VP Ratio of the Infant Cry Sounds and Its Influences on Their Perception

Tohoku Psychologica Folia
1987, **46** (1-4), 36-42

VP RATIO OF THE INFANT CRY SOUNDS AND ITS INFLUENCES ON THEIR PERCEPTION

By

HITOSHI OKADA (岡田 斉)[1], NORIO MURAI (村井憲男)[2]

(*Tohoku University*)

and

TOMOAKI ADACHI (足立智昭)[3]

(*Niigata Chuoh Junior Collage*)

Our previous study (Adachi et al., 1985) reported that the VP ratio (the ratio of the voiced phonation to the total phonation) of infant cries effected the maternal perception. In the present study we synthesized the cries which had lower VP ratio than the original cries to investigate the influences on the perception of the subjects (20 female undergraduates). Semantic Differencial technique was employed to assess the perception of the subjects. The results showed that the cries of low VP ratio were perceived by them as more urgent, grating, sick, arousing, piercing, discomforting, aversive and distressing ones than the original cries. These results suggested that VP ratio significantly affected to the perception of the subjects.

Key words: infant cries, acoustic features, maternal perception, Semantic Differential technique, speech synthesis, VP ratio

Introduction

Some researchers have paid an attention to the relation between the perception of the cry sound of an infant and its physical characteristics (Zeskind & Lester, 1978; Lounsbury & Bates, 1982; Adachi, Murai, Okada & Nihei 1985; Porter, Miller & Marshall, 1986). For example, Zeskind and Lester (1978) suggested that the fundamental frequency was an important parameter for the perception of the cry sound. Our study (Adachi et al., 1985) used the multiple regression analysis between physical characteristics and the rating of the subjects. In the regression equation, VP ratio played an important role in addition to the fundamental frequency. In short, VP ratio is the auto-correlation of the wave form. When the maximum value of the auto-correlation of the analysis frame (time window) is above 0.25, this frame is judged as the voiced phonation (Markel & Gray, 1976). We defined VP ratio as the ratio of

1. Department of Psychology, Faculty of Art and Letters Tohoku University, Kawauchi, Sendai 980, Japan.
2. Department of Developmental Disorders, Faculty of Education, Tohoku University, Kawauchi, Sendai 980, Japan.
3. Niigata Chuoh Junior College, Kamo, 959-13, Japan.

the frames of voiced phonation to the total frames of phonation.

Both fundamental frequency and the VP ratio are the informations based on the wave form of the cry sound. In contrast, formant is on the spectrum which is calculated from the wave form. The fundamental frequency is the frequency of the wave form. The VP ratio is the "regularity" of the wave form. Both these two are also related to the condition of the voice source, especially that of the vocal cords in seeing from the speech production system.

When we investigate the relation between the perception of the cry and its physical characteristics, it is ideal to change only one physical index and measure the reaction of the subjects. Application of the speech synthesis technique is adequate to such study. However, a few studies have been done by this method (for example Simner 1971). The reason why we were interested in the linear prediction technique is that the synthesis and analysis are equivalent in its procedure : when a physical characteristic is extracted from the analysis we can operate this parameter and retest the meaning of it clearly. It is well known that biological acoustics has got many findings by this method, which is called feedback method (Bright, 1986).

Our study reported that the cry sound of low VP ratio was perceived by the subjects as more urgent, grating, sick, arousing, piercing, discomforting, aversive, and distressing than the cry with high VP ratio(Adachi et al. 1985). In our present study we synthesized a low VP ratio infant cry sound to examine the relation between perception of an infant cry sound and VP ratio. And with a rating method, the relationship of VP ratio to the perception of the cry sound was discussed.

Method

Subjects

Twenty female undergraduates of Tohoku University served as subjects. They were single women and had no professional experience with infants.

Stimulus

We chose 4 cries as stimulus out of the 16 taken out of our previous experiments (Adachi et al. 1985). Three cries were used for the synthesis. Cry A is from a high-risk infant and B and C are from low-risk infants. The physical characteristics of the cries are shown in Table 1. They were selected because each of them represented high, middle and low fundamental frequency respectively among the 16 cries. Cry D, which had 0.6 VP ratio, was not used perse for the synthesis.

The present synthesis is the reverse process of the analysis of the cry that we previously reported. Linear prediction method was used. Three sample cries were passed through a low pass filter (cut off frequency 10 kHs, 48 dB/oct) and digitized by an AD converter (sample rate 20 kHz, 12 bit). All digitized data were processed by a computer (NEC ACOSS 1000). The synthesis was done as follows :

Table 1. Physical characteristics of the four sample cries

	Cry			
	A	B	C	D
Mean fundamental frequency in 10 sec (Hz)	689	445	643	687
(first Phonation (Hz))	778	406	638	687
VP ratio	0.80	0.89	0.84	0.63
Duration of the first phonation (msec)	666	2496	2700	844
Number of utterance	10	4	3	12

(1) The frame was set at 512 points (25.6 msec).

