INFANT PREFERENCES FOR INFANT-DIRECTED SPEECH: IS VOCAL AFFECT MORE SALIENT THAN PITCH?

Christine Kitamura & Denis Burnham
School of Psychology
University of New South Wales

ABSTRACT - The aim in the following experiments was to ascertain whether infants are more responsive to the pitch or vocal affect in infant-directed speech (IDS). In Experiments 1 and 2, infant preferences were tested for high vs low vocal affect with the level of pitch equated (HiAffect vs LoAffect IDS) and in both experiments, infants preferred to listen to HiAffect IDS. In Experiment 3, high vs low pitch was presented with the level of vocal affect equated (HiPitch vs LoPitch IDS) and it was found that infants preferred LoPitch over HiPitch IDS. This result was unanticipated and when a different procedure was used to rate the vocal affect of the speech exemplars in Experiment 4, there was no difference in infant preferences for Hi or LoPitch IDS. Taken together these two sets of results suggest that it is the affective salience of IDS that is important to infant responsiveness and not necessarily the pitch characteristics alone. In the final experiment infants showed no differential preferences for normal or low-pass filtered IDS, confirming they are as responsive to the intonation or pitch characteristics as they are to full spectral versions of speech. Therefore it is suggested that pitch is used as a means of conveying affective intent to infants.

INTRODUCTION

Infant-directed speech is the speech style used to address infants and young children. Among its most notable prosodic characteristics are its high pitch and exaggerated pitch modulation. Past research has found that infants are more responsive to IDS than adult-directed speech (ADS) (Fernald, 1985; Werker & McLeod, 1989; Cooper & Aslin, 1990). From this research it is often concluded that infants are responding to the higher pitch and increased pitch modulation in parents' voices. Although Fernald & Kuhl (1987) found infants respond to the pitch rather than the amplitude or duration patterns in sinewave analogues of speech, to date, no study has directly investigated this assumption with natural speech. The affective tone of IDS is also discussed as being important to infants' preferences and two recent studies (Papousek, Bornstein, Nuzzo, Papousek & Symmes, 1990 and Fernald, 1993) have investigated this, both using approving and disapproving contours. Papousek et al. (1990) found infants showed a *preference* for approving contours and in Fernald's (1993) study infants were more *affectively responsive* to the same type of contour. However, neither study controlled the level of pitch in the speakers' voices. Thus the primary aim in the following experiments was to investigate infants' responses to (1) high vs low vocal affect controlling the level of pitch and (2) high vs low pitch controlling the level of vocal affect in IDS.

EXPERIMENT 1: HiAffect IDS vs LoAffect IDS

In this study infants were presented with two speech samples in which mean fundamental frequency (F_0) and standard deviation F_0 were equated and the degree of affection expressed by the speaker manipulated; on one side the speech sample available to the infant had high affective tone and on the other side, low affective tone. Each speech sample was matched with a visual target which consisted of three concentric red circles resembling a target. Based on past research and theory it was expected that infants would prefer the HiAffect to the LoAffect IDS stimuli.

Method

Twenty 6-month-old infants participated in the study. Infants were seated on their parent's lap in front of, and facing midline of two video monitors with separate speakers. Each of the two types of speech were paired with the same visual target, one presented to the infants' left and the other to their right. Thus if infants fixated the left visual target they heard one style and if they fixated the right visual target they heard the other style. Initially infants were familiarised for 30-seconds with both speech styles, played alternately. In the test phase, there were six 20-second trials in which presentation of the speech stimulus was contingent on the infants' direction of fixation. Side of presentation was counterbalanced so that half the infants heard HiAffect IDS on the left side and LoAffect IDS on the right side and this was reversed for the other half of the subjects. In addition, order of presentation in the familiarisation phase was counterbalanced: half the infants began trials on the left side and the other half on the right side. During familiarisation and testing the infants' head and eye movements were

recorded by video camera and observed on a video monitor in an adjacent room. The observer viewed the infant's image and used a left-right toggle switch to turn on the left or right speech stimuli depending on the infants' looking direction. It should be noted that the toggle switch directly controlled the speakers rather than the cassette players (which played continuously). Hence any delay in the onset of speech was avoided.

