Internal Use Only
非公開

TR－A－0087 20

The effect of tempo on infant cry categorization

（泣き声の知覚におけるテンポの役割）

Taeko TSUKAMOTO *Yoh'ichi TOHKURA*
塚本 妙子 東倉 洋一

1990. 8. 1

ATR視聴覚機構研究所

〒619-02 京都府相楽郡精華町乾谷 ☎07749-5-1411

ATR Auditory and Visual Perception Research Laboratories

Inuidani, Seika-cho, Soraku-gun, Kyoto 619-02 Japan

Telephone: +81-7749-5-1411
Facsimile: +81-7749-5-1408
Telex: 5452-516 ATR J

The effect of tempo on infant cry categorization

TAEKO TSUKAMOTO
Department of Psychology, Konan Women's University, Higashinada-ku, Kobe 658 and Hearing & Speech Perception Department, ATR Auditory and Visual Perception Research Laboratories, Seika-cho, Kyoto 619-02, Japan

YOH'ICHI TOHKURA and SEIICHI TENPAKU
Hearing & Speech Perception Department, ATR Auditory and Visual Perception Research Laboratories, Seika-cho, Kyoto 619-02, Japan

A perceptual experiment was performed to examine the effect of tempo on the categorization of infant cry stimuli using three cry labels: hunger, anger, and call cries (i.e., calling for infant-mother interaction). In this experiment, cry stimuli with a variety of tempo were generated from original cry stimuli by changing the duration of cry stimuli using linear time-scale conversion. Subjects were instructed to make a forced choice among the three categories in labeling the cry stimuli. The results showed that the correct response rates were reduced significantly when the original cry tempo was changed. It was also found that anger cry stimuli, when time-scale expanded, tended to be perceived as hunger cries. On the whole, it was shown that fast tempo (i.e., time-compression of the original stimuli) was a perceptual cue for the anger cry and slow tempo (i.e., time-expansion of the original stimuli), on the other hand, was a perceptual cue for the call cry. These results indicate that the cry tempo is one reliable cue for cry category identification.

1. INTRODUCTION

The infant cry has been investigated from the viewpoint of communication behavior, since it has a powerful effect on eliciting infant care from the mother and is a strong trigger for mother-infant interaction. In the attachment theory by Ainsworth, Bell, and colleague (1972, 1973, 1978), they have stressed that the mother's sensitivity in perceiving and in interpreting the infant cry accurately, and her prompt response to them are the most important factors in an infant's social and mental development. How good is the mother's ability to judge the cry? In previous studies on infant cry recognition, one of the major concerns was whether the mother can judge the cry accurately or not. However, there has been no agreement among the experiment results, because of different research standpoints and the use of different types of cry samples, categories and subjects in the experiments.

From psychiatric evidences, it is assumed that crying is a part of the biobehavioral affect system which is related to the central and peripheral nervous systems, the respiratory system, and a variety of muscles (Lester, 1978, 1984; Michelsson & Wasz-Höckert, 1980; Boukydis & Burgess, 1985; Golb & Corwin, 1985). It is also assumed that, with development of the infant's organizations, the cry is differentiated from simple reflex responses to homeostatic regulation in the neonate period, and that different acoustic properties appear in the cry, depending upon the infant's needs or emotions, about one month after birth (Illingworth, 1955; Wolff, 1969; Futatsugi, 1979). Based upon these assumptions, it is hypothesized that the mother could categorize the infant cry and what it means based upon perceptual cues characterized by different acoustic features.

Taking account of this hypothesis, Tsukamoto (1983) performed perceptual experiments composed of both identification and discrimination tasks, using cry stimuli labeled with hunger, pain, sleepiness, discomfort, and call (i.e., calling for mother-infant interaction). For the identification task, subjects who were mothers scored higher than those who had had no experience of infant care. Mothers could not identify cries very accurately, but their identification rates were significantly higher than the chance level. Additionally, in the mothers' responses, there was a systematic error property in cry categorization with respect to the four categories: hunger, sleepiness, discomfort, and call. Namely, the mothers tended to perceive discomfort as sleepiness, sleepiness as call, call as hunger, and hunger as discomfort. It was also shown that there was a positive correlation between identification and discrimination characteristics obtained from the mothers' responses. These results indicated that mothers could significantly categorize the cries although their judgements were not always accurate. This finding is supported by Wolff's study (1969) in which he reported that mothers could discriminate pain from hunger cries produced in experimental situations. It was also found by Tsukamoto that lying-in or pregnant women who had had no experience in infant care showed a response pattern similar to the mothers but different from university students. However, pain cries were identified correctly regardless of the subjects' experience in infant care. These findings lead us to two hypotheses: (1) the mothers' sensitivity to the cry is related to degree of their motivation for infant care and (2) the cry has some distinctive features which can be perceived as related to infant's emotional state.

As the first step toward proving these hypotheses, it would be necessary to investigate infant cry perception by those who have had no experience in infant care and to reveal how experience affects the infant cry perception. Here, our questions are: What acoustic features are important in cry perception? and How are they related to cry category identification?

From previous studies based upon acoustic analyses of infant cries, a relationship between the acoustic features and the reason for the infant's cry, or disease has been suggested (Wasz-Höckert, et al.,1968; Rosenhouse, 1977; D'Odorico, 1984). Tsukamoto et al., (1988, 1989) also found that there are some differences among cry category in temporal structures such as temporal patterns of phonation/non-phonation and the average duration of cry segment units defined based upon the infant's breath group vocalizations. Thus, among a variety of acoustic features, fundamental frequency, amplitude, duration, and their temporal patterns all of which are related to prosody, were predominated for differentiating the cry.

In speech perception, it is generally understood that prosody, which is composed of tempo, rhythm and melody (i.e., pitch contour), plays an important role to convey attitudes and emotional states of speakers. Do these prosodic features also play an important role on cry perception? How are they related to perceptual cues in identifying cry categories? To answer these questions, we focused on "tempo", which is one of the most predominant prosodic features, and studied how it affects cry categorization. Our hypothesis in this study is that, a change of tempo causes a category shift in cry perception.

2. METHOD

2.1. Subjects

Subjects were 36 females and 32 males ranging in age from 19 to 42. None had had experience in infant care, and most were university students. Preceding the perceptual experiment, all subjects were trained in cry category identification in order to recognize the distinctive features of each cry category.

2.2. Samples

It is important to note how we chose cry samples when we conducted infant cry perceptual experiments. Most previous investigations used cry samples recorded in experimental situations. From the viewpoint of obtaining typical cry samples, this is a practical and useful method since it is not difficult for the experimenter to define the reason for the cries in the samples elicited operationally. However, the problem is that cry categories produced in experimental situations are limited to those such as pain, hunger, or startle, and that the samples thus obtained might not always be produced by the experimental situations even though that was the experimenter is intent. Taking this into consideration, we preferred to obtain cry samples produced spontaneously by infants in observational situations even though there are some difficulties in defining the reasons for the cries, and in controlling the recording conditions.

