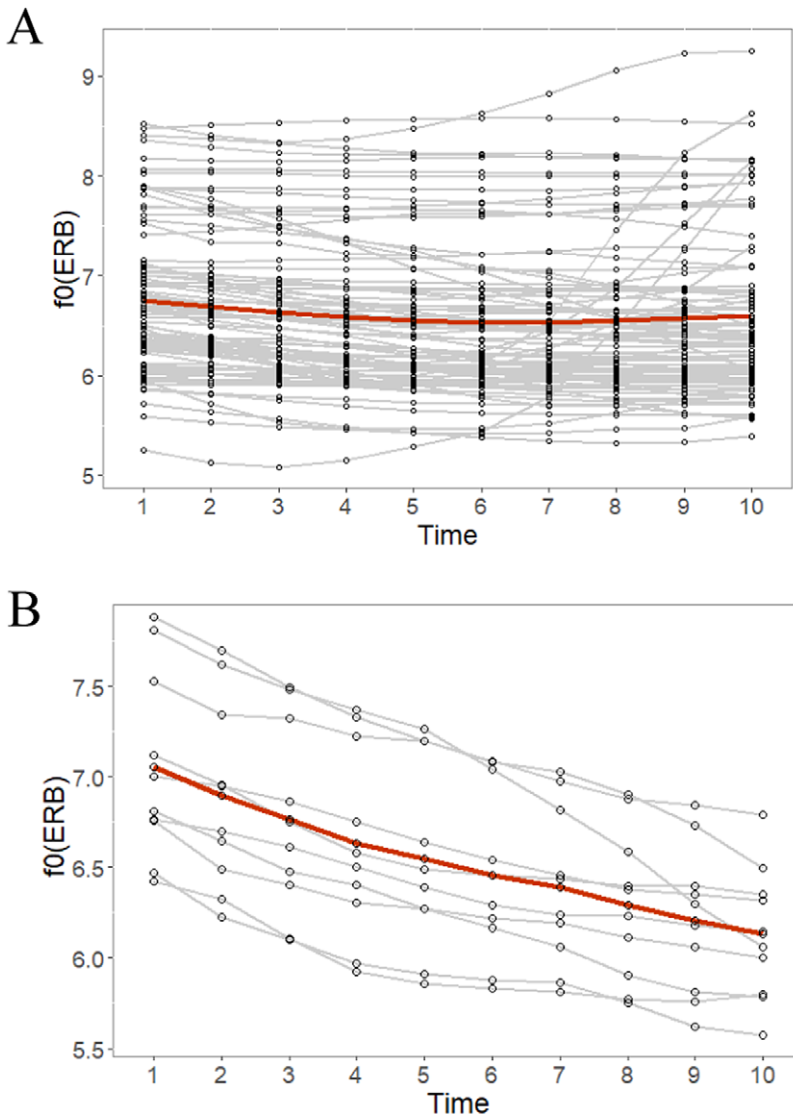


Figure 2. An illustration of the overall central tone contour and the central tone contour of a tone category with data from one participant. (A) The red line demonstrates the participant's overall central tone contour. The grey lines represent all the tone contours produced by this participant, with 10 dots along each contour signifying the 10 equally spaced time points. (B) The red line denotes the central tone contour of Tone4. The grey lines represent all the tone contours of Tone4 produced by this participant, with 10 dots along each contour signifying the 10 equally spaced time points.

Note. f_0 : fundamental frequency; ERB: Equivalent-rectangular-bandwidth-rate.

(R Core Team, 2021), using the *lm4* (Bates et al., 2015), *lmerTest* (Kuznetsova et al., 2017), and *emmeans* (Lenth, 2023) packages. LME models were fitted using the *lmer* function for each metric of lexical tones and intonation. We first included as many fixed and random factors as possible. Specifically, for individual tonal dispersion and within-category tone

variation, the initial model included fixed effects of Condition (IDS-15m vs. IDS-23m vs. ADS) and Tone Category (six lexical tones), their interaction, as well as the random effect of Participant (including the intercept and the slope for the effect of Condition). For the overall tonal space (i.e., the tone triangle area) and the three metrics of intonation, the initial model included the fixed effect of Condition (IDS-15m vs. IDS-23m vs. ADS) and the random intercept for Participant. These initial models did not contain maximal random effects because including all random effects resulted in overparameterization relative to the data available. Then, the *step* function was employed to eliminate the non-significant factors and interactions from the initial models. Once the final models were obtained (see [Appendix](#) for the output of all the final models), the *anova* function with Satterthwaite's method was applied to test their significance. If a significant fixed effect of Condition was observed, pairwise comparisons would be conducted among the three conditions (i.e., IDS-15m, IDS-23m, and ADS) using the *emmeans* function with the Tukey method for multiple comparison corrections. If, meanwhile, a significant interaction between Condition and Tone Category was found, pairwise comparisons would be conducted within each tone category.

Spearman correlation tests were performed to investigate the relations between the age-related changes in lexical tones and intonation. The analysis encompassed all three intonational metrics and the metric of overall tonal space (i.e., the tone triangle area). A single value was derived for each participant with respect to these metrics. Tone variation was excluded from the correlation tests due to its potential inherent association with intonational pitch variability. To quantify the age-related change in each metric, the value at 23 months was divided by the value at 15 months (for example, the change in pitch height of utterances = the pitch height in IDS-23m/the pitch height in IDS-15m).

Results

Lexical-tone modifications and age effects

The results concerning lexical-tone modifications have been shown in [Figure 3](#). The final LME model for the overall tonal space (the tone triangle area) included Condition as the fixed effect and an intercept for Participant as the random effect. For the individual tonal dispersion, the final model included Condition and Tone category as the fixed effects, the interaction between Condition and Tone category, an intercept for Participant, and a by-Participant random slope for the effect of Condition. As for the within-category tone variation, the final model included Condition and Tone Category as the fixed effects, an intercept for Participant, and a by-Participant random slope for the effect of Condition. The interaction between Condition and Tone category was eliminated. The output of the full models can be found in [Table 1-3](#) in [Appendix](#).