It was necessary to use a female actor to record the speech stimuli because the required manipulations of affect and pitch do not occur frequently in natural settings, e.g., low pitch with high affect or high pitch with low affect. Furthermore, it was necessary to control the semantic content which is not possible in natural speech samples. For these reasons a female actor with good voice control was used to construct the auditory stimuli. Multiple instances of the actor saying 'Would you like to go for a walk with mummy? Yes? Lets go for a little walk.' were recorded. For the experiments in which pitch is equated, two sets of stimuli were recorded, one with high vocal affect and one with low affect. Similarily, in the experiments where affect was equated two additional sets were recorded: one with high pitch and and the other with low pitch. The pitch of these utterances was then measured using the same procedure as in Kitamura & Burnham (1996). To determine the level of expressed vocal affect a ratings study was conducted. Twenty-five utterances were selected and rated by fifteen post-graduate students for degree of vocal affect on a scale ranging from 1 (neutral) to 12 (very high).

	HI AFFECT SPEECH EXEMPLARS			LO AFFECT SPEECH EXEMPLARS		
Pair No	Mean F_0	SD F_0	Rated Affect	Mean F_0	SD F_0	Rated Affect
1	306	119	11.4	319	109	7.4
2	316	76	11.7	307	76	7.2
3	306	103	11.2	306	103	4.9
4	342	92	10.9	367	92	5.4
5	346	107	12	399	107	5.7
mean	339.6	98.4	11.4	323.4	99.4	6.1

Table 1: Measures of pitch and rated affect for the five HiAffect and five LoAffect exemplar pairs used in Experiment 1 and Experiment 2.

For the speech stimuli in Experiment 1, five exemplars of the HiAffect utterances and five exemplars of the LoAffect utterances were selected based on matching the level of pitch as closely as possible. Table 1 shows the mean-F_0, standard deviation-F_0 and the rated affect for the five LoAffect and five HiAffect IDS exemplar pairs used. The speech stimuli were recorded onto five audio cassettes in different orders with each exemplar being presented at a different position in each of the five sequences. The length of each exemplar was 5 seconds. The sound level was adjusted to 60-65 dB SPL (B scale) for presentation to the infants.

Results and Discussion

Infant looking times were averaged across the six trials and the data analysed using a 2 x 2 x (2) analysis of variance (ANOVA) with side of presentation (HiAffect left or HiAffect right) and order of presentation (HiAffect first or LoAffect first) as the between subjects factors and preference (HiAffect or LoAffect) as the within subjects factor. As shown in Figure 1, infants showed significantly greater preference for HiAffect IDS than for LoAffect IDS ($F=5.1$). There was a main effect for order of presentation indicating that infants looked longer overall in test trials when LoAffect IDS was presented first in familiarisation trials than when it was presented second ($F=4.58$). Generally it seems HiAffect IDS has more attentional characteristics than LoAffect IDS not only because infants show a preference for HiAffect IDS, but also because when it was the last speech style heard in familiarisation trials infants show longer overall fixation times. It made no difference whether HiAffect was presented to the left or the right of the infant.

EXPERIMENT 2: HiAffect IDS vs LoAffect IDS (with Face)

In Experiment 2, a female face was used as a visual target instead of the three concentric red circles, to determine whether the presentation of a face as a visual target would have an impact on infant looking times. If so, it should increase infants' attentiveness.

Method

Twenty 6-month-old infants participated in the study. The same method and speech stimuli from Experiment 1 were used in this experiment. The only exception was that a face was presented as a visual target on the video monitors during testing. The visual target was a colour image of a woman with a slight smile; her expression being between happy and emotionally neutral. This intermediate expression was compatible with both HiAffect and LoAffect IDS as an emotionally neutral face would not have been congruent with the affection being

expressed in either speech style while a happy expression would not have been consistent with the LoAffect speech stimulus. The female face was recorded onto two video-tapes for presentation during the experiment.

Results and Discussion

As in Experiment 1, the data were analysed in a 2 x 2 x (2) ANOVA with order of presentation and side of presentation as between subjects factors and infant preferences as the within subjects factor. The results are shown in Figure 1 and substantiate the findings from Experiment 1. They indicate infants' preferences for HiAffect IDS are greater than for LoAffect IDS ($F=12.1$). No other main effects or interactions were significant. The data from this study were also analysed in conjunction with the data from Experiment 1 in a 2 x 2 x 2 x (2) ANOVA where visual target (face or circles), order of presentation and side of presentation were between subjects factors and preference was the within subjects factor. As would be expected from the previous results, infants showed a stronger preference for HiAffect IDS than for LoAffect IDS ($F=15.55$). There was also a main effect for visual target which showed that overall infants looked longer when the face was presented as the visual target than the three concentric circles ($F=24.64$). Infant attention appears to be greater when the face is presented rather than the circles not only because the face is a more complex pattern but it may have more emotive associations for the infant. No other main effects or interactions were significant. Thus while the face results in increased attention, the degree of preference for HiAffect IDS remained unaltered.

