Auditory-Visual Infant Directed Speech in Japanese and English

Takaaki Shochi [1] *Kaoru Sekiyama* [1], *Nicole Lees*[2], *Mark Boyce*[2], *Roland Göcke*[3], *Denis Burnham*[2]

[1] Division of Cognitive Psychology, Kumamoto University, Japan
[2] MARCS, University of Western Sydney, Australia
[3] HCC Lab, Faculty of Information Sciences & Engineering, University of Canberra, Australia

shochi38@gmail.com, sekiyama@kumamoto-u.ac.jp, {n.lees, m.boyce, d.burnham}@uws.edu.au, roland.goecke@ieee.org

Abstract

The aim of this project is to compare (i) the acoustic vs. visual characteristics of infant-directed speech (IDS), (ii) IDS vs adult-directed speech (ADS), and (iii) the acoustic/visual characteristics of IDS and ADS cross-linguistically, in Australian English (AusE) vs Japanese. Acoustic data are presented along with preliminary visual data. Native AusE and Japanese speaking mothers spoke to their 4- to 9-month-old infants and another adult using target words containing one of four vowels [a,i,u,o]. Results show higher F_0 mean and greater F_0 variation for IDS than ADS in both language groups, and longer vowel duration in IDS than ADS but this was only significantly so for the AusE mothers. Finally, there was a tendency for vowel *hyper*-articulation in AusE mothers' IDS, but for vowel *hypo*-articulation in Japanese mothers' IDS, and overall vowel hyperarticulation was greater in AusE than Japanese IDS. Preliminary visual data suggest that there appears to be, contrary to what would be expected, a substantial *decrement* in visual vowel lip area in Japanese IDS compared to ADS, a finding that is in concert with tendency for vowel hypo-articulation in Japanese than English IDS.

Index Terms: Infant Directed Speech, Adult Directed Speech, hyper articulation, cross cultural comparison

1. Introduction

Compared with ADS, IDS has (i) higher mean fundamental frequency (F_0) [1,2] (ii) greater variation of F_0 range [3,4] (iii) longer syllable duration [5], (iv) greater rated affect in the voice [6], and (v) hyperarticulation of vowels [1,7] and hyperarticulation of lexical tones [8]. There is evidence that these aspects of IDS are universal or at least widespread [7,9,10,11], and both English [12,13,14] and Japanese [15] language infants have been found to prefer IDS to ADS. In addition, the hyperarticulation of vowels in IDS appears to serve a linguistic purpose – mothers who show greater vowel [10] and tone [8] hyperarticulation have infants with better speech perception abilities.

The focus of this experiment is the production of vowels in IDS in English and Japanese. Japanese has just 5 vowels [a,e,i,o,u] but each can occur in two phonemically distinct forms - short (1 mora) or long (2 mora) forms). On the other hand, English has 12-14 vowels which are all distinguished by vowel colour. In a cross-language study Werker and her colleagues [16] recorded IDS from Canadian English and Japanese mothers and analysed data from two sets of vowel pairs common to the two languages ([ɪ]-[i] and [ɛ]-[e]). Their results reflect those expected from the language-specific distributional properties of the two languages: dichotomous category membership for each of the vowel pairs was better predicted by the duration than the colour of the vowels produced by Japanese mothers; and better predicted by the colour than the duration of the vowels produced by Canadian English mothers. This is a striking demonstration of the differential distribution of phonetic features based on the phonology of particular languages. Unfortunately no ADS was collected in their study so no conclusions about hyperarticulation can be drawn. Nevertheless, a working hypothesis here might be, extrapolating from their results, that hyperarticulation should be more evident in vowel lengthening in Japanese, and more evident in vowel colour in AusE.

