A Computer Model for Calculating the Speech Transmission Index Using the Direct STIPA Method

Paweł DZIECHCIŃSKI
Wrocław University of Science and Technology, Department of Acoustics and Multimedia, Wybrzeże Wyspiańskiego 27, 50-370 Wrocław,
pawel.dziechcinski@pwr.edu.pl

Abstract

Computer models currently used for the simulation of the speech transmission index (STI) calculate the STI using the statistical method or are based on numerically determined impulse response of the transmission channel. The limitation of both these computational methods is that they do not allow to take into account the nonlinear properties of the transmission channel and fluctuating background noise. This paper presents a proposition of a model based on the direct STIPA method. This model allows computer simulations of STIPA for distributed sound systems, and enables analysis to include both changes in signal dynamics and fluctuating background noise. The paper presents the idea of the model and validation of its basic elements - the generator and the analyser. The possibilities of using the model for computer simulation of outdoor public address systems were also discussed.

Keywords: speech transmission index, STIPA, attenuation of sound by the atmosphere

1. Introduction

The most often used method of rating the speech intelligibility of sound systems is the STI (speech transmission index) method, which is the subject to Houtgast and Steeneken works [5, 11], Brachmański papers [1, 2] and IEC 60268-16 standard [6]. This method is based on the fact that the speech signal can be divided into two spectra – audible spectrum (octave bands from 125 Hz to 8 kHz) and modulation spectrum, including the phonemes (phoneme – the smallest unit of speech distinguishable for users of a given language) – from 0.63 to 12.5 Hz. When the signal passes through the transmission channel (electroacoustic system and chamber), the signal level in each octave band and the modulation depth for different modulation frequencies change, which is mainly associated with the presence of noises and reverberation in the channel. The single-number STI value is determined based on the modulation transfer function, which is the ratio of the respective modulation indexes at the output to the indexes at the input of the channel. Measurements of speech transmission index can be conducted using the direct method – analysing the modulation depth of a special measurement signal, or indirect – determining modulation indexes based on the impulse response of the transmission channel – by the Schroeder equation [10]. Computationally the speech transmission index can be determined by the statistical method or on the basis of the numerically determined channel impulse response [6]. The limitation of both calculation methods is that they allow including neither the nonlinear properties of the transmission channel nor the fluctuating background noise. Although properly adjusted sound system should not be a source of large nonlinear distortion, but in practice commonly used nonlinear elements are e.g. dynamics processors. The paper proposes a computer model that allows taking into account in the STI determination both changes in the dynamic range of the signal and temporally fluctuating background noise.

The influence of various factors on speech quality (including intelligibility) was discussed, among others, by Brachmański [1]. In his book, however, he did not deal with sound systems, whose special feature is the transmission of sound over long distances. This is partly related to the problem of sound attenuation by the atmosphere. The proposed model allowed to assess the impact of this phenomenon on STI.

2. Model concept

The proposed model is based on a direct measurement method, a simplified STI version, which is STIPA (speech transmission index for public address systems). The general scheme of the model (Figure 1) assumes that the STIPA signal can come from $n \geq 1$ number of loudspeakers. In general, the sound reaching the observation point will therefore be a superposition of signals coming from individual loudspeakers and noises occurring at this point.

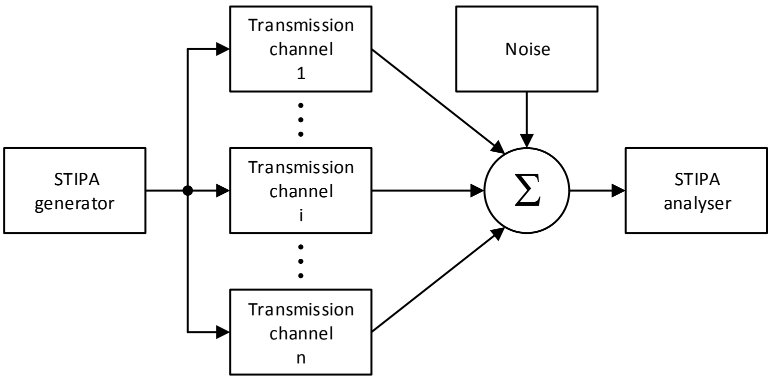

Figure 1. General structure of the proposed calculation model

In a typical electrical input sound system, the transmission channel consists of devices for processing the audio signal, loudspeakers, and sound propagation environment (Figure 2). The audio signal transmission chain can be divided into two parts:

1) A-chain – from the signal source to the main gain control of the loudspeaker channel,
2) B-chain – from the main gain control of a given loudspeaker channel to a loudspeaker.

