Internal Use Only
非公開

TR－A－0139 17

The effects of tempo and pitch on the judgment of infant cries

Taeko TSUKAMOTO and Yoh'ichi TOHKURA

1992. 3.17

ATR視聴覚機構研究所
〒619-02 京都府相楽郡精華町光台2-2 ☎07749-5-1411

ATR Auditory and Visual Perception Research Laboratories

2-2, Hikaridai, Seika-cho, Soraku-gun, Kyoto 619-02 Japan
Telephone: +81-7749-5-1411
Facsimile: +81-7749-5-1408

The effects of tempo and pitch on the judgment of infant cries

Taeko TSUKAMOTO and Yoh'ichi TOHKURA
ATR Auditory and Visual Perception Research Laboratories

A perceptual experiment was performed to examine the effects of pitch and the interactions of pitch and tempo in three categories of cries: hunger, anger, and call. In each category, the duration and fundamental frequency (F_0) of the cry stimuli were systematically varied by signal processing so as to keep their formant structure unchanged. After the training in classifying the cries, the subjects (60 women and 29 men) who had had no experience of infant care were instructed to identify the cry stimuli by making a forced choice among the three categories. The data for 38 women and 17 men who successfully learned the cry classification were analyzed. Results showed that the correct response rates shifted significantly by varying the tempo but rather less when F_0 was varied across all categories. This indicates that tempo is a predominant perceptual cue for the judgment of cries.

At preverbal stages of infancy, cries are an important form of communication and play an important role in the relationships between infants and those who care for them, including their mothers. For a better understanding of this field, it is necessary to study the mechanisms of cry judgement and to reveal how the mother perceives cries and identifies their reasons. There has been no agreement on the issue that mothers are able to identify these cries because of methodological difficulties and inconsistencies of the cry samples used in such investigations. However, a number of previous studies suggest that some distinctive patterns or features are perceived in the cries. In previous studies on the relationships between auditory impressions and the acoustic features of infant cries, it was indicated that cries with a high fundamental frequency (F_0) and an irregular temporal pattern are perceived as unpleasant, annoying, or irritating (Bates, et al., 1979; Wiesenfeld, et al., 1981; Lounsbury & Bates, 1982). Cries with a much higher F_0 are related to urgency, arousal, or distress (Zeskind & Lester, 1978; Zeskind & Marshall, 1985) by the listeners. Most of these studies focused on aversiveness or stress of the mother caused by the cries from a clinical viewpoint. They also aimed at determining what kinds of acoustic feature are important cues for the judgment of cries. Previous acoustic studies of infant cries in established that cries can be characterized by differences in the mean and dynamic range of F_0 (Murry et al., 1983), the temporal patterns of F_0 (Wasz-Höckert et al., 1968; Rosenhouse, 1977; D'Odorico, 1984), and rhythmic or durational patterns in phonation and non-phonation (Wolff, 1967, 1969; Futatsugi, 1979, Kobayashi et al., 1986; Tsukamoto & Katagiri, 1988). These findings suggest that F_0 and the duration are the predominant acoustic parameters for differentiating cries, and that variations of these two parameters produce various perceptual acoustic features, such as pitch, intonation (melody), tempo, or rhythm, which may be used as cues for cry judgment by the mother. However it is unclear how these can be systematically related to the judgment of infant cries.

To discuss this issue in more detail, a useful technique is to produce a variety of infant cries with varying acoustic features by manipulating the fundamental frequency and/or duration systematically by signal processing. This allows the investigation of how manipulating these acoustic parameters affects the judgment of cries. In our previous study (Tsukamoto & Tohkura, 1992), the effects of tempo on cry judgment were examined using cry stimuli with a variety of tempi produced by expanding or compressing the duration of original cries labeled as hunger, anger, and call (i.e., calling for mother-infant interaction or demanding the mother's attention). The results showed that correct response rates shifted significantly corresponding to the degree of time-scale expansion or compression in each category. For the anger stimuli, when the time scale was expanded more than twice, the correct response rate decreased and more than half were classified as hunger cries. On the other hand, when the hunger stimuli were compressed, more than half of the responses were anger cries. In contrast, when the time scale was expanded, the correct response rate for the hunger stimuli increased and became even higher than for the original stimuli. As for the call stimuli, the correct response rate reduced when the time scale was compressed, although the stimuli were still predominantly identified as call cries (about 70%). These results indicate that the cry tempo is a reliable cue for discriminating between hunger and anger cries, but that there are other predominant perceptual cues for the judgment of call cries.

In this study, we designed an experiment to investigate the effects of pitch and the interactive effects of pitch and tempo for judgment of infant cries. The purpose of this experiment is to examine the effects of pitch in different temporal conditions and to establish the relationships between pitch and tempo as perceptual cues. We used natural cries for the original stimuli and employed both men and women's university students as subjects. These subjects were chosen based upon the findings of previous studies which were as follows. In our preliminary experiments which examined the listeners' impressions of cries and their identification of the cry reasons (Tsukamoto, 1991), it was shown that university students could categorize the cries in the same dimensions (violent, urgent, and earnest) as the mothers, whose responses were extracted from their verbal responses to the acoustic features of the cry stimuli. The mothers could identify the cry reasons significantly well even when given a free choice, but showed a tendency to attribute the cries to physiological causes. On the other hand, the university students could also identify anger and call cries as well as or even better than the mothers, although they were not so good at identifying cries produced by physiological conditions of the infant, such as hunger or sleepy. Several studies on the physical responses of both women and men suggest that there is no gender difference in the overall skin potential response (Boukydis, 1985) or in the patterns of skin potential arousal responses (Frodi et al., 1978) for the three types of cries; average, easy, and difficult infants' cries. These physical responses are related to perception and are thought to support our preliminary findings indirectly. We thus considered, at the first step in this study, that it was necessary to examine the responses of subjects with no experience in infant care other than the mothers' responses biased more or less due to their own infant care experiences.

