Journal of Child Language (2021), **48**, 1235–1261
doi:10.1017/S0305000920000707

CAMBRIDGE UNIVERSITY PRESS

ARTICLE

Lexical and Prosodic Pitch Modifications in Cantonese Infant-directed Speech

Luchang WANG[1,2], Marina KALASHNIKOVA[3,4], René KAGER[5], Regine LAI[1], and Patrick C.M. WONG[1,2]

[1]Department of Linguistics and Modern Languages, The Chinese University of Hong Kong, Hong Kong, China, [2]Brain and Mind Institute, The Chinese University of Hong Kong, Hong Kong, China, [3]Basque Center on Cognition, Brain and Language, San Sebastian, Spain, [4]IKERBASQUE, Basque Foundation for Science, Bilbao, Spain, and [5]Utrecht Institute of Linguistics OTS, Utrecht University, Utrecht, Netherlands
Address for correspondence:
Patrick C.M. Wong, G03, Leung Kau Kui Building, Department of Linguistics and Modern Languages, The Chinese University of Hong Kong, Hong Kong. Email: p.wong@cuhk.edu.hk

(Received 8 February 2020; revised 3 November 2020; accepted 9 November 2020;
first published online 3 February 2021)

Abstract

The functions of acoustic-phonetic modifications in infant-directed speech (IDS) remain a question: do they specifically serve to facilitate language learning via enhanced phonemic contrasts (the hyperarticulation hypothesis) or primarily to improve communication via prosodic exaggeration (the prosodic hypothesis)? The study of lexical tones provides a unique opportunity to shed light on this, as lexical tones are phonemically contrastive, yet their primary cue, pitch, is also a prosodic cue. This study investigated Cantonese IDS and found increased intra-talker variation of lexical tones, which more likely posed a challenge to rather than facilitated phonetic learning. Although tonal space was expanded which could facilitate phonetic learning, its expansion was a function of overall intonational modifications. Similar findings were observed in speech to pets who should not benefit from larger phonemic distinction. We conclude that lexical-tone adjustments in IDS mainly serve to broadly enhance communication rather than specifically increase phonemic contrast for learners.

Keywords: infant-directed speech; pet-directed speech; the hyperarticulation hypothesis; the prosodic hypothesis; lexical tones

Introduction

When speaking to infants, a special speech register, infant-directed speech (IDS), is widely employed. IDS is distinctive from adult-directed speech (ADS) in both prosodic and phonemic characteristics. Prosodically, IDS intonation is characterized by higher overall pitch, greater pitch variability within and across utterances, and simplified and smoothed pitch contours (Fernald et al., 1989; Fernald & Kuhl, 1987; Fernald & Simon, 1984; Jacobson, Boersma, Fields & Olson, 1983; Kitamura & Burnham, 2003; Papoušek, Papoušek & Symmes, 1991; Stern, Spieker, Barnett &

© The Author(s), 2021. Published by Cambridge University Press

https://doi.org/10.1017/S0305000920000707 Published online by Cambridge University Press

MacKain, 1983). Phonemically, the most widely discussed distinction of IDS is the expansion of vowel space, manifested in a variety of languages including English, French, Russian, Swedish, Japanese, and Mandarin Chinese (Burnham, Kitamura & Vollmer-Conna, 2002; Dodane & Al-Tamimi, 2007; Kuhl, Andruski, Chistovich, Chistovich, Kozhevnikova, Ryskina, Stolyarova, Sundberg & Lacerda, 1997; Liu, Kuhl & Tsao, 2003).

IDS constitutes an important component of the language input that infants receive from the environment (Soderstrom, 2007). However, its primary functions remain an open question. Different acoustic-phonetic components of IDS have been proposed to serve different purposes. The prosodic exaggeration in IDS has generally been proposed to serve a communicative function, to attract, maintain and regulate the attention of infants, and to convey communicative intent, especially positive emotions, to boost infant social communication (Fernald, 1989; Fernald & Simon, 1984; Grieser & Kuhl, 1988). When considering the phonemic modification, two hypotheses have been proposed. The hyperarticulation hypothesis (HH) postulates that IDS serves a didactic linguistic function to facilitate infant language learning by providing enhanced phonemic contrasts (Burnham et al., 2002; Kuhl et al., 1997; Liu et al., 2003; Werker et al., 2007; Xu Rattanasone, Burnham & Reilly, 2013). Alternatively, the prosodic hypothesis (PH) argues that IDS serves a communicative function by exaggerating prosody, and modifications of phonemes are unintended byproducts of prosodic modulations (Benders, 2013; McMurray, Kovack-Lesh, Goodwin & McEchron, 2013; Tang, Xu Rattanasone, Yuen & Demuth, 2017; Wong & Ng, 2018).

The large body of IDS research has primarily examined non-tone languages, in which prosodic modulation and phonemic contrasts are marked separately by suprasegmental and segmental cues. By contrast, in tone languages, lexical tones mark phonemic changes in a similar manner to vowels and consonants (see Yip, 2002), but share the prosodic cue pitch (or acoustically speaking f0, i.e., fundamental frequency) with intonation, which serves paralinguistic purposes, thus resulting in a unique test case for identifying the primary function of acoustic-phonetic modifications in IDS. Previous studies on lexical tone adjustments in IDS have drawn inconsistent conclusions on the functions that they primarily serve. A number of studies have reported an acoustic enhancement of lexical tones in IDS, which has been considered as evidence for a specific didactic linguistic function of IDS, as argued by HH (Cheng & Chang, 2014; Han, de Jong & Kager, 2018; Liu, Tsao & Kuhl, 2007; Xu Rattanasone et al., 2013). However, a few other studies have argued for the possibility that lexical tone adjustments in IDS are driven by intonational effects, in line with PH (Papoušek & Hwang, 1991; Tang et al., 2017; Wong & Ng, 2018), suggesting that a non-linguistic communicative function may be prioritized over a didactic linguistic function. To gain a clearer insight into this issue, the current study conducted a more comprehensive investigation into the adjustments of lexical tones in IDS of Cantonese and their relations with intonational modifications.

The hyperarticulation hypothesis (HH)

HH is mainly based on the findings of a larger acoustic vowel space in IDS than ADS (Burnham et al., 2002; Dodane & Al-Tamimi, 2007; Kuhl et al., 1997; Liu et al., 2003). The vowel space is defined by the area of the vowel triangle formed when the first and second formants (F1 and F2) of the three most peripheral vowels /a/, /i/ and /u/ are plotted in two-dimensional space. The findings of vowel space expansion in IDS have

https://doi.org/10.1017/S0305000920000707 Published online by Cambridge University Press

led to the postulation that speakers hyper-articulate vowels in IDS to enlarge cross-category acoustic contrasts to facilitate infants' learning (Kuhl et al., 2008).

Further evidence for HH was provided by a series of studies conducted mainly by Burnham and colleagues, which suggested that vowel space expansion in IDS was closely related to the audience's linguistic competence. They reported that, in speech to foreigners, vowel space expansion was found as in IDS (Uther, Knoll & Burnham, 2007), while in speech to pets, there was no sign of such exaggeration (Burnham et al., 2002; Xu, Burnham, Kitamura & Vollmer-Conna, 2013; and see Gergely, Faragó, Galambos & Topál, 2017 for more recent evidence). However, there was vowel exaggeration in speech to parrots, a pet with a perceived potential to learn language (Xu et al., 2013). More evidence came from cross-language research, showing that acoustic realization of vowels in IDS was language-specific. In IDS of Japanese and English, mothers distinguished the vowels using a language-specific vowel-contrasting cue: which is vowel duration for Japanese, and vowel color (spectral differences) for English (Werker et al., 2007). Empirical studies on the influence of IDS on infant language development have also provided evidence for HH. A positive association was reported between the degree of the expansion of vowel space in individual caregivers' IDS and their infants' ability to discriminate consonants in the first year of life (Liu et al., 2003), and expressive and receptive vocabulary size in the second year of life (Hartman, Bernstein Ratner & Newman, 2017; Kalashnikova & Burnham, 2018). In a lexical processing task, typical-IDS vowel exaggeration in the speech stimuli immediately led to an improvement in the ability of 19-month-old infants to recognize familiar words (Song, Demuth & Morgan, 2010).

However, HH has also been challenged in a number of respects. A few studies failed to observe any expansion of the acoustic vowel space in IDS of some languages such as Norwegian and Dutch (Benders, 2013; Englund & Behne, 2005), suggesting that the enhancement of vowel space may not be universal. An investigation into Norwegian IDS even showed hypoarticulation of non-peripheral vowels with larger overlap between vowel contrasts (Englund, 2018). Moreover, it has been shown that having an expanded vowel space does not necessarily imply that IDS provides clearer and more distinguishable phonemic categories. When paying attention to non-peripheral vowels, no consistent increase of cross-category separation was observed in English IDS, even though an expansion of vowel space determined by the peripheral vowels was found (Cristia & Seidl, 2014). In addition, when an algorithm was employed to perform a vowel classification task, vowels in IDS were found to be less clear and more difficult to discriminate than those in ADS (Martin et al., 2015).

More importantly, a number of studies demonstrated that within-category vowel variation produced by an individual speaker was increased in IDS compared to ADS, and vowel differentiation when considering within-category variation was not enhanced in IDS (Cristia & Seidl, 2014; McMurray et al., 2013; Miyazawa, Shinya, Martin, Kikuchi & Mazuka, 2017). Larger vowel variation can lead to a greater overlap between vowel categories and thus less distinct contrasts between them, which is assumed to counteract the benefit of the expansion of vowel space and hinder infants' distributional category learning (Cristia & Seidl, 2014; McMurray et al., 2013). For instance, the Native Language Magnet (NLM) model claims that the prototypes of the phonemic category, i.e., the exemplars most frequently activated, serve as referents for infants' phonetic learning; specifically, infants showed greater generalization from the prototype to other tokens in the category than the

https://doi.org/10.1017/S0305000920000707 Published online by Cambridge University Press

other way around (Kuhl, 1991; Kuhl, Conboy, Coffey-Corina, Padden, Rivera-Gaxiola & Nelson, 2008). Therefore, while prototypical vowel tokens separated by larger acoustic distance in IDS would support phonetic category learning (as per HH), increased within-category variation which leads to more tokens deviating from the prototypes and greater between-category overlap would instead be detrimental. It is worth noting that another possibility has also been proposed that larger within-category variation may not reduce the learnability of phonemic categories (see the General Discussion for more details). For instance, Eaves, Feldman, Griffiths and Shafto (2016) found that phonemic variation within categories increased in the ideal data that they generated for teaching phonetic categories; therefore, they argued that the larger within-category variation in IDS may be beneficial for phonetic learning. Despite the alternative possibility, the findings of increased within-category vowel variation in IDS further cast doubt on the linguistic function of this register.