In the A-chain, signal processing processes common to all transmission channels are carried out, such as dynamic compression or frequency characteristic equalisation. The B-chain is a loudspeaker processor and power amplifier. The audio processing path can be modelled by applying signal processing procedures that correspond to those used in reality, or if their algorithms are not known, by generating a file with the STIPA signal and processing it in a real processor. A similar processing procedure can be used to process not only the STIPA signal, but also other audio samples. This possibility was designed for processing samples to be used in subjective tests.

Depending on the available data, the loudspeaker will be modelled as a filter with appropriate frequency characteristics or by convolution of a loudspeaker impulse response with a signal.

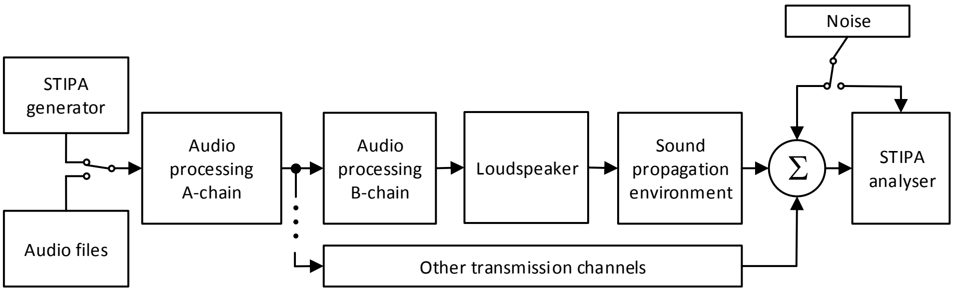

Figure 2. Detailed structure of the proposed calculation model

The sound propagation environment can be a room, an external environment or a free field. In the case of a room, it is assumed that the sound propagation environment will be modelled by convolution of the room impulse response with the signal. The impulse response can be obtained by means of measurements or as a result of modelling in an external programme. The algorithms from ISO 9613-1 [7] and ISO 9613-2 [8] standards will be used to model the external environment. The current version of the sound propagation model in the environment has been limited to the phenomenon of sound attenuation by the atmosphere. The simplest way to take into account the effect of this phenomenon on STIPA is to determine sound attenuation for octave bands from 125 to 8000 Hz and appropriate attenuation of the signal spectrum at the generation stage. Due to the fact that the model was also planned for processing signals other than STIPA, a different algorithm was chosen. It consisted of:

1) determining the frequency response of sound attenuation by the atmosphere on the basis of ISO 9613-1 standard,
2) designing a digital filter with determined frequency characteristics,
3) performing signal filtration.

The frequency characteristics of the filter resulting from sound attenuation by the atmosphere for a source-microphone distance *d* equal to 200 m, for two of the four analysed cases are shown in Figure 3.

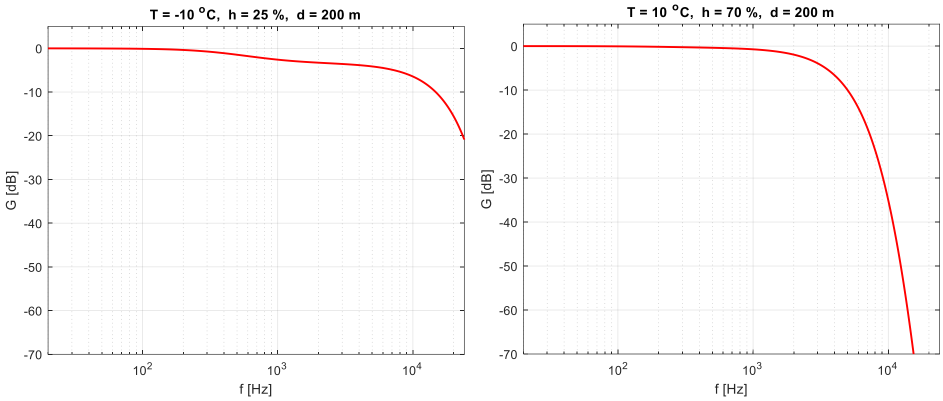

Figure 3. Frequency characteristics of the filter resulting from sound attenuation by the atmosphere for a source-microphone distance of 200 m

