STUDIA PHONOLOGICA XXV (1991)

Speech Intelligibility for Acoustic Simulation of Cochlear Implant System

Shigeyoshi KITAZAWA, Mitsuaki NAITO, Kazuyuki MURAMOTO, Koshi YAMADA
Junji SAKAKIHARA, Juichi ITO and Shuji DOSHITA

ABSTRACT

For evaluating our digital signal processing system which is compatible to the WSP (Wearable Speech Processor) of the Nucleus 22 channel cochlear implant system, we performed listening experiments with our evaluation system. We carried out the experiment with normal hearing subjects using standardized video recording and an originally developed acoustic simulator. By analyzing experimental results we confirmed improvements of our digital speech processing algorithm.

1. INTRODUCTION

The nucleus 22 channel cochlear implant system consists of a cochlear implant RSU (Receiver Stimulator Unit) and a WSP(Wearable Speech Processor). The WSP extracts F0(pitch), F1(first formant), F2(second formant) of speech sound using analog electric circuit and transmits these information to the cochlear implant RSU. In our previous report[1], we described a digital system (DSP: Digital Speech Processor) which is compatible to the analog WSP, and an originally developed acoustic simulator. We were now evaluating this DSP and WSP with normal hearing subjects using an acoustic simulator.

Using the results of these experiments, we improved the DSP algorithm and the acoustic simulator, and the results were confirmed by listening experiments.

2. METHOD

We carried out the listening experiment, using standardized video recording[a)], as follows:

1) Evaluation speech sound from video tape was taken into each SP.

Shigeyoshi KITAZAWA (北澤茂良): Associate professor, Department of Computer Science, Faculty of Engineering, Shizuoka University
Mitsuaki NAITO (内藤光明), Kazuyuki MURAMOTO (村本一之): Graduate student, Department of Computer Science, Faculty of Engineering, Shizuoka University
Koshi YAMADA (山田耕史): Engineer, Department of Computer Science, Faculty of Engineering, Shizuoka University
Junji SAKAKIHARA (榊原淳二): Department of Otolaryngology, Faculty of Medicine, Kyoto University
Juichi ITO (伊藤壽一): Ohtsu Red Cross Hospital
Shuji DOSHITA (堂下修司): Professor, Department of Information Science, Faculty of Engineering, Kyoto University

Table 1. Frequency of each channel of acoustic simulator.

Channel	Frequency (Hz)	Channel	Frequency (Hz)
ch1	3796	ch11	1309
ch2	3411	ch12	1176
ch3	3066	ch13	1059
ch4	2753	ch14	917
ch5	2478	ch15	760
ch6	2235	ch16	635
ch7	2008	ch17	525
ch8	1804	ch18	431
ch9	1623	ch19	360
ch10	1458	ch20	164

2) At the same time, sounds generated by the acoustic simulator were recorded on a tape.
3) Human subjects listened to the tape and wrote down their perception.

2.1 Condition

Experiments were performed under open condition; listeners were not given any concrete information about syllables, words and sentences, and were not given any preliminary trainings.

2.2 Subjects

Normal hearing 12 or 13 male subjects with ages from 22 to 24, participated in this experiments.

2.3 Speech Samples

50 syllables, 50 words and 29 sentences from a female speaker were picked up from the video tape.

2.4 Frequency of acoustic simulator

Frequencies for each channel of the acoustic simulator are disposed as shown in Table 1.

2.5 Improvement for DSP algorithm

After carrying out the listening experiments with the first version of DSP system, we imporved the DSP algorithm as follows;

1) Addition of 3.0 kHz LPF(Low Pass Filter) of Chebyshev IIR (Infinite Impulse Response) for extracting F2(second formant) correctly.
2) Change of the filter order for estimating F0 (pitch) from 4 to order 2.
See section 4 for further explanations.

3. Experimental Results

3.1 Results of the first DSP system

a) Video tape provided from Y. Fukuda was played by herself: entitled "*Evaluation of speech sound recognition for cochlear implant patients*".

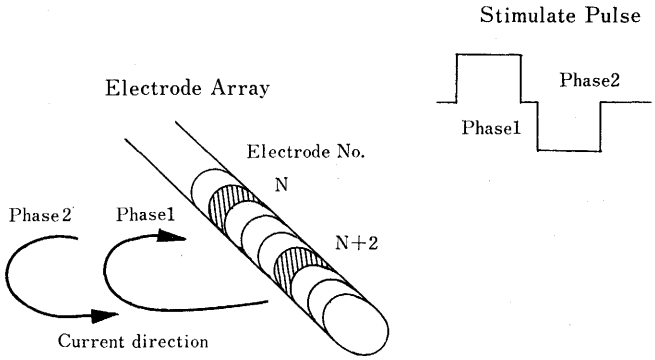

Fig. 1. Illustration of electrodes and electric current flow. Electric stimulate pulses flow between the No. N electrode and the No. N+2 electrode.

