inter.noise
27-30 AUGUST HONG KONG 2017

Application of the Speech Transmission Index (STI) for planning communication areas

Michael BOEHM[1]; Wolfgang PROBST[2];

[1] DataKustik GmbH, Germany

[2] DataKustik GmbH, Germany

ABSTRACT

The STI (Speech Transmission Index) quantifies the transmission of a speech signal between a speaker and a listener. It takes into account the remaining signal strength of the speech-relevant frequency bands, the detailed energetic impulse response of the room – or generally speaking the surrounding of the source and receiver position – as well as the reduced modulation depth due to background noise. Therefore it is an excellent parameter to identify problems occurring in open plan offices, restaurants or any other communication area and assess alternative planning scenarios. The workflow of such a planning is demonstrated using a restaurant or an open plan office as an example. The general idea is to use the persons as sources for the background noise as well as the desired speech signal. Although standards often set upper limits for the reverberation time, taking into account the STI would be a major improvement for acoustic planning.

Keywords: STI, acoustic planning, simulation   I-INCE Classification of Subjects Number(s): 76

1. INTRODUCTION

Existing standards about room acoustic requirements due to disturbance or intelligibility of communication usually set upper limits for the reverberation time. Nevertheless a stronger integration of the speech intelligibility - which is also important for extra-auditory effects - would mean a major improvement of the acoustical planning of any communication area.

The Speech Transmission Index (STI) is described in IEC 60268-16 (1) and quantifies the transmission of the speech signal between a speaker and a listener, which makes it an important parameter for the planning of restaurants, open plan offices and communication areas. As an example a concept for the application of the STI during the planning stage of a restaurant is introduced in this paper. The basic idea is to use the persons in the room as sources for both background noise as well as speech signal. The results presented in this paper were obtained using the simulation software CadnaR (2).

2. Speech source

In acoustic terms a speaker is a sound source with a certain sound power level defining its emission. It makes sense to relate this sound power level to the speech effort as it is expressed figure 1.

If we assume a "normal relaxed" talking we see that an Aweighted sound power level of 65 dB can be assumed and this again can be resolved into the seven levels in octave bands shown in the figure 2. For sound prediction with speech sources seven calculations for the octave frequency bands from 125 Hz up to 8000 Hz have to be performed generally for each single source.

[1] Michael.Boehm@datakustik.com
[2] Wolfgang.Probst@ datakustik.com

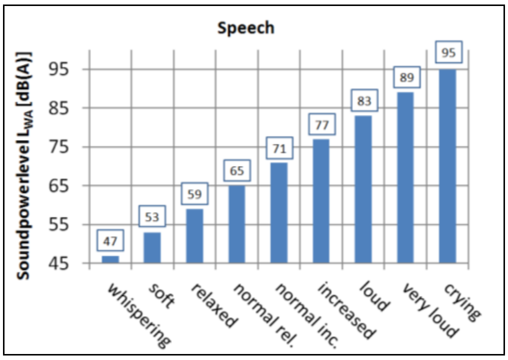

Figure 1 – Sound power level of a speaking person for different speech efforts

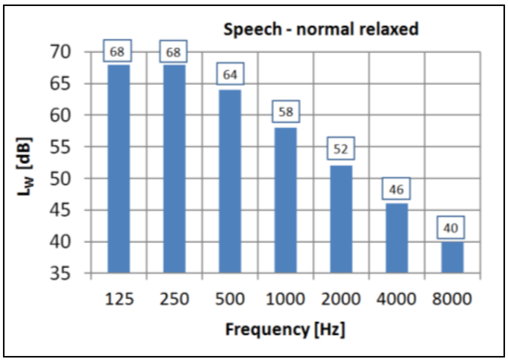

Figure 2 – Octave band levels of a "normal relaxed" speech with a total sound power level of 65 dB(A)

3. Sound propagation, detection and STI target values

In simulations applying the SERT-method (Stochastic Energy Ray Tracing) the propagation path of each particle or ray is determined up to a certain length - taking into account the speed of sound we then know the time shift between radiation of a sound-particle and each impact in a detector volume. Even if the calculations are performed sequentially according to the computer-technique applied, we can assume that they have been radiated at the same time and sort them in classes of the time needed till impact in the regarded detector volume. The result is - for each detector volume separate - the energy related impulse response as shown in figure 3.

