The Effect of Amplitude Compression on Wide Band Telephone Speech for Hearing-Impaired Elderly People

Mutsumi Saito, Kimio Shiraishi** and Kimitoshi Fukudome***

*Fujitsu Kyushu Digital Technology, Ltd.
**Department of Acoustic Design, Kyushu Institute of Design

saito@qdt.fujitsu.com
{kimio,fukudome}@kyushu-id.ac.jp

Abstract

Recently, high-speed multimedia communication systems have become widespread. Not only conventional narrow band speech signal (up to 3.4 kHz) but also wide band speech signal (up to 7 kHz) can be transmitted through high-speed communication lines. Generally, the quality of wide band speech signal is high and its articulation score is good for normal-hearing people, but for elderly people who have hearing losses in higher frequencies, the effect of wide band speech is doubtful. Therefore, we investigated the effect of wide band phone speech on the elderly people's speech perception in terms of articulation. And we also considered the effect of amplitude compression method that is used for hearing aids.

Japanese 62 CV syllables were used as test speech samples. The original speech samples were re-sampled to narrow band speech (8 kHz sampling) and wide band speech (16 kHz sampling). All speech samples were processed with AMR-CODEC (Adaptive Multi Rate COder-DECoder), which is a voice coding system available to both narrow band and wide band speech signals. Then, coded speech signals were processed with a multi-band amplitude compression method. Ratios of the compression in each frequency bands were determined according to the average value of subjects' hearing levels. All subjects were native Japanese speakers, aged 68 to 72 years, and have hearing losses (more than 40 dB HL).

From the results of the test, we found that combination of wide band speech and amplitude compression showed significant improvement of the articulation.

1. Introduction

Communication systems have advanced a great deal, and high-speed multimedia communication service is now possible [1]. Even portable mobile phones have capability of transmitting high quality video data [2]. On high-speed communication lines, not only conventional narrow band speech signal but also wide band speech signal can be transmitted thanks to the high-speed communication system and speech coding algorithms. The quality of wide band speech is high and its articulation score is good because it contains higher frequency signals that are useful for consonant recognition.

However, for hearing-impaired individuals, such as elderly people, the effectiveness of wide band communication system to speech perception is doubtful, because elderly people have hearing losses in general and it is hard for them to hear high frequency sounds. A study about the relation between comprehensibility of speech over the telephone and aging reported that aging degrades comprehensibility of speech [3]. But even for hearing-impaired people, intelligibility of speech signal that has wide frequency range is better than that of low pass filtered narrow band speech signal [4][5]. And in the study of hearing aids, wide frequency range (up to 10 kHz) is suggested for high-fidelity hearing aid [6].

From these preceding studies, we assumed that wide band communication is also beneficial to hearing impaired elderly people. Therefore, we studied the effect of wide band speech on the elderly people's speech perception in terms of articulation.

Furthermore we introduced amplitude compression method. Advantage of using amplitude compression to telephone speech signal was reported [7], but this study was done for telephone band (300 Hz to 3000 Hz) speech and the effect for wide band speech has not been examined. We supposed that amplitude compression is more effective for wide band speech and we tried to confirm this hypothesis.

2. Signal Processing

We arranged two software simulators of signal processing for this study. One was a speech CODEC simulator, and another was an amplitude compression simulator. Both simulation programs were executed on a workstation.

2.1. Speech CODEC and Frequency range

To compare narrow band speech and wide band speech in terms of actual telephone speech quality, we introduced AMR (Adaptive Multi Rate) CODEC system, which was adopted as a standard speech coding system of IMT-2000 (International Mobile Telecommunication-2000). AMR has two types of coding methods, NB (Narrow-Band)-AMR [8][9] and WB (Wide-Band)-AMR [10][11]. NB-AMR is used for 8 kHz sampling speech signals (frequency range is up to 3.4 kHz), meanwhile WB-AMR is used for 16 kHz sampling speech signals (frequency range is up to 7 kHz).

Fig. 1 shows test speech stimuli conditions. Input speech signals (48 kHz sampling) were down-sampled and processed with NB-AMR or WB-AMR. Then the coded signals were processed with multi-band amplitude compression system. Before the amplitude compression, level adjustment was done and the level of speech samples was set to 60 dB SPL or 80 dB SPL (in this study, the level of sinusoidal wave which has the maximal amplitude of the 16bit PCM was assumed as 110 dB SPL).