Method

Cry Samples

Cries were videotaped (SONY SL-F1; SONY CCD-V8AF2) when observing mother-infant interactions. The observations were carried out by visiting their home biweekly for one hour on infants ranging from 3 weeks to 12 months. The observers were several university students majoring in psychology, and one of the authors. Records were made by each observer of the behavior of the mothers and infants, and records were made of the infant's "life rhythm" in which the mothers checked the time of changing diapers, suckling or feeding, sleep, awareness, and physical conditions of their infants the day before observation. In order to classify the cries, the observers evaluated them according to the infants' situations by watching repeated playbacks of the videotapes and investigating all the other material. Cry category classification was performed by a consensus (more than 85%) between the observers. Thus, the cries were eventually classified into several categories according to the situations which caused them. However, the cries were generally observed to be produced not only by simple situations, but also for complex physical or psychological reasons. from these several categories, we finally selected three; hunger, anger, and call, for our perceptual experiments. These categories were frequently observed in the infants' situations at their homes.

Regarding the infants' ages, we chose cries produced by infants of six months or younger, because various changes occur at around six months of age, such as a changing role and function of the cry in communication behavior (Tsukamoto, 1985), maturational changing of the brain structure (Lester, 1985), and developmental advances in voluntary motor control (Kent & Hodge, 1990).

Acoustic Features of Samples

The samples were digitized by a computer (MASSCOMP 5600) at a sampling frequency of 20 kHz (LPF=10 kHz, - 96 dB/oct) and were analyzed to extract F_0 and the segmental duration. F_0 was measured every 2.5 ms by applying an autocorrelation pitch extractor to the cry waveforms. The duration of the segmental units was measured at a time resolution of 2.5 ms by observing digital spectrograms and the cry waveforms. A segmental unit is defined as a breath group of the infant cry, consisting of all the vocalization occurring during a single respiration. The statistical distributions (maximum, minimum, mean and SD) of F_0 and the segmental durations are shown in Table 1 for the nine cry samples used in this experiment.

The results show that the mean segmental durations of call cries are longer than those of anger and hunger cries, which are almost equal. The mean F_0s of anger and hunger cries are higher than those of call cries on the whole, though there are individual differences in F_0 between the samples.

To allow a more intuitive interpretation of these statistical distributions, histograms of F_0 and the duration are shown in Figures 1 and 2, respectively. In Figure 1 F_0 is quantized every 50 Hz, and its occurrences are normalized by the total counts. Figure 2 shows the distribution of the pitch duration, which is defined as the segmental duration of vocalized parts in which clear harmonics of F_0 can be continuously observed. The pitch duration is quantized every 250 ms and its occurrences are normalized by the total counts. In both figures, it is interesting to note the similarities between the distributions in each category, even across cry samples. These results suggest that there might be acoustic characteristics peculiar to each category, although they are observed globally in their statistical distributions. However, we cannot draw any general conclusions from this because the number of cry samples we used was rather low.

Subjects

The subjects were 60 women and 29 men ranging in age from 18 to 26. None of them had experience of infant care, and most of them were university students. Before the perceptual experiment, all the subjects were trained in cry category identification so as to learn the distinctive features of each category.

Training

Stimuli.

The stimuli used in training were two samples from each category: call cries of YU (15 weeks) and MA (17 weeks), hunger cries of TE (11 weeks) and TE (17 weeks), and anger cries of TO (9 weeks) and MA (12 weeks) which were selected from the nine samples in Table 1. Their overall durations were between 10.6 and 12.6 seconds.

Training procedure.

The training was consisted of three trials: (1) learning, (2) identification, and (3) confirmation of learning. In the learning and confirmation trials, the stimuli were presented twice in succession with their category labels given to the subjects prior to listening. In the identification trial, the same stimuli as those used in the learning procedure were presented twice in random order (12 stimuli in total, inter-stimulus interval (ISI)=3 s). The subjects were instructed to identify them by making a forced choice between the three categories: hunger, anger and call cries. After training, an identification test was carried out on all the subjects. In this test, the same procedure as the identification trial was used, except that the stimuli were presented in a different random order. As a result of this identification test, 55 subjects (38 women and 17 men) who could accurately identify all 12 stimuli were regarded as having been successful in learning the category identification.

Experiment

Original stimuli.

In the experiment, three samples (one from each category) produced by one infant but different from those used in the training were selected from the nine samples in Table 1; call: MA (15 weeks), hunger: MA (12 weeks), and anger: MA (13 weeks). Two original stimuli were generated for each category from each of the three samples with segment units from 6.6 to 8.1 seconds in duration excised from the sample in each category (see Table 2). These stimuli were assumed to carry sufficient acoustic features for the judgment of cries, because their identification rates proved to exceed 85% in our previous experiment (Tsukamoto & Tohkura, 1990).

Stimulus series.

To examine the effects of pitch under different temporal conditions, F_0 and the duration of the original stimuli were changed using a signal reconstruction method based on a short-time Fourier transform (Abe et al., 1989; Griffin & Lim, 1984). The method is capable of compressing or expanding the duration (time scale of the signals) and also modifying the F_0 of the original stimuli. It is important to note that this method is capable of modifying the acoustic features independently. For instance, when transforming the time scale to obtain stimuli with a variety of tempi, all the temporal structures of the stimuli are linearly compressed or expanded, while keeping the segmental frequency structures unchanged. Furthermore, it is also capable of modifying F_0 while leaving both the formant and temporal structures unchanged. Thus, this method is quite useful in producing stimuli with modified durations and/or F_0.

