Evaluation of speech intelligibility of sound fields in underground stations

Ryota Shimokura and Yoshiharu Soeta*

Health Research Institute, National Institute of Advanced Industrial Science and Technology (AIST),
1-8-31 Midorigaoka, Ikeda, 563-8577 Japan

(Received 30 March 2010, Accepted for publication 6 September 2010)

Keywords: Speech intelligibility, Public address system, Underground station, Impulse response measurement, Sound field

PACS number: 43.28.Lv [doi:10.1250/ast.32.73]

1. Introduction

The speech transmission quality of public address (PA) systems in railway stations is not usually very high [1]. In underground stations, the speech intelligibility of the PA system may be degraded by reflection and reverberation. In a reverberant sound field, hearing the content of PA announcements can be particularly difficult for elderly people [2], and poorly heard emergency announcements may cause increased safety risks when instructing passengers to exit the station [3].

Underground railway stations are normally long enclosures in which sound does not diffuse radially. Thus, the modeling of the sound field in such a long enclosure is different from that of a diffused field [4–9]. Although models can simulate the sound fields in a long enclosure, there have not been any studies on evaluating the sound field in underground railway stations with respect to speech intelligibility.

The aim of this study is to evaluate the speech intelligibility of the sound field in underground railway stations. Impulse responses were measured in two underground stations by positioning a sound source and receiver on the ceiling and the platform, simulating the loudspeaker of a PA system and a passenger, respectively. From the impulse response, the acoustical parameters were calculated to evaluate the speech intelligibility of the PA system.

2. Impulse response measurements

2.1. Measurement stations

The stations (denoted as “A” and “B”), at which measurements were taken, are on the same underground line and have island-type platforms (one platform at the center and railway tracks on either side). The columns and ceiling are covered by metallic panels, and the platforms and railway floors are covered by stone panels and concrete, respectively. However, the lateral walls of the platforms are covered by metallic panels in station A and wooden panels in station B. The platform in station B (length: 332 m, maximum width: 8 m) is longer than that in station A (length: 162 m, maximum width: 7.3 m). Since the underground line had not yet begun operations at that time, there were no passengers or trains in the station. Therefore, the background noise was small (46 dB).

2.2. Sound source and receiver

One sound source and four receiver positions (r1 to r4) were located at the center of the platform (Fig. 1(a)). To represent the various PA loudspeakers in use in railway stations, omnidirectional sources were used to measure sound fields of the underground station [7] and train compartment [10], because the speech intelligibility and loudspeaker's directivity were not directly related to each other [11]. The sound source we used was an omnidirectional loudspeaker (Type 4292, B&K), as shown in Fig. 2(a). It was located close to the loudspeaker of the PA system on the ceiling of the stations (Fig. 1(b)). The receiver was a dummy head microphone (KU100, Neumann), as shown in Fig. 2(b). The dummy head microphone is a model of a human head with two microphones inserted at the locations of the ear canals. It can simulate the frequency-dependent distortions of phase and amplitude in sound reaching the left and right ears. The dummy head microphone was arranged facing the sound source, and the height of the microphones in the dummy head was 1.6 m from the floor. In the stations, there are columns at 7.5 m intervals. Because of the locations of PA system loudspeakers, the sound source was close to the column (distance of 0.3 m) in station B (Fig. 2(b)). The distance was 1.2 m in station A (Fig. 2(a)).

2.3. Procedure

A sinusoidal signal with an exponentially varying frequency swept from 40 Hz to 20 kHz over a period of 18 s was used to measure the impulse responses. The responses of the receivers were deconvoluted so that the impulse responses could be obtained [12]. A laptop computer (Let's Note, Panasonic) generated the signals and saved the responses via an AD/DA converter (AudioFire8, Echo Digital Audio) at a sampling rate of 48 kHz and a sampling resolution of 32 bits.

From the impulse responses of the dummy head microphone, the strength (G), early decay time (EDT), clarity (C_{50}) [13], speech transmission index (STI) [14,15], and interaural cross-correlation coefficient ($IACC$) [13,16] were calculated to evaluate the speech intelligibility in the sound field.