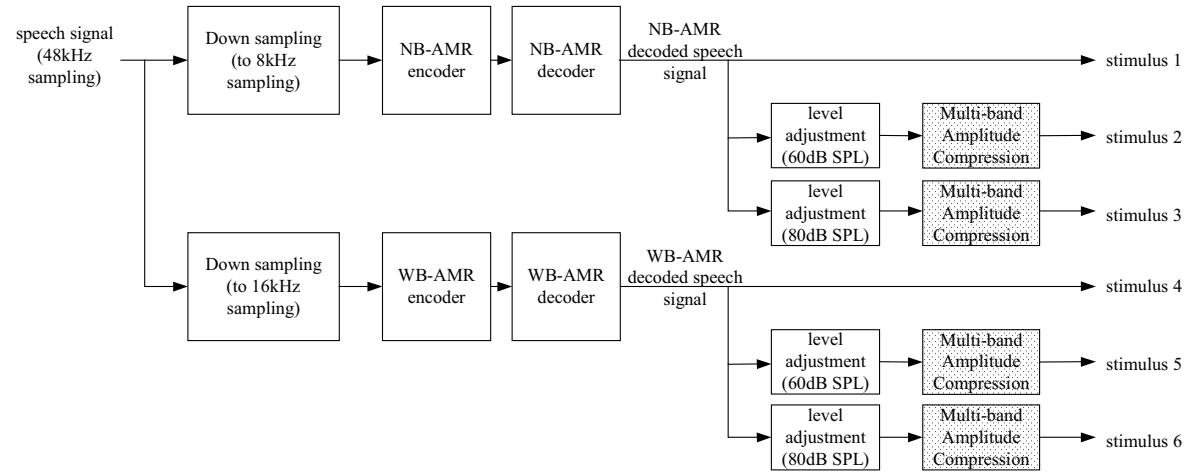

Figure 1: Test speech stimuli conditions.

2.2. Amplitude Compression

In this study, we tried an amplitude compression method to compensate for hearing losses of elderly people. Some compression method has been proposed and developed [12][13][14]. We introduced Wide Dynamic Range Compression (WDRC) method, which was adopted for consumer hearing aids of Philips Hearing Instruments [15].

2.2.1. Multi-band amplitude compression

In the WDRC, speech signals were divided into some frequency bands and their amplitudes are compressed independently (Fig. 2). We modified the number and width of frequency bands so as to connect amplitude compression to speech CODEC. BPFs (includes LPF and HPF) are 512 tap FIR filters.

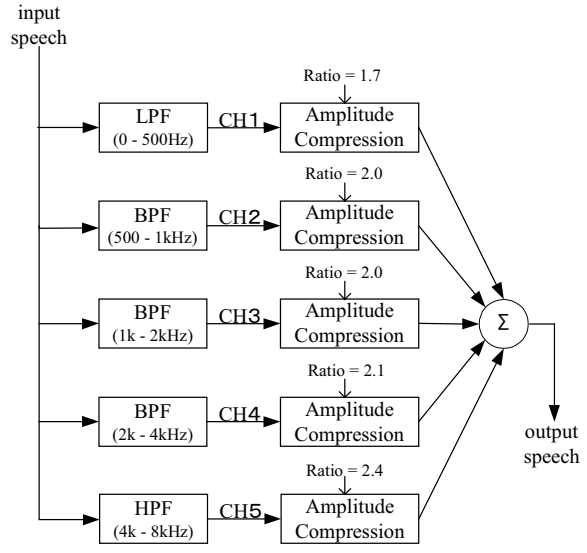

Figure 2: Block diagram of WDRC (multi-band amplitude compression). CH5 exists only in wide band condition.

2.2.2. Compression ratio determination

The compression ratios for each frequency band were determined according to subjects' hearing levels. Table 1 shows average hearing levels of three subjects whose hearing levels were measured beforehand. Dynamic ranges of impaired ear were assumed from average hearing levels according to Fig. 3. Then the compression ratio of each frequency band was calculated (Table 2). Level functions are shown in Fig. 4.

Table 1: Average hearing levels of three subjects (subjects A, B, and C).

