Ann. Bull. RILP (1983) No. 17, 201-209

DISCRIMINABILITY OF SECOND FORMANT FREQUENCIES IN HEARING-IMPAIRED CHILDREN

Tatsuo Nakagawa, Shuzo Saito, and Tomoyoshi Yoshino***

Introduction

When human voices are reproduced at unusually high levels, their intelligibility is lowered. There are two reasons for this phenomenon. First, from the shapes of equal-loudness countours[1] it can be concluded that the loudness balance among each frequency component in a speech sound will change as a function of the reproduced level. With an increase in intensity, the loudness growth is greater for the lower frequencies than for the middle frequencies. Second, from the masked threshold curves[2] it can be concluded that the lower frequencies become more effective in masking the higher frequencies.

Normally, persons with a hearing loss wear hearing-aids and hear sounds at a high intensity level. A few studies have reported that some hearing-impaired persons can not perceive important information carried in the middle to high frequency components. The results of experiments performed at Gallaudet College[3,4] implicate the "upward spread of masking" as a contributing factor in the poor identification of place of articulation. But other studies[5,6] could not detect an upward spread of masking effects in their experimental paradigms.

The purpose of the present experiment was to determine the effects of first formant (F1) amplitude attenuation on the discrimination of second formant (F2) frequencies using the five synthetic vowels, synthetically produced.

Method

1. Subjects

Four hearing-impaired subjects between the ages of 8 and 12 participated in this experiment. Pure tone audiograms for these subjects (S1-S4) are shown in Fig. 1. All of the subjects had bilateral flat or gently sloping sensorineural hearing loss. The test ear for each subject was the ear in which he or she wore the monaural hearing-aid. The subjects were selected from the hard-of-hearing program in the normal public schools.

* Graduate Program in Special Education, Tsukuba University
** Institute of Special Education, Tsukuba University

