Perceptual Evaluation of Emotional Aspects of Infants' Vocalizations

Yoko Shimura,* Satoshi Imaizumi, and Toshimoto Shutou*

Introduction

Infants' vocalizations have been investigated mainly to address some linguistic issues in speech acquisition. On the other hand, few studies have been made on the emotional aspects of infants' vocalizations. Exceptionally, vocalizations in "cry" have been investigated for medical purposes to detect brain dysfunctions by analyzing acoustic characteristics[1-5]. Development of emotional expression through voice has not been addressed in the studies on "cry".

Several reports have been made on the emotional aspects of adult speech[6-8], and these have indicated that there are little variations in emotional expression/perception through voice among different languages[9]. Several other studies, analyzing the prosodic patterns of infant-oriented mothers' speech (motherese) across different languages, found some common characteristics[10-14], for instance, 1) high and wide range of the fundamental frequency of voice, 2) slow speaking rate, and so on. Several studies found that infants prefer those acoustic properties of the "motherese"[15,16], which are believed to be relating to "affection", "favor" and "tender."

Those results suggest that it may be possible to hypothesize that emotional aspects of voice have higher universality when compared to the phonotactic aspects of voice. If the emotional aspects of voice are rather universal, their prototype might be easily found in infants' vocalizations.

In this paper, infants' vocalizations recorded during mother/infant interactions were analyzed through perceptual/acoustical experiments to examine how adult hearing subjects perceived emotional profile of infants' vocalizations.

Method

The voice samples analyzed here were recorded from a Japanese male infant once a week in the period of 6 to 17 months after birth. The infant has been brought up by his parents who speak standard Japanese. Recording was carried out when the infant played with his mother. No crying was included.

Two voice corpora, A and B were prepared. Corpus A consisted of 14 voice

* Department of Education, Saitama University

Table I. Five conditions for selecting voice samples

1) short single vowel
2) long single vowel
3) monosyllable with one consonant
4) repetition of monosyllables
5) repetition of polysyllables

Table II. Terms used for rating by the successive category method

Japanese	English equivalent
1) Gokigen:	Pleasure
2) Hanasu:	Speaking
3) Ureshii:	Happy
4) Warau:	Laughing
5) Amaeru:	Requesting affection
6) Utau:	Singing
7) Youkyusuru:	Requiring
8) Kyohisuru:	Refutation
9) Kowagaru:	Fear
10)Kanashii:	Sad
11)Naku:	Crying
12)Okoru:	Angry

1	2	3	4	5
None	Slightly	Rather	Very	Extremely

Figure 1. A rating scale with 5 successive categories.

samples: 7 of which were selected from recordings at 6 months and 7 selected at 17 months of age. Corpus B consisted of 28 voice samples: 7 at 6 months, 9 at 9 months, 5 at 12 months and 7 at 17 months of age. At least one sample satisfying one of the conditions listed in Table I was selected at each month of age.

Perceptual judgments were carried out using a 5-point successive category method. The terms used were shown in Table II. The rating scale for "Pleasure", for instance, is shown in Figure 1.

The hearing subjects were 92 normally hearing students. Half of them judged the first corpus, and other half judged the second corpus.

The perceptual judgments were carried out in a quiet class room. Each voice sample was presented five times through a high-fidelity speaker system at the most comfortable loudness level. The hearing subjects judged each sample by selecting one category out of five (None, Slightly, Rather, Very, Extremely) for each term.

Results

Extracted Principal Factors

A factor analysis was carried out for the ratings by the principal factor analysis method. Figure 2(a) shows the results for the first corpus consisting of 14 voice samples recorded at 6 and 17 months of age. Two significant factors, whose eigen values were larger than 1.0, were extracted. The rates accounting for were 49.6% and 21.4% respectively, and 71.1% totally. Figure 2(a) shows that Factor 1 highly correlates with the positive terms such as "Pleasure", "Happy", "Laugh", "Amaeru", "Singing" and "Speaking", whereas Factor 2 correlates with the negative terms such as "Refutation", "Fear", "Sad", "Cry", and "Angry." "Request" correlated with both factors.

Figure 2(b) shows the result of the factor analysis for the second corpus consisting of 28 voice samples recorded at 6, 9, 12 and 17 months of age. For this corpus, three significant factors were extracted, and their rates accounting for were 56.3%, 14.6% and 8.4% respectively, and 79.3% totally. Factor 1 correlated with the positive terms, whereas, Factor 2 correlated with the negative terms just as in the results obtained for the first corpus. Factor 3 correlated only with "Sad".