In our study, cries were recorded with a Video Tape Recorder (SONY SL-F1) in observational situations of infants in a mother-infant interaction. The observations were carried out at infant ages of 3 weeks to 12 months. The cries were identified by observers and were classified into several categories according to the observational situations of the infants. The observers were several university students majoring in psychology and one of the authors. Cry category classification was performed based on an evaluation agreement (more than 85 %) among the observers. Among several categories, three categories were selected for our study: hunger, anger and call (i.e., the cry calling for mother-infant interaction) cries. The cries from these three categories were used as cry samples for perceptual experiment and training in category learning in this study. They were digitized using an A/D converter with a sampling frequency of 20 kHz and edited using a computer (MASSCOMP 5600) to generate cry samples and stimuli.

2.3. Training

Preceding our perceptual experiment, training in category learning was carried out for all subjects to learn some distinctive features of the three cry categories.

2.3.1. Stimuli

The six samples in Table 1 were used as stimuli. They were produced by four infants (MA, YU, TO, TE) whose ages ranged from 9 to 17 weeks. Their durations were from 10.6 to 12.6 sec.

2.3.2. Training procedure

Training was composed of four procedures; (1) Category learning, (2) Identification trials (first), (3) Confirmation of category learning, and (4) Identification trials (second). In the category learning and the confirmation of category learning procedures, the stimuli were presented twice in succession with their category labels given to the subjects prior to listening. In the two identification trials, the same stimuli used in the category learning procedures were presented twice (12 stimuli in total) in random order. Subjects were instructed to identify them by a forced choice among three categories: hunger, anger and call cries. As a result of this training, 40 subjects (58.8 %) achieved cry category learning by accurately identifying all 12 stimuli in the second identification trials (see details in Appendix I).

2.4. Experiment

2.4.1. Stimuli

Cry samples were produced by one of the infants (MA) used in training, but different samples were selected for this perceptual experiment. Segment units excised from these cry samples were used as original stimuli. They were two for each category (Hunger: H1, H2; Anger: A1, A2; Call: C1, C2 in Table 2) which were from 4- to 6-segment units in length and from 6.11 to 16.69 seconds in duration. According to our previous experiments, the identification rates for these original stimuli exceeded 85% (see details in Appendix II).

The duration of the original stimuli was changed using a linear time-scale conversion method (Morita & Itakura, 1986) to generate stimulus series including cry stimuli with a variety of tempo for each category. Note that the linear time-scale conversion method we used does not produce any undesirable change in the segmental features of the spectrum or in the fundamental frequency, but varies only the tempo of stimuli. This property of time-scale conversion is very important in satisfying our experimental purpose and is clearly seen in examples of cry stimulus spectra (Fig.1). In Fig. 1, A is a spectrogram of the original stimulus, B is of the stimulus expanded 1.5 times, and C is of the stimulus compressed to 0.5. With respect to the tempo of these three stimuli (i.e., A, B, and C) B is 1.5 times slower and C is twice as fast as A. When the expansion/compression rates are so high, some acoustic distortions produced by time-scale conversion become so large as to be significant. To avoid these undesirable conditions and to use stimuli without any perceived distortion, the range of expansion/compression was set at 0.4 to 2.5, based upon the preliminary experiment results. Also, taking into account the perceptual experiment efficiency, the effective ranges of expansion/compression for each cry category were determined by the preliminary experiment. Thus, an anger cry stimulus series was generated by expansion only and a call cry stimulus series by compression only. Consequently the expansion rates for the anger cry series were 1.0 (original), 1.5, 2.0, and 2.5 (4 stimuli in each stimulus series) and the compression rates for the call cry series were 1.0 (original), 0.67, 0.5, and 0.4 (4 stimuli in each stimulus series). For the hunger cry, the compression/expansion rates were 0.5, 0.67, 1.0 (original), 1.5 and 2.0 (5 stimuli in each stimulus series). As two stimulus series were generated for each cry category, 26 stimuli were generated in all. In addition to these 26 stimuli, another 14 stimuli, excised from the same original samples but from different portions, were also prepared for embedding in the stimulus series as dummy stimuli. The number of the stimuli for the perceptual experiments was 40 in total.

2.4.2. Experiment procedure

The 40 stimuli were randomized and presented to the subjects with 10 dummy stimuli at the beginning (ISI = 3 sec). The subjects were required to identify each stimulus by a forced choice among the three categories: hunger, anger, and call cries.

2.5. Procedure

All stimuli in the training and the perceptual experiment were recorded once using a DAT (Digital Audio Tape, SONY DTC-1000ES) and presented to the subjects with headphones in a sound booth. Training lasted about 15 minutes. Then there was a 5 to 10 minute break followed by the experiment itself which lasted about 15 minutes.

3. RESULTS

The results were discussed based upon the response data for the 40 subjects who achieved cry category learning in their training (as described in Section 2.3.). The responses of 40 subjects for all stimuli were calculated and are shown in Tables 3, 4, and 5 for each category. Here, correct responses to the stimuli were defined as responses to the same category as that of their original stimuli. First, according to these data, the relationships between the response characteristics and the time-scale conversion were analyzed in each stimulus series. The results showed that there were significant correlations ($p < .05$ or $.001$) between them in all stimulus series: the anger cry stimuli (A1: $\chi^2 = 14.4$, $p < .05$; A2: $\chi^2 = 24.96$, $p < .001$, $df = 6$), the call cry stimuli (C1: $\chi^2 = 14.08$, $p < .05$; C2: $\chi^2 = 27.2$, $p < .001$, $df = 6$), and the hunger cry stimuli (H1: $\chi^2 = 33.66$, $p < .001$; H2: $\chi^2 = 19.8$, $p < .05$, $df = 8$). Second, the differences in the correct response rates in each stimulus series were analyzed. In chi-square analysis, the results showed that the correct responses shifted significantly according to the time-scale expansion/compression rates for each anger stimulus series (A1: $\chi^2 = 59.15$; A2: $\chi^2 = 38.7$; $df = 3$ $p < .001$) and hunger stimulus series (H1: $\chi^2 = 64.22$; H2: $\chi^2 = 78.31$; $df = 4$, $p < .001$). However, different analysis results were shown between two stimulus series for the call; differences in correct response rates were significant for one series ($\chi^2 = 9.91$, $p < .05$, $df = 3$), but not significant for the other.

Average response rates of two stimulus series, as a function of the time-scale conversion rates, were illustrated for each category in Figs. 2, 3, and 4, as similar response properties were observed between the two stimulus series for each category. As shown in Fig. 2, when time-scale expanded, the correct response rates for the anger stimuli decreased and more than half of the anger stimuli were perceived as hunger cries. When time-scale expanded more than 2.0 times, call cry responses appeared in significant numbers. In Fig. 3, the correct response rates of the call stimuli were also reduced when time-scale compressed, although the stimuli were still predominantly perceived as call cries. When the compression rates were 0.5 and 0.4, the call stimuli were to some extent perceived as hunger or anger cries. For the hunger stimuli shown in Fig. 4, when the time-scale was compressed, more than half of the stimuli were perceived as anger cries. However, when the time-scale was expanded, the correct response rates increased and became higher than those of the original stimuli.