The analysis of variance (ANOVA) demonstrated significant main effects of Condition on the overall tonal space ($F(2,36)=11.25, p<.001$). Post-hoc pairwise comparisons among the three conditions, i.e., IDS-15m, IDS-23m, and ADS generated results in [Table 3](#). Compared to ADS, the overall tonal space was significantly larger in both IDS-15m and IDS-23m, while there was no significant difference between IDS-15m and IDS-23m. An illustration of the tone triangles and the overall tonal space in the three conditions can be found in [Figure 3A](#) and [Figure 3B](#) respectively.

For the individual tonal dispersion, the analysis of variance (ANOVA) showed significant main effects of Condition ($F(2,22)=27.46, p<.001$) and Tone category ($F(5,288)=74.31, p<.001$), as well as significant interactions between Condition and Tone category

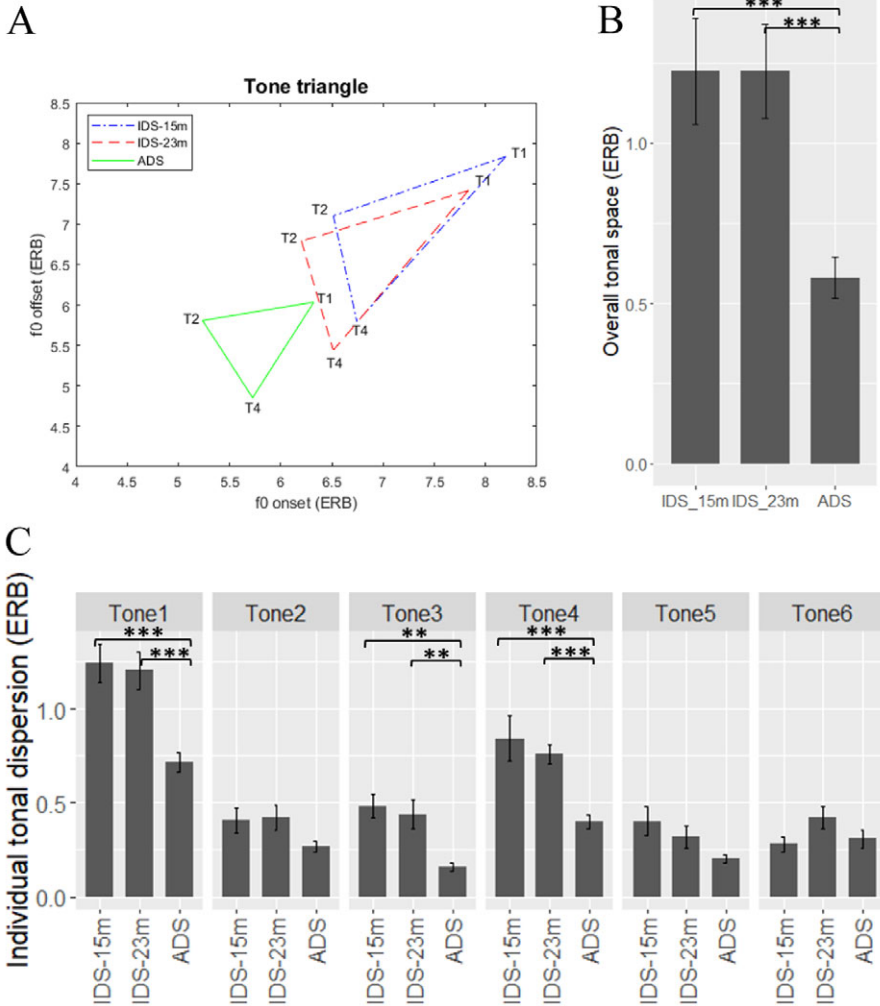


Figure 3. (A) The tone triangles in IDS at 15 months, IDS at 23 months, and ADS (using data averaged across participants). (B) The average overall tonal space (the tone triangle area) in IDS at 15 months, IDS at 23 months, and ADS. (C) The average individual dispersion of each tone category in IDS at 15 months, IDS at 23 months, and ADS. (D) The average intra-talker variation within tone categories in IDS at 15 months, IDS at 23 months, and ADS (averaged across the six categories). The error bars in the charts indicate the standard errors, and the asterisks denote groups that were significantly different from each other.

Note. f0: fundamental frequency; ERB: Equivalent-rectangular-bandwidth-rate. T1: Tone1; T2: Tone2; T4: Tone4.

** $p_{adj} < .01$, *** $p_{adj} < .001$.

($F(10,288)=3.45$, $p<.001$). Based on the results, post-hoc pairwise comparisons among the three conditions, i.e., IDS-15m, IDS-23m, and ADS were conducted within each tone category. The results have been reported in Table 3. For Tone1, Tone3, and Tone4, their dispersion from the center of tonal space significantly increased in IDS at both age points compared to ADS, while there was no significant difference between the two ages,

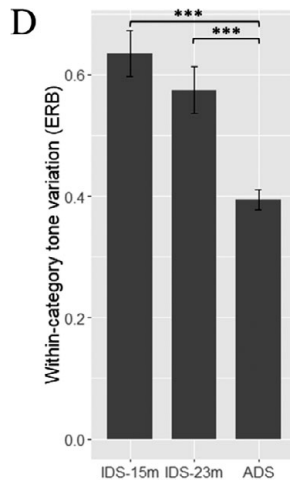


Figure 3. (Continued)

consistent with the findings on the overall tonal space. The dispersion of Tone2, Tone5, and Tone6, on the other hand, did not significantly differ between any two of the conditions. The dispersion of each tone category in the three conditions has been demonstrated in Figure 3C.

As with the within-category tone variation, the analysis of variance (ANOVA) reported significant main effects of Condition ($F(2,18)=22.03, p<.001$) and Tone category ($F(5,280)=20.55, p<.001$). Post-hoc pairwise comparisons among the three conditions, i.e., IDS-15m, IDS-23m, and ADS generated the results presented in Table 3. The variation of tone tokens was greater in IDS at both age points compared to ADS, while there was no significant difference between the two ages, similar to the findings on the overall tonal space and the dispersion of Tone1, Tone3, and Tone4. The tone variation averaged across the six categories in the three conditions has been presented in Figure 3D.

To sum up, the mothers produced lexical tones with an overall expanded tonal space, more dispersed tone categories (of Tone1, Tone3, and Tone4 specifically) from the center of the tonal space, and greater within-category tone variation in IDS when the children were both 15 and 23 months of age. None of the metrics showed a statistically significant change in IDS between the two age points.

