SIMPLE APPLICATION OF STI-METHOD IN PREDICTING SPEECH TRANSMISSION IN CLASSROOMS

Jukka Keränen, Petra Larm, Valtteri Hongisto

Finnish Institute of Occupational Health
Laboratory of Ventilation and Acoustics
Lemminkäisenkatu 14-18 B, 20520 Turku, Finland
`jukka.keranen@ttl.fi`

1. INTRODUCTION

Adequate speech intelligibility should be the primary goal in the acoustical design of classrooms. Typical design parameters are reverberation time and background noise level. Unfortunately, these parameters do not directly tell much about the speech intelligibility in the room. One better parameter for describing speech intelligibility is Speech Transmission Index, STI, described in, e.g. IEC 268-16. However, for predicting the STI in practical rooms, the designer should be able to predict early decay time, T_{10}, at listening points. This may require the use of acoustical modelling, because T_{10} may strongly deviate from the reverberation time, T_{60}, in practical non-diffuse rooms. In this paper a simple application of STI using predicted T_{10}, background noise level and speech sound level is presented.

2. METHODS

When the sound power level of the speaker's voice is known conforming e.g. ANSI S3.5, the speech sound level, L_S, at listening point can be predicted using any acoustical modelling method. In the simplest case, L_S includes only the reverberant speech. At short distances, direct sound improves STI and the directivity of the sound source has to be considered. If the sound power level of the noise source(s), e.g. ventilation equipment, is known, the background noise level, L_N, at the listening point can also be predicted. However, the background noise level predictions are uncertain, because the sound power level of the noise source(s) is usually not exactly known and the background noise usually varies. However, there are certain requirements for the background noise produced by the HVAC-systems, which HVAC designers should obey, so that the noise level L_N at listening point may be supposed to be equal with them. The uncertainty of the noise level, L_N, is not so important if the speech level, L_S, is at least 15 dB higher than the noise level. T_{10} can be predicted using acoustical modelling e.g. ray-tracing or image-source-method. L_S, L_N and T_{10} are used in STI calculation at the listening point.

2.1. Modelling Speech and Noise level

A simple small class room (6.15 x 6.45 x 3.0 m^3) is modelled using Odeon 3.1 software. Four different acoustical designs are examined. The total area of absorbent is the same (80 pieces of 0.6 x 0.6 m^2 boards of 50 mm mineral wool with surface treatment) in all of the designs. The floor, the ceiling and the side walls are supposed to be 10 % absorbent. The front and end walls are smooth painted concrete (sound absorption coefficients 0.02,…, 0.05). The room is empty and no other details are included in the models. The models are presented in Figure 1. The absorbents are shown in grey.

The modelled parameters are T_{10}, L_S, and L_N at listening points 1-4 (Figure 2). The parameters are modelled at octave bands of 125, …, 8000 Hz. First, the sound pressure level produced by a normally speaking person at the listening points is modelled. The directivity pattern of the speaking person is pre-defined in the software. The sound power level of the speaking person is presented in Table 1. T_{10} is obtained at every listening point among the other modelling results. The value of T_{10} is calculated in the software (result parameter EDT) using the modelled point response, which is calculated using hybrid of image-source and ray-tracing methods. The first two reflections are modelled using image-source method (Transition order 2). The later reflections are modelled using ray-tracing. The number of rays is 2000.

Secondly, the sound pressure level produced by the noise source alone at the listening points is modelled. The noise source is an omnidirectional point source close to the ceiling. The noise source is modelled with two sound levels: "normal" and "+ 15 dB". The sound power level of the noise sources are presented in Table 1.