There is little if any work on visual hyperarticulation of speech [17], especially in IDS. With respect to visual speech *per se*, there are there are differences across English and Japanese in the degree that visual and auditory speech information are used in speech perception. Japanese language perceivers use less visual speech information (from the lips and face) in a McGurk effect task than do English language perceivers [18,19,20], and pay less attention to mouth movements when observing infant vocalisations [21]. As there is hyperarticulation of vowels and tones in acoustic speech[10,8], and as acoustic and optic speech are manifestations of the same underlying articulatory gestures, it is hypothesised here that there will indeed be hyperarticulation of visual, lips and face, aspects of IDS compared to ADS; and, given the oblique evidence from McGurk studies, that there would be greater (or at least different) visual hyperarticulation in AusE than Japanese IDS,

Here the acoustic properties (mean F_0, range F_0, F_1, F_2, duration) and preliminary data on the visual properties (mean lip width and height) of (Sydney) AusE and (Kumamoto) Japanese IDS and ADS are compared, in particular the characteristics of four vowels, [a,i,u and o]. Previous studies of hyperarticulation have usually used just three vowels, [a,i,u]. However, [u] is more centred in Japanese than in English, and the Tokyo Japanese dialect has no rounding on [u]. (It is difficult to find data on the Kyushu dialect, that of the Kumamoto participants in this study, but we observed similar articulatory characteristics in our potential participants.) For these reasons in this study we added [o] to the usual repertoire of vowels in IDS studies.

2. Method

2.1. Participants

Eleven native Japanese speakers from Kumamoto region, Japan, and 10 AusE native speakers from Sydney, all mothers of infants between 4 and 9 months, participated. Japanese mothers were recruited by newspaper advertisement and AusE mothers from the MARCS Labs Baby Register.

2.2. Procedure

Four teddy bears named [baba], [bibi], [bubu], and [bobo] were used to examine the four vowels of interest [a,i,u,o]. In the IDS session mothers sat on a chair in front of their baby in a baby chair (see Figure 1), and told them a short story or fictitious character details about each teddy bear. They were instructed to say each bear's name at least 12 times in the session. In the ADS session mothers were instructed to talk about four people with whom they had a close relationship, replacing their names with the target words ([baba], [bibi], [bubu], [bobo]). Sessions were recorded by digital camera (JVC-Everio GZ-MG505). Audio information was digitally recorded at sampling rate 44.1kHz by clip microphone (Sennheiser ew 112-p G2 Wireless Lapel Microphone) attached near the throat of mothers.

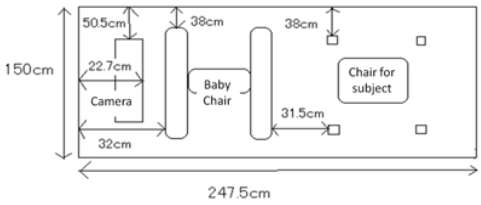

Figure 1: Experimental set-up for mother and baby including recording materials

3. Auditory Speech Results

3.1. Acoustic aspects of IDS and ADS

3.1.1. Statistical analysis

From utterances produced by mothers in IDS and ADS, three acoustic values: F_0 mean, F_0 range (F_0 *max.* - F_0 *min.*), and duration of each vowel (1st or 2nd in each of the [bVbV] strings) were extracted using Praat ver.5.1. The data of one AusE and one Japanese mother were not included, because of difficulties in extraction of acoustic values from their vowels, leaving data from 10 Japanese and 9 AusE mothers. Three analyses of variance (ANOVAs) were conducted for each of the three dependent variables: the first two analyses were separate ANOVAs for the AusE and Japanese mothers, each designed with 2 experimental conditions (IDS/ADS), 2 vowel positions (1st/2nd), and 4 vowel types ([a, i,u, o]), with repeated measures on each factor. The third ANOVA was conducted to compare Japanese and AusE mothers. As in this analysis it was hyperarticulation, i.e., the relative difference between IDS and ADS that was of interest, and due to possible intrinsic phonological factors (e.g., rhythm, accent etc.) which may result in absolute differences between acoustic measures of the two languages, for this third ANOVA scores were normalised by dividing each mother's IDS values by her ADS values. This gives a hyperarticulation score in which unity indicates equal levels in IDS and ADS, scores > 1 indicate hyperarticulation for the dependent variable in question, and scores < 1 indicate hypoarticulation for the dependent variable in question.

3.1.2. F_0 mean

For the within-language ANOVAs of F_0 mean, there was a main effect of speech register in both language groups, F_{AusE} (1, 24) = 31.34, $p < .001$; F_{Jpse} (1, 27) = 21.62, $p < .005$ (see Figure 1), showing that pitch is generally higher in IDS than in ADS. The main effect of vowel type was significant for both language groups, F_{Jpse} (4, 27) = 5.73, $p < .005$; F_{AusE} (4, 24) = 3.28, $p < .05$, but there was no interaction with other factors. The main effect of vowel position was significant only for Japanese mothers, F (1, 27) = 39.36, $p < .0001$, showing a higher F_0 mean in the 1st compared to the 2nd vowel.

