Tone Hyperarticulation and Intonation in Cantonese Infant Directed Speech

Nan Xu [1], *Denis Burnham*[1]

[1] MARCS Auditory Laboratories, University of Western Sydney, Australia

n.xu@uws.edu.au, d.burnham@uws.edu.au

Abstract

Vowel hyperarticulation in infant-directed speech (IDS) has been found consistently across both tone (Mandarin [1]) and non-tone (Russian, Swedish, American [2] and Australian English [3]) languages and has been posited as a possible bootstrapping mechanism for early language development in infancy [1]. Here we investigated (1) IDS in Cantonese to examine whether tones, like vowels, are hyperarticulated in a tone language and (2) the interaction of F0 measures (mean and range) in tone and intonation. Our results show there is tone hyperarticulation in Cantonese IDS compared to Cantonese adult-directed speech (ADS). Regarding the interaction of tone and intonation, F0 mean was elevated in IDS compared with ADS especially for level tones. F0 range is greater in intonation over utterances than in tones in words, and greater in ADS than IDS. These results suggest that pitch in IDS tone hyperarticulation and IDS intonation is manipulated relatively independently and tone fidelity is not affected by the exaggerated intonation of IDS.

Index Terms: infant-directed speech, prosody, Cantonese tone

1. Introduction

Research in infant-directed speech (IDS) has found consistent suprasegmental and segmental modifications that distinguish it from adult-directed speech (ADS). Studies examining the acoustic characteristics of IDS across a range of languages (including French, Italian, German, Japanese, British, and American English) have found that IDS contains higher F0, wider pitch excursions, shorter utterances, longer pauses, and more prosodic repetitions than ADS [4,6,7]; and that higher pitch is particularly associated with utterance-final positions and focus words [8]. Moreover, factor analysis on adults' ratings of low-pass filtered IDS and ADS reveals higher attentional and affective components in IDS than ADS [9,10]. This suggests that the increased pitch and greater prosodic modulations found in IDS is related to gaining the infant's attention and communicating affective information.

In addition to pitch and prosodic modifications, IDS also exhibits modifications at the segmental level, namely vowel-hyperarticulation — the stretching of articulatory space so that vowel tokens are more separated from each other than they are in ADS [1,2,3,11]. Such vowel hyperarticulation in IDS has been found for American English, Russian and Swedish [2], Australian English [3], and Mandarin [1]. In addition, it appears that vowel hyperarticulation has a didactic influence as there is a positive correlation between vowel hyperarticulation in mothers' IDS and their infants' speech perception, such that mothers who hyperarticulate vowels more have infants who are better able to discriminate native consonant contrasts [1].

Although investigations have revealed vowel hyperarticulation in tone and non-tone languages alike, it is not known if a similar phenomenon occurs for lexical tones in tone language IDS. In tone languages F0 height and contour are used to distinguish among lexical items in a similar way as do consonant and vowel variations. Recently some evidence for segmental level pitch modifications related to tone information has been found in that there are exaggerations of pitch height, F0 range, and duration in the four tones of Mandarin IDS compared to ADS [12]. Nevertheless, a definitive study of tone hyperarticulation in tone language IDS is yet to be conducted.

Similarly, there is also evidence that supersegmental modifications are made in tone IDS as in non-tone languages. Intonation in pitch languages is not all that dissimilar from English and is used to mark boundaries, express emotional information and show sentence-final declination [13]. Grieser and Kuhl (1988) compared Mandarin IDS to English and German IDS, and found that in all cases IDS had significantly higher F0, exhibited larger F0 range over the entire sample, and larger F0 range per phrase than ADS [4]. Similarly Papousek and Hwang (1991) found that when native Mandarin speakers were asked to produce IDS, they increased peak and minimum F0, reduced the rate of F0 fluctuations and, in foreign language instruction, also expanded F0 patterns and F0 range in comparison to ADS [5].