Intonational modifications and age effects

The pitch height of utterances, pitch variability across utterances, and pitch variability within utterances in IDS to children of the two ages and ADS have been demonstrated in Figure 4. The final LME models for all three metrics included Condition as the fixed effect and an intercept for Participant as the random effect. The output of the full models can be found in Table 4-6 in Appendix. The analysis of variance (ANOVA) revealed significant main effects of Condition on the pitch height of utterances ($F(2,36)=61.15, p<.001$), pitch variability across utterances ($F(2,36)=38.83, p<.001$), and pitch variability within utterances ($F(2,36)=4.08, p=.025$). Hence, post-hoc pairwise comparisons among the three

Table 3. The results of pairwise comparisons among IDS-15m, IDS-23m, and ADS for the lexical-tone metrics.

Overall tonal space					
	Estimate	SE	Df	t.ratio	p _{adj}
ADS vs. IDS-15m	-.64	.16	36	-4.11	<.001
ADS vs. IDS-23m	-.64	.16	36	-4.11	<.001
IDS-15m vs. IDS-23m	-.00035	.16	36	-.002	1
Individual tonal dispersion					
	Estimate	SE	Df	z.ratio	p _{adj}
Tone1					
ADS vs. IDS-15m	-.53	.089	168	-5.91	<.001
ADS vs. IDS-23m	-.49	.089	168	-5.49	<.001
IDS-15m vs. IDS-23m	.037	.097	106	.39	.92
Tone2					
ADS vs. IDS-15m	-.14	.089	168	-1.56	.27
ADS vs. IDS-23m	-.15	.089	168	-1.71	.21
IDS-15m vs. IDS-23m	-.013	.097	106	-.13	.99
Tone3					
ADS vs. IDS-15m	-.32	.089	168	-3.62	.0011
ADS vs. IDS-23m	-.28	.089	168	-3.13	.0059
IDS-15m vs. IDS-23m	.045	.097	106	.46	.89
Tone4					
ADS vs. IDS-15m	-.44	.089	168	-4.97	<.001
ADS vs. IDS-23m	-.36	.089	168	-4.05	<.001
IDS-15m vs. IDS-23m	.082	.097	106	.85	.67
Tone5					
ADS vs. IDS-15m	-.2	.089	168	-2.25	.066
ADS vs. IDS-23m	-.12	.089	168	-1.33	.38
IDS-15m vs. IDS-23m	.082	.097	106	.85	.68
Tone6					
ADS vs. IDS-15m	.029	.089	168	.33	.94
ADS vs. IDS-23m	-.11	.089	168	-1.27	.42
IDS-15m vs. IDS-23m	-.14	.097	106	-1.47	.31
Within-category tone variation					
	Estimate	SE	Df	t.ratio	p _{adj}
ADS vs. IDS-15m	-.24	.039	18	-6.14	<.001

Table 3. (Continued)

Within-category tone variation					
	Estimate	SE	Df	t.ratio	p _{adj}
ADS vs. IDS–23m	–.18	.039	18	–4.66	<.001
IDS–15m vs. IDS–23m	.06	.044	18	1.37	.38

Notes. For individual tonal dispersion, pairwise comparisons were conducted within each lexical tone.

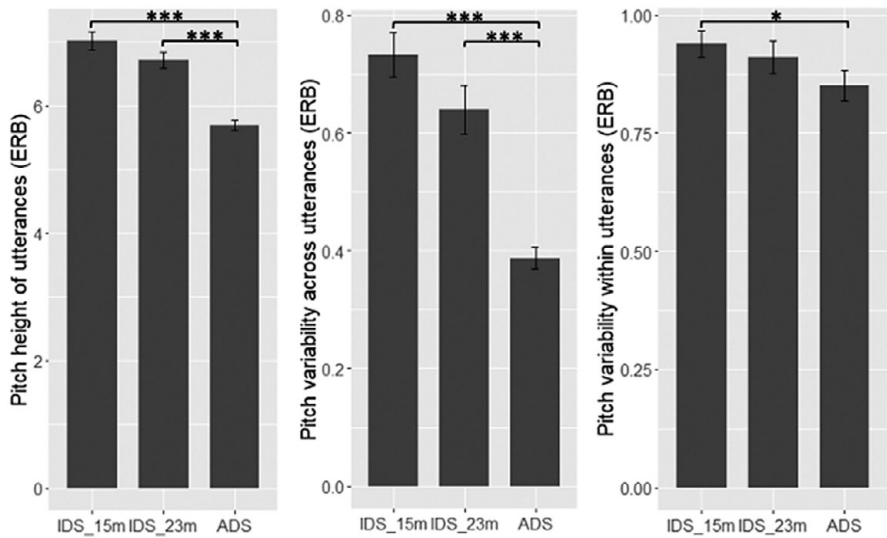


Figure 4. The average pitch height of utterances, pitch variability across utterances, and pitch variability within utterances in IDS at 15 months, IDS at 23 months, and ADS. The error bars in the chart indicate the standard errors. The asterisks denote groups that were significantly different from each other. Note. ERB: Equivalent-rectangular-bandwidth-rate. * $p_{adj} < .05$, *** $p_{adj} < .001$.

conditions, i.e., IDS-15m, IDS-23m, and ADS were conducted for the three measurements. The results have been reported in Table 4 and summarized below:

- 1) Pitch height of utterances: in comparison to ADS, it was significantly larger in both IDS-15m and IDS-23m, while there was no significant difference between IDS-15m and IDS-23m;
- 2) Pitch variability across utterances: consistent with the findings on pitch height, it was significantly greater in both IDS-15m and IDS-23m than in ADS, whereas there was no significant difference between IDS-15m and IDS-23m;
- 3) Pitch variability within utterances: it was only significantly greater in IDS-15m compared to ADS, while no significant difference was found between IDS-23m and ADS, and between IDS-15m and IDS-23m.

In brief, pitch height and pitch variability across utterances increased in the mothers' IDS at both age points compared to their ADS. In contrast, pitch variability within utterances was found to increase in IDS at 15 months only. Regarding the age effects, the changes of

Table 4. The results of pairwise comparisons among IDS-15m, IDS-23m, and ADS for the intonation metrics.