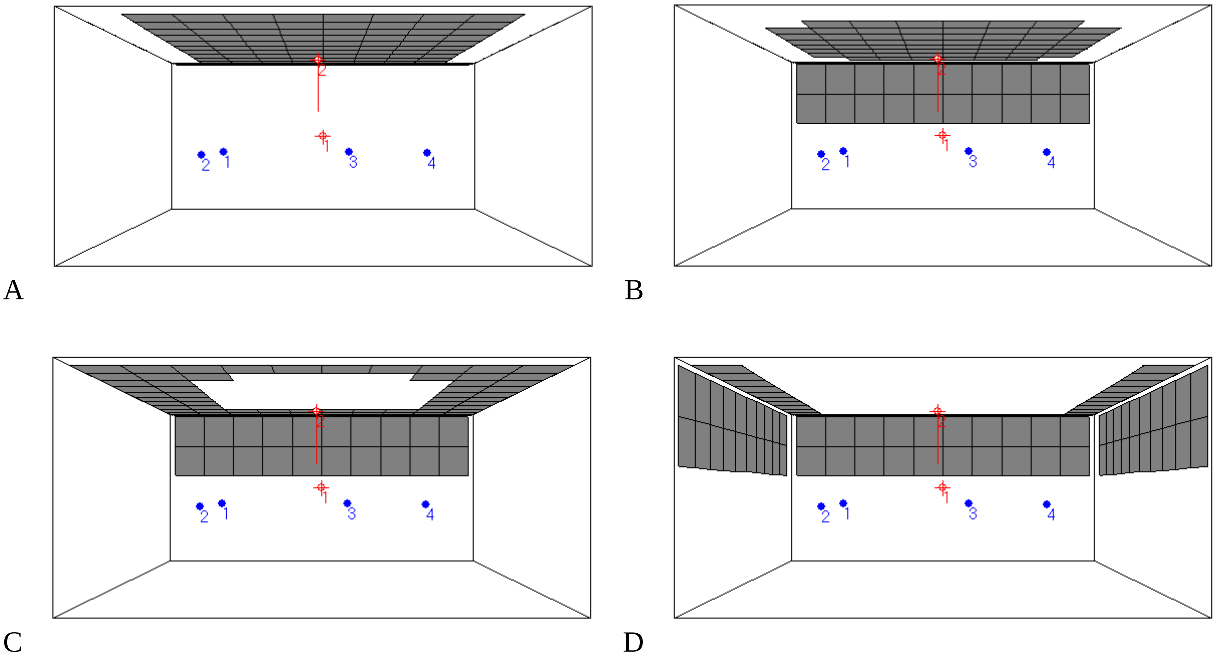


Figure 1. *The models of the 4 acoustical designs of the class room (Cases A-D).*

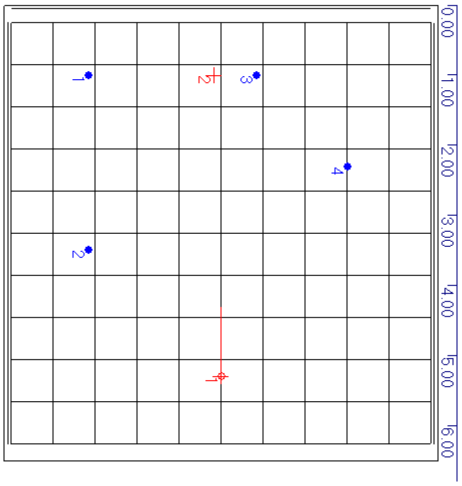


Figure 2. *The speaking person (cross 1), the noise source (cross 2) and the listening points (dots 1-4).*

Table 1. *The sound power level of the speaking person and the noise sources.*

Sound Power level	**125**	**250**	**500**	**1000**	**2000**	**4000**	**8000**
The speaking person	61.9	64.1	67.8	62.0	54.6	49.9	49.9
The noise source	50.0	45.0	40.0	35.0	30.0	25.0	20.0
The noise source +15 dB	65.0	60.0	55.0	50.0	45.0	40.0	35.0

2.2. Simple prediction of speech and background noise level

In small rooms like this example, the use of acoustical modelling may be quite expensive. That is why an alternative method to predict the needed parameters for the STI-calculation is presented. In this case, the early decay time, T_{10}, is replaced by the reverberation time, T_{60}, which is calculated at each octave band 125,…, 8000 Hz using the Sabine's equation

$$T_{10} \approx T_{60} = \frac{0.16V}{\alpha S} \tag{1}$$

where V is the room volume, S total room surface area and α total absorption coefficient at the octave band.
The sound pressure level at the listening point is calculated by equation

$$L_p = L_w + 10\log_{10}\left(\frac{Q}{4\pi r^2} + 4\frac{1-\alpha}{\alpha S}\right) \tag{2}$$

where L_w is the sound power level of the sound source, r is the distance between the listening point and the sound source and Q is the directivity factor of the sound source. This equation is used for both the speech level and the noise level. In this study, the directivity factor of the speaker is $Q = 3$ and directivity factor of the noise source $Q = 1$.