The prosodic hypothesis (PH)

The alternative PH argues that acoustic-phonetic modifications in IDS, specifically exaggerated prosody, primarily serve the communicative function to transmit positive affect to infants and capture their attention, which may incidentally bring about the adjustments in phonemes. According to this hypothesis, the acoustic adjustments in IDS that lead to less distinct phonemic contrasts, such as the hypoarticulation of non-peripheral vowels and increase of within-category vowel variation could be explained as byproducts of prosodic exaggeration. A number of studies have demonstrated that the phonemic adjustments observed in IDS may originate from prosodic modifications for communicative purposes.

McMurray et al. (2013) found that the effects of IDS register on vowel formants resembled the effects of the prosodic position; vowels in the prosodically strongest position exhibit acoustic enhancement similar to vowels in IDS. Adriaans and Swingley (2017) further showed a co-occurrence of vowel space expansion and prosodic exaggeration in IDS. Vowel tokens in IDS were perceptually judged to be prosodically exaggerated or not. The vowels in the prosodically exaggerated position were found to be hyperarticulated with larger overall space and between-category distance as compared to the vowels in the prosodically non-exaggerated position.

A second line of evidence comes from studies relating acoustic changes in vowel formants in IDS to emotion expression. Benders (2013) observed higher F2 and F3 of vowels and a higher spectral mean of fricatives in Dutch IDS, which are all acoustic markers of positive affect. Similarly, Tang et al. (2017) demonstrated that vowel space expansion in Mandarin IDS resulted from an increase of F1 and F2 for low and back vowels, which is similar to the acoustic characteristics of happy speech. Furthermore, Benders (2016) simultaneously rated IDS utterances perceptually on emotion-related factors and acoustically analyzed the low-back vowel /ɑ/ contained in the utterances. The study found that a higher F1 of the vowel can be predicted from a higher percept of emotional energy of the utterance, and a higher F2 of the vowel can result in the percept of more child-like utterances produced with more smiling.

In addition, from the articulatory perspective, Kalashnikova, Carignan, and Burnham (2017) reported that mothers did not exaggerate tongue and lip movements during IDS production. Instead, they shortened their vocal tract by raising the larynx to signal non-threatening attitude, and thereby express emotion and maintain infant attention, which acoustically resulted in higher pitch as well as

https://doi.org/10.1017/S0305000920000707 Published online by Cambridge University Press

larger vowel space. In other words, the acoustic expansion of vowel space appears to be a side effect of shortening the vocal tract which mainly serves communicative purposes rather than a consequence of an (unconscious) desire to teach infants about vowel categories.

It is worth noting that PH and HH are not mutually exclusive. It is possible that IDS serves both communicative and linguistic functions, and certain phonemic modification such as the vowel space expansion results from both didactic hyperarticulation and prosodic exaggeration. Moreover, even if the vowel space expansion is merely a byproduct of prosodic modifications, it does not exclude the possibility that infants benefit from it in phonetic learning (Adriaans & Swingley, 2017; Kalashnikova et al., 2017).

Lexical tones in IDS

The two hypotheses were tested in the case of lexical tones in the current study. Most previous research on lexical tones in IDS of different tone languages has demonstrated some signs of lexical tone exaggeration, which have been interpreted in support of HH. Xu Rattanasone et al. (2013) found that Cantonese IDS to 3-, 6-, and 9-month-old infants was produced with a larger tonal space than ADS. The tonal space was measured as the area of the tone triangle determined by the averaged onset and offset f0 of three peripheral lexical tones of Cantonese, i.e., T1, T2, and T4. More importantly, an age-related reduction was observed for such expansion of tonal space for infants from 3 months to the end of the first year of life, which seemingly corresponded to the timeline of infants' perceptual attunement for lexical tones. In addition, an enhancement of pairwise differences between lexical tones was observed in IDS of Hakka (Cheng & Chang, 2014) and Mandarin for infants in both the first and second year of life (Han, de Jong & Kager, 2018; Liu, Tsao & Kuhl, 2007).

However, these findings do not provide sufficient evidence to determine whether tone discriminability increases in IDS compared to ADS, since they do not account for the degree of within-category tone variation. An enlarged tonal space would not guarantee a better separation of tone categories given that within-category variation may increase simultaneously. This would lead to more varied tokens deviating from the prototypes and thus more overlap among tone categories, especially for languages with a large tone inventory and hence a crowded tonal space like Cantonese. According to the above discussion, the increase of tone variation may also be detrimental to infants' distributional learning of tone categories. Taking the three level tones of Cantonese as an example, tokens of T3 (mid-level) produced with too high or low pitch are likely to be identified as T1 (high-level) or T6 (low-level) respectively (Wong & Diehl, 2003), thus posing challenges for infants to form representations of the three tone categories.

If larger within-category tone variation is found in IDS as with vowels (Cristia & Seidl, 2014; McMurray et al., 2013; Miyazawa et al., 2017), this adjustment is also very likely to be explained by the intonational modifications in IDS for communicative purposes, as predicted by PH. When intonation is exaggerated, in particular when pitch variability becomes larger at the utterance level, it may force the pitch realization of tone tokens at the subordinate syllabic level to vary more and perhaps also provides a larger acoustic space for tones to disperse from each other. Several studies have addressed the possibility that lexical tone adjustments in IDS are intonational effects, providing evidence for PH. Papoušek and Hwang (1991) claimed that tone contrasts in Mandarin IDS were reduced and the pitch realization of lexical

https://doi.org/10.1017/S0305000920000707 Published online by Cambridge University Press

Table 1. The six Cantonese lexical tones.

Tone	Pitch pattern	Example
Tone 1 (T1)	High-level	/si1/ *poem*
Tone 2 (T2)	High-rising	/si2/ *history*
Tone 3 (T3)	Mid-level	/si3/ *to try*
Tone 4 (T4)	Low-falling	/si4/ *time*
Tone 5 (T5)	Low-rising	/si5/ *market*
Tone 6 (T6)	Low-level	/si6/ *yes*

Notes. The 'high', 'mid' and 'low' in 'pitch pattern' indicate the relative pitch height of the tones, and 'level' (i.e., relatively constant), 'rising' and 'falling' signify the pitch contours of the tones.

tones was modified in correspondence with the expanded global intonation contour; by contrast, foreigner-directed speech maintained the contrasts between tone categories. Unfortunately, these conclusions regarding lexical tones were made through visual inspections of the pitch contours. Two more recent studies employed more objective acoustic and perceptual measures. Wong and Ng (2018) reported that lexical tones from IDS of Cantonese were identified by adult native speakers with a lower accuracy rate than those from ADS. Furthermore, the error patterns in identifying lexical tones in IDS resembled the errors made in perception of lexical tones in the utterance-final position where lexical tones were influenced by intonational pitch modulation. Tang et al. (2017) found that tonal space in Mandarin IDS to 12-month-old infants was only expanded in the utterance-final position, indicating that lexical tone modifications may be driven by the mother's communicative goal to convey positive affect, since the findings were consistent with the exaggeration of pitch contour in the utterance-final position in happy speech. In addition, they found no expansion of tonal space in Lombard speech, further casting doubt on the linguistic function of tonal space expansion.

The current study

The current study aimed to provide new evidence for the debate on the primary function of acoustic-phonetic modifications in IDS by testing the two hypotheses, HH and PH, specifically with lexical tones in IDS of Cantonese, a tone language with one of the most complex lexical tone systems (Wong & Chan, 2018) (see Table 1 and Figure 1 for the six Cantonese lexical tones). Since the two hypotheses are not mutually exclusive, our goal is not to completely reject one and accept the other. Instead, by testing lexical tones, we intend to explore which function would be the priority in producing IDS, when the acoustic-phonetic realization needs to consider both phonemic and prosodic factors.

Two specific research aims were pursued. The first aim was to test whether contrasts between lexical tones are enhanced in Cantonese IDS when taking intra-talker within-category tone variation into consideration. In Experiment 1, Cantonese-speaking caregivers' IDS was compared to their ADS to examine the change of tonal space, tone variation, and tone differentiation considering both tonal space and tone variation.

The second aim was to test whether lexical tone adjustments in Cantonese IDS can be explained as byproducts of intonational modifications. Although previous studies

https://doi.org/10.1017/S0305000920000707 Published online by Cambridge University Press

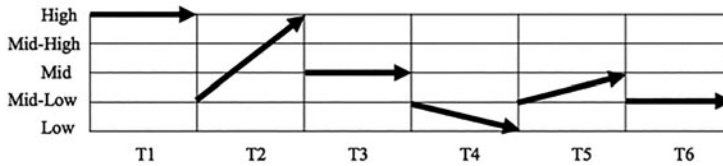


Figure 1. An illustration of the pitch patterns of the Cantonese lexical tones and their relative positions in the tonal space.

have indicated some potential relations of lexical tone adjustments in IDS to intonational exaggeration (Papoušek & Hwang, 1991; Tang et al., 2017; Wong & Ng, 2018), they did not reveal how a specific aspect of lexical tone adjustments in IDS is influenced by certain aspects of intonational modifications. To promote understanding of this issue, two steps were taken in the current study.