A three way analysis of variance was carried out for three factors: voice sample, month and term. The results showed that these three factors and their interactive effects were all significant ($p<0.01$).

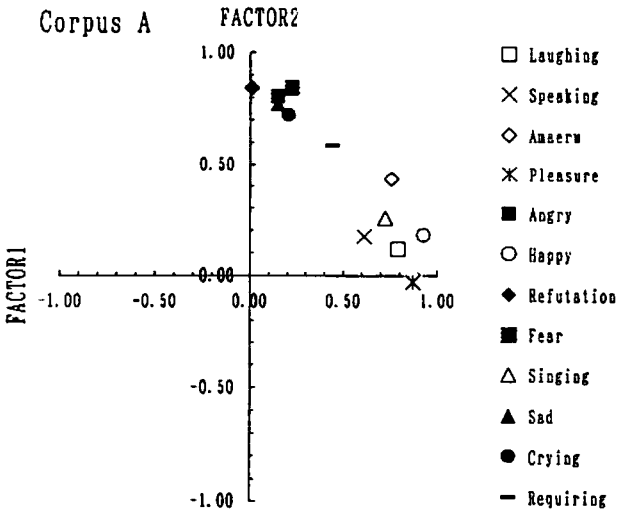

Figure 2 (a)

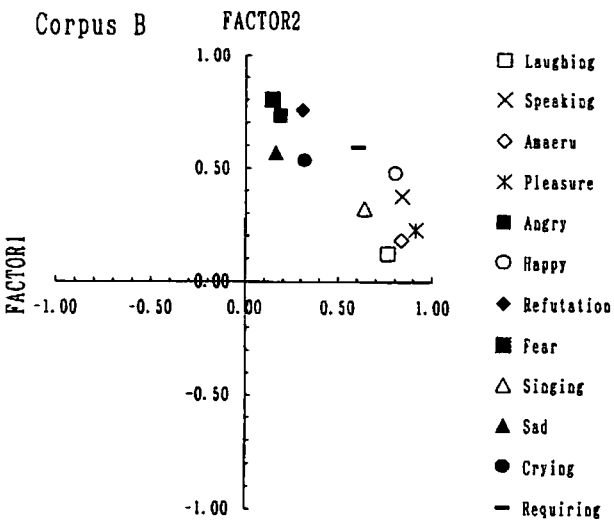

Figure 2 (b)

Figure 2. Factors extracted from the factor analysis for Corpus A (a) and B (b).

Emotional Profile of Infants' Voice

Figures 3(a) and 3(b) show the ratings averaged between 46 hearing subjects for 7 voice samples recorded at 6 and 17 months of age obtained from corpus A. Figure 3(c) shows the ratings averaged between 7 voice samples at 6 and 17 months respectively for the 12 terms. As shown in these figures, the ratings for the voice samples at 17 months of age tend to be higher and their variations tend to be larger than for those at 6 months of age. There were significant differences ($p<0.01$) between the two months for all the terms except "Singing" and "Sad."

For the voice samples at 6 months of age, the ratings were high for "Speaking"(2.99), "Amaeru"(2.45) and "Pleasure"(2.17), whereas the ratings were relatively low for negative terms such as "Angry"(1.31), "Fear"(1.26) and "Refutation"(1.43). At 17 months of age, although the rating variations between the voice samples were large, there were some voice samples having high ratings for negative terms such as "Angry"(1.77), "Fear"(1.45) and "Refutation"(1.88). The ratings for "Amaeru"(1.99) tended to decrease compared whit samples from 6 months of age.

Figure 4 shows the ratings averaged between the voice samples at each month of recording from corpus B.

At 6 months of age, the highest three terms were "Speaking" (2.21), "Requiring" (2.16) and "Amaeru" (1.89). At 9 months of age, they were "Speaking" (2.71), "Pleasure" (2.21) and "Amaeru" (2.17). At 12 months of age, they were "Speaking" (2.70), "Requiring" (1.89) and "Refutation" (1.78). Whereas, at 17 months of age, they were "Speaking" (2.52), "Angry" (2.10) and "Pleasure" (2.02).

The "Angry" and "Cry" scores were significantly higher ($p<0.01$) at 17 months of age than at 6, 9, and 12 months of age. On the other hand, the "Amaeru" scores decreased across the months (6>12, 9>12, 6>17, 9>17, 12>17, $p<0.01$).

