CLASSTALK SYSTEM FOR PREDICTING AND VISUALISING SPEECH IN NOISE IN CLASSROOMS

PACS REFERENCE: 43.55Hy, 43.55Ka

Hodgson, Murray
School of Occupational and Environmental Hygiene, University of British Columbia
3rd Floor, 2206 East Mall
Vancouver, BC V6T 1Z3
Canada
Tel: +1-604-822-3073
Fax: +1-604-822-9588
E-mail: hodgson@mech.ubc.ca

ABSTRACT

This paper discusses the *ClassTalk* system for modeling, predicting and visualizing speech in noise in classrooms. Modeling involves defining the classroom geometry, sources, sound-absorbing features and receiver positions. Empirical models, used to predict speech and noise levels, and reverberation times, are described. Surface absorption coefficient contributions are assigned based on previous work. Male or female speech sources, and overhead-, slide- or LCD-projector, or ventilation-outlet, noise sources can have four output levels; values are assigned based on ranges of values found from published data and on measurements. *ClassTalk* visualizes the floor-plan, speech- and noise-source positions, and the receiver position. The user can 'walk through' the room at will. In real time, five quantities - background-noise level, speech level, signal-to-noise level difference, Speech Transmission Index and speech intelligibility - are predicted and displayed, along with occupied and unoccupied reverberation times. An example of a large classroom before and after treatment is presented. The future development of improved prediction models and of the sound module, which will auralize speech in noise with reverberation, is discussed.

INTRODUCTION

A major acoustical concern in classrooms is the quality of verbal communication [1]. This can be quantified by the Speech Intelligibility (*SI*), the percentage of speech material correctly transmitted from a speaker to a listener. Here we are considering normal-hearing, adult listeners working in their first language. Non-optimal classroom design, acoustical conditions and *SI* can result in impaired verbal communication between a teacher and a student, with detrimental effects on learning. Classrooms must be designed to optimize *SI*. Prediction models are invaluable tools in the design process. They allow the design of a new classroom to be tested and optimized before construction. In the case of an existing classroom with non-optimum *SI*, they also allow control measures to be tested and optimized.

The objective of the work reported here was to develop a classroom speech-intelligibility prediction tool, applicable to typical classrooms, which is accessible to non-specialist users. It was required to be simple, fast, accurate and interactive, running on standard PC computers. The system developed - a Windows system - is called *ClassTalk*.

FUNCTIONALITY

ClassTalk allows a classroom to be modeled, sound-field descriptors to be predicted, and for the classroom and predicted quantities to be visualized on the computer monitor. In real time, the user can 'walk-through' the classroom on the monitor, exploring the sound field and associated speech quality. Using *ClassTalk*, the user can define the physical and acoustical characteristics of a classroom and its noise sources. *ClassTalk* takes into account occupant sound absorption and student-activity noise, both crucial to the accurate prediction of classroom acoustics. The classroom floor-plan is visualized on the monitor, along with the speech- and noise-source positions, and a receiver icon indicating the receiver position. The user can 'walk-through' the classroom at will by moving the receiver icon. At every walk-through position, in real time, five outputs - speech level, background-noise level, signal-to-noise level difference, Speech Transmission Index (*STI*) and speech intelligibility - are calculated and displayed. Contour maps of the five predicted quantities can be plotted. Also calculated and displayed are reverberation times for both the unoccupied (i.e., as designed by the architect) and occupied (i.e., as experienced by the occupants) classroom. The classroom floor-plan display can be printed, along with the input data, in the form of a one-page prediction report. The values of the five predicted quantities at the receiver-grid positions can be exported to a file for further processing.

The input characteristics of the classroom that must be entered are its description, dimensions, sound-absorbing features, and the coordinates and sound-power levels of the sources. Student-activity noise can be considered or not, as desired. Sound-absorbing features include carpeting, acoustical treatment of the walls and ceilings, upholstered seating and students, who absorb propagating sound. Sources can be male or female speakers, three types of projector or ventilation outlets. At any time, the classroom data can be edited - for example, to introduce control measures. Control can be achieved by modifying the room geometry, adding surface absorption, and reducing ventilation-outlet and equipment output levels. The modified classroom and the new values of the five acoustical quantities are displayed. The classroom can be 'walked-through' with real-time update of the visual display. New noise contours can be drawn.

