



## INFANT'S EXPRESSION AND PERCEPTION OF EMOTION THROUGH VOCALIZATIONS

*Yoko Shimura\* and Satoshi Imaizumi\*\**

**\*Department of Education, Saitama University  
255 Shimo-Okubo, Urawa, Saitama, 338 JAPAN**

**\*\*Research Institute of Logopedics and Phoniatics  
Faculty of Medicine, University of Tokyo  
7-3-1 Hongo, Bunkyo-ku, Tokyo, 113 JAPAN**

### ABSTRACT

This paper examined two hypotheses; (1) infants produce voices necessary for emotional communication; and (2) Infants interpret emotions contained in their own vocalizations. Voice samples produced by six infants at 6, 9, 12 and 17 months of age were perceptually evaluated by 79 adults (aged 20-22) and 31 children (aged 2-6) using rating scales representing emotions. The following results were obtained. 1) Even six month old infants who have not yet developed a language could produce various voices necessary for emotional communication through nonlinguistic aspects of voice. 2) Both the adult and child listeners perceived rich contents of emotions from the voice samples recorded even at 6 month of age. 3) More factors were needed to account for the rating scores given by the children than by the adults via factor analyses. 4) Children seem to be more sensitive in perception of emotions from infants' vocalization than the adults. These results support the hypothesis that "infants begin to communicate through nonlinguistic aspects of voice at very early stage of their life."

### I. INTRODUCTION

There is a hypothesis that "infants begin to communicate through nonlinguistic aspects of voice rather than linguistic aspects at very early stage of their life." One way to test this hypothesis is to observe scenes of interaction between the infant and the parent and collate the contents of such interaction with the infant's vocalization or expression. Several attempts to explore this field have been made so far. Another way to test the hypothesis is to determine experimentally whether conditions necessary for communicating through nonlinguistic aspects of voice exist at an early stage of the infant's development. We adopted the latter method, defined "emotion" as "information communicated through nonlinguistic aspects of voice" and have conducted several experiments. The object of our study was to confirm some of the following conditions for communication through nonlinguistic aspects; (1) infants can produce voices necessary for communication, (2) the parents as well as surrounding adults can interpret emotions contained in the infant's vocalizations with a certain degree of regularity, and (3) Infants can also interpret emotions contained in their own vocalizations at some stage of their life.

For this purpose, we conducted an experiment on adult perception of an infant's voice samples excluding crying voices, and reported that adults can perceive a rich variety of emotions such as "pleasure vs. discomfort" even

in voices made by a six-month-old infant. We also tried to find out whether an adult's experience in child-rearing influences her perception and whether her assessment varies depending on her being informed in advance of the context in which the infant vocalized. We found that, although there was specificity dependency on experience, that is, a significant difference due to the adult's child-rearing experience in the evaluation of specific items, there was also universality in the perception of emotions. The reports showed that there were infants who satisfied at least condition (1) and that condition (2) was also satisfied although there was a difference between adults in perception and in dependency on the vocalization situation.

In our present study, we increased the number of infants to demonstrate that condition (1) is not the special ability of special infants, and examined the development of the infant's expression of emotions with age and individual differences. And then, as a preliminary study of the condition (3), we tested how young children interpret emotions contained in infants' vocalizations.

### II. METHOD

#### 2.1 Experiment I

The voice samples used were recorded once a month from six Japanese infants (five male, one female) while they grew from 6 months to 17 months old. They were being raised in households where standard Japanese was spoken. It was confirmed each month through a questionnaire that the infants showed normal development and behavior.

Recording was made when they were playing in interaction with the mother in their homes.

Among 517 voice samples used, 153 samples were recorded at 6 months of age, 165 samples at 9 months of age, 120 samples at 12 months of age and 250 samples at 17 months of age.

The listeners participated in Experiment I were 16 university students with normal hearing whose mother tongue was Japanese. They rated each voice sample using 12 5-point rating scales related to emotions, listed in Table 1. In selecting these scales, several voice researchers preliminarily assessed some of voice samples and wrote down terms describing non linguistic contents conveyed by each voice sample. Of the terms listed out, nine relatively independent items (Items 1-9 in Table 1) were selected as the basic rating scales, and three terms, "speaking", "singing" and "crying", representing the manner of vocalization, were included.

