2nd International Conference on
Spoken Language Processing (ICSLP 92)
Banff, Alberta, Canada
October 12-16, 1992

THE EFFECT OF FUNDAMENTAL FREQUENCY FOR VOWEL PERCEPTION IN INFANTS

Toshisada DEGUCHI[1)], Shigeru KIRITANI[2)], Akiko HAYASHI[1)] and Fumi KATOH[2)]

1)Faculty of Education, Tokyo Gakugei Univ., Koganei-shi, Tokyo, 184 Japan
2)Faculty of Medicine, Univ. of Tokyo, Hongo, Bunkyo-ku, Tokyo, 113 Japan

10.21437/ICSLP.1992-91

ABSTRACT

Human speech perception shows remarkable abilities to normalize wide variations in speech sounds such as due to speaker difference and context difference. Among such abilities is perceptual normalization of vocal tract size effects. We have been conducting series of experiments on the developmental study of this ability.

In our previous paper(K. Sakata et al., 1990), we reported the results of experiments comparing the ability of adults and young children aged 3-5. In this paper, preliminary results on the infants aged 5-7 month will be presented.

Experiments on infants have been conducted using the technique of conditioned head-turn responses. The experiments consist of two parts, experiment 1 and 2. Experiment 1 was conducted essentially following the procedures developed by Kuhl et al. Based on the experience in the experiment 1, slightly modified procedures were employed in the experiment 2. At the same time, control experiment on the adult subjects were conducted under experimental conditions similar to that of infants.

1.INTRODUCTION

In our previous paper, we reported the results of experiment comparing the ability of adults and young children aged 3-5. the results obtained from this experiments showed that children aged 3-5 already had an accuracy level of identification almost equal to that of adults. And it could be seen that all the subject group showed a similar shift in the boundary of the formant frequencies with changes in the fundamental frequency and higher formants. It was noted, however, that the amount of boundary shift appeared to be greater for young children than for adults.

These results suggest that children aged 3-5 show a perceptual normalization of vocal tract size in response to changes in fundamental frequency which are similar to but more native than those of adults. It appears that the effect of the fundamental frequency becomes smaller for adults.

The present study aims at investigating whether a similar perceptual normalization of vocal tract size exist in infant auditory perception.

2.Experiment 1

2.1 Procedure

Synthetic vowels were produced using a computer program simulating a terminal-analog speech synthesizer. A series of vowels which varied from /o/ to /a/ were synthesized by changing the first and second formant frequencies as follows.

F1=400,450,500,.........,900 Hz
F2=1.1F1+350 Hz

The parameter values were selected based on the results of the analysis in Fujisaki and Kawashima's study (Fujisaki and Kawashima,1968). Two set of vowels were synthesized using different fundamental frequencies (F0=100Hz,220Hz). For a given F0,
the frequencies of the higher formants were varied concurrently.

F0	100, 220 Hz
F3	6.3(F0+270) Hz
F4	1.4F3 Hz
F5	1.8F3 Hz

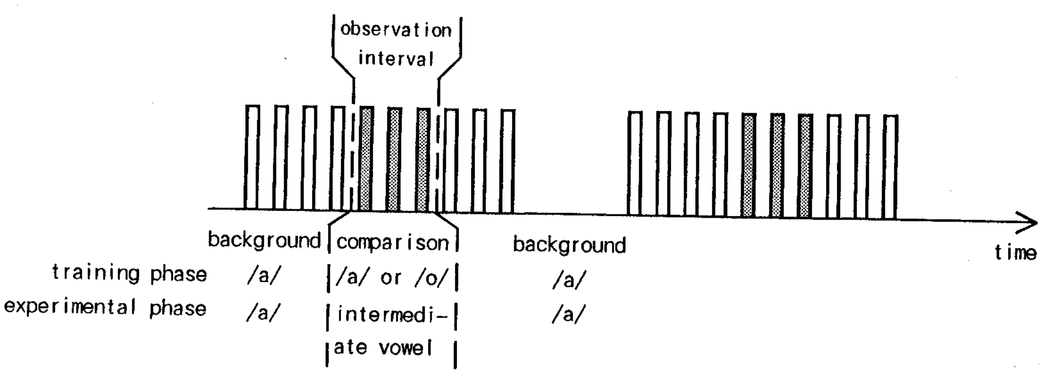

Figure 1. Stimulus presentation format in experiment 1 . Head-turn responses were measured during the observation interval.