(2) These data were passed through a hamming window.

(3) Using the Durbin method, partial correlation coefficients were calculated until the order reached 15.

(4) A filter was shaped with these coefficients. The original signals were introduced in this filter and the prediction residue was extracted.

(5) The partial of this residue was changed into random series. The VP ratio of the cry sound was reduced by this operation. The VP ratio was set at 0.3 and 0.6 for each cry. By this operation, data of the voice source were changed but the vocal tract data (formant) were preserved.

(6) The synthesized residue was entered into the inverse filter created by the correlation coefficients. In this manner, the synthesized cry sound of one frame was created.

(7) The frame was shifted every 25.6 msec and the same operation was continued until the end of the data.

(8) The synthesized data of the cry were passed through the digital low pass filter, the same as the input, and converted into analog data by a DA converter.

To examine the perceptual quality of these cries, we synthesized the cry which was delivered from the only analysis-synthesis of the original cry. The residue from the PACOR filter was not changed but the orders of the prediction were set at 8, 15 and 30 for each cry. With these stimulus, the similarity between the original cry and the synthesized cry was rated by 10 subjects. The results showed that it was difficult to distinguish the synthesized cry from the original cry regardless of the order of the prediction. The synthesized cries which have VP ratio 0.3 or 0.6 were heard as a natural hoarse voice but they were sounded like an infant's cry.

All cry signals were recorded in a cassette tape and used for the experiment.

Apparatus

The stimulus was presented by a cassette taperecorder (Pioneer CT620) and a speaker with an amplifier (Sony APMO90). The cries were presented monaurally 1.5 m apart from the subject in the sound proofroom. The level of the cry sound was

67dBSPL at the peak.

Procedure

The perceptual responses to cry sounds were measured by using the eight 7-point scales which Zeskind and Lester (1978) developed: (1) urgent-not urgent, (2) pleasing-grating, (3) sick-healthy, (4) soothing-arousing, (5) piercing-not piercing, (6) comforting-discomforting, (7) aversive-nonaversive, and (8) distressing-not distressing.

After the subjects practiced the rating task for two cry sounds which were not involved in the stimulus of the experiment, the rating of the synthesized cries began. They rated each cry on four different scales during the 20-sec intercry intervals. After 5-minute intervals they rated on the remaining four scales. The cry sounds were presented in the random order. The mean ratings of each cry for the eight scales were computed.

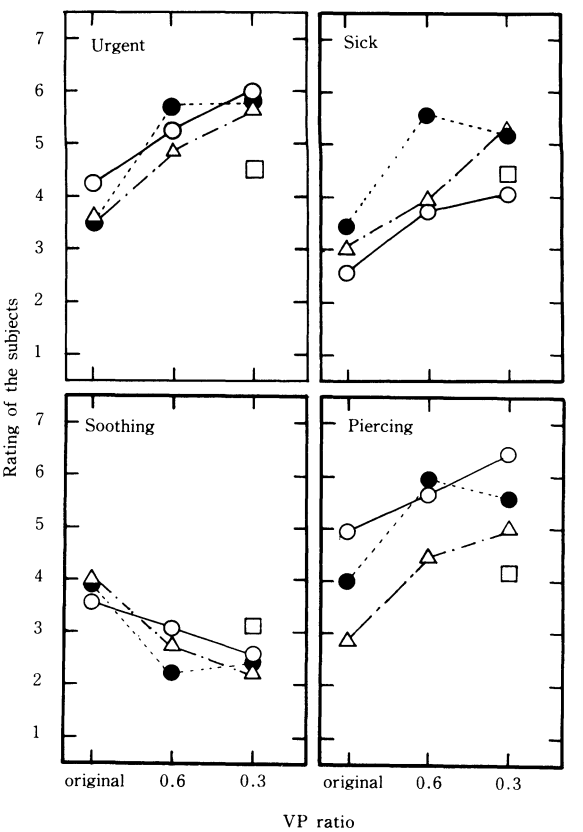

Fig. 1. Mean rating of each cry as the function of VP ratio. The open circle represents the mean rating of cry A. the open triangle is B, the filled circle is C and the open square is D.

Results

Fig. 1 shows the mean rating of each cry as the function of VP ratio. The open circle represents the mean rating of the cry A, the open triangle is B, the filled circle is C and the open square is D. The remaining four scales show the same pattern as the scale "urgent".

The rating of the subjects increased in all scales as the VP ratio went down. But the scale "soothing" displays opposite tendency. This result is not inconsistant with the other results because on the response sheat the direction of the meaning was inverse to that in the other scales.