The Japanese vs AusE ANOVA for mean F_0 revealed no main effect of language, F (1, 17) = 3.45, $p > .05$, vowel type, F (3, 51) = 1.00, $p > .05$, nor vowel position, F (1, 17) = 0.04, $p > .05$, but the language x vowel position interaction was, F (1, 17) = 11.42, $p < .05$, showing that the overall heightening of mean F_0 in IDS compared to ADS was accentuated in the 1st vowel for AusE mothers, F (1, 34) = 7.76, $p < .01$.

3.1.3. F_0 range

There was greater F_0 variation in IDS than ADS in both languages, F_{AusE} (1, 24) = 31.80, $p < .005$; F_{Jpse} (1, 27) = 13.84, $p < .01$. There was no main effect of vowel type for either language, F_{Jpse} (4, 27) = 1.73, $p > .05$; F_{AusE} (4, 24) = 0.49, $p > .05$. There was a main effect of vowel position but only for AusE mothers, F (1, 24) = 24.84, $p < 0.01$, indicating greater pitch variation for the 2nd vowel compared with the 1st.

For Japanese vs AusE there was greater pitch variation in IDS for AusE than Japanese mothers, F (1, 17) = 6.19, $p < .05$, and generally greater pitch variation in IDS in the 1st than the 2nd syllable, F (1, 17) = 12.99, $p < .005$, though the interaction between the language and vowel position was not significant.

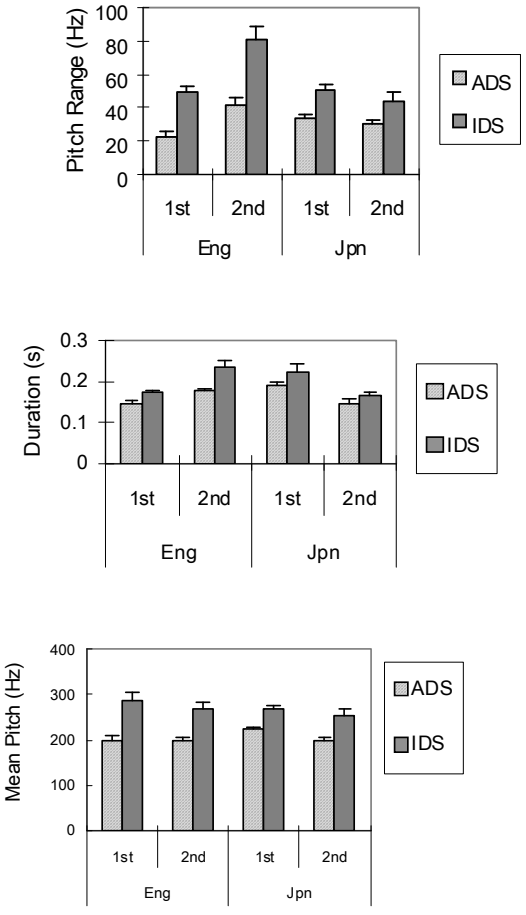

Figure 2: Mean F_0, F_0 range & duration (upper, central, lower) for Japanese and AusE mothers' 1st & 2nd vowels in ADS & IDS. Bars indicate standard errors.

The ANOVA comparing Japanese and AusE mothers' F_0 mean showed a main effect of language, F (1, 17) = 6.19, p < .05, as well as vowel position, F (1, 17) = 12.99, p < .005, showing greater pitch variation for AusE than Japanese mothers. There was no effect of vowel type, F (3, 51) = 0.74, p > .05, and the interaction between the language factor and vowel position was not significant.