Pitch height of utterances					
	Estimate	SE	Df	t.ratio	p _{adj}
ADS vs. IDS-15m	-1.33	.13	36	-10.55	<.001
ADS vs. IDS-23m	-1.03	.13	36	-8.15	<.001
IDS-15m vs. IDS-23m	.303	.13	36	2.4	.055
Pitch variability across utterances					
	Estimate	SE	Df	t.ratio	p _{adj}
ADS vs. IDS-15m	-.34	.041	36	-8.51	<.001
ADS vs. IDS-23m	-.25	.041	36	-6.23	<.001
IDS-15m vs. IDS-23m	.093	.041	36	2.29	.071
Pitch variability within utterances					
	Estimate	SE	Df	t.ratio	p _{adj}
ADS vs. IDS-15m	-.089	.032	36	-2.79	.02
ADS vs. IDS-23m	-.062	.032	36	-1.94	.14
IDS-15m vs. IDS-23m	.027	.032	36	.84	.68

all the three measurements between the two age points did not reach statistical significance, consistent with the findings on lexical tones.

Lexical tone – intonation correlations

The Spearman correlation tests found significant positive correlations between the change in overall tonal space from 15 to 23 months of age (IDS-23m/IDS-15m) and the changes in all three intonational metrics over this period. Specifically, significant correlations were observed with the pitch height of utterances ($r_s=.49$, $p=.034$), pitch variability across utterances ($r_s=.57$, $p=.012$), and pitch variability within utterances ($r_s=.48$, $p=.038$), indicating medium to large relationships (Cohen, 1992). Scatter plots in Figure 5 visualize these correlations, demonstrating a similar changing pattern in the overall tonal space and intonational measurements. As the children developed from 15 to 23 months old, the overall tonal space and intonational modifications decreased in some mothers' IDS while increased in others' IDS. There appeared to be a general tendency for those mothers whose intonational modifications in IDS decreased to show a corresponding decrease in the overall tonal space, while for those whose intonational modifications increased to also have an increase in the overall tonal space.

Discussion

To deepen understanding of lexical and prosodic pitch modifications in IDS of tone languages, the current study 1) examined the lexical-tone and intonational modifications

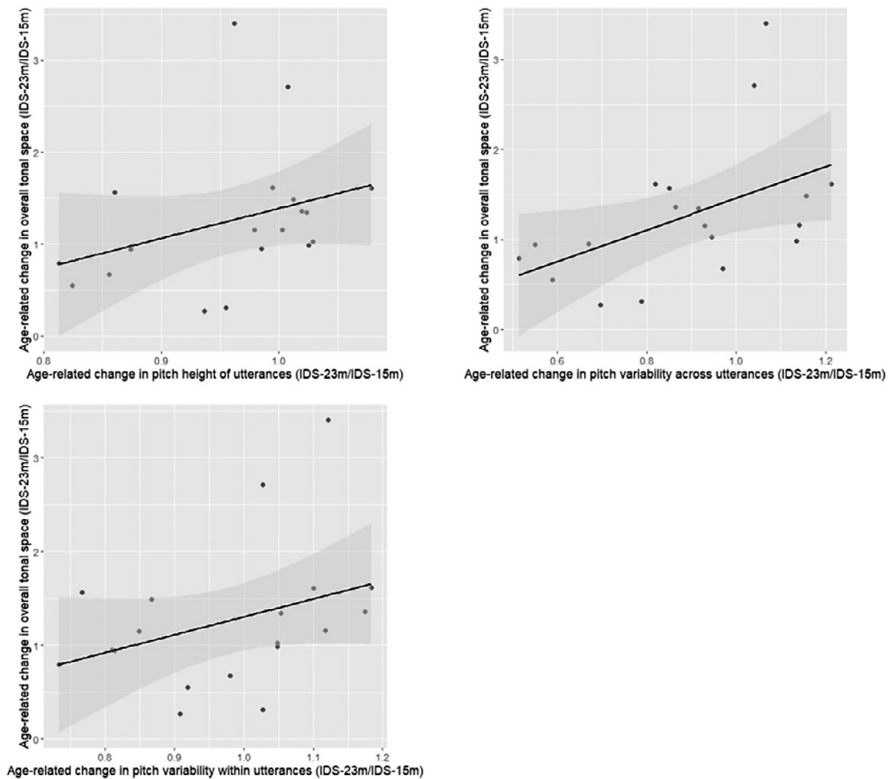


Figure 5. The scatter plots for the correlations between the change in the overall tonal space from 15 to 23 months of age (IDS-23m/IDS-15m) and the changes in the three intonational metrics over this period. Each dot denotes the data of one participant. The linear regression line is plotted with a 95% confidence interval.

in Cantonese IDS to children at 15 and 23 months of age, in comparison to ADS; 2) explored the changes in lexical tones and intonation in IDS from the early to late age points; 3) tested the relationships between the age-related changes in lexical tones and intonation. The findings are summarized and discussed in this section.

Lexical-tone and intonational modifications in IDS-15m

When addressing their 15-month-old children, Cantonese-speaking mothers raised the intonational pitch height, pitch variability across utterances, and pitch variability within utterances in IDS. They also had overall expanded tonal space, more dispersed tone categories (of Tone1, Tone3, and Tone4), and increased tone variation within categories in IDS. These results largely replicate the previous findings on Cantonese IDS to children of the same age (Wang et al., 2021), with one exception of the pitch variability within utterances. In the earlier study, the increase in pitch variability within utterances in IDS did not reach statistical significance. The discrepancy between the two studies is not surprising, because an increase in pitch variability within utterances appears not to be a robust prosodic feature in IDS. There is a substantial inconsistency in findings regarding

this feature in previous studies. A cross-language investigation (Fernald et al., 1989) observed a significant increase in within-utterance pitch range in IDS of American English but not in IDS of other languages such as German, French, and so on. Different studies on IDS of the same language also reported divergent findings on modifications in pitch range within utterances (cf. Fernald et al., 1989 vs. Fernald & Simon, 1984 on German; Grieser & Kuhl, 1988 vs. Papoušek & Hwang, 1991 on Mandarin Chinese). In addition, although the earlier study reported an increase in overall tonal dispersion in IDS compared to ADS, it did not examine potential differences in dispersion among tone categories. In the present study, by calculating the dispersion data for each tone separately, we found evidence for increased dispersion from the center of the tonal space for Tone1, Tone3, and Tone4 in mothers' IDS to their 15-month-old children, but not for Tone2, Tone5, and Tone6. The findings suggest that the previously observed enhancement in overall tonal dispersion might be primarily attributed to the increased dispersion of Tone1, Tone3, and Tone4.