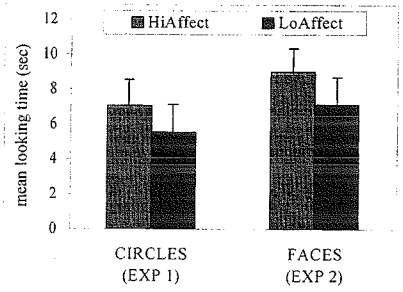

Figure 1: Mean looking time (seconds) for HiAffect IDS and LoAffect IDS for Experiment 1 (circles visual target) and Experiment 2 (faces visual target)

EXPERIMENT 3: HiPitch IDS Vs LoPitch IDS (I)

In this study vocal affect was equated and the level of mean pitch and pitch modulation in the speaker's voice manipulated so there was HiPitch IDS on one side and LoPitch IDS on the other side. Note that the level of pitch in LoPitch IDS is comparable to the pitch level of ADS. It was expected that infants would show a stronger preference for HiPitch IDS than LoPitch IDS.

Method

Twenty 6-month-old infants were tested using the same method as in Experiment 2 with the female face presented as a visual stimulus to ensure optimal infant attention. However, the speech stimuli were replaced by five exemplars of HiPitch IDS and five exemplars of LoPitch IDS with the pairings based on the ratings of vocal affect. The F_0 means, F_0 standard deviations and rated affect of each pair of speech exemplars are shown in Table 2. However, unlike Experiments 1 and 2, the length of the speech exemplars was different. Here, the length of each of the LoPitch speech exemplars was approximately 5.4 seconds while the length of the HiPitch speech exemplars was approximately 5.0 seconds. It seems when positive affect is expressed with lowered pitch, increasing duration rather than pitch conveys heightened vocal affect. The sound level was adjusted to 60-65 dB SPL (B scale) for both sets of speech stimuli.

	HiPitch Speech Exemplars			LoPitch Speech Exemplars		
No	Mean F0	SD F0	Rated Affect	Mean F0	SD F0	Rated Affect
1	334	104	11.3	208	78	11.4
2	316	76	10.9	174	69	11.0
3	378	120	9.4	171	35	9.4
4	387	92	9.1	192	67	9.6
5	367	103	9.0	208	67	8.9
Mean	356.4	99	9.94	190.6	63.2	10.06

Table 2: Measures of pitch and rated affect for the five HiAffect and five LoAffect exemplar pairs used in Experiment 3

Results and Discussion

As in the previous experiments, data were analysed in a 2 x 2 x (2) ANOVA with order of presentation and side of presentation as between subjects factors and infant preferences as the within subjects factor. The results, as shown in Figure 2a, indicate that when vocal affect is similar, infants prefer to listen to LoPitch than HiPitch IDS (F=6.97). This finding is unexpected as, based on past research (Fernald, 1985; Werker & McLeod, 1989; Cooper & Aslin, 1990), it was expected that infants would prefer HiPitch to LoPitch IDS as the latter is more analogous to the pitch of ADS. The difference in this study was that the low pitch speech conveyed a high level of vocal affect. No other main effects or interactions were significant.

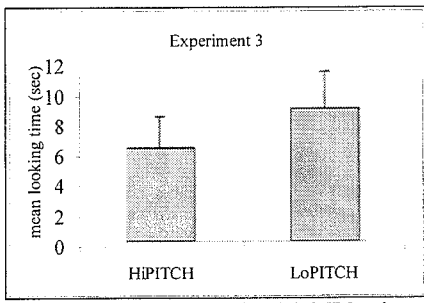

Figure 2a: Mean looking times for HiPitch IDS and LoPitch IDS with vocal affect equated for Exp. 3.

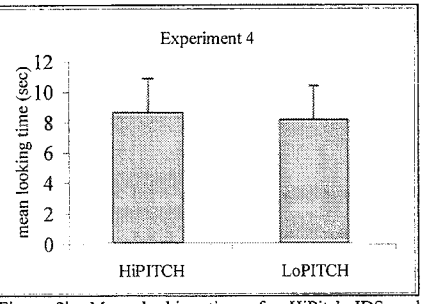

Figure 2b: Mean looking times for HiPitch IDS and LoPitch IDS with vocal affect equated for Exp. 4.