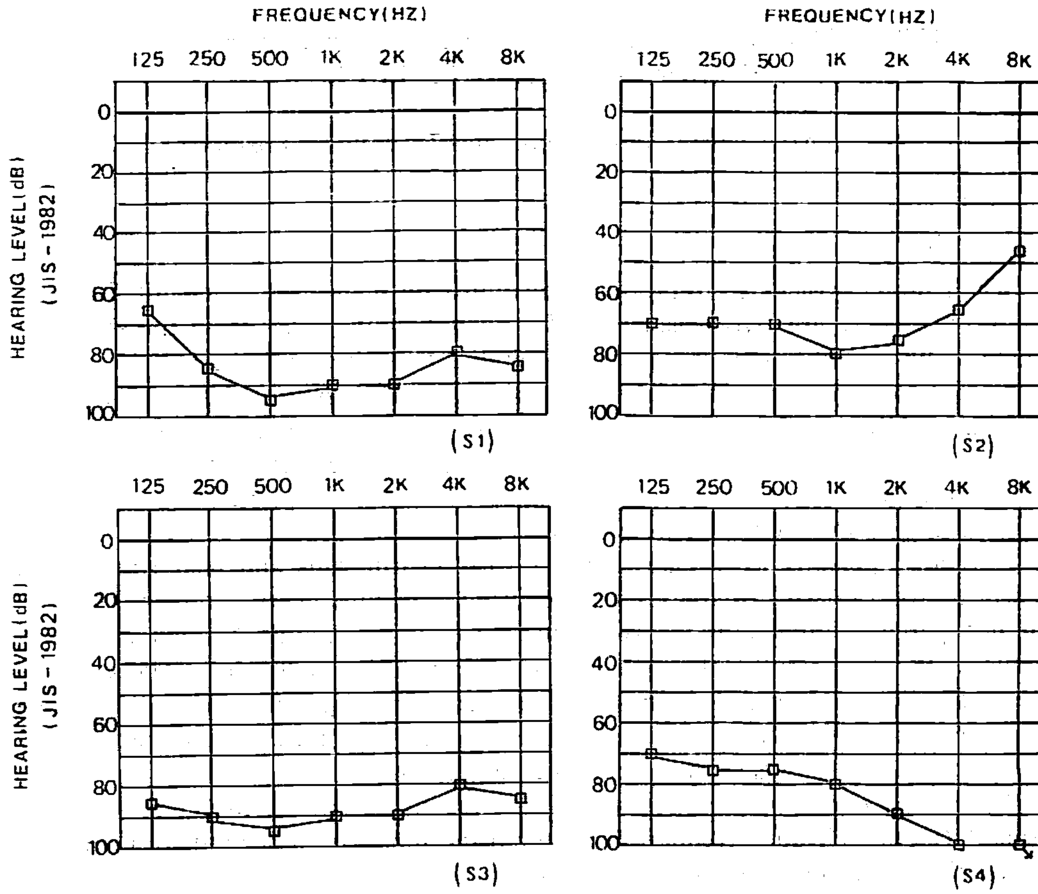


Fig. 1 *Audiograms for each of the four hearing-impaired subjects used in this experiment. Thresholds for the impaired ear (□) are plotted.*

2. *Stimuli*

The stimuli were synthesized using a computer-programmed, cascade-connected formant synthesizer on the basis of data analyzed by the linear prediction method. Table 1 shows the parameter values used in synthesizing each reference stimulus. The temporal fluctuation in both fundamental frequency and overall amplitude was simulated based on an analysis of natural utterance data.

In order to yield stimuli with an attenuated F1 amplitude, the stimuli were recorded using the apparatus diagrammed in Fig. 2. The synthetic stimuli were played back by tape recorder 1. Then the stimuli were passed through low-pass filter 1 and a high-pass filter with limiting attenuation slopes steeper than 60 dB/ octave beyond the cut-off frequency. After being passed through the low-pass filter 1, the signals were attenuated. The outputs from the two channels were mixed; low-passed at 5 kHz (low-pass filter 2); amplified, and recorded on tape recorder 2.

Table 1 *The formant frequencies, bandwidths, and duration for each synthesized reference stimulus*

	PHONE	/a/	/i/	/u/	/e/	/o/
F1	FREQUENCY(HZ)	690	305	335	515	500
	BANDWIDTH(HZ)	100	20	50	80	100
F2	FREQUENCY(HZ)	1100	2240	1250	1985	800
	BANDWIDTH(HZ)	130	90	120	70	150
F3	FREQUENCY(HZ)	2750	3285	2280	2600	2800
	BANDWIDTH(HZ)	180	200	150	230	200
F4	FREQUENCY(HZ)	3750	3700	3550	3750	3600
	BANDWIDTH(HZ)	200	200	200	170	200
F5	FREQUENCY(HZ)	4050	-	-	4100	-
	BANDWIDTH(HZ)	180	-	-	220	-
	DURATION (MS)	245	230	240	245	240