The experiments were conducted in a quiet room where sound was transmitted via a loudspeaker.

Table I. Twelve 5-point-rating unipole scales used in Experiment I.

1 Happy	7 Rejecting
2 Pleased	8 Angry
3 Demanding	9 Sad
4 Amaeru	10 Singing
5 Laugh	11 Speaking
6 Frightened	12 Crying

+---+---+---+---+---+ Pleased  
0 1 2 3 4 5

## 2.2 Experiment II

Out of 517 voice samples used in Experiment I, 14 voice samples produced by an infants recorded at 6, 9, 12 and 17 months of age were selected. Those samples were selected because those covered a wide variety of emotional profile based on the results of Experiment I.

Perceptual rating was carried out using seven 9-point-rating dipole scales as shown in Table II.

Table II. seven 9- or 3-point-rating dipole scales used in Experiment II.

Happy	vs	Sad
Laugh	vs	Crying
Pleased	vs	Frightened
Demanding	vs	Rejecting
Amaeru	vs	Angry
Singing	vs	Speaking
Secret talk	vs	Shout

Happy +---+---+---+---+---+ Sad  
-4 -3 -2 -1 0 1 2 3 4

Listening subjects participated in this experiment were 79 adults (aged 20-22) and 32 children (aged 4.1).

For the children, a two-stage rating was carried out through an interview for each children. Presenting one voice sample through a loud speaker, an experimenter asked the child if he/she assume that the baby producing the voice was expressing "happy" or "sad" or "neutral" emotion. If the child judged "happy", then the experimenter asked how strongly "happy" the baby was. The child was asked to select one circle whose diameter represents the degree of "happiness" from "1(slightly)" to "4(extremely)" grades.. Among 32 children, only 13 children aged 4.8 in average could judge the degree for the 9 scales.

Rating scores given by the adult and child listeners were analyzed by a factor analysis and an analysis of variance.

## III. RESULTS & DISCUSSION

### 3.1 Experiment I

All the rating scores were analyzed by an ANOVA to test the significance of individual differences between infants, differences due to their age, and their interaction for each rating scale. Differences between infants were significant with respect to all the rating scales. The differences due to age were significant only with respect to two scales: "seeking affection (Amaeru)" ( $p < 0.0044$ ) and "speaking" ( $p < 0.0001$ ). The interaction was significant for all the scales ( $p < 0.01$ ).

The above results suggest that the principal factor affecting the rating of emotion-related characteristics of the infants vocalization is individual differences among the

infants and that the differences due to their age are not very large.

A principal component analysis was carried out to extract a few essential components from the rating scores on the 12 scales. Three principal components were extracted. Their accounting for rates in percentage were 54%, 13% and 1%, adding up to 68% in total.

Factor loading of the rating scales after Varimax orthogonal rotation is shown in Fig. 1. From Fig. 1, it can be seen that Factor 1 represents emotion relevant to "crying (0.89) / frightened (0.86) / sad (0.83) vs. happy (-0.49) / speaking (-0.46)" and that Factor 2 represents emotion relevant to "laughing (0.93) / pleased (0.83) / happy (0.69) vs. demanding (-0.47) / speaking (-0.24)". From factor loading, Factor 3 represents emotion relevant to "speaking (0.75) / seeking affection (0.69) / singing (0.63) vs. rejecting (-0.27) / frightened (-0.25) / angry (-0.22)."

Main effects of individual infant, infant's age and their interaction on those factors were tested by an analysis of variance. With respect to Factor 1, the effect of individual infant ( $p > 0.0001$ ) and in the interaction ( $p < 0.0001$ ) between age and infant were significant. With respect to Factor 2, the effects of infant ( $p > 0.0001$ ) and the interaction ( $p < 0.0002$ ) were significant. With respect to Factor 3, the effects of infants ( $p > 0.0001$ ) and the interaction ( $p < 0.0001$ ) were significant. Age has significant effects only through the interaction with individual infant.

The above results suggest the following: (1) Factor 1 can be considered as information pertaining to "crying / frightened / sad vs. happy / speaking"; (2) Factor 2 can be considered as information pertaining to "laughing / pleased / happy vs. demanding / speaking"; (3) Factor 3 can be considered as information pertaining to "speaking / seeking affection / singing vs. rejecting / frightened / angry"; and (4) For all the three factors, the main effect of individual infants is significant, but infant's age itself has no significant effects. Infant's age had a significant effect only through the interaction with individual infants. (5) These results suggest that there are significant differences among individual infants, in emotional expression through vocalizations. However, age related changes in emotional expression are not common among individual infants.