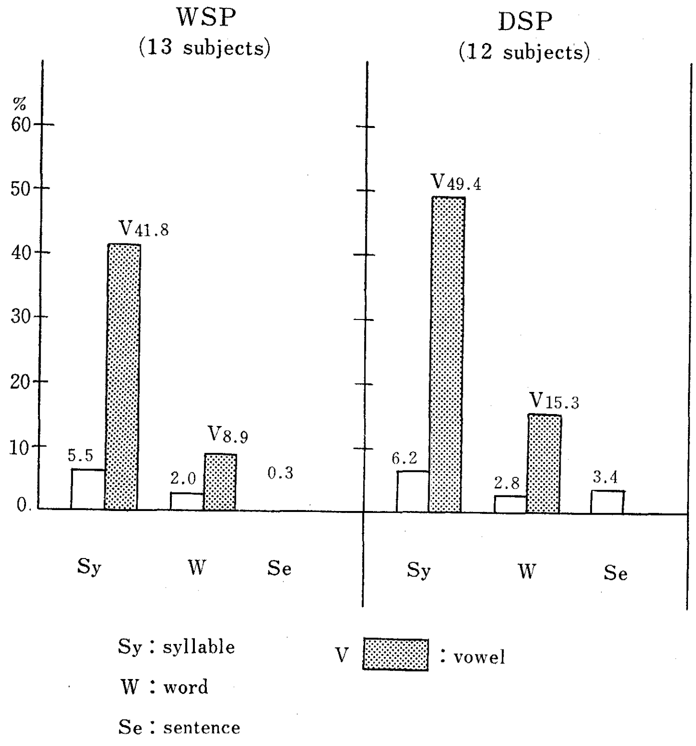

Fig. 2. Experimental results of the first version of our DSP and the WSP.

Fig. 2 shows the result of the first version of our DSP and the WSP using the first version of simulator. Fig. 2 shows that the recognition rate of the DSP is higher than that of the WSP.

The results of vowel recognition of syllables in the DSP and the WSP are shown in Table 2 and Table 3, respectively. In these tables both tendencies of error were different. /a/ was perceived corerctly more than 50% in both systems, but sometimes /a/ was misperceived as /o/ in the WSP. /i/ was perceived correctly more than 50% with both systems, but sometimes /i/ was misperceived as /e/ in the WSP.

Table 2. Vowel recognition of syllable using the first version of acoustic simulator and the DSP (12 subjects).

		Answer (%)					
		a	i	u	e	o	No answer
Stimuli	a	78.2	0.6	1.3	1.3	5.1	13.5
	i	0	57.4	6.5	8.3	5.6	22.2
	u	0.9	42.6	14.8	9.3	9.3	23.1
	e	0	4.2	7.3	20.8	52.1	15.6
	o	9.1	0	9.8	7.6	57.6	15.9

Table 3. Vowel recognition of syllable using the first version of acoustic simulator and the WSP (13 subjects).

		Answer (%)					
		a	i	u	e	o	No answer
Stimuli	a	50.3	1.2	1.8	4.7	35.5	6.5
	i	0.9	56.4	5.1	24.8	6.0	6.8
	u	41.0	0.9	20.5	4.3	18.8	14.5
	e	3.9	0	19.2	20.2	44.2	12.5
	o	5.6	3.5	14.0	13.3	55.2	8.4

/u/ was perceived correctly less than 50% in both systems, and /u/ was misperceived as /a/ in the WSP on the other hand as /i/ in the DSP system. /e/ was perceived correctly about 20% and /e/ was mispercieved as /o/ more than 50% in both systems.[b] /o/ was perceived correctly more than 50%.

3.2 Results of the improved DSP system

Fig. 3 shows the results of the revised version of our DSP and the WSP using the revised version of the simulator.[c] These figures show that the recognition rate of the DSP is higher than that of the WSP.

The results of vowel recognition of syllables in the DSP and the WSP are shown in Table 4 and Table 5, respectively. In these tables errors like /u/ was misperceived as /i/ and /e/ was mispercived as /o/ decreased as compared with the previous systems. The WSP showed the same error tendency as before, but correct recognition for /i/ and /o/ decreased compared to before improvement.

