PROCEEDINGS of the
23rd International Congress on Acoustics

9 to 13 September 2019 in Aachen, Germany

A pilot study of the relationship between mandarin Chinese word and sentence recognition for the elderly

Jianxin PENG[1]; Jiazhong ZENG[2]; Jiaming LIU[3];

South China University of Technology, China

ABSTRACT

In present study, the simulated binaural room impulse responses with different reverberation times (RTs), were convolved with the test word and sentence signals, and noise signal, and then reproduced by headphone. The subjective word and sentence recognition scores for elderly under different signal-to-noise ratio and RT were obtained. The results show that the shorter the RT, the higher the scores of word recognition and sentence recognition for the elderly. The higher the signal-to-noise ratio, the higher the scores of word recognition and sentence recognition. There is high correlation between the scores of word recognition and sentence recognition for the elderly.

Keywords: Reverberation time, Signal-to-noise ratio, Word recognition, Sentence recognition, Ederly

1. INTRODUCTION

The auditory function of the elderly is gradually degenerative with the increase of age, and their hearing sensitivity and speech comprehension ability in noisy environment are gradually reduced. Compared with young adults with normal hearing, most of the elderly are difficult to hear speech signals in unfavorable listening environment (1-3). Nabelek and Robinson (4) pointed out that speech intelligibility scores obtained at different reverberation times (RTs) required 10-20 dB sound pressure levels higher for old adults than in young adults through Modified Rhyme Test. Helfer and Huntley (5) studied the recognition of English consonants for the elderly in reverberation and noise. The subjects were divided into the elderly with impaired hearing, the elderly with slight hearing loss and the young normal hearing. The results show that the elderly had a considerable barrier to speech understanding in the non-ideal auditory environment. A study by Sato (6) showed that speech intelligibility scores of the older listeners were 25% lower than those of young adults for a series of speech intelligibility test involving 20 younger and 20 older subjects. He pointed out that the effect of this difference was equal to the 5 dB increase of ambient noise. These studies demonstrated that the auditory function of elderly are gradually impaired when the elderly get older and the elderly need a higher speech intelligibility.

STI was first proposed by Houtgast and Steeneken (7) and combined the effects of reverberation and noise on speech intelligibility in rooms into an objective parameter, which is used by many researchers and is good at predicting the speech intelligibility in room. Peng and Zhang (8, 9) conducted a subjective listening experiment on the correlation between Chinese speech intelligibility and STI in room. The results show that STI can predict and evaluate the speech intelligibility in room well.

In this work, subjective evaluation of Chinese the word and sentence recognition under different reverberation and signal-to-noise ratio (SNR) conditions were conducted. The the word and sentence recognition in different acoustical environments were analyzed and compared. The relationship between the word and sentence recognition scores was also analyzed. The aims are to explore the speech communication effect of elderly in different reverberation and SNR conditions.

[1] phjxpeng@163.com
[2] zczjz@scut.edu.cn
[3] 942807917@qq.com

2. METHODS

2.1 Rooms and sound fields

Six rooms were modelled using ODEON. In each model, a source and a receiver (listening position) were set. The listening position in each room was located outside of the reverberation radius of the room and the minimum distance was 5.2 m and the maximum distance was 12.6 m between the source and listening position among these six listening positions. At each listening position, the RT and a binaural room impulse response were obtained by ODEON acoustical simulation through adjusted the absorption material in the room. The RT in 500-1000Hz octave frequency band at these six listening positions were given in Table 1. The room No .1 is the case in the anechoic chamber with no reverberation.

