

INFANT PREFERENCES FOR AGE-RELATED INFANT-DIRECTED SPEECH: THE SALIENCE OF VOCAL AFFECT

CHRISTINE M. KITAMURA
School of Psychology
University of New South Wales

Three explanations of the potential benefit that the exaggerated intonation patterns of infant-directed speech (IDS) may have for infants are: (i) attentional, (ii) affective/social and (iii) linguistic/didactic. These were investigated using an auditory preference procedure that tested infants aged 5 and 12 months of age for their preferences for speech directed to 5-month-old infants (5DS) and to 12-month-old infants (12DS). In addition to age, attentional and affective variables were manipulated by using four speaker combinations created by pairing each of two 5DS with two 12DS speakers. Prior to the auditory preference study these speaker combinations were analysed for their pitch characteristics and affective salience. Contrary to previous findings, the 5DS stimuli did not always have higher pitch and more pitch modulation than the 12DS stimuli. Nevertheless 5DS samples were always rated to be warmer and more affectionate than the 12DS samples. The results of the auditory preference study showed that both 5- and 12-month-old infants preferred to listen to 5DS. This was found for both normal and low-pass filtered versions of the speech stimuli. These results suggest IDS preferences are based not only on pitch but also on expressed affect.

When adults and even children address infants they use a particular speech style which has been termed 'parentese' or infant directed speech (IDS). IDS is characterised by shorter utterances, a more restricted vocabulary and more utterance repetition (Snow & Ferguson, 1977). More notably, acoustic analyses of the prosodic features of IDS show higher overall pitch, more pitch modulation, slower tempo and longer pauses (Stern, Spieker & McKain, 1982; Fernald & Simon, 1984 & Fernald & Mazzie, 1991).

Several studies show infants at various ages prefer to listen to IDS rather than ADS (Fernald, 1985; Cooper & Aslin, 1990; Werker & McLeod, 1989; Pegg, Werker & McLeod, 1992). In IDS, it seems that pitch is of considerable importance in communicating with the infant rather than duration or amplitude patterns (Fernald & Kuhl, 1987). Three possible functions for IDS have been identified: firstly it engages and maintains infant attention (Fernald et al., 1987; Stern et al., 1982; Papousek, Papousek & Symmes, 1991; Sullivan & Horowitz, 1983); secondly, it communicates affect and facilitates social interaction (Fernald, 1989; Werker & McLeod, 1989; Papousek, Bornstein, Nuzzo, Papousek & Symmes, 1990; Fernald, 1993) and thirdly, it facilitates language acquisition (Kemler Nelson, Hirsh-Pasek, Jusczyk & Wright Cassidy, 1989; Hirsh-Pasek, Kemler Nelson, Jusczyk, Wright Cassidy, Druss & Kennedy, 1987; Fernald & Mazzie, 1991; Jusczyk, Cutler & Redanz, 1993). As there is evidence to support each of these hypotheses, it seems the role of IDS may be multifaceted: the attentional and affective functions may render IDS salient to the infant without diminishing its role of facilitating the acquisition of communication and language skills.

It has been suggested that it is through the attentional and affective qualities of IDS that language first becomes meaningful to the prelinguistic infant but that the linguistic function of IDS may become more dominant with increasing age (Fernald, 1985; Cooper & Aslin, 1989; Fernald & Mazzie, 1991). Therefore it is important to examine age-related modifications to IDS as the infant develops, both with respect to its production by caregivers and its perception by infants. Studies involving children from 16 months to 5 years show that pitch and pitch modulation decrease with the increasing age of the child (Garnica, 1977; Warren-Leubecker & Bohannon, 1984; Remick, 1976). With younger prelinguistic infants, only one study has examined age-related adjustments to the acoustic characteristics of IDS. Stern, Spieker, Barnett & McKain (1983) found greater pitch modulation in speech addressed to 4-month-old infants compared to speech directed to neonates, 12-month-olds and 2 year olds. While these studies indicate that aspects of IDS are modified ontogenetically, they do not reveal what specific features of IDS infants preferentially perceive and respond to at different ages.