First, correlations were assessed between lexical tone and intonational modifications in individual caregivers' IDS collected in Experiment 1. Instead of subjective perceptual assessments conducted by the previous studies for emotional-prosodic information of individual utterances (Adriaans & Swingley, 2017; Benders, 2016), the current study applied multiple objective acoustic measures, including not only measures of intonational modulation within individual utterances but also a global measure of intonational variability across utterances. Lexical tones were also assessed by global measures based on a number of tone tokens to reflect the overall adjustments of the whole tonal space and tone variation in a speaker's IDS. Due to these differences in measurement, the current study correlated lexical tone and intonational modifications across speakers instead of using data from individual utterances and tone tokens within speakers as done in previous studies.

Next, Experiment 2 was conducted, in which the caregivers were recorded producing pet-directed speech (PDS) to test whether IDS-like lexical tone adjustments can be observed in PDS when there are intonational effects alone without the need for a didactic linguistic function. On the one hand, like IDS, PDS serves a communicative function: it can attract the attention of pets better than ADS; pets show a preference for PDS over ADS (Benjamin & Slocombe, 2018; Jeannin, Gilbert, Amy & Leboucher, 2017); and PDS has higher positive affect than ADS (Burnham et al., 2002). Correspondingly, intonation in PDS undergoes modifications similar to IDS, showing an increase of pitch height and pitch range compared to ADS (Burnham et al., 1998; Burnham et al., 2002; Gergely et al., 2017). On the other hand, unlike IDS, there should not be any hyperarticulation of phonemic categories for didactic purposes in PDS since speakers should not expect pets to learn language (Burnham et al., 2002). Previous studies have reported no expansion of vowel space in PDS as mentioned above. PDS thus offers an approach to test how IDS-like intonational modifications may drive adjustments of lexical tones, with the factor of didactic hyperarticulation excluded.

Pursuing the two research aims could enlighten us about whether the communicative function argued by PH or the linguistic function argued by HH is the priority in producing IDS (despite that the two functions may co-exist). We would tend to prefer PH, if 1) the increase of tone variation counteracted the tonal space expansion resulting in little enhancement of tone differentiation (based on the assumption that an increase of within-category variation reduces the learnability of

https://doi.org/10.1017/S0305000920000707 Published online by Cambridge University Press

phonemic contrasts); 2) all lexical tone adjustments were found to be explained by intonational effects: that is, they were positively correlated with intonational exaggeration and present in PDS. In contrast, we would be more inclined to favor HH, if 1) tone differentiation as a ratio of tonal space to tone variation was enhanced; 2) lexical tone adjustments, at least regarding the tonal space expansion, could not be explained by intonational effects: that is, they were not positively related to intonational modifications and not present in PDS.

Experiment 1

In this experiment, we first compared lexical tones in caregivers' speech to their 15-month-old infants (IDS) and an adult experimenter (ADS) to address the first research aim. Then the lexical tone adjustments in IDS were correlated to intonational modifications produced by individual caregivers to address the second research aim. The 15-month-olds were recruited to promote understanding of lexical tones in Cantonese IDS to infants at an understudied developmental stage. Previously, a careful examination of lexical tones in IDS of Cantonese has been conducted for infants within the first year of life (Xu Rattanasone et al., 2013), whereas little is known about the subsequent developmental course in the second year of life when infants face an increasing challenge of associating sounds with objects. In addition, out of methodological consideration, the inclusion of 15-month-olds rather than younger infants enabled the caregivers to more naturally produce all the target words required for lexical tone analyses in their interaction with the infants.

Method

Participants

Data were collected from 28 native Cantonese-speaking caregivers and their 15-month-old Cantonese-learning monolingual infants (M = 1;03.03; range: 1;02.04–1;03.22; 13 males and 15 females). The participants were all Chinese recruited in Hong Kong. The caregivers were the primary caregivers of the infants from birth up until the time of the study. Twenty-seven of these caregivers were mothers and one was a grandmother. According to the caregivers' reports, their infants and themselves were not affected by any mental, sensory or language deficits. As far as the families' socioeconomic status was concerned, parental education ranged from high school to graduate school, and their occupation ranged from semiskilled workers to major professionals. The participants were recruited by advertising in WhatsApp groups. Written informed consent approved by The Joint Chinese University of Hong Kong - New Territories East Cluster Clinical Research Ethics Committee was obtained from the caregivers.

Stimuli and materials

Six target words (see Table 2), corresponding to the six Cantonese lexical tones, were selected to be elicited for both IDS and ADS. Three criteria were followed in selection of the target words. First, all target words should have the same vowel, so that the analysis of lexical tones would not be affected by the difference of vowels due to their intrinsic pitch (i.e., high vowels intrinsically have higher pitch than low vowels) (Lehiste & Peterson, 1961). Second, the consonants were chosen to be all

https://doi.org/10.1017/S0305000920000707 Published online by Cambridge University Press

Table 2. The six target words.

Tone	Tone1	Tone2	Tone3	Tone4	Tone5	Tone6
Logograph	飛	畀(俾)	四	旗	被	鼻
IPA	[feɪ]	[peɪ]	[seɪ]	[k^heɪ]	[p^heɪ]	[peɪ]
Gloss	Fly	Give	Four	Flag	Quilt	Nose
Word class	Verb	Verb	Quantifier	Noun	Noun	Noun

Notes. The six target words selected to be elicited in recordings of both IDS and ADS. Each of the target words carries one of the six Cantonese lexical tones. IPA: International Phonetic Alphabet.

voiceless fricatives or stops. Finally, from the semantic perspective, we selected words that were as appropriate as possible for the caregivers to use with their infants, in order to ensure the naturalness of their speech. The final six words belong to several different word classes, including three nouns, two verbs and a quantifier.

Toys were prepared to elicit each target word in caregiver-infant interaction (see Table 3), and each toy had a small label with its corresponding Chinese character.

Recording procedure

For all participants, caregiver-infant interaction was collected first for IDS, and then a conversation between the caregiver and an adult experimenter was recorded for ADS.

IDS recording

The caregivers interacted with their infants in a sound-attenuated booth in our laboratory. During the recording, the infants sat in a baby chair, facing their caregiver. The caregivers wore a small, skin-colored cardioid condenser head-mounted microphone (Audio-Technica BP894). The microphone was connected to a laptop (MacBook Air) through an audio interface (Roland Quad-Capture) which were both put outside the booth for speech recording. Recordings were made with 44100Hz sampling rate, and 16 bits sampling precision.

The caregivers were instructed to naturally interact with their infant in a play session with the toys as they normally did at home. They were encouraged to use the words labelled on the toys in the interaction. Instructions like "please remember to speak the words when appropriate during the interaction" were given to the caregivers. For each word, all corresponding toys were put in a cloth bag (in total six bags). The caregivers were provided with one bag at a time, and the six bags were given in a random order. The experimenter monitored the caregiver-infant interaction through a headphone connected to the audio interface outside the booth and counted the number of tokens produced by the caregivers. A minimum of 10 utterances containing the target words needed to be met for each word before the next bag was given. The caregivers in general took little effort to produce enough tokens during the interaction. On average, about 1-2 minutes were needed for a caregiver to meet the 10-utterance criterion. The caregivers would be allowed to continue playing with the toys for a little longer, if their infant still showed great interest in these toys (but no longer than 5 minutes for each target word). During IDS recordings, only caregivers and infants were in the booth, and no experimenter was present.

https://doi.org/10.1017/S0305000920000707 Published online by Cambridge University Press

Table 3. The toys prepared to elicit the target words in caregiver-infant interaction.

Target words	Toys prepared
飛 [feɪ] fly	Two toy planes; two toy birds
畀 [peɪ] give	Three plush toys
四 [seɪ] four	Two sets of rubber toys with four different animals and a family of four people respectively; a set of plush toys with a family of four pigs
旗 [k^heɪ] flag	Four small flags of different countries; five flashcards demonstrating different national flags; a string of colorful pennants
被 [p^heɪ] quilt	A small toy quilt for the doll
鼻 [peɪ] nose	Five animal cards with removable noses; a plush elephant with a long nose

ADS recording

ADS was recorded in the same sound-attenuated booth, by the same equipment with the same parameters. An adult experimenter who was a native speaker of Cantonese conducted a conversation with the caregiver. The experimenter managed to elicit the target words from the caregivers by asking questions about the toys, such as their infant's interest in and previous experience with the toys (e.g., "Does the infant like the toy plane?" "Has the infant played with toy planes before?"). The caregivers were encouraged to produce the target words during the conversation. It proved to be difficult to elicit as many tokens as for IDS. Efforts were made to elicit at least 8 utterances for each target word instead (Tang et al., 2017). However, for some caregivers, this criterion still could not be met for some target words. In this case, tokens were elicited – as many as the experimenter was able to. The experimenter would end the conversation about a target word and moved on to the next if she judged that the caregiver was clearly tired of the current topic, or the conversation about this target word had exceeded 5 minutes. On average, about 3 minutes were spent on every target word.

Data analysis

Pre-processing

For each target word, the first 10 utterances of good quality produced by every caregiver containing the word were extracted from both IDS and ADS recordings for analysis. A section of speech was segmented as an utterance if it was separated from previous and following speech by more than 300 ms pause or non-speech, following the criterion previously used (Fernald & Simon, 1984, Fernald et al., 1989, Kitamura, Thanavishuth, Burnham & Luksaneeyanawin, 2002). Utterances with noises (including interruption from infants such as crying and vocalizations), and one-word utterances (i.e., containing only the target word) were excluded. For some participants who did not produce a sufficient number of utterances for certain target words especially in ADS, as many utterances were extracted as possible (see Table 4 for the number of utterances extracted).

The extracted utterances of all six target words were collapsed for the analysis of intonation, and the target words in these utterances were labelled for the analysis of lexical tones (see Table 4 for the number of target words). A small portion of the

https://doi.org/10.1017/S0305000920000707 Published online by Cambridge University Press

https://doi.org/10.1017/S0305000920000707 Published online by Cambridge University Press

Table 4. The number of utterances, target words, and target words per utterance for IDS and ADS analysis.