In order to see how each infant's utterances change with age, the factor scores of voice samples were analyzed. Fig. 1 (a) and (b) show two examples. In the figures, the circle represents the 75% confidence area of voice samples recorded at a particular months of age.

In the case of Infant 1, shown in Fig. 1 (a), Factor 1 varies little with age. With regard to Factor 2, both Factors 1 and 2 tend to shift to the negative domain. Although the number of voice samples perceived as "speaking" increased when the infant's months of age increased, there are many voice samples related the emotional contrast between "pleasant vs. discomfort" is perceived even at six months of age.

In the case of Infant 2, shown in Fig. 1 (b), the scatter diagram expands with age with regard to Factor 1. With regard to Factor 2, there is some dispersion due to age difference. However, a comparison of samplings from infants at 6 or 9 months of age and those at 12 or 17 months reveals that, in the scatter diagram, the domain in which Factor 1 is negative and Factor 2 is positive decreases with age and progresses to a domain in which both Factors 1 and 2 are negative. The difference in the scatter diagrams between 6 months and 17 months is particularly noticeable.

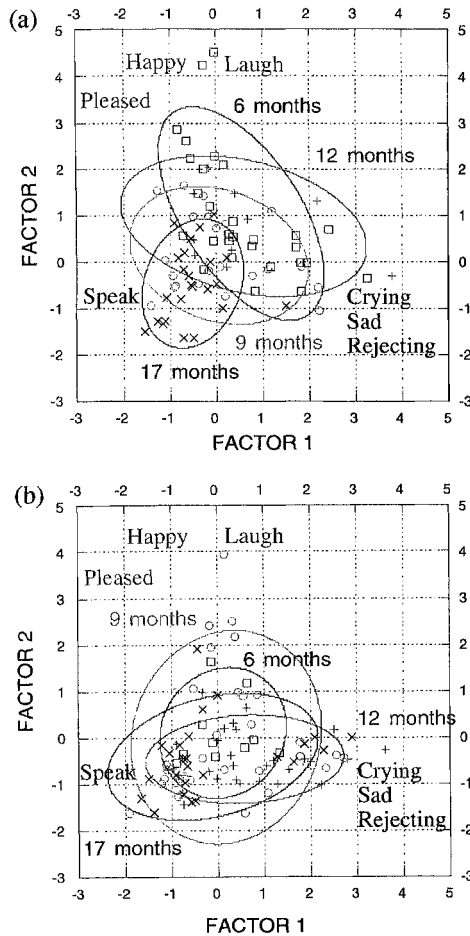


Figure 1. Scatter diagram of the factor scores on F1 and F2 extracted from adult listeners. (a) Infant 1. (b) Infant 2.

The results of analysis of the other four infants showed that for each infant, there was only small shift or expansion of the scatter diagram with age, although the pattern of change differed with each infant.

### 3.2 Experiment II

A factor analysis was carried out to test how many factors are necessary to account for the rating scores given by the adult and child listeners.

Fig. 2 shows the relationship between the number of factors extracted by the factor analyses and the accumulated accounting for rate. The adult listeners put 61% weight (accounting for rate) on Factor 1, while the children groups put only 33%. Principal three factors for the adult listeners accounted for 80% out of the total variance in the rating scores, while four factors for the child listeners accounted for only 71%.

This result suggests two possibilities. The first possibility is that more factors are needed to account for the total variance in the rating scores given by the children than by the adults. The second possibility is that the child than by the adults. The second possibility is that the child listeners gave more unstable responses in rating experiment.

Figures 3 (a) and (b) show the factors extracted from the rating scores given by the adult listeners. Fig. 3 (a) indicates that Factor 1 represents "happy, pleased, laugh" vs "sad, frightened, crying." Factor 2 represents "singing" vs "speaking."

