ITU-T SG12 회의 기고서

SOURCE : REPUBLIC OF KOREA

TITLE : SUBJECTIVE TESTS ON DIFFERENCES IN LOUDNESS AND SPEECH QUALITY BETWEEN WIDE-BAND AND NARROW-BAND TELEPHONY

Contact:Jeong-Hwan Kim
Acoustic Information Processing Section
ETRI **Tel:+82-42-860-6142**
P.O.Box 106, Yusong **Fax:+82-42-861-5597**
Taejon, 305-350, KOREA **E-mail:jhkim@audio.etri.re.kr**

1. Introduction

The purpose of this contribution is to report the results of three experiments examining differences in loudenss and speech quality between wide-band(150~7,000Hz;WB) and narrow-band(300~3,400Hz;NB) speech signals. For our preferred and compared loudness level tests, a method of adjustment is applied in order to make a direct comparison between WB and NB speech. The intelligibility test using 100 Korean meaningless monosyllables is also performed to examine the speech quality difference between WB and NB speech.

2. Conclusion

The results of our adjustment test indicate that the difference are approximately 2dB and 3.5dB for preferred and compared loudness between WB and NB, respectively. These values are rather smaller than those reported by other laboratories[1][2].

And, the results of intelligibility test show that there is only a negligible difference between WB and NB in total intelligibility score of using 100 monosyllables, while WB speech produces higher

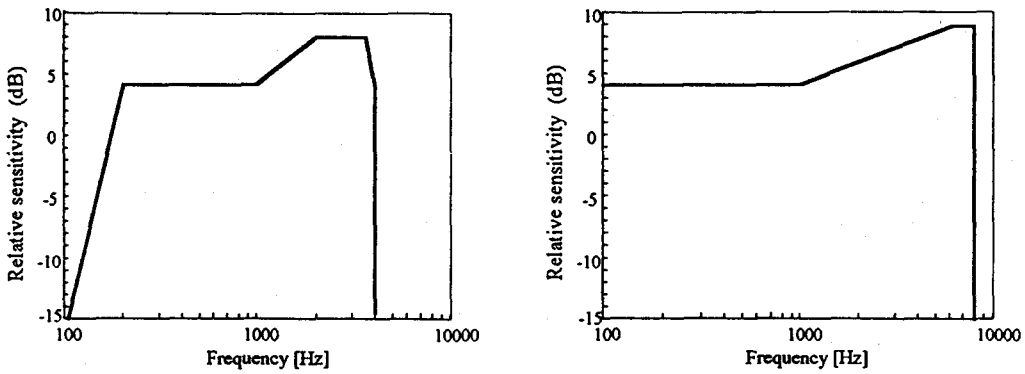
score than NB speech by 17 points for a subset of 20 monosyllables consisting of fricatives, affricates and plosives. These Korean consonants usually have a relatively large hige frequency energy above 3,400Hz.

3. Experiments

3.1 Speech Material

A paragraph of about 1 minute 45 second duration was used as a speech source the preferred loudness level test and a 7 second sentence, in the compared loudness level test. In the intelligibility test, as mentioned above, 100 Korean meaningless monosyllables were utilixzed. The speech was recorded in a sound-proof room while the microphone was kept 30cm apart from the speaker's lips. The speech was sampled at 15KHz with the 12-bit resolution.

To obtain NB speech which satisfies the national sending frequency characteristics for ISDN telephones, the speech source was digitally filtered using a FIR filter. The WB speech which satisfies the sending frequency characteristics of the draft recommendation P.WBHS[3] was also obtained using the FIR filter. Both of the NB and WB frequency charateristics are presented in Figure 1.



(a) NB frequency characteristics (b) WB frequency characteristics

Figure 1. NB and WB frequency characteristics

3.2 Subjects

Fifteen subjects(6 male and 9 female) participated in our experiments. All subjects were university students and could be considered to be inxperienced.

3.3 Test Environment

The block diagram of tests is illustrated in Figure 2. The speech was reproduced over a headphone(Sennheiser HD540). To compensate for the effect of frequency characteristics of headphone, we equalized the relative sensitivity of the filter using the Head and Torso Simulator(B & K type 4128) and frequency analyzer(B & K type 2133).

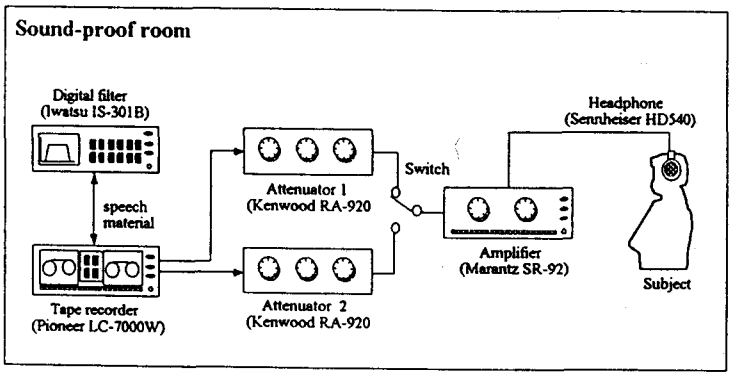


Figure 2. The block diagram of tests

3.4 Test Procedure

The test can be divided into three subsets. The first is for the adjustment of preferred loudness level, the second, for the adjustment of compared loudness level, and finally, the third is for the dictation of monosyllables. These tests were carried out separately.

