DTU Library

# Classroom acoustics design guidelines based on the optimization of speaker conditions

**Pelegrin Garcia, David; Brunskog, Jonas**

*Published in:*
Proceedings - European Conference on Noise Control

*Publication date:*
2012

*Document Version*
Peer reviewed version

# Classroom acoustics design guidelines based on the optimization of speaker conditions

David Pelegrin-Garcia
Laboratory of Acoustics and Thermal Physics, Katholieke Universiteit Leuven, Heverlee, Belgium

Jonas Brunskog
Acoustic Technology, Department of Electrical Engineering, Technical University of Denmark, Lyngby, Denmark

**Summary**
School teachers suffer frequently from voice problems due to the high vocal load that they experience and the not-always-ideal conditions under which they have to teach. Traditionally, the purpose of the acoustic design of classrooms has been to optimize speech intelligibility. New guidelines are suggested in order to optimize the vocal comfort and the vocal load experienced by speakers. Theoretical prediction models of room-averaged speaker-oriented parameters like voice support or reverberation time derived from an oral-binaural impulse response are combined with empirical models of actual voice and noise level measurements in classrooms. Requirements of optimum vocal comfort, average A-weighted speech levels across the audience higher than 50 dB, and a physical volume higher than 6 $m^3$/student are combined to extract optimum acoustic conditions, which depend on the number of students. These conditions, which are independent on the position of the speaker, cannot be optimum for more than 50 students. For classrooms with 10 students, the reverberation time in occupied conditions shall be between 0.5 and 0.65 s, and the volume between 60 and 170 $m^3$. For classrooms with 40 students, the reverberation times shall be between 0.7 and 0.75 s and the volume between 240 and 280 $m^3$.

**PACS no. 43.55.Fw, 43.55.Hy**

## 1. Introduction

The acoustic design of school classrooms, both in terms of noise control and room acoustics, is relevant because it affects the quality of oral communication between teachers and students, which is still the most common way of teaching and learning, and has an effect on the overall performance of pupils. Excessive noise and late reverberation degrade speech intelligibility. The degradation is higher for younger students and those with impairments than for the general population. Therefore special attention is given to acoustics in classrooms. More information about the importance of classroom acoustics is found in [1, 2].

Besides affecting speech intelligibility, noise and classroom acoustics affect the voice of a speaker. The variation of voice level with noise has been described as the Lombard effect [3]. In a scenario with different talkers, the absorption of a room has an influence on the voice power level of each talker, which is known as the café effect [4]. Moreover, in situations with only one talker, this talker adjusts the voice power level according to the amplification that a room produces on his voice at his own ears [5, 6]. When a teacher speaks in a classroom, besides being heard, he wants to talk comfortably and not to overstrain his voice [7].

The present paper briefly introduces the room acoustics parameters relevant for a speaker and their connection to subjective attributes (vocal effort and vocal comfort). It also provides models of the room-averaged acoustics parameters for a speaker in terms of volume and statistical reverberation time, together with models of activity noise and voice levels. Finally, the models and different requirements are combined to derive a set of recommendations for speaker-oriented classroom acoustics design.

## 2. Room acoustics parameters for a speaker

Room acoustics parameters for a speaker are derived from an *oral-binaural room impulse response* (OBRIR), i.e. an impulse response measured at a microphone located at the end of the ear canal of a


ISSN 2226-5147
ISBN 978-80-01-05013-2

dummy head when a loudspeaker inside its mouth acts as the source. For all measurements and derivations, a head and torso simulator (HATS) from Brüel & Kjær type 4128 has been used. Two parameters are found relevant: the *voice support* and the *reverberation time mouth-to-ears*.

## 2.1. Voice support

### 2.1.1. Definition

The voice support $\mathrm{ST}_V$ is a measure of the degree of amplification of a room to the voice of a speaker at his own ears. More specifically, it is defined as the difference between the reflected sound level ($L_R$) and the airborne direct sound level ($L_D$) of the voice of a speaker, as found in an OBRIR.

$$\mathrm{ST}_V = L_R - L_D \text{ (dB)}. \quad (1)$$

The level of direct sound is calculated by windowing the first 5 ms of an OBRIR while the reflected sound is calculated from the rest of the signal (5 ms to $\infty$). A more complete description of $\mathrm{ST}_V$ is to be found in [8]. $\mathrm{ST}_V$ is also related to the measure *room gain* $G_{\mathrm{RG}}$ described by Brunskog *et al.* [5].