Table 1 –Objective acoustical parameters

Room No.	RT /s	STI
1	No reverberation	
2	0.4	0.61
3	0.6	0.57
4	1.0	0.52
5	1.5	0.46
6	2.2	0.44
7	5.2	0.37

2.2 Word and sentence recognition test

The word recognition test conducted by using Mandarin Chinese test word lists as specified by GB 4959 (10). The test uses 25 five-word rows of similar-sounding Chinese monosyllabic words and is similar to modified rhyme test of English. The five words in each row are randomly arranged and differ only in the initial consonant sound (for example, hao, sao, gao, zao, kao). Fifteen word lists were derived from five basic word lists. One list is used for students to practice, the rest fourteen word lists are used for the test. The test word in carrier phrase is "The × row reads – ". The "×" stands for row number and "—" stands for a test word. On the other hand, Chinese Mandarin Hearing in Noise Test (MHINT) (11) is used for the sentence recognition test. Fifteen sentence lists are used in the test. One list is used for students to practice, the rest fourteen lists are used for the test. Every list contains 10 sentences. And each sentence is composed of 10 target words, for example, "这个球队终于打入决赛 (This team enters the finals at last)" . All word and sentence lists are recorded at the rate of 4.0 words per second in an anechoic chamber by using an omni-directional precision microphone at a distance of 0.5m from the male speaker. The recording is edited by CoolEdit Pro. A ten-second (/thirty second) interval of silence between two adjacent carrier sentences (/test sentence) allowing listeners to mark the test word (/ write the test sentence) which he (/she) heard on the test paper has been inserted.

Based on the average speech spectrum from test word and sentence signals, the corresponding speech-shaped noises are selected for the tests. The SNR was made equal for all selected frequency bands using speech-shaped noise with a frequency spectrum equivalent to the long-term speech spectrum. Since a speech-shaped noise with a frequency spectrum equivalent to the long-term speech spectrum was used, the SNR was approximately equal for all selected frequency bands (12). All the noise, Chinese word and sentence test signal recorded in the anechoic chamber were convolved with the binaural room impulse response respectively, then mixed according to a certain SNR (-5dBA and 5dBA), and then reproduced through the Sennheiser HD580 headphone for the subjective Word and sentence recognition evaluation..

2.3 Participants

Six elderly aged more than 60 years old who could hear and speak Mandarin were randomly selected to participate in the test. The old participants received a few minutes of instruction before the test. There was an interval of more than seven days after the participants performed a test in order to avoid their remembering the word lists. Presented by playback system in closed-response format, the participants' task is to identify the correct answer from five choices. The average of the subjective Chinese Mandarin word and sentence recognition scores across all six participants are obtained for each testing condition. Each participant participated in Chinese Mandarin word and sentence recognition test of a total of 14 conditions which included 7 different RTs (including a test condition with no reverberation and six reverberation conditions) and 2 SNRs conditions. During the test, the participant adjusted the sound pressure level to his/her favorite level. Because the test recording error, the Chinese word test score was deleted in room 4 under 5 dBA SNR.

3. RESULTS AND DISCUSSION

The subjective Chinese word and sentence recognition in all rooms is conducted with the different SNRs. Figures 1 show the Chinese word and sentence recognition scores and their standard deviations in different rooms and different SNRs, respectively. It can be seen from the Figure 1 that, both Chinese word and sentence recognition scores for elderly in these rooms increased with the increase of SNRs and decreased with the increase of RTs.

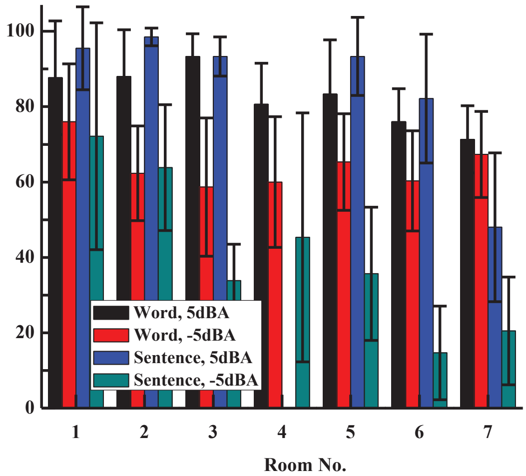

Figure 1 –Chinese word and sentence recognition scores and their standard deviations

The Chinese word and sentence recognition scores of elderly are plotted against the corresponding STI from different test conditions in Figures 2. The correlation coefficients R and standard deviations (SDs) are 0.85, 6.5%; 0.94, 10.8%. To compare the difference in scores from different age group, the relationship between STI and Chinese word and sentence recognition scores of adults is also plotted in Figures 2. The data came from Peng (13) and Peng (14), respectively. Figures 2 show that both Chinese word and sentence recognition scores increase as the STI and the age increases.