4. DISCUSSION

In this paper, we focused on tempo and examined the effect of tempo on cry categorization. Observed in the relationships between the response characteristics of the subjects and the time-scale conversion rates, it was shown that correct response rates varied in all categories, and that the perceptual categories shifted between the hunger and the anger cries; i.e., the anger stimuli were perceived as hunger cries when the time-scale was expanded and the hunger stimuli were perceived as anger cries when it was compressed. Although the absolute number of call cry responses is not particularly large, they appeared to some considerable extent for anger stimuli when time-scale expanded more than 2.0 times (i.e., twice as slow as the original stimulus tempo). On the other hand, anger cry responses appeared for call stimuli when time-scale compressed more than 0.5 times (i.e., twice as fast as the original stimulus tempo). Taking all this into consideration, it is assumed that (1) fast tempo is a perceptual cue for the anger cry, and (2) slow tempo is a perceptual cue for the call cry on cry category identification.

Here, let us consider general response characteristics for the three kinds of cries used in the experiments. Summarizing the experiment results, the correct response rates for the call cry are markedly higher than those for the anger and hunger cries. This suggests that the call cry is perceived as quite different from the other two. The anger and hunger cries are two kinds of cries which are easily confused with each other. Similar results were also observed in the previous experiment (Appendix II). Based upon these experiment results, it can be assumed that the call cry has the acoustic features which allow it to be clearly differentiated from the other two cries. The anger and hunger cries seem rather similar in their acoustic features.

As is seen in both Fig. 2 and Fig. 4, the effect of tempo on categorizing the anger and hunger cries is fairly strong and a category shift owing to change of tempo between these two cries is clear. When the hunger cry stimuli were compressed at the rate of 0.5, they were perceived as anger cries by more than 70% of the subjects. From these results, there is no doubt that fast tempo is one of the most predominant cues for anger cry categorization. This finding is also supported by the fact, observed in emotional speech, that fast tempo is one of the parameters which characterize anger (Scherer, 1979). Here, it is necessary to further consider the detailed response characteristics of the hunger cry. As is described in Section 2.4.1 the original hunger stimuli used in the experiment were segment units excised from the cry samples. Their correct response rates were shown to exceed 85% in the previous experiment. However, for the same stimuli, the correct response rates stayed as low as around 40%, which is only a little bit better than the chance level. One reason for this seems to be related to the differences in the stimulus sets in the two experiments. In the previous experiment, stimulus sets consisted of a number of segment units with a variety of acoustic features. In this experiment, on the other hand, six stimuli (two for each of the three categories) which differed in various acoustic features were used as the original stimuli. All other stimuli had the same segmental features as the original and could be differentiated only in tempo. To avoid a perceptual bias which may occur when using stimuli which are similar to each other, 14 dummy stimuli were embedded among the 26 stimuli (Section 2.4.1). However, this might not be sufficient to normalize the perceptual bias entirely and to reduce the subjects excessive concentration on differences in tempo. As a result, the anger and hunger stimuli were differentiated mainly in tempo, regardless of the differences in some other acoustic features. Eventually, perceptual cues, used to identify the hunger cry in the previous experiment, were less effective in this identification test. On the contrary, the subjects tended to rely too much on tempo as a predominant perceptual cue to differentiate the anger and hunger cries. Due to this, the subjects perceived hunger stimuli with fast tempo as anger cries, even when the hunger stimuli are original. In fact, both of the two hunger stimuli used in the experiment are composed of repetitions of relatively short segments which presumably cause perceptual fast tempo (Table 2). Summarizing these discussions, it can be understood that the correct response rates for the hunger stimuli stayed fairly low and that they became higher when time-scale expanded.

Response characteristics for the call cry are also considered. Taking a look at the experiment results in Figs. 2, 3 and 4, the effects of tempo on call cry perception were observed to some extent. As is seen in Fig. 3, the more the call cry stimuli were compressed, the more the correct recognition rates decreased. In addition, Fig. 4 showed that the more the anger stimuli were expanded, the more they were perceived as call cries. However, this tendency was not so remarkable as that seen in the category shift between the anger and the hunger cries. Summarizing, slow tempo is a perceptual cue for the call cry, but is not as predominant as the perceptual cue of fast tempo for the anger cry. As the call cry seems to have acoustic features which are distinctive from those of the other two cries, there must be better cues to characterize the call cry. These cues, though they might have been weakened by time-scale compression, seemed to work to preserve the call cry characteristics fairly well (i.e., the correct response rates of about 70% as shown in Fig. 3) when the call cry stimuli were compressed, even at the rate of 0.4.

The linear time-scale conversion we used is a method to change the time-scale uniformly and linearly for all acoustic parameters. From a global view, the "tempo" of cry stimuli is varied linearly using this method without any change in segmental acoustic features. However, it is obvious that some dynamics, such as formant frequency transitions and amplitude envelopes, vary due to time-scale conversion. Note that "tempo" in this paper includes all of these factors. When the contributions of these dynamic features are considered as perceptual cues for the cry, a nonlinear time-scale conversion is needed to change some temporal localized features.

In this paper we focused on tempo and studied its effect on infant cry categorization. As a variety of acoustic features which could be perceptual cues for the cry were observed, further research is needed to reveal each of their characteristics.

5. ACKNOWLEDGMENTS

The authors would like to thank Professors, Dr. Shoshiro Kuromaru, Dr. Shin-ichi Ohno, and Dr. Takao Umemoto of Konan Women's University for their helpful comments and support, and also Dr. Shigeru Katagiri, ATR Auditory & Visual Research Laboratories, for valuable comments and discussions throughout this study.

6. REFERENCES

Ainsworth, M.D.S. and Bell, S. Mother-infant interaction and the development of competence. In K.J. Connolly & J.S. Bruner (Eds.), The growth of competence. London: Academic Press, 1973.

Ainsworth, M.D., Blehar, M.C., Waters, E. and Wall, S. Patterns of attachment. Hillsdale, New Jersey: Erlbaum, 1978.

Bell, S. and Ainsworth, M.D.S. Infant crying and maternal responsiveness. Child Development, 1972, 43, 1171-1190.

Boukydis, C.F.Z. and Burgess, Perception of infant crying as an interpersonal event. In B.M. Lester & C.F.Z. Boukydis (Eds.) Infant crying: Theoretical and research perspectives. New York: Plenum Press, 1985.

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

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

Golb and Corwin A physioacoustic model of the infant cry. In B.M. Lester & C.F.Z. Boukydis (Eds.) Infant crying: Theoretical and research perspectives. New York: Plenum Press, 1985.

Illingworth, R.S. Crying in infants and children. British Medical Journal, 1955, 8, 75-78.

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

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

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

Morita, N. and Itakura, F. Time-scale modification algorithm for speech by use of autocorrelation method and evaluation. The Institute of Electronics, Information and Communication Engineers of Japan, 1986, EA86-5, 9-16.

Muller, E., Hollien, H., and Murry, T. Perceptual responses to infant crying: identification of cry types. Journal of Child Language, 1974, 1, 1, 89-95.