Here we investigate two questions: (1) tone hyperarticulation in Cantonese (2) the interaction of tone and intonation in Cantonese IDS. To investigate tone hyperarticulation Barry and Blamey's (2004) method for plotting Cantonese 'tone space' will be used. By identifying the corner tones (i.e., the three most extreme tones in F0 onset/offset space), plotting them in F0 onset/F0 offset space, and calculating tone triangle areas by joining the centroids of each tone (in an analogous fashion to that used to calculate vowel space), we can compare different speech registers (i.e., IDS and ADS) produced by the same speaker [14]. To investigate IDS intonation and tone, mean F0 and F0 range will be used to measure and compare pitch changes in intonation of utterances and tone of words in IDS and ADS. Speech data were collected from Cantonese mothers speaking to their 6-month-old infants (IDS) and to another adult (ADS). For the first question we expect tone hyperarticulation to be evident in Cantonese IDS just as vowel hyperarticulation has being found in IDS. The second question is much more speculative as it is unknown how tone hyperarticulation, if it is indeed found to occur, might interact with the exaggerated intonation of IDS

2. Methods

2.1. Participants

Eleven mothers were recorded speaking to their 6-month-old infants (IDS) and to a female native Cantonese adult (ADS). All mothers were native Cantonese speakers with Cantonese as their dominant and preferred language and the only language used with their infants.

2.2. Material

Recordings were made in the infants' home using a Sony digital recorder (TCD-D100) with a unidirectional lapel microphone. Sampling rate was 48kHz and recording level fixed (level 7) to accommodate a relatively loud female voice. The resultant recordings were normalised to 90% of amplitude using Cool Edit 2000. Formant and pitch information were extracted using the speech analysis software, Praat.

2.2.1. *Stimuli*

Nine target words were used for eliciting the three corner vowels /a/, /i/ and /u/, and the six Cantonese tones. There are six tones in Cantonese three so-called level tones—high-level (HL), mid-level (ML) and low-level (LL) each with minimal F0 change over time; and three contour tones—high-rising (HR, mid-low rising to high), mid-rising (MR, mid-low rising to mid) and low-falling (LF, mid-low falling to low), that change markedly over time.

To elicit the target words, toys labeled with Chinese logographs were used during IDS recording sessions. The toys were a lion, a plastic number four, a clock, a coaster with a picture of the city of Sydney, various plastic food items, a butterfly, plastic eggs, and a snake with the word 'dead' elicited by the snake toy (see Table 1).

Table 1. *Words and Tones Used*

Tone	HL	HR	ML	LF	MR	LL
Word*	/si/	/sei/	/sei/	/si/	/si/	/sik/
Gloss	LION	DEAD	FOUR	TIME	CITY	FOOD

*Words appropriate for use with infants were chosen. For four of the six tones, words using the vowel [i] were available while for the HR and ML tones words with spectrally similar [ei] were used [14].

2.3. Procedure

IDS recordings were collected when the infants were 6 months old (mean = 186 days). Mothers were instructed to use six toys labeled with logographs of target words on the six Cantonese tones while speaking to their infants (IDS) and with a native Cantonese-speaking adult (ADS).

Recordings were normalized using Cool Edit 2000, and pitch information (i.e., F0 onset, offset, mean and range) was extracted using Praat. For the tone hyperarticulation analysis, only F0 onset and offset were used. For investigation of the relationship between tone and intonation, F0 mean and range were extracted: for tone both from the portion of the target word vowel at maximum amplitude and at 50% drop in amplitude; and for intonation of the utterance in which the target words were spoken.

3. Results

3.1. Tone Hyperarticulation

Tone triangles were created by plotting the F0 onset and offset values for the three corner tones, HL, HR and LF, then joining the centriods of each tone to form a triangle [13]. The triangle areas were transformed into Mels^2 (perceptual units) ahead of a repeated measures ANOVA. Data screening was satisfactory and alpha was set at .05.

The results showed tone triangle areas for IDS were significantly larger than for ADS, F(1, 10) = 6.328, p = .049, η_p^2 = .335, (see Figure 1).