3. Results

Figure 3 shows the acoustical parameters obtained in stations A and B. The values of G , EDT , and C_{50} were the mean of the values calculated from the impulse responses at the left and right ear positions. The G was attenuated more than 4 kHz and less than 250 Hz, respectively. Peaks of EDT were observed in the 500 Hz octave band. The C_{50} began to

*e-mail: y.soeta@aist.go.jp

Fig. 1 (a) Plan and (b) cross-sectional views of underground station, with sound source (○) and receiver (●) positions.

Fig. 2 Photographs of measurement environments: (a) station A with sound source and (b) station B with receiver.

increase at the 500-Hz octave band. The *IACC* decreased up to the 500-Hz octave band, and remained constant or increased beyond the band. The *G*, *C*₅₀, and *IACC* decreased, and the *EDT* increased as the distance between the source and receivers became greater. The *EDT* in station A was longer than that in station B for r3 and r4.

The speech intelligibility is ranked according to the speech-weighted *C*₅₀ and *STI* [17,18]. Figure 4 shows the speech-weighted *C*₅₀ and *STI* as a factor of the distance from the source to the receiver and the rank order of speech intelligibility, respectively. The speech-weighted *C*₅₀ is respectively weighted and averaged ($0.15C_{50}(500\text{ Hz}) + 0.25C_{50}(1\text{ kHz}) + 0.35C_{50}(2\text{ kHz}) + 0.25C_{50}(4\text{ kHz})$). In terms of speech-weighted *C*₅₀ and *STI*, the speech intelligibility was “good” or “fair” when the distance was within 15 m. However, the speech intelligibility became “poor” when

Fig. 3 (a) *G*, (b) *EDT*, (c) *C*₅₀, and (d) *IACC* in A-weighted overall sound and one-octave filtered bands. The solid symbols indicate receiver positions r1 (●), r2 (▲), r3 (■), and r4 (◆) in station A, and the non solid symbols indicate r1 (○), r2 (△), r3 (□), and r4 (◇) in station B.

Fig. 4 (a) Speech-weighted *C*₅₀ and (b) *STI* as a function of distance from the source in stations A (●) and B (○).

the distance was more than 15 m. Also, the speech-weighted *C*₅₀ and *STI* in station A were smaller than those in station B.

4. Discussion

The acoustical parameters as a function of the distance from the sound source to the receiver, except for *G* and *IACC*, were different in the two stations. At receiver positions (r3 and r4) far from the sound source, the *EDT* was longer and the speech-weighted *C*₅₀ and *STI* were smaller in station A. One possible reason is the difference in the interior materials at the lateral walls (metallic and wooden panels in stations A and B, respectively). Therefore, the average absorption coefficients of the stations were calculated according to the image source model for a long rectangular enclosure, because Sabine's reverberation equation is not appropriate for a nondiffused sound field such as an underground station [5]. When the geometrical data and reverberation times at r4 were assigned in the image source model, the average sound absorption coefficients were similar for stations A (0.17) and B (0.18).

Fig. 5 Impulse response at r4 in (a) station A and (b) station B. The arrival time of the direct sound was fixed at 0 ms for all impulse responses.

Thus, the difference in interior materials may not influence the sound fields in these stations significantly. Another possible reason is reflection from the other boundary. Figure 5 shows the impulse responses at r4 in stations A and B. In station A, the reflected sound came 20 ms after the direct sound. In contrast, in station B, strong early reflected sound came soon after the direct sound. Since the receiver position r4 in station B was in front of the closed shutter of an escalator, the sound reflected from it strengthened the early reflections, and hence the speech-weighted C_{50} and STI might become larger.

The speech intelligibility of the sound field was maintained at “good” or “fair” at the receiver position within 15 m from the sound source, according to the speech-weighted C_{50} and STI [17,18], and was maintained at “good” at all receiver positions, according to the EDT [19]. When the sound source and receiver were separated by more than 15 m, the subsequent reflections from the floor, ceiling, and lateral wall impaired the speech intelligibility. In station B, the speech intelligibility at the distance of 20 m from the sound source was better in terms of speech-weighted C_{50} and STI . As mentioned above, the improvement might be an effect of the reflective wall constructed on the platform. Therefore, to maintain speech intelligibility, our results suggest that PA system loudspeakers should be placed at intervals no greater than 15 m, or reflective walls should be installed on the platform, such as platform screens with doors. Such a transparent platform screen is often used to separate the track and platform for safety reasons. However, the platform screen, which insulates train noise and strengthens early reflection, would also help improve the speech intelligibility of the PA system in an underground station.