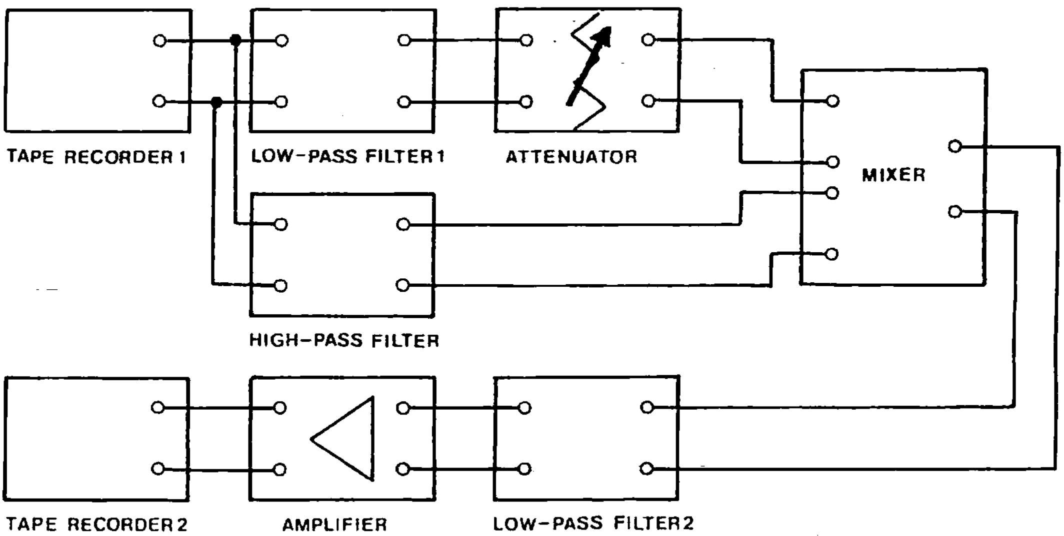


Fig. 2 *Block diagram of the equipment used for recording the test stimuli and attenuating the F1 amplitude.*

3. *Procedure*

The subjects, sitting in a sound-proof room, were presented pairs of the stimuli in which one stimulus, the reference, had either the same or a different F1 frequency compared to the other, the test stimulus. After the presentation of each stimulus pair the subjects responded by pressing one of two response buttons on the basis of whether the sounds in the pair seemed to be the same or different.

For each subject, the stimulus presentation level was her or his most comfortable listening level (MCL). The MCL was determined by presentating the five reference stimuli to each subject and determining the level at which each subject felt best hearing them.

Each vowels F1 amplitude was changed in three steps (Table 2). In the 0 dB condition, the F1 was at its normal amplitude relation to F2.

For each attenuation condition, the 60 stimulus pairs were arranged in 6 blocks of ten pairs. Each block contained three identical pairs and seven different pairs. The maximum shift in the F2 frequency for /i/, /e/, /a/, /o/, and /u/ was 845, 650, 385, 325 and 650 Hz, respectively. These intervals were equally divided into 7 steps. Therefore, the frequency variation steps used in determining the DL's for each stimulus was as follows: −120 Hz for /i/; −93 Hz for /e/; +55 Hz for /a/; +46 Hz for /o/: and +93 Hz for /u/. The presentation order of the pairs within each block was randomized. The time interval between the stimuli was 0.5 sec: the interval between the pairs was 6 sec; and the interval between the blocks was 12 sec.

For each test session, one block of each attenuation condition was presented to the subjects in random order. The sessions were scheduled for one day out of each school week and continued for six weeks.

