The Speech Recognition Ability for Different Age Groups on the Chinese Language System

Linghua Ran[1(✉)], Ling Luo[2], Xin Zhang[1], Taijie Liu[1], and Chaoyi Zhao[1]

[1] Ergonomics Laboratory, China National Institute of Standardization, Beijing 100191, China
{ranlh, Zhangx, liutj, zhaochy}@cnis.gov.cn
[2] China Standard Certification Co. Ltd., Beijing 100088, China
luoling@csc.org.cn

1 Introduction

Public address system can provide useful information for the audience, which are especially important at the railway station or the subway station. Effective public address can offer help to evacuate people at these places.

At present, most speech recognition researches are to predict and assess effects of hearing devices, to do relevant medical legal identification and evaluate the injury degree with low hearing function. In China, large public places basically broadcast through functional address system. It is very important for the public barrier free design to try to take the old people into account.

2 Experimental Method

2.1 Experiment Design

The experiment is conducted in the anechoic room of ergonomics laboratory. The audio files are the length of each sentence is about 4 s, the fade in and out time of each sentence is about 0.15 s and the interval of word front and back is 0.2 s and 0.3 s respectively.

44 subjects were divided into the old group and the young group. The young group includes 21 subjects with the average age of 25.4, and the old group includes 22 subjects with the average of 62.3.

2.2 Experimental Method

The speech recognition experiment adopts the method of subjective assessment of the public sound system [1]. The experiment is divided into two tests. The first one is the hearing level, and after that is the speech recognition test.

© Springer International Publishing Switzerland 2015
C. Stephanidis (Ed.): HCII 2015 Posters, Part I, CCIS 528, pp. 317–320, 2015.
DOI: 10.1007/978-3-319-21380-4_54

Table 1. Definition of difficulty degree

Degree of difficulty	Degree of difficulty
1	Not difficulty
2	Slightly difficult
3	Moderate difficult
4	Very difficult

The hearing level test is to measure the hearing threshold of subjects to guarantee that their hearing belongs to the normal range of their age (Table 1).

3 Experiment Results and Discussion

3.1 Hearing Level

The hearing level test showed that the hearing attenuation of Chinese young group and the old group have no significant frequency differences with the data according to the international standards of ISO 7029 [2], except in the frequency of 1000 Hz and 4000 Hz for the old group.

3.2 Speech Recognition Threshold

Under the same 50 dB environmental noise, both for the young group and for the old group, a single factor analysis was conducted and there was a significant difference between the two kinds of voices for the two age groups. It is showed in Figs. 1 and 2 that under the same noisy environment, the threshold value of male voice is lower than that of female voice in the both young and old groups. It can be seen that threshold value for the old group for the male voice is lower than to the female voice.

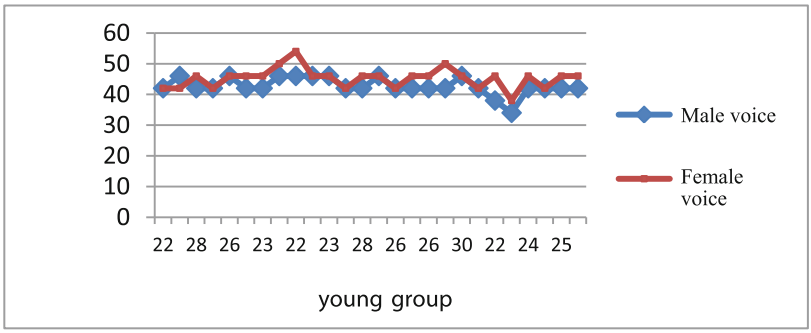

Fig. 1. The speech recognition threshold of the young group under 50 dB environmental noise

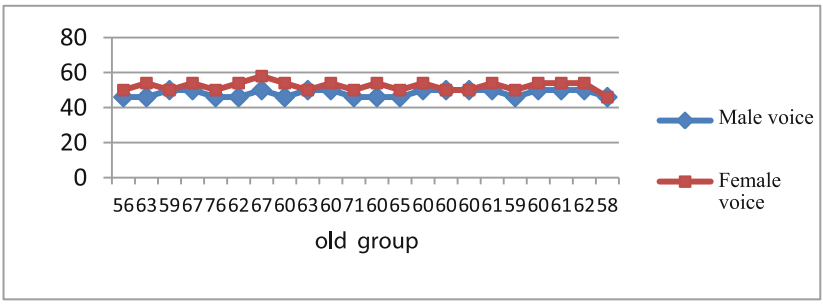

Fig. 2. The speech recognition threshold of the old group under 50 dB environmental noise

Under the same 50 dB environmental noise, the average speech recognition threshold value for the young group to the male voice is 43 dB and the signal noise ratio is −7 dB while the average value to the female voice is 45 dB and the signal noise ratio is −5 dB. The average speech recognition threshold value for the old group to the male voice is 48 dB and the signal noise ratio is −2 dB while the average value to the female voice is 52 dB and the signal noise ratio is 2 dB. So under the same noisy environment, the signal noise ratio should be at least 2 dB to make sure that more than 50 % people can identify the language correctly.