Lexical-tone and intonational modifications in IDS-23m

The present study further provided new findings on Cantonese IDS to 23-month-olds, a developmental stage that has received little attention for this language. Cantonese-speaking mothers maintained their use of exaggerated intonation in IDS to their children at the end of the second year of life, as demonstrated by larger pitch height and pitch variability across utterances compared to ADS. They did not significantly increase pitch variability within utterances in IDS at this age though. Meanwhile, their IDS still contained lexical tones with larger overall space, more dispersed tone categories (of Tone1, Tone3, and Tone4), and greater variation within categories than ADS.

Age effects

Statistical comparisons of IDS to children at 15 and 23 months old indicated no significant age-related changes in all the metrics of both intonation and lexical tones in maternal IDS. As stated in the introduction, there is currently limited knowledge about the age effects on lexical tones and intonation in IDS of tone languages. By utilizing longitudinal data on Cantonese IDS, our findings could help draw a more comprehensive picture of the developmental pathways for lexical-tone and intonational modifications in IDS of tone languages.

The tonal space had been the focus of most previous studies. The one study that collected longitudinal data on Cantonese IDS (Xu Rattanasone et al., 2013) reported a significant decrease in overall tonal space in IDS to infants from 3 months to the end of the first year of life. As introduced earlier, this study assessed overall tonal space by calculating the tone triangle area. To extend this study, we initially examined the age-related changes in lexical tones using the same measurement. The results suggest that Cantonese-speaking mothers seem to maintain an overall expanded tonal space in IDS as their children enter the second year of life, with no significant changes detected over that year. Furthermore, we explored age-related changes in tonal space with a new measurement, individual tonal dispersion, which quantifies the magnitude of tonal space with data from each tone. Similarly, we observed no significant changes in the degree of dispersion across all tone categories in IDS to children at the two ages. The absence of findings for significant age-related changes in tonal space aligns with vowel research that has demonstrated the stability of vowel space in IDS throughout the first

and second years of children's lives (Cox et al., 2023; Hartman et al., 2017; Kalashnikova & Burnham, 2018).

More importantly, our longitudinal investigation was also conducted on the variation within each tone category - a dimension that has been substantially overlooked. In Cantonese-speaking mothers' IDS, the within-category tone variation appears to follow the same developmental trajectory as the overall tonal space and the individual tonal dispersion, i.e., showing no signs of significant changes through the second year of children's lives. Future research on other tone languages with children of different ages is needed to enrich understanding of this dimension of lexical tones, particularly with regard to its changes in relation to children's age.

Regarding intonation, the absence of evidence for significant age-related changes is in line with those observed in IDS of Tamil, Tagalog, and Korean with infants aged from 4 to 15 months (Narayan & McDermott, 2016), and in IDS of English with children aged from 7 to 19 months (Kalashnikova & Burnham, 2018). When compared to the findings of the study on a different tone language, Mandarin, with children of comparable ages (Han et al., 2020), some discrepancies were noted. Han et al. (2020) analyzed intonation with respect to pitch height and pitch range within utterances, and found that they were both increased only in IDS to 18-month-olds but not in IDS to 24-month-olds, compared to ADS. By contrast, while the current study also observed no increase in pitch variability within utterances in IDS to 23-month-olds, it revealed a rise in pitch height in IDS to both 15- and 23-month-olds. The inconsistent findings might result from the methodological differences between the two studies. The previous study collected IDS data through a semi-spontaneous storybook-telling task, different from the toy-playing task employed by the present study. Alternatively, there could indeed be a difference in the use of exaggerated intonation in IDS between mothers speaking Cantonese and Mandarin. Further research may be conducted to explore the possibilities.

The current study only focused on a limited age range. To obtain a more comprehensive understanding of age effects on lexical tones and intonation in IDS of tone languages, future studies are expected to investigate children at various ages. By collecting data from younger infants within the first year of life using similar methods, direct comparisons can be made with the previous longitudinal study on lexical tones in Cantonese IDS (Xu Rattanasone et al., 2013). More importantly, investigations on IDS to older children are necessary to determine the milestone at which Cantonese-speaking caregivers shift to ADS-like lexical tones and intonation when speaking to their children.

Within-category tone variation

Our findings on tone variation summarized above align with the prevailing evidence indicating an increase in vowel variation in IDS to young learners during their first two years of life (Cristia & Seidl, 2014; McMurray et al., 2013; Miyazawa et al., 2017; Rosslund et al., 2022). The increase in variation can lead to more overlap among phonemic categories and thus less differentiated phonemic contrasts, counteracting the expansion of the overall acoustic space. It may pose challenges for young children in discerning phonemes and constructing mental representations for phonemic categories, thereby impeding child language learning. Correlation studies on vowels have reported negative correlations between within-category vowel variability in parental speech and child lexical development (Hartman et al., 2017; Rosslund et al., 2022). Regarding lexical tones, there are only some preliminary findings with a limited sample size, showing negative

correlations between within-category tone variation in caregivers' IDS and their children's concurrent and later vocabulary size (Wang, 2020). The current findings on tone variation, like the past studies on vowel variation, cast some doubt on the hyperarticulation hypothesis which posits that phonemic contrasts are hyperarticulated in IDS to serve a didactic function in facilitating child language learning (Kuhl et al., 1997; Werker et al., 2007; Xu Rattanasone et al., 2013). If mothers, whether consciously or unconsciously, modify their production of lexical tones for a didactic purpose, it may be less likely for them to produce more varied tone tokens for individual lexical tones.

To further investigate the hyperarticulation hypothesis from the perspective of within-category phonemic variation, it is necessary to combine it with cross-category mean differences to quantify the phonemic contrasts in IDS. In addition, the impact of greater variation within categories in IDS on child language learning needs to be further examined. More correlation studies may be conducted with a larger sample size to confirm the negative correlations observed previously. Laboratory experiments can also be designed to test the immediate effect of increased tone variation in speech input on child language processing.