2.3. STI-calculations

The predicted speech and noise sound levels are used to calculate STI at the listening points. STI is calculated using modulation reduction factor

$$m(F) = \frac{1}{\sqrt{1+\left[2\pi F\frac{T}{13.8}\right]^2}} \cdot \frac{1}{1+10^{-L_{SN}/10}} \tag{3}$$

where T is the early decay time, T_{10}, and $L_{SN}=L_S-L_N$. The modulation frequencies F are 0.63, 0.80, 1.00, 1.25, 1.60, 2.00, 2.50, 3.15, 4.00, 5.00, 6.30, 8.00, 10.00 and 12.50. The modulation reduction factor is calculated at each octave band 125, …, 8000 Hz.
The m values are converted into an apparent signal-to-noise ratio

$$(S/N)_{app} = 10\log\frac{m}{1-m}\,\text{dB} \tag{4}$$

The values above +15 dB are replaced by +15 dB and similarly below -15 dB by -15 dB. After that an arithmetic average of $(S/N)_{app}$ is calculated at each octave band

$$(S/N)_{app,k} = \frac{1}{14}\sum_{i=1}^{14}(S/N)_{app,i} \tag{5}$$

where k is the index of the octave band 125,…, 8000 Hz.

The weighted average of the octave band $(S/N)_{\mathrm{app,k}}$ values is determined by

$$\overline{(S/N)}_{app} = \sum_{k=1}^{7} w_k\,(S/N)_{app,k} \tag{6}$$

where w_k is 0.13, 0.14, 0.11, 0.12, 0.19, 0.17, 0.14.
Finally, STI is determined by equation

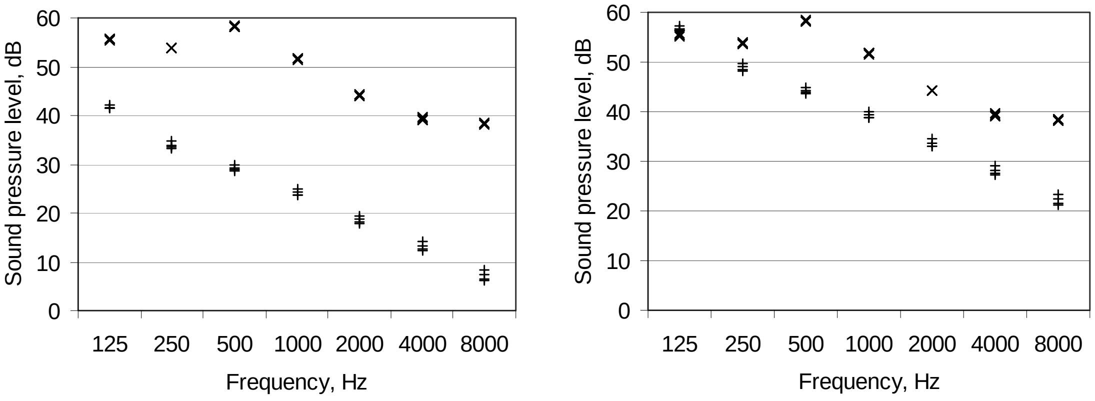
$$STI = \frac{\overline{(S/N)}_{app} + 15}{30} \tag{7}$$

2.4. RASTI-measurements

RASTI is a simplified version of STI-method and the comparability of the results of STI and RASTI are presented elsewhere [1]. RASTI measurements were measured before this study in a real class room. The measurement situation is modelled in the case A.

3. RESULTS

The modelled speech level L_S and background noise level L_N at the listening points in the cases A and E are presented in Figure 3. The speech and noise levels are very similar in the rest of the cases.
The modelled L_{SN} and T_{10} are presented in Table 2. The reverberation times T_{60} calculated using equation 1 are also presented in Table 2. The cases A to D are presented in Figure 1. The acoustical design in the case E is the same as in the case A and in the case F the same as in the case C. In the cases E and F the sound power level of the noise source is 15 dB higher than in the cases A and C.
The calculated STI values are presented in Table 3. In addition to the modelled cases A-F the STI calculation results based on Sabine's reverberation time, T_{60}, and simply predicted Speech and Noise levels at listening points are presented in Table 3.
The measured RASTI values in the case A are presented in Table 4.