vowel boundary due to the fundamental frequency change, there are some inter-subject variations. Subjects 1, 3 and 4 show effects similar to those of adults, namely, boundary formant frequency becomes higher with a higher fundamental frequency. Subject 2 does not show a significant shift in the vowel boundary. Subject 5 shows a vowel boundary shift which is opposite to that shown by adults. However, this infant required an exceptionally large number of training trails, compared to other infants, to be able to categorize the stimulus vowels, and thus may be considered a special case.

3.Experiment

3.1 Procedure

Based on the experience in the experiment 1, experiment 2 was conducted with slight modifications in the stimulus sounds and the method stimulus presentation.

The procedure in the Experiment 1, stimulus sounds were synthesized with constant pith which appears to have resulted in somewhat unnatural sound quality. Thus, in experiment 2, the stimulus sounds were synthesized with pith inflection as follows.

	0msec	100msec	140msec	500msec
F0=100Hz	100Hz	118Hz	118Hz	82Hz
F0=220Hz	220Hz	260Hz	260Hz	181Hz

As can be seen in Fig.5(a), a considerably large number of trials were required for training phases in the experiment 1. With an aim to reduce the number of trials in the training phase, several types of modifications in the stimulus presentation procedures were tested and the following procedure was adopted (Fig.6). The background stimulus is presented constantly at regular intervals until the experimenter judges the infant is ready and signals the computer to start the observation period. Then the control program randomly determines whether or not to present comparison stimulus. After presentation of the comparison stimulus the control program goes back to the conscious presentation background stimulus.

As for the reference data with adults, the data were recollected using new stimulus sounds . At the same time, considering the possible contrast effect of the background stimulus in the infants experiment, the method of the identification test with the adult was modified to present precursor stimuli preceding target stimuli. Precursor stimuli was the same as the background stimuli in infants' experiment and repeated 4 times.

3.2 Results

Fig.5(b) shows that number of trials in the training phase was greatly reduced with the new method, at nearly the same accuracy of discrimination with that in experiment 1(Fig.7). Fig.8 shows the boundary shift associated with the fundamental frequency change obtained for 5 infants. Average results for the 7 adults are also shown in the figure. It can be seen in the figure that subjects on the average show the shift in vowel boundary due to the fundamental frequency change. Although there are some inter-subject variations.

Above results suggest that infants show the effect of boundary change due to the fundamental frequency similar to that of adults. However, within above experimental paradigm, the results may be interpreted as reflecting the change in the discriminability. Further experiments are now being considered to confirm the change in he vowel category boundary including mixed presentation of vowels with different fundamental frequency within one experimental session.

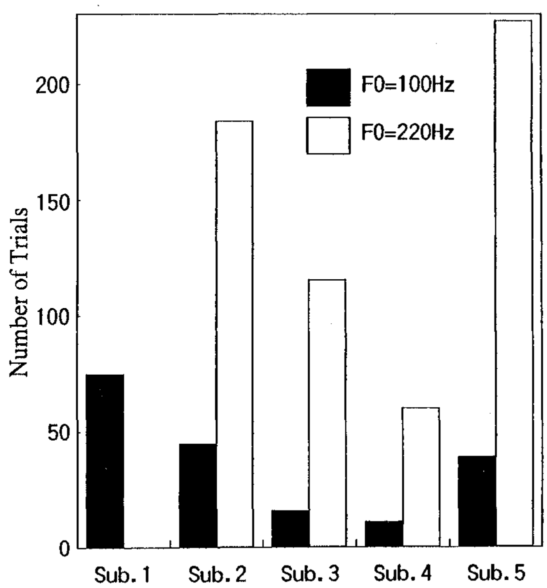

Figure 5(a). Number of the trials required to pass the criterion at the training stage in experiment 1.