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

Scherer, K. R. Nonlinguistic vocal indicators of emotion and psychopathology. In C.E. Izard (Ed.), Emotions in personality and psychopathology. New York: Plenumn Press, 1979.

Tsukamoto, T. Mother's recognition of the infant cry. Annual Report of Konan Women's University, 1983, 2, 31-44 (In Japanese).
Tsukamoto, T. Development and mechanism of the infant cry. Annual Report of Konan Women's University, 1985, 4, 1-20 (In Japanese).
Tsukamoto, T. and Katagiri, S. The quantitative acoustic feature analysis of the infant cry. Proceedings of the 52nd Japanese Psychological Association, 1988 (In Japanese with English summary).
Tsukamoto, T. and Tohkura, Y., Perceptual units of the infant cry. 117th Meeting of the ASA, Paper U-19, 1989.
Wasz-Höckert, O., Lind, J., Vuorenkoski, V., Partanen, T., and Valanne, E. The infant cry: A spectrographic and auditory analysis. Clinics in Developmental Medicine 29. London: Spastics International Medical Publications Heinemann. 1968.
Wolff, P.H. The natural history of crying and other infant vocalization. In Foss, B.M. (Ed.) The Determinants of infant behavior Vol. 4. London:Methuen, 1969.

Table 1

Cry stimuli used in training

Category	Infant	Age (weeks)	Duration (sec)
Hunger	TE	11	12.3
	TE	17	12.6
Anger	TO	9	12.6
	MA	12	10.6
Call	YU	15	10.8
	MA	17	12.6

Table 2

Original stimuli (MA) used in experiment

Category	Stimulus	Number of Segments	Duration (sec)	Segmental Duration (sec)			Age (weeks)
				Ave.	Max.	Min.	
Hunger	H1	6	6.11	1.02	1.39	0.31	12
	H2	5	6.17	1.23	1.74	0.62	
Anger	A1	4	6.77	1.69	2.31	0.64	13
	A2	4	6.99	1.75	3.02	0.51	
Call	C1	5	9.70	1.94	5.04	0.53	15
	C2	5	16.69	3.38	4.75	1.50	

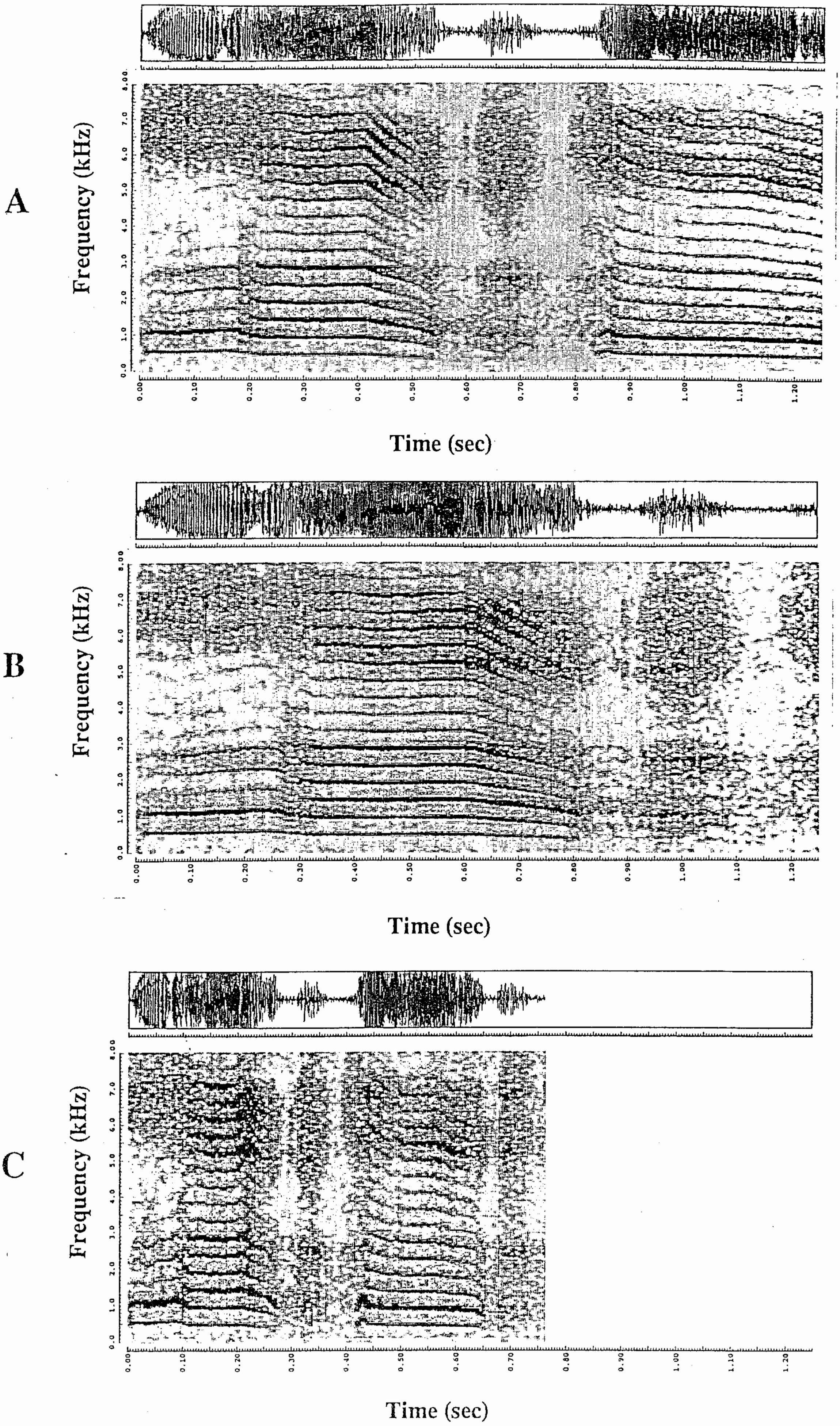

Fig. 1. Examples of cry stimulus spectrograms. A: an original stimulus B: a stimulus expanded 1.5 times C: a stimulus compressed 0.5 times. Upper and lower windows show cry waveforms and sound spectrograms respectively, for each example.