EXPERIMENT 4: HiPitch IDS Vs LoPitch IDS (I I)

The results in Experiment 3 were unexpected. Of course, it may be that infant preferences for IDS over ADS might be better explained by the greater affect in IDS rather than its pitch characteristics but if vocal affect is more important to infant preferences than pitch, it would be more likely that infants would show no preference for either speech style. An uncontrolled feature of the speech stimuli which might have affected testing was the duration of exemplars: the HiPitch exemplars (5.0 seconds) were shorter than the LoPitch exemplars (5.4 seconds). Remember because both tapes were playing continuously (with the speakers off), when infants fixated the video monitors with exemplars of the same duration, the same part of either exemplar in any pair should play as was the case in Experiments 1 and 2. However, in Experiment 3, HiPitch IDS exemplars were shorter than the LoPitch IDS exemplars and therefore would have been accessed at a point further along the tape than the LoPitch IDS. As the range of rated vocal affect in these two sets of exemplar pairs was quite substantial (8.9 to 11.4), the HiPitch speech exemplars might sometimes play at a point where they actually had lower rated vocal affect than LoPitch exemplars. It may be that HiPitch exemplars with lower rated vocal affect may not have been attractive to the infants, given the perceptual subtleties of this voice quality (see Table 2). A more cohesive set of stimuli might improve this situation.

Method

Twenty-four 6-month-old infants were tested with the same method as described in previous experiments. However, for the speech stimuli, a smaller set of exemplar pairs (n=3) was created with a more restricted range of vocal affect. This time, instead of relying solely on the results from the previous ratings studies which had used isolated utterances, utterances were rated in pairs. Nine pairs of speech exemplars were selected from the previous ratings study for possible inclusion as stimuli with the level of vocal affect ranging from 10.4 to 11.4. These were played to four expert raters who were asked to indicate to what extent they agreed the two utterances were a good match. A practice pair was included and each pair was played in both orders, that is, HiPitch then LoPitch and the reverse. Three of the nine pairs were rated the best matches overall: Pair 1 had 87% agreement; Pair 2, 100% agreement; and Pair 3, 85% agreement were chosen. The ratings are shown in Table 3.

	HiPitch Speech Exemplars			LoPitch Speech Exemplars		
Pair No.	Mean F_0	SD F_0	Rated Affect	Mean F_0	SD F_0	Rated Affect
1	346	117	11.4	208	78	11.3
2	316	76	10.9	192	67	10.2
3	344	92	10.8	174	69	10.5
Mean	335.33	95	11.03	191.33	71.33	10.67

Table 3: Measures of pitch and rated affect for each of the three exemplar pairs used in Experiment 4

Results

As in Experiment 4, data was analysed in a 2 x 2 x (2) ANOVA with order of presentation and side of presentation as between subjects factors and infant preferences as the within subjects factor. No main effects or interactions were significant. The results, as shown in Figure 3b indicate that when vocal affect is equated in the manner described above, infants show no preference for HiPitch or LoPitch IDS. Together with the findings from the previous experiments, the results indicate that pitch is less important in gaining infant attention than vocal affect. Yet in IDS, the intonation or pitch contour is probably an important means of communicating vocal affect to infants. To explore this further, infants' preferences to listen to the full-spectum of speech or just the intonation pattern was investigated in Experiment 5.

EXPERIMENT 5: Filtered IDS vs Unfiltered IDS

This experiment compared normal speech to filtered speech and was conducted to determine whether infants are more responsive to the full spectrum of speech or the intonation contour alone. Filtered speech removes the upper frequencies but leaves the intonation contour intact, rendering speech unintelligible. Cooper and Aslin (1994) found 1-month-old infants showed no preference when filtered IDS was paired with filtered ADS but preferred normal IDS to filtered IDS. They argue that spectral complexity is important to speech perception and preferences in very young infants, as the ability to perceive pitch is not sufficiently developed. However they claim there may be a perceptual shift later in infancy whereby older infants attend more to pitch contour information and spectral information becomes less important. In light of this, they predict that infants of 4 months would show no preference for either filtered speech or normal speech. Supposedly by this age the intonation contour should be a significant communicative signal carrying the implicit message that the speaker is attempting to convey. Furthermore, the findings from the previous four experiments suggest that infants are attending to mothers' emotional messages rather than the semantic content of speech. Therefore it could be expected that infants would show no differential preference for normal or low-pass filterered IDS.

