THE VALIDITY TESTS OF ARTICULATION LISTS USING THE EFFECTS OF MONAURAL AND BINAURAL FACTORS

Jeong-Hwan Kim, Seong-Hoon Kang, Dae-Young Jang

Electronics and Telecommunications Research Institute
Yusong P.O. Box 106
Taejon 305-600
Republic of Korea

ABSTRACT In this study we developed monosyllable lists for articulation test for Korean. We sampled 103,581 colloquial monosyllables, applied them to five selection rules that based on Korean linguistic characteristics, and finally constructed five different lists with fifty monosyllables. The validity test using the monaural impairment factors such as S/N ratio and cut-off frequency showed that articulation scores were changed systematically according to the level of impairment factors.

In addition, we investigated the effect of azimuth of a single competing sound source upon articulation scores. The syllables were always reproduced by the loudspeaker in front of the subject, while Hoth noise were reproduced by the loudspeaker with varying azimuth around subject. The result indicated that the articulation depended on the azimuth of competing sound sources.

Finally, no significant differences among lists were found in all experimental conditions.

1. INTRODUCTION

Early studies on speech intelligibility began with the need for designing telephone and radio communication systems. The results of manipulating some of the temporal-monaural factors associated with transmission systems showed that in a quiet environment, speech level constitutes a factor of speech intelligibility and that S/N ratio also constitutes a factor of speech intelligibility if the speech is heard against background noise. And the studies on the effect of reducing the band width of a transmission line showed that the frequency range of a transmission line has an influence on speech intelligibility.

Speech intelligibility may also be influenced by spatial-binaural factors for sound fields. The masking of speech by speech is very common in everyday situations, varying from the case of a single competing talker to the case of a babble of many talkers in a cocktail party. It is well recognized that in binaural hearing the spatial divergence of the direction of the voices can increase intelligibility substantially. For the multi-talker situation, this effect has sometimes been referred to as "the cocktail party effect", after Cherry[1].

After Santon's study[2] which was the first one to investigate the relationship between speech

intelligibility in a room and the spatial-binaural factors, the factors connected with our binaural hearing have been generally recognized to have an influence on speech intelligibility in rooms according to some studies.

Recently, teleconference systems using loudspeakers or headphones are being designed and introduced to the telecommunication network. These systems are expected to give a more natural and spatial impression to listener than traditional monaural handset telephones by utilizing binaural effects. However, methods for assessing the speech quality in binaural telecommunications such as teleconferences have not yet been established, and thus become to be an urgent problem to solve.

In Korea there are not any standardized speech units for articulation tests and test procedures to measure speech intelligibility of acoustic environments and telecommunication systems. The purpose of this study is to test the validity of monosyllable lists which we developed using the effects of temporal-monaural and spatial-binaural factors upon speech intelligibility. These studies are concerned with the effects of masking due to noise, the effects of band-pass filtering, and the effects of azimuth angle of competing noise. And unilateral deafness was simulated by occluding one ear in order to investigate the advantage of binaural over monaural hearing.

2. THE DEVELOPMENT OF KOREAN MONOSYLLABLE LISTS

As described previously, it must be need standardized speech units for articulation tests and test procedures to assess speech transmission quality of acoustic spaces and telecommunication systems. Some attempts[3][4] failed of establishment standardized monosyllable PB-lists for articulation tests because of the problem of validity between lists. We, therefore, concluded that it is impossible to make completely homogeneous Korean monosyllable PB-lists, and developed new monosyllable lists, not PB-lists, which was applied following selection rules refer to Korean linguistic characteristics;

(1) sampling from among colloquial monosyllables.
(2) application of Korean pronunciation rules(the Ministry of Education rule 88-1).
(3) including only meaningless monosyllables.
(4) excluding ambiguous pronunciations attributed to dialects.
(5) excluding monosyllables which have very low discrimination by pretest.

We sampled colloquial 103,581 monosyllables from TV drama, news, report, and talk show and finally derived 250 monosyllables by means of above five selection rules. And we constructed 5 lists, each consisting of 50 monosyllables, using random allocation procedure.

Speech source of monosyllable lists was recorded with digital audio tape by professional announcer. The distance between talker and microphone was 30cm and speech level was fixed 60dB(A). Moreover, the recording interval between monosyllables was 3 second.