Tone		Tone1	Tone2	Tone3	Tone4	Tone5	Tone6
Number of utterances	IDS	9.96 ± .19	9.93 ± .26	10 ± 0	9.96 ± .19	9.93 ± .37	9.96 ± .19
	ADS	9.54 ± .78	8.86 ± 1.41	8.93 ± 1.39	7.54 ± 1.88	8.43 ± 1.66	8.29 ± 1.67
Number of target words	IDS	13.18 ± 3.21	12.14 ± 2.01	10.89 ± .9	11.14 ± 1.53	13.93 ± 3.92	15.29 ± 4.14
	ADS	11.32 ± 2.07	10.71 ± 2.1	10.11 ± 1.93	8.25 ± 2.46	9.61 ± 1.99	10 ± 2.67
Number of target words per utterance	IDS	1.32 ± .32	1.22 ± .21	1.09 ± .09	1.12 ± .15	1.4 ± .39	1.53 ± .41
	ADS	1.19 ± .19	1.22 ± .18	1.13 ± .12	1.09 ± .13	1.15 ± .11	1.2 ± .18

Notes. These are all numbers averaged across the 28 caregivers (± standard deviation).

target words was produced in the utterance-final position; these tokens were included in the analyses of lexical tones since it was found that excluding them did not change the results (see Appendix 2).

Acoustic analysis was done in Praat (Boersma & Weenink, 2017) using the 'prosodypro' script (Xu, 2013). For analysis of lexical tones, the voiced portion of the target words was labelled by hand. The onset and offset were marked at the zero crossing point of the first and last pulse respectively that extended through F1 and or F2 (Liu et al., 2007). The onset and offset of the utterances, also labelled by hand, were the onset of the voiced portion of the first syllable and the offset of the voiced portion of the last syllable respectively. The f0 data were provided by prosodypro with a sampling rate of 100Hz. The vocal cycle marks generated by Praat for f0 calculation were checked and corrected manually. All f0 measurements were converted to the Equivalent-rectangular-bandwidth-rate (ERB) scale, since this psychoacoustic measure was believed to best characterize f0 changes from the perspective of speech perception (Hermes & Van Gestel, 1991; Liu et al., 2007).

Acoustic measurements

All the measurements of lexical tones and intonation were made for both IDS and ADS, and the IDS/ADS ratios were then calculated to index the IDS changes.

For lexical tones, tonal space dispersion, intra-talker tone variation within categories, and tone differentiation were quantified for every caregiver using methods developed from Zhao and Jurafsky (2009). F0 data at 10 equidistant time points along each tone contour were used throughout the measurements. Equations with more details are provided for the following measurements in Appendix 1.

1) Tonal space dispersion (see Figure 2A): an overall central tone contour was first calculated for each caregiver by averaging the contours of all tone tokens produced by this caregiver at every time point. Then the Euclidean distance between the contour of every tone token and the overall central tone contour was calculated for each time point. The tonal space dispersion was finally calculated by averaging the Euclidean distances across all tone tokens and all 10 time points.
2) Intra-talker tone variation within categories (see Figure 2B): a central tone contour was first calculated for each lexical tone by averaging the contours of all tone tokens within that tone category at every time point. Then the Euclidean distance between the contour of every tone token within the tone category and the central tone contour was calculated for each time point. The variation of this tone category was calculated by averaging the Euclidean distances across all tone tokens within the tone category and all 10 time points. Finally, the variation of all six lexical tones was averaged to index the overall tone variation of a caregiver.
3) Tone differentiation: it was computed as the ratio of tonal space dispersion to tone variation. For an individual caregiver, the larger the tonal space dispersion, and the smaller the tone variation, the greater the tone differentiation would be.

The method above has not been used in IDS research of Cantonese. The tonal space dispersion resembles the tonal space based on tone triangle measured in Xu

https://doi.org/10.1017/S0305000920000707 Published online by Cambridge University Press

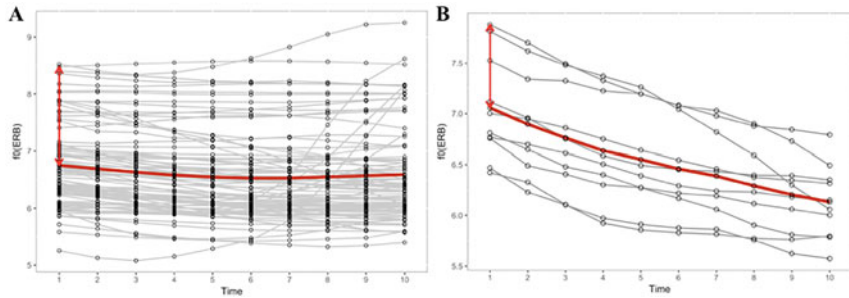


Figure 2. (A) An illustration of the overall central tone contour and the measuring of tonal space dispersion, using IDS data from one caregiver. The dots on the lines are the 10 equidistant time points sampled along each tone contour. The grey lines represent the pitch contour of all tone tokens produced by this caregiver, and the red line represents the overall central tone contour of this caregiver. The red double-headed arrow exemplifies the calculation of the Euclidean distance between the pitch contour of one tone token and the overall central tone contour at the first time point; (B) An illustration of the central tone contour of one lexical tone and the measuring of variation of this tone category, taking Tone 4 as an example, using IDS data from one caregiver. The dots on the lines are the 10 equidistant time points sampled along each tone contour. The grey lines represent the pitch contour of all tokens of Tone 4, and the red line represents the central tone contour of Tone 4. The red double-headed arrow exemplifies the calculation of the Euclidean distance between the pitch contour of one tone token and the central tone contour of Tone 4 at the first time point.

Note. f0: fundamental frequency; ERB: Equivalent-rectangular-bandwidth-rate.

Rattanasone et al. (2013) for lexical tones in Cantonese IDS, but has advantages over the latter one in two respects. On the one hand, it considers all six lexical tones rather than only the three peripheral tones. On the other hand, it is based on the data of 10 sampling points along the tone contour instead of only the onset and offset f0, which can capture more information of the temporal f0 changes that are important for lexical tones. This measurement, however, has a shortcoming that it mainly demonstrates the absolute f0 distances at which tone categories disperse from the centroid, but does not take into consideration whether each tone token is well realized; thus it alone cannot quantify the acoustic contrasts between tone categories, in particular with respect to the contour tones.

Alongside the new measurements, to compare to previous findings from Xu Rattanasone and colleagues' study, we also measured the tonal space which was the area of the tone triangle formed when the averaged onset and offset f0 of the three peripheral lexical tones of Cantonese, i.e., T1 (high-level), T2 (high-rising), and T4 (low-falling), were plotted in two-dimensional space.

Intonation was also measured to prepare for the test of correlation between lexical tone and intonational modifications. With the sampled f0 data provided by prosodypro, four measurements were made:

1) Mean pitch of utterances. The mean f0 was calculated for each utterance, and then averaged across all selected utterances.
2) Pitch range of utterances. The difference between the maximum and minimum f0 within each utterance was calculated and then averaged across utterances.
3) Pitch variability within utterance. The standard deviation (SD) of f0 across all sampling points was calculated within each utterance and then averaged across utterances. Although pitch range also reflects pitch variation within utterance,

https://doi.org/10.1017/S0305000920000707 Published online by Cambridge University Press

this measurement has the advantage of considering not only the maximum and minimum f0. However, it only has been employed in IDS research of non-tone languages (Jacobson et al., 1983), whereas pitch range has been measured in tone languages including Cantonese (Xu & Burnham, 2010). Thus, pitch range was still measured in the current study to compare to previous findings.

4) Pitch variability across utterances. The mean f0 was calculated for each utterance, and then the SD of mean f0 was calculated across utterances.

Statistical analysis

Comparisons were made between IDS and ADS for all the measurements of lexical tones and intonation using paired t-tests. Then, Pearson correlation tests were conducted between the IDS/ADS ratios of lexical tone measurements and intonation measurements. To avoid redundancy, the tonal space based on tone triangle and intonational pitch range were not included in the correlation tests. Tone differentiation was also not included as it was covered by the tonal space dispersion and tone variation measures.

Results and discussion

Lexical tone changes

By calculating the area of tone triangle (see Figure 3A) for every caregiver, our results showed a significantly larger tonal space in IDS compared to ADS ($t(27) = 4.26$, $p < .001$, $d = .97$), replicating previous findings for Cantonese IDS to infants within the first year of life (Xu Rattanasone et al., 2013). Moreover, the results from the new measurement of tonal space dispersion (see Figure 3B) of every caregiver further confirmed the expansion of tonal space in IDS with all six lexical tones considered. Tonal space dispersion was significantly larger in IDS than in ADS ($t(27) = 8.88$, $p < .001$, $d = 1.88$). However, the intra-talker tone variation of every caregiver was also found to be significantly greater in IDS than in ADS ($t(27) = 7.33$, $p < .001$, $d = 1.69$). Consequently, tone differentiation, as the ratio of tonal space dispersion to tone variation for every caregiver, was not found to be significantly different between IDS and ADS ($t(27) = 1.02$, $p = .32$, $d = .29$) (see Figure 3C).

Intonational changes

Compared to ADS, IDS showed higher mean pitch ($t(27) = 12.06$, $p < .001$, $d = 2.51$), reduced pitch range ($t(27) = -3.46$, $p = .002$, $d = -.66$), and greater pitch variability across utterances ($t(27) = 10.34$, $p < .001$, $d = 2.23$), while no significant difference was found for pitch variability within utterance ($t(27) = 1.29$, $p = .21$, $d = .27$) (see Figure 4A). Among the four measurements, the mean pitch and pitch range have been measured for Cantonese IDS to 6-month-old infants (Xu & Burnham, 2010), and the current results replicated previous findings. The pitch range was determined by the minimum and maximum pitch in every utterance. We found that, in IDS, the minimum pitch of utterances increased to a much larger degree than did the maximum pitch (see Figure 4B), which could explain why the pitch range appeared to be reduced in IDS compared to ADS.

