Evaluation of a short reverberant sound field using syllable intelligibility and listening difficulty under multitalker babble

Toshie Matsui1,*, Shuichi Yanai2, Ryota Shimokura3 and Hiroshi Hosoi3

1University of Tsukuba, Life Science Center for Tsukuba Advanced Research Alliance (TARA), 1–1–1 Tennodai, Tsukuba, 305–8577 Japan
2Tokyo Metropolitan Institute of Gerontology, Aging Neuroscience Research Team, 35–2 Sakae-cho, Itabashi-ku, Tokyo, 173–0015 Japan
3Nara Medical University, Department of Otorhinolaryngology—Head and Neck Surgery, 840 Shijo-cho, Kashihara, 634–8521 Japan

(Received 7 February 2014, Accepted for publication 9 April 2014)

Keywords: Syllable intelligibility, Reverberation time, Multitalker babble, Listening difficulty

PACS number: 43.66.Ba, 43.66.Lj, 43.71.Bp [doi:10.1250/ast.35.279]

1. Introduction

Various methods have been developed to evaluate reverberant sound fields in terms of the quality of speech communication. Researchers have mainly used listening performance, such as word intelligibility or syllable intelligibility, for evaluation. However, there was the difficulty that listening performance scores could not reflect the differences among sound fields because listening performance scores were often similarly high in sound fields under good conditions such as very short reverberation time. To overcome this difficulty, the use of degraded speech sounds [1], and the introduction of the measurement of the ease of conversation rather than the measurement of listening performance only [2] have been adopted.

The evaluation of speech communication in a short reverberant sound field can be useful for an acoustic evaluation of a standard Japanese house. One study about Japanese housing revealed that the most annoying noise reported by residents was speech sounds heard through a wall [3]. Given the frequent exposure to speech sounds, produced by both humans and various types of media, in daily life, it is possible that such speech sounds are sufficiently distracting that they degrade speech intelligibility in short reverberant sound fields.

Additionally, it has been reported that “listening difficulty” [4] could be a more precise index for the evaluation of a sound field under good conditions of reverberation [5]. The systematic demonstration of the difference between syllable intelligibility as a listening performance and listening difficulty in a short reverberant sound field with distracting noises should be considered.

In this study, we measured syllable intelligibility and listening difficulty by exposing participants to external noise composed of speech sounds in a short reverberant sound field. Our aim was to investigate the potential utility of such noise and listening difficulty as a basis of assessing sound environments especially with a short reverberation time.

---

* e-mail: tmatsui@tara.tsukuba.ac.jp
The syllables were either presented independently or 1.0–1.9 s after the onset of a multitalker babble stimulus. The listeners were instructed (a) to write down or type the syllables that they heard, and (b) to evaluate “listening difficulty” for each syllable. No feedback was given. In the listening difficulty task, the listeners were instructed to use numbers ranging from 0 (easiest to understand) to 100 (most difficult to understand) to rate listening difficulty. The listeners were allowed to proceed through the two tasks at their own pace.

The experiment consisted of 240 trials (20 syllables × six T60 levels × with/without multitalker babble). The participants were given 40 practice trials before the experiment. The trials were separated into blocks with multitalker babble stimuli and blocks without multitalker babble stimuli.

All the experiments were performed using a GUI generated by MATLAB (Mathworks) on a computer (Apple PowerBook Pro). The stimuli were diotically presented through a set of headphones (Sennheiser HD650). The experiment took place in an adequately quiet room.

3. Results

We performed an analysis of variance (ANOVA) to test individual syllable intelligibility for all the conditions [Fig. 1(a)]. The effects of both T60 and multitalker babble were significant \([F(1,108) = 76.77, p < 0.0001]\). The main effects of T60 and multitalker babble were significant \([F(5,108) = 10.96, p < 0.0001\) and \(F(1,108) = 725.67, p < 0.0001\), respectively]. The interaction between these factors was also significant \([F(5,108) = 12.79, p < 0.0001]\). A Tukey-Kramer HSD test confirmed significant differences between the five T60 levels from 0 s to 0.4 s and that at 0.5 s, and between the two shortest T60 levels and the T60 of 0.4 s \((\alpha = 0.05)\).

As with the syllable intelligibility, we performed an ANOVA to test individual listening difficulties for all conditions [Fig. 1(b)]. The effects of two factors were significant \([F(11,108) = 73.21, p < 0.0001]\). The main effects of the T60 and multitalker babble were significant \([F(5,108) = 2.62, p = 0.03\) and \(F(1,108) = 790.75, p < 0.0001\), respectively]. However, the interaction between the factors was not significant \([F(5,108) = 0.29, p = 0.92]\). A Tukey-Kramer HSD test for the effect of T60 revealed a significant difference between no reverberation and 0.5 s \((\alpha = 0.05)\).