Three time scale transformations were employed: original (the duration of the original stimuli was not changed), compressed (fast tempo), and expanded (slow tempo). The transformation ratios for time scale compression and expansion were chosen differently for the 3 categories based on the results of our previous study on the effect of tempo on cry judgment (Tsukamoto & Tohkura, 1992). For compressed stimuli, the transformation ratios were set to 0.5, 0.5, and 0.67 for call, hunger, and anger stimuli, respectively. For expanded stimuli, the transformation ratios were 1.5, 1.5 and, 2.0, respectively.

As for the pitch, it was found that the F_0 of original stimuli could be increased or decreased by a factor of up to 0.3 (30%) without any noticeable acoustic distortion. Consequently, seven levels were constructed with shift ratios of 1.3, 1.2, 1.1, 1.0 (original), 0.9, 0.8, and 0.7. For each original stimulus, three stimulus series were generated by combining the seven pitch levels and three tempi. Thus, there were 126 stimuli (7 x 3 variations x 2 original stimuli x 3 categories) in total.

Procedure

It was presumed that, in general, the subjects might tend to notice which acoustic parameters of the stimuli have been manipulated when repeatedly listening to stimuli with similar acoustic features. To avoid this experimental distortion, 54 dummy stimuli were added to the 126 stimuli, thus 180 stimuli were prepared in total. These were randomized and divided into two series of 90 stimuli each. Another 10 dummy stimuli were added at the beginning of each series and the total of 100 stimuli was recorded on digital audio tape (SONY DTC-1000ES). For each series, the ISI was set at 3 s, and a beep was presented at the beginning of each stimulus. Three beeps were presented every 10 stimuli as a block sign. Taking the order of stimulus presentation into account, two tapes with different random orders were prepared for each series, and presented to balanced groups of subjects. The subjects were required to identify each stimulus by making a forced choice between the three categories: hunger, anger, and call. The training and the experiment were conducted in a sound booth, and all the stimuli were presented to the subjects with headphones (STAX SRM-1/MK-2).

Results

Effect of Fundamental Frequency

The results were analyzed based upon the response data given by the 55 subjects who successfully learned the cry category identification. The correct response rates for the three stimulus series with different tempi were compared. Here, the correct response to a stimulus was defined as the category of the original stimulus. For each category, there were two kinds of stimulus series which were generated from two different original stimuli with the same tempo. These two series showed similar responses both to variations of tempo and between categories. There were no significant differences in the responses of women and men subjects. To clarify the relationship between the correct response rates and the shift of F_0, all the responses for each stimulus were added up and illustrated in graphs.

Figure 3 shows the correct response rates for call stimulus series. In this figure, the correct response rates were highest for expanded stimuli followed by the original and compressed stimuli. For original and expanded stimuli, the correct response patterns are quite similar; the correct response rates were very high throughout the stimulus series, regardless of the F_0 shift. For compressed stimuli, however, correct response rates decreased (by about 20%) at higher values of F_0. On the whole, the correct response rates were lower than for the other two conditions. In a chi-squared analysis, the correct response rates for compressed stimuli were significantly different with respect to F0, χ^2 $(6, N = 55) = 27.2, p < .001$.

The correct response rates for the hunger stimulus series shown in Figure 4 were highest for expanded stimuli followed by the original and compressed stimuli, similar to the results for call stimulus series. However, the correct response rates increased with F_0 for all three different conditions. This tendency was most apparent for the original condition. Consequently, the F_0 dependency of the correct response rates was shown to be significantly in chi-squared analysis, χ^2 (6, N = 55) = 14.19, $p < .05$.

Figure 5 shows the results for the anger stimulus series. This figure shows that the correct response rates were highest for compressed stimuli followed by the original and the expanded stimuli which both increased as F_0 increased. In chi-squared analysis, the correct response rates differed significantly with respect to F_0 for the original stimuli, χ^2 (6, N = 55) = 19.03, $p < .005$, and for the expanded stimuli, χ^2 (6, N = 55) = 37.57, $p < .001$.

The response properties show that the effects of varying F_0 differ between the three tempi in the judgment of cries. For the call stimulus series, when F_0 increased, the correct response rates decreased whereas the hunger responses increased. As shown in Figure 6, the hunger and anger responses varied according to the tempo.

Relationships between Pitch and Tempo

To examine the relationships between pitch and tempo as perceptual cues for cry judgment, their individual and interactive effects were analyzed for each category. Chi-squared analysis of the results showed that there were significant differences in the main effects of the tempo factor in all three categories: call, χ^2 (6, N = 55) = 14.86, $p < .001$; hunger, χ^2 (6, N = 55) = 101.00, $p < .001$; anger, χ^2 (6, N = 55) = 148.60, $p < .001$. However, the main effects of varying F_0 were only significantly different for the anger stimulus series, χ^2 (6, N = 55) = 14.60, $p < .05$. No interactive effects were noted in any of the categories.

Multiple comparison tests using Ryan's procedure showed significant differences ($p < .05$) as follows: For the call, there were significant differences between the expanded and compressed stimuli, χ^2 (1, N = 55) = 18.29, and also between the original and compressed stimuli, χ^2 (1, N = 55) = 15.69. For the hunger, there were significant differences between the original and compressed stimuli, χ^2 (1, N = 55) = 27.53, between the expanded and original stimuli, χ^2 (1, N = 55) = 6.6, and between the expanded and compressed stimuli, χ^2 (1, N = 55) = 53.94. For the anger, there were significant differences between the compressed and expanded stimuli, χ^2 (1, N = 55) = 39.32, and between the compressed and original stimuli, χ^2 (1, N = 55) = 17.93. There were also significant differences between the shift ratios of 1 (original) and 0.8, and between 1.2 and 0.8 ($p < .05$) for the anger stimulus series. On the whole, these results show that the tempo is more predominant than the pitch in each category.