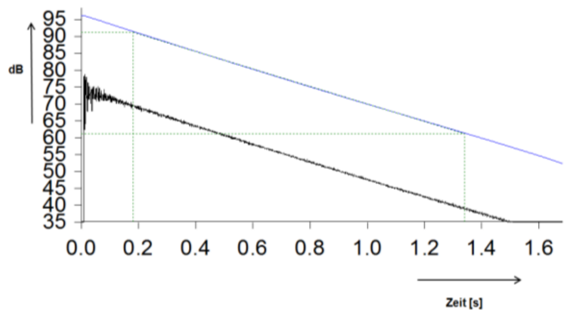

Figure 3 - Energy related impulse response at the listener position

The intelligibility of speech depends on the modulation of the sound signal due to the forming of syllables, words and sentences. The time shifted impact of the sound emitted at the same time represented by the energetic impulse response "smears" the speech signal and reduces the modulation depth - thus reducing its intelligibility. Therefore the speech signal emitted by the speaker is not only reduced in its intensity or level when it arrives at the listener, but also the modulation depth is decreased. This reduction in modulation depth can be calculated and qualified with the strategies of IEC 60268-16. The modulation depth of a speech signal is not only influenced by the acoustic response of the environment, but also by existing background noise from other sources. If the sound pressure level of this background is known for the relevant 7 frequency bands, the resulting further reduction of the STI can be determined with the procedure of IEC 60268-16.

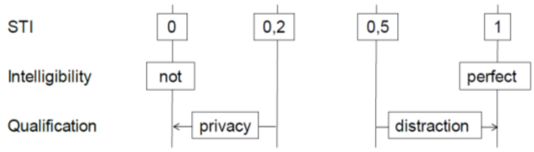

Figure 4 – Intelligibility and its qualification in dependence of the STI value

In all applications where the intelligibility of speech sounds may be important the qualification can be oriented at the STI value determined between a speaker and a listener position as shown in figure 4. At distances where the STI is around or below 0.2, intelligibility is not given and privacy can be assumed. This should be checked between areas where confidential talks shall be possible and all other workplaces or other positions of persons. If the STI is about 0.5 or larger then intelligibility is probable. This may cause distraction and has to be checked between different workplaces. If relaxed communication shall be ensured then values of the STI at or above 0.5 are the target.

4. STI for planning of a restaurant

The planning procedure in a simple case is demonstrated with the model shown in figure 5.

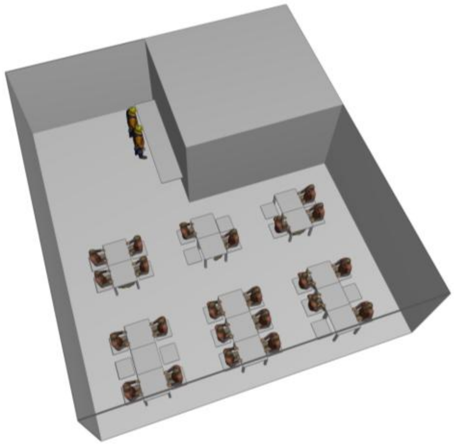

Figure 5 – Model of a restaurant

An acceptable acoustical planning should ensure that a conversation with normal relaxed speaking effort is possible. This means an STI of at least 0.5 for the conversation between any pair of persons at each table. Obviously the (background) noise level inside the restaurant is essential for the intelligibility of any conversation in the restaurant. In case of a full restaurant the speakers at all other tables are sources for the background noise at a specific table.

In the concept introduced in this paper the background noise at the listener position at one table is therefore determined by the speakers at the other tables. Provided that the model is complete and includes all absorptions, screens and acoustically relevant furniture and fittings, the assessment for each table is performed in two steps.