Method

Twenty 6-month-old infants were tested using the same method as previously described. The speech stimulus was a single HiPitch HiAffect exemplar typical of IDS. There were two versions: one version was a full spectral version and the other version was low-pass filtered at 400 Hz (Hamming filter and filtered order of 0.8 using the CSL package). Normal and low-pass filtered versions were recorded onto separate audio tapes. As filtering speech also decreases intensity, the low-pass filtered speech was recorded at 64-66 dB SPL and the normal speech at 56-58 dB SPL (B-scale). Two naive adults judged the speech to be of equal intensity at this degree of loudness.

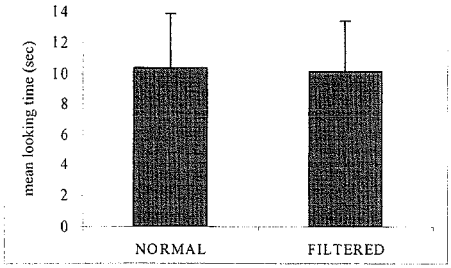

Figure 3: Mean looking times (seconds) for normal IDS and low-pass filtered IDS

Results and Discussion

As in the previous experiments, data was analysed in a 2 x 2 x (2) ANOVA with order of presentation and side of presentation as between subjects factors and infant preferences as the within subjects factor. There were no significant main effects or interactions as shown in Figure 3. The means averaged across trials were Normal (M=10.37) and Filtered (M=10.12). Infants showed no preference for either normal or filtered speech. This suggests that 6-month-old infants find the intonational characteristics of filtered IDS as interesting as normal IDS which contains spectral and segmental information. It also suggests that the information contained in

the intonation contour may have some priority in the infants' speech processing, that is, infants are responding to the tone of maternal utterances as communicative signals.

CONCLUSIONS

The critical finding from this series of experiments is that infants prefer to listen to infant-directed speech that expresses high rather than low vocal affect. This is complemented by the results from the experiments in which the confound of affect was controlled, showing that infants display no differential preference for Hi or LoPitch speech. Taken together these two sets of results suggest that it is the affective salience of IDS per se that is important to infant responsiveness and not necessarily its pitch charactistics alone. However, higher pitch and increased pitch modulation are usually associated with expressions of happiness and joy (Scherer, 1986). Thus it seems that higher pitch and increased pitch modulation are used to *convey* affect to infants: pitch provides the vehicle for successful social interaction but vocal affect is the key feature. Confirmation of this was suggested by the results of the final study which showed that 6-month-old infants are as responsive to the intonation contours in speech as they are to the full spectrum of speech, and thus may be attending to speech as a communicative rather than a linguistic signal. Thus, language acquisition in prelinguistic infants may be primarily concerned with learning about the communicative aspects of language rather than gaining specific linguistic skills.

REFERENCES

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

Cooper, R.P., & Aslin, R.N. (1994). Developmental differences in infant attention to the spectral characteristics of infant-directed speech. *Child Development, 65*, 1663-1677.

Fernald, A. (1985). Four-month-old infants prefer to listen to motherese. *Infant Behaviour and Development, 8*, 181-195.

Fernald, A., & Kuhl, P. (1987). Acoustic determinants of infant preference for motherese speech. *Infant Behaviour and Development, 10*, 279-293.

Fernald, A., (1993). Approval and disapproval: infant responsiveness to vocal affect in familiar and unfamiliar languages. *Child Development, 64*, 657-674.

Kitamura, C.M., & Burnham, D. (1996). Infants preferences for age-related infant directed speech: The salience of vocal affect. In *Proceedings of the Sixth Australian International Conference on Speech Science and Technology.*.

Papousek, M., & Papousek, H. (1991). The meaning of melodies in motherese in tone and stress languages. *Infant Behaviour and Development, 14*, 415-440.

Papousek, M., Bornstein, M.H., Nuzzo, C., Papousek, H., & Symmes, D. (1990). Infant responses to prototypical melodic contours in parental speech. *Infant Behaviour and Development, 13*, 539-545.

Scherer, K.R. (1986). Vocal affect expression: A review and model for future research. *Psychological Bulletin, 99*, 143-165.

Werker, J.F., & McLeod, P.J. (1989). Infant preferences for both male and female infant-directed talk: a developmental study of attentional affective responsiveness. *Canadian Journal of Psychology, 43*, 230-246.