Acoustic Characteristics

Preliminary acoustic analyses showed the following tendencies. 1) Voice samples assigned high "Speaking" scores tended to be polysyllable repetitions having a falling pattern for the fundamental frequency. 2) Voice samples assigned high "Amaeru" scores tend to have a slow rising and slow falling pattern for the fundamental frequency. The "Amaeru" scores tended to decrease for the voice samples of polysyllabic repetitions. 3) High scores for negative emotions such as "Angry" or "Refutation" seemed to be assigned to voice samples with abrupt changes in fundamental frequency. However, there were some exceptions. Further concrete results on acoustic characteristics conveying emotions will be obtained through a synthesis-by-analysis experiment which is now being carried out.

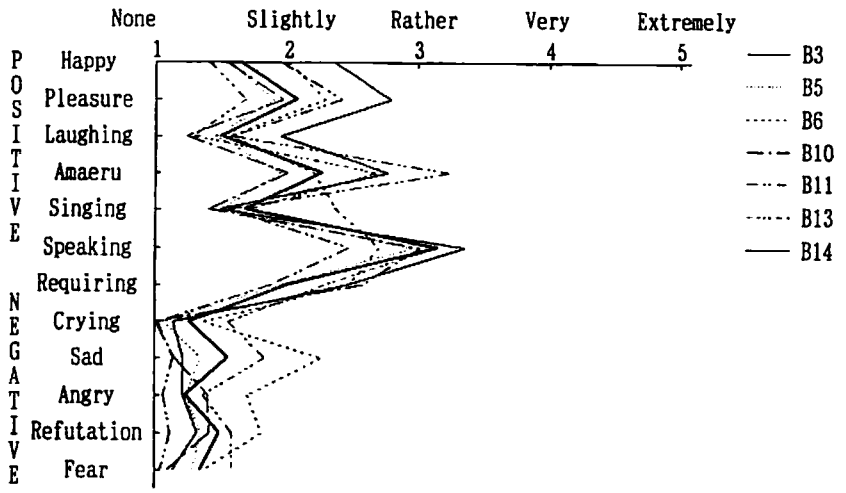

Figure 3(a)

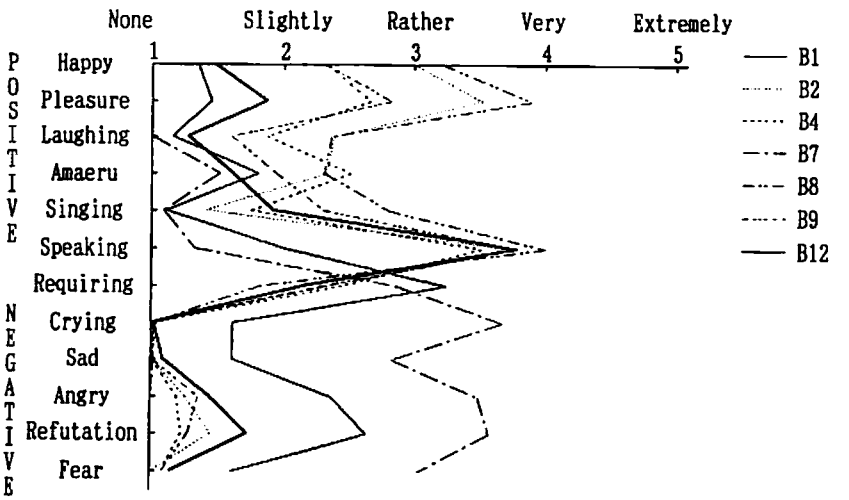

Figure 3(b)

Figure 3. Emotional profile for Corpus A. (a) 7 voice samples recorded at 6 months of age. (b) 7 voice samples at 17 months of age.

