Local variations of speaker-oriented acoustic parameters in typical classrooms: a simulation study

David Pelegrín-García[1], Monika Rychtáriková[1,2], Christ Glorieux[1]
[1] Laboratory of Acoustics, Division Soft Matter & Biophysics, Dept. Physics & Astronomy, KU Leuven, 3001 Heverlee, Belgium.
[2] STU Bratislava, Faculty of Civil Engineering, Dept. of Building Structures, Radlinskeho 11, Bratislava, 813 68, Slovak Republic.

Summary
Classrooms are important spaces were teaching and learning takes place primarily through acoustic communication, which ideally should be enhanced by the acoustic conditions (geometry, materials and low background noise). Therefore, acoustic design is important to optimize speech intelligibility and vocal comfort, while limiting the vocal effort required to talk in. Two speaker-oriented parameters have been proposed: the Voice Support, linked to vocal effort, and the Decay Time at the Ears, linked to vocal comfort. Theoretical models exist for the prediction of room-averaged values of these parameters which overlook important local variations that teachers can use in their own benefit, as e.g. getting closer to a reflecting surface to increase the voice support. The present paper presents a method to calculate these acoustic parameters from commercial acoustic simulation software and studies the local variations of these parameters in typical classrooms. Results show that speaking close to the walls increases the voice support. Nevertheless, limitations in the calculation algorithms and the characterization of boundary conditions lead to a remarkable uncertainty in the prediction of Voice Support and Decay Time at the Ears.

PACS no. 43.55.Ka

1. Introduction

Classrooms are spaces where pupils and teachers devote a long time within their lives in order to receive or give an education, which takes place most of the time in the form of acoustic communication. Therefore, the acoustic conditions of the classrooms should allow the message to be clearly understood and at the same time, teachers and students should be able to speak comfortably without straining their voices [1].

While research on classroom acoustics has traditionally focused on speech intelligibility (e.g. [2, 3, 4]), recent studies have also focused on teachers' vocal behavior during lessons (e.g. [5, 6]) and on the classroom acoustics design for speakers' comfort [7].

Two room acoustics parameters have been proposed to describe the conditions for speakers: the voice support ST_V [8] and the decay time $DT_{40,ME}$ [9]. These parameters are derived from an Oral-Binaural Room Impulse Response (OBRIR) [10], which characterizes the airborne sound propagation between the mouth and the ears and is typically measured with a dummy head with a loudspeaker at its mouth and microphones at its ears.

The voice support is calculated from the OBRIR as the difference between the level of the reflected sound and the level of the direct sound. It is calculated in frequency bands and an overall value is calculated using a typical speech spectrum measured at the ears as a weighting function. The voice support is linked to the vocal intensity variations that occur when speaking in different physical environments [11, 12].

The decay time $DT_{40,ME}$ is a reverberation parameter but with the particularity that it is derived explicitly from the OBRIR (ME in the subindex refer to Mouth-to-Ears). It is defined as 1.5 times the time required for the backwards integrated energy curve on the OBRIR to fall from 0 dB to -40 dB.

Both ST_V and $DT_{40,ME}$ are averaged for the responses at the two ears.

While these speaker-oriented parameters have shown a potential link with vocal behaviour and subjective preference, their utility in classroom acoustic

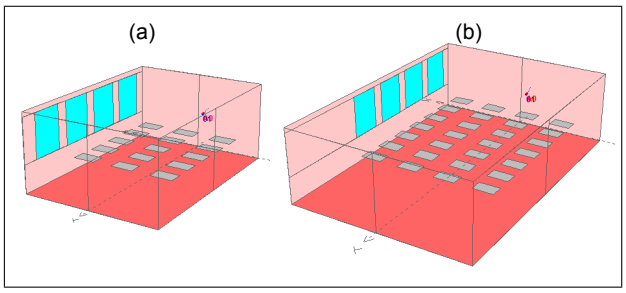

Figure 1. 3D models of the simulated small (a) and large (b) classrooms.

design is bound to the availability of a method to reliably predict these parameters. Simple prediction models of room-averaged values have been proposed in [7]. However, room-averaged values may not be suitable for e.g. auditoria or other rooms where the speaking area is a small percentage of the total floor area. In this article, we propose a method to predict ST_V and $DT_{40,ME}$ using commercial room acoustics software and attempt to estimate the spatial variability of the parameters with the method.