A E

Figure 3. *The modelled speech level (x) and the modelled noise level (+) at the listening points in the cases A and E.*

Table 2. *The modelled* L_{SN} *and* T_{10} *and Sabine's reverberation time* T_{60} *in the cases A-F.*

Case A		**125**	**250**	**500**	**1000**	**2000**	**4000**	**8000**
1	L_{SN}	13.7	19.7	28.8	27.2	25.5	25.9	30.8
	T_{10}, s	1.24	0.53	0.47	0.49	0.38	0.39	0.29
2	L_{SN}	14.4	20.6	29.8	28.1	26.1	26.9	32.2
	T_{10}, s	1.34	0.56	0.61	0.62	0.52	0.46	0.35
3	L_{SN}	13.3	19.1	28.2	26.8	24.8	25.5	30.3
	T_{10}, s	1.36	0.57	0.62	0.62	0.52	0.45	0.35
4	L_{SN}	14.1	20.2	29.3	27.9	26.0	26.5	31.8
	T_{10}, s	1.37	0.56	0.61	0.62	0.50	0.45	0.36
	T_{60}, s	1.31	0.67	0.72	0.71	0.61	0.54	0.42

Case B		**125**	**250**	**500**	**1000**	**2000**	**4000**	**8000**
1	L_{SN}	13.3	20.4	29.3	27.8	26.1	26.5	31.3
	T_{10}, s	1.25	0.49	0.54	0.56	0.43	0.39	0.31
2	L_{SN}	13.9	21.8	30.6	29.0	27.3	27.7	33.1
	T_{10}, s	1.27	0.53	0.58	0.60	0.47	0.42	0.35
3	L_{SN}	12.9	19.9	28.8	27.4	25.5	26.0	30.9
	T_{10}, s	1.32	0.55	0.60	0.61	0.49	0.42	0.35
4	L_{SN}	13.6	21.0	29.9	28.5	26.8	27.3	32.4
	T_{10}, s	1.32	0.54	0.60	0.61	0.47	0.42	0.35
	T_{60}, s	1.26	0.64	0.69	0.69	0.59	0.52	0.41

Case C		**125**	**250**	**500**	**1000**	**2000**	**4000**	**8000**
1	L_{SN}	13.8	20.8	29.7	28.5	26.7	27.0	31.9
	T_{10}, s	1.22	0.49	0.56	0.57	0.45	0.41	0.31
2	L_{SN}	14.4	22.2	31.0	29.7	27.9	28.5	33.7
	T_{10}, s	1.24	0.56	0.61	0.62	0.51	0.47	0.36
3	L_{SN}	13.6	20.5	29.3	28.4	26.3	26.6	31.4
	T_{10}, s	1.27	0.57	0.63	0.62	0.51	0.45	0.35
4	L_{SN}	14.1	21.6	30.4	29.3	27.6	28.1	33.1
	T_{10}, s	1.27	0.56	0.62	0.62	0.50	0.45	0.36
	T_{60}, s	1.22	0.62	0.67	0.67	0.57	0.51	0.40

Case D		**125**	**250**	**500**	**1000**	**2000**	**4000**	**8000**
1	L_{SN}	11.7	19.2	28.0	26.9	25.0	25.3	30.0
	T_{10}, s	1.18	0.59	0.63	0.64	0.52	0.46	0.35
2	L_{SN}	12.3	20.4	29.1	28.0	26.2	26.8	31.9
	T_{10}, s	1.21	0.64	0.67	0.68	0.57	0.51	0.39
3	L_{SN}	11.4	18.6	27.5	26.7	24.5	24.8	29.3
	T_{10}, s	1.26	0.66	0.69	0.69	0.58	0.50	0.39
4	L_{SN}	12.0	19.8	28.5	27.5	25.7	26.0	31.1
	T_{10}, s	1.26	0.63	0.67	0.68	0.57	0.50	0.39
	T_{60}, s	1.12	0.58	0.62	0.62	0.53	0.47	0.38