### 2.1.2. Relation with voice

Interaction between the voice of a speaker and $\mathrm{ST}_V$ or $G_{\mathrm{RG}}$ has been documented in a number of studies [5, 6, 9, 10, 11]. In summary, $\mathrm{ST}_V$ is related to the *vocal effort* experienced by a speaker in different rooms. A model that represents the average variations in voice power level $L_W$ as a function of $\mathrm{ST}_V$ in a teaching setting with a silent audience is

$$\Delta L_W = -13 - 0.78\mathrm{ST}_V \text{ (dB)}, \quad (2)$$

which is only valid in typical rooms within a limited range of $\mathrm{ST}_V$, approximately between -18 dB and -8 dB.

It is important to point out that the variations of $L_W$ in the presence of activity noise do not follow the model of Eq. (2), as the activity noise depends itself on $\mathrm{ST}_V$, as will be shown later.

### 2.1.3. Prediction model

A prediction model for the average $\mathrm{ST}_V$ in a room is presented in [8]. The model disregards the importance of the surroundings of the speaker in determining the actual $\mathrm{ST}_V$ at the speaker position and provides a unique value for a room, averaged across positions. The final prediction model is formulated as

$$\mathrm{ST}_V = 10\log\left[\left(\frac{cT_{60}}{\ln(10^6)V} - \frac{4}{S} + \frac{Q^*}{4\pi(2d)^2}\right)S_{\mathrm{ref}}\right] + \Delta L_{\mathrm{HRTF}} - K \text{ (dB)}, \quad (3)$$

where $c$ is the speed of sound in the air ($\approx 343$ m/s), $T_{60}$ is the statistical reverberation time, $V$ is the volume, $S$ is the total surface area, $Q^*$ is the directivity of

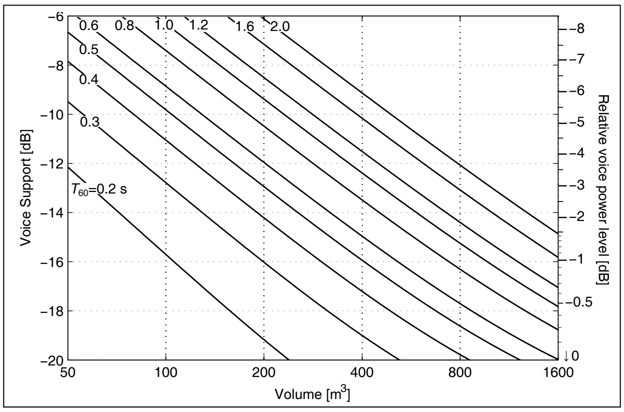


Figure 1. Average voice support versus room volume for different values of $T_{60}$, considering a flat $T_{60}$ across frequency. The axis on the right edge shows the average voice power level variations experienced by speakers.

a speaker in the downward direction, $d$ is the distance from the mouth to the floor ($=1.5$ m), $S_{\mathrm{ref}}$ is the reference area ($\approx 1\mathrm{m}^2$), $\Delta L_{\mathrm{HRTF}}$ is the magnitude of the diffuse-field head-related transfer function (HRTF, in dB), $K$ is the difference between SPL at the eardrum and the source sound power level (in dB).

This model contains the following terms:

- Diffuse-field attenuation of sound, indicated by the term $[(cT_{60})/(6V\ln 10) - 4/S]$ inside the $10\log$, which is written sometimes as $4/R$ in the context of room acoustics.
- Floor reflection, given by the term $Q^*/[4\pi(2d)^2]$ inside the $10\log$. The floor reflection is considered present in all measurements, and it is assumed that the floor is totally reflective and that the mouth and the ears are at a height of 1.5 m above the floor. All the early reflections from the walls, when averaged across positions in a room, are included in the diffuse-field attenuation term. The reflection from the ceiling is included in the diffuse-field attenuation term because it is attenuated by the typical presence of an absorbing ceiling in classrooms and because the height varies across rooms.
- Diffuse-field HRTF ($\Delta L_{\mathrm{HRTF}}$), accounting for the increase in level associated to the use of a dummy head instead of a small microphone for the measurement of the sound reflections.
- Direct sound characterization with the term $-K$.