3.3 The Relationship Between Speech Recognition Score and the Signal Noise Ratio

The speech recognition score refers to the percentage of the speech signals heard and understood by subjects. The hearing loss caused by the age influences the speech recognition score, which is a function of the signal noise ration.

From the Fig. 3, it could be found that when the signal noise ratio reaches 4 dB, the accuracy of the young group to words recognition exceeds 90 % and has no obvious

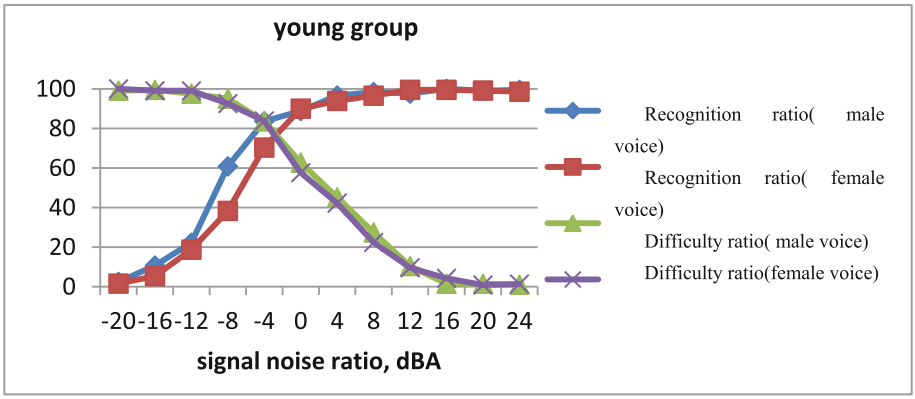

Fig. 3. The speech recognition accuracy and difficulty of the young group

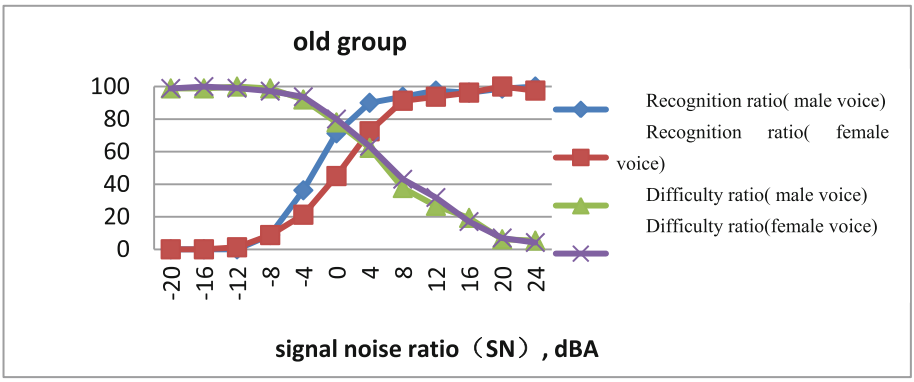

Fig. 4. The speech recognition accuracy and difficulty of the old group

change. When the signal noise ratio is under −4 dB, the difficulty of target words is over 80 % When the signal noise ratio reaches 8 dB and $F(2, 35) = 2.111$, there is no significant difference in statistics for the young group.

From the Fig. 4, it could be found that when the signal noise ratio of the old group reaches over 8 dB, their accuracy to words recognition is over 90 % and has no significant change. For the old group, when the signal noise ratio reaches 12 dB and $F(2, 37) = 2.408$, there is no significant difference in statistics.

When the environmental noise is 50 dB, in order to meet the requirements of speech recognition of the young people and to let the old people gain better hearing effects, in the sound pressure level design of audio equipment the signal noise ratio should reach 12 dB to make sure that most people can hear harmonious and beautiful sound level.

Acknowledgment. This work is supported by the National Key Technology R&D Program (project number: 2012BAK28B03-2) and China National Institute of Standardization through the "special funds for the basic R&D undertakings by welfare research institutions" (project number: 522013Y-3055).

References

1. Kobayashi, M., Morimoto, M., Sato, H., Sato, H.: Optimum speech level to minimize listening difficulty in public spaces. J. Acoust. Soc. Am. **121**(1), 251–256 (2007)
2. ISO 7029:2000: Acoustics statistical distribution of hearing thresholds as a function of age (2000)