3. EXPERIMENT I

3.1 Experimental method and conditions

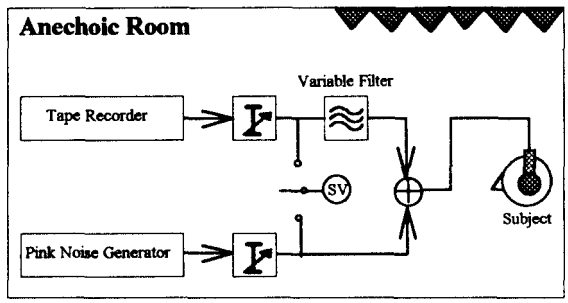

Figure 1. Block diagram of the apparatus.

Articulation tests were administered to ten female subjects(university students) with normal hearing acuity seated in an anechoic room. Speech sources were presented to the subjects through the headphones at 65dB(A) to be considered typical for conversational levels in everyday situations. For present purpose, we will report only the percentage of syllables correctly identified by the listeners.

Fig. 1 represents block diagram of the apparatus. S/N ratio condition was designed to test the speech intelligibility to evaluated the masking effects upon speech of pink noise and had 3 sub-conditions, monaural, binaural, and interaural conditions. In monaural condition monosyllables and noise were presented to subjects at the only left ear with S/N ratio of 15, 10, 5, 0, and -5dB. In binaural condition monosyllables and noise were presented to subjects at the left and the right ear respectively with S/N ratio of 0, -5, -10, -15, and -20dB. And in interaural condition monosyllables were presented at the left ear and noise at two ears, in which total S/N ratio, 10, 5, 0, -5 and -10dB, was maintained.

Band-pass filtering condition was also designed to test the speech intelligibility to evaluated the cut-off frequency effects upon speech and had 2 sub-conditions, high pass filtering and low pass filtering conditions. High pass filtering condition was consisted of cut-off frequencies of 300, 500, 1k, 2k, and 3kHz and low pass filtering condition was consisted of cut-off frequencies of 500, 1k, 2k, 3k and 4kHz. 5 subjects were anticipated in the S/N ratio condition and band-pass filtering condition respectively according to Latin-Square design .

3.2 Results

The articulation scores obtained at several S/N ratio are plotted in Fig. 2. In binaural condition the results indicated that the noise could not have the masking effects upon speech. In monaural condition the noise had the significant masking effects upon speech($F=67.54$; $p<0.01$). Above results are in agreement with previous research[5]. In interaural condition the noise also had the significant masking effects upon speech($F=12.52$; $p<0.01$), but significant differencies were found at S/N ratio 10dB vs. 5dB and -5dB vs. -10dB through Scheffe tests. In addition, no significant differencies among lists were found in all conditions.

The articulation scores obtained in band-pass filtering condition are plotted in Fig. 3. In low-pass($F=100.02$; $p<0.01$) and high-pass filtering condition($F=214.65$; $p<0.01$) the results indicated that there were very significant cut-off frequency effects upon speech. And we also could not find any significant differencies among lists in band-pass filtering conditions.

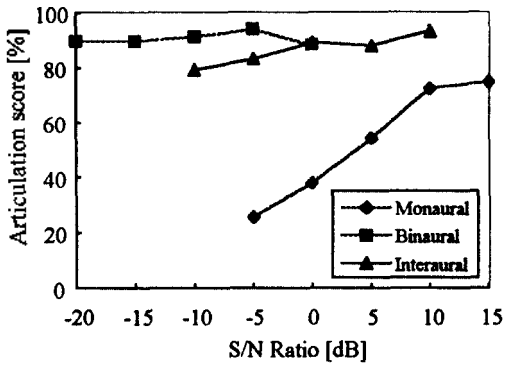

Figure 2. Measured speech intelligibility scores versus S/N ratio.

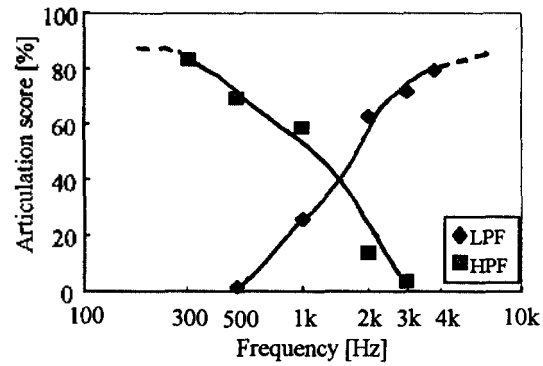

Figure 3. Measured speech intelligibility scores versus band pass filter.