Table 2 *The amount of F1 amplitude attenuation for each vowel*

PHONE	F1 AMPLITUDE ATTENUATION (dB)		
/a/	0	5	10
/i/	0	8	16
/u/	0	8	16
/e/	0	7	14
/o/	0	7	14

Results and Discussion

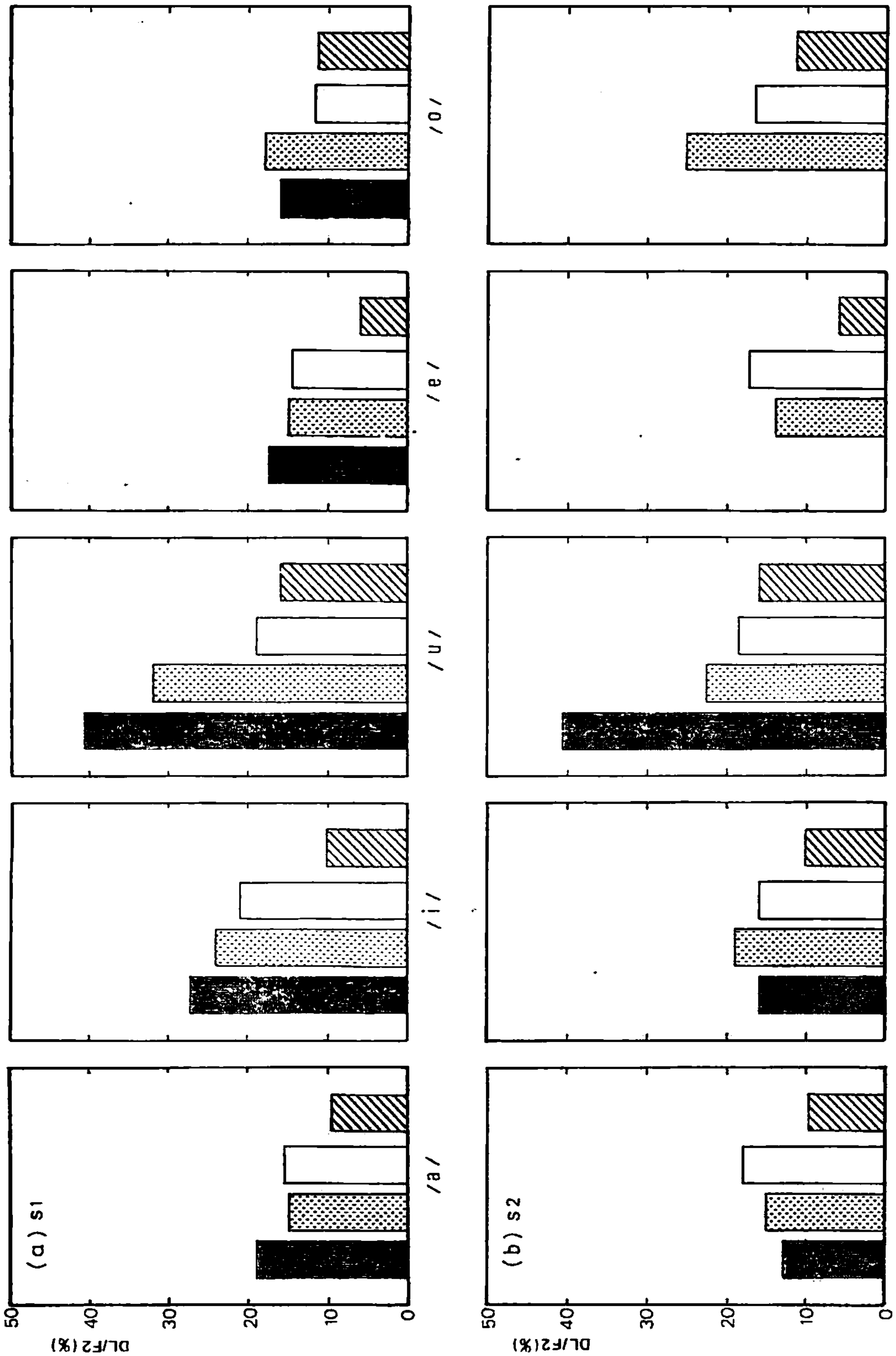
The percentage of pairs judged different was calculated separately for each condition. From these data, psychometric functions were obtained for each subject. The difference limen (DL) for the F2 frequency was defined as the point along the F2 frequency continuum at which 50% of the judgments were different and 50% were the same. This was estimated by the Müller-Urban process.

Table 3 shows the DL's and the relative DL's defined by the ratio of the DL's to the corresponding F2 reference frequencies. For one subject (S4), the DL's could not be determined. For the remaining three subjects, decreasing the energy of F1 improved the discrimination of the F2 frequencies for /u/. Particularly, attenuating the energy of F1 by 16 dB made it possible for the subjects to attain a discrimination performance equal to normal-hearing children[7].