Table 3

Response rates for anger cry stimuli

Stimuli & conversion rates	Response rates (%) (N = 40)			Correlation between response rates & time-scale conversion χ^2 $df = 6$
	Hunger	Anger	Call	
A1 - 1	45.0	55.0	0.0	
A1 - 1.5	52.0	47.5	0.0	14.4 *
A1 - 2.0	60.0	30.0	10.0	
A1 - 2.5	67.5	25.0	7.5	
χ^2 $df = 3$		59.15 ***		
A2 - 1	22.5	77.5	0.0	
A2 - 1.5	42.5	55.0	2.5	
A2 - 2.0	50.0	45.0	5.0	24.96 ***
A2 - 2.5	55.0	30.0	15.0	
χ^2 $df = 3$		38.7 ***		

A1 - 1 and A2 - 1 are original stimuli. * $p < .05$ *** $p < .001$

Table 4

Response rates for call cry stimuli

Stimuli & conversion rates	Response rates (%)	(N = 40)		Correlation between response rates & time-scale conversion χ^2 $df = 6$
	Hunger	Anger	Call	
C1 - 1	0.0	5.0	95.0	
C1 - 0.67	5.0	0.0	95.0	
C1 - 0.5	7.5	10.0	82.5	14.08 *
C1 - 0.4	15.0	12.5	72.5	
χ^2 $df = 3$			n.s.	
C2 - 1	7.5	0.0	92.5	
C2 - 0.67	20.0	0.0	80.0	
C2 - 0.5	27.5	12.5	60.0	27.2 ***
C2 - 0.4	20.0	12.5	67.5	
χ^2 $df = 3$			9.91 *	

C1 - 1 and C2 - 1 are original stimuli. * $p < .05$ *** $p < .001$

Table 5
Response rates for hunger cry stimuli

Stimuli & conversion rates	Response rates (%) (N = 40)			Correlation between response rates & time-scale conversion χ^2 $df = 8$
	Hunger	Anger	Call	
H1 - 0.5	20.0	67.5	12.5	
H1 - 0.67	27.5	60.0	12.5	
H1 - 1	40.0	52.5	7.5	33.66 ***
H1 - 1.5	60.0	25.0	15.0	
H1 - 2.0	65.0	17.5	17.5	
χ^2 $df = 4$	64.22 ***			
H2 - 0.5	17.5	77.5	5.0	
H2 - 0.67	22.5	72.5	5.0	
H2 - 1	35.0	52.5	12.5	19.8 *
H2 - 1.5	50.0	45.0	5.0	
H2 - 2.0	45.0	42.5	12.5	
χ^2 $df = 4$	78.31 ***			

H1 - 1 and H2 - 1 are original stimuli. * $p < .05$ *** $p < .001$

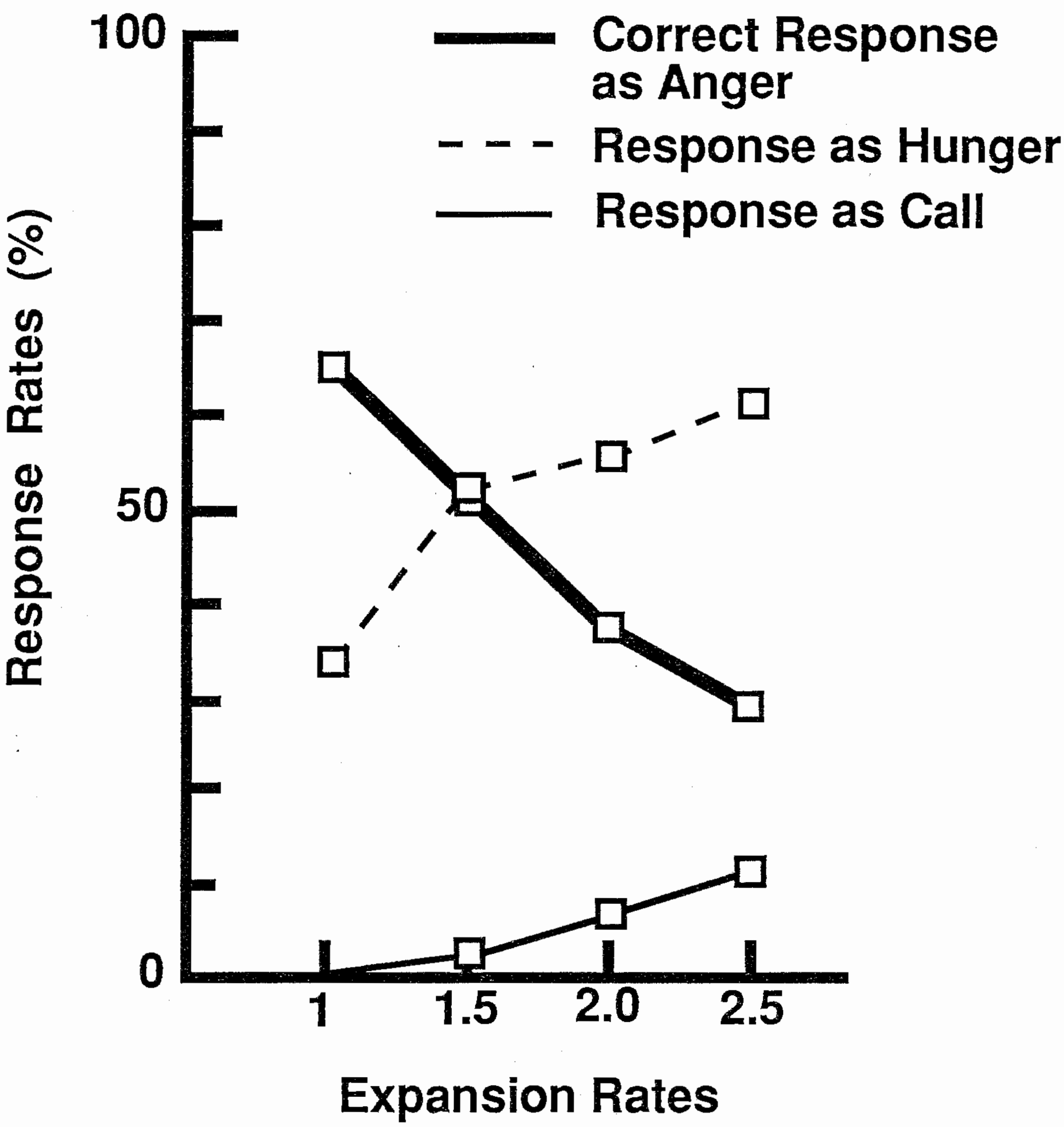

Fig. 2. Average response rates for anger cry stimuli. When the time-scale was expanded, anger cry stimuli tended to be perceived as a hunger cry.