4. EXPERIMENT II

4.1 Experimental method and conditions

Fig. 4 represents block diagram of the apparatus. One channel for speech source is connected to loudspeaker S0, whereas the other channel for noise can be connected via switch to each of the five loudspeakers S1,...,S5 distributed in a horizontal plane at equal distances (1.5m) from the center of the subject's head. The loudspeakers S0 and S1 are mounted one directly on top of the other to have coinciding directions.

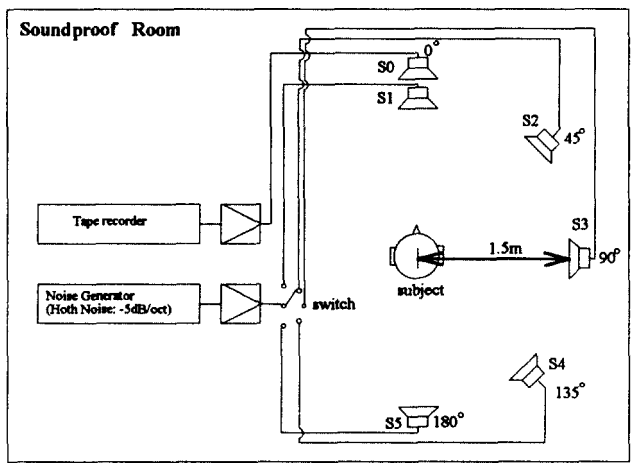

Figure 4. Block diagram of the apparatus.

The azimuths of the loudspeakers S1,...,S5 were 0 °, 45 °, 90 °, 135 °, and 180 ° relative to S0, respectively. The subject, seated on a chair in a soundproof room, was instructed to look always straight ahead, facing the loudspeaker, S0, and to write down reproduced target monosyllables on the response sheets.

There were 3 listening modes(binaural, monaural occluding right ear, and monaural occluding left ear), 5 azimuth angles(0 °, 45 °, 90 °, 135 °, 180 °), 3 Hoth noise levels(55, 65, 75dB) and 5 monosyllable lists conditions in experiment. Since there were too many conditions, we constructed Split-Plot Factorial design(SPF-35.35) to reduce subjects' load. As experiment were carried out with 15 subjects with normal hearing acuity, five subjects are allocated in each listening mode conditions, azimuth angle conditions nested, respectively. And each conditions consisted of 15 trials(5 lists by 3 noise level). With about 3 min for each trial, a session in which the subject anticipated lasted about 45 min. Speech level had in all sessions a constant value of about 65dB, to be considered typical for conversational levels in everyday situations.

Unilateral deafness was simulated by occluding one or other ear by an ear plug and a muff which have a good acoustic seal of 25dB in full band. This parameter was obtained by using Head and Torso Simulator(B & K Type 4128) and Signal Analyzer Unit(B & K Type 2035).

4.2 Results

The articulation scores obtained at each azimuth angle in binaural, monaural occluding right ear, and monaural occluding left ear condition are plotted in Fig. 5, 6, and 7 respectively.

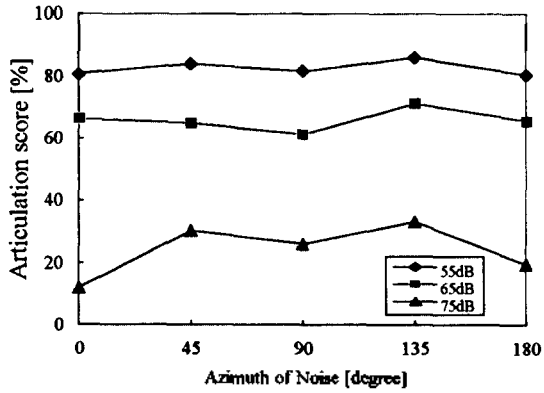

Figure 5. Measured speech intelligibility scores in binaural condition.