Lexical tone - intonation correlation

Significant positive correlations were found between the changes (IDS/ADS) of lexical tones and intonation (see Table 5). Not only was the increase of tone variation

https://doi.org/10.1017/S0305000920000707 Published online by Cambridge University Press

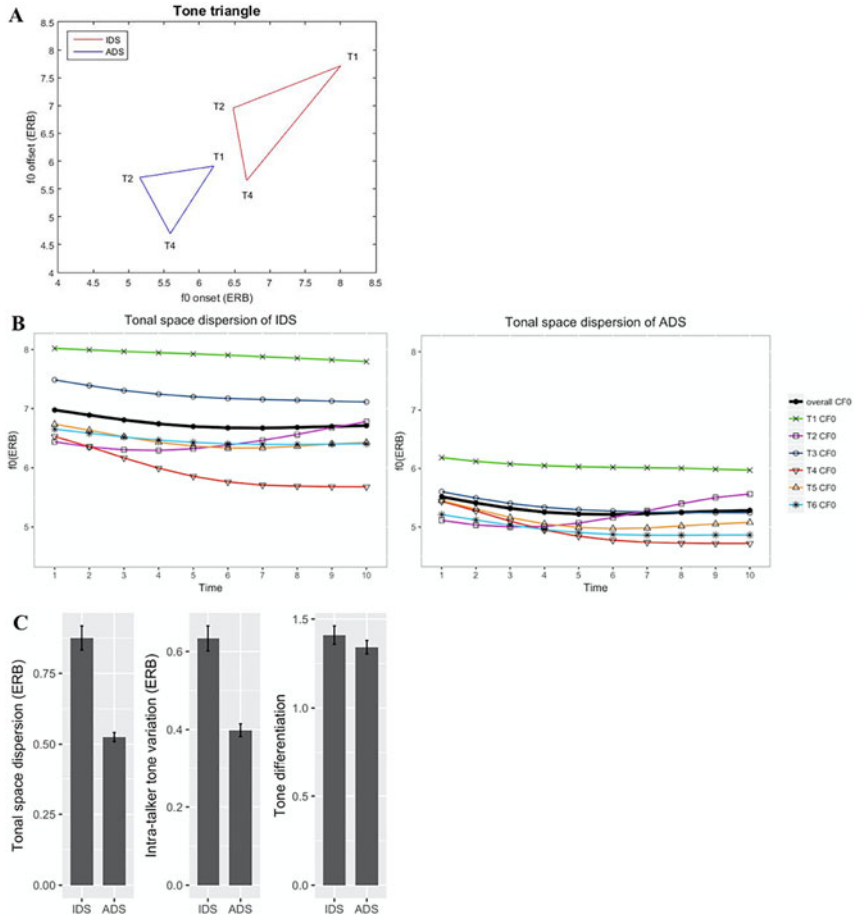


Figure 3. (A) The tone triangle of IDS and ADS determined by the averaged onset and offset f0 of the three corner lexical tones of Cantonese, i.e., T1, T2, and T4, using the average data of all subjects; (B) The tonal space dispersion of IDS and ADS shown by the central contours *CF0* of the six tones and the overall central tone contour, using the average data of all subjects; (C) The tonal space dispersion, tone variation, and tone differentiation of lexical tones in IDS vs. ADS (averaged across participants). Error bars represent standard errors. *Note.* f0: fundamental frequency; ERB: Equivalent-rectangular-bandwidth-rate.

positively correlated with pitch variability within and across utterances, but the expansion of tonal space was also positively correlated with intonational mean pitch, as well as pitch variability within and across utterances. In a word, caregivers who exaggerated intonation more in IDS tended to produce lexical tones with both more expanded space and greater variation within categories.

The results first verified tonal space expansion in Cantonese IDS with infants aged 15 months, using both the commonly used measurement of tone triangle area and the new measurement of tonal space dispersion introduced here. More importantly, an increase of intra-talker tone variation within categories in IDS was revealed. It may provide an explanation from the acoustic perspective for the lower identification

https://doi.org/10.1017/S0305000920000707 Published online by Cambridge University Press

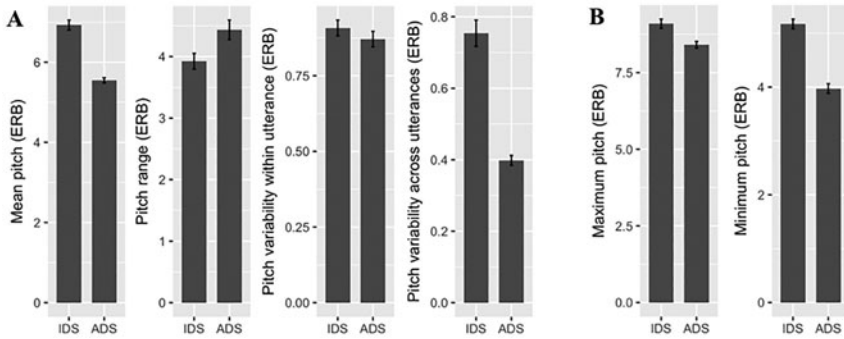


Figure 4. (A) The mean pitch, pitch range, pitch variability within and across utterances of intonation in IDS vs. ADS (averaged across participants); (B) The maximum and minimum pitch of utterances in IDS vs. ADS (averaged across participants). Error bars represent standard errors.
Note. ERB: Equivalent-rectangular-bandwidth-rate.

Table 5. The correlation matrix (uncorrected) between lexical tone and intonational changes in IDS compared to ADS.

		Intonational changes (IDS/ADS)		
		Mean pitch	Pitch variability across utterances	Pitch variability within utterance
Lexical tone changes (IDS/ADS)	**Tonal space dispersion**	$r = .51^{**}$	$r = .79^{***}$	$r = .65^{***}$
	Tone variation	$r = .22$	$r = .76^{***}$	$r = .51^{**}$

Notes. r: Pearson correlation coefficient. ** $p < .01$, *** $p < .001$.

accuracy of lexical tones in Cantonese IDS reported by Wong and Ng (2018). The failure of detecting any significant difference in tone differentiation between IDS and ADS further suggested that the increase of tone variation counteracted the expansion of tonal space, and acoustic contrasts between tone categories may not be enhanced in IDS. Moreover, the strong positive relations between lexical tone and intonational modifications in IDS appeared to be in accordance with the argument that lexical tone adjustments in IDS are byproducts of intonational modifications.

So far, the findings about lexical tone adjustments in Cantonese IDS and their relations with intonational modifications tend to favor PH. Nevertheless, since correlation itself does not necessarily mean causation and cannot determine the direction of the influence, extra efforts are needed to explore the underlying mechanism of the correlations detected. One approach is provided by pet-directed speech which offers a condition in which IDS-like intonational pitch modifications for a communicative function but no hyperarticulation with a linguistic function is expected (Burnham et al., 2002). In this way, any lexical tone adjustments detected in PDS should be driven by intonational effects, while any adjustments not present in this register but uniquely found in IDS should reflect hyperarticulation for didactic purposes.

https://doi.org/10.1017/S0305000920000707 Published online by Cambridge University Press

Experiment 2

In this experiment, Cantonese PDS, specifically dog-directed speech, was recorded and analyzed in the same way as IDS in Experiment 1. If lexical tone modifications found in IDS were also observed in PDS, it was highly likely that these adjustments in IDS were byproducts of intonational effects. On the contrary, if any lexical tone modification in IDS was not observed in PDS, this adjustment in IDS could be attributed to hyperarticulation for didactic purposes. To our knowledge, this was the first time that PDS had been examined in a tone language with regard to lexical tones.

Method

Participants

Seventeen of the 28 caregivers from Experiment 1 agreed to come back for the PDS recording, several months after the IDS session. Five caregivers were dog owners while the others did not have any pets at home (statistical analyses showed no significant difference between these two sub-groups in their PDS production).

Stimuli and materials

To ensure that PDS could be comparable to previously recorded IDS and ADS, the same target words and corresponding toys were used.

Recording procedure

The recording was made in the same sound-attenuated booth, with the same devices and parameters. Out of consideration for safety, pictures of puppies instead of a real dog were provided for the caregivers, presented on an iPad put in front of them. A previous study using pictures of dogs to elicit PDS succeeded in detecting an increase of intonational pitch height and pitch variation similar to what was found in IDS (Ben-Aderet, Gallego-Abenza, Reby & Mathevon, 2017), although it differed slightly from the current study, in that it used fixed sentences, rather than semi-spontaneous speech. The six bags of toys were provided one by one in a random order to elicit the target words. The caregivers were asked to imagine playing with the puppies using the toys and speak to them. As in Experiment 1, they were encouraged to speak the words labelled on the toys. During the recording, there was only the caregiver in the booth. It turned out to be more difficult to elicit target words here than previously in the IDS recording. Thus as with ADS, efforts were made to elicit at least 8 utterances instead for each target word. Again, a very small number of participants, who were still unable to meet this criterion for some target words, produced as many tokens as they could. The experimenter would end the recording of a target word and move on to the next if she judged that the caregiver was clearly tired of the current target word, or the recording of this word had exceeded 5 minutes. On average, about 2.5 minutes were spent on every target word.

Data analysis

The PDS recordings underwent similar pre-processing (see Table 6 for the number of utterances extracted and the number of target words), acoustic measurement, and statistical analysis to the IDS recordings in Experiment 1. Since PDS was compared to both ADS and IDS in this experiment using separate sets of analyses, BH (Benjamini–Hochberg) adjusted p-values were also reported.

https://doi.org/10.1017/S0305000920000707 Published online by Cambridge University Press

Results and discussion

Before focusing on lexical tones, we need to ensure that intonation in PDS had been modified in a similar way to that of IDS. Thus, intonation was analyzed first. As expected, like the IDS findings in Experiment 1, intonation in PDS showed higher mean pitch ($t(16) = 6.4$, $p < .001$, $p_{adj} < .001$, $d = 2.06$) and larger pitch variability across utterances ($t(16) = 4.72$, $p < .001$, $p_{adj} < .001$, $d = 1.52$) compared to ADS; no significant difference was found for pitch variability within utterance ($t(16) = -.26$, $p = .8$, $p_{adj} = .8$, $d = -.07$) (see Figure 5A).