2. Method

A number of OBRIRs were calculated in two simulated classrooms using room acoustics software CATT-Acoustic™v9.0c. From these OBRIRs, speaker-oriented acoustic parameters ST_V and $DT_{40,ME}$ were derived in order to study the spatial variability and the influence of different simulation approaches based on using alternatively diffraction or scattering on the desks.

2.1. Description of the simulated classrooms

Two example classrooms were chosen for simulation: a standard-sized classroom ($V = 196$ m^3) of dimensions $7 \times 10 \times 2.8$m and a smaller classroom ($V = 98$m^3) of dimensions $5 \times 7 \times 2.8$m. The geometry used for the simulations is shown in Figure 1.

In the small classroom, 15 desks with a surface of 0.5×0.8m were drawn, arranged in 3 columns and 5 rows, with a spacing of 0.5 m between rows and 0.7 m between columns. In the large classroom, 28 desks were placed in 4 columns and 7 rows, with the same dimensions and spacing as in the small classroom.

The absorption coefficients of the materials used to simulate the classroom, and the corresponding surface areas, are shown in Table I.

2.2. Calculation points

In a small classroom, there were 56 calculation positions, with a grid spacing of 1 m (between rows) x 0.7 m (between columns) in between the desks and 0.5 m between rows close to the teacher lecturing position. In the larger one there were 108 positions. The grid space was 1 m (between rows) x 0.75 m (between columns) in between the desks and 0.5 m between rows close to the teacher lecturing position.

At each calculation point, regarded as the center point between the two ears, a receiver was placed, pointing towards the center of the room. A sound source, with a speech-like-directivity pattern was located at 10 cm in front of the receiver (in the aiming direction), and aimed towards the same direction as the receiver. It is assumed that, while the reflected sound is acceptably calculated (although with the limitations of commercial geometrical acoustics software), the direct sound ignores the diffraction effects of the head, though it can be corrected with the post-processing detailed in subsection 2.4.

2.3. Simulation parameters

In CATT-Acoustic, the calculations were run using algorithm 2 (which performs an actual split-up of rays after diffuse reflections), 100.000 rays(40.000 for the small classroom), and an IR length of 1000 ms.

The HRTF dataset used was the one of the artificial head developed and measured at ITA (RWTH Aachen, Germany) at full resolution (file ITA1_plain_48.DAT in CATT Acoustic). The sampling frequency was 48 kHz. No headphone compensation filter was applied.

Two different sets of simulations were performed: one modelling the finite size of the desks with scattering; the other set modelled them with diffraction edges:

- When using only scattering, the tables had a scattering of 50% at 125 Hz progressively decreasing to 20% at 1 kHz and keeping 20% in higher frequency bands.
- In the case of simulations with diffraction, the diffraction edges were only those of the desks (modelled as double-sided plains) and 0% of scattering was assigned to the desks. Only 1st order diffraction was considered, together with conversion of specular-to-diffracted and diffracted-to-specular energy components.

The rest of the surfaces had a default scattering of 10%.

2.4. Post-processing

Given that the direct sound $h_{D,CATT,spk,bin}$ (with sub-index D standing for direct sound, $CATT$ for the calculation method, *spk* for using a speaker-like directive source, *bin* for binaural receiver) in the oral binaural room impulse responses (OBRIRs) calculated from CATT Acoustic was not correct, a correction was applied. The direct sound $h_{D,CATT,spk,bin}$ (obtained from an anechoic simulation with only direct sound) was removed from the OBRIR $h_{CATT,spk,bin}$ and a

Table I. Absorption coefficients as a function of frequency and total surface area of the different materials in the simulated classrooms (SC = small classroom, LC = large classroom). Notice that the desk surface area counts both the top and bottom faces.