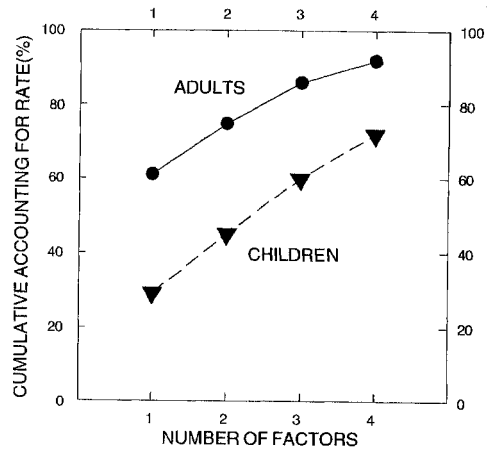


Figure 2. Relationship between the number of factors extracted by factor analyses and the accumulated accounting for rate.

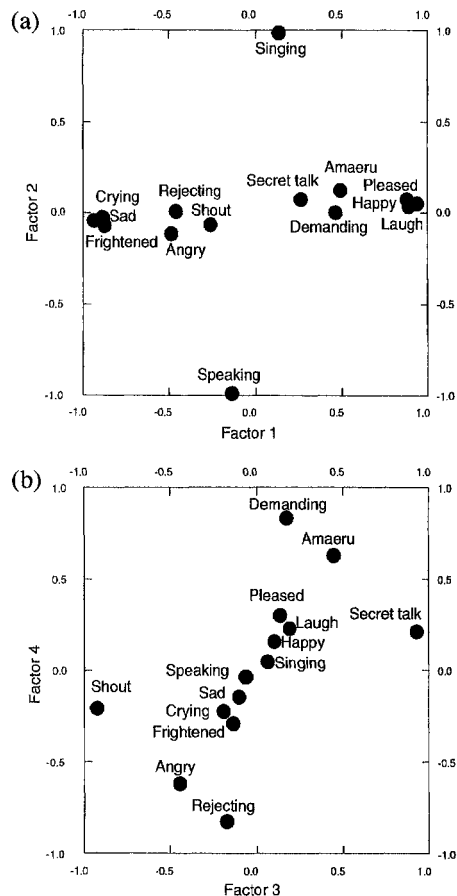


Figure 3 Factors extracted from the rating scores given by the adult listeners. (a) Factor 1 vs. 2. (b) Factor 3 vs. 4.

represents "secret talk" vs "shout," and Factor 4 represents "demanding, amaeru" vs "rejecting, angry."

On the other hand, for the child listeners, Factor 1 as shown in Fig. 4 (a), represents "singing, laugh, pleased, demanding" vs "speaking, crying, frightened, rejecting" Factor 2 represents "secret talk, demanding" vs "shout, rejecting." Factor 3, as shown in Fig. 4 (b), represents "amaeru" vs "angry." Factor 4 represents "happy, pleased" vs "sad, frightened."

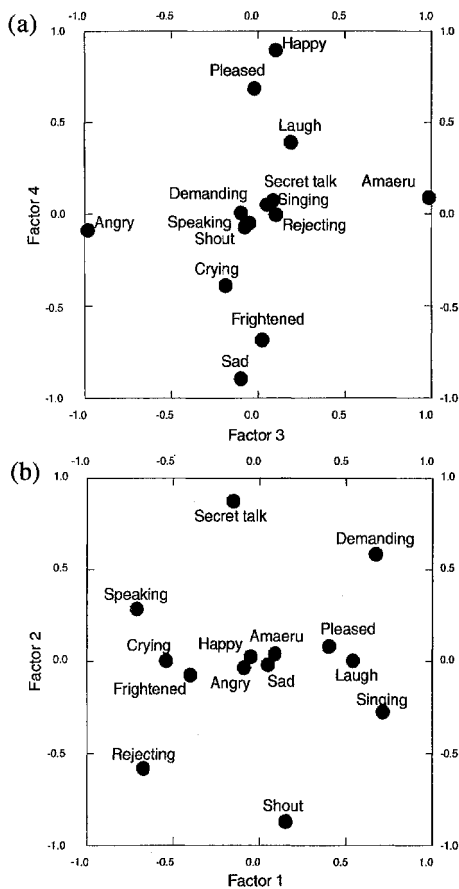


Figure 4. Factors extracted from the rating scores given by the child listeners. (a) Factor 1 vs. 2. (b) Factor 3 vs. 4.