Table 3 *DL's and relative DL's (DL/F2) for each F2 frequency for the subjects*

PHONE	F1-ATT. (dB)	S1 DL (HZ)	S1 DL/F2 (%)	S2 DL (HZ)	S2 DL/F2 (%)	S3 DL (HZ)	S3 DL/F2 (%)	S4 DL (HZ)	S4 DL/F2 (%)
/a/	0	205.2	18.7	141.9	12.9	183.6	16.7	–	
	5	168.6	15.3	166.5	15.1	162.0	14.7	–	
	10	172.8	15.7	195.0	17.7	318.0	28.9	–	
/i/	0	606.5	27.1	366.6	16.4	–		–	
	8	538.9	24.1	425.1	19.0	523.3	23.3	–	
	16	472.6	21.1	366.6	16.4	–		–	
/u/	0	510.0	40.8	510.0	40.8	530.0	42.4	–	
	8	397.5	31.8	281.0	22.5	218.5	17.5	–	
	16	233.0	18.6	235.0	18.8	182.5	14.6	–	
/e/	0	341.0	17.2	–		510.0	25.7	–	
	7	306.5	15.4	271.0	13.6	435.0	21.9	–	
	14	292.0	14.7	343.5	17.3	368.5	18.6	–	
/o/	0	130.3	16.3	–		–		–	
	7	142.5	17.8	201.3	25.2	178.0	22.3	–	
	14	93.3	11.6	134.0	16.8	–		–	

However, for the other vowels, there was an enormous between-subject variability in discrimination performance. In one subject (S1), discrimination improved as the attenuation of F1 was increased (Fig. 3(a)). For this subject the effects of the attenuation were clearer for the back vowels than for the front ones. Of the two subjects who performed in a different manner on the discrimination of the F2 frequencies for /i/, /e/, and /a/, one (S2) could discriminate better in the 0 dB condition for /a/ (Fig. 3(b)) and the other (S3) could discriminate better in the intermediate attenuation condition for /a/, /i/ and /o/ (Fig. 3(c)).