	Frequency (Hz)						S (m^2)	
Material	125	250	500	1000	2000	4000	SC	LC
Glass window	0.18	0.12	0.08	0.05	0.02	0.02	7.7	7.7
Hard Wall	0.1	0.1	0.1	0.1	0.1	0.1	59.5	87.5
Hard floor	0.08	0.08	0.08	0.05	0.05	0.05	35	70
Absorbing ceiling	0.2	0.4	0.7	0.85	0.88	0.9	35	70
Desks	0.1	0.1	0.1	0.1	0.1	0.1	12	22.4

new corrected direct sound $h_{D,bin}$ was then added to obtain the final OBRIRs h_{bin}.

$$h_{bin}(t) = h_{CATT,spk,bin}(t) - h_{D,CATT,spk,bin}(t) + h_{D,bin}(t) \quad (1)$$

The anechoic IR $h_{D,CATT,spk,bin}$ contains the effect of the HRTF at frontal incidence and a directivity filtering due to the emission in the backward direction from the source. These two effects are accounted for by calculating the anechoic IR $h_{D,CATT,omni,mo}$ with an omnidirectional source and a monaural omnidirectional receiver. However, the true direct sound is to be calculated taking into account the diffraction of sound around the head. Using the Boundary Element Method and a three-dimensional model of a human head taken from the OpenHear database [13], the sound pressure generated by a monopole source at the mouth (at the middle point between the lips) at a pair of points located at the entrance of the blocked ear canals P_{BEC} was calculated. In addition, a reference sound pressure P_{ref} generated by the same source at a point 10 cm away from it in the absence of the head was calculated (this point can be considered as the middle of the head). A *head gain* filter H_{HG} may be calculated as the ratio of both quantities, i.e.

$$H_{HG}(f) = P_{BEC}(f)/P_{ref}(f). \quad (2)$$

The meaning of this filter is the gain seen observed in the direct sound at the ears compared to the direct sound that would be obtained at a receiver in the middle position between the ears if the head were not present. The frequency characteristic of this filter pair (one for the left ear and one for the right ear) is shown in Figure 2. The two filters are not exactly identical due to geometrical asymmetry in the actual head.

Thus, the correct direct sound $h_{D,bin}$ is obtained by convolution of the monaural/omnidirectional IR calculated in CATT $h_{D,CATT,omni,mo}$ and the head gain filter h_{HG}.

$$h_{D,bin}(t) = h_{D,CATT,omni,mo}(t) * h_{HG}(t). \quad (3)$$

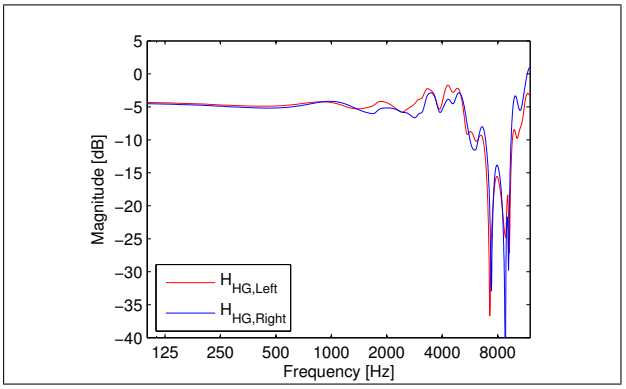

Figure 2. Head gain filter pair describing the sound pressure level at the entrance of the ear canal relative to the sound pressure level at the centre of the head (in its absence) in anechoic conditions.