Discussion

In this experiment, the effect of pitch at different tempo conditions, and the relationships between pitch and tempo as perceptual cues for the judgment of cries, were examined psychophysically. The results show that the correct response rates change significantly when F_0 is shifted. The effects of F_0 on category identification of the cries, however, were different between the three tempi. In particular, the hunger and anger responses shifted depending on whether the tempo was compressed or expanded. This tendency is quite consistent with the results of our previous experiment (Tsukamoto & Tohkura, 1992) which examined the effects of tempo; when the time scale was compressed, the responses shifted to anger cries, and when the time scale was expanded, the responses shifted to hunger cries. The correct response rates for call stimuli were still above 60%, though they reduced significantly when F_0 increased for the compressed stimuli. These results suggest that, as far as the judgment of hunger and anger cries is concerned, tempo is a much more predominant perceptual cue than pitch, which might be used as a secondary cue when the acoustic features were atypical or indistinct. It is assumed that there might be other perceptual cues which are much more predominant for recognizing call cries.

As seen in Figures 1 and 2, acoustic analysis of the cry samples used in this experiment showed that there is a great deal of similarity in the acoustic features of hunger and anger cries, though the anger cries have a slightly higher F_0 and a shorter average pitch duration than the hunger cries. Note that the pitch duration is not consistent with the segmental duration defined by the cycle of infant respiration. As described in our previous study, which used the same cry samples (Tsukamoto & Tohkura, 1992), the hunger cries have a slightly shorter average segmental duration than the anger cries. Accordingly, the tempo of hunger cries is a little faster than that of anger cries. However, the subjects judged a faster tempo as the anger and a slower tempo as hunger cries, both in the previous study and in this experiment. It is of considerable interest that these tendencies are partially consistent with the findings that a high pitch level, a wide pitch range, a loud voice, and a fast tempo are all distinctive features characterizing the emotion of anger (Sherer, 1979, 1982). It is assumed that the subjects might have judged the cry stimuli by the same criterions by which they judge the attitude or emotion of other people. Are there effects related to the fact that the subjects in this experiment had no experience of infant care?

From psychiatric evidence, it is inferred that crying is a part of the biobehavioral affect system which is related to the central and peripheral nervous systems (Wolff, 1967, 1969; Zeskind, 1978; Lester, 1978, 1984; Lester & Zeskind, 1982; Michelsson et al., 1980; Boukydis & Burgess, 1985; Golb & Corwin, 1985). Taking this into account, it is assumed that the subjects discriminate the cries using to the differences in tempo as one of the cues to infer the arousal level or stress of the infants.

In this experiment, we used three categories of natural cries which were produced by one infant and investigated the effects of pitch and tempo on the judgment of infant cries by changing F_0 and the duration of the original stimuli. It is noted that, besides these two factors, there might be other factors in the original acoustic features which should have been manipulated in this study. It is also noted that it is too early to conclude that the effects of pitch and tempo obtained experimentally can be generalized for the judgment of cries. In this study, we focused on some of acoustic features which can be manipulated globally. The tempo of cries was increased or decreased by linearly compressing or expanding the time scale of the stimuli. The pitch of the cries was shifted by fixed ratios for each of the stimuli. Further research should deal with segmental acoustic features, such as formant transitions and pitch contours. In order to manipulate these segmental acoustic features in a sophisticated way, we must develop a high quality cry synthesizer based upon cry production models. These issues should be discussed in further research.

In this view, the further investigation of the mechanisms of infant cry judgement will be important for studying the development of communication patterns between mothers and infants.

References

Abe, M., Tamura, S., & Kuwabara, H. (1989). A speech modification method by signal reconstruction using short-time Fourier transforms. *The Transactions of the Institute of Electronics, Information and Communication Engineers of Japan, J72-D-II, 8*, 1180-1186. (In Japanese)

Bates, J. E., Freeland, C. A. B., & Lounsbury, M. L. (1979). Measurement of infant difficultness. *Child Development, 50*, 794-802.

Boukydid, C. F. Z. (1985). Perception of infant crying as an interpersonal event. In B. M. Lester & C. F. Z. Boukydis (Eds.), *Infant crying: theoretical and research perspective* (pp. 187-215). New york: Plenum Press.

Boukydis, C. F. Z. & Burgess, R. L. (1982). Adult physiological response to infant cries: effects of temperament of infant, parental status and gender. *Child Development, 53*, 1291-1298.

D'Odorico, L. (1984). Non-segmental features in prelinguistic communications: an analysis of some types of infant cry and non-cry vocalization. *Journal of Child Language, 11*(1), 17-27.

Frodi, A. M., Lamb, M., Leavitt, L., Donovan, W., Neff, C, & Sherry, D. (1978). Fathers' and mothers' responses to the appearance and cries of premature and normal infants. *Developmental Psychology, 14*, 490-498.

Futatsugi, T. (1979). Cry characteristics in newborn infants and developmental meanings of the infant cry. *Japanese Journal of Child Psychiatrica, 20*(3), 161-177 (In Japanese).

Golb, H. L. & Corwin, M. J. (1985). A physioacoustic model of the infant cry. In B.M. Lester & C.F.Z. Boukydis (Eds.), *Infant crying: theoretical and research perspectives* (pp. 59-82). New York: Plenum Press.