Correlations between age-related changes in lexical tones and intonation

As children developed from 15 to 23 months old, significant positive correlations were uncovered between the age-related changes in overall tonal space and intonational modifications (including pitch height of utterances, and pitch variability within and across utterances) in their mothers' IDS, although comparisons between IDS at the two age points found no significant differences. With increasing age of the children, mothers seem to adopt varying approaches in altering the way they spoke to them, with some increasing while others decreased their exaggerations of lexical tones and intonation. In general, it appears that the changes in tonal space and intonation in mothers' IDS follow a similar pattern. Mothers who exhibit an age-related decrease in intonational modifications in their IDS are more likely to show a corresponding decrease in tonal space; and those who demonstrate an increase in intonational modifications tend to also increase tonal space.

The current longitudinal findings have extended our understanding beyond the prior research on the association between lexical and prosodic pitch realizations in IDS at a single age point (Wang et al., 2021). Not only do the degree of lexical-tone and intonational modifications in individual parents' IDS, as compared to their ADS, positively correlate with each other, but the age-related changes in lexical tones (specifically tonal space) and intonation in IDS with children's development also exhibit positive correlations. The correlation findings are consistent with the prosodic hypothesis which argues that modifications of phonemes (in this case lexical tones) in IDS may be unintended byproducts of the prosodic modulations.

Previously, when a significant decrease in lexical-tone modifications, in particular tonal space, was observed in IDS with increasing age of children, it was argued to be a manifestation of parents adjusting their speech to accommodate children's developing language abilities and learning needs (Liu et al., 2009; Xu Rattanasone et al., 2013). The findings were thus considered as evidence for the hyperarticulation hypothesis that phonemic exaggerations in IDS are made for didactic purposes. Our findings of positive correlations between age-related changes in tonal space and intonational exaggerations have introduced another potential explanation. The previously observed decline in tonal

space in IDS as children grow older might be driven by a concurrent decrease in intonational exaggerations during that period. To test this possibility, it is necessary to simultaneously examine the age-related changes in lexical tones and intonation in IDS at the children's age when a significant decrease in lexical-tone modifications is observed.

In conclusion, the current study first demonstrates that compared to ADS, Cantonese-speaking mothers exaggerate intonation (with higher overall pitch and larger pitch variability across utterances) and modify lexical tones (with generally larger tonal space and greater tone variation within categories) in IDS when interacting with their children at both 15 and 23 months of age, with no significant changes observed between the two ages. The increase in tone variation in IDS, coinciding with the expansion of tonal space, seems to cast some doubt on the hyperarticulation hypothesis. In addition, the study shows positive correlations between age-related changes in tonal space and intonational exaggerations in IDS as children grow older, consistent with the proposition of the prosodic hypothesis. The findings on Cantonese enhance our understanding of lexical and prosodic pitch modifications in IDS of tone languages, particularly concerning three aspects: the within-category tone variation, the age-related changes in lexical tones and intonation over the course of children's development, and the associations between the lexical-tone and intonational modifications. The current findings were based on data from a single language. Cross-linguistic studies are necessary to validate and extend these findings, or to explore potential differences between languages (Kalashnikova et al., 2023). Future research should also consider how these IDS features may facilitate tone acquisition and language development more broadly, especially in relation to individual differences in language functions (Antoniou & Wong, 2015; Ingvalson et al., 2013; P. C. M. Wong et al., 2012, 2013, 2020, 2021).

Supplementary material. The supplementary material for this article can be found at <http://doi.org/10.1017/S0305000924000333>.

Acknowledgements. The current research data were collected during the first author's Ph.D. study at The Chinese University of Hong Kong.

Funding. This work was supported by the University Grants Committee (HKSAR) [grant number RGC4024-21G and RGC4055-19G].

Competing interest. The authors declare no competing interests.

References

- Antoniou, M., & Wong, P. C. M. (2015). Poor phonetic perceivers are affected by cognitive load when resolving talker variability. *The Journal of the Acoustical Society of America*, 138(2), 571–574. <https://doi.org/10.1121/1.4923362>
- Bates, D., Mächler, M., Bolker, B. M., & Walker, S. C. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, 67(1), 1–48. <https://doi.org/10.18637/jss.v067.i01>
- Benders, T. (2013). Mommy is only happy! Dutch mothers' realisation of speech sounds in infant-directed speech expresses emotion, not didactic intent. *Infant Behavior and Development*, 36(4), 847–862. <https://doi.org/10.1016/j.infbeh.2013.09.001>
- Boersma, P., & Weenink, D. J. M. (2017). *Praat: Doing Phonetics by Computer [Computer program]*.
- Burnham, D., Kitamura, C., & Vollmer-Conna, U. (2002). What's new, pussycat? On talking to babies and animals. *Science*, 296(5572), 1435. <https://doi.org/10.1126/science.1069587>
- Cheng, M. C., & Chang, K. C. (2014). Tones in hakka infant-directed speech: An acoustic perspective. *Language and Linguistics*, 15(3), 341–390. <https://doi.org/10.1177/1606822X14520662>
- Cohen, J. (1992). A power primer. *Psychological Bulletin*, 112(1), 155–159.