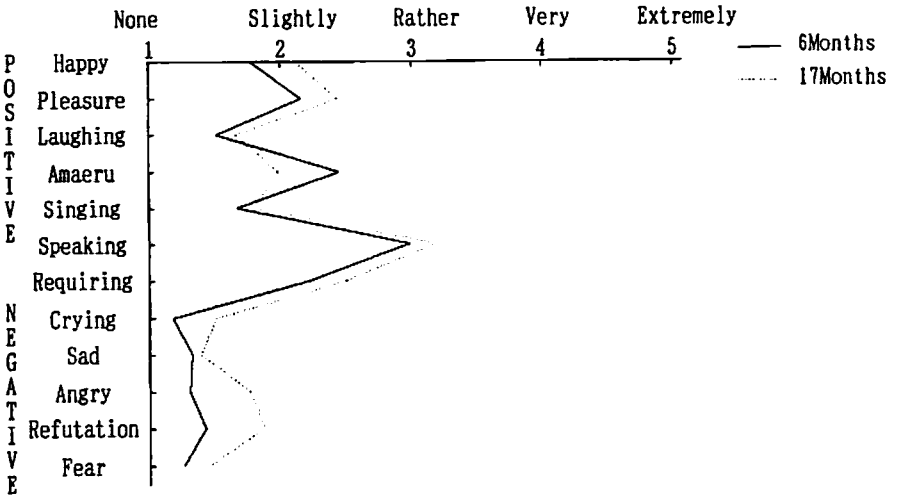

Figure 3 (c). Average profiles at 6 and 17 months of age for Corpus A.

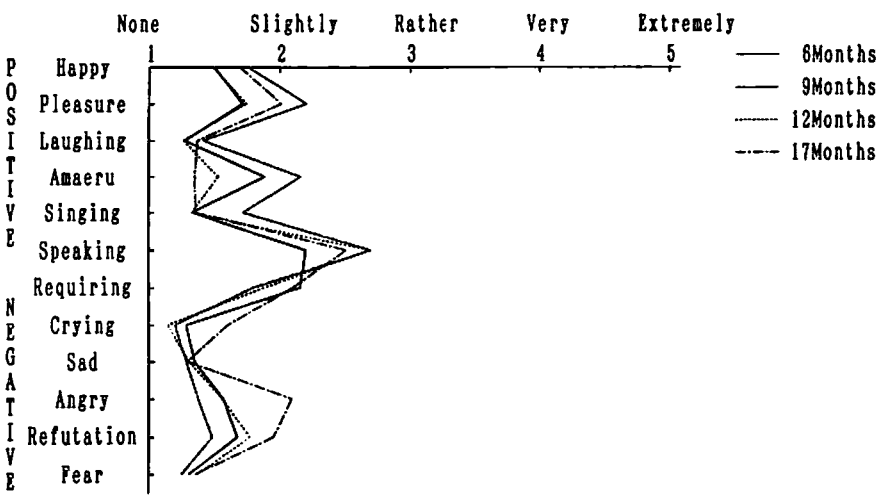

Figure 4. Average emotional profiles at 6, 9, 12 and 17 months of age for Corpus B.

Discussion

The present study of the emotional aspects of infant voice showed the following results.

1) From corpus B, consisting of voice samples recorded at 6, 9, 12 and 17 months of age, three principal factors were extracted. Factor 1 represented the positive emotions such as "Pleasure", "Happy", "Laugh", "Amaeru", "Singing" and "Speaking", whereas Factor 2 represented the negative emotions such as "Refutation", "Fear" and "Angry." Factor 3 represented the "Sad" emotion. "Requiring" correlated with Factors 1 and 2.

2) The results of the three way analysis of variance showed that there were significant differences ($p<0.01$) between the voice samples, age months and the emotional terms used for rating.

3) For the voice samples at 6 months of age, the average rating score was high for "Speaking", "Amaeru" and "Pleasure", whereas it was relatively low for the negative terms such as "Angry", "Fear", and "Refutation." At 17 months of age, the variation in ratings between the voice samples became large, and some voice samples were rated high also for negative terms such as "Angry", "Fear" and "Refutation." However, the ratings for "Amaeru" decreased compared to those at 6 months of age. In summary, the "Angry" and "Cry" scores were significantly higher ($p<0.01$) at 17 months of age than at 6, 9, and 12 months of age, whereas the "Amaeru" scores decreased across the months (6>12, 9>12, 6>17, 9>17, 12>17, $p<0.01$).

These results indicate that there are rich acoustic characteristics conveying emotions even in infant voice samples at 6 months of age, although this richness seems to develop at least during the first 6/9 months of age. Especially, the expression of negative emotions such as "Refutation" and "Angry" through non-crying voice seems to develop. On the other hand, positive emotional aspects such as "Pleasure" and "Happy" and "Amaeru" are exhibited in the voice at 6 months of age. The "Amaeru" aspect of voice decreases at 12/17 months of age compared to 6/9 months of age.

These results indicate that not all aspects of emotional expression through voice develop synchronously or simultaneously.