The dependence of $\mathrm{ST}_V$ with $V$ and $T_{60}$ is illustrated in Fig. 1, considering a flat $T_{60}$ across frequency. $\mathrm{ST}_V$ decreases almost linearly with the logarithm of $V$ (except for the largest volumes at low reverberation times) and increases with $T_{60}$. The axis on the right edge shows the average voice power level variations experienced by speakers, according to Eq. (2). The values on the axis are derived from Eq. (2) for $\mathrm{ST}_V$ between -14.5 and -6.5 dB and from Eq. (12) in ref. [6] for $\mathrm{ST}_V < -14.5$ dB.

### 2.2. Reverberation time mouth-to-ears

#### 2.2.1. Definition

The reverberation time mouth-to-ears $T_{30,ME}$ is a classical measurement of reverberation time $T_{30}$ derived from an OBRIR. Thus, it is twice the difference between the times when the backwards integrated energy curve is -35 dB and -5 dB. Differently from traditional impulse responses in which the receiver is far away from the source, in an OBRIR source and receiver are located very close to each other, so the direct sound has much more energy than the reflected sound. Therefore, the $T_{30,ME}$ is very sensitive to the direct-to-reflected sound ratio. The value of $T_{30,ME}$ is always smaller than that of $T_{30}$ with distant source-receiver, and it does not represent the slope of the decay.

#### 2.2.2. Relation with voice

Laboratory experiments reported in [7], which investigated relevant subjective attributes when speaking in different rooms, found that $T_{30,ME}$ was linearly related to the sensation of reverberance. More interestingly, the general sensation of *vocal comfort* $\hat{C}$ was non-linearly related to $T_{30,ME}$. The regression model for the pooled results of $\hat{C}$ for all healthy speakers was

$$\hat{C} = -4.25\, T_{30,ME}^2 + 4.37\, T_{30,ME} - 0.81, \quad (4)$$

where the units of $\hat{C}$ are of no practical relevance (for information, 0 corresponds to the average comfort, and +1 is one standard deviation across the experimental results). The vocal comfort had a maximum at 0.51 s. When groups of subjects were considered separately, their regression models had maxima between 0.45 s and 0.55 s, which will be considered reference values for later design requirements. As an alternative approach, regression models based on energy ratios could have been used as well.

#### 2.2.3. Prediction model

Another model predicts the average $T_{30,ME}$ in a room as a function of $V$ and $T_{60}$. This model makes the same assumptions and includes the same elements as the prediction model for $\mathrm{ST}_V$, but with temporal considerations. One difference is that the prediction model for $T_{30,ME}$ does not have a closed mathematical expression and has to be calculated by means of an algorithm with the following steps:

1. Modeling of a parametric OBRIR
2. Calculation of the backward integrated energy curve
3. Searching the time instants when the backwards integrated energy curve decays -5 dB and -35 dB relative to the level at the time of arrival of the direct sound.
4. Finally, $T_{30,ME}$ is calculated as twice the difference between the two time instants of the previous step.

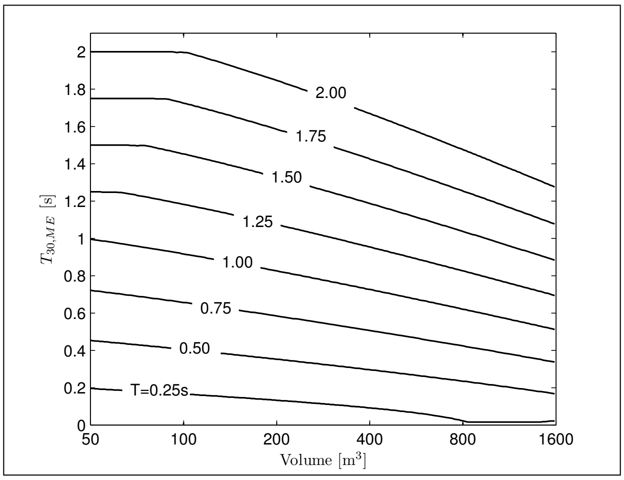


Figure 2. Average reverberation time mouth-to-ears versus volume for different values of $T_{60}$.