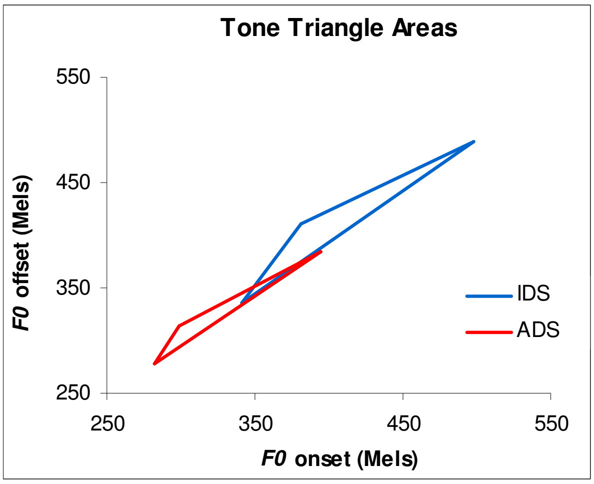

Figure 1: Tone triangle areas of 11 mothers for Cantonese IDS and ADS

3.2. Intonation and Tone

Mean F0 and F0 range were extracted for both the target tone words (see 3.1) and the utterance containing the target tone words. F0 values were transformed into Mels (perceptual units) before performing a repeated measures, speech register (IDS/ADS) x speech unit (tone/intonation) x tones (1-6) ANOVA. Planned comparisons were used to compare the 6 tones as follows: (i) level vs. contour tones;, (ii) within level tones, HL vs. ML and LL, and ML vs. LL; and (iii) within contour tones, falling (LF) vs. rising (HR & LR), and HR vs. LR. Data screening was satisfactory and alpha set at .05, with Bonferoni adjustments made for multiple comparisons. Results are plotted for the six tones for mean F0 in the utterance of the target tone words (Figure 2) and the tone words themselves (Figure 3); and for F0 range for the utterances (Figure 4) and target tones (Figure 5) in IDS and ADS.

3.2.1. *Mean F0 Analysis*

All three main effects were significant. For speech register, mean F0 was significantly higher for IDS than ADS, F(1, 10) = 22.353, p = .001, η_p^2 = .691. For speech unit, mean F0 was significantly higher for tone than intonation in the utterance, F(1, 10) = 10.291, p = .009, η_p^2 = .507. Finally, there was a main effect of tones, F(1, 10) = 19.826, p < .000, η_p^2 = .665), and planned comparisons showed significantly greater mean F0 for level (HL, ML, & LL) than contour (HR, MR & LF) tones, F(1, 10) = 76.13, p < .000, η_p^2 = .884; greater for HL than the lower level tones (ML & LL), F(1, 10) = 17.789, p = .002, η_p^2 = .640; and higher for the two rising tones (HR & MR) than the LF tone, F(1, 10) = 26.462, p < .000, η_p^2 = .726. This suggests that mean F0 differentiates well between tone shape (level and contour) and also between high vs. low tones.

In addition there was a significant register by tones interaction (F(1, 10) = 4.466, p = .002, η_p^2 = .309), a speech unit by tones interaction (F(1, 10) = 24.094, p < .001, η_p^2 = .707) and a three-way interaction between speech register, speech unit and tones (F(1, 10) = 3.430, p = .01, η_p^2 = .255). Together these interactions suggest that the difference in mean F0 between contour vs. level, and high vs. low tones is greater for target words than for utterances, and this difference is more pronounced in IDS than ADS.

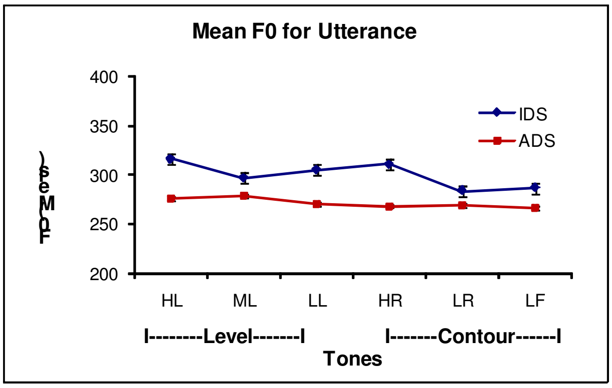

Figure 2. Mean F0 for utterance surrounding the target words in IDS and ADS for each tone

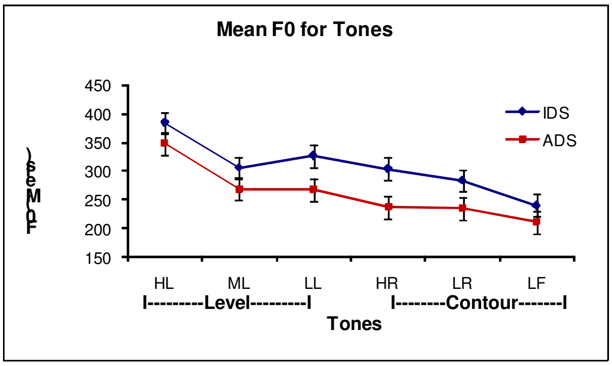

Figure 3. Mean F0 for the vocalic portion of the target word in IDS and ADS for each tone