The difference corresponded to the materials appeared in the two scales-sick and piercing. In piercing, this difference accorded with that of fundamental frequency, i.e. When the fundamental frequency got higher, the rating increased. But in "sick" it is not always so. The open square is the cry sound which had 0.6 VP ratio without manipulation. The rating to this cry sound is between the synthesized cries (VP=0.6) and the original cries. This suggests that the synthesize and natural decrease of the VP ratio brought clear effects in this rating task.

We did the three way analysis of variance (ANOVA) to scrutinize these results. All three main effects and first order interactions show significant effect but second order interaction is not significant (Table 2).

Table 2. The results of the analysis of variance.

Source	*SS*	*df*	*MS*	*F*	*P*
A materials	25.31	2	12.655	7.798	0.001
B VP ratio	589.08	2	294.539	189.495	0.000
C Scales	771.34	7	110.192	67.900	0.000
A*B	27.80	4	6.951	4.283	0.002
A*C	117.21	14	8.372	5.159	0.000
B*C	358.45	14	25.603	15.777	0.000
A*B*C	29.14	28	1.041	0.641	0.926
Erorr	2220.05	1368	1.623		

Discussion

In the present study we employed the digital speech synthesis technique to the infant cry and retested the meaning of the index which were extracted from our previous study. In all eight scales the ratings increased in proportion to the decrease of VP ratio. Only in the scale of piercing, the rating increased as the function of fundamental frequency. The artificial manipulation of the VP ratio effects the rating of subjects more than the fundamental frequency. This result supports our results of the multiple regression analysis.

The rating were changed according to the materials in the two scales-piercing and sick. We reported piercing could be explained by the VP ratio and the fundamental frequency but sick could not be explained by the multiple regression analysis between these physical characteristics and the rating of the subjects. On Figure 1 the rating of piercing increased as the VP ratio went down and fundamental frequency went up. However the scale of sick showed a different tendency. In this scale when the VP ratio decreased, the ratings increased. But the rating has no relation to the fundamental frequency. This difference according to the material could be explained by another physical characteristics — especially temporal cues such as duration or pause.

We did not inform the subjects as to whether the cries were natural or synthesized sounds. In spite of this, all subjects reported that all cries were heard as the natural cry sound. And they said that the baby sounded to have a cold or it had something wrong with its throat. These reports indicated that they perceived and judged the decrease of the VP ratio as the abnormal state of the voice source. This suggests that the PACOR method fairly reflected the cry production process.

The sound spectrograph has been used for the analysis of the infant cries. This equipment is useful for the quantitative analysis but it requires extensive work for the analysis and quantitative descriptions are limited. So we have tried to introduce the digital signal processing techniques into the analysis of the cry sounds. Among many methods we chose linear prediction technique — especially PACOR method — because firstly the model of PACOR is said to simulate the mechanism of the speech production system well and secondly it is useful not only for the analysis but also for the synthesis of speech. The operation of the VP ratio satisfied both these aspects.

From the observation of speech production system, we find that VP ratio can reflect the stability of the control of the vocal cords by the central nervous system. If the control of the vocal cords is instable or the vocal cords is damaged, the VP ratio will go down. We investigated the relation between VP ratio and the state of the crying infant and reported that some obstetric factors influences VP ratio (Okada, Murai, & Adachi., 1985, 1986). Fundamental frequency is also related to the condition of the vocal cords but it is not clear what kind of mechanisms are at work.

References

Adachi, T., Murai, N., Okada H., & Nihei, Y. **1985** Acoustic properties of infant cries and maternal perception. *Tohoku Psychologica Foria*, **44**, 51-58.

Bright, M. **1986** *Animal language*. Tokyo : Doubutu-shya (Transelated into Japanese).

Lounsbury, M.L. & Bates, J.E. **1982** The cries of infants of differing levels of percieved temperamental difficulties : Acoustic properties and effects on listeners. *Child Development*, **53**, 677-686.

Markel, J.D. & Gray, A.H. Jr. **1976** *Linear prediction of speech*. New York : Springer-Verlag Berlin Heidelberg.

Okada, H., Murai, N. & Adachi, T. **1985** An application of the linear prediction technique to the

analysis of infant cries (2). *Paper read at the 21th Congress of the Japan Society of Neonatology*. (in Japanese).

Okada, H., Murai, N., & Adachi, T. **1986** The relation between the physical characteristics of infant cries and their obstetric factors. *Paper read at the 22th Congress of the Japan Society of Neonatology*. (in Japanese)

Poter, F.L., Miller, R.H., & Marshall, R.E. **1986** Neonatal pain cries : Effect of circumcision on acoustic features and percieved urgency. *Child Development*, **57**, 790-802.

Simner, M.L. **1971** Newborn's response to the cry of another infant. *Developmental Psychology*, **5**, 136-150.

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.

(Received December 20, 1987)