The study of IDS reported here combines both a production and perception experiment. By analysing the features of mothers' speech to infants of two different ages and testing infants' response to these features it is hoped to determine not only if the function of IDS changes over age but also to determine whether the basis of IDS preferences at each age is attentional, affective or linguistic/didactic.

EXPERIMENT 1:
PART A: THE PRODUCTION OF AGE-RELATED IDS

The aim in Part A of the production study was to determine any differences there might be between speech directed to 5-month-old infants (5DS) and to 12-month-old infants (12DS), firstly, in the pitch characteristics and secondly, with adult ratings of expressed warmth and affection. In regard to these two measures it was expected that the pitch level and pitch variation would be higher in 5DS than in 12DS, and that 5DS would express more warmth and affection than 12DS.

METHOD

Speech samples were collected from four mothers of 5-month-old infants and four mothers of 12-month-old infants.

Acoustic Analysis: Ten second excerpts of speech were sampled approximately every 30 seconds from each of the eight speakers and the pitch extraction function used to calculate mean F_0 and standard deviation F_0 .

Affective Analysis: Rating of warmth and affection were collected from 70 adults using a Likert scale ranging from 1 (neutral warmth and affection) to 9 (extreme warmth and affection). This was done for both normal speech samples and low-pass filtered versions, obtained by applying a Hanning filter set to an upper limit of 400 Hz.

RESULTS

Acoustic Data: Mean F_0 and standard deviation F_0 were analysed in 2×4 ANOVAs with IDS age, either 5- or 12-month-old, and the four speakers for each age group, as between subjects factors. The results are shown in Table 1. The difference between the pitch characteristics of maternal speech to the two ages of infants was not significant for either mean F_0 or SD F_0 . The speaker variation in mean F_0 was significantly greater between the four mothers speaking to 5-month-old infants than that between the four mothers speaking to 12-month-old infants. It is possible that these may have obscured any differences between ages.

Affective Data: Warmth and affection ratings were analysed in a $(2 \times 2 \times 4)$ ANOVA with age of the infant being spoken to, 5DS or 12DS; version, normal or filtered; and speaker, the four speakers from each age group, as within subjects factors. The ratings of warmth and affection for the normal and filtered versions of the 5DS and 12DS are shown in Table 1. Adults rated 5DS as being significantly warmer and more affectionate than 12DS. Overall the normal versions were rated as being significantly warmer and more affectionate than low-pass filtered versions. Further, the interaction between age and version was significant

IDS SPEAKER	PITCH		WARMTH& AFFECTION	
	Mean	SD	Normal	Filtered
5DS				
Speaker 5-1	321	104	6.93	5.39
Speaker 5-2	244	77	6.40	5.69
Speaker 5-3	268	81	6.41	6.50
Speaker 5-4	246	68	5.91	5.27
12DS				
Speaker 12-1	249	79	4.54	3.73
Speaker 12-2	281	100	5.13	5.31
Speaker 12-3	278	88	5.30	5.10
Speaker 12-4	296	104	5.76	4.90

Table 1: Results for Pitch and Ratings of Warmth & Affection

indicating that for 5DS there was a larger difference between normal and filtered versions than for 12DS. This suggests that in 5DS warmth and affection is expressed both by the prosody and by the semantic content as rated warmth and affection was greater for normal than filtered speech. On the other hand for 12DS, warmth and affection appears to be conveyed predominantly by the prosody.

In summary the significant interactions between speaker and age; speaker and version; and speaker, version and age respectively showed more variation between the ratings of warmth and affection of 12DS speakers than of 5DS speakers, more variation in ratings of the filtered versions than the normal versions of the stimuli, and finally that filtered 12DS were the most variable while those of normal 5DS were the least variable. Most importantly, though, irrespective of this individual variation, 5DS is consistently rated as warmer and more affectionate than 12DS. These results contrast with those from the acoustic analysis where individual variation may have obscured any pitch differences. Therefore it is possible that the expression of warmth and affection may be the more salient feature that distinguishes speech directed to younger infants from that addressed to older infants.