Frequency (Hz)	125	250	500	1k	2k	4k	8k
Hearing Level (dB)	39	44	52	55	58	61	73

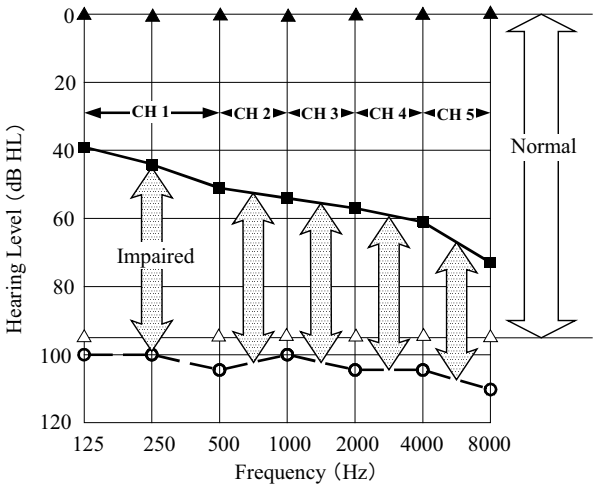

Figure 3: Comparisons of dynamic ranges for five frequency bands between normal ear and impaired ear. Solid line denotes the average hearing levels of subjects.

Table 2: Compression ratio calculation based on average hearing levels.

	CH1	CH2	CH3	CH4	CH5
Hearing Level (dB HL)	44	54	57	60	67
Dynamic range of impaired ear (dB)	56	47	45	45	40
Dynamic range of normal ear (dB)	95	95	95	95	95
Calculated compression ratio	1.7	2.0	2.1	2.1	2.4

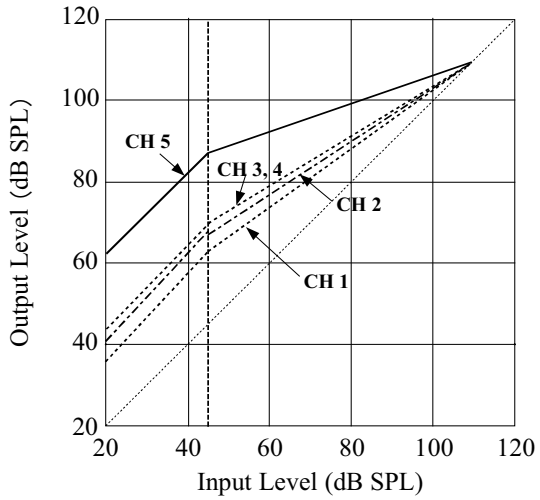

Figure 4: Input-Output level functions.

2.2.3. *Level adjustment*

The level of input speech to multi-band amplitude compression block was set to two types of level (60 dB SPL and 80 dB SPL) in order to test the difference between two types of input level, because the effect of compression is level-dependant. After the compression processing, the levels of all test stimuli were adjusted to equal level.

2.2.4. *Time constant*

Time constant of compression were set to 10ms for attack and 100ms for release.

3. Experiment

3.1. Subjects

Six subjects (age: 68-72) participated in the experiment. They were all native Japanese speakers and have hearing losses. Table 3 shows each subject's hearing level.

Table 3: Subjects' hearing levels (dB HL).

Subject	age	ear	Frequency (Hz)						
			125	250	500	1k	2k	4k	8k
A	69	R	50	45	45	45	55	65	85
B	68	R	35	35	45	50	50	55	65
C	71	L	30	50	60	60	60	50	75
D	72	L	50	55	45	55	55	85	100
E	72	L	35	40	55	55	60	70	95
F	71	L	55	50	45	35	45	55	55

3.2. Stimuli

The test stimuli presented in the experiment were Japanese 62 CV (consonant-vowel) syllables. Vowels were /**a, i, u, e** and **o**/, and consonants were /**k, s, t, n, h, m, y, l, w, g, z, d, p** and **b**/. All stimuli were spoken by a male speaker and recorded (48 kHz sampling, 16-bit quantization).

3.3. Procedure

In the experiment, the six conditions shown in Fig. 1 were presented randomly. The stimuli were delivered directly (without hearing aid) to right or left ear (shown in Table 3) via Sennheiser HDA 200 headphone. Each subject adjusted the presentation level to a comfortable level. The listening tests were carried out in a soundproof room.