3.1.4. Duration

For duration there was a main effect of speech register for AusE mothers, F (1, 24) = 12.67, p < .01, who produced longer vowels in IDS than in ADS. The result was in the same direction for Japanese mothers but failed to reach significance, F (1, 27) = 3.42, p < .098. The main effect of vowel type was significant for AusE but not Japanese mothers, F_{AusE} (4, 24) = 4.18, p < .05; F_{Jpse} (4, 27) = 0.58, p > .05, but there was no interaction with other factors. There was a main effect of vowel position for each language group, F_{AusE} (1, 24) = 16.57, p < .05; F_{Jpse} (1, 27) = 39.36, p < .001, but in opposite directions, suggesting an overall stress difference - Japanese mothers showing significantly longer duration for the 1st syllable vowel, and AusE mothers for the vowel in the 2nd syllable. There was also a tendency for these syallables (1st for Japanese, 2nd for AusE) to be the ones that were hyperarticulated in each language, but neither of the relevant interactions were significant.

The Japanese vs AusE ANOVA for vowel duration showed no main effects or interactions for any of the factors, language, F (1, 17) = 0.35, p > .05, vowel position, F (1, 17) = 3.09, p > .05, or vowel type, F (3, 51) = 1.09, p > .05.

3.2. Articulatory aspects of IDS and ADS

3.2.1. Vowel triangle and quadrilateral area derivation

To investigate the articulation of ADS and IDS in each language group, 1st formant (F_1) and 2nd formant (F_2) were measured at the temporal midpoint of the target vowels. The mean value of each subject's samples was calculated within the same category of vowel and separately for vowel position (i.e. 1st or 2nd vowel). Formant values from the 10 Japanese subjects and 9 English Australian subjects were used to create both vowel triangles (based on the three vowels [a,i,u,] as in other studies of vowel hyperarticulation [1,7,9,10] and also vowel quadrilaterals based on the four vowels used here [a,i,u,o]. Methods to calculate vowel triangles have already been documented [1,7], but not so for vowel quadrilaterals. To calculate the vowel quadrilaterals here, Bretschneider's formula [22] was used in which the area of a general quadrilateral with sides of length a, b, c, and d is given by: ¼ $\sqrt{(4p^2q^2 - (b^2 + d^2 - a^2 - c^2)^2}$, where p and q = the lengths of the diagonals.

3.2.2. Articulation in IDS and ADS

Vowel Triangles: Figures 3 and 4 illustrate vowel triangles based on the mean F_1 and F_2 values of the three vowels: [a, i, u,] for Japanese and AusE mothers respectively. As for the acoustic variables three ANOVAs were conducted, one each on each language separately and one comparing the languages. First, individual participants' triangle areas were subjected to an ANOVA with two within-subjects factors: speech register (IDS / ADS) and vowel position (1st,/ 2nd). The results showed no hyperarticulation for IDS, either for the Japanese, F (1, 9) =

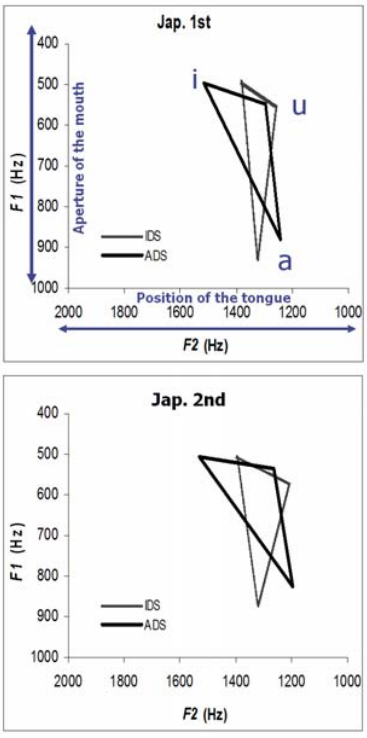

Figure 3: Triangles of mean F_1 and F_2 values for three vowels , [u], [i]) in Japanese mothers' IDS and ADS (1st vowel, upper panel; 2nd vowel lower panel).

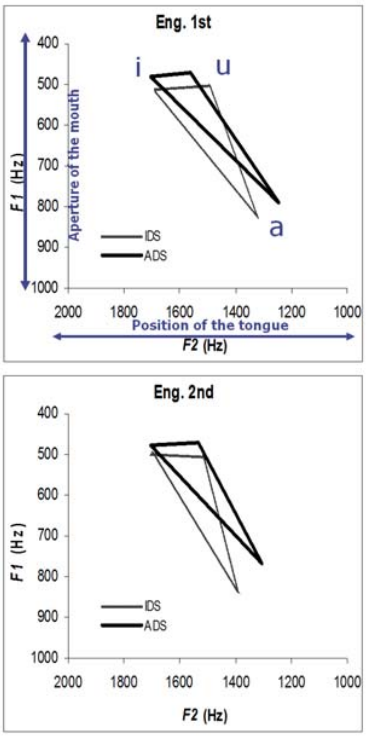

Figure 4: Triangles of mean F_1 and F_2 values for three vowels ([a], [u], [i]) in AusE mothers' IDS and ADS (1st vowel, upper panel; 2nd vowel lower panel).