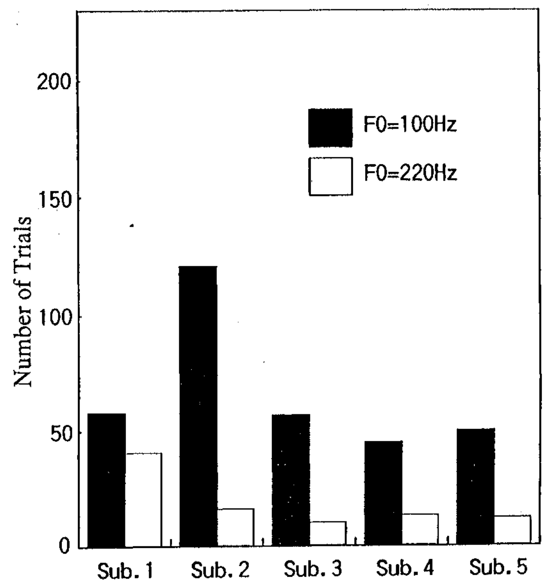

Figure 5(b). Number of the trials required to pass the criterion at the training stage in experiment 2.

The effect of changes in the fundamental frequency and the formant frequencies on vowel boundaries in infants were measured using the technique of the conditioned head-turn response.

The experimental procedure essentially follows that developed by Kuhl (1979). An infant is held by a caretaker sitting in a booth. A loudspeaker is located at a 90-degree-angle to the left of the infant and a visual reinforcer (a mechanical animal toy) is placed on the speaker. The stimulus sounds are delivered through the speaker, and the infant is conditioned to make a head-turn response toward the visual reinforcer when the stimulus is changed from one speech sound to another.

In each trial, a series of ten stimulus sounds is presented. The method of stimulus presentation is explained in Figure 1. In the training phase, two synthetic vowels, a representative /o/ and /a/, are used as stimulus sounds. One of the two synthetic vowels is presented four times as the background stimulus with intervals of two second. The stimulus sound is then changed to the other vowel and presented three times. Following the change of the stimulus sound, the reinforcer is activated. The infant learns to anticipate the reinforcer and make a head-turn with the change in the stimulus sound.

In the second stage of the training phase, the stimulus sound is changed in half of the trials but not changed in the remaining half. The infant learns to make a head-turn only when the stimulus sound is changed. The infant remained in the second stage of training, until the performance criterion, nine out of ten consecutive trials correct, was met.

In the experimental phase, the background stimulus is one of the representative vowels used in the training phase, and the test stimuli are the intermediate vowels (See Fig.1 for illustration). The frequency of the head-turn for each intermediate vowel is recorded.

Five infants aged 5-months to 7-months served as subjects.

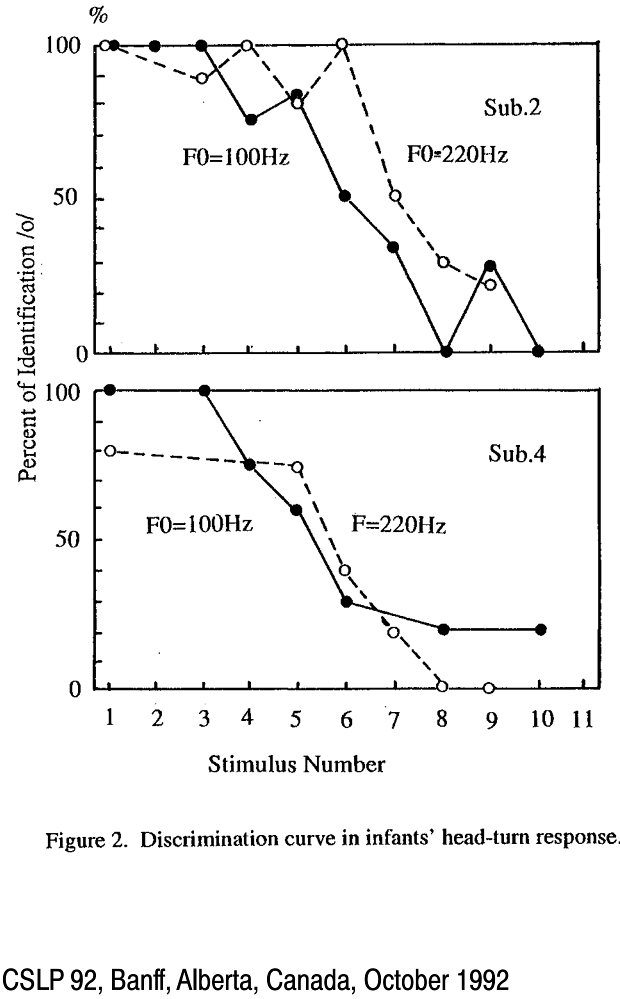

Figure 2. Discrimination curve in infants' head-turn response.