PART B: DERIVATION OF IDS STIMULI

In Part B the speech of two 5DS speakers and two 12DS speakers was selected to be used as stimuli in a perception experiment. Here it was hoped to identify the relative contributions of pitch and rated affect to infants' preferences

for IDS. For example, given that rated warmth and affection is consistently greater in 5DS than 12DS, which do infants prefer when they hear 5DS and 12DS spoken with similar pitch characteristics? On the other hand, what do infants prefer when they are presented with 5DS which is higher pitched than 12DS, or in another test with 12DS which is higher pitched than 5DS?

METHOD

The speakers to be included were 5-1 ($F_0=321$) and 5-2 ($F_0=244$) from the 5DS group and speaker 12-1 ($F_0=249$) and speaker 12-2 ($F_0=281$) and from the 12DS group (see Table 1). As 5DS was always to be paired with 12DS, four speaker combinations were possible. These are distinguished by abbreviations that reflect the differences in the mean pitch level between the 5DS and 12DS speakers. Thus, 5VHI (speaker 5-1 & speaker 12-1) indicates that the pitch level in 5DS is very high compared to 12DS, 5HI (speaker 5-1 & speaker 12-2) indicates that the pitch level in 5DS is higher than in 12DS, 5EQ12 (speaker 5-2 & speaker 12-1) reflects equivalent pitch levels in 5DS and 12DS and the 5LO combination (speaker 5-2 & speaker 12-2) indicates that the pitch level in 5DS is lower than 12DS.

Acoustic Analysis: For the purpose of testing infants' preferences in the perception study, sixty-second speech stimuli were created from each of the four speakers and the mean and standard deviation F_0 calculated using the pitch extraction profile.

Affective Analysis: Normal and low-pass filtered (400Hz) versions of the IDS stimuli were presented to a new group of 90 undergraduate students for rating.

RESULTS

Acoustic Data: Data were analysed using a 4-level ANOVA which tested the four pairwise comparisons of speakers to be used in Experiment 2. See Table 2 for results. For both the 5VHI and 5HI combination both mean F_0 and SD F_0 were higher for speaker 5-1 than speaker 12-1. In the 5EQ12 combination there was no significant difference between speaker 5-2 and speaker 12-1 on either mean F_0 or SD F_0 . For the 5LO combination, speaker 5-2 had a significantly lower mean F_0 than speaker 12-2 but there was no significant difference in SD F_0 .

Affective Data: Warmth and affection ratings from the 1-9 Likert scale were analysed using a 4-level ANOVA which tested the effect of version, normal or filtered, and the four pairwise comparisons of speakers. Table 2 shows the ratings of warmth and affection for each speaker in both normal and filtered versions. Overall adults rated the normal speech versions to be significantly warmer and more affectionate than the filtered versions. The pairwise comparisons constructed to test the four

IDS SPEAKER	PITCH		WARMTH& AFFECTION	
	Mean	SD	Normal	Filtered
5DS				
Speaker 5-1	332	122	7.63	5.48
Speaker 5-2	239	74	7.77	5.57
12DS				
Speaker 12-1	237	69	6.39	3.88
Speaker 12-2	284	89	6.86	5.14

Table 2: Results for Pitch and Ratings of Warmth & Affection for IDS Stimuli

speaker combinations to be used in Experiment 2 revealed that, despite the relative differences in pitch, the 5DS speakers were rated as warmer and more affectionate than the 12DS speakers in all four combinations, irrespective of whether speech was normal or filtered. None of these contrasts interacted significantly with version.

EXPERIMENT 2: INFANTS' RESPONSE TO AGE-RELATED IDS

In Experiment 2, and auditory preference method was used to test sixteen 5-month-old and sixteen 12-month-old infants. Infants were presented with two conditions: in one, normal versions of 5DS and 12DS were available and in the other, versions of the 5DS and 12DS stimuli that had been low-pass filtered at 400 Hz were available.