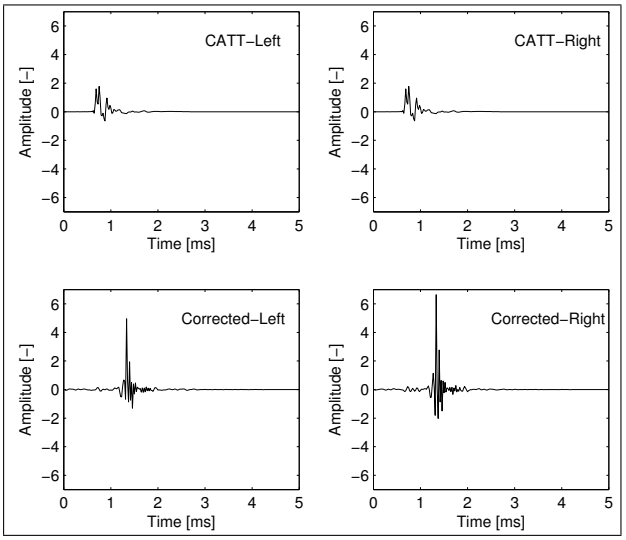

Figure 3. Impulse responses corresponding to the airborne sound propagation between the mouth and the ears, erroneously obtained with CATT $h_{D,CATT,spk,bin}$ (top) and after correction $h_{D,bin}$ (bottom).

The direct sound of the OBRIRs calculated in CATT $h_{D,CATT,spk,bin}$ and the corrected ones $h_{D,bin}$ are shown in Figure 3.

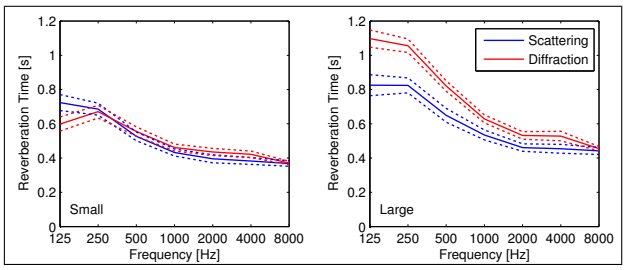

Figure 4. Reverberation time (average in solid line, ±1SD above and below in dotted line) extracted from the late part of the simulated OBRIRs with diffraction or with scattering, in the small (left) and the large (right) classrooms.

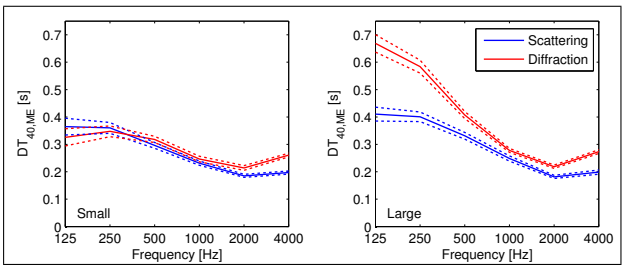

Figure 5. Decay Time DT_{40ME} (average in solid line, ±1SD above and below in dotted line) derived from the simulated OBRIRs with diffraction or with scattering, in the small (left) and the large (right) classrooms.

For the calculation of ST_V, the method of Brunskog et al. [11] was chosen i.e. the anechoic direct sound has been subtracted from each OBRIR in order to obtain the reflections. In fact, many calculation points in the present study had surfaces closer than 1 m and thus the method of windowing the OBRIR to separate direct and reflected sound [12] could not be applied.

2.5. Limitations of the calculation program

At a number of points (16 out of 108 in the large classroom, 9 out of 56 in the small classroom), the early part of the impulse response could not be calculated, probably due to conflicts of the simulation program in handling diffraction in close proximity of multiple diffraction edges and the source.

These problematic points were treated as outliers and thus excluded from the analysis. In addition, it is relevant to point out that the program had only one set of HRTFs to model sound incidence from a source in far-field; however there are reflections coming from close surfaces which theoretically result in an increase of SPL at low-frequencies, not accounted for by the simulation program.

3. Results

3.1. Reverberation Time

The values of reverberation time T, extracted from the decay between -30 and -60 dB in the backwards integrated energy curve in the OBRIRs and averaged in both ears, averaged across all positions in the classrooms and ±1 standard deviation (SD) around the mean are shown in Figure 4 as a function of frequency. The decay values for evaluation of T were chosen in order to avoid the influence of the direct sound, since source and receiver were only 10 cm away from each other. The reverberation time presents limited spread except for the lowest frequency bands. Whereas the simulation method does not introduce much variability in the small room, it does in the large room. Simulations with diffraction increase the values of reverberation time, specially at low frequencies.