4. Results

In Fig. 5, intelligibility scores are plotted. Although average score of wide band without compression was similar to that of narrow band without compression, the average score of wide band with compression showed significant improvement than narrow band with compression. We found that for subject (E) compression was not so effective, and for subject (D) the effect of wide band was not observed, but for four subjects (A, B, C and F) compression was very effective on wide band conditions.

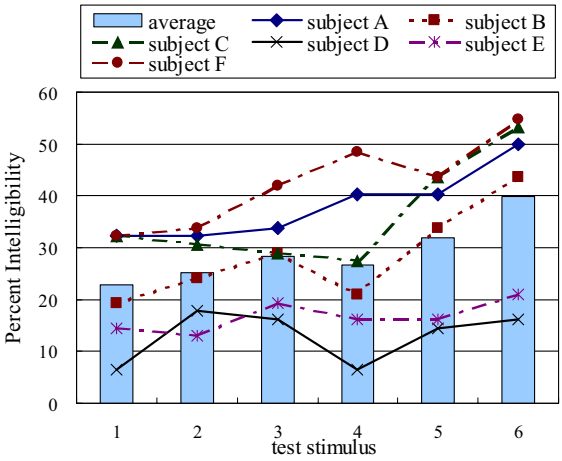

Figure 5: Results of the experiment. Bars show average intelligibility scores, and lines show each subject's intelligibility scores. The test stimuli conditions are: (1) narrow band; (2) narrow band with compression (input level: 60 dB SPL); (3) narrow band with compression (input level: 80 dB SPL); (4) wide band; (5) wide band with compression (input level: 60 dB SPL); (6) wide band with compression (input level: 80 dB SPL)

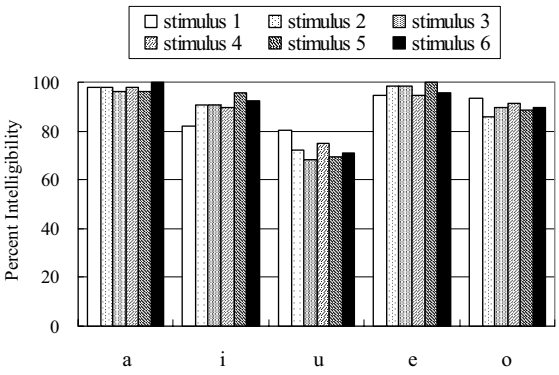

Figure 6: Results of the experiment by vowel (averaged for all subjects). Bars show average intelligibility scores for each stimulus condition.

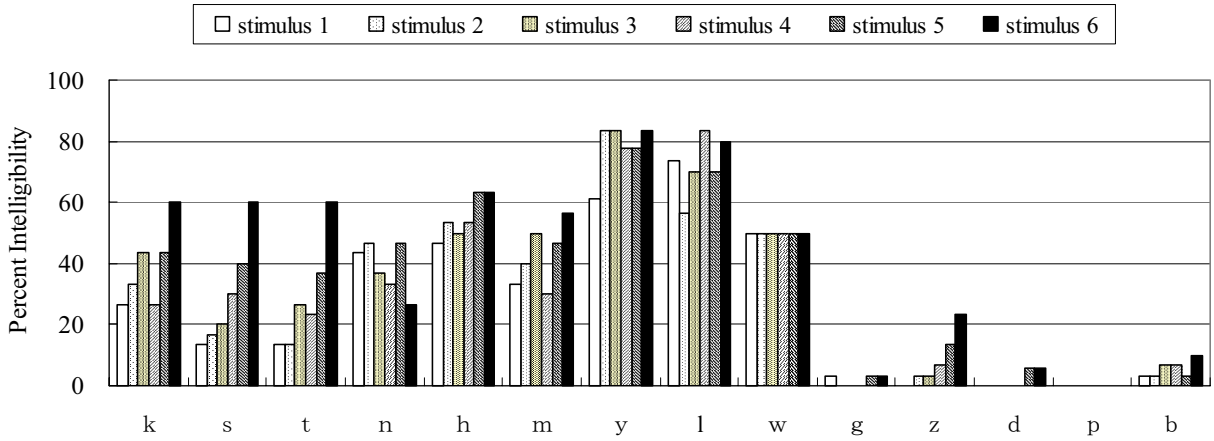

Figure 7: Results of the experiment by consonant (averaged for all subjects). Bars show average intelligibility scores for each stimulus condition.