2.2 Results

Figure 2 shows an example of the response pattern obtained for a 6-month-old infant who could categorize the synthetic vowels in this experimental paradigm. Generally, all five infants have been able to discriminate between given stimulus vowels, although, as shown in Fig.3, the accuracy of identification is worse than that of adults. The response curves in Figure 2 were approximated by the method of maximum-likelihood and the position of the vowel boundary were estimated. Fig.4 shows vowel boundaries estimated for 5 infants. As for the shift in the

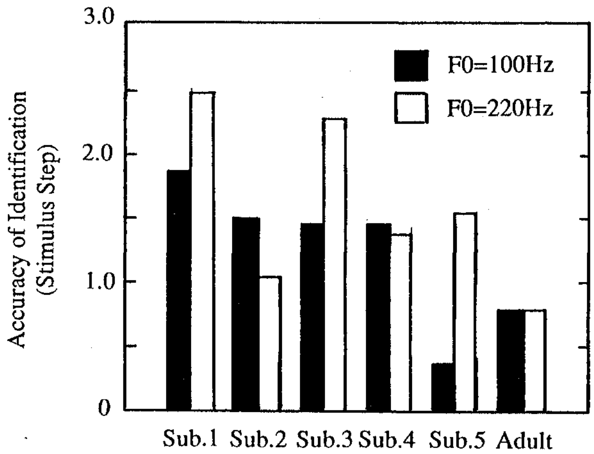

Figure 3. Accuracy of discrimination for 5 infants in experiment 1 and average result for adults.

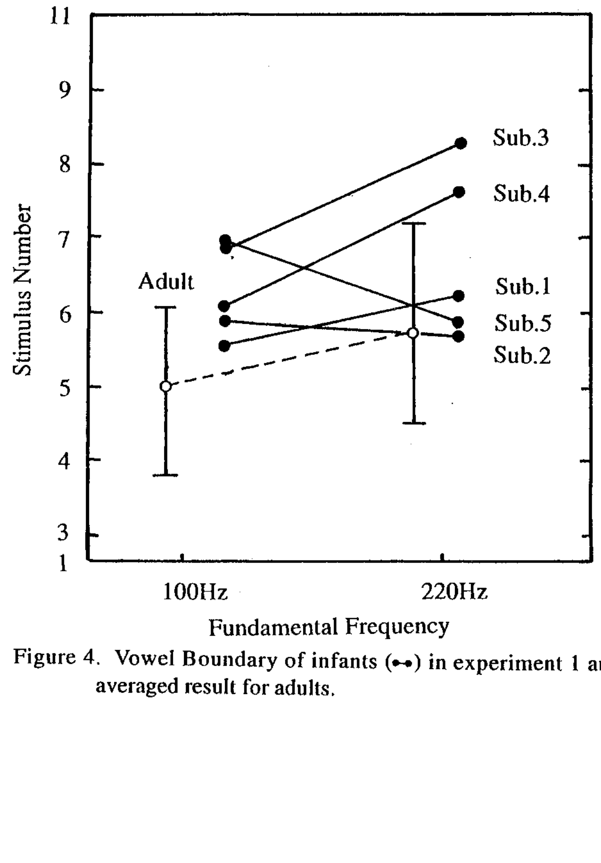

Figure 4. Vowel Boundary of infants (•–•) in experiment 1 and averaged result for adults.