As in the prediction model for $\mathrm{ST}_V$, the prediction model for $T_{30,ME}$ assumes an OBRIR with the following components: direct sound, a floor reflection, and a reverberation tail. The direct sound and the floor reflection are modeled as Dirac delta functions and the reverberation tail as a decaying exponential function, which depends on $V$ and $T_{60}$ (both decay rate and total energy).

Figure 2 shows the output of the prediction model for different values of $V$ and $T_{60}$. The predicted $T_{30,ME}$ decreases with $V$ and increases with $T_{60}$. More detailed information about the prediction model can be found in [7].

### 2.3. Combined prediction models

Figure 3 shows the mutual relationship between $\mathrm{ST}_V$ and $T_{30,ME}$, for equal values of $V$ (dotted lines) and $T_{60}$ (solid lines). This chart is meaningful because the abscissa is linked to the *vocal effort*, whereas the ordinate is related to the *vocal comfort*. On the bottom axis of the figure, there is an indication of $\Delta L_W$ experienced by a speaker in the presence of low background noise levels. The values in this axis illustrate how different classroom acoustic designs affect the voice levels of teachers while the audience is silent.

## 3. A model for classroom activity noise and speech levels

Hodgson *et al.* [12] made measurements of noise levels and speech levels in university classrooms with different number of students $N$ and proposed prediction models for student activity noise levels (SA) and instructor sound power level ($L_W$).

### 3.1. Student activity noise levels

The A-weighted student-activity noise [12] (SA or $L_{\mathrm{SA}}$) is described as:

$$L_{\mathrm{SA}} = 83 + 10\log N - 34.4\log A_0 + 0.08 A_0 \text{ (dB)} \quad (5)$$