Green, J. A. & Gustafson, G. E. (1983). Individual recognition of human infants on the basis of cries alone. *Developmental Psychobiology, 16*(6), 485-493.

Griffin, D. W., & Lim, J. S. (1984). Signal estimation from modified short-time fourier transform. *IEEE Transactions on Acoustic Speech & Signal Process, 32*,(2), 236-242.

Kent, R. D. & Hodge, M. (1990). The biogenesis of speech: continuity and process in early speech and language development. In J. F. Miller (Ed.), *Progress in research on child language disorders* (pp. 25-53). Austin, Texas: Pro-Ed.

Kobayashi, K., Oda, T., & Murooka, H. (1986). Characteristic pattern in time series of crying vocal early infant. *Japanese Society of Medical Electronics & Biological Electronics* (In Japanese).

Lester, B. M. (1978). The organization of crying in the neonate. *Journal of Pediatric Psychology, 3*(3), 122-130.

Lester, B. M. (1984). Infant crying and the development of communication. In N. A. Fox & R. J. Davidson (Eds.), *The psychobiology of affective development* (pp. 231-258). Hillsdale, New Jersey: London.

Lester, B. M., & Zeskind, P. S. (1982). A biobehavioral perspective on crying in early infancy. In H. E. Fitzgerald, B. M. Lester & M. W. Yogman (Eds.), *Theory and research in behavioral pediatrics* (pp. 133-180), Plenum Press: New York.

Lounsbury, M. L., & Bates, J. E. (1982). The cries of infants of differing levels of perceived temperamental difficultness: acoustic properties and effects on listeners. *Child Development, 53,* 677-686.

Michelsson, K. & Wasz-Höckert, O. (1980). The value of cry analysis in neonatology and early infancy. In T. Murry & J. Murry (Eds.), *Infant communication: cry and early speech* (pp. 152-182). Houston: College-Hill Press.

Murry, T., Amundson, P., & Hollien, H. (1977). Acoustical characteristics of infant cries: fundamental frequency. *Journal of Child Language, 4,* 321-328.

Murry, T., Hoit-Dalgaard, J., & Gracco, V. L. (1983). Infant vocalization: a longitudinal study of acoustic and temporal parameters. *Folia Phoniatrica, 35,* 245-253.

Rosenhouse, J. (1977). A preliminary report: an analysis of some types of a baby's cries. *Journal of Phonetics, 5,* 299-312.

Scherer, K. R. (1979). Nonlinguistic vocal indicators of emotion and psychopathology. In C. E. Izard (Ed.), *Emotions in personality and psychopathology* (pp. 495-529). New York: Plenum Press.

Scherer, K. R. (1982). The assessment of vocal expression in infants and children. In C. E. Izard & L. M. Dougherty (Eds.), *Measuring emotions in infants and children* (pp. 127-163). Cambridge University Press.

Tsukamoto, T. (1985). Development and mechanism of the infant cry. *Annual Reports of Konan Women's University, 4,* 1-20 (In Japanese).

Tsukamoto, T. (1991). Acoustical impressions of infant cries and the effect of infant care experience on cry judgment. *Proceedings of the 33 Japanese Association of Educational Psychology* , (In Japanese).

Tsukamoto, T. & Katagiri, S. (1988). The quantitative acoustic feature analysis of the infant cry. *Proceedings of the 52nd Japanese Psychological Association*, (In Japanese with English summary).

Tsukamoto, T. & Tohkura, Y. (1990). Perceptual units of the infant cry. *Early Child Development and Care, 65*, 167-178.

Tsukamoto, T. & Tohkura, Y. (1992). Tempo as a perceptual cue for judgment of infant cries. *Perceptual and Motor Skills, 74*, 258.

Wasz-Höckert, O., Lind, J., Vuorenkoski, V., Partanen, T., & Valanne, E. (1968). The infant cry: A spectrographic and auditory analysis. *Clinics in Developmental Medicine* 29 (pp. 1-42). London: Spastics International Medical Publications Heinemann.

Wiesenfeld, A.R., Malatesta, C. Z., & DeLoach, L. L. (1981). Differential parental response to familiar and unfamiliar infant distress signals. *Infant Behavior and Development, 4*, 281-295.

Wolff, P.H. (1967). The role of biological rhythms in early psychological development. *Bulletin of the Menninger Clinic, 31*(4), 197-218.

Wolff, P. H. (1969). The natural history of crying and other infant vocalization. In Foss, B. M. (Ed.), *The determinants of infant behavior Vol. 4.* (pp. 81-109), London: Methuen.

Zeskind, P. S. (1985). A developmental perspective of infant crying. In B. M. Lester, and C. F. Z. Boukydis (Eds.), *Infant crying: theoretical and research perspectives* (pp. 159-185), Plenum Press.

Zeskind, P. S., & Lester, B. M. (1978). Acoustic features and auditory perceptions of the cries of newborns with prenatal and perinatal complications. *Child Development, 49*, 580-589.

Zeskind, P. S., & Marshall, T. R. (1985) The relation between variations in pitch and maternal perceptions of infant crying. *Child Development, 59*, 193-196.

Acknowledgements

The authors would like to thank Dr. Shigeru Katagiri for his stimulating discussion at the earlier stages of this work and his great help in the acoustic analysis of cries. Thanks are also due to Mr. Masanobu Abe, who made his program available for producing transforms of the fundamental frequency and duration of cries.