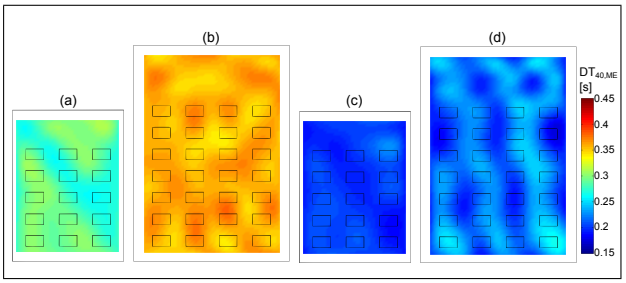

Figure 6. Spatial distribution of overall Decay Time DT_{40ME} in the small [(a), (c)] and the large [(b), (d)] classrooms in the simulations without [(a), (b)] and with [(c), (d)] diffraction.

3.2. Decay Time DT_{40ME}

The decay time DT_{40ME}, extracted from the decay between 0 and -40 dB in the backwards integrated energy curve in the OBRIRs and averaged for the two ears, averaged across calculation points and with ±1SD above and below the mean, is shown in Figure 5 as a function of frequency. Similarly to the reverberation times, the differences between the calculation methods are more remarkable for the large room.

The overall DT_{40ME}, derived from the low-pass filtered OBRIRs (f_c = 10kHz), is plotted as maps in Figure 6. The observed values are relatively smooth with variations within 0.05 s in the same room.

In order to analyze spatial variations, the overall DT_{40ME} is plotted as a function of the distance to the closest wall in Figure 7. The points closest to the walls tend to have shorter decay times than those further away (except in the case of calculation with diffraction in the large classroom). This is probably due to stronger first reflections from these walls that strengthen the very early energy and produce an apparently faster subsequent decay. This fact would lead to very different subjective experiences if these responses were to be used in auralization.

The differences in overall DT_{40ME} between the simulation methods, summarized in Table II, are between 0.08 and 0.13 s, which constitutes about a 50% in relation to the absolute values of the parameter.

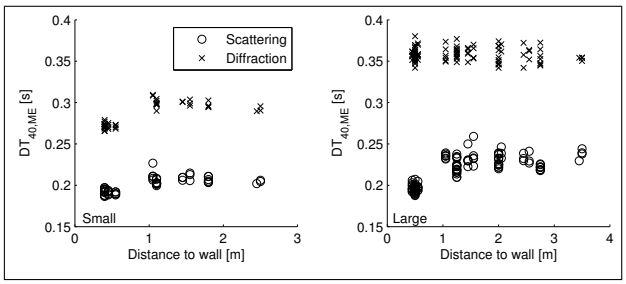

Figure 7. Overall Decay Time DT_{40ME} as a function of distance to the closest wall in the small (left) and large (right) classrooms.

Table II. Average (SD) values of $DT_{40,ME}$ in the two classrooms simulated with scattering or diffraction.

	Small	Large
Scattering	0.20 (0.01)	0.22 (0.02)
Diffraction	0.28 (0.02)	0.35 (0.01)

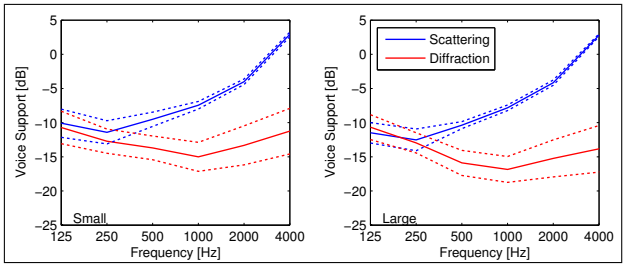

Figure 8. Voice support (average in solid line, ±1SD above and below in dotted line) derived from the simulated OBRIRs with diffraction or with scattering, in the small (left) and the large (right) classrooms.

3.3. Voice Support

The voice support ST_V, averaged across positions in the classrooms, with ±1SD above and below the mean values, is shown in Figure 8 as a function of frequency. Differently from the values in the previous parameters, there are large variations between the two simulation techniques in the two classrooms, as the simulations with scattering predict values that are more than 15 dB higher than the simulations with diffraction in some frequency bands. In addition, the spread in voice support increases with frequency but only in the simulations with diffraction.