In order to distinguish between the attentional and affective functions of IDS, specific predictions were made in respect of the individual speaker combinations. These were generated from the analyses of pitch and affect summarised in Table 3. If infants respond on the basis of the mean level of pitch in the mother's voice then it would be expected that infants should prefer 5DS in the 5VHI and 5HI combination, show no preference in the 5EQ12 combination and prefer 12DS in the 5LO combination. If infants respond on the basis of pitch modulation they should show similar results for the first three pairings but no preference in the 5LO combination. If infants respond on the basis of the rated warmth and affection in the speakers' voices then they should prefer the 5DS speaker in all speaker combinations. Preferences consistent with predictions from the pitch analyses would suggest IDS has a primarily attentional role while preferences consistent with predictions from the affective analyses would suggest an affective one.

Other outcomes are possible in terms of the age of the infant. Firstly, both age groups of infants may prefer their own-age IDS. If so then it could be assumed that mothers are fine-tuning their speech input to the infants' maturing level of perceptual and linguistic competence and infants are responding to such fine-tuning. In terms of the three functions suggested earlier such results would support the linguistic/didactic role of IDS. Secondly, both 5- and 12-month-old infants may prefer 5DS. This would imply that infants are responding to the heightened pitch characteristics and/or increased warmth and affection in speech to younger infants and would suggest either an affective or an attentional role for IDS.

For the low-pass filtered versions of the speech samples of 5DS and 12DS two hypotheses were considered: Firstly, it is possible that infants would show no preference between filtered 5DS and filtered 12DS. If so, and if they do show a preference in the normal versions, then it could be assumed that the segmental features of speech are more important in determining preferences than the suprasegmental features. Secondly, infants may show a preference for one filtered IDS sample over the other. If so, and if they show similar preferences in the normal pairings, it could be assumed the suprasegmental information is at least as important as the segmental in determining infant preferences.

METHOD

Infants were tested seated on their parent's lap in front and on midline of two video monitors. Each of the two speech stimuli (5DS or 12DS) were matched with one of the video monitors (left or right) each displaying identical visual stimulus. The infants' head and eye

movements were video-taped by a video camera connected to a video monitor in the adjacent room on which the observer viewed the infant's image and used a left-right toggle switch to turn on the left or right speech stimuli. Initially, infants were familiarised with the speech stimuli: the 5DS and the 12DS sample were alternatively played for 12 seconds. In the 60-second test phase, presentation of the speech stimuli was contingent on the infant's direction of fixation: infants could 'turn on' the left or right speech stimulus by fixating the appropriate video monitor.

RESULTS

The data were analysed in a 2 x 4 x (2 x 2) ANOVA with age group, 5 or 12 months, and the four speaker combinations as between subjects factors and version, normal or filtered, and IDS age, 5DS or 12DS, as within subject factors.

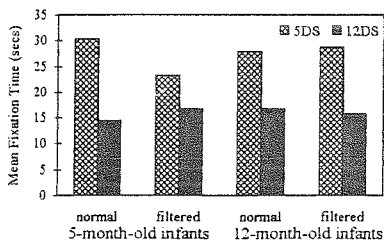


Figure 1: Mean Fixation Time of Infants to 5DS & 12DS for Normal and Filtered Versions

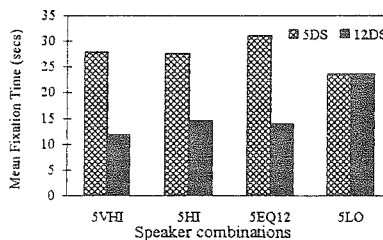


Figure 2: Mean Fixation Time for the Four Speaker Combinations

Age-related IDS effects: Figure 1 shows the preferences for 5DS and 12DS for normal and filtered versions in both age groups. Overall infants listened to 5DS for longer than 12DS. The interaction between IDS age and age group was not significant indicating a statistically equivalent preference for 5DS was shown by both 5- and 12-month-old infants.