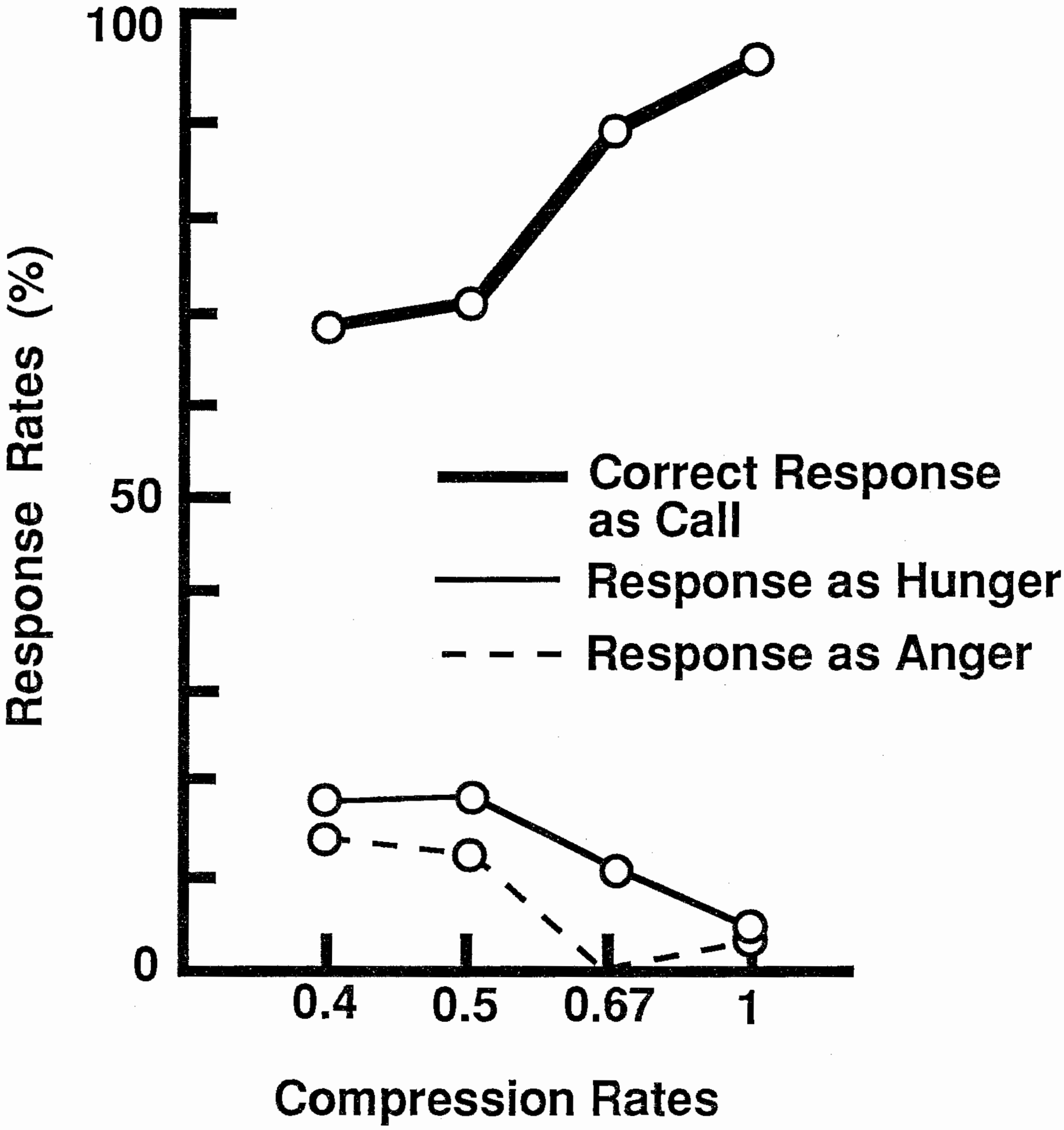

Fig. 3. Average response rates for call cry stimuli. When the time-scale was compressed, the correct response rates for the call cry stimuli were reduced significantly.

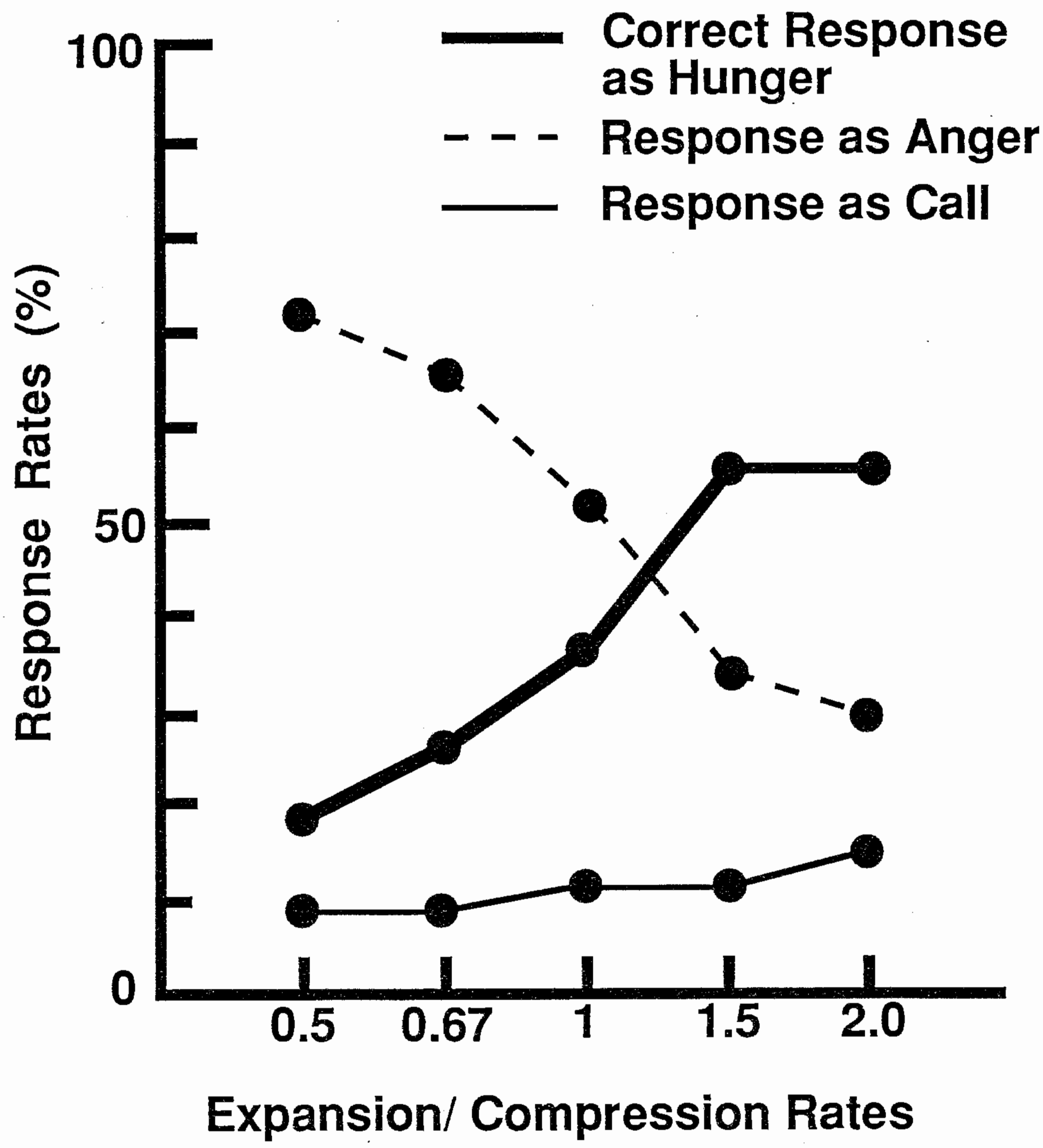

Fig. 4. Average response rates for hunger cry stimuli. The correct response rates for the original cry stimuli were chance levels and were confused with an anger cry. When time-scale compressed, the hunger cry stimuli were perceived as an anger cry and when time-scale expanded, the response rates for the hunger cry were increased significantly.

Appendix I

Training procedure

Training is composed of four procedures;

(1) Category learning

Subjects are instructed to learn some distinctive features of each cry category. Stimuli with their category labels are presented twice in succession to the subjects.