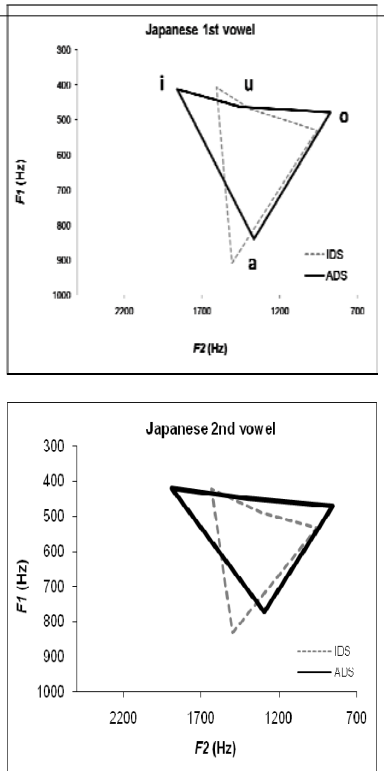

Figure 5: Quadrilaterals of mean F_1 and F_2 values for four vowels ([a], [i], [u], [o]) in AusE mothers' IDS and ADS (1st vowel, upper panel; 2nd vowel lower panel).

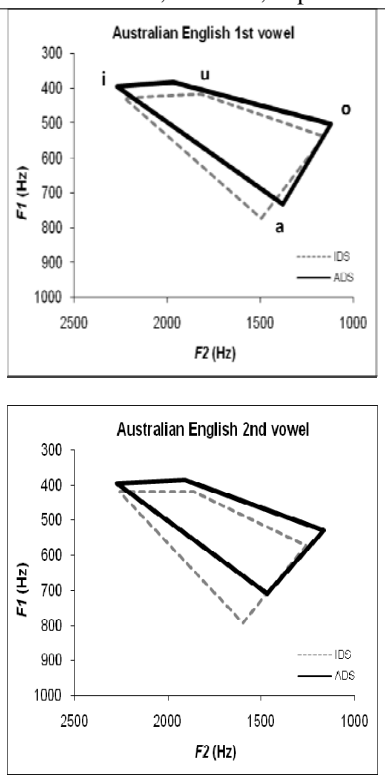

Figure 6: Quadrilaterals of mean F_1 and F_2 values for four vowels ([a], [i], [u], [o]) in AusE mothers' IDS and ADS (1st vowel, upper panel; 2nd vowel lower panel).

1.52, p > .05, or the AusE mothers, F (1, 8) = 0.86, p > .05, and no main effect of vowel position for the Japanese group, F (1, 9) = 2.40, p > .05; or for the AusE group, F (1, 8) = 0.73, p > .05.

To compare the level of vowel hyperarticulation in Japanese and AusE mothers an ANOVA was conducted with the IDS/ADS ratio for each mother in each language group as the dependent variable. A language group (Japanese/AusE) x vowel position (1st/2nd) ANOVA revealed neither a main effect of language (F (1, 17) = 0.001, p > .05) nor vowel position (F (1, 17) = 0.55, p > .05).

Vowel Quadrilaterals: Figures 5 and 6 illustrate vowel quadrilaterals based on the mean F_1 and F_2 values of the four vowels [i, u, a, o] for the Japanese and AusE mothers respectively. ANOVAs for each language with two within-subjects factors: experimental condition factor (IDS / ADS) and vowel position (1st,/ 2nd) showed a close to significant effect for IDS vs ADS for the Japanese mothers, F (1, 9) = 4.32, p > .05, but in the *opposite* direction than expected, i.e., there was vowel *hypo*articulation in Japanese IDS compared with Japanese ADS. On the other hand, for the AusE mothers there was a slight degree of hyperarticulation, but this was not significant, F (1, 8) = 2.31, p > .05. There was no significant effect of vowel position for either the Japanese, F (1, 9) = 1.98, p > .05, or the AusE vowels, F (1, 8) = 0.13, p > .05.

