Applied Acoustics 81 (2014) 1–14

Contents lists available at ScienceDirect

Applied Acoustics

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

Accuracy of speech transmission index predictions based on the reverberation time and signal-to-noise ratio

Laurent Galbrun *, Kivanc Kitapci

School of the Built Environment, Heriot-Watt University, Edinburgh EH14 4AS, UK

ARTICLE INFO

Article history:
Received 6 August 2013
Received in revised form 11 November 2013
Accepted 4 February 2014
Available online 28 February 2014

Keywords:
Acoustic measurements
Room acoustics
Speech intelligibility
Sound quality

ABSTRACT

This paper examines the accuracy of the speech transmission index (STI) calculated from the reverberation time (T) and signal-to-noise ratio (L_{SN}) of enclosed spaces. Differences between measured and predicted STIs have been analysed in two rooms (reverberant vs. absorbent), for a wide range of absorption conditions and signal-to-noise ratios (sixteen tests). The STI was measured using maximum length sequence analysis and predictions were calculated using either measured or predicted values of T and L_{SN}, the latter assuming diffuse sound field conditions. The results obtained for all the conditions tested showed that STI predictions based on T and L_{SN} tend to underestimate the STI, with differences between measured and predicted STIs always lower than 0.1 (on a 0.0–1.0 scale), and on average lower than 0.06. According to previous research, these differences are noticeable and therefore non-negligible, as 0.03 is the just noticeable difference in STI. The use of either measured or predicted values of T and L_{SN} provided similar STI predictions (i.e. non-noticeable changes), with differences between predictions that are on average lower than 0.03 for the absorbent room, and lower than 0.01 for the reverberant room.

1. Introduction

Speech intelligibility is a sound quality descriptor that can be used to analyse the suitability of spaces where speech is crucial (e.g. teaching rooms, meeting and conference rooms). The speech intelligibility properties of a space can be quantified either through listening tests or through physical measurements. The latter method is normally used in room acoustics, as it is an objective measurement and is also much faster than listening tests which are subject based. In particular, the speech transmission index (STI) is commonly used to measure the speech intelligibility of enclosures. The STI is an electronic method which was developed by Houtgast and Steeneken [1] and which normally requires specialist equipment or software to calculate the modulation transfer function (MTF). The MTF forms the basis of the STI method and is typically determined from impulse responses [2], but can also be estimated from the reverberation time and signal-to-noise ratio present in the space [3]. However, the accuracy of this simple acoustic method is not documented in the literature, and is therefore examined in this paper.

The importance of being able to quantify the STI from simple room acoustic parameters lies in the fact that this allows determining a fundamental design parameter without the need of specialist equipment or software (e.g. maximum length sequence software or ray tracing software), as a simple spreadsheet can be used. This method can therefore be used by non-specialists for design purposes or acoustic assessments. Its accuracy needs however to be known, in order to define its applicability and limitations. This is achieved in the current study by comparing STI values obtained from the impulse response method based on maximum length sequence analysis [2] and for which accuracy is known, with STI values calculated from the reverberation time and signal-to-noise ratio [3]. Two rooms have been tested under sixteen different acoustic conditions (different reverberation times and signal-to-noise ratios), allowing to examine a wide range of STI values (0.1–0.8) and carrying out a detailed analysis.

The paper begins with the background theory to the STI and a description of the methodology used. Results are then presented and analysed, followed by conclusions where the main findings are summarised.

2. Background theory

In the 1970s, Houtgast and Steeneken [1] developed an electronic method which has since then been used to measure speech intelligibility. In this method, speech is modelled as modulated bands of noise, and the signal's distortion between the source

---

* Corresponding author. Tel.: +44 131 4513145.
E-mail address: L.G.U.Galbrun@hw.ac.uk (L. Galbrun).

http://dx.doi.org/10.1016/j.apacoust.2014.02.001
0003-682X/

and receiver can be quantified using the modulation transfer function (MTF). The MTF effectively mimics the behaviour of speech and quantifies the reduction in modulation between the source and receiver using the modulation reduction factor $m(f_m)$, where f_m is the modulation frequency. This method is well established and its details can be found in the literature [1,4,5].

In practice, the MTF can be calculated by producing an electronic signal and measuring the reduction in modulation of the signal at a receiving position; this is normally achieved through the use of the impulse response of the signal. A common method used to calculate the MTF from impulse responses, and which is used in this study, is to apply maximum length sequence analysis [2]. Maximum length sequences are binary periodic sequences which correlate with impulses; the accuracy of maximum length sequence methods has been demonstrated [2,6], and Rife [2] provided a calculation example in which the STI value was found to be within 0.6% of the exact STI value. This method is applied here through the use of the MLSSA software (Maximum Length Sequence System Analyzer, DRA Laboratories (Sarasota, USA)).