Fig. 6 shows results by vowel (We ignored consonant error and counted only correctness of vowel). Fig. 7 shows results by consonant (We ignored vowel error and counted only correctness of consonant). As for results by vowel, no differences between conditions were seen, whereas for results by consonant, differences between conditions of stimuli were observed. And for results by consonants, differences between consonants were also observed. For some consonants no improvement of intelligibility was seen (/**n, l, w, g, d, p** and **b**/), meanwhile for consonants /**k, s, t** and **z**/, significant improvement for wide band and amplitude-compressed speech was observed.

We assumed that the differences between consonants come from spectrum differences of consonants. The consonants /**k, s, t** and **z**/ contain certain amount of speech energy in higher frequency region (above 3.4 kHz), whereas consonants /**n, l** or **w**/ have less energy in higher frequency region above 3.4 kHz. It is assumed that with amplitude compression, the energy of higher frequency region was amplified enough, and hearing-impaired subjects recognized the consonants.

Overall, these results support the idea that amplitude compression is more effective for wide band speech.

5. Conclusions

In this study, we investigated the effect of wide band phone speech and amplitude compression on the elderly people's speech perception. Intelligibilities of the narrow band and wide band speech signals were evaluated. And amplitude compression method that is used for hearing aids was introduced. The results showed that combination of wide band speech and amplitude compression showed significant improvement of articulation. And the idea that amplitude compression is more effective for wide band speech was confirmed. However, the effect was subject dependent, requiring us to investigate further with more subjects and conditions.

6. Acknowledgement

This research was supported by Fukuoka Industry, Science & Technology Foundation, Fukuoka, Japan.

References

[1] http://www.imt-2000.org
[2] http://foma.nttdocomo.co.jp/english/index.html
[3] Eskenazi M. and Black A., "A study on speech over the telephone and aging", *EuroSpeech2001*, Vol.1, p171-174, 2001.
[4] Patricia G. Stelmachowicz, et.al., "Effect of stimulus bandwidth on the perception of /s/ in normal- and hearing-impaired children and adults", *J. Acoust. Soc. Amer*., 110(4):2183-2190, 2001.
[5] Thomas Baer, Brian C. J. Moore, and Karolina Kluk, " Effects of low pass filtering on the intelligibility of speech in noise for people with and without dead regions at high frequencies", *J. Acoust. Soc. Amer*., 112(3): 1133-1144, 2002.
[6] Boothroyd A. and Medwetsky L., "Spectral distribution of /s/ and the frequency response of hearing aids", *Ear Hear.*, 13(3):150-157, 1992.
[7] Terry M. et.al., "Processing the telephone speech signal for the hearing impaired", *Ear Hear.*, 13(2):70-79, 1992.
[8] 3GPP TS26.090 (V4.0.0), "*AMR Speech Codec; Speech transcoding*", Mar. 2001.
[9] 3GPP TS26.073 (V4.1.0), "*AMR Speech Codec; ANSI C source code*", Dec. 2001.
[10] 3GPP TS26.190 (V5.1.0), "*AMR Wideband Speech Codec; Speech transcoding*", Dec. 2001.
[11] 3GPP TS26.173 (V5.4.0), "*AMR Wideband Speech Codec; ANSI C source code*", Mar. 2002.
[12] Hedrick MS. and Rice T., "Effect of a single-channel wide dynamic range compression circuit on perception of stop consonant place of articulation", *J. Speech Lang. Hear. Res.*, 43 (5):1174-1184, 2000.
[13] Stone MA., Moore BC., Alcantara JI. and Glasberg BR., "Comparison of different forms of compression using wearable digital hearing aids", *J. Acoust. Soc. Amer*., 106 (6):3603-3619, 1999.
[14] Hidaka H. et.al, "Clinical evaluation of a portable digital hearing aid with narrow-band loudness compensation", *Scand. Audiol.*, 27(4):225-236, 1998.
[15] Philips Hearing Instruments, *Multi-mode Compression Architecture*, 2001 (in Japanese)