Conclusion

The emotional aspects of infants' vocalizations recorded in mother/infant interactions were analyzed through perceptual experiments. Crying was not included. Results showed that most vocalizations recorded at 6/9 months of age were rated highly for terms expressing positive emotions such as "amaeru" and "pleasure". On the other hand,

those recorded at 12/17 months were rated highly not only for positive emotions like "happy" and "pleasure", but also for negative emotions such as "refutation" and "anger." The "Amaeru" scores decreased across the months. These results indicate that infants' vocalizations at very early months of age contain a wide variety of acoustic characteristics which may convey various emotional expressions, but that the ability to express negative emotion without crying might still be developing at least during the first 6/9 months of age. Not all aspects of emotional expression through voice seemed to develop synchronously or simultaneously.

Acknowledgement

This report was partially supported by a Grant-in-Aid for Scientific Research from the Japanese Ministry of Education, Science and Culture.

References

1) Wasz-Hockert, O., Lind, J., Vuorenkoski, V., Partanen, V., and Valanne, E.: The Infant Cry: A Spectrographic and Auditory Analysis. Clinics in Developmental Medicine, No. 29, Spastics International Medical Publications in Association with William Hainemann Medical Books Ltd., 1968.
2) Michelsson, K. and Wasz-Hockert, O.: The Value of Cry Analysis in Neonatology and Eary Infancy. Infant Communication : Cry and Early Speech, College-Hill Press, 152-182, 1980.
3) Toden, C-J. and Koivist, M.: Acoustic Analysis of the Normal Pain Cry. Infant Communication : Cry and Early Speech, College-Hill Press, 124-151, 1980.
4) Michelsson, K., Raes, J., Toden, C-J. and Wasz-Hockert, O.: Sound Spectrographic Cry Analysis in Neonatal Diagnostics. Anevaluative Study. Journal of Phonetics, 10: 79-88, 1982.
5) Freudenberg, R. P., Driscoll, J. W., and Stern, G.S.: Reactions of Adult Humans to Cries of Normal and Abnormal Infant. Infant Behavior and Development, 1: 224-227, 1978.
6) Sherer, K. R., Koivumaki, J., and Rosentahal, R.: Minimal Cues in the Vocal Communication of Affect: Judging Emotions from Contentmasked Speech. Jounal of Psycholinguistic Research, 1: 57-64, 1972.
7) Williams, C. E. and Stevens, K.N.: Emotions and Speech : Some Acoustical Correlates. The Journal of Acoustical Society of America, 52, 4: 1238-1250, 1972.
8) Sherer, K. R.: Vocal Correlates of Emotion. Handbook of Psychophysiology: Emotion and Social Behavior, Chichester: Wiley, UK. 1989.
9) Sherer, K. R.: Vocal Affect Expression : A Review and a Model for Future Research. Psychological Bulletin, 99, 2: 143-165, 1986.
10) Fernald, A.: Intonation and Comminucation Intent in Mothers' Speech to Infants: Is the Melody the Message?. Child Development, 60: 1497-1510, 1989.
11) Fernald, A., Taeschner, T., Dunn, T., Papousek, M., Boysson-Bardies, B., and Fukui,I.: A Cross-language Study of Prosodic Modifications in Mothers' and Fathers' Speech to Preverbal Infants. Journal of Child Language, 16: 477-501, 1989.
12) Grieser, D.L., and Kuhl P.K.: Maternal Speech to Infants in a Tonal Language: Support for Universal Prosodic Features in Motherese. Developmental Psychology,

24: 14-20, 1988.
13) Shimura, Y., Imaizumi, S., Saito, K., Ichijima, T., Gauffin, J., Halle, P., and Yamanouchi, I.: Infants' Vocalization Observed in Verbal Communication: Acoustic Analysis. International Conference on Spoken Language Processing. Proceeding, 2: 1329-1332, 1990.
14) Shimura, Y. and Yamanouchi, I. : Sound Spectrographic Studies on the Relation Between Motherese and Pleasure Vocalizarion in Early Infancy: Acta Paediatrica Japonica, 34, 3: 259-266, 1992.
15) Fernald, A.: Four-month-old Infants Prefer to Listen to Motherese. Infant Behavior and Development, 8: 181-195, 1985.
16) Fernald, A., and Kuhl P.K.: Acoustic Determinants of Infan Perference for Motherese Speech. Infant Behavior and Development, 10: 279-293, 1987.