PREDICTION MODELS

ClassTalk uses novel, simplified empirical models to predict total A-weighted speech levels (*SLA*) and noise levels - *BGNA*, due to ventilation (*VNA*) and student-activity (*SANA*) noises, as well as 1000-Hz early-decay times (*EDT1*) [2, 3]. Noise levels at a receiver position are calculated by summing the contributions of the individual noise sources. Individual speech- and noise-level contributions are determined from the predicted sound-propagation curves describing the rate of decrease of levels with increasing distance from a source, as well as from their absolute levels. The empirical models were developed from a database of 'typical' university classrooms measured when unoccupied and occupied. These account for speaker voice level (and how it adapts to the prevailing acoustical conditions), noise-source output levels, classroom shape and absorption (including occupants), as well as student-activity noise, as they occur in real classrooms. Following is a brief outline of the *SI* prediction algorithm (u = unoccupied classroom; o=occupied classroom); full details can be found elsewhere [2, 3, 4]:

- Predict *EDT1*u - correct for occupancy to *EDT1*o by diffuse-field theory;
- Predict *SLA*u - correct for occupancy to *SLA*o by diffuse-field theory;
- Predict *VNA*u - correct for occupancy to *VNA*o by diffuse-field theory;
- Predict *SANA*o;
- Calculate *BGNA*o by adding *VNA*o and *SANA*o energetically;
- Calculate signal-to-noise level difference, *SNA*o=*SLA*o-*BGNA*o;
- Calculate *STI*o from *EDT1*o and *SNA*o using the procedure described by Steeneken and Houtgast [5];
- Calculate *SI*o from *STI*o - assign the corresponding speech-quality descriptor [4].

Table 1. 1000-Hz surface-absorption-coefficient contributions of untreated and sound-absorbing surfaces [2].

Untreated surface	Carpeted floor	Wall or ceiling absorption	Upholstered seating
0.09	0.04	0.10	0.12

INPUT DATA

Sound-Absorbing Features

Entering the extent of the sound-absorbing features in the classroom involves inputting the percentage of the side walls and of the ceiling covered with acoustical treatment, as well as the percentage floor area that is carpeted. Since students absorb sound, the number of students inside the classroom is specified; the resulting absorption is determined from the 1000-Hz absorption per student of 0.81 m^2 [4, 6]. Finally, seating can be defined as hard or acoustically-absorbent upholstered seating. The total 1000-Hz surface absorption coefficient is found by summing the contributions of the untreated and sound-absorbing surfaces, shown in Table 1 [2].

Speech and Noise Sources

The typical ranges of total A-weighted output sound-power levels of male and female speakers was determined from published data [7]. Those of the four types of noise source - overhead, slide and LCD projectors, and of ventilation outlets - were determined from extensive sound-intensity measurements [8]. From the ranges, values corresponding to four output levels (referred to as quiet, normal, raised and loud) were assigned. The ranges and four levels for each source type are shown in Table 2.

EXAMPLE PREDICTIONS

Figure 1 shows *ClassTalk* prediction results for a large classroom, before treatment. Figure 2 shows the corresponding results after treatment. Student-activity noise was not considered. The figures show contour maps of four predicted quantities; from top to bottom, they are the total A-weighted speech level (*SLA*o), the total A-weighted background-noise level (*BGNA*o), the total A-weighted signal-to-noise level difference (*SNA*o) and speech intelligibility (*SI*o). The lower figures show the full *ClassTalk* visual displays for the two cases. They include the unoccupied and occupied reverberation times, and the values of the predicted quantities at the walk-through position indicated by the square receiver icon.