Case E		**125**	**250**	**500**	**1000**	**2000**	**4000**	**8000**
1	L_{SN}	-1.3	4.7	13.8	12.2	10.5	10.9	15.8
	T_{10}, s	1.29	0.53	0.58	0.59	0.47	0.42	0.33
2	L_{SN}	-0.6	5.6	14.8	13.1	11.1	11.9	17.2
	T_{10}, s	1.34	0.56	0.61	0.62	0.52	0.46	0.35
3	L_{SN}	-1.7	4.1	13.2	11.8	9.8	10.5	15.3
	T_{10}, s	1.36	0.57	0.62	0.62	0.52	0.45	0.35
4	L_{SN}	-0.9	5.2	14.3	12.9	11.0	11.5	16.8
	T_{10}, s	1.37	0.56	0.61	0.62	0.50	0.45	0.36
	T_{60}, s	1.31	0.67	0.72	0.71	0.61	0.54	0.42

Case F		**125**	**250**	**500**	**1000**	**2000**	**4000**	**8000**
1	L_{SN}	-1.2	5.8	14.7	13.5	11.7	12.0	16.9
	T_{10}, s	1.22	0.49	0.56	0.57	0.45	0.41	0.31
2	L_{SN}	-0.6	7.2	16.0	14.7	12.9	13.5	18.7
	T_{10}, s	1.24	0.56	0.61	0.62	0.51	0.47	0.36
3	L_{SN}	-1.4	5.5	14.3	13.4	11.3	11.6	16.4
	T_{10}, s	1.27	0.57	0.63	0.62	0.51	0.45	0.35
4	L_{SN}	-0.9	6.6	15.4	14.3	12.6	13.1	18.1
	T_{10}, s	1.27	0.56	0.62	0.62	0.50	0.45	0.36
	T_{60}, s	1.22	0.62	0.67	0.67	0.57	0.51	0.40

Table 3. *The predicted STI values using acoustical models (Odeon) and simple predictions.*

	Listener1		**Listener2**		**Listener3**		**Listener4**		**Average**	
Case	simple	Odeon	simple	Odeon	simple	Odeon	simple	Odeon	simple	Odeon
A	0.67	0.74	0.67	0.70	0.67	0.70	0.67	0.70	0.67	0.71
B	0.68	0.73	0.68	0.71	0.68	0.71	0.68	0.71	0.68	0.72
C	0.68	0.72	0.68	0.70	0.68	0.70	0.68	0.70	0.68	0.71
D	0.70	0.70	0.70	0.69	0.70	0.68	0.70	0.68	0.70	0.69
E	0.56	0.61	0.57	0.62	0.56	0.60	0.56	0.61	0.56	0.61
F	0.57	0.63	0.58	0.63	0.56	0.61	0.57	0.63	0.57	0.63

Table 4. *The measured RASTI values in the case A.*

A	0.76	0.78	0.75	0.77	0.77

The differences between STI results using acoustical modelling (Odeon) or simple prediction method are quite small. The predicted STI values are lower, when reverberation time and simple sound level predictions are used, because T_{60} values are higher than T_{10} values (except in case D). Both the simple prediction method and the acoustical modelling method underestimated STI when compared to the measured RASTI values in the case A. Apparently, the models do not accurately take into account the directivity of the speaker. Unfortunately, there are no measured results of the other cases.

4. CONCLUSIONS

In small rooms STI can be adequately predicted using even the simple method. In larger and more complex rooms, where the ratio of early and late reflections is important, STI values should be predicted applying acoustical modelling methods. However, the modelling of the early decay time at listening points may be difficult e.g. in auditoria, where large area of the ceiling is often treated with absorptive material and the placing of it has much stronger influence on the early decay time than the conventional reverberation time, which depend more on late reflections. The model rooms were all empty, so that the influence of audience and obstacles is not included. However, the effect of them may be very strong especially in the rooms, where most of the reflections are prevented using effective absorption materials. Thus, more research is needed to optimize the acoustical design of rooms, where speech intelligibility is important.

5. REFERENCES

[1] Houtgast, T., Steeneken H. J. M., "RASTI: A Tool for Evaluation Auditoria ", *B&K Technical Review No.3*: 13-29, 1985.