The impact of noise on STIPA can be included in the model in two ways. The first way is to take it into account computationally, in accordance with the algorithm specified in IEC 60268-16 standard. Taking into account this type of noise requires including in the model its sound pressure levels in the 1/1-octave bands. The second method is to sum the analysed measurement signal with real noise saved in the file and properly calibrated as to level.

It is assumed that in the model it will be possible to study the influence of individual transmission channel elements independently. An additional functionality of the model will be the ability to process any sound files that can be used, for example, for subjective research.

3. Validation of the STIPA generator and analyser

The IEC 60268-16 standard describes the algorithm for generating and analysing the STIPA signal, however, the requirements for the filters used in the generator and the analyser are usually described in terms of minimum requirements. For the generator, ½ octave filters with a stop-band attenuation of at least 60 dB/octave and the pass-band ripples not exceeding 0.3 dB are required. As an example, the standard proposes 12th order 0.2 dB ripple type I Chebyshev. In the analyser, the standard requires the use of class 0 or 1 filters in accordance with IEC 61260 standard with phase characteristics as close to linear as possible. The analyser should also use a low pass filter to obtain the signal envelope. The standard states only that its cut-off frequency should be approximately 100 Hz.

In the case of the generator, it is possible to use advanced filters requiring high computing power due to the fact that signal preparation is usually done only once (saving to a file or a carrier, etc.). In the case of the analyser, one should adapt to obtaining filters similar to those implemented in sound level metres. It is therefore important to optimise the computing power necessary for their implementation. The selection of optimal filters is made more difficult by the fact that preliminary analyses have shown that different types of analyser filters are more suitable for different generator filters. Furthermore, due to the fact that the STIPA signal is a noise signal, the results obtained are variable in time. This variation also depends on the filters used in the analyser and generator. Finally, it was decided to perform a comparative analysis of STIPA results obtained using a generator realised with filters of adjacent band attenuation from 36 to 96 dB in 12 dB steps and acceptable irregularities in the 0.1, 0.2, and 0.3 dB pass-band. Type 1 Butterworth and Chebyshev filters were checked. In the case of the analyser, Butterworth filters of the orders 10, 12, 14, 16 and 18 were tested. As a smoothing filter, a Butterworth filters with a cut-off frequency of 100 Hz of orders of 1, 2 and 3 as well as 95 Hz cut-off frequency of order 2 were used. For each type of generator filter, 10 files with STIPA signal were generated with a length enabling to perform analysis for 15 seconds. Each of 720 combinations of filter types was tested. Including 10 replications, it gave 7200 results. The optimal configuration was chosen based on the following criteria:

1) the highest maximum of STIPA value (as close as possible to 1,000),
2) the highest minimum of STIPA value,
3) the highest average of STIPA value,
4) the lowest standard deviation of STIPA.

For a computer model, ensuring the generator's repeatability is not necessary because the selected excitation can be saved on a file that can be used later for analysis instead of the generator. Repeatability was also included in the analyses due to the fact that on the basis of the model, STIPA method was also planned to be implement on the platform of the sound level meter.

Out of 720 filter configurations analysed and 10 replications:

1) 81 combinations failed to obtain a STIPA value of 1.00,
2) for 226 combinations, the minimum STIPA value was lower than 0.97,
3) the average STIPA value for 293 combinations was less lower 0.99.

For 12 out of 720 tested filter configurations there were obtained:

1) the maximum STIPA value for 10 replications 1,000,
2) the minimum STIPA value for 10 replications higher than 0.97,
3) the average STIPA value for 10 replications higher than 0.99,
4) STIPA standard deviation for 10 replications lower than 0.007.

Of these 12 cases, the smallest order combination of filters was used in the model.

The validation of the computational part of the STI algorithm, i.e. determining the single-number final value based on the modulation transfer function, was performed using the data from the example in the IEC 60268-16 standard [6]. After simplifying the STI algorithm to STIPA, validation was based on comparing the results with the results of two measuring systems existing on the market.