First the background noise at the listener position of the table under test is determined by the speakers at the others tables. This is done by taking into account one source at each table with an emission due to normal relaxed speech. This is the recommended procedure generally and independent from the number of persons at a table in restaurants - one person is speaking while the others are listening. The SERT-calculation is applied and the level spectrum at the listener position at the table under test is calculated. The setup of this calculation is displayed in figure 6: the table under test is marked in blue, the speakers at the other tables are marked in red.

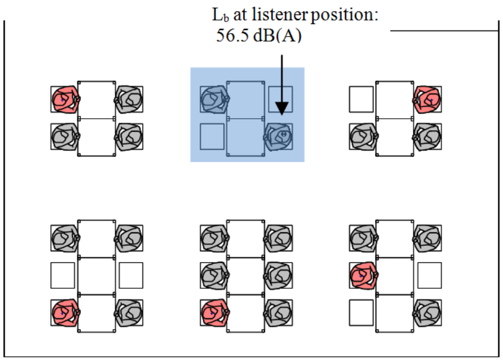

Figure 6 – Calculation of the background noise (L_b) for the conversation at the table marked in blue: The speakers at the other tables (marked in red) determine the background noise for the STI calculation at the table marked in blue

In this case the speakers at the other tables cause a background noise of 56.5 dB(A) at the listener position of the table marked in blue.

As a second step the frequency spectrum of the background noise (determined in step one) is now taken as background noise for the STI calculation with only the person speaking on the left side of the table under test. As shown in figure 7, the speech signal level is 53.2 dB(A) at the listener and the STI is 0.37 for the conversation at this table.

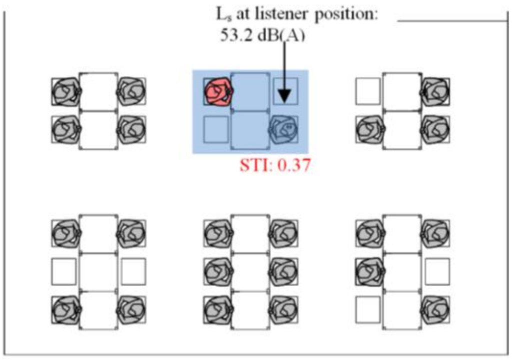

Figure 7– Calculation setup for the speech signal level (L_s) at the listener position and the STI for the conversation at the table marked in blue. The speaker is marked in red.

The two calculations (calculation of background noise and STI calculation) are performed for each of the 6 tables separately. The resulting STI values for the conversations at each table are displayed in figure 8.

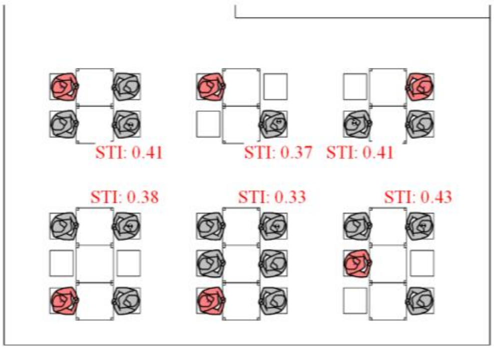

Figure 8 – STI values for the conversations at each table. The speaker at each table is marked in red, the listener is marked with a small receiver symbol

The STI values show that the intelligibility is not good enough for a conversation with normal relaxed speaking effort. At this point the speaking person will increase the speaking effort to help the listener. This would increase the signal to noise ratio and the resulting STI – but this only works if the other speakers do not speak up themselves. Since the other speakers will also increase their speaking effort the signal to noise ratio and the STI will effectively remain as bad as before. This well known Lombard effect can only be avoided by improving the acoustic fittings and layout in the room.

Such an improvement is indicated in figure 9. Absorbing barriers are inserted between the tables to create a damped individual space at each table. Additionally a baffle ceiling is installed.