5. Conclusion

In the measured underground stations, the sound field provides good or fair speech intelligibility for PA system announcements when the distance between the passenger and loudspeaker is within 15 m. The acceptable distance can be lengthened by installing reflective walls on the platform to strengthen early reflection. In further studies, we must measure the acceptable distance between the passenger and

loudspeaker in an underground station with a platform screen to investigate the effect of early reflection on the speech intelligibility.

Acknowledgments

This work was supported by a Grant-in-Aid for Young Scientists (A) from the Japan Society for the Promotion of Science (18680025), a grant from Ono Acoustics Research Fund, and a Sasagawa Scientific Research Grant.

References

- [1] P. A. A. Kootwijk, “The speech intelligibility of the public address systems at 14 Dutch railway stations,” *J. Sound Vib.*, **193**, 433–434 (1996).
- [2] A. K. Nábělek and P. K. Robinson, “Monaural and binaural speech perception in reverberation for listeners of various ages,” *J. Acoust. Soc. Am.*, **71**, 1242–1248 (1982).
- [3] S. J. Wijngaarden and J. A. Verhave, “Prediction of speech intelligibility for public address systems in traffic tunnels,” *Appl. Acoust.*, **67**, 306–323 (2006).
- [4] J. Kang, “Acoustics in long enclosures with multiple sources,” *J. Acoust. Soc. Am.*, **99**, 985–989 (1996).
- [5] J. Kang, “Reverberation in rectangular long enclosures with geometrically reflecting boundaries,” *Acta Acust. Acust.*, **82**, 509–516 (1996).
- [6] J. Kang, “A method for predicting acoustics indices in long enclosures,” *Appl. Acoust.*, **51**, 169–180 (1997).
- [7] L. Yang and B. M. Shield, “The prediction of speech intelligibility in underground stations of rectangular cross section,” *J. Acoust. Soc. Am.*, **109**, 266–273 (2001).
- [8] K. M. Li and K. K. Lu, “Propagation of sound in long enclosures,” *J. Acoust. Soc. Am.*, **116**, 2759–2770 (2004).
- [9] K. M. Li and P. M. Lam, “Prediction of reverberation time and speech transmission index in long enclosures,” *J. Acoust. Soc. Am.*, **117**, 3716–3726 (2005).
- [10] R. Shimokura and Y. Soeta, “Evaluation of speech intelligibility of sound fields in passenger train compartments,” *Acoust. Sci. & Tech.*, **30**, 379–382 (2009).
- [11] K. D. Jacob, “Subjective and predictive measures of speech intelligibility — The role of loudspeaker directivity,” *J. Audio Eng. Soc.*, **33**, 950–955 (1985).
- [12] A. J. Berkhout, D. De Vries and M. M. Boone, “A new method to acquire response in concert halls,” *J. Acoust. Soc. Am.*, **68**, 179–183 (1980).
- [13] ISO 3382, *Acoustics — Measurement of the Reverberation Time of Rooms with Reference to Other Acoustical Parameters* (1997).
- [14] H. J. M. Steeneken and T. Houtgast, “A physical method for measuring speech-transmission quality,” *J. Acoust. Soc. Am.*, **67**, 318–326 (1980).
- [15] M. R. Schroeder, “Modulation transfer functions: Definition and measurement,” *Acustica*, **49**, 179–192 (1981).
- [16] Y. Ando and Y. Kurihara, “Nonlinear response in evaluating the subjective diffuseness of sound field,” *J. Acoust. Soc. Am.*, **80**, 833–836 (1986).
- [17] L. G. Marshall, “An acoustics measurement program for evaluating auditoriums based on the early/late sound energy ratio,” *J. Acoust. Soc. Am.*, **96**, 2251–2261 (1994).
- [18] ISO 9921, *Ergonomics — Assessment of Speech Communication* (2003).
- [19] Ha. Sato, M. Morimoto, Hi. Sato and M. Wada, “Relationship between listening difficulty and acoustical objective measures in reverberant sound fields,” *J. Acoust. Soc. Am.*, **123**, 2087–2093 (2008).