Table 1

Acoustic Features of Cry Samples

Category	Infant (weeks)	Duration of Segmental Unit Mean (s)	F0 (Hz) Max.	Min.	Mean	SD
Call	YU (15)	2.7	448.7	220.8	354.8	34.5
	MA (15)	2.9	466.9	169.5	371.0	40.7
	MA (17)	3.2	456.7	216.2	371.0	40.0
Hunger	TE (11)	1.4	627.8	230.5	446.2	80.9
	MA (12)	1.3	689.1	181.3	393.1	63.5
	TE (17)	1.6	618.0	304.7	468.8	60.1
Anger	TO (9)	1.4	771.5	295.9	423.1	65.8
	MA (12)	1.8	527.8	283.1	418.5	50.8
	MA (13)	1.6	557.8	264.3	437.8	57.5

Table 2

Cry Stimuli (MA) Used in the Experiment

Category / Stimulus	Age (weeks)	Segmental Duration (s) Total	Mean	F0 (Hz) Max.	Min.	Mean
Call						
C1	15	7.8	3.9	430.3	215.3	353.7
C2		8.1	4.1	466.9	169.5	372.0
Hunger						
H1	12	7.2	1.2	502.5	247.9	398.1
H2		6.7	1.0	689.1	230.1	427.7
Anger						
A1	13	7.4	1.9	557.8	285.1	427.2
A2		6.6	1.7	539.3	301.9	440.8

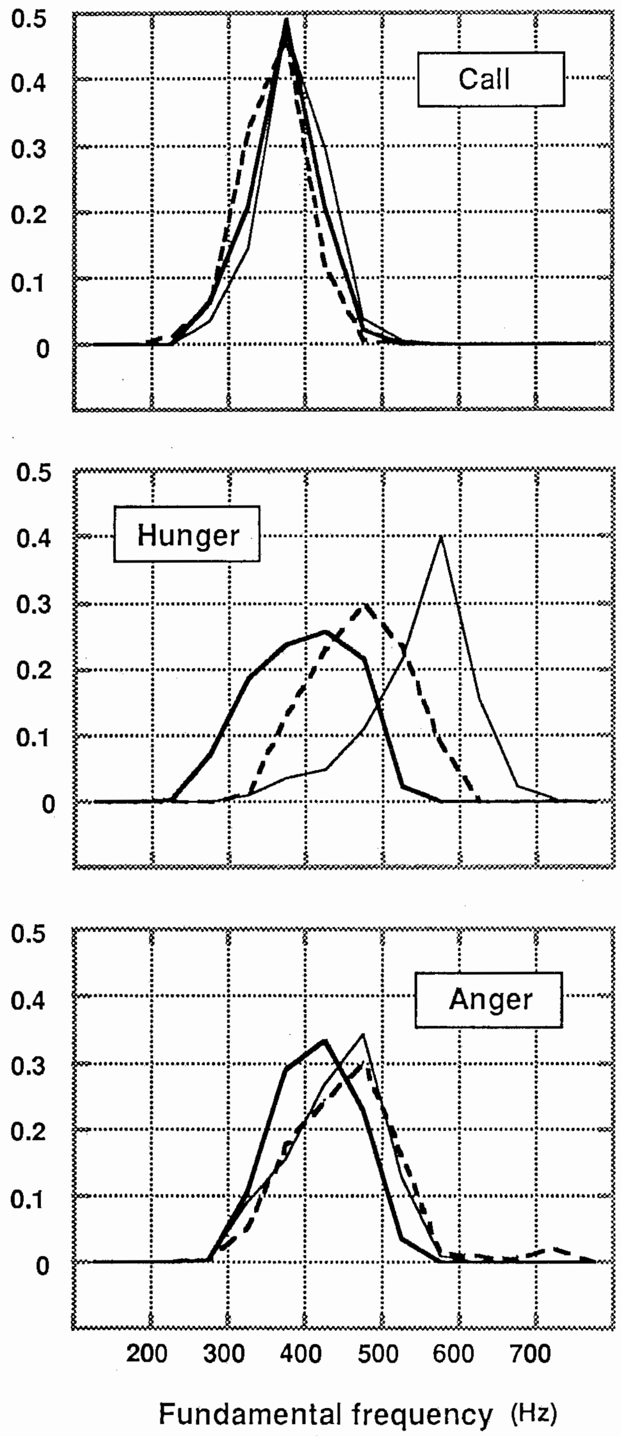

Figure 1. F0 histograms of three cry samples in each category. F0 is measured every 2.5 ms by applying an autocorrelation pitch extractor to the waveforms. They are quantized every 50 Hz and its occurrences are normalized by the total counts.