The spatial distribution of overall voice support values is shown in Figure 9. As a reminder, the voice support is calculated from the frequency band values applying a typical speech spectrum as a weighting function. Since the maps in the two calculation methods share the same scale, the absolute differences are clearly observed: blue colors (i.e. low ST_V) in the calculations with diffraction and orange colors (i.e. high ST_V) in the calculations with scattering. In the same

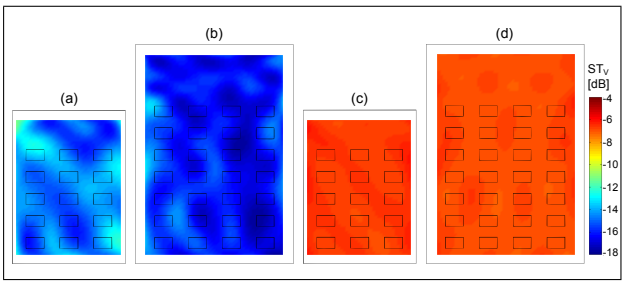

Figure 9. Spatial distribution of speech-averaged Voice Support ST_V in the small [(a), (c)] and the large [(b), (d)] classrooms in the simulations without [(a), (b)] and with [(c), (d)] diffraction.

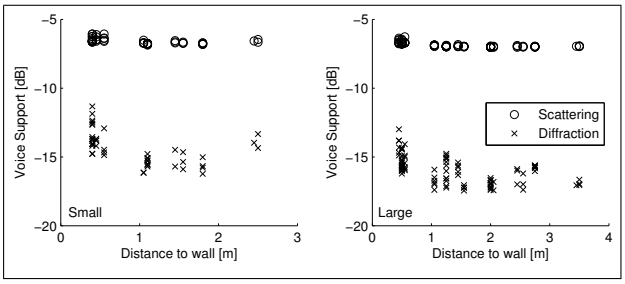

Figure 10. Overall voice support as a function of distance to the closest wall in the small (left) and large (right) classrooms.

Table III. Average (SD) values of ST_V in the two classrooms simulated with scattering or diffraction.

	Small	Large
Scattering	-6.3 (0.56)	-6.7 (0.36)
Diffraction	-13.5 (2.0)	-15.1 (1.9)

figure, larger variability is observed in the calculations with diffraction.

The overall values of voice support are also shown in Figure 10 as a function of the distance to the nearest wall. Specially in the simulations with diffraction, one can see that points closer to the walls tend to result in higher voice support values, thanks to the increase in energy produced by the closest surfaces.

Table III summarizes the spatially averaged values (and the standard deviation) for the two calculation methods. As happens with $DT_{40,ME}$, the voice support shows very large—and worrying—variations (more than 7 dB) depending on the calculation method.

4. Discussion and concluding remarks

A method for deriving OBRIRs from commercial room acoustics software has been proposed. It has been used to calculate the OBRIRs in two different classrooms.

The values of ST_V and $DT_{40,ME}$ obtained by two different modelling methods (either with scattering

or diffraction) of the desks are in strong disagreement. Using the prediction model of [7] (validated with actual measurements) the ST_V in a classroom of $V = 100m^3$ and $T = 0.45$ s is -10 dB and in a classroom of $V = 200m^3$ and $T = 0.6$ s, it is -11.5 dB. These values roughly correspond to the average between the two methods (as seen in Table III).

However, the decay time $DT_{40,ME}$ that follows from the prediction model in [7] would be 0.35 s for the small classroom and 0.48 s for the large classroom. These values are higher than those derived from the simulations (see Table II).

It has also been supported the hypothesis that the voice support increases close to the walls. By talking in the proximity of a wall, a teacher will indeed feel more supported by the room and will tend to limit the vocal effort.

The accuracy of currently existing geometrical acoustics software and the acoustic characterization of boundary conditions still has to improve, specially in complex environments like furnished classrooms, which requires a mixed modelling of diffraction and scattering by diverse architectural elements. While it seems apparent that the use of diffraction in CATT-Acoustic is relevant to model actual acoustic effects present in the classroom, predictions with diffraction give an underestimation of voice support and decay time (whereas just scattering modelling gives an overestimation of voice support and a further underestimation of decay time). The implementation of higher diffraction orders or conversion of higher reflection order to diffracted sound may give an advantage, at the expense of longer computation times.