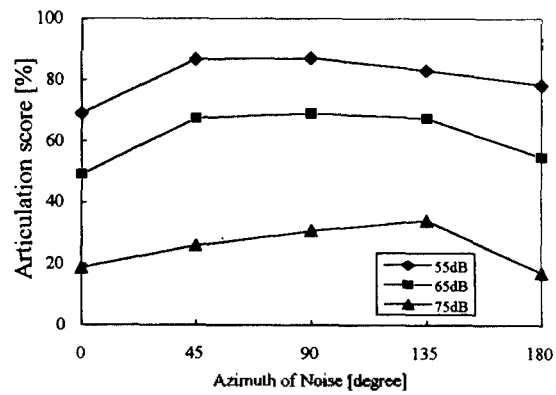

Figure 6. Measured speech intelligibility scores in monaural condition occluding right ear

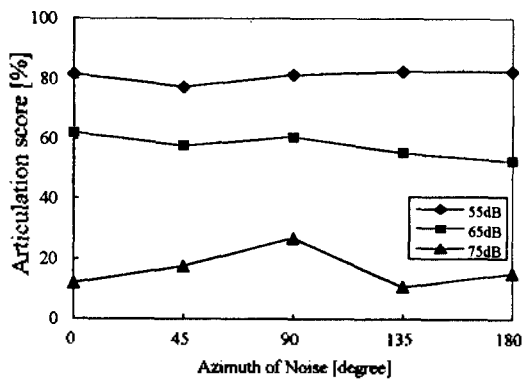

Figure 7. Measured speech intelligibility scores in monaural condition occluding left ear

In all of the conditions the results showed that there were very significant differences of noise levels. For binaural condition, the effect of azimuth of the Hoth noise on the articulation score was not significant in each noise level. However, the results of Scheffe test of mean score irrespective of noise level indicated that there was significant difference at 135 ° vs. other angles($F=3.44$; $p<0.05$). For right ear occluded monaural condition, the effect of azimuth of the Hoth noise on the articulation score was very significant($F=17.17$; $p<0.01$). The results of Scheffe test indicated that there was significant difference at 45 °, 90 °, 135 ° vs. 0 ° and 135 ° vs. 180 °. For left ear occluded monaural condition, the effect of azimuth of the Hoth noise on the articulation score was not significant. In addition, the slight advantage of binaural over monaural condition occluding right ear appeared at 0 ° and 180 °, and the great advantage of binaural over monaural condition occluding left ear appeared at most angles. These results are in agreement with Plomp's[6] which presented by masked threshold.

No significant differencies among lists were found in all conditions such as experiment I.

5. DISCUSSION AND CONCLUSION

As described in introduction, the main purpose of the present paper is to test the validity of monosyllable lists we have developed using the effects of spatial-binaural and temporal-monaural factors upon speech intelligibility. In the results of experiment I and II we could not find any significant differencies among 5 monosyllable lists. In the result 5 monosyllable lists were quite strongly related to each other and very nearly as difficult as each other, that is, the homogeneity of monosyllable lists was obviously confirmed. We, therefore, concluded that we can use these lists to assess transmission quality of acoustic spaces and telecommunication systems, for example, a teleconference room, as one of standardized speech units in Korea.

The results of experiment II demonstrated the ability of humans to listen to the object sound in a noise environment. In the experiment, intelligibility of the object speech is improved when the noise location is separated from the speech, although the total signal to noise ratio is kept constant. This increase in intelligibility is not estimated with the conventional monaural articulation index, since it is to be contributed to binaural hearing ability. This phenomenon can be used in a teleconference system using headphone or loudspeakers so that the system provides the listener with talker identification information and spatial impression by artificial localization. If properly used, articulation test would provide a quantitative and common frame of reference for evaluating the relative abilities of most communication systems to transmit intelligible speech.

ACKNOWLEDGEMENTS

We are pleased to acknowledge the considerable guidance and encouragement of other research staffs of Acoustical Information Processing Section in ETRI. And we also thank the volunteers who participated in the experiments.

REFERENCES

1. E. C. Cherry and B. McA. Sayers, "Human 'cross-correlator'-A technique for measuring certain parameters of speech perception," J. Acoust. Soc. Am. 28(5), 889-901 (1956).
2. F. Santon, "Numerical prediction of echograms and of the intelligibility of speech in rooms," J. coust. Soc. Am. 59(6), 1399-1405 (1976).
3. M. Y. Lee, "The study on articulation test(I): The development of the standardized monosyllable lists," Korea Univ. (1990), in Korean.
4. M. Y. Lee, "The study on articulation test(II): The validity test of monosyllable lists," Korea Univ. (1991), in Korean.
5. S. Hayasaka, "Acoustic engineering," Kansen, 11 (1986), in Japanese.
6. R. Plomp, "Binaural and monaural speech intelligibility of connected discourse in reverberation as a fuction of azimuth of a single competing sound source (speech or noise)," Acustica 34, 201-211 (1976).