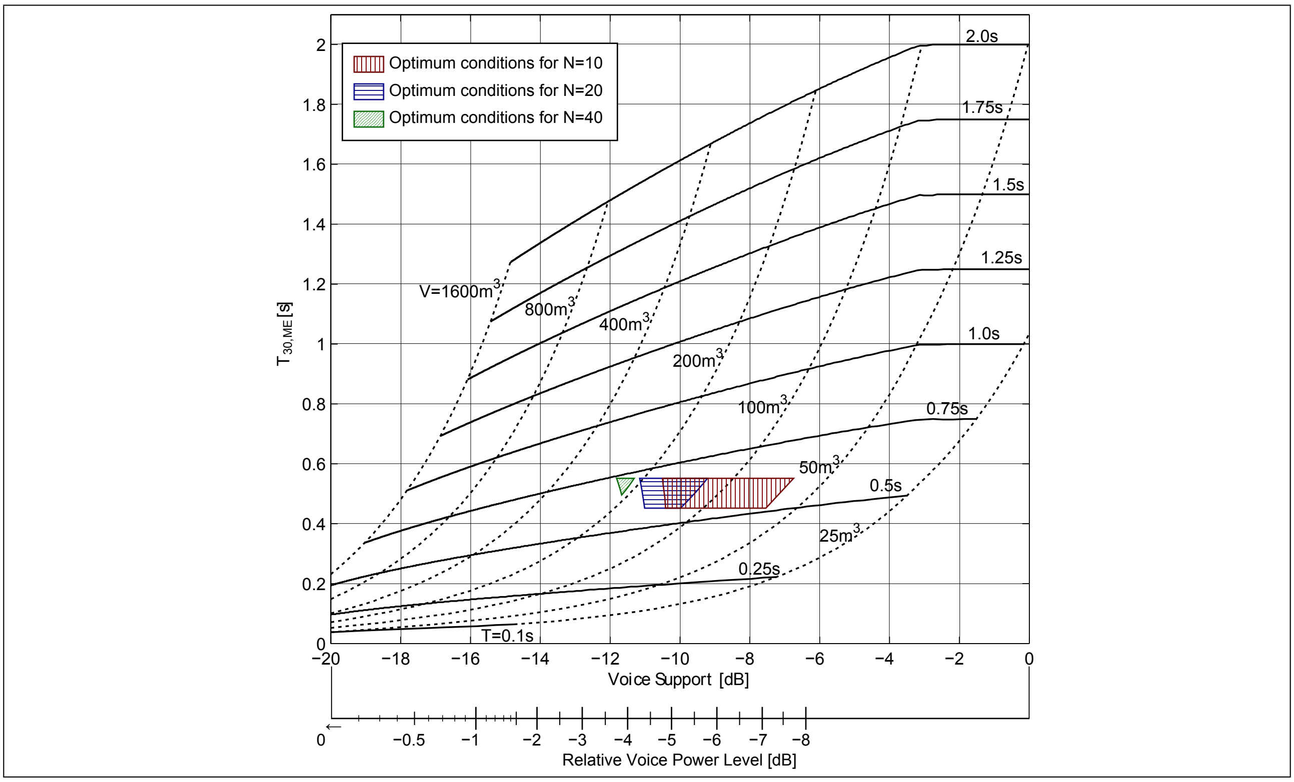


Figure 3. Reverberation time mouth-ears $T_{30,ME}$ versus voice support $\mathrm{ST}_V$ for different values of $T_{60}$ (solid lines) and $V$ (dotted lines). The optimum acoustic conditions for a speaker are indicated in hatched areas, as a function of the number of students in the classroom. All the values correspond to occupied conditions.

where $A_0$ is the total absorption area in the room, which can be calculated from Sabine's formula as $A_0 = 0.161V/T_{60}$.

The student-activity noise levels are shown in Fig. 4 as a function of $\mathrm{ST}_V$ and $T_{30,ME}$, for different $N$. The lines corresponding to equal SA are almost vertical, except for the lowest range of $T_{60}$, where Sabine's formula does not hold. The prediction model shows that, when $\mathrm{ST}_V$ increases, SA increases due to conversational feedback or café effect [4] (even though the university students were not particularly noisy). When increasing the number of students, the same value of $\mathrm{ST}_V$ results in higher SA values.

### 3.2. Instructor speech level

The A-weighted instructor speech level (SL or $L_{\mathrm{SL}}$) is derived from the A-weighted $L_W$ which, averaged for male and female instructors and according to [12], is

$$L_W = 53.5 + 0.5L_{\mathrm{SA}} + 0.016V - 9.6\log A_0 \quad \text{(dB)}. \quad (6)$$

As can be seen, $L_W$ increases with SA, due to the Lombard effect. $L_W$ depends also on the volume $V$ and the total absorption area $A_0$. As an empirical model, Eq. (6) is not related to Eq. (2).

It is assumed that, in a diffuse sound field, the SL at the most unfavorable listener position is

$$L_{\mathrm{SL}} = L_W + 10\log\left(\frac{4(1-\bar{\alpha})}{A_0}\right) \text{ (dB)}, \quad (7)$$

where $\bar{\alpha}$ is the mean absorption coefficient. The equal SL contours predicted with this model are shown in Fig. 5 for $N = 10$ and $N = 80$ students. An increase in $\mathrm{ST}_V$ results in increased SL due to increased SA and increased reverberant energy.

## 4. Derivation of optimum values for a speaker

The derivation of the optimum classroom acoustics conditions for a speaker is illustrated in Fig. 6. This derivation is based in three requirements:

1. The vocal comfort should be optimal (i.e. $0.45\text{ s} \leq T_{30,ME} \leq 0.55\text{ s}$).
2. SL $> 50$ dB in order to provide a signal-to-noise ratio of at least 15 dB for the students, assuming that the background noise level is not higher than 35 dB.
3. The volume of the room should be at least 6 $\mathrm{m}^3$ per student.

As can be seen in Fig. 6, the area of optimum design decreases with the number of students. There are more design possibilities for a classroom with 10 students than for a classroom with 40. In the latter case, there is almost no flexibility to choose the conditions that will result in optimum design. For 80 students, the different requirements do not intersect and therefore there are no optimum conditions. The optimum