(2) Identification trials (first)

The stimuli are presented twice in random order. The subjects are instructed to identify them by a forced choice among the three categories: hunger, anger, and call cries.

(3) Confirmation of category learning

The stimuli are presented once again using the same procedure used in category learning.

(4) Identification trials (second)

The same procedure used in the first identification trials is carried out. As a result of these trials, subjects who identify all 12 stimuli are regarded as having achieved cry category learning.

Appendix II

Perceptual units of the infant cry

This appendix is a review of our previous study (Tsukamoto and Tohkura, 1989) on perceptual units of the infant cry to give a clear idea about what segment units of cry samples are and why these are important for generating cry stimuli in the perceptual experiments.

The purpose of the study on "perceptual units of the infant cry" is to examine how long cry samples must be for cry category identification. Cry samples in three categories: hunger, anger, and call cries were chosen, and employed in perceptual experiments.

1. Method

In order to generate stimuli, each of cry samples was first segmented into single segment units according to the infant's breath group vocalization (Fig. A1). The single segment units were combined with each other in temporal order to generate two-, three-, five-, and seven-segment unit stimuli (Fig. A2). In addition to these multi-segment unit stimuli, the single segment units and the three original samples were used as one-segment unit stimuli and full-segment unit stimuli, respectively. 52 subjects who were trained by category learning were instructed to identify by a forced choice among the three cry categories. Experiment results given by 20 subjects who achieved category learning in their training were analyzed and discussed.

2. Results

Category identification rates increased significantly dependent upon the number of segments, and reached a plateau at two-segment units in the call cry and three to five-segment units in the anger and hunger cries (Fig. A3). The number of segments at the plateau is considered to be the minimum perceptual units for cry category identification.

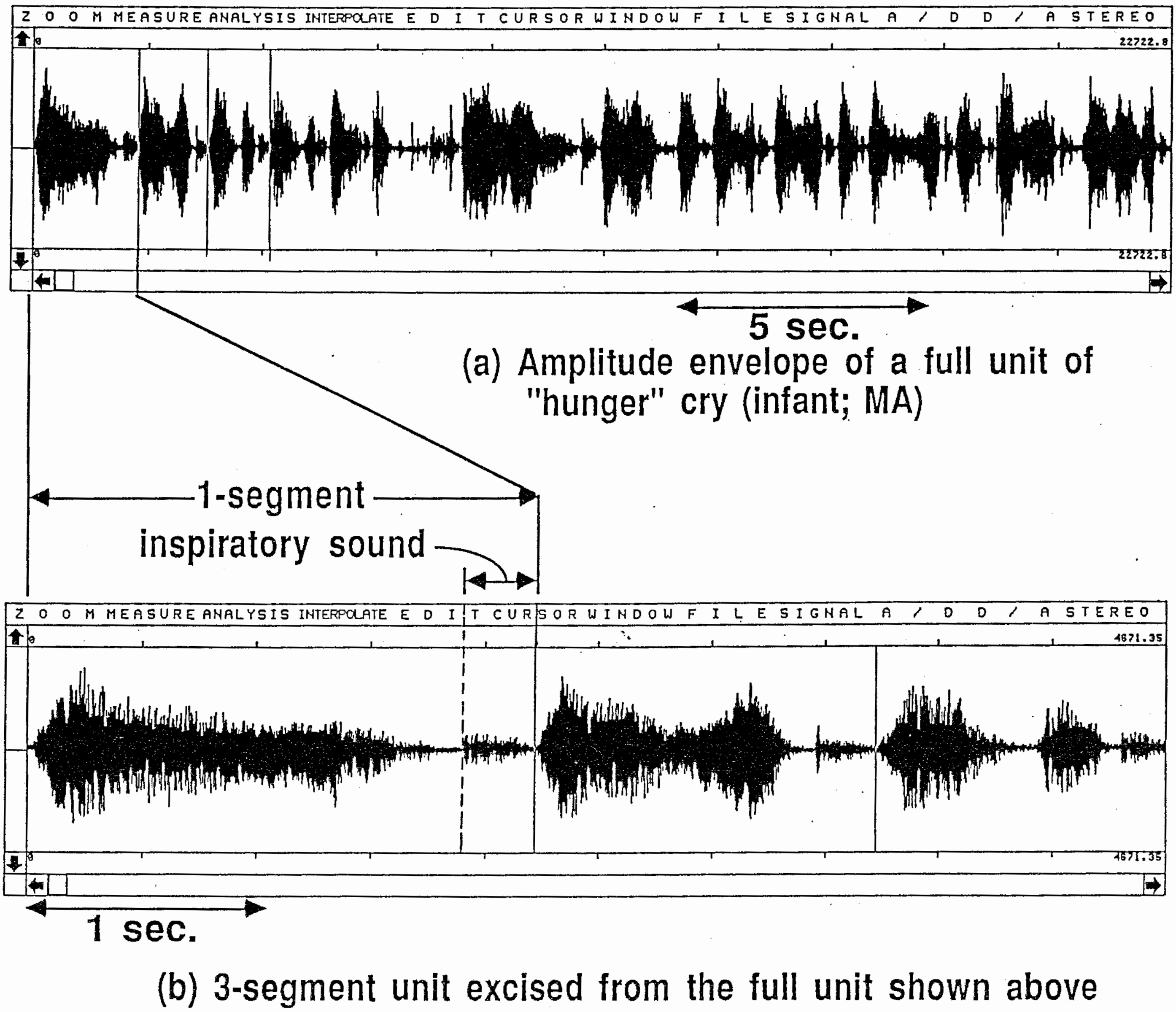

Fig. A1. Segmentation of cry waveformes.