Turning to lexical tones, PDS showed larger tonal space dispersion ($t(16) = 5.14$, $p < .001$, $p_{adj} < .001$, $d = 1.7$) and greater tone variation ($t(16) = 3.54$, $p = .003$, $p_{adj} = .004$, $d = 1.12$) compared to ADS; no significant difference was found for tone differentiation ($t(16) = 1.92$, $p = .07$, $p_{adj} = .19$, $d = .58$) (see Figure 5B). An expansion of tone triangle in PDS is also demonstrated in Figure 5C. All adjustments of lexical tones found in IDS were present in PDS. In addition, in PDS, the same patterns of positive correlations were observed between the changes (PDS/ADS) of lexical tones and intonation as in IDS with comparable correlation coefficients (see Table 7).

Unlike the findings for vowels (Burnham et al., 2002; Xu et al., 2013), with respect to lexical tones, not only was there an increase of tone variation but also an expansion of tonal space in PDS, compared to ADS. This would indicate that these adjustments to lexical tones found in IDS can indeed be driven by intonational effects *per se*. These findings went a step further than the results of correlation tests, providing stronger evidence for PH that lexical tone adjustments in IDS may be unintended byproducts of intonational modifications.

Finally, it is noteworthy that, when directly comparing PDS to IDS produced by the same caregivers, significant differences were observed for both intonation and lexical tones. In terms of intonation, mean pitch ($t(16) = -2.92$, $p = .01$, $p_{adj} = .01$, $d = -.75$), pitch variability within ($t(16) = -3.67$, $p = .002$, $p_{adj} = .006$, $d = -.69$) and across utterances ($t(16) = -2.91$, $p = .01$, $p_{adj} = .01$, $d = -.98$) were lower in PDS than in IDS (see Figure 5A). Correspondingly, with regard to lexical tones, PDS showed smaller tonal space dispersion ($t(16) = -2.79$, $p = .01$, $p_{adj} = .01$, $d = -.91$) and tone variation ($t(16) = -2.75$, $p = .01$, $p_{adj} = .01$, $d = -.78$) than IDS (see Figure 5B).

In view of the positive correlations detected between lexical tone and intonational modifications in PDS, the smaller size of lexical tone adjustments was very likely due to the lower exaggeration of intonation, in line with the postulation of PH that lexical tone adjustments are byproducts of intonational modifications. The lower exaggeration of intonation found in the current experiment, inconsistent with previous findings (Burnham et al., 2002), was highly likely due to our use of pictures instead of real dogs. Without a real dog present, there was no feedback from the addressee, and the caregivers' intention to attract attention and express emotion may have decreased automatically, resulting in less exaggerated intonation than when talking to real pets and infants.

General Discussion

The current study tested the two hypotheses about the function of IDS, HH (arguing for a linguistic function for the acoustic-phonetic modifications in IDS) and PH (arguing for a non-linguistic communicative function for the acoustic-phonetic modifications in IDS), by investigating lexical tones in Cantonese IDS to 15-month-old infants. Although tonal space expansion that is typically considered as evidence for HH was

https://doi.org/10.1017/S0305000920000707 Published online by Cambridge University Press

https://doi.org/10.1017/S0305000920000707 Published online by Cambridge University Press

Table 6. The number of utterances, target words, and target words per utterance for PDS, IDS and ADS analysis.

Tone		Tone1	Tone2	Tone3	Tone4	Tone5	Tone6
Number of utterances	IDS	10 ± 0	9.88 ± 0.32	10 ± 0	9.94 ± 0.24	10 ± 0	10 ± 0
	ADS	9.24 ± 0.88	9 ± 1.37	9.18 ± 1.25	7.53 ± 1.75	8.94 ± 1.26	8.65 ± 1.28
	PDS	9.47 ± 0.61	9.41 ± 0.77	9.82 ± 0.51	8.71 ± 1.13	9.71 ± 0.57	9.65 ± 0.59
Number of target words	IDS	13.88 ± 3.66	12.29 ± 1.96	11.06 ± 0.87	11.35 ± 1.78	14.65 ± 4.13	16.24 ± 4.35
	ADS	10.76 ± 1.93	11.06 ± 2.01	10.59 ± 1.72	8.41 ± 2.47	10 ± 1.61	10.71 ± 2.49
	PDS	11.65 ± 1.41	11.35 ± 2.17	11.35 ± 1.45	9.65 ± 1.75	12.29 ± 3.44	13.41 ± 3.63
Number of target words per utterance	IDS	1.39 ± 0.37	1.25 ± 0.2	1.11 ± 0.09	1.14 ± 0.17	1.46 ± 0.41	1.62 ± 0.43
	ADS	1.17 ± 0.18	1.24 ± 0.19	1.15 ± 0.11	1.11 ± 0.16	1.12 ± 0.12	1.23 ± 0.2
	PDS	1.23 ± 0.12	1.2 ± 0.19	1.16 ± 0.13	1.11 ± 0.14	1.26 ± 0.33	1.39 ± 0.35

Notes. These are all numbers averaged across the 17 caregivers (± standard deviation), who are a subset of the 28 caregivers for IDS and ADS recording.

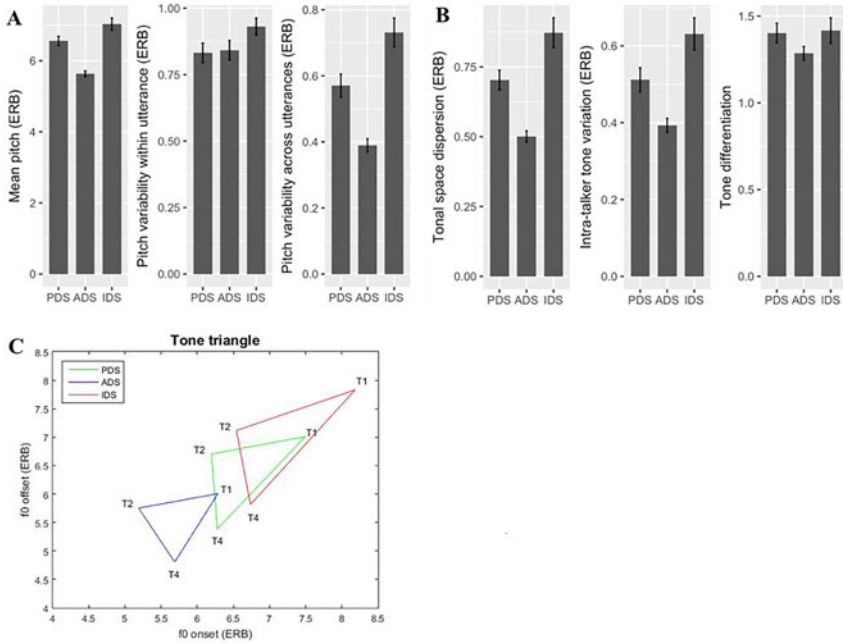


Figure 5. (A) The mean pitch, pitch variability within and across utterances of intonation in PDS, ADS and IDS (averaged across participants); (B) The tonal space dispersion, tone variation, and tone differentiation of lexical tones in PDS, ADS and IDS (averaged across participants); (C) The tone triangle of PDS, ADS and IDS determined by the averaged onset and offset f0 of the three corner lexical tones of Cantonese, i.e., T1, T2, and T4, using the average data of all subjects. Error bars represent standard errors.
Note. f0: fundamental frequency; ERB: Equivalent-rectangular-bandwidth-rate.

Table 7. The correlation matrix (uncorrected) between lexical tone and intonational changes in PDS compared to ADS.

		Intonational changes (PDS/ADS)		
		Mean pitch	Pitch variability across utterances	Pitch variability within utterance
Lexical tone changes (PDS/ADS)	**Tonal space dispersion**	$r = .64^{**}$	$r = .76^{***}$	$r = .53^{*}$
	Tone variation	$r = .32$	$r = .51^{*}$	$r = .62^{**}$

Notes. *r*: Pearson correlation coefficient. * $p < .05$, ** $p < .01$, *** $p < .001$.

observed in IDS compared to ADS, it seemed to be counteracted by the increase of intra-talker tone variation in IDS, leading to no significant enhancement of differentiation among tone categories. In other words, acoustic contrasts of lexical tones appeared not to be enhanced in IDS. More importantly, the tonal space expansion along with the increase of tone variation in IDS could be explained by intonational effects, as evidenced by their strong positive relations with intonational exaggeration in IDS as well as their presence in PDS.

https://doi.org/10.1017/S0305000920000707 Published online by Cambridge University Press

These new findings in general provide consistent evidence for PH, suggesting that lexical tone adjustments in IDS are likely unintended byproducts of intonational modifications, which serve a communicative function. Considering the uniqueness of lexical tones that they are instantiated by changes of pitch, which carry both prosodic and phonemic information, the evidence for PH in lexical tones indicates to some extent that a non-linguistic communicative function may be prioritized over a didactic linguistic function in producing IDS, when the acoustic-phonetic realization of a phoneme needs to consider the two functions. However, the support for PH does not imply a complete rejection of HH (Kalashnikova et al., 2017). As discussed in the introduction, it is possible that a certain lexical tone adjustment in IDS, in particular the tonal space expansion, is a result of both prosodic exaggeration and didactic hyperarticulation. The hyperarticulation-driven modification might get entangled and masked by the similar adjustment caused by intonational effects in the current study.

More importantly, the findings on lexical tones do not generalize to previous evidence on other phonemes like vowels. In particular, with respect to PDS, despite our findings of IDS-like tonal space expansion, most of previous studies found no expansion of vowel space (Burnham et al., 2002; Xu et al., 2013; Gergely et al., 2017). It seems that the prosodic factors alone are sufficient to drive an expansion of tonal space but not vowel space. Such discrepancy further suggests that while tonal space expansion in IDS appears to be largely driven by intonational exaggeration, hyperarticulation for linguistic purposes should play an important role in the expansion of vowel space instead.