- Cooper, R. P., Abraham, J., Berman, S., & Staska, M. (1997). The development of infants' preference for motherese. *Infant Behavior and Development*, 20(4), 477–488. [https://doi.org/10.1016/S0163-6383\(97\)90037-0](https://doi.org/10.1016/S0163-6383(97)90037-0)
- Cox, C., Bergmann, C., Fowler, E., Keren-Portnoy, T., Roepstorff, A., Bryant, G., & Fusaroli, R. (2023). A systematic review and Bayesian meta-analysis of the acoustic features of infant-directed speech. *Nature Human Behaviour*, 7, 114–133. <https://doi.org/10.1038/s41562-022-01452-1>
- Cristia, A., & Seidl, A. (2014). The hyperarticulation hypothesis of infant-directed speech. *Journal of Child Language*, 41(4), 913–934. <https://doi.org/10.1017/S0305000912000669>
- Englund, K. T., & Behne, D. M. (2005). Infant directed speech in natural interaction-norwegian vowel quantity and quality. *Journal of Psycholinguistic Research*, 34(3), 259–280. <https://doi.org/10.1007/s10936-005-3640-7>
- Fernald, A. (1989). Intonation and communicative intent in mothers' speech to infants: Is the melody the message? *Child Development*, 60(6), 1497–1510.
- Fernald, A., & Simon, T. (1984). Expanded intonation contours in mothers' speech to newborns. *Developmental Psychology*, 20(1), 104–113. <https://doi.org/10.1037/0012-1649.20.1.104>
- Fernald, A., Taeschner, T., Dunn, J., Papousek, M., De Boysson-Bardies, B., & Fukui, I. (1989). A cross-language study of prosodic modifications in mothers' and fathers' speech to preverbal infants. *Journal of Child Language*, 16(3), 477–501. <https://doi.org/10.1017/S0305000900010679>
- Goldfield, B. A., & Reznick, J. S. (1990). Early lexical acquisition: Rate, content, and the vocabulary spurt. *Journal of Child Language*, 17(1), 171–183. <https://doi.org/10.1017/S0305000900013167>
- Grieser, D. A. L., & Kuhl, P. K. (1988). Maternal speech to infants in a tonal language: Support for universal prosodic features in motherese. *Developmental Psychology*, 24(1), 14–20. <https://doi.org/10.1037/0012-1649.24.1.14>
- Han, M., de Jong, N. H., & Kager, R. (2018). Lexical tones in Mandarin Chinese infant-directed speech: Age-related changes in the second year of life. *Frontiers in Psychology*, 9, 434. <https://doi.org/10.3389/fpsyg.2018.00434>
- Han, M., De Jong, N. H., & Kager, R. (2020). Pitch properties of infant-directed speech specific to word-learning contexts: A cross-linguistic investigation of Mandarin Chinese and Dutch. *Journal of Child Language*, 47(1), 85–111. <https://doi.org/10.1017/S0305000919000813>
- Hartman, K. M., Ratner, N. B., & Newman, R. S. (2017). Infant-directed speech (IDS) vowel clarity and child language outcomes. *Journal of Child Language*, 44(5), 1140–1162. <https://doi.org/10.1017/S0305000916000520>
- Hermes, D. J., & van Gestel, J. C. (1991). The frequency scale of speech intonation. *Journal of the Acoustical Society of America*, 90(1), 97–102. <https://doi.org/10.1121/1.402397>
- Ingvallson, E. M., Barr, A. M., & Wong, P. C. M. (2013). Poorer phonetic perceivers show greater benefit in phonetic-phonological speech learning. *Journal of Speech, Language, and Hearing Research*, 56(3), 1045–1050. [https://doi.org/10.1044/1092-4388\(2012/12-0024](https://doi.org/10.1044/1092-4388(2012/12-0024)
- Jacobson, J. L., Boersma, D. C., Fields, R. B., & Olson, K. L. (1983). Paralinguistic features of adult speech to infants and small children. *Child Development*, 54(2), 436–442. <https://doi.org/10.2307/1129704>
- Kalashnikova, M., & Burnham, D. (2018). Infant-directed speech from seven to nineteen months has similar acoustic properties but different functions. *Journal of Child Language*, 45(5), 1035–1053. <https://doi.org/10.1017/S0305000917000629>
- Kalashnikova, M., Singh, L., Burnham, R., Cannistraci, R., Chen, H., Ng, B. C., Dos Santos, M., Dwyer, A., Feng, Y., Gisvold, A. K., Gustavsson, L., Hui, O. S., Hay, J., Kager, R., de Klerk, M., Lai, R., Liu, L., Marklund, E., Nazzi, T., ... Woo, P.-J. (2023). The development of tone categories in infancy: Evidence from a cross-linguistic, multi-lab report. *Developmental Science*, 27(3), e13459. <https://doi.org/10.1111/desc.13459>
- Kitamura, C., Thanavishuth, C., Burnham, D., & Luksaneeyanawin, S. (2002). Universality and specificity in infant-directed speech: Pitch modifications as a function of infant age and sex in a tonal and non-tonal language. *Infant Behavior and Development*, 24(4), 372–392. [https://doi.org/10.1016/S0163-6383\(02\)00086-3](https://doi.org/10.1016/S0163-6383(02)00086-3)
- Kuhl, P. K., Andruski, J. E., Chistovich, I. A., Chistovich, L. A., Kozhevnikova, E. V., Ryskina, V. L., Stolyarova, E. I., Sundberg, U., & Lacerda, F. (1997). Cross-language analysis of phonetic units in language addressed to infants. *Science*, 277(5326), 684–686. <https://doi.org/10.1126/science.277.5326.684>

- Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2017). lmerTest package: tests in linear mixed effects models. *Journal of Statistical Software*, 82(13), 1–26. <https://doi.org/10.18637/jss.v082.i13>
- Lenth, R. (2023). *emmeans: Estimated Marginal Means, aka Least-Squares Means*. <https://cran.r-project.org/package=emmeans>
- Li, Y. F. (2022). *Manifestation of Cantonese Lexical Tones in Speech Registers*. [Doctoral dissertation, The Chinese University of Hong Kong]. ProQuest Dissertations and Theses database.
- Liu, H. M., Tsao, F. M., & Kuhl, P. K. (2007). Acoustic analysis of lexical tone in mandarin infant-directed speech. *Developmental Psychology*, 43(4), 912–917. <https://doi.org/10.1037/0012-1649.43.4.912>
- Liu, H. M., Tsao, F. M., & Kuhl, P. K. (2009). Age-related changes in acoustic modifications of Mandarin maternal speech to preverbal infants and five-year-old children: A longitudinal study. *Journal of Child Language*, 36(4), 909–922. <https://doi.org/10.1017/S030500090800929X>
- Ma, J. K.-Y., Ciocca, V., & Whitehill, T. L. (2006). Effect of intonation on Cantonese lexical tones. *The Journal of the Acoustical Society of America*, 120(6), 3978–3987. <https://doi.org/10.1121/1.2363927>
- Ma, W., Zhou, P., Singh, L., & Gao, L. (2017). Spoken word recognition in young tone language learners: Age-dependent effects of segmental and suprasegmental variation. *Cognition*, 159, 139–155. <https://doi.org/10.1016/j.cognition.2016.11.011>
- Marklund, E., & Gustavsson, L. (2020). The dynamics of vowel hypo- and hyperarticulation in Swedish infant-directed speech to 12-month-olds. *Frontiers in Communication*, 5(October), 1–15. <https://doi.org/10.3389/fcomm.2020.523768>
- McMurray, B., Kovack-Lesh, K. A., Goodwin, D., & McEchron, W. (2013). Infant directed speech and the development of speech perception: Enhancing development or an unintended consequence? *Cognition*, 129(2), 362–378. <https://doi.org/10.1016/j.cognition.2013.07.015>
- Miyazawa, K., Shinya, T., Martin, A., Kikuchi, H., & Mazuka, R. (2017). Vowels in infant-directed speech: More breathy and more variable, but not clearer. *Cognition*, 166, 84–93. <https://doi.org/10.1016/j.cognition.2017.05.003>
- Narayan, C. R., & McDermott, L. C. (2016). Speech rate and pitch characteristics of infant-directed speech: Longitudinal and cross-linguistic observations. *The Journal of the Acoustical Society of America*, 139(3), 1272–1281. <https://doi.org/10.1121/1.4944634>
- Papoušek, M., & Hwang, S. F. C. (1991). Tone and intonation in Mandarin babytalk to presyllabic infants: Comparison with registers of adult conversation and foreign language instruction. *Applied Psycholinguistics*, 12(4), 481–504. <https://doi.org/10.1017/S0142716400005889>
- R Core Team. (2021). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing. <https://www.r-project.org/>
- Rosslund, A., Mayor, J., Óturai, G., & Kartushina, N. (2022). Parents' hyper-pitch and low vowel category variability in infant-directed speech are associated with 18-month-old toddlers' expressive vocabulary. In *Language Development Research* (Vol. 2, Issue 1). <https://doi.org/10.34842/2022.0547>
- Soderstrom, M. (2007). Beyond babytalk: Re-evaluating the nature and content of speech input to preverbal infants. *Developmental Review*, 27(4), 501–532. <https://doi.org/10.1016/j.dr.2007.06.002>
- Stern, D. N., Spieker, S., Barnett, R., & Mackain, K. (1983). The prosody of maternal speech: Infant age and context related changes. *Journal of Child Language*, 10(1), 1–15. <https://doi.org/10.1017/S0305000900005092>
- Tang, P., Xu Rattanasone, N., Yuen, I., & Demuth, K. (2017). Phonetic enhancement of Mandarin vowels and tones: Infant-directed speech and Lombard speech. *The Journal of the Acoustical Society of America*, 142(2), 493–503. <https://doi.org/10.1121/1.4995998>
- Wang, L. (2020). *The Pitch Modifications of Lexical Tones in Cantonese Infant-directed Speech: The Influence from Intonational Factors and Their Associations with Infant Language Development* [Doctoral dissertation, The Chinese University of Hong Kong]. ProQuest Dissertations and Theses database.
- Wang, L., Kager, R., & Wong, P. C. M. (2022). The effect of tone hyperarticulation in Cantonese infant-directed speech on toddlers' word recognition in the second year of life. *First Language*, 42(5), 670–692. <https://doi.org/10.1177/01427237221109342>
- Wang, L., Kalashnikova, M., Kager, R., Lai, R., & Wong, P. C. M. (2021). Lexical and prosodic pitch modifications in Cantonese infant-directed speech. *Journal of Child Language*, 48(6), 1235–1261. <https://doi.org/10.1017/S0305000920000707>

- Werker, J. F., Pons, F., Dietrich, C., Kajikawa, S., Fais, L., & Amano, S. (2007). Infant-directed speech supports phonetic category learning in English and Japanese. *Cognition*, **103**(1), 147–162. <https://doi.org/10.1016/j.cognition.2006.03.006>
- Wong, P., & Ng, K. W. S. (2018). Testing the hyperarticulation and prosodic hypotheses of child-directed speech: Insights from the perceptual and acoustic characteristics of child-directed cantonese tones. *Journal of Speech, Language, and Hearing Research*, **61**(8), 1907–1925.
- Wong, P. C. M., Chandrasekaran, B., & Zheng, J. (2012). The derived allele of ASPM is associated with lexical tone perception. *PLoS ONE*, **7**(4). <https://doi.org/10.1371/journal.pone.0034243>
- Wong, P. C. M., Ettlinger, M., & Zheng, J. (2013). Linguistic grammar learning and DRD2-TAQ-IA polymorphism. *PLoS ONE*, **8**(5), 1–9. <https://doi.org/10.1371/journal.pone.0064983>
- Wong, P. C. M., Kang, X., Wong, K. H. Y., So, H.-C., Choy, K. W., & Geng, X. (2020). ASPM-lexical tone association in speakers of a tone language: Direct evidence for the genetic-biasing hypothesis of language evolution. *Science Advances*, **6**(22). <https://doi.org/10.1126/sciadv.aba5090>
- Wong, P. C. M., Lai, C. M., Chan, P. H. Y., Leung, T. F., Lam, H. S., Feng, G., Maggu, A. R., & Novitskiy, N. (2021). Neural speech encoding in infancy predicts future language and communication difficulties. *American Journal of Speech-Language Pathology*, **30**(5), 2241–2250. https://doi.org/10.1044/2021_AJSLP-21-00077
- Xu, N., & Burnham, D. (2010). Tone hyperarticulation and intonation in Cantonese infant directed speech. *Speech Prosody 2010-Fifth International Conference*.
- Xu Rattanasone, N., Burnham, D., & Reilly, R. G. (2013). Tone and vowel enhancement in Cantonese infant-directed speech at 3, 6, 9, and 12 months of age. *Journal of Phonetics*, **41**(5), 332–343. <https://doi.org/10.1016/j.wocn.2013.06.001>
- Xu, Y. (2013). ProsodyPro - A tool for large-scale systematic prosody analysis. *Proceedings of Tools and Resources for the Analysis of Speech Prosody*, 7–10.
- Zhao, Y., & Jurafsky, D. (2009). The effect of lexical frequency and Lombard reflex on tone hyperarticulation. *Journal of Phonetics*, **37**(2), 231–247. <https://doi.org/10.1016/j.wocn.2009.03.002>