The ANOVA to compare the level of vowel hyperarticulation (the IDS/ADS ratio) in Japanese and AusE mothers revealed significantly greater vowel hyperarticulation in AusE than in Japanese, F (1, 17) = 7.40, p < .05. The effect of vowel position was not significant, F (1, 17) = 0.05, p > .05.

3.2.3. Summary of Acoustic and Articulation Measures

Table 1 summarises the acoustic and articulation measures in Japanese and English IDS. As can be seen AusE and Japanese IDS are similar in terms of the exaggeration of acoustic variables, mean F_0, range F_0, and duration, but not vowel articulation. Vowel hyperarticulation is greater in AusE than Japanese IDS, and if anything there is hypo-articulation in Japanese IDS compared with hyper-articulation in AusE IDS. With respect to the hypotheses put forward earlier, there is indeed greater hyperarticulation in AusE than Japanese IDS, and in addition a tendency for hypo-articulation in Japanese IDS has been uncovered. These results are in accord with the hypothesis of greater exaggeration of vowel colour in AusE than Japanese IDS, but not with the hypothesis that Japanese IDS would be more evident in vowel lengthening.

Table 1: IDS within & between English & Japanese mothers

Measure	**AusE IDS**	**Japse IDS**	**Eng vs Jpse**
Mean F_0	Height'd	Height'd	AusE=Jpse
Range F_0	Height'd	Height'd	AusE>Jpse
Duration	Sig longer	Longer	AusE=Jpse
Vow Tri'le	IDS=ADS	IDS=ADS	AusE=Jpse
Vow Quad'l	HYPERart'n	HYPOart'n	AusE>Jpse

4. Visual Speech Results

Using the 3D Lip Tracking software [23] the following measurements were made for each mother in each condition: lip width, lip height, lip protrusion (i.e., distances in x, y, z aspects of three-dimensional space). Lip width measured the distance between lip corners, lip height the distance between

inner contours of the lower and upper lip midpoints, and lip protrusion the distance from a conceptual plane behind the mouth. In contrast to the auditory analysis, only the first vowel in each[/baba], [bibi], [bobo], [bubu] utterance was analysed.

For initial comparison purposes, the lip width and lip height values were derived and plotted for one Japanese (J01) and one AusE (E02) mother's IDS and ADS. These are presented in Figure 7 in order to compare the mouth area over vowels in ADS and IDS (in mimicry of the FI/F2 space presented in Figures 3, 4, 5 and 6). As can be seen there appears to be no hyperarticulation of mouth area in IDS vs. ADS for either Japanese or AusE. However it may be noted that (a) there is rotation of the vowel quadrilateral between ADS and IDS in both languages, which can be described as $+90^0$ rotation; (b) while this rotation is generally the case, for AusE there is a non-transitive shift of the position of the [u] vowel to move from between [o] and [i] and opposite [a] in ADS to between [a] and [i] and opposite [o] in IDS such that in AusE IDS [u] has similar lip height to [a] and [i]; and (c) there appears to be, contrary to what would be expected, a substantial *decrement* in visual vowel lip area in Japanese IDS compared to ADS.

Inspection of Figure 7 reveals that for AusE IDS, despite no apparent overall change in lip height by width area, visual feature hyperarticulation *is* indicated, as there is, as might be expected, increased lip width for [i] and increased lip height for [a] in IDS. In addition, there is increased lip width for both [u] and [o] in IDS vs ADS. For Japanese IDS however there is the opposite, *decreased* height for [o] and [u], no change in height for [a] and [i] and decreased width for all the vowels except [o]. These results seem to suggest a different visual style in Japanese IDS, perhaps greater protrusion and/or more whispered speech.