The findings on lexical tones in IDS may have implications for research on infant tone acquisition from the perspective of language input. Infant tone acquisition shows a complicated developmental course (see a review: Singh & Fu, 2016). Compared to vowels and consonants, tone categories emerge precociously (Yeung, Chen & Werker, 2013), but they take a longer time to mature and stabilize (Liu & Kager, 2014; Tsao, 2017; Wong, 2013). According to our findings, intonational effects appear to take precedence over the need to maintain and enhance phonemic contrasts in the realization of lexical tones in IDS. Under the intonational effects, lexical tones in IDS may become even less discriminable due to the increase of tone variation. In other words, IDS may provide infants with less informative input for lexical tones, thus making it harder for infants to stabilize mental representations for tone categories. Yet, the puzzle of tone acquisition regarding the precocious emergence of tone categories remains unsolved. With the present findings on the language input, it appears to be necessary to address how infants can decouple lexical tones and intonation in their language acquisition.

PH may provide an alternative and even more satisfactory explanation for the abovementioned age-related reduction of tonal space expansion found in Cantonese IDS, which has been considered as evidence for HH (Xu Rattanasone et al., 2013). HH is weak in explaining why tonal space expansion in IDS reduces for infants by the end of the first year when they continue developing more accurate representations for native lexical tones (Tsao, 2017). By contrast, it has been found in Thai, another tone language, that IDS intonational pitch exaggeration decreases by the end of the infant's first year of life (Kitamura et al., 2002), displaying a developmental pattern quite similar to the observed reduction of tonal space expansion in Cantonese IDS. Thus, it is likely that the tonal space expansion reduces as a result of the decrease of intonational modifications. Future research is needed to

https://doi.org/10.1017/S0305000920000707 Published online by Cambridge University Press

verify such explanation by simultaneously testing longitudinal changes of IDS modifications on intonation and lexical tones in the same tone language.

Despite the evidence for PH in the current study, several efforts can be made to further test the underlying mechanisms for acoustic adjustments of lexical tones in IDS. First of all, as mentioned above, our emphasis on the increase of intra-talker tone variation in IDS as evidence preferring PH over HH is based on one assumption that larger intra-talker variation reduces the learnability of phonemic categories (Cristia & Seidl, 2014; McMurray et al., 2013). According to the NLM model (Kuhl, 1991; Kuhl et al., 2008), infants prefer and are more sensitive to the more prototypical tokens of a phoneme (the exemplars most frequently activated) which serve as referents in phonetic learning. Thus, an increase of variation within phonemic categories would imply that more tokens would deviate from the prototypes increasing overlap across categories, which would be detrimental to infants' phonetic category learning. However, it is also plausible that increased within-category variation does not counteract the positive effects of the acoustic space expansion on early phonetic category acquisition. Theoretically, the Natural Referent Vowel (NRV) framework has been proposed (Polka & Bohn, 2011), which emphasizes the role of peripheral vowels as the referent in speech perception. English speakers were found to favor the more peripheral though less prototypical vowel tokens over those that are less peripheral but more prototypical (Masapollo, Polka, Molnar & Ménard, 2017; Zhao, Masapollo, Polka, Ménard & Kuhl, 2019). From the NRV perspective, the space expansion of phonemes alone might constitute a means of hyperarticulation to facilitate infant language learning by providing more peripheral tokens, irrespective of the increase of variation within categories. However, whether this framework is applicable to lexical tones remains to be seen.

Empirically, as mentioned above, Eaves et al. (2016), based on a formal theory of teaching from pedagogy that has been widely used to capture human learning, generated ideal data for teaching phonetic categories by manipulating formant values of vowels; these optimal teaching data were found to show increases of within-category variation, which was very similar to IDS. It was thus argued that the increase of within-category variation of phonemes in IDS might be beneficial for phonetic learning. However, little research has been conducted to directly compare the learning performance when infants were exposed to speech input with larger vs. smaller within-category phonemic variation. It is worth noting that a group of studies showing advantages of greater variability (e.g., Rost & McMurray, 2009) actually investigated talker variability: that is, to compare infants' language performance with speech input produced by multiple talks vs. a single talker, which is quite different from the intra-talker variation of contrastive cues for phonemic categories discussed in the current study. A study conducted by the same research team later on demonstrates that the multi-talker advantage is not likely to be attributed to the phonetic variability in contrastive acoustic cues, but should benefit from phonologically irrelevant information (Rost & McMurray, 2010). Specifically, when 14-month-old infants were tested in a word-learning task with the minimal pair /buk/-/puk/, they succeeded when trained by stimuli produced by multiple speakers even though the within-category variability of the contrastive cue VOT (voice onset time) was eliminated, but failed when the stimuli were manipulated to have large variability in terms of the contrastive voicing cues but produced by a single speaker. A follow-up study (Galle, Apfelbaum & McMurray, 2015) further found that infants would succeed in this word-learning task in the single-talker

https://doi.org/10.1017/S0305000920000707 Published online by Cambridge University Press

situation when the stimuli contained variability with regard to phonologically irrelevant non-contrastive cues such as pitch and duration. The facilitation effect of phonologically irrelevant variability in speech input was also observed by Singh (2008), who showed that infants performed better in word recognition when provided with speech stimuli containing higher variability of vocal affect (acoustically instantiated by pitch cues). In summary, to better understand the function of lexical tone modifications in IDS, it may be necessary to identify how young tone-language learners' perception of lexical tones is affected by peripherality vs. prototypicality of tone tokens, and whether their phonetic learning benefits from or is hindered by an increase of intra-talker variation within phonemic categories with regard to the contrastive cues.

Second, PDS collected in the current study showed a smaller size of tonal space expansion than IDS. This may well be explained as a result of the less exaggerated intonation in PDS elicited with pictures of pets instead of real pets. Nevertheless, it is still possible that the difference in the degree of tonal space expansion between IDS and PDS is driven by a linguistic function that is unique to IDS. In this case, it would be intonational effects together with a didactic linguistic motivation that lead to tonal space expansion in IDS. To test this possibility, PDS to real pets should be examined, in which intonation is expected to undergo comparable exaggeration to IDS (Burnham et al., 2002). If the difference between PDS and IDS in the size of tonal space expansion becomes no longer significant, this possibility can be excluded, and PH will be further supported.

Last but not least, the current evidence for PH was obtained from IDS to 15-month-old infants only. A possibility remains that for younger infants at an early stage of language acquisition, caregivers may actually produce clearer and more discriminable tone categories in IDS to help the infants' learning of lexical tones, as HH argues, whereas such hyperarticulation becomes no longer necessary for infants by the age of 15 months when they have acquired mental representations for tone categories (Tsao, 2017). Future work is needed to test the two hypotheses with infants at different developmental stages, especially younger infants.

In conclusion, the present study has for the first time confirmed an increase of intra-talker variation of lexical tones in Cantonese IDS, and revealed positive relations between lexical tone adjustments and intonational modifications as well as the similar patterns of lexical tone adjustments in PDS as in IDS. The results demonstrate that acoustic contrasts of lexical tones may not be enhanced in IDS, and the lexical tone adjustments in IDS are likely unintended byproducts of intonational modifications, in line with the prosodic hypothesis. These new findings add important evidence to the debate about the primary function of this register with respect to its acoustic-phonetic modifications.

Acknowledgments. Patrick C.M. Wong is the founder of a technological startup company in Hong Kong supported by a Hong Kong government technological startup scheme for universities; the research reported here has no association with the company. All other authors declare no conflict of interest.

We thank Angela Xiaoxue He for valuable comments and suggestions on drafts of this manuscript. This work was supported by the University Grants Committee (HKSAR) (RGC34000118), the Innovation and Technology Fund (HKSAR) (ITS/067/18), Dr. Stanley Ho Medical Development Foundation, and the Global Parent Child Resource Centre Limited. The second author's work is supported by the Basque Government through the BERC 2018-2021 program and by the Spanish Ministry of Science and Innovation through the Ramon y Cajal Research Fellowship, PID2019-105528GA-I00.

https://doi.org/10.1017/S0305000920000707 Published online by Cambridge University Press

Supplementary Material. For supplementary material accompanying this paper, visit https://doi.org/10.1017/S0305000920000707

References

Adriaans, F., & Swingley, D. (2017). Prosodic exaggeration within infant-directed speech: Consequences for vowel learnability. *The Journal of the Acoustical Society of America*, *141*(5), 3070–3078. doi: 10.1121/1.4982246

Ben-Aderet, T., Gallego-Abenza, M., Reby, D., & Mathevon, N. (2017). Dog-directed speech: why do we use it and do dogs pay attention to it?. *Proceedings of the Royal Society B: Biological Sciences*, *284*(1846), 20162429. doi: 10.1098/rspb.2016.2429

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. doi: 10.1016/j.infbeh.2013.09.001

Benders, T. (2016). Emotion-related explanations of the vowel variability in infant-directed speech. In *Australasian International Conference on Speech Science and Technology* (16*th*: 2016), 233–236.

Benjamin, A., & Slocombe, K. (2018). 'Who'sa good boy?!'Dogs prefer naturalistic dog-directed speech. *Animal cognition*, *21*(3), 353–364. doi: 10.1007/s10071-018-1172-4

Boersma, P., & Weenink, D. J. M. (2017). Praat: Doing Phonetics by Computer [Computer program]. Available at: http://www.praat.org/

Burnham, D., Francis, E., Vollmer-Conna, U., Kitamura, C., Averkiou, V., Olley, A., Nguyen, M., & Paterson, C. (1998). *Are you my little pussy-cat? acoustic, phonetic and affective qualities of infant-and pet-directed speech*. Paper presented at the Fifth International Conference on Spoken Language Processing, Sydney, Australia.

Burnham, D., Kitamura, C., & Vollmer-Conna, U. (2002). What's new, pussycat? On talking to babies and animals. *Science*, *296*(5572), 1435–1435. doi: 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. doi: 10.1177/1606822X14520662

Cristia, A., & Seidl, A. (2014). The hyperarticulation hypothesis of infant-directed speech. *Journal of Child Language*, *41*(4), 913–934. doi: 10.1017/S0305000912000669

Dodane, C., & Al-Tamimi, J. (2007). *An acoustic comparison of vowel systems in adult directed-speech and child-directed speech: evidence from French, English & Japanese*. Paper presented at the 16th International Congress of Phonetics Sciences, Saarbrücken, Germany.