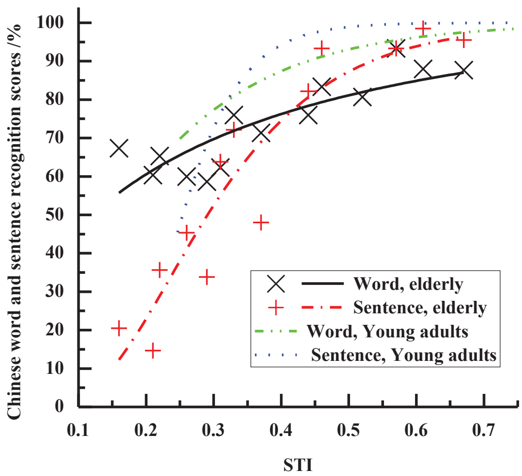

Figure 2 -Relationship between Chinese word and sentence recognition scores and STI

Both RT and SNR impact critical communication for the elderly. In this study, speech-shape noise which has the similar spectrum with speech test signals is mixed. The speech-shape noise can mask the Chinese word and sentence by energetic masking and decrease the recognition score of word and sentence. Moreover, reverberation sounds also interfered with the word recognition by smeared the time gap between syllables. The longer the RT is, the lower the word and sentence recognition scores are. And more, the reverberation also interferes with the recognition of subsequent words and reduces the word and sentence recognition scores (15). When the reverberation and noise combines, especially the long RT, the reverberation not only strongly smear the word, but also enhance the noise level that increases the masking effect of noise on the word. All these factors result in the word and sentence recognition scores for the elderly in these rooms decreased with the increase of RTs and increased with the increase of SNRs.

Sentence recognition requires the listener to extract both word and semantic information from the test speech signal, as well as acoustical information. There are two basic operations involved in the recognition of words with in sentences (16). One is the initial processing of acoustic-phonetic information, the other is the utilization of linguistic-situational information of speech (16, 17). The word recognition in a sentence is constrained by phonetic input and context. Listeners can utilize different sources of knowledge about the structure of spoken language, including phonological, prosodic, lexical, semantic, syntactic, and situational, to constrain word choice. When STI is more than 0.42, the elderly can associate many Chinese words in the whole sentence according to several key Chinese words, and the score of sentence recognition is higher than that of word recognition. When STI is less than 0.42, the Chinese test words that can hear clearly are limited and the elderly cannot associate the whole sentence, so the score of sentence recognition is lower than that of word recognition. On the other hand, the elderly's hearing and speech perception reduced gradually (18, 19), Figure 2 demonstrates that the word recognition scores and sentence recognition scores are lower than those of young adults. Figure 3 shows the relationship between the word recognition score and sentence recognition score of the elderly. There is high correlation between the word recognition score and sentence recognition score of the elderly and the correlation coefficient R is 0.86. When the word recognition score is lower than 80%, the sentence recognition score is lower than the word recognition score. When the word recognition score is higher than 80%, the sentence recognition score is higher than the word recognition score.

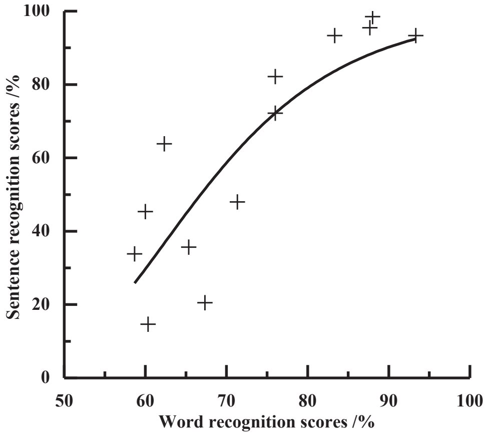

Figure 3 -Relationship between the word recognition scores and sentence recognition scores of the elderly

4. CONCLUSIONS

In the present work, six rooms were modelled by ODEON and the simulated binaural room impulse responses with different RT were obtained. The binaural room impulse responses were convolved with the test word and sentence signals, and noise signal and reproduced by headphone. The subjective word and sentence recognition scores for elderly under different SNR and RT were obtained. The results show that the shorter the RT, the higher the scores of word recognition and sentence recognition for the elderly. The higher the SNR, the higher the scores of word recognition and sentence recognition. There is high correlation between the scores of word recognition and sentence recognition for the elderly.