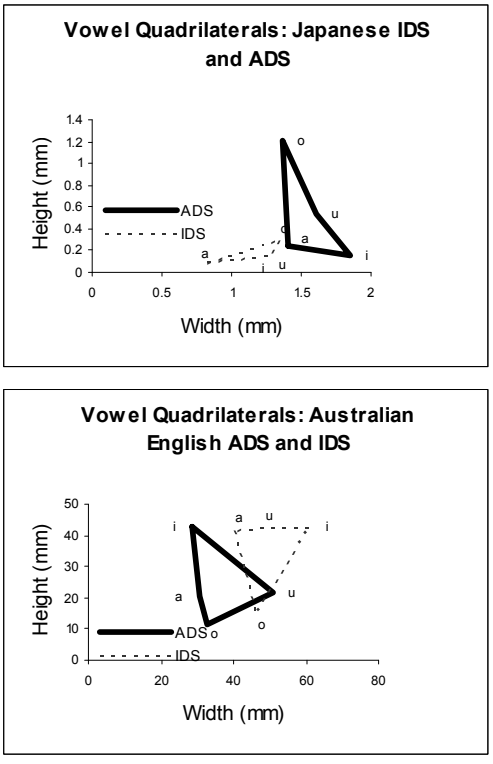

Figure 7: Quadrilaterals of mean lip height and width for four vowels ([a, i, o, u]) in Japanese (top panel) and Australian (lower panel) mothers' IDS and ADS

5. Discussion and Conclusions

In concert with previous findings, e.g., [1,2,7,9], there was greater pitch variation and higher mean F_0 in IDS compared to ADS in both AusE and Japanese. Lengthening of vowel duration was also evident in both languages, but only significantly so for AusE mothers. The lack of significant lengthening of vowels in Japanese IDS may be due to the stricter control of vowel length imposed by the moraic structure of Japanese and its use of vowel length as a phonemic discriminator, compared with the stress-timed nature of English. Indeed, the working hypothesis put forward earlier that there should be more vowel lengthening in Japanese could be modified if it is considered that the presence of vowel length as a phonemic marker in Japanese actually militates against an overall lengthening of vowels. Further research is required to investigate this more refined hypothesis that 2-mora vowels are lengthened, whereas 1-mora vowels are shortened in Japanese IDS to accentuate vowel duration.

In vowel articulation, there was surprisingly little hyperarticulation in IDS in this study, either when vowel triangles or vowel quadrilaterals are considered. This may have been due to the peculiarities of the particular test procedure we used here in which the experimenters and assistants often needed to interact with mothers to ensure target words were uttered, which appeared to make the IDS less spontaneous than otherwise might have been the case. In addition, the 2-syllable target words may have led to any hyperarticulation of vowels being masked by the stress pattern of the words — greater increase in mean F_0 in IDS on the first syllable for AusE and on the second syllable for Japanese.

Over and above these constraints, it is interesting to note that vowel quadrilaterals appear to be more sensitive to variations in vowel space as a function of speech register than are vowel triangles. Of even greater interest is that, in addition to support for the hypothesis that there would be greater vowel hyperarticulation for AusE than Japanese IDS, there was in fact slight *hypo*articulation of vowels by Japanese mothers.

With respect to visual speech, although mouth area (lip height x width) is about equal for AusE IDS and ADS there is, nevertheless hyperarticulation of lip movements in expected directions for AusE IDS – increases in lip height for [a] and lip width for [i]. This shows that while the measure of visual hyperarticulation employed here may have some value, it could in future be augmented in some way to reflect, numerically, the rotations of the quadrilaterals in space. Turning to Japanese, the lip height x width mouth area for vowels is substantially *less* in IDS than in ADS. The nature of the IDS → ADS changes plus anecdotal evidence suggests that it is possible that there is a change in voice *quality* in Japanese IDS, possibly involving greater lip protrusion and/or greater use of whispered speech. Further analyses in which a protrusion measure is included, and in which more mothers' data are included should clarify this issue.

The cross-language articulation evidence so far in the visual realm (hypo-articulation for Japanese and hyper-articulation for AusE) parallels the cross-language articulation evidence in the auditory realm (hypo-articulation for Japanese and hyper-articulation for AusE). Thus we have here preliminary evidence for concordance of changes between ADS and IDS in the auditory and the visual manifestations of articulation. Of course, further visual data are required and auditory vs. visual correlations must be conducted before definitive conclusions can be drawn, but these results are promising and point to a method by which auditory and visual comparisons of IDS and ADS may be made.