Spearman’s rank correlation coefficients \((\rho)\) were calculated between T60 and syllable intelligibility as well as between T60 and listening difficulty. There was no significant correlation between T60 and syllable intelligibility under the no multitalker babble condition \((\rho = 0.15, p = 0.24)\). With multitalker babble, a strong and significant correlation was found between T60 and syllable intelligibility \((\rho = -0.73, p < 0.0001)\) \((\alpha = 0.05)\). In the presence and absence of multitalker babble, listening difficulty correlated positively with T60, although the \(p\) value for the no multitalker babble condition did not reach the significance level [with multitalker babble: \(\rho = 0.47, p = 0.0001\); without multitalker babble: \(\rho = 0.19, p = 0.14\)].

4. Discussion

The results represented a different trend between listening difficulty and syllable intelligibility. The T60 under 0.5 s had little effect on syllable intelligibility when there was no multitalker babble. We found that there was no significant difference among the four levels from 0 s to 0.3 s in terms of T60. The 0.3 s as a boundary is roughly consistent with the threshold of “reverberation annoyance” in an utterance task in the previous study [2]. It is conceivable that such a type of annoyance would come out into the open as a decrease in syllable intelligibility by adding an external noise.

Listening difficulty was influenced by T60 in the presence or absence of multitalker babble, as indicated by ANOVA. Listening difficulty differs from syllable intelligibility in that no ceiling effect was observed without multitalker babble. It can be considered that listening difficulty reflected the changes in reverberation time in contrast to syllable intelligibility. However, the multiple comparison showed a significant difference only between 0 s and 0.5 s. The correlation coefficient was not sufficiently high and not significant under the no multitalker babble condition. There is a possibility that the large standard deviation caused those results of statistical tests. Individual differences must be reduced to use listening difficulty as an index of speech communication in a very short reverberant sound field regardless of external noise.
Note that the impulse responses used in this experiment are atypical for a room in a standard house. The duration from direct sound to early-reflected sound of the impulse response used in the experiment appears long (about 50 ms). We predict that a similar evaluation could result in higher syllable intelligibility and lower listening difficulty with the impulse response of a room in a real house.

5. Summary

We used multitalker babble to degrade syllable intelligibility and listening difficulty while evaluating a short reverberant sound field. We found that listening difficulty increased with T60 with or without multitalker babble, while syllable intelligibility decreased with T60 only under the multitalker babble condition. The post hoc test using multiple comparison showed that syllable intelligibility decreased with T60 beyond 0.4 s with multitalker babble.

Acknowledgments

This study was supported by Daiwa House Industry Co., Ltd. and JSPS KAKENHI Nos. 21791636, 23730715, 25861583.

References

[1] Y. Nagatani, T. Sakaguchi and H. Hosoi, “Evaluation of acoustic environments using deteriorated speech sound,” Proc. Acoustics ’08, pp. 1249–1253 (2008).
[2] H. Sato, M. Morimoto and K. Kobuki, “The effect of reverberation sound on conversation in living rooms,” J. Acoust. Soc. Jpn. (J), 66, 541–551 (2010) (in Japanese).
[3] S. Yanai, T. Sakaguchi, H. Hosoi and N. Suzuki, “The factor structure of noisy sound through walls inside residences,” Behav. Sci. Res., 48, 115–122 (2010).
[4] M. Morimoto, H. Sato and M. Kobayashi, “Listening difficulty as a subjective measure for evaluation of speech transmission performance in public spaces,” J. Acoust. Soc. Am., 116, 1607–1613 (2004).
[5] H. Sato, J. S. Bradley and M. Morimoto, “Using listening difficulty ratings of conditions for speech communication in rooms,” J. Acoust. Soc. Am., 117, 1157–1167 (2005).
[6] Architectural Institute of Japan, Ed., Sound Materials in Living Environment (SMILE2004) (Gihodo, Tokyo, 2004).
[7] Japan Audiological Society, Ed., “The 67-S syllable list” (1987).
[8] S. Amano, S. Sakamoto, T. Kondo and Y. Suzuki, “NTT-Tohoku University familiarity controlled word lists (FW03), speech resources consortium. Japan,” National Institute of Informatics (2006).
[9] Y. Nagatani, R. Tachibana, T. Sakaguchi and H. Hosoi, “Loudness calibration of monosyllabic speech sounds in FW03,” J. Acoust. Soc. Jpn. (J), 64, 647–649 (2008) (in Japanese).
[10] K. Takahashi, “The speaking rate variation speech database (SRV-DB),” http://www.it.ice.uet.ac.jp/SRV-DB/ (2014-02-07).