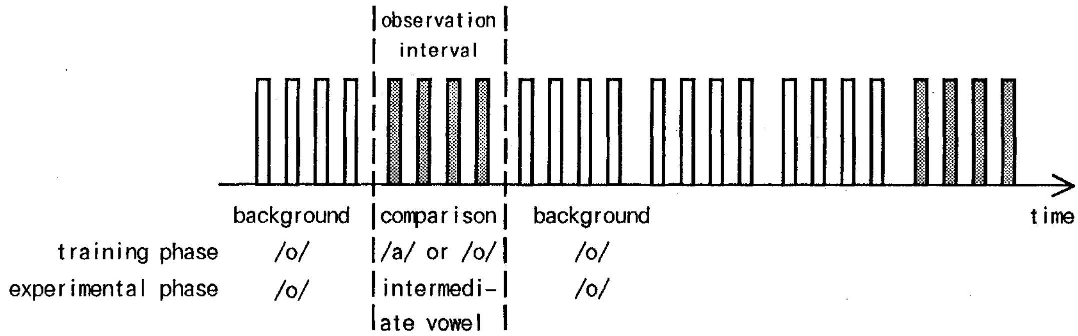

Figure 6. Stimulus presentation format in experiment 2.

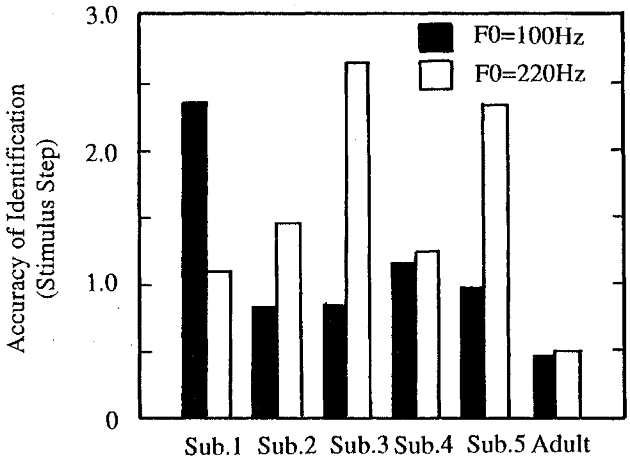

Figure 7. Accuracy of discrimination for 5 infants in experiment 2 and averaged result for 7 adults.

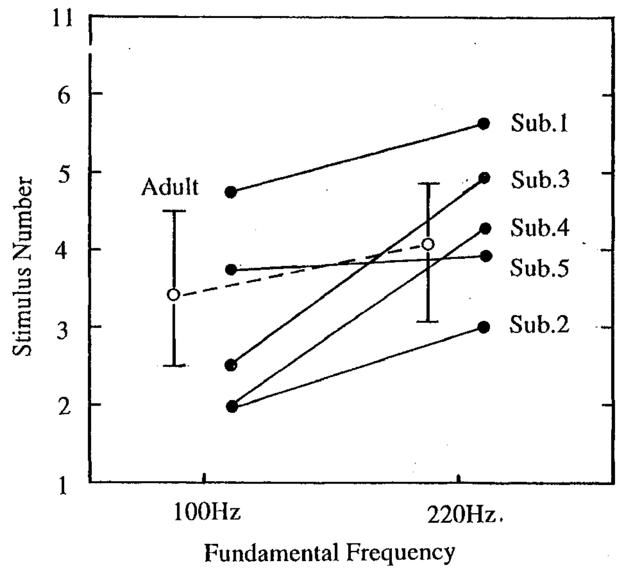

Figure 8. Vowel Boundary of 5 infants (•–•) in experiment 2 and averaged result for 7 adults.

References

1)Kuhl,P.K. :Perception of speech and sound in early infancy, In Salapatek, P. and Cohen, L., (eds.), Handbook of Infant Perception, 2, From Perception to Cognition, New York Academic Press, 275-373, 1987.

2)Fujisaki, H. and Kawashima, T.:The role of pith and higher-formants in the perception of vowels, IEEE Transactions on Audio and Electroacoustics, AU-16(1), 1988.

3)Kuhl, P. K.:Speech perception in early infancy:Perceptual constancy for spectrally dissimilar vowel categories, J. Accost. Soc. Am., 66,1668-1679, 1979.

4)Sakata, K. et al.,:Ability of perceptual normalization of vocal tract size in young children, ANN. Bull. RILP, No.24, 1990.