6. Acknowledgements

We thank Dr Roland Göcke for his 3D Lip Tracking software and his persistent and effective assistance. We thank Ms. Motoko Takao for her contribution to the auditory aspects, and Dr Kimiko Tsukada for her contribution in an earlier stage of the study. This work was supported by an Australian Research Council Discovery grant (DP0558698) to the 6th and 2nd authors, and a Grant-in-Aid for Scientific Research from Japan Society for the Promotion of Science (21243040) to the 2nd author.

7. References

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

[2] Kitamura, C. and Burnham, D. (2003). Pitch and communicative intent in mother's speech: Adjustment for age and sex in the first year. *Infancy, 41*(1), 85-110.

[3] Fernald, A. (1989). Intonation and communicative intent in mothers' speech to infants: Is the melody a message? *Child Development, 60,* 1497-1519.

[4] Fernald, A. (1992). Meaningful melodies in mothers' speech. In H. Papousek, U. Jurgens, M. Papousek (Eds.), *Nonverbal communication. Comparative and developmental approaches.* Paris: Cambridge University Press. pp 262-282.

[5] Stern, D., Spieker, S., Barnett, R.K. and MacKain, K. (1983). The prosody of maternal speech: Infant age and context related changes. *Journal of Child Language, 10*, 1-15.

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

[7] 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.

[8] Xu, N., and Burnham, D. (under review) Tone hyperarticulation in Cantonese infant-directed speech

[9] Andruski, J.E., Kuhl, P.K., and Hayashi, A. (1999). The acoustics of vowels in Japanese womens' speech to infants and adults. Paper presented at ICPhS99, San Francisco, 2177-2179.

[10] Liu, H.M., Kuhl, P.K, and Tsao, F.M. (2003). An association between mothers' speech clarity and infants' speech discrimination skills. *Developmental Science, 6*(3), F1-F10.

[11] Kitamura, C., Thanavisuth, C., Burnham, D., & Luksaneeyanawin, S. (2002). Universal pitch modifications in infant directed speech: a prelinguistic longitudinal study in a tonal and non-tonal language. *Infant Behavior and Development,* 24 (4), 372-392.

[12] Cooper, R. P., & Aslin, R. N. (1990). Preference for infant-directed speech in the first month after birth. *Child Development, 61,* 1584–1595.

[13] Fernald, A. (1985) Four-month-old infants prefer to listen to motherese. *Infant Behavior and Development, 8,* 181–195.

[14] Pegg, J. E., Werker, J. F., & McLeod, P. J., (1992). Preference for infant-directed over adult-directed speech: Evidence from 7-week-old infants. *Infant Behavior and Development, 15,* 325–345.

[15] Hayashi, A., Tamekawa, Y., and Kiritani, S. (2001). Developmental changes in auditory preferences for speech stimuli in Japanese infants. *J. Speech, Lang., Hear. Res*., *44*, 1189-1200.

[16] Werker, J.F., Pons, F., Dietrich, C., Kajikawa, S., Fais, L., and Amano, S. (2007). Infant-directed speech supports phonetic category learning in English and Japanese. *Cognition*, *103*, 147-162.

[17] Lees, N.C. (2008) Vocalisations with a better view: Hyperarticulation augments the auditory-visual advantage for the detection of speech in noise. Doctoral Thesis, MARCS Labs, University of Western Sydney.

[18] Sekiyama, K., and Tohkura, Y. i. (1993). Inter-language differences in the influence of visual cues in speech perception. *Journal of Phonetics, 21*(4), 427-444.

[19] Sekiyama, K. and Tohkura, Y. (1991) McGurk effect in non-English listeners: Few visual effects for Japanese subjects hearing Japanese syllables of high auditory intelligibility. *J. Acoustical Society of America*, *90*, 1797-1805.

[20] Sekiyama, K., and Burnham, D. (2008) Impact of language on development of auditory-visual speech perception. *Developmental Science*, 11(2), 303–317.

[21] Bloom, K., & Masataka, N. (1996). Japanese and Canadian impressions of vocalizing infants. *International Journal of Behavioral Development, 19*(1), 89-99.

[22] Coxeter, H. S. M., & Greitzer, S. L. (1967). *Geometry Revisited.* Washington, DC: Math. Assoc. Amer., pp. 56-60, 1967.

[23] Göcke, R (2005) 3D lip tracking and co-inertia analysis for improved robustness of audio-video automatic speech recognition. *Proceedings of the Auditory-Visual Speech Processing Workshop, Canada,* 109-114.