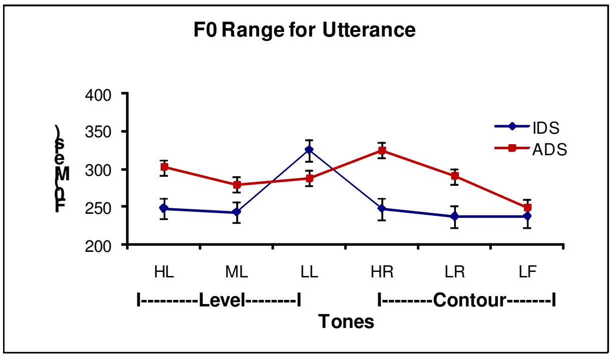

Figure 4. F0 range for utterance surrounding the target words in IDS and ADS for each tone

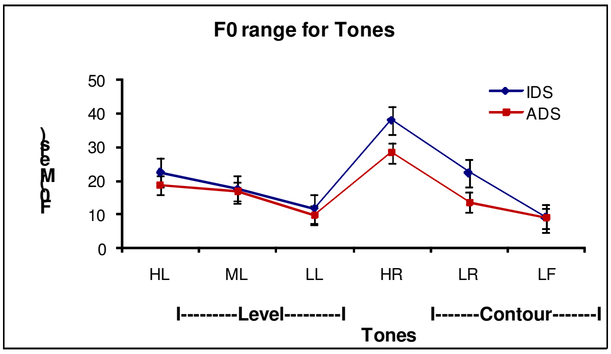

Figure 5. F0 range for utterance surrounding the target words in IDS and ADS for each tone

3.2.2. *F0 Range Analysis*

Two main effects of speech unit and tones were found. For speech units F0 range was significantly greater in utterances surrounding target words than in tone words, F(1.10) = 618.335, p < .000, η_p^2 = .984. For the main effect of tones (F(1.10) = 3.427, p = .01, η_p^2 = .255) planned comparisons showed only one significant result — that F0 range is greater in rising tones (HR & MR) than the LF tone (F(1.10) = 7.444, p = .021, η_p^2 = .427). This suggests F0 range is a good measure for specifying tone contours (rising vs. falling).

In addition, there is a significant speech unit by tones interaction (F(1.10) = 2.714, p = .03, η_p^2 = .213) and a three-way interaction between speech register, speech unit and tones (F(1.10) = 2.836, p = .025, η_p^2 = .221). These interactions show first that F0 range is far greater for the utterances than the tones by a factor of 10. Second they show that the tone words with the greatest range are those with HR tone and this is the case for both IDS and ADS (see Figure 5), whereas for the utterances, while in ADS those surrounding the HR tone also have the greatest range, for IDS it is the utterances surrounding the LL tone that have the greatest range. Over and above these fluctuations, the most striking aspect is the relative ranges in IDS and ADS. For the words F0 range for IDS is just slightly greater than for ADS with a somewhat larger range difference for HR and LR. On the other hand, F0 range is far greater in utterances for most tones with a slight reversal for LL

4. Discussion

The results of the hyperarticulation analyses show that tone triangles are larger for Cantonese IDS than Cantonese ADS, so it can be concluded that there is indeed tone hyperarticulation in Cantonese IDS. This is consistent with previous studies that have found vowel hyperarticulation in IDS [1,2,3] and suggests that exaggerated tones as well as vowels in IDS may bootstrap early language development.

The results of the intonation and tone analysis show firstly that speech registers (IDS and ADS) are differentiated by their relative mean levels of F0 both in words and utterances. In IDS mean F0 is much higher than in ADS, in concert with previous findings. The specific patterns of greater mean F0 between level and contour tones, and within level and contour tones are particularly interesting for IDS research. It appears to be the case that mothers use mean F0 to emphasize the difference between level (HL, ML, LL) and contour (HR, MR, LF) tones, and between high vs. low tones in both level (HL vs. ML & LL) and contour (HR & MR vs. LF) tones.