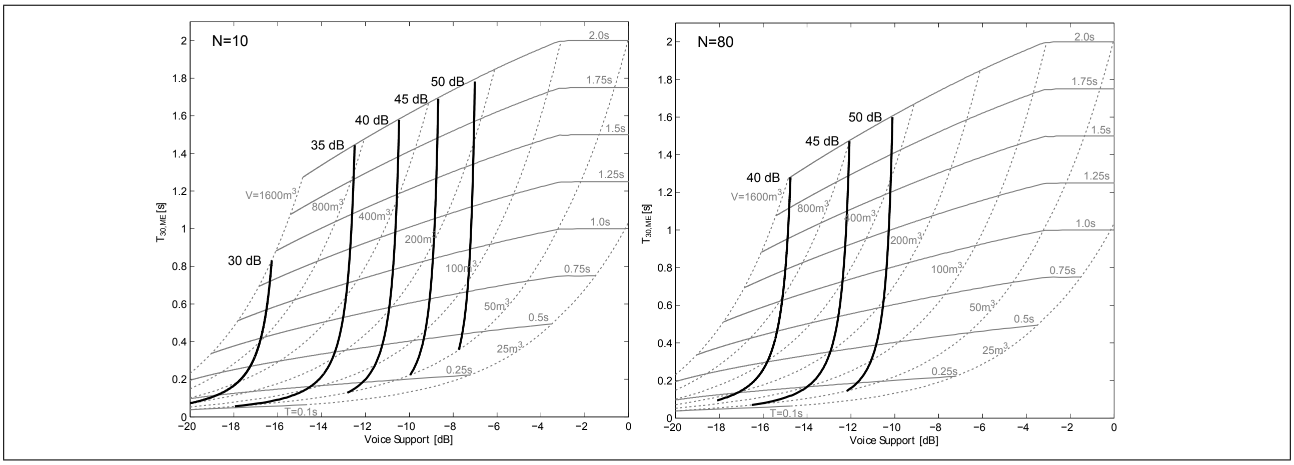


Figure 4. A-weighted student activity noise level iso-contours for different number of students in a classroom (left $N = 10$, right $N = 80$) as a function of $\mathrm{ST}_V$ and $T_{30,ME}$. In gray: guide lines for equal $T_{60}$ and $V$.

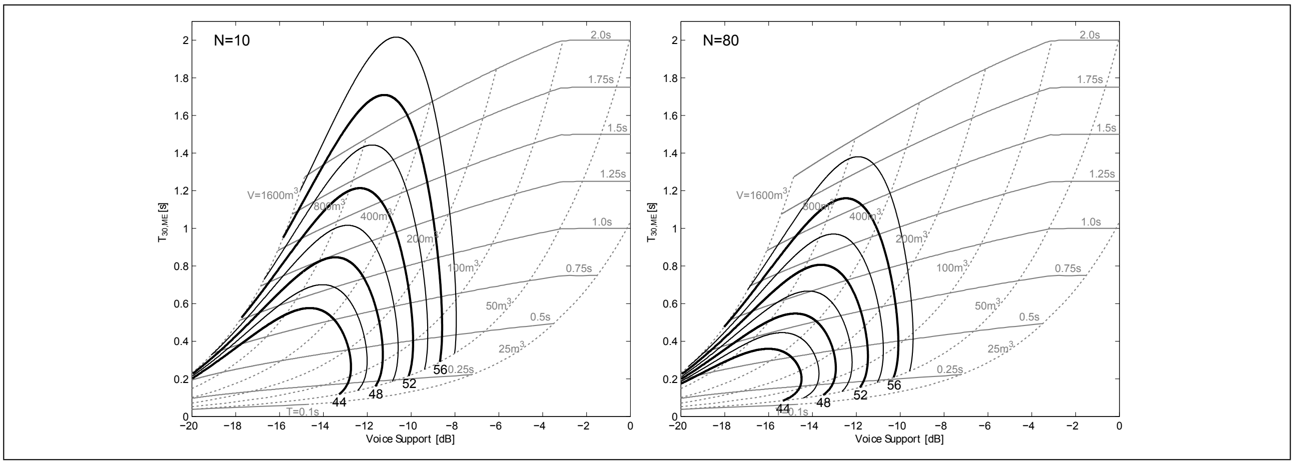


Figure 5. A-weighted instructor speech level iso-contours for different number of students in a classroom (left $N = 10$, right $N = 80$) as a function of $\mathrm{ST}_V$ and $T_{30,ME}$. In gray: guide lines for equal $T_{60}$ and $V$.

Table I. Recommended ranges for the parameters voice support $\mathrm{ST}_V$, reverberation time $T_{60}$ and volume $V$ for a speaker-oriented classroom acoustic design, as a function of the number of students $N$.

| Students $N$ | $V$ (m$^3$) | $T_{60}$ (s) | $\mathrm{ST}_V$ (dB) |
|---|---|---|---|
| 10 | 70 to 170 | 0.5 to 0.65 | -10.5 to -7.0 |
| 20 | 120 to 210 | 0.55 to 0.7 | -11.5 to -9.5 |
| 40 | 240 to 280 | 0.7 to 0.75 | -12.0 to -11.5 |

acoustic conditions, which do not exist for more than approximately 50 students, are summarized on table I and illustrated on Fig. 3, as a function of the number of students in a classroom.