4. Impact of sound attenuation by the atmosphere on STIPA

The impact of sound attenuation by the atmosphere on STIPA will be discussed for the sound system emitting sound from the central cluster, not introducing linear and nonlinear distortions, and ensuring an even distribution of sound levels. The propagation concerns a free field. The analyses were carried out for four combinations of temperature

T and relative humidity *h*. The expected effect of sound attenuation by the atmosphere is a decrease in sound pressure levels of the sound pressure level signal in the high frequency range. The effect of such a change in the signal spectrum is the increasing influence of the masking effect on the STIPA result and the decreasing signal-to-noise ratio. Figure 4 presents changes in STI as a function of the source-microphone distance *d*, assuming that the noise has an insignificant impact, and the equivalent continuous sound level of the L_{Aeq} measurement signal is 60 dB. For a signal of this level, the effect of masking is minimal, therefore changes in STI begin to be noticeable at distances above 60 m in the case of the combination of *T* and *h* providing the highest attenuation in high frequencies. The lowest of the analysed attenuations occurs for T = 10ºC and $h = 25\%$. For distances *d* up to 200 m such attenuation does not introduce any change to STIPA.

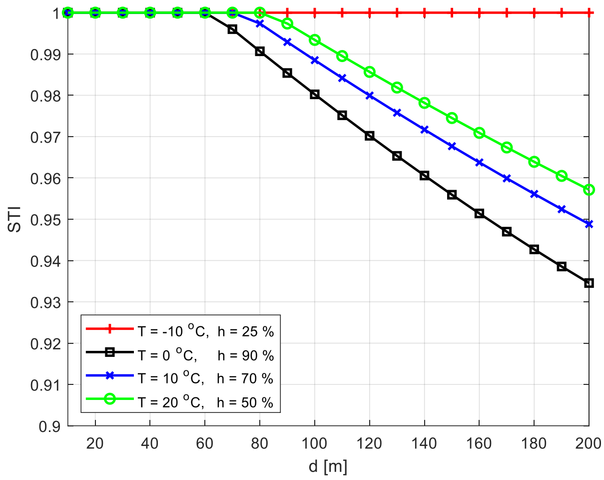

Figure 4. Impact of sound attenuation by the atmosphere on STIPA-signal level L_{Aeq} = 60 dB, no noise

In practice, sound systems are rarely designed to work with sound levels close to 60 dB. Most often, the required STIPA value and the level of noise determine the sound level of the signal. In the case of a sound propagation environment with a small reverberation time or for a free field, to obtain an STI in the range of 0.50 - 0.60 the sound level of the signal is chosen to be about 10 dB higher than the noise level. Figure 5 shows the noise results with the sound level L_{Aeq} = 70 dB and the "transport noise" spectrum [4, 9] and the signal with the level L_{Aeq} = 70 dB. In this case, the effect of sound attenuation for long distances is already significant and for *d* = 200 m STI for the value in the source proximity decreases by 0.08 - 0.13.

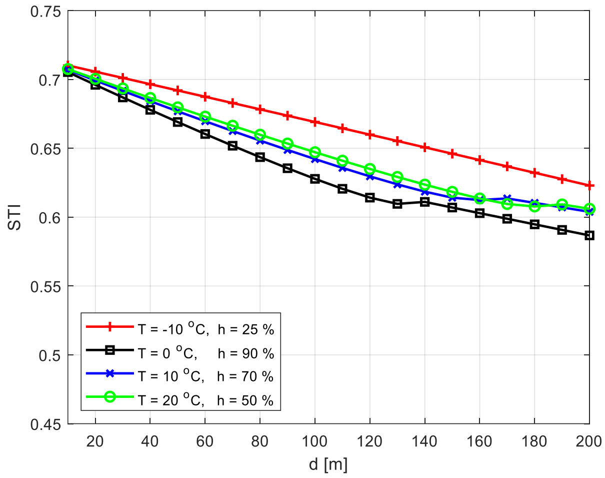

Figure 5. Impact of sound attenuation by the atmosphere on STIPA - signal level L_{Aeq} = 80 dB, noise level L_{Aeq} = 70 dB ("transport noise" spectrum)

Figure 6 shows a typical situation for a sound system in an environment with a high level of noise. In this case, the signal-to-noise ratio is fixed at about 6 dB. Compared to the previous case, the STI decrease is similar or slightly smaller, but the STI values obtained for large distances from the source are close to the minima required for many systems (0.45-0.50). It should be remembered that in reality the STI values obtained will be smaller due to factors that were not included in the analyses, i.e. linear distortions and reverberation.