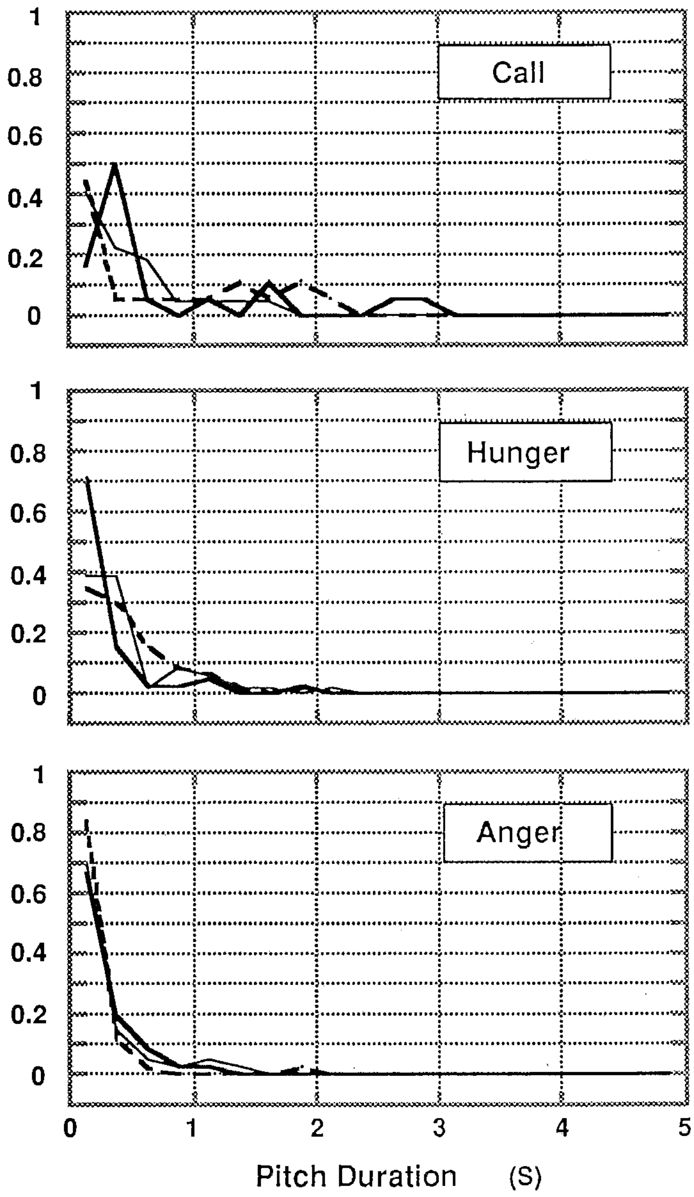

Figure 2. Pitch duration histograms of three cry samples for each category. The pitch duration is measured at a time resolution of 2.5 ms and quantized every 250 ms and its occurrences are normalized by the total counts.

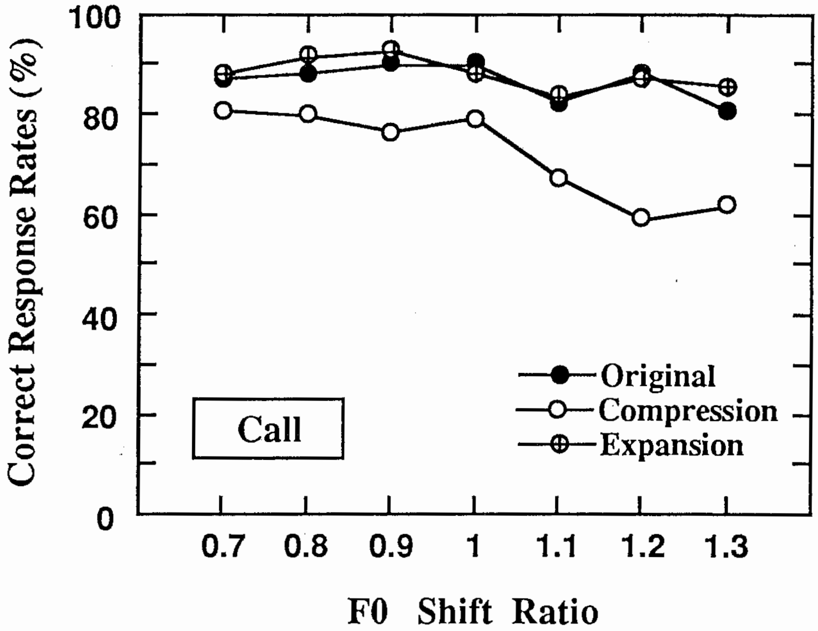

Figure 3. Correct response rates for call stimulus series. They were highest for expanded stimuli followed by the original and compressed stimuli.

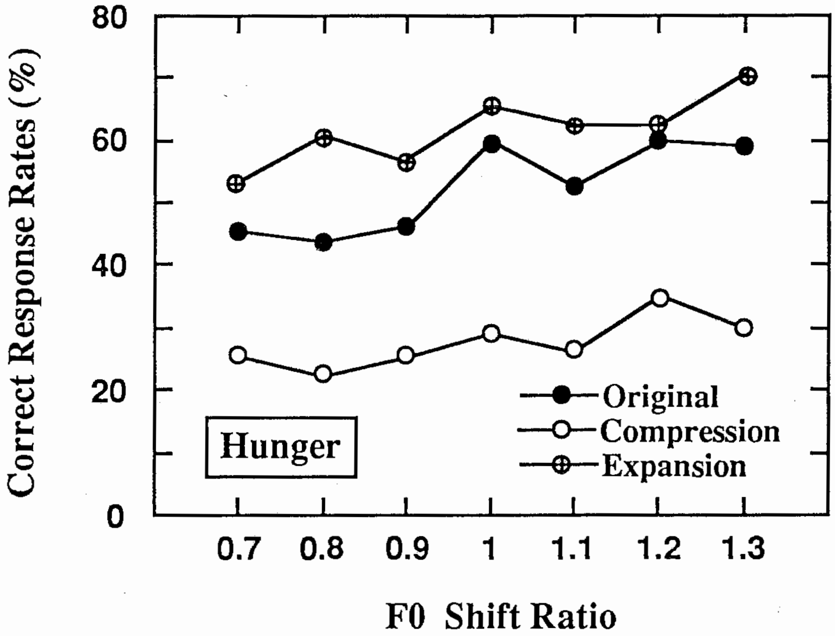

Figure 4. Correct response rates for hunger stimulus series. They were highest for expanded cstimuli followed by the original and compressed stimuli, and increased with F0 for all three conditions.