Version effects: The main effects for version, for age group and for the speaker combinations were not significant nor were any of their interactions with each other. This indicates that 5- and 12-month-old infants were no more responsive to the normal than the filtered versions of the speech stimuli in any of the speaker combinations.

Speaker combination effects: As can be seen in Figure 2 fixation times for 5DS were greater than those for 12DS for the speaker combinations 5VHI, 5HI and 5EQ12 but for the 5LO combination there was no difference between 5DS and 12DS. This indicates that infants in both age groups showed a similar pattern of 5DS/12DS preferences for the four speaker combinations.

CONCLUSIONS

In Experiment 1, the production study, no overall difference was found in the level of pitch or pitch variation between the 5DS speakers and the 12DS speakers. However, the variation among individual speakers within each IDS age may have overshadowed the variation between IDS ages. The measure which gave rise to the most interesting results was the expressed warmth and affection in mothers' speech. This was uniformly higher for all four 5DS speakers, that is, there was no individual 5DS speaker for whom the ratings of warmth and affection was less than any 12DS speaker. This contrasts with the results of the pitch analysis where some 5DS speakers (e.g. speaker 5-2) had lower mean pitch and standard deviation than some of the 12DS speakers. Generally high pitch and high pitch modulation are associated with positive affect in speech (Scherer, 1986), however, while these two factors usually co-vary, the results of the current study suggest that the acoustic salience of IDS may vary independently of its affective salience.

The most important finding in Experiment 2, the perception study, was that both 5- and 12-month-old infants generally preferred to listen to 5DS than 12DS. This was the case in both the normal and filtered speech conditions. Moreover, as shown in Table 3, this 5DS preference occurred not only when the mean pitch and rated warmth and affection of 5DS were greater than that of 12DS (see 5VHI and 5HI speaker combinations) but also when only warmth and affection was higher in the 5DS sample (see 5EQ12 combination). It was only in the 5LO combination, in which mean pitch of 5DS was lower than 12DS but rated warmth & affection higher, that infants' preference for 5DS disappeared. Thus it seems that both pitch and affective salience independently contribute to infants' preferences for IDS and that these preferences are not simply the consequence of variations along a single dimension.

This preference for 5DS over 12DS occurred for infants of both ages, 5 and 12 months. This implies that IDS plays an affective and attentional role in language acquisition rather than a linguistic/didactic role which would have been supported had infants shown own-age IDS preferences. As the results discussed above show that the affective features contribute to these preferences independently of the attention-getting pitch characteristics, it may be that the enhanced affective tone of IDS (especially 5DS) plays a role in the infant's acquisition of language skills and the protocols of interpersonal interaction (Snow, 1989). Infants of 12 months of age seem to respond to the higher pitch and positive affect in 5DS despite the lower pitch level (Stern et al., 1983) and lower affective salience in 12DS, the speech they normally hear. It may be that the attentional and affective salience of speech normally addressed to 5-month-olds, is necessary in order that these young infants learn the basic skills of socialisation and communication, such as turntaking.

The preference for 5DS over 12DS occurred for both the normal and the filtered versions of the mothers' speech. Moreover, the duration of fixation for both 5- and 12-month-old infants was statistically equivalent for the normal and filtered versions. Thus, even though it might have been expected that infants would have been more interested in the more familiar normal speech rather than the low-pass filtered speech, this was not the case. Although no real parallels can be drawn between the different methods, it should be noted that the adult raters in the production study perceived the filtered versions to express less warmth and affection than the normal versions. Thus it seems that filtered speech is less attractive to adults than normal speech while for 5- and 12-month-old infants the low-frequency prosodic contours alone are just as appealing as speech that also contains segmental information. These findings suggest that infants of both age groups do not necessarily attend preferentially to the segmental information.

The combination of a production and perception experiment in the one study has proven to be a useful method for identifying the features of speech that infants respond to in their preference for IDS. This study has found that the affective salience of speech directed to 5-month-old infants is a significant contributor to infants' preferences for this speech style. The affective component of IDS may have important consequences for infants' development, not only generally of socialisation but also more specifically of language acquisition.