(a) s1
(b) s2
/a/
/i/
/u/
/e/
/o/
DL/F2 (%)
DL/F2 (%)
0
10
20
30
40
50

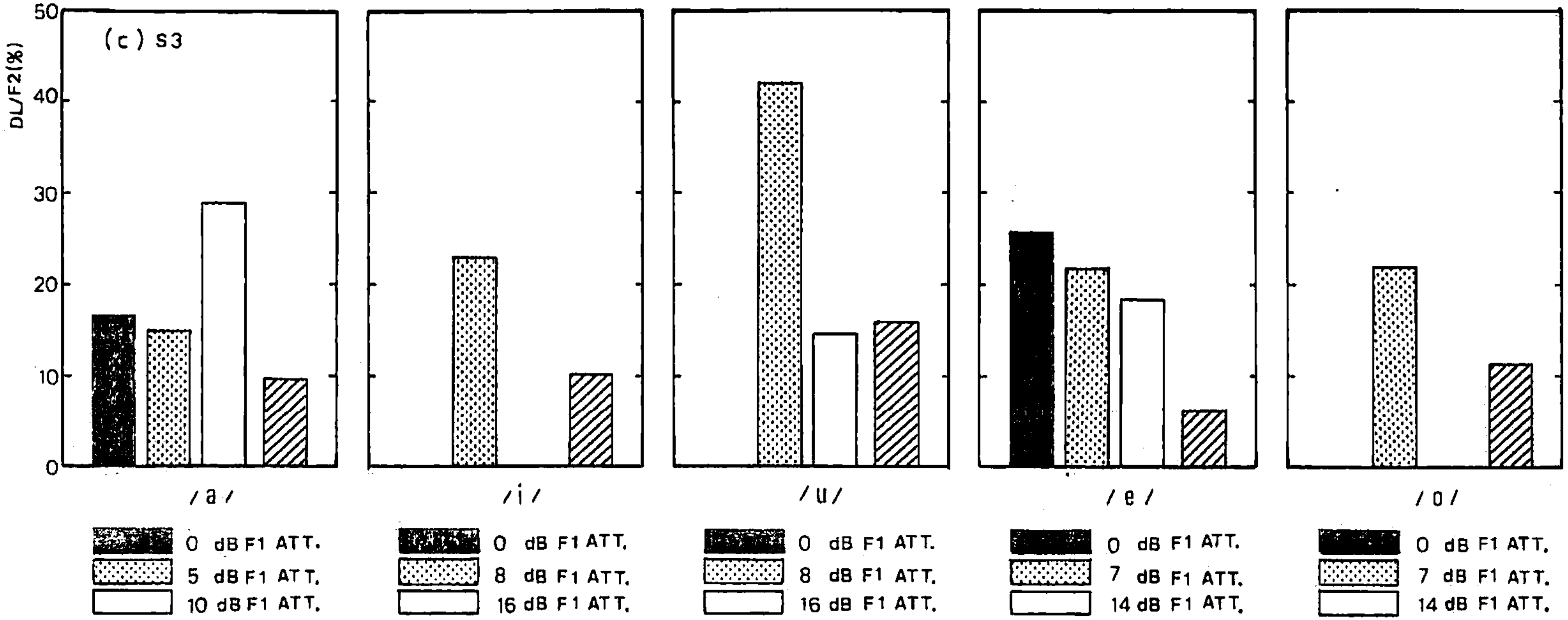


Fig. 3 *The effects of F1 amplitude attenuation on the discrimination of F2 frequencies in each condition. Results from normal-hearing children*[7] are also shown to the extreme right of each panel in this figure (▨).

Subjects with a similar audiogram configuration and a similar amount of hearing sometimes performed in a radically different manner. A similar result has been noted in many experiments using the hearing-imapired. This means that threshold audiograms do not correlate with suprathreshold processing such as that used in this task. Consequently, it is not necessarily reasonable to make a frequency analysis of the ear based on audiograms.

For subjects who can perform this task, there is a general tendency for the effect of the F1 amplitude attenuation to have a greater effect on F2 frequency discrimination for /u/ than for other vowels. A spectral analysis of the vowels indicates that in the vowel /u/, the amplitude of F1 is greater than that of F2. Therefore, it seems that a greater F1 masking effect occurs in this vowel than in the other vowels. The relative amplitude of F1 and F2 seems to have a great effect on F2 discrimination. The implication of this study is that low-frequency speech elements can not always be expected to mask high-frequency elements at high listening levels.

It is necessary to continue our study using a greater number of subjects. In addition, other studies such as psychoacoustical measures of frequency analysis and frequency selectivity may be important. We are now developing a more complete investigation.

Summary

The DL's for second formant frequencies were measured in four hearing-impaired children. To assess whether an upward spread of masking is responsible for poor speech perception, the subjects were tested using stimuli in which the first formant amplitude was attenuated. There was a high variability in DL's among the subjects with a similar audiometric configuration. The results were interpreted as showing that low-frequency speech elements can not always be expected to mask high-frequency elements at high listening levels.

Acknowledgments

We would like to express our deep gratitude to Sotaro Sekimoto and Hiroshi Imagawa (University of Tokyo), to Hiroshi Irii (Musashino Electrical Communication Laboratory), and to Sanetomi Eguchi (Tsukuba University) for making it possible for us to do the present research. We are also greatly indebted to Hisashi Kado (Electrotechnical Laboratory) for his statistical advice.

References

1. Robinson, D.W. and R.S. Dadson (1956): "A redetermination of the equal-loudness relations for pure tones", Br. J. Appl. Phys. 7, 166-181.
2. Egan, J.P. and H.W. Hake (1950): "On the masking pattern of a simple auditory stimulus", J. Acoust. Soc. Am. 22, 99-109.
3. Danaher, E.M., M.J. Osberger and J.M. Pickett (1973): "Discrimination of formant frequency transitions in synthetic vowels", J. Speech Hear. Res. 16, 439-451.
4. Danaher, E.M. and J.M. Pickett (1975): "Some masking effects produced by low-frequency vowel formants in persons with sensorineural hearing loss", J. Speech Hear. Res. 18, 261-271.

5. Van de Grift, S., M.F. Dorman and J.R. Franks (1980); "Identification of synthetic /bdg/ by hearing-impaired listeners under monotic and dichotic formant presentation", J. Acoust. Soc. Am. 67, 1031-1040.
6. Van Tasell, D. (1980); "Perception of second-formant transitions by hearing-impaired persons", Ear Hear. 1, 130-136.
7. Nakagawa, T., S. Saito and T. Yoshino (1982); "Tonal difference limens for second formant frequencies of synthesized Japanese vowels", Ann. Bull. RILP, 16, 81-88.