These optimum conditions represent average values in classrooms and are meant for situations where the position of the speaker is variable, as in flexible teaching methods. For these situations, no optimum conditions for a talker can be achieved for more than approximately 50 students without exposing the teacher to talk uncomfortably or overstraining his voice. The optimum values of reverberation time, even though slightly higher than those recommended in standards like ANSI S12.60 [13] (maximum of 0.6 s in unoccupied, fully furnished conditions), are in good agreement with recent investigations on optimum classroom acoustic conditions to maximize speech intelligibility [14], who proposed optimum values in occupied conditions between 0.5 and 0.7 s and acceptable values between 0.4 s and 0.8 s. In lecturing style, the position of the talker is rather stable and therefore other acoustic designs taking into account geometrical considerations might be more efficient. In the last case, the limitation of a maximum of 50 students does not necessarily apply.

## Acknowledgement

This project has been funded by the Swedish organization AFA Försäkring.

## References

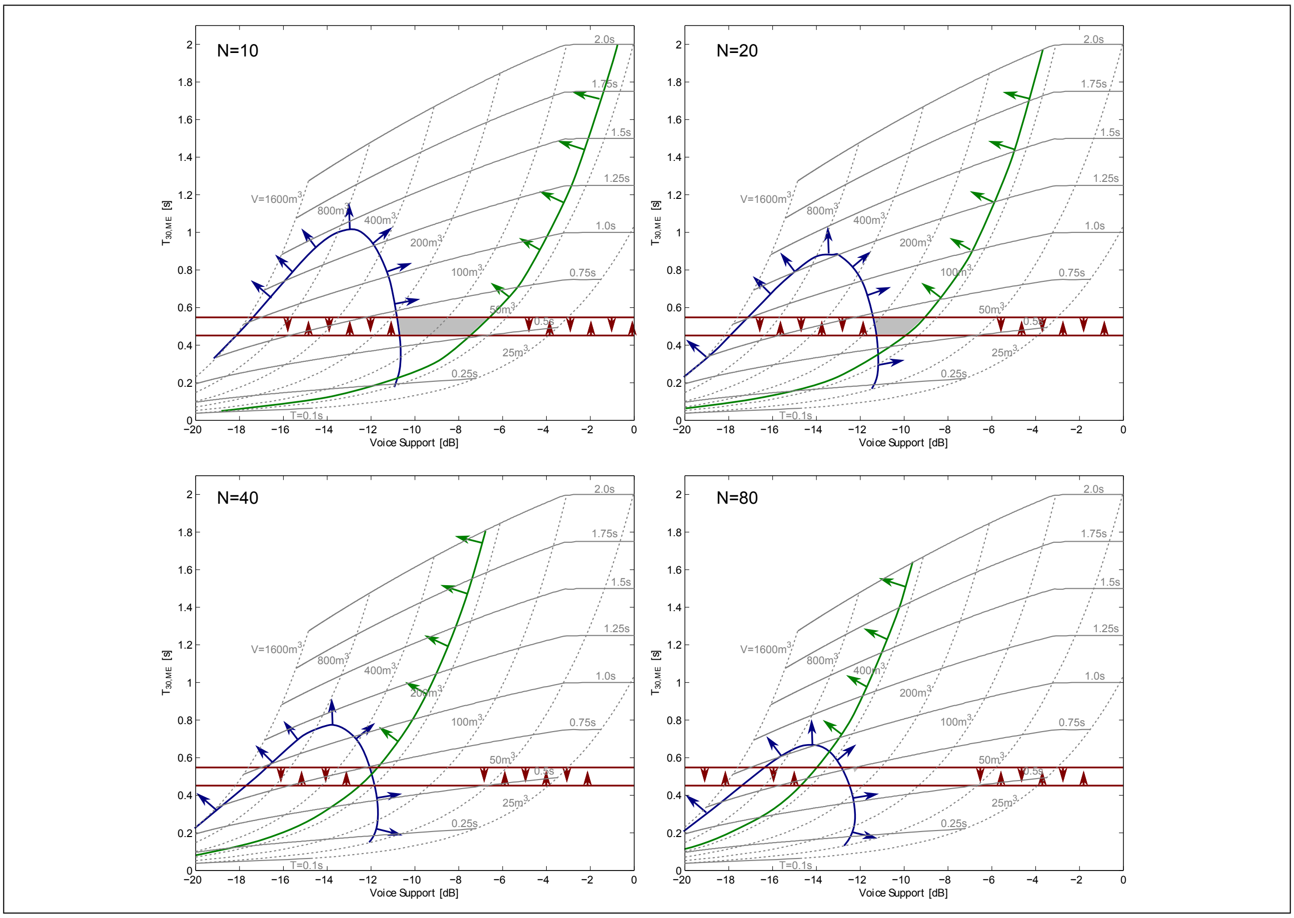

Figure 6. Derivation of the optimum classroom acoustics conditions according to the requirements of maximum vocal comfort (dark red), average SL of at least 50 dB (in dark blue), and more than 6 $m^3$ per student (in dark green). The different plots show the derivations for $N = 10$ (top-left), $N = 20$ (top-right), $N = 40$ (bottom-left), and $N = 80$ (bottom-right). The filled gray areas, as the intersection of the 3 requirements, show the optimum design values.

[1] P. Nelson, S. Soli, A. Seltz: Acoustical barriers to learning. Classroom acoustics 2. Acoustical Society of America, Melville, NY, 2002.

[2] V. Lyberg-Åhlander: Voice use in teaching environments. Speakers' comfort. PhD Thesis, Dept. of Logopedics, Phoniatrics and Audiology, Lund University, Sweden, 2011.

[3] H. Lane, B.Tranel: The Lombard sign and the role of hearing in speech. J.Speech Lang.Hear.Res. **14** (1971) 677–709.