REFERENCES

- Cooper, R.P., & Aslin, R.N. (1990). Preference for infant-directed speech in the first month after birth. *Child Development, 61*, 1584-1595.
- Fernald, A. (1984). The perceptual and affective salience of mothers' speech to infants. In L. Feagans, C. Garvey & R. Golinkoff (Eds.). *The origins and growth of communication*. Ablex: Norwood, N.J.
- Fernald, A. (1985). Four-month-old infants prefer to listen to motherese. *Infant Behaviour and Development, 8*, 181-195.
- Fernald, A. (1989). Intonation and communicative intent in mothers' speech to infants: Is the melody the message? *Child Development, 60*, 1497-1510.
- Fernald, A., & Kuhl, P. (1987). Acoustic determinants of infant preference for motherese speech. *Infant Behaviour and Development, 10*, 279-293.
- Fernald, A., & Mazzie, C. (1991). Prosody and focus in speech to infants and adults. *Developmental Psychology, 27*, 209-221.
- Fernald, A., & Simon, T., (1984). Expanded intonation contours in mother's speech to newborns. *Developmental Psychology, 20*, 104-113.
- Garnica, O. (1977). Some prosodic and paralinguistic features of speech to young children. In C. E. Snow & C. A. Ferguson (Eds.). *Talking to children: Language input and acquisition*. New York: Cambridge University Press.
- Hirsh-Pasek, K., Kemler Nelson, D.G., Jusczyk, P.W., Wright Cassidy, K., Druss, B., Kennedy, L. (1987). Clauses are perceptual units for young infants. *Cognition, 26*, 269-286.
- Jusczyk, P., Cutler, A., & Redanz, N. (1993). Infants' preference for the predominant stress patterns of English words. *Child Development, 64*, 675-687.
- Kemler Nelson, D.G., Hirsh-Pasek, K., Jusczyk, P.W., Wright Cassidy, K. (1989). How the prosodic cues in motherese might assist language learning. *Journal of Child Language, 16*, 55-68.
- Papousek, M., Papousek, H., & Symmes, H. (1991). The meanings of melodies in motherese in tone and stress languages. *Infant Behaviour and Development, 14*, 415-440.
- Papousek, M., Bornstein, H., Nuzzo, C., Papousek, H., & Symmes, D. (1990). Infant responses to prototypical melodic contours in parental speech. *Infant Behaviour and Development, 13*, 539-545.
- Pegg, J., Werker, J., & McLeod, P. (1992). Preference for infant-directed over adult-directed speech: Evidence from 7-week-old infants. *Infant Behaviour and Development, 15*, 325-345.
- Rernick, H., (1976). The maternal environment of linguistic development. Paper (1971) abstracted in C. Snow & C. Ferguson (Eds.) *Talking to children*. New York: Cambridge University Press.
- Scherer, K.R. (1986). Vocal affect expression: A review and model for future research. *Psychological Bulletin, 99*, 143-165.
- Snow, C.E. (1989). Understanding social interaction and language acquisition; sentences are not enough. In M.H. Bornstein & J.S. Bruner (Eds.). *Human interaction*. Lawrence Erlbaum Associates: New Jersey.
- Snow, C.E., & Ferguson, C.A. (1977). *Talking to children: Language input and acquisition*. New York: Cambridge University Press.
- Stern, D.N., Spieker, S., & MacKain, K. (1982). Intonation contours as signals in maternal speech to prelinguistic infants. *Developmental Psychology, 18*, 727-735.
- 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*, 1-15.
- Sullivan, J.W., & Horowitz, F.D., (1983). The effects of intonation on infant attention: the role of the rising intonation contour. *Journal of Child Language, 10*, 521-534.
- Warren-Leubecker, A., & Bohannon, J.N. (1984). Intonation patterns in child-directed speech: Mother-father differences. *Child Development, 55*, 1379-1385.
- Werker, J.F., & McLeod, P.J. (1989). Infant preference for both male and female infant-directed talk: a developmental study of attentional affective responsiveness. *Canadian Journal of Psychology, 43*, 230-246.