Applied Acoustics 97 (2015) 65–77

Contents lists available at ScienceDirect

Applied Acoustics

journal homepage: www.elsevier.com/locate/apacoust

Comparisons between simulated and in-situ measured speech intelligibility based on (binaural) room impulse responses

Peisheng Zhu [a], Fangshuo Mo [b,*], Jian Kang [c,*], Guofeng Zhu [a]

[a] School of Architecture and Fine Art, Dalian University of Technology, Dalian 116023, China
[b] Institute of Acoustics, Tongji University, Shanghai 200092, China
[c] School of Architecture, University of Sheffield, Sheffield S10 2TN, UK

ARTICLE INFO

Article history:
Received 17 January 2014
Received in revised form 27 March 2015
Accepted 10 April 2015
Available online 24 April 2015

Keywords:
Speech intelligibility
Simulation
In-situ measurement
(Binaural) room impulse responses

ABSTRACT

This study systematically compares acoustic simulation and in-situ measurement in terms of speech transmission index (STI), speech intelligibility scores and relationship curves when considering (binaural) room impulse response and four general room conditions, namely, an office, a laboratory, a multimedia lecture hall and a semi-anechoic chamber. The results reveal that STI can be predicted accurately by acoustic simulation (using room acoustics software ODEON) when there is a good agreement between the virtual models and the real rooms and that different reverberation time (RT) and signal-to-noise ratio (SNR) may exert less significant influence on the simulated STI. However, subjective intelligibility may be overestimated when using acoustic simulation due to the head-related transfer function (HRTF) filter used, and the score bias may be minimal and difficult to detect in everyday situations. There is no obvious score tendency caused by different RT, though with the decrease in the SNR, score bias may increase. Overall, considering that the accurate acoustic modelling of rooms is often problematic, it is difficult to obtain accurate speech intelligibility prediction results using a simulation technique, especially when the room has not yet been built.

1. Introduction

Speech intelligibility is an important metric and can be used to evaluate the sound transmission quality of auditoriums. The assessment of speech intelligibility mainly includes subjective evaluation and objective evaluation [1,2]. However, performing such measurements in real rooms has limitations, such as schedule conflicts, and it is difficult to perform a speech intelligibility test simultaneously with a large number of subjects at a single receiver position. In recent years, the rapid progress in the acoustic simulation technique offers a potential solution to these limitations and provides an unlimited capacity to reproduce the same listening environments while also making it possible for speech intelligibility to be assessed in a room before it is built [3–8]. However, before it can be used with confidence, the acoustic simulation technique must be validated in comparison with in-situ measurement in real rooms.

Subjective intelligibility tests were performed in virtual and real classrooms, and the results were compared by Yang and Hodgson [3] using the CATT-Acoustics prediction and auralization system. The results showed that auralized subjective intelligibility tests were found to be reliable if the classroom was neither very absorptive nor noisy. However, in their study, the comparison of the objective evaluation metric speech transmission index (STI) was not involved. Subjective intelligibility tests were also performed in virtual and real classrooms, and the results were compared by Hodgson et al. [4] using the CATT-Acoustics and ODEON prediction and auralization system. The results suggested that auralization is not accurate in the case of high noise or low reverberation. The comparison of the objective evaluation metric STI, however, was still not involved in this study. Peng et al. [5–8] made many meaningful attempts on using acoustic simulation technique to assess the speech intelligibility of Chinese. The results showed that the relationship between the subjective intelligibility scores and STI can be better reflected based on acoustic simulation, which is an effective method for the evaluation of speech intelligibility. However, their conclusions obtained are mainly based on simulation, and in-depth comparison and validation with the in-situ measurement are still needed. Overall, there is still a lack of study on the systematic comparison and validation of simulation technique for the evaluation of speech intelligibility.

The aim of this study is therefore to systematically compare the simulated speech intelligibility scores, STI and the curve thus

---

* Corresponding authors.
E-mail addresses: mfs@tongji.edu.cn (F. Mo), j.kang@sheffield.ac.uk (J. Kang).

http://dx.doi.org/10.1016/j.apacoust.2015.04.005

produced with that of the in-situ measured. Finally, the influence factors for simulation bias, if any, should be considered carefully.

2. Methods

This section starts with selecting four general rooms and then, based on the room impulse response (IR) and binaural room impulse response (BRIR) and the STI and Chinese speech intelligibility scores of 12 receiver positions, a total of 48 listening environments in the four general rooms were obtained by the two types of methods: in-situ measurement and acoustic simulation. A general flowchart of this study is given in Fig. 1.

2.1. Experimental arrangement

Four general rooms were selected as the test rooms in this study, including an office, a laboratory, a multimedia lecture hall and a semi-anechoic chamber (with one desk and four chairs inside), of which, the office, laboratory and semi-anechoic chamber are rectangular, and the multimedia lecture hall is octagon. There are two receiver positions arranged in the office, three receiver positions in the laboratory, six receiver positions in the multimedia lecture hall and one receiver position in the semi-anechoic chamber. The layout of the receiver positions and the sound sources are shown in Fig. 2.