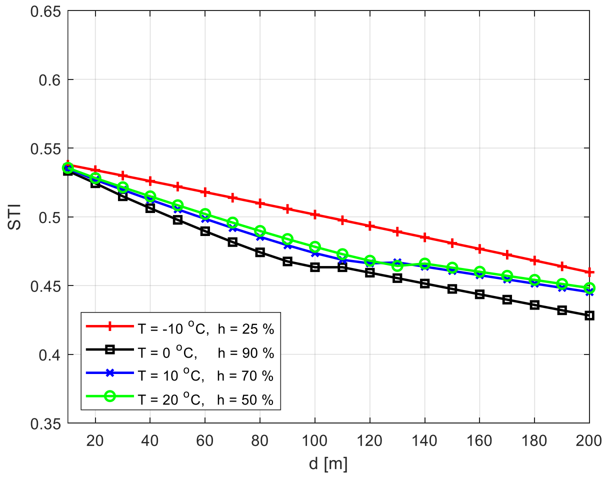

Figure 6. Impact of sound attenuation by the atmosphere on STIPA - signal level L_{Aeq} = 100 dB, noise level L_{Aeq} = 94 dB ("transport noise" spectrum)

5. Conclusions

The paper presents the idea of a computer model enabling the determination of the speech transmission index using the direct STIPA method. The advantage of this model is primarily the ability to take into account the presence of dynamics processors in the electroacoustic chain.

The analyses show that the selection of generator and analyser filters has a significant impact on the final result of the direct STIPA measurement method. Despite of the use of filters meeting the requirements of the standard IEC 60268-16, it is possible to obtain a dispersion of results in a series of 10 replications that are higher than 0.03 uncertainty of STIPA measurement.

The impact of sound attenuation by the air on STIPA values can be significant in the case of long distance sound transmission. It is particularly important for signals with a high sound level and a relatively small signal-to-noise ratio.

References

1. S. Brachmanski, *Estimation of logatom intelligibility with the STI method for polish speech transmitted via communication channels*, Archives of Acoustics, **29**(4) (2004) 555 – 562.
2. S. Brachmanski, *Experimental comparison between speech transmission index (STI) and mean opinion scores (MOS) in rooms*, Archives of Acoustics, **31**(4) (2006) 171 – 176.
3. S. Brachmański, *Wybrane zagadnienia oceny jakości transmisji sygnału mowy*, Oficyna Wydawnicza Politechniki Wrocławskiej, Wrocław 2015.
4. P. Dziechciński, M. Hadyk, *Projektowanie systemów rozgłoszeniowych na otwartych peronach kolejowych*, Postępy akustyki 2017, red. D. Bismor, Gliwice: Polskie Towarzystwo Akustyczne Oddział Górnośląski, (2017) 197 – 208.
5. T. Houtgast, H. J. M. Steeneken, *The Modulation Transfer Function in rooms acoustics as a predictor of speech intelligibility*, Acustica, **28**(1) (1973) 66 – 73.
6. IEC 60268-16:2011 - *Sound system equipment - Part 16: Objective rating of speech intelligibility by speech transmission index*.
7. ISO 9613-1:1993, *Acoustics — Attenuation of sound during propagation outdoors – Part 1: Calculation of the absorption of sound by the atmosphere*.
8. ISO 9613-2:1996, *Acoustics — Attenuation of sound during propagation outdoors – Part 2: General method of calculation*.
9. ISO 717-1:2013, *Acoustics — Rating of sound insulation in buildings and of building elements – Part 1: Airborne sound insulation*.
10. M. Schroeder, *Modulation Transfer Functions: Definition and Measurement*, Acustica, **49** (1981).
11. H. J. M. Steeneken, T. Houtgast, *A physical method for measuring speech transmission quality*, J. Acoust. Soc. Am., **67**(1) (1980) 318 – 326.