[4] J. Whitlock, G. Dodd: Speech intelligibility in classrooms: specific acoustical needs for primary school children. Build. Acoust. **15** (2008) 35–47.

[5] J. Brunskog, A.C. Gade, G.P. Ballester, L.R. Calbo: Increase in voice level and speaker comfort in lecture rooms. J.Acoust.Soc.Am. **125** (2009) 2072–2082.

[6] D. Pelegrin-Garcia: Comment on "Increase in voice level and speaker comfort in lecture rooms" J.Acoust.Soc.Am. 125, 2072–2082 (2009)]. J.Acoust.Soc.Am. **129** (2011) 1161–1164.

[7] D. Pelegrin-Garcia: The role of classroom acoustics on vocal intensity regulation and speakers' comfort. PhD Thesis. Technical University of Denmark, Lyngby, Denmark, 2011.

[8] D. Pelegrin-Garcia, V. Lyberg-Åhlander, J. Brunskog, A. Löfqvist: Measurement and prediction of voice support and room gain in school classrooms. J.Acoust.Soc.Am. **131** (2012) 194–204.

[9] D. Pelegrin-Garcia, B. Smits, J. Brunskog, and C.H. Jeong: Vocal effort with changing talker-to-listener distance in different acoustic environments. J.Acoust.Soc.Am. **129** (2011) 1981–1990.

[10] D. Pelegrin-Garcia, O. Fuentes-Mendizabal, J. Brunskog, C.-H. Jeong: Equal autophonic level curves under different room acoustic conditions. J.Acoust.Soc.Am. **130** (2011) 228–238.

[11] D. Pelegrin-Garcia, V. Lyberg-Åhlander, R. Rydell, J. Brunskog, A. Löfqvist: Influence of classroom acoustics on the voice levels of teachers with and without voice problems: A field study. Proc.Meet.Acoust. **11** (2010).

[12] M. Hodgson, R. Rempel, S. Kennedy: Measurement and prediction of typical speech and background-noise levels in university classrooms during lectures. J.Acoust.Soc.Am. **105** (1999) 226–233.

[13] ANSI/ASA S12.60-2010/Part 1: American National Standard Acoustical Performance Criteria, Design Requirements and Guidelines for Schools, Part 1: Permanent Schools. Acoustical Society of America, Melville, NY, 2010.

[14] J. Bradley: A new look at acoustical criteria for classrooms. Proc. Inter-Noise 2009, Ottawa, Canada.