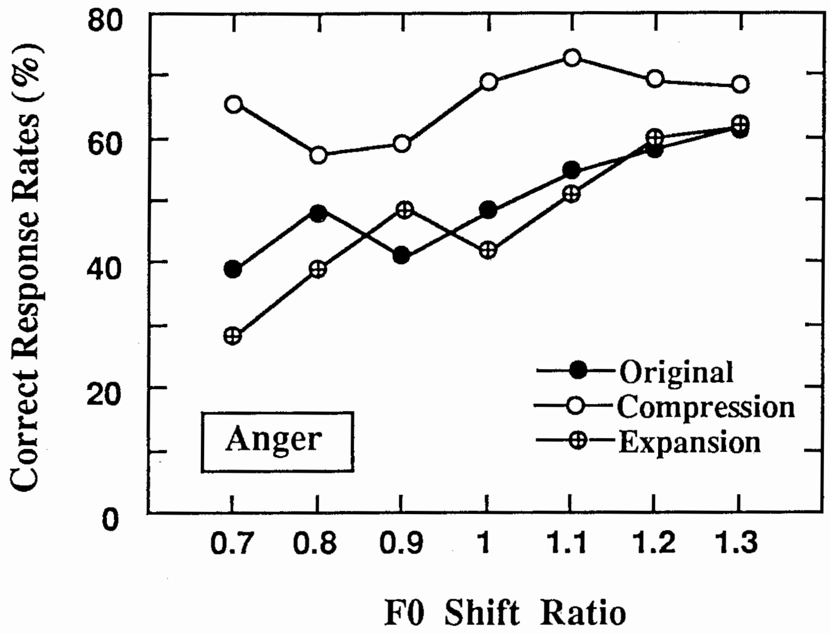

Figure 5. Correct response rates for anger stimulus series. They were highest for compressed stimuli followed by the original and expanded stimuli for which the correct response rates increased.

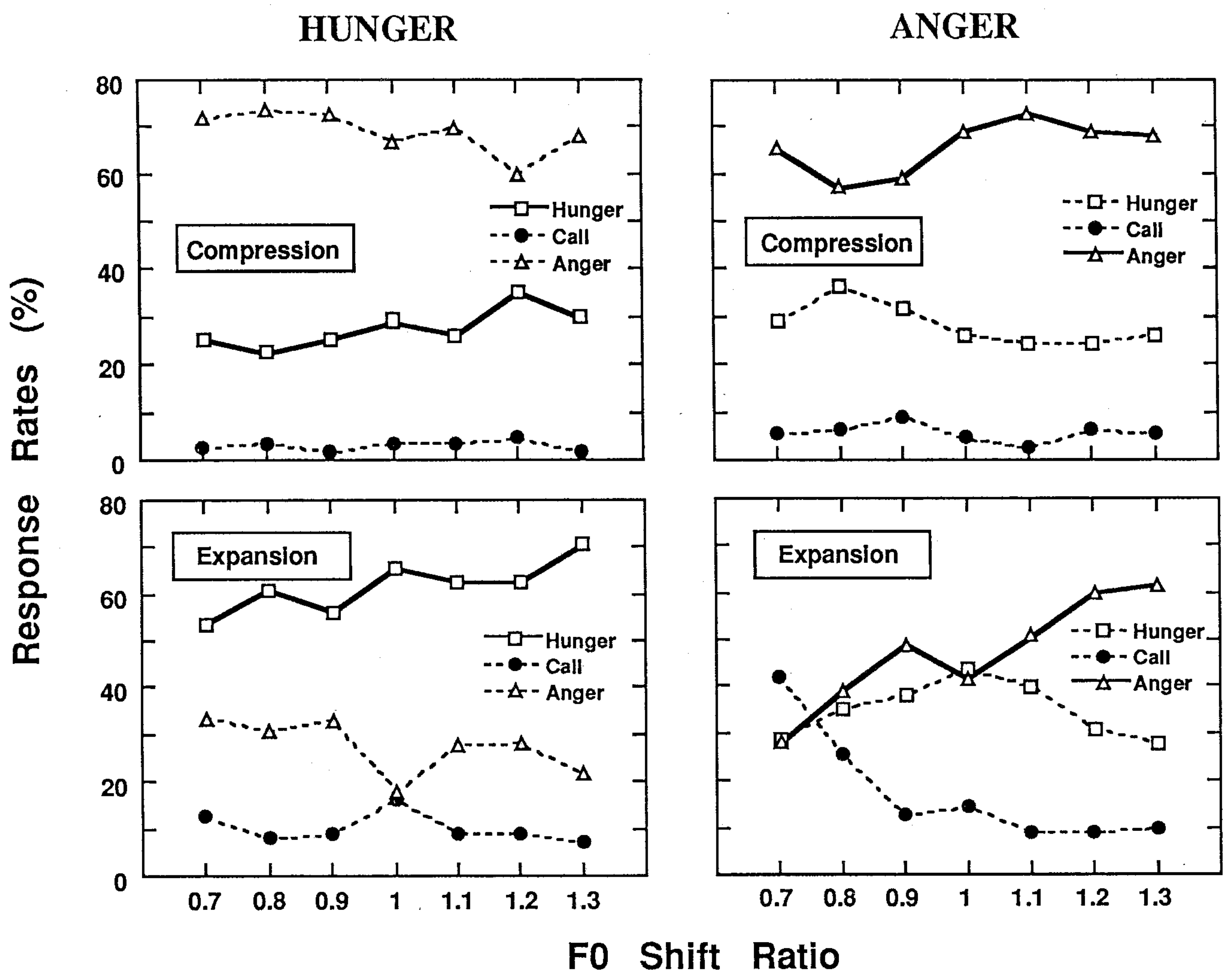

Figure 6. Response rates for compressed and expanded for hunger and anger stimuli. Correct responses are represented by solid lines and error responses are represented by dashed lines. It can be seen that the hunger and anger responses change according to tempo.