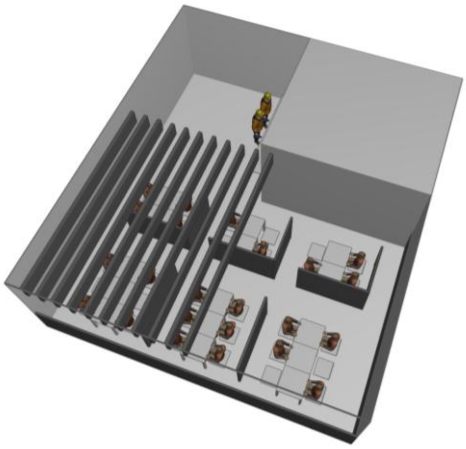

Figure 9 – Model of a restaurant with acoustic optimization

The calculation steps are repeated. For the first table under test the background noise is determined by the speakers at the other tables as shown in figure 10. At the listener position a background noise level of 46.4 dB(A) results from the speakers at the other tables.

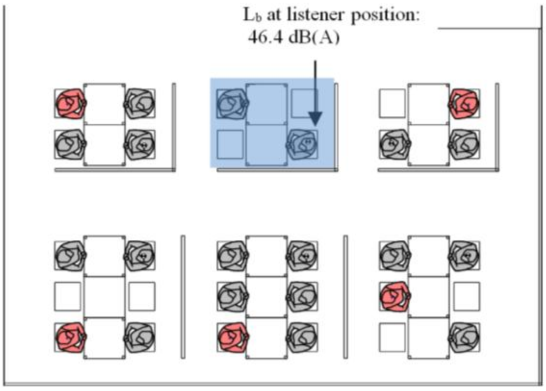

Figure 10 – Calculation of the background noise (L_b) with acoustic optimizations for the conversation at the table marked in blue: The speakers at the other tables (marked in red) determine the background noise for the STI calculation at the table marked in blue

In the second step the STI for the conversation is calculated. The resulting STI values after performing the background noise and STI calculation for each table are displayed in figure 11. The STI values indicate that with the acoustic fittings a conversation with normal relaxed speaking effort is possible at each table.

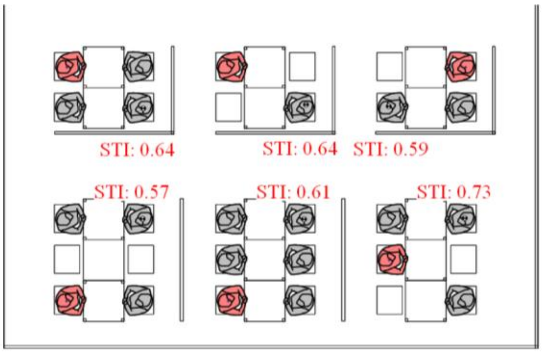

Figure 11 – STI values for the conversations at each table after acoustic optimization.

5. STI for planning of open plan offices

This concept for using the STI during the planning stage can also be applied for open plan offices. The idea is to define areas or zones inside the open plan office and then meet the requirements for the intelligibility inside and between the zones. For example zones can be defined by workgoups, i.e. persons working in the same team on the same project. In this case the intelligibility between the team members in the same zone should be good enough for conversations (STI > 0.5) but the persons outside the zone should not be disturbed, therefore the STI should be lower than 0.5 or ideally lower than 0.2. For the background noise calculation, the speakers outside the zone of investigation are considered as noise sources.

In general the zoning itself and number of speakers considered for the background noise calculation obviously depend on the designated use of the open plan office. Call centers surely have to be handled differently than engineering offices.

6. Conclusion

In this paper it was demonstrated that simulation techniques offer excellent means to assess the acoustic relations between subareas with different usage and requirements. Especially the concept of the STI according to IEC 60268-16 includes the influence of the existing background noise and of the energetic impulse response independent of the existence of a diffuse sound field - therefore all acoustic requirements based on the intelligibility of speech can be investigated. Integrating the speech intelligibility in the planning process vastly improves it and allows assessments independent of diffuse sound fields.

REFERENCES

1. IEC 60268-16: Objective rating of speech intelligibility by speech transmission index (2011), Beuth Verlag
2. CadnaR, URL: http://www.datakustik.de/