Alongside signal processing analysis, Houtgast et al. [3] have shown that the MTF can be calculated from the room acoustic properties of a space, and more specifically from the reverberation time and signal-to-noise ratio present in the space, using the equation

$$m(f_m) = \frac{1}{\sqrt{1 + \left(2\pi f_m \frac{T}{13.8}\right)^2}} \times \frac{1}{1 + 10^{-0.1 L_{SN}}} \qquad (1)$$

where $m(f_m)$ is the modulation reduction factor, L_{SN} is the signal-to-noise level (dB), f_m is the modulation frequency (Hz) and T is the room's reverberation time (s). Eq. (1) shows that simple room acoustic parameters are sufficient for calculating the MTF, thus removing the need for using impulse responses. The comparison of this method with impulse response methods, as well as its accuracy, is however not documented in the literature.

Table 1
Materials and furniture present in the reverberant and absorbent rooms.

	Reverberant room	Absorbent room
Walls	Brickwork, plasterboard	Concrete blocks, single glazed windows, whiteboards, screen, pin board
Floor	Concrete	Thin carpet
Ceiling	Concrete	Suspended ceiling with mineral fibre tiles
Furniture	None	Desks (veneered chipboard), upholstered chairs

Regardless of the MTF method used, speech intelligibility results are expressed using the speech transmission index (STI) [4]. The STI is calculated from the modulation reduction factor $m(f_m)$, with f_m ranging from 0.63 to12.5 Hz in 1/3 octave intervals, and each $m(f_m)$ is calculated for octave bands from 125 Hz to 8 kHz. To obtain the STI, the apparent signal-to-noise ratio, L_{SNapp} (dB), should first be calculated from

$$L_{SNapp} = 10 \log \frac{m(f_m)}{1 - m(f_m)} \qquad (2)$$

where L_{SNapp} is the signal-to-noise ratio that would have produced the modulation reduction factor $m(f_m)$, had all the distortion been caused by interfering noise [5]. L_{SNapp} is then averaged over all modulation frequencies for each octave band frequency (125 Hz to 8 kHz) to give seven average L_{SNapp} values. These average L_{SNapp} values are then summed to give a single weighted average apparent signal-to-noise ratio, $\overline{L_{SNapp}}$, (dB) according to

$$\overline{L_{SNapp}} = \sum_{i=1}^{7} w_i (L_{SNapp}) \qquad (3)$$

where w_i is the weighting used for octave bands from 125 Hz to 8 kHz (= 0.13, 0.14, 0.11, 0.12, 0.19, 0.17, 0.14) [5]. Lastly, the STI can be calculated from this single $\overline{L_{SNapp}}$ using the formula [5]

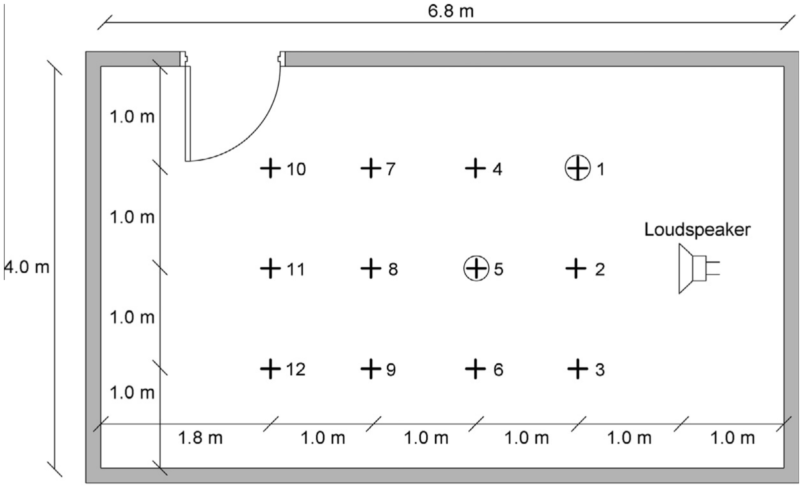
$$STI = (\overline{L_{SNapp}} + 15)/30 \qquad (4)$$

noting that STI = 1 when $\overline{L_{SNapp}} \geq 15$ dB and STI = 0 when $\overline{L_{SNapp}} \leq -15$ dB.

The results presented in this paper are based on this STI calculation procedure, and further details about the methodology applied are described in the following section.

3. Methodology

Two medium sized rooms were selected within the School of the Built Environment of Heriot-Watt University (material properties listed in Table 1). One was an empty laboratory chamber (Fig. 1) with reflective surfaces and no windows (named 'reverber-



+ Measurement positions

(central and off-axis positions analysed in the results' section are circled)

Fig. 1. Floor plan of the reverberant room showing the sound source and 12 receiver positions tested.