ACKNOWLEDGEMENTS

The authors thank the elderly who participated in this work. This work is supported by National Natural Science Foundation of China (Grant No. 11674104).

REFERENCES

1. Bergman M. Hearing and aging, Audiology, 1971; 10(3):164-171.
2. Garstecki DC, Mulac A. Effects of test material and competing message on speech discrimination, Journal of Auditory Research, 1974; 14(3):171-178.
3. Dubno JR, Dirks DD, Morgan DE. Effects of age and mild hearing loss on speech recognition in noise, J Acoust Soc Am. 1984; 76(1): 87-96.
4. Nabelek AK, Robinson PK. Monaural and binaural speech perception in reverberation for listeners of various ages, J. Acoust. Soc. Am. 1982; 71(5): 1242-1248.
5. Helfer KS, Huntley RA. Aging and consonant errors in reverberation and noise, J.Acoust.Soc.Am. 1991; 90(4, Pt 1) : 1786-1796.
6. Sato H. Effect of aging of hearing on speech recognition in rooms. In: Proceedings of the 5th International Conference on Gerontechnology, 2005, Nagoya, CD-rom.
7. Steeneken HJM, Houtgast T, A physical method for measuring speech-transmission quality. J. Acoust. Soc. Am., 1980; 67(1):318-326.
8. Peng J., Bei C., Sun H., Relationship between chinese speech intelligibility and speech transmission index in rooms based on auralization, Speech Communication, 2011; 53(7): 986-990.
9. Zhang H.H., Xie H., Applicability of weighting methods in calculating speech transmission index for assessing speech intelligibility of chinese in rooms, Journal of Zhejiang University. Engineering Science, 2012; 46 (3) : 463-469.
10. GB 4959-85: Methods of measurement for the characteristics of sound reinforcement in auditoria, Standard of P. R. China, 1985.
11. Wong L. N., Soli S. D., Liu S., Han N., Huang M.W. , Development of the Mandarin Hearing in Noise Test (MHINT), Ear & Hearing, 2007;28(s2), 70-74.
12. Steeneken HJM, Houtgast T. Mutual dependence of the octave-band weights in predicting speech intelligibility, Speech Commun. 1999; 28(2):109-123.
13. Peng J.X., Relationship between Chinese speech intelligibility and speech transmission index in rooms using dichotic listening, Chinese Science Bulletin, 2008; 53(14): 2748-2752
14. Peng J.X., Lü Y.X., Liu X.J., Lu X.D., Comparison of speech intelligibility between standard Chinese and English by native Chinese listener in rooms, Journal of South China University of Technology (Natural Science), 2011; 39(3): 130-134.
15. Tillery K.H., Brown C. A., Bacon S. P., Comparing the effects of reverberation and of noise on speech recognition in simulated electric-acoustic listening, J. Acoust. Soc. Am., 2012; 131(1) : 416-423.
16. Stuart A, Reception Thresholds for Sentences in Quiet, Continuous Noise, and Interrupted Noise in School-Age Children, J Am Acad Audiol, 2008; 19(1):135–146.
17. Kalikow D. N., Stevens K. N. and Elliott L. L., Development of a test of speech intelligibility in noise using sentence materials with controlled word predictability, J. Acoust Soc Am., 197; 61(5): 1337-51.
18. Füllgrabe C, Moore BCJ, Stone MA. Age-group differences in speech identification despite matched audiometrically normal hearing: Contributions from auditory temporal processing and cognition, Front. Aging Neurosci. 2015; 6(1):1-25.
19. Helfer KS. Older adults in complex listening environments. Hearing Care for Adults—The Challenge of Aging. In: Proceedings of the Second International Adult Conference, 2009, pp. 237-245