To obtain a wide range of the STI, an interference noise source (monitor loudspeaker GENELEC 8020B) was placed at a distance of 0.5 m beside the signal source. A dodecahedral sound source was not used as an interference noise source in this experiment because, for the dodecahedral sound source, there was no main radiation and the directivity changed with orientations, the equalisation and calibration was difficult [9], and room acoustics software could hardly simulate a real dodecahedral sound source. Accordingly, these factors may exert significant influence on the comparison results. In an anechoic chamber, the sound pressure level (SPL) on the front axis at 1 m of the signal source (artificial mouth GRAS 44AA) was set at 60 dBA [10]. The noise source reproduced a males spectra shaped [10] pink noise, and the SPL was adjusted simultaneously to make the positions 1 m away from the two sound sources correspond to four distinct relative background noise levels (RBNLs): 5 dB, 0 dB, −10 dB, and −20 dB. The SPL on the front axis at 1 m of the monitor loudspeaker 8020B was set as 65 dBA, 60 dBA, 50 dBA, and 40 dBA, respectively. The RBNL equals the signal-to-noise ratio (SNR) in a noiseless anechoic chamber; however, due to the influence of different reflections and sound source directivity patterns, and possible environmental noise, the RBNL does not equal the actual SNR at the R_1–R_{12} receiver positions. The signal source and the noise source, pre-set in an anechoic chamber, were placed in the corresponding sound source positions in the test rooms and at each receiver position, the STI, IR and BRIR as well as the operational speech level and background noise levels were each measured in turn. Both of the sound source systems were equalised using their inverse filter systems calculated from the impulse responses measured on the front axis of the sources in an anechoic chamber [9].

2.2. Virtual room modelling

Four room models were erected corresponding to the four real rooms using the ODEON version 12.0 [11] room acoustics software. During the simulation, the virtual signal source and the virtual interference noise source, namely, virtual-44AA and virtual-8020B, respectively, were erected in ODEON using the horizontal and vertical directivity patterns of the artificial mouth GRAS 44AA and the monitor loudspeaker 8020B. In the ‘Directivity Polar Plot Editor’ menu, both the virtual-44AA and the virtual-8020B were marked with ‘Natural’, the horizontal and vertical directivity patterns of each octave band were established, and the ‘+EQ’ of each octave band was adjusted to ensure that, in the ‘Point Source Editor’ menu, the SPL of the virtual-44AA on the front axis at 1 m (which should be 20 dB higher than the SPL on the front axis at 10 m) was set to the same octave band (from 125 to 8000 Hz) SPL as that measured on the front axis at 1 m in an anechoic chamber when reproducing a composite signal of seven half-octave band carriers without modulation with a SPL of 60 dBA using the artificial mouth 44AA. The SPL of the virtual-8020B on the front axis at 1 m was set to the same octave band (from 125 to 8000 Hz) SPL as that measured on the front axis at 1 m in an anechoic chamber when reproducing a male spectra shaped [10] pink noise with a SPL of 60 dBA using the monitor loudspeaker 8020B. In addition, in the ‘Point Source Editor’ menu, the ‘ + EQ’ of each octave band was set to 0 dB for both virtual-44AA and virtual-8020B, the ‘Overall gain’ was set to 0 dB for the virtual-44AA, and 5 dB, 0 dB, −10 dB, and −20 dB, respectively, for the virtual-8020B. The octave band SPL for the artificial mouth GRAS 44AA and the monitor loudspeaker 8020B measured on the front axis at 1 m in an anechoic chamber are presented in Table 1. In Fig. 3, the horizontal and vertical directivity patterns of the monitor loudspeaker 8020B and the artificial mouth GRAS 44AA at 500, 1000, 2000, and 4000 Hz are shown based on the data provided by the manufacturer of the monitor loudspeaker 8020B, while the data for the artificial mouth 44AA were obtained through

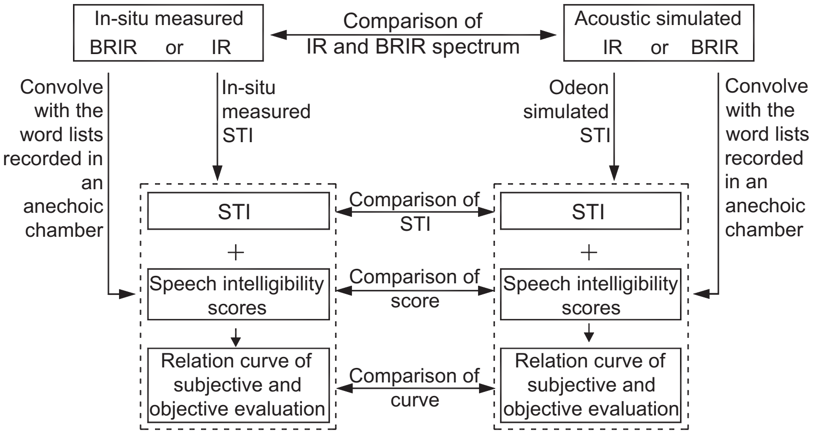

Fig. 1. Flowchart of the general experimental procedure.