Eaves, B. S. Jr., Feldman, N. H., Griffiths, T. L., & Shafto, P. (2016). Infant-directed speech is consistent with teaching. *Psychological review*, *123*(6), 758. doi: 10.1037/rev0000031

Englund, K. T. (2018). Hypoarticulation in infant-directed speech. *Applied Psycholinguistics*, *39*(1), 67–87. doi:10.1017/S0142716417000480

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. doi: 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*, 1497–1510. doi: 10.2307/1130938

Fernald, A., & Kuhl, P. (1987). Acoustic determinants of infant preference for motherese speech. *Infant behavior and development*, *10*(3), 279–293. doi: 10.1016/0163-6383(87)90017-8

Fernald, A., & Simon, T. (1984). Expanded intonation contours in mothers' speech to newborns. *Developmental psychology*, *20*(1), 104. doi: 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. doi: 10.1017/S0305000900010679

Galle, M. E., Apfelbaum, K. S., & McMurray, B. (2015). The role of single talker acoustic variation in early word learning. *Language Learning and Development*, *11*(1), 66–79. doi: 10.1080/15475441.2014.895249

Gergely, A., Faragó, T., Galambos, Á., & Topál, J. (2017). Differential effects of speech situations on mothers' and fathers' infant-directed and dog-directed speech: An acoustic analysis. *Scientific reports*, *7*(1), 13739. doi: 10.1038/s41598-017-13883-2

Grieser, D. 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. doi: 10.1037/0012-1649.24.1.14

https://doi.org/10.1017/S0305000920000707 Published online by Cambridge University Press

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. doi: 10.3389/fpsyg.2018.00434

Hartman, K. M., Bernstein Ratner, N., & Newman, R. S. (2017). Infant-directed speech (IDS) vowel clarity and child language outcomes. *Journal of child language*, *44*(5), 1140–1162. doi: 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. doi: 10.1121/1.402397

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*, 436–442. doi: 10.2307/1129704

Jeannin, S., Gilbert, C., Amy, M., & Leboucher, G. (2017). Pet-directed speech draws adult dogs' attention more efficiently than Adult-directed speech. *Scientific reports*, *7*(1), 4980. doi: 10.1038/s41598-017-04671-z

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. doi: 10.1017/S0305000917000629

Kalashnikova, M., Carignan, C., & Burnham, D. (2017). The origins of babytalk: smiling, teaching or social convergence?. *Royal Society open science*, *4*(8), 170306. doi: 10.1098/rsos.170306

Kitamura, C., & Burnham, D. (2003). Pitch and communicative intent in mother's speech: Adjustments for age and sex in the first year. *Infancy*, *4*(1), 85–110. doi: 10.1207/S15327078IN0401_5

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. doi: 10.1016/S0163-6383(02)00086-3

Kuhl, P. K. (1991). Human adults and human infants show a "perceptual magnet effect" for the prototypes of speech categories, monkeys do not. *Perception & psychophysics*, *50*(2), 93–107. doi: 10.3758/BF03212211

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. doi: 10.1126/science.277.5326.684

Kuhl, P. K., Conboy, B. T., Coffey-Corina, S., Padden, D., Rivera-Gaxiola, M., & Nelson, T. (2008). Phonetic learning as a pathway to language: new data and native language magnet theory expanded (NLM-e). *Philosophical Transactions of the Royal Society B: Biological Sciences*, *363*(1493), 979–1000. doi: 10.1098/rstb.2007.2154

Lehiste, I., & Peterson, G. E. (1961). Some basic considerations in the analysis of intonation. *The Journal of the Acoustical Society of America*, *33*(4), 419–425. doi: 10.1121/1.1908681

Liu, L., & Kager, R. (2014). Perception of tones by infants learning a non-tone language. *Cognition*, *133*(2), 385–394. doi: 10.1016/j.cognition.2014.06.004

Liu, H. M., Kuhl, P. K., & Tsao, F. M. (2003). An association between mothers' speech clarity and infants' speech discrimination skills. *Developmental Science*, *6*(3), F1-F10. doi: 10.1111/1467-7687.00275

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. doi: 10.1037/0012-1649.43.4.912

Martin, A., Schatz, T., Versteegh, M., Miyazawa, K., Mazuka, R., Dupoux, E., & Cristia, A. (2015). Mothers speak less clearly to infants than to adults: A comprehensive test of the hyperarticulation hypothesis. *Psychological science*, *26*(3), 341–347. doi: 10.1177/0956797614562453

Masapollo, M., Polka, L., Molnar, M., & Ménard, L. (2017). Directional asymmetries reveal a universal bias in adult vowel perception. *The Journal of the Acoustical Society of America*, *141*(4), 2857–2869. doi: 10.1121/1.4981006

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. doi: 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. doi: 10.1016/j.cognition.2017.05.003

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. doi: 10.1017/S0142716400005889

https://doi.org/10.1017/S0305000920000707 Published online by Cambridge University Press

Papoušek, M., Papoušek, H., & Symmes, D. (1991). The meanings of melodies in motherese in tone and stress languages. *Infant behavior and development*, *14*(4), 415–440. doi: 10.1016/0163-6383(91)90031-M

Polka, L., & Bohn, O. S. (2011). Natural Referent Vowel (NRV) framework: An emerging view of early phonetic development. *Journal of Phonetics*, *39*(4), 467–478. doi: 10.1016/j.wocn.2010.08.007

Rost, G. C., & McMurray, B. (2009). Speaker variability augments phonological processing in early word learning. *Developmental science*, *12*(2), 339–349. doi: 10.1111/j.1467-7687.2008.00786.x

Rost, G. C., & McMurray, B. (2010). Finding the signal by adding noise: The role of noncontrastive phonetic variability in early word learning. *Infancy*, *15*(6), 608–635. doi: 10.1111/j.1532-7078.2010.00033.x

Singh, L. (2008). Influences of high and low variability on infant word recognition. *Cognition*, *106*(2), 833–870. doi:10.1016/j.cognition.2007.05.002

Singh, L., & Fu, C. S. (2016). A new view of language development: the acquisition of lexical tone. *Child development*, *87*(3), 834–854. doi: 10.1111/cdev.12512

Soderstrom, M. (2007). Beyond babytalk: Re-evaluating the nature and content of speech input to preverbal infants. *Developmental Review*, *27*(4), 501–532. doi: 10.1016/j.dr.2007.06.002

Song, J. Y., Demuth, K., & Morgan, J. (2010). Effects of the acoustic properties of infant-directed speech on infant word recognition. *The Journal of the Acoustical Society of America*, *128*(1), 389–400. doi: 10.1121/1.3419786

Stern, D. N., Spieker, S., Barnett, R. K., & MacKain, K. (1983). The prosody of maternal speech: Infant age and context related changes. *Journal of child language*, *10*(1), 1–15. doi: 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. doi: 10.1121/1.4995998

Tsao, F. M. (2017). Perceptual improvement of lexical tones in infants: effects of tone language experience. *Frontiers in psychology*, *8*, 558. doi: 10.3389/fpsyg.2017.00558

Uther, M., Knoll, M. A., & Burnham, D. (2007). Do you speak E-NG-LI-SH? A comparison of foreigner-and infant-directed speech. *Speech Communication*, *49*(1), 2–7. doi: 10.1016/j.specom.2006.10.003

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. doi: 10.1016/j.cognition.2006.03.006

Wong, P. (2013). Perceptual evidence for protracted development in monosyllabic Mandarin lexical tone production in preschool children in Taiwan. *The Journal of the Acoustical Society of America*, *133*(1), 434–443. doi: 10.1121/1.4768883

Wong, P., & Chan, H. Y. (2018). Acoustic characteristics of highly distinguishable Cantonese entering and non-entering tones. *The Journal of the Acoustical Society of America*, *143*(2), 765–779. doi: 10.1121/1.5021251

Wong, P. C., & Diehl, R. L. (2003). Perceptual normalization for inter-and intratalker variation in Cantonese level tones. *Journal of Speech, Language, and Hearing Research*. doi: 10.1044/1092-4388(2003/034)

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. doi: 10.1044/2018_JSLHR-S-17-0375

Xu, Y. (2013). ProsodyPro - A Tool for Large-scale Systematic Prosody Analysis. In *Proceedings of Tools and Resources for the Analysis of Speech Prosody* (*TRASP* 2013), 7–10.

Xu, N., & Burnham, D. (2010). *Tone hyperarticulation and intonation in Cantonese infant directed speech*. Paper presented at the Speech Prosody 2010-Fifth International Conference, Chicago, USA.

Xu, N., Burnham, D., Kitamura, C., & Vollmer-Conna, U. (2013) Vowel hyperarticulation in parrot-, dog- and infant-directed speech. *Anthrozoos*, *26*(3), 373–380. doi: 10.2752/175303713X13697429463592

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. doi: 10.1016/j.wocn.2013.06.001

Yeung, H. H., Chen, K. H., & Werker, J. F. (2013). When does native language input affect phonetic perception? The precocious case of lexical tone. *Journal of Memory and Language*, *68*(2), 123–139. doi: 10.1016/j.jml.2012.09.004

https://doi.org/10.1017/S0305000920000707 Published online by Cambridge University Press

Yip, M. (2002). *Tone*. Cambridge University Press. doi: 10.1017/CBO9781139164559

Zhao, Y., & Jurafsky, D. (2009). The effect of lexical frequency and Lombard reflex on tone hyperarticulation. *Journal of Phonetics*, *37*(2), 231–247. doi: 10.1016/j.wocn.2009.03.002

Zhao, T. C., Masapollo, M., Polka, L., Ménard, L., & Kuhl, P. K. (2019). Effects of formant proximity and stimulus prototypicality on the neural discrimination of vowels: Evidence from the auditory frequency-following response. *Brain and language*, *194*, 77–83. doi: 10.1016/j.bandl.2019.05.002

Cite this article: Wang L, Kalashnikova M, Kager R, Lai R, Wong PCM (2021). Lexical and Prosodic Pitch Modifications in Cantonese Infant-directed Speech. *Journal of Child Language* **48**, 1235–1261. https://doi.org/10.1017/S0305000920000707

https://doi.org/10.1017/S0305000920000707 Published online by Cambridge University Press