F0 range, on the other hand, appears to differentiate tone and intonation — F0 range is much wider for intonation units than for tone units. Given this, there are also interesting differences between IDS and ADS in intonation vs. tone units. For intonation, F0 range is consistently greater in ADS than IDS, except for LL; for lexical tone, F0 range in IDS and ADS are quite similar although IDS range is always just slightly greater in IDS than ADS for most tones and notably greater for HR and LR. It appears that mothers may particularly emphasis the pitch characteristics of these two tones in IDS. It is possible that mothers may use F0 range to emphasise differences in tone contour between rising and falling tones, but mean F0 to emphasise differences in pitch level (i.e., high from low level tones).

Therefore in tone language IDS, there appears to be less interplay in the use of F0 information for specifying intonation versus lexical tone than in ADS. A possible explanation of these results is that in ADS (normal speech) F0 range can be compromised in the production of tones, for adults have robust lexical entries such that veridical word identification can occur despite narrowing of F0 range and under-specification of tone identity. On the other hand since infants have less well-established lexical entries than adults mothers may strive to preserve the integrity of tone information in IDS by extending F0 range and emphasising tone identity. This emphasis of tone identity is heightened even further if the overall range of the utterance is narrowed.

5. Conclusions

In Cantonese IDS, tone is hyperarticulated compared to ADS just as vowels are in other languages. In addition, exaggerations in intonation in Cantonese IDS appear to be driven by greater differences in mean F0 between level versus contour tones, and high versus low tones, while F0 range appears to be important for specifying contour information in rising and falling tones. In Cantonese there is higher mean F0 in IDS than ADS in both individual words and utterances. However, while pitch range for words is always slightly greater for IDS than ADS, for utterances pitch range is narrower in IDS than ADS. The combined effect of these modifications is to emphasise F0 height and movement. In the latter case, pitch movement, F0 range in tones in words and intonation across utterances interact in IDS to differentiate tones and emphasise tone identity, input that is extremely important for building up lexical representations which, in tone languages, depend upon pitch information. Once strong lexical representations are set up (in adulthood) pitch range in words can be under-specified with little or no loss of meaning.

6. Acknowledgements

We thank the ethnic community health nurses for recruiting mothers and infants.

7. References

[1] 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.

[2] Kuhl, P. K., Andruski, J. E., Chistovich, I. A., Chistovich, L. A., Kozhevnikova, E. V., Ryskina, V. L., et al. (1997). Cross-language analysis of phonetic units in language addressed to infants. *Science, 277*(5326), 684-686.

[3] Burnham, D., Kitamura, C., & Vollmer-Conna, U. (2002). What's new, pussycat? On talking to babies and animals. *Science, 296*(5572), 1435.

[4] Grieser, D. L., & Kuhl, P. K. (1988). Maternal Speech to Infants in a Tonal Language: Support for Universal Prosodic Features in Motherese. *Source: Developmental Psychology, 24*(1), 14-20.

[5] Papousek, 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.Andruski, J. E., Kuhl, P. K., & Hayashi, A. (1999). Point vowels in Japanese mothers' speech to infants and adults. *The Journal of the Acoustical Society of America, 105*, 1095.

[6] Fernald, A., & Simon, T. (1984). Expanded intonation contours in mothers' speech to newborns. *Developmental Psychology, 20*(1), 104-113.

[7] Fernald, A., Taeschner, T., Dunn, J., Papousek, M., 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*, 477-501.

[8] Fernald, A., & Mazzie, C. (1991). Prosody and focus in speech to infants and adults. *Developmental Psychology, 27*(2), 209-221.

[9] Kitamura, C., & Burnham, D. (1998). The infant's response to vocal affect in maternal speech. In C. Rovee-Collier (Ed.), *Advances in infancy research* (Vol. 12, pp. 221-236).

[10] 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.

[11] Rvachew, S., Mattock, K., Polka, L., & Menard, L. (2006). Developmental and cross-linguistic variation in the infant vowel space: The case of Canadian English and Canadian French. *Journal of the Acoustical Society of America, 120*, 2250-2259.

[12] 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.

[13] Ladefoged, P. (2001a). A course in Phonetics (4th ed.) (Fourth Edition. ed.). Boston: Heinle & Heinle, a division of Thomson Learning, Inc.

[14] Barry, J. G., & Blamey, P. J. (2004). The acoustic analysis of tone differentiation as a means for assessing tone production in speakers of Cantonese. *Acoustical Society of America, 116*(3), 1739–1748.

[15] Zee, E. (1999). *An acoustical analysis of the diphthongs in Cantonese.* Paper presented at the Conference proceedings from ICPhS, San Francisco, USA.