Further measurement studies in actual classrooms will shed more light about the variability of ST_V and $DT_{40,ME}$ within and across rooms and will be a valuable input to better calibrate computational room models and algorithms for a more accurate prediction of these parameters.

Acknowledgement

DPG is grateful to FWO-V for supporting this research by financing his postdoctoral grant no. 1280413N.

References

[1] P. B. Nelson, S. D. Soli, and A. Seltz, "Classroom Acoustics II: Acoustical Barriers to Learning," tech. rep., Acoustical Society of America, Melville, NY, 2002.

[2] J. Bradley and H. Sato, "The intelligibility of speech in elementary school classrooms.," *The Journal of the Acoustical Society of America*, vol. 123, pp. 2078–2086, Apr. 2008.

[3] A. Astolfi, P. Bottalico, and G. Barbato, "Subjective and objective speech intelligibility investigations in primary school classrooms.," *The Journal of the Acoustical Society of America*, vol. 131, pp. 247–257, Jan. 2012.

[4] L. Nijs and M. Rychtáriková, "Calculating the Optimum Reverberation Time and Absorption Coefficient for Good Speech Intelligibility in Classroom Design Using U50," *Acta Acustica united with Acustica*, vol. 97, pp. 93 – 102, 2011.

[5] V. Lyberg Åhlander, D. Pelegrín García, S. Whitling, R. Rydell, and A. Löfqvist, "Teachers' Voice Use in Teaching Environments: A Field Study Using Ambulatory Phonation Monitor," *Journal of Voice*, vol. 28, no. 6, pp. 841.e5–841.e15, 2014.

[6] J. Kristiansen, S. r. P. Lund, R. Persson, H. Shibuya, P. M. b. Nielsen, and M. Scholz, "A study of classroom acoustics and school teachers' noise exposure, voice load and speaking time during teaching, and the effects on vocal and mental fatigue development.," *International archives of occupational and environmental health*, vol. 87, pp. 851–60, Nov. 2014.

[7] D. Pelegrín-García, J. Brunskog, and B. Rasmussen, "Speaker-Oriented Classroom Acoustics Design Guidelines in the Context of Current Regulations in European Countries," *Acta Acustica united with Acustica*, vol. 100, pp. 1073–1089, 2014.

[8] D. Pelegrín-García, J. Brunskog, V. Lyberg-Å hlander, and A. Löfqvist, "Measurement and prediction of voice support and room gain in school classrooms," *The Journal of the Acoustical Society of America*, vol. 131, pp. 194–204, Jan. 2012.

[9] D. Pelegrín-García and J. Brunskog, "Speakers' comfort and voice level variation in classrooms: Laboratory research," *The Journal of the Acoustical Society of America*, vol. 132, pp. 249–260, July 2012.

[10] D. Cabrera, H. Sato, W. Martens, and D. Lee, "Binaural measurement and simulation of the room acoustical response from a person's mouth to their ears," *Acoustics Australia*, vol. 37, no. 3, pp. 98–103, 2009.

[11] J. Brunskog, A. C. Gade, G. P. Bellester, and L. R. Calbo, "Increase in voice level and speaker comfort in lecture rooms.," *Journal of the Acoustical Society of America*, vol. 125, no. 4, pp. 2072–2082, 2009.

[12] D. Pelegrín-García, "Comment on ŞIncrease in voice level and speaker comfort in lecture roomsŤ [J. Acoust. Soc.Am. 125, 2072Ű2082 (2009)]," *Journal of the Acoustical Society of America*, vol. 129, pp. 1161–1164, Mar. 2011.

[13] R. R. Paulsen, J. A. Baerentzen, and R. Larsen, "Markov random field surface reconstruction.," *IEEE transactions on visualization and computer graphics*, vol. 16, pp. 636–646, Jan. 2010.