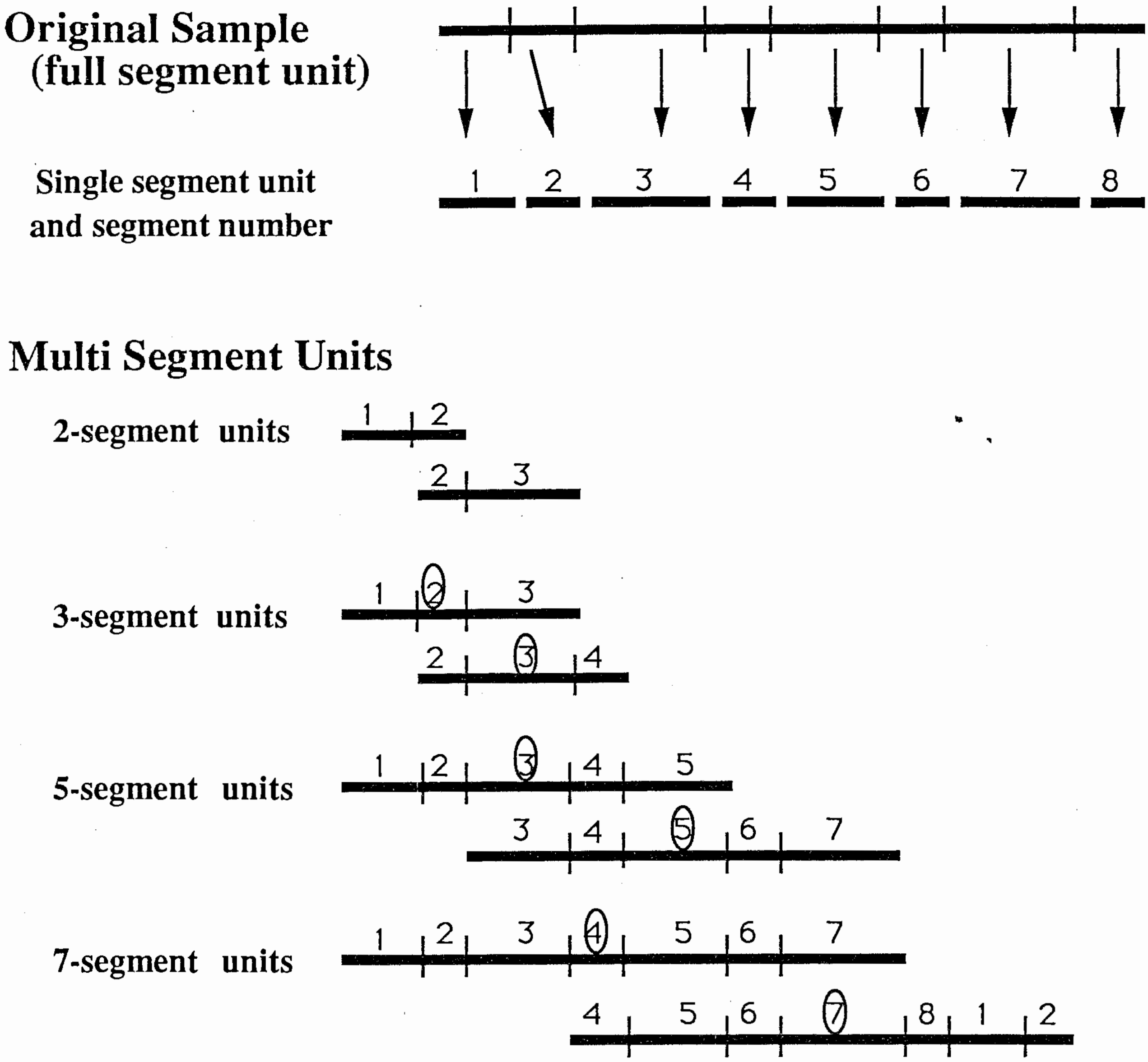

Fig. A2. Illustration of cry segment units used in category identification tests. A full segment unit is divided into *n* cry segments. A 1-segment unit is a unit with only one segment. An *i*-segment unit is composed of succesive *i* segments. In the identification test, 1-, 2-, 3-, 5-, 7-segment and full segment units are used as stimuli. The numbers marked by a circle represent central segment numbers of each stimulus.

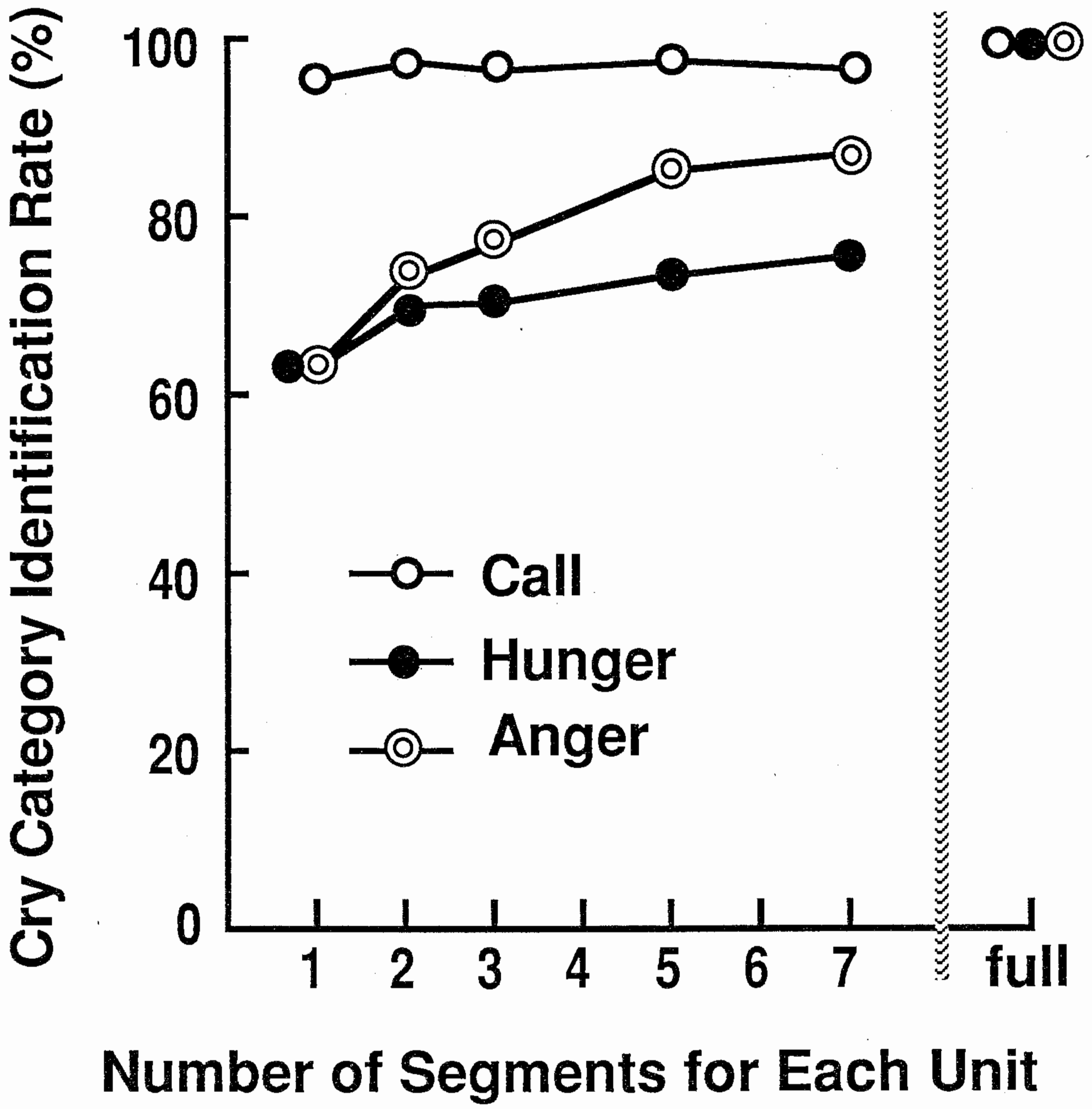

Fig. A3. Cry category identification rates for 20 subjects who correctly identified each full segment unit of three cry categories by 100%. Category identification rates increased significantly dependent upon the number of segments, and reached a plateau at 2-segment units in the call cry and 3- to 5-segment units in the anger and hunger cries.