---

b) It was considered that vowel /e/ misperceived as /o/ resulted from the polarity of the biphasic pulse of acoustic simulator. So we improved the acoustic simulator.

c) Improvement of acoustic simulator: Electirc stimulate pulses flow between the No. N electrodes and the No. N+2 electrode as shown in Fig. 1. The first version of the acoustic simulator generates two sounds with different frequency, according to the polarity of the biphasic pulse between N and N+2. We carried out the first listening experiments using this acoustic simulator. Analyzing the results of these experiments, we concluded that this method affected speech recognition. Further we improved the acoustic simulator to generate only one sound for each burst signal. This was done by inhibiting the 2nd sound caused by the reverse pulse.

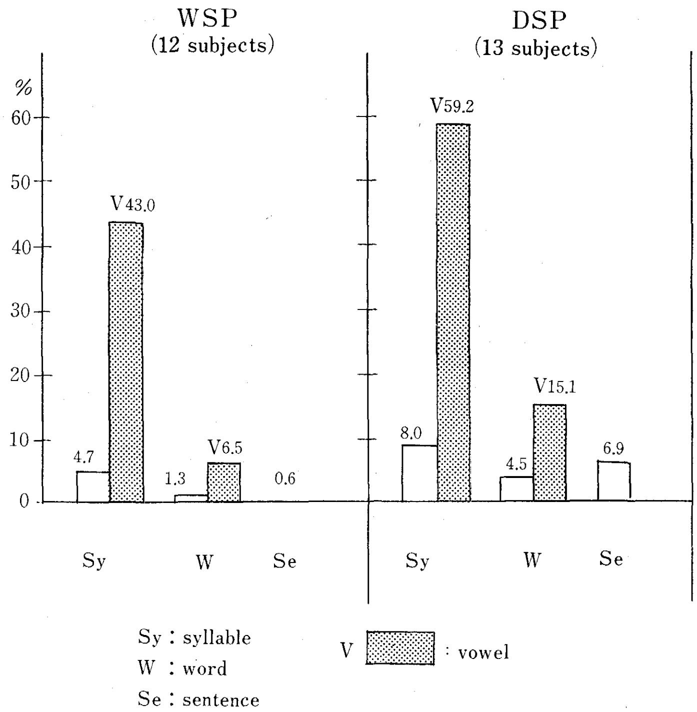

Fig. 3. Experimental results of the revised version of our DSP and the WSP.

Table 4. Vowel recognition ot syllable using revised version of acoustic simulator in revised version of DSP (13 subjects).

		Answer (%)					
		a	i	u	e	o	No answer
Stimuli	a	75.7	4.1	5.3	1.8	8.3	4.7
	i	2.6	66.7	11.1	9.4	6.0	4.3
	u	11.1	6.0	39.3	2.6	39.3	1.7
	e	4.8	12.5	6.7	56.7	15.4	3.8
	o	25.2	1.4	10.5	7.7	51.7	3.5

Table 5. Vowel recognition of syllable using revised version of acoustic simulator in the WSP (12 subjects).

		Answer (%)					
		a	i	u	e	o	No answer
Stimuli	a	69.9	2.5	4.5	4.5	13.5	5.1
	i	2.8	47.2	8.3	22.2	2.8	16.7
	u	30.6	14.8	23.1	5.6	7.4	18.5
	e	5.2	14.6	11.5	33.3	18.7	16.7
	o	14.4	15.9	6.8	25.0	28.8	9.1

4. Considerations

We evaluated the results of the first version of the DSP system. Table 6 shows the number of correct answers for each syllable. From this table and Table 2, we found that the vowel part of syllable /u/ was misperceived as /i/ and /e/ was mispercieved as /o/.

Fig. 4 shows the sound spectrogram of sample /ru/. Fig. 5 shows the plot of F1 (first formant) and F2 (second formant) extracted by using the first algorithm. In these figures we found that F3 (third formant) was extracted instead of F2 (second formant). Fig. 6 shows the plot of F1 (first formant) and F2 (second formant) using 3.0 kHz LPF. In this figure, F2 was extracted as F2. For investigating the 3.0 kHz LPF influence on F2 extraction from vowel /i/, a new algorithm was tested on vowel /i/. Fig. 7 shows the plot of F1 and F2 as obtained by the first algorithm, and Fig. 8 shows the results from the second algorithm. It can be seen from these figures, that the 3.0 kHz LPF have no significant influence on the extraction of formants in vowel /i/.