In the preferred loudness test, the subjust attenuators until their preferred loudness level is achieved. In the comparedloudness test, the subjects were asked to listen to the reference NB speech and then adjust the attenuator assigned for WB speech until the subjects cannot perceive the loudness difference between the reference and compared speech. In case of using WB speech as a reference, the same procedure was applied. the reference NB or WB reference speech level was varied to 60, 65, and 70dB(A) according to random allocation. The adjustment tests were repeated twice with a reasonable break.

In intelligibility test, the subjects were instructed to listen to NB or WB 100 monosyllables and then, to write them down on response sheets while the speech was fixed at 65dB(A).

4. Results

4.1 Preferred Loudness Level Test

The results in Table 1 indicate that the difference in preferred loudness between WB and NB speech is approximately 2dB. This difference is significant at the level of $\alpha=0.01$($t=-3.87$, $p<0.0003$).

Table 1. Preferred loudness level for NB and WB speech

	NB speech	WB speech
mean	70.7dB(A)	68.6dB(A)
stdev	4.85	4.08

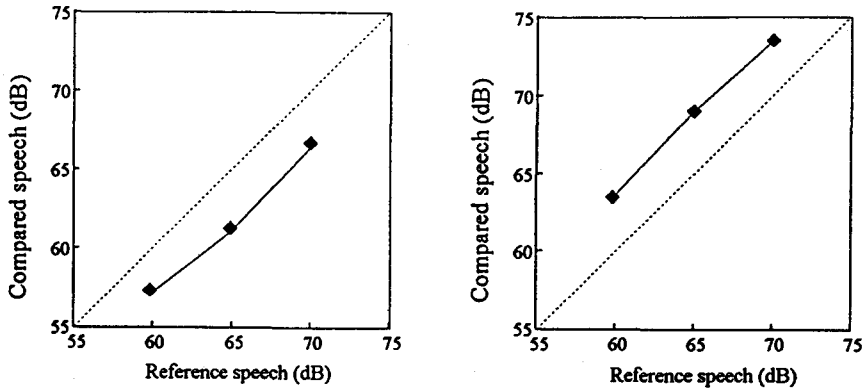
$t=3.87$, $p<0.0003$

4.2 Compared Loudness Level Test

The results in Table 2 indicate that the difference in compared loudness between WB and NB speech is approximately 3.5dB. And in Figure 3, the loudness difference between the reference and the compared speech is presented graphically.

Table 2. Adjusted loudness level for NB and WB reference speech

	NB 60dB(A)	NB 65dB(A)	NB 70dB(A)	WB 60dB(A)	WB 65dB(A)	WB 70dB(A)
mean	57.3	61.3	66.7	63.7	68.6	73.5
stdev	1.98	1.17	1.25	1.77	1.62	1.84
diff. of level	2.7	3.7	3.3	3.7	3.6	3.5
mean diff. of level	3.42					



(a) in case of NB reference (b) in case of WB reference

Figure 3. Loudness difference between the reference and the compared speech

4.3 Intelligibility Test

The results in Table 3 and 4 indicate that only a negligible difference was produced between WB and NB in the intelligibility score when we utilized the whole 100 monosyllables. On the other hand, WB speech scores higher by 17 points than NB speech for the subset of 20 monosyllables including fricatives, affricates and plosives. This difference is believed to be caused by the different bandwidwidths of WB and NB because the above Korean consonants have considerably large energy at frequencies above 3,400Hz. This difference is very significant at the level of $=0.01$)$t=3.209$, $p<0.003$).

Table 3. Total intelligibility score for NB and WB conditions

	NB condition	WB condition
mean	78.1	81.1
stdev	6.44	6.96

$t=1.224$, $p<0.121$

Table 4. Partial intelligibility score for NB and WB conditions

	NB condition	WB condition
mean	67.1	84.1
stdev	15.0	10.8

$t=3.209$, $p<0.003$

References

[1] ITU-T COM XIID. 58, "Wideband Loudness Balances", Geneva, Oct. 24~31, 1990
[2] ITU-T COM 12-11-E, "Difference in Loudness and Speech Quality for 3.1KHz and 7KHz Handset Telephony", Geneva, March 1993.
[3] ITU-T com 12 TD. 27-E, Status of Question, Geneva, May 10~19, 1993.