If the child listeners responded in an unstable way, then the factor loading must be small and scattered only around the center of Figure 3. Fig. 3 shows that most scales have large factor loading with large inter relationships among the 12 rating scales. This result supports, among the two possibilities mentioned above, the first possibility, that is, more factors are needed to account for the total variance in the rating scores given by the children than by the adults.

One difference between the adult and child listeners is the structure of Factor 1. The emotional contrast between "happy," vs "sad" is included in Factor 1 for the adults, but it is independent and form Factor 4 for the children. Furthermore, for the adults, contrasts between the manners of vocalization are extracted as the factors, such as Factor 2 representing "singing" vs "speaking" and as Factor 3 representing "secret talk" vs "shout." On the other hand, for the children, contrasts between the manners of vocalization are always associated with some emotional contrasts. For instance, Factor 1 represents "singing, laugh, pleased, demanding" vs "speaking and crying with frightened and rejecting emotions", and Factor 2 represents "secret talk to demand" vs "shout to reject."

These results suggest the following. 1) The adult listeners tend to judge emotional contents in a categorical way as a "pleasant" sign or a "discomfort" sign. This contrast was represented by Factor 1, which accounted for 61% of the total variance in the rating scores. Two other factors related to the contrasts between the manners of vocalizations without any strong association with emotions. 2) The child listeners, on the other hand, seemed not to classify the emotional contents in a categorical way.

into a "pleasant" sign or a "discomfort" sign. Although the children also perceived the contrasts between the manners of vocalization, those contrasts were always associated with some emotional contrasts. The children seemed to be more sensitive than the adult listeners in perception of emotions. They might not be good at categorizing the perceived emotion compared to the adults.

Even six-month-old infants can produce various voices which induce consistent interpretations or responses in adult or child listeners about the baby's emotional states. Even if the induced responses or interpretation of the emotional contents might not be the same as the baby's actual emotional state, we can conclude that surrounding people always can interpret baby's vocalizations in a lawful way. This is important because lawful responses by the surrounding people may induce baby's notice on the social meanings of his/her own vocalization.

#### IV. CONCLUSION

Developmental aspects of the infant's ability to express and perceive emotions through vocalization are examined through perceptual rating experiments by adult and child listeners. Following results were obtained. 1) Even six-month-old infants who are yet to develop their language-using ability can vocalize various voices necessary to express the emotional contrast at least between "pleasant" and "discomfort". 2) The adult listeners tend to judge emotional contents in a categorical way as a "pleasant" sign or a "discomfort" sign. This contrast was represented by the first principal factor, which accounted for 61% of the total variance in the rating scores. The adults were sensitive to the contrasts between the vocalization manners without any strong association with emotions. 3) The child listeners, on the other hand, seemed not to classify the emotional contents in a categorical way, into a "pleasant" sign or a "discomfort" sign. Although the children also perceived the contrasts between the vocalization manners, those contrasts were always associated with some emotional contrasts.

These results suggest that children seem to be more sensitive than the adult listeners in perception of emotions. Even six-month-old infants can produce various voices necessary for communication through emotional aspects of voice.

#### Acknowledgment

We are greatly indebted to Dr. Itsuro Yamanouchi, honorary director of Okayama National Hospital, for his valuable guidance. This research was partially supported by Grant-in-Aid for Scientific Research for Priority Areas on "Cognitive and Linguistic Development," No. 06205102, Ministry of Education, Science and Culture, Japan.

#### Reference

- 1) Shimura, Y., Imaizumi, S., et. al. "Infant' Vocalization Observed in Verbal Communication: -Acoustic Analysis-" ICSLP90, Proceeding, pp.1329-1332, 1990.
- 2) Shimura, Y. and Imaizumi, S. "Emotional Expression by Infant through Vocalization -Perceptual Evaluation-" The Japan Journal of Logopedics and Phoniatrics, Vol. 33, No. 4, pp. 325-332. 1992.
- 3) Shimura, Y. and Imaizumi, S. "Listener and Context Effects in Perception of Emotional Aspects of Infants' Vocalization " The Japan Journal of Logopedics and Phoniatrics, Vol. 34, No. 4, pp. 417-424. 1993.
- 4) Toda, S., Bornstein, M., et. al. "The Relationship Between Mother's Behavior and Child's Play and Language Development," Proceeding, Japan Society of Development Psychology, 1214, 1991.