Table 6. Recognition of 50 syllables in the first version of DSP system (12 subjects). In this table Vowel is correct only vowel part of the syllable.

Syllable		Correct	Vowel	Syllable		Correct	Vowel
1	ga	4	11	26	ni	0	6
2	ma	0	6	27	i	5	7
3	ra	2	6	28	ri	0	5
4	mi	1	7	29	ji	0	5
5	do	1	2	30	de	0	2
6	de	0	0	31	te	0	1
7	no	0	9	32	su	0	1
8	mo	0	6	33	da	0	11
9	mu	0	3	34	ba	3	10
10	ne	0	3	35	go	0	6
11	wa	0	9	36	to	0	8
12	o	0	7	37	ki	0	9
13	ru	0	0	38	yo	2	6
14	fu	0	3	39	ro	2	10
15	ya	4	9	40	ho	1	9
16	ko	0	7	41	ka	0	10
17	ta	0	11	42	sa	0	10
18	a	3	9	43	chi	1	9
19	hi	2	10	44	se	0	4
20	so	0	6	45	yu	1	2
21	ku	0	2	46	na	0	11
22	shi	0	4	47	u	2	4
23	tsu	0	0	48	ha	2	9
24	me	0	3	49	ke	0	4
25	e	1	2	50	zu	0	1

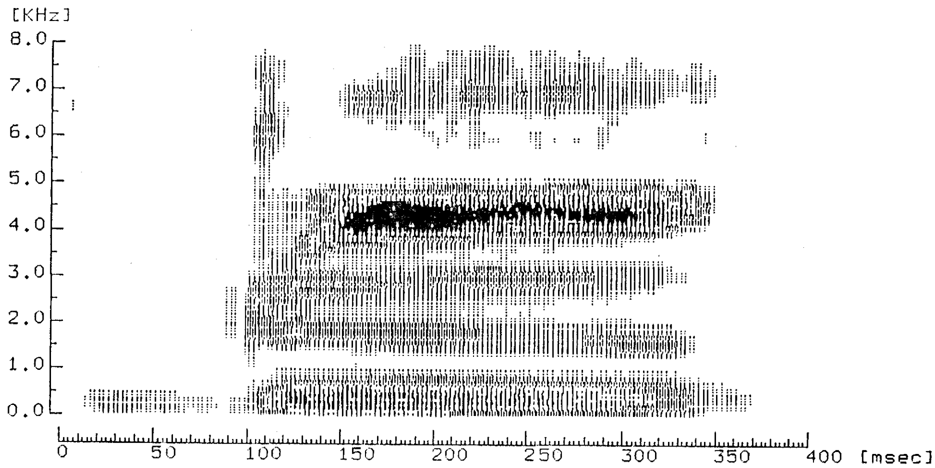

Fig. 4. Sound spectrogram of sample /ru/.

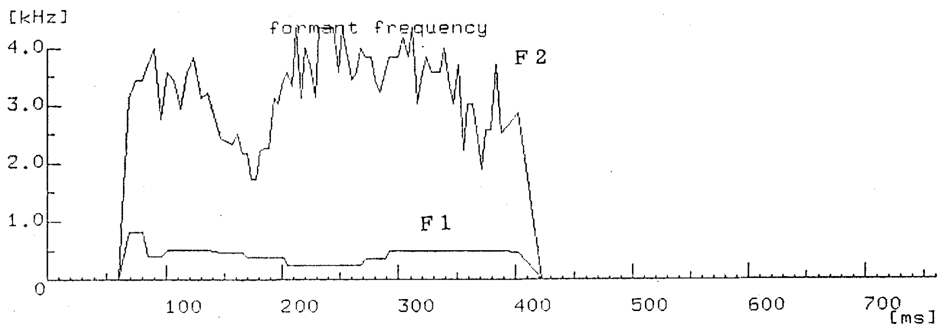

Fig. 5. Plot of F1 (first formant) and F2 (second formant) extracted by using the first algorithm.

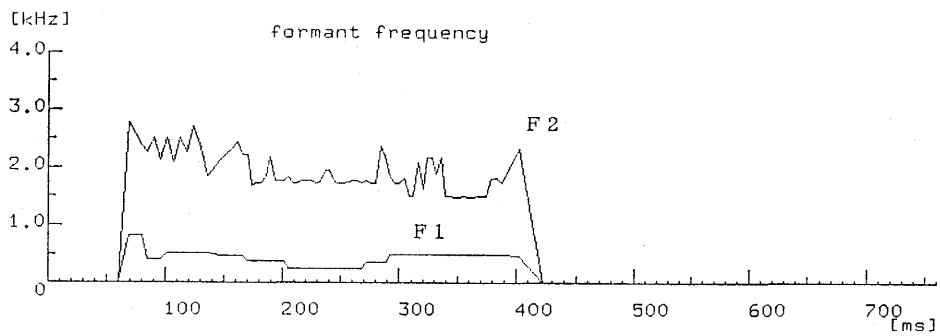

Fig. 6. Plot of F1 (first formant) and F2 (second formant) using 3.0 kHz LPF.

5. CONCLUSIONS

Listening Experiments of speech processing using digital signal processing instead of analog processing were carried out. By analyzing the experimental results, we found defects of speech processing algorithm in WSP and improved all over performance

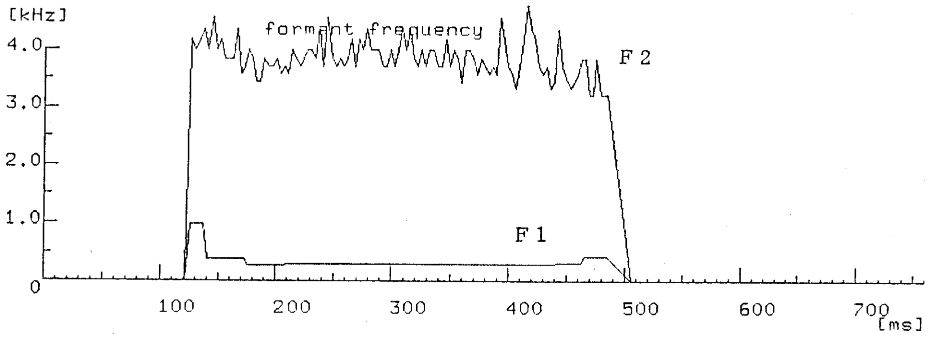

Fig. 7. Plot of F1 and F2 of vowel /i/,as obtained by the first algorithm.

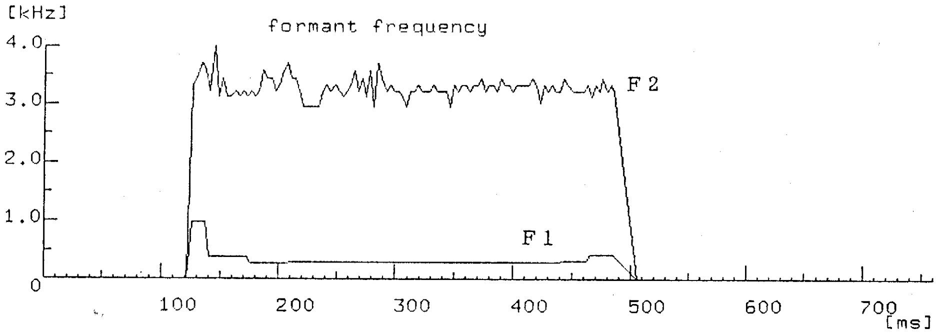

Fig. 8. Result from the second algorithm using 3.0 kHz LPF.

by removing these defects accordingly. We showed the availability of our system for the analysis of speech processing.

6. Acknowledgement

We would like to thank Prof. Iwao Honjo and Michio Kawano of faculty of medicine of Kyoto University and Dr. Tomoji Haji of Kohchi Medical College, who suggested direction of our research. We thank to Dr. Masatake Dantsuji for his helpful suggestions. We would also like to thank Prof. Tatsuo Nishida who have supported our research in many ways. And we would like to thank Y. Fukuda for providing us her video tapes.

References

[1] Shigeyoshi Kitazawa, Mitsuaki Naito, Koushi Yamada and Shuji Doshita: *Application of Digital Signal Processor for the Nucleus 22 Channel Cochlear Implant System,* Studia Phonologica XXIV, Kyoto University, (1990) 61–69.

[2] Mitsuaki Naito, Kazuyuki Muramoto, Koshi Yamada, and Shigeyoshi Kitazawa: *Application of digital processor for Nucleus 22 channel Cochlear Implants,* The 1990 autumn meeting of the Acoustical Society of Japan, 2-7-7, 289–290.

[3] Blamey P.J., Dowell R.C., Brown A.M., Clark G.M. and Seligman P.M.: *Formant estimating speech processor for cochlear implant patients,* Speech Communication, 6 (1987) 293–298.

[4] *Nucleus 22 Channel Cochlear Implant System Audiologist's Handbook,* No. 2008 issue 2, by Cochlear Ltd., November (1986) 80–91.