The classroom has dimensions of 24 m by 22 m by 6 m high, and contains 400 occupants. Before treatment, it has non-absorptive seats, walls and ceiling; the floor is 80% carpeted. A female instructor, speaking in a normal voice, is located 2.5 m from the front wall. An overhead projector, with normal output level, is located near the instructor. An LCD projector, with normal output level, hangs from the ceiling in the middle of the classroom. Three ventilation outlets, with loud output levels, are located on one side-wall. Speech levels vary with distance from the instructor from 55 to 39 dB. Noise levels vary from 49 dB near the noisy ventilation outlets, to 39 dB far from them. Signal-to-noise level differences vary from 12 dB near the instructor to -9

Table 2. Total A-weighted sound-power levels (in dB) used in ClassTalk for the different types of speech and noise source and the four volume settings.

Source	Minimum	Maximum	Quiet	Normal	Raised	Loud
Male Speaker	-	-	62.2	69.2	76.2	85.6
Female Speaker	-	-	59.6	66.6	73.6	80.6
Overhead Projector	47	57	42	48	54	60
Slide Projector	58	62	57	59	61	63
LCD Projector	49	64	40	50	60	70
Ventilation	41	60	35	45	55	65

dB near the ventilation outlets. Speech intelligibility varies from 94% ('good' quality) near the instructor to 39% ('bad' quality) near the vents. Reverberation times vary from 1.1 to 2.8 s in the unoccupied classroom, and from 0.6 to 1.7 s in the occupied classroom.

Acoustical treatment of this classroom aimed to reduce noise, and to control reverberation to near optimal values [9]. It consisted of HVAC noise control (reducing the ventilation outlets to 'normal' output levels), applying sound-absorptive materials to 50% of the ceiling and side walls, and installing sound-absorbing upholstered seating. After treatment, speech levels vary with distance from the instructor from 50 to 36 dB (reductions of 3 to 5 dB). Noise levels vary from about 35 dB near the overhead and video projectors - now the dominant noise sources - to 28 dB far from noise sources (reductions of 10 to 13 dB). Signal-to-noise level differences vary from 18 dB (an optimal value) near the instructor, to 6 dB under the video projector, and to 8 dB near the ventilation outlet farthest from the instructor (increases of 6 to 17 dB). Speech intelligibility varies from 94 to 96% ('very-good' quality) throughout the room (increase of 3 to 55 %). Reverberation times vary from 0.6 to 0.8 s in the unoccupied classroom, and from 0.5 to 0.7 s in the occupied classroom. The sound absorption - in particular, the upholstered seating - has reduced the sensitivity of the reverberation times to occupancy.

CONCLUSIONS

ClassTalk achieves the objective of developing a classroom prediction tool that is accessible to the non-specialist. A demo version is available from the author. A major objective of future work is to use simplified auralization techniques to auralize speech in noise with reverberation, corresponding to the predicted classroom sound field. However, realizing this within the constraints of readily-accessible hardware (computer, sound card, etc.) and real-time update is a challenge. Mixing speech with noise at the appropriate relative levels is straight-forward, but superimposing reverberation, with the predicted frequency variation, is a challenge which is currently being tackled. A further future objective is to improve the *ClassTalk* prediction models. More accurate, frequency-varying models are under development. The temporal effect of reverberation on speech intelligibility is better described by early-to-late energy fractions (*C50*) than by reverberation times. An improved *SI* model, involving a new *C50* empirical model, is currently being developed. *ClassTalk* also must be validated experimentally.

REFERENCES

[1] M. Picard and J. S. Bradley, “Revisiting speech interference in classrooms”, *Audiology* **40**, 221-244 (2001).

[2] M. R. Hodgson, "Empirical prediction of speech levels and reverberation in classrooms", *J. Build. Acoust.* **8**(1) 1-14 (2001).

[3] M. R. Hodgson, R. Rempel and S. M. Kennedy, “Measurement and prediction of typical speech and background noise levels in university classrooms during lectures”, *J. Acoust. Soc. Am*. **105**(1) 226-233 (1999).

[4] M. R. Hodgson, "Rating and ranking classroom acoustical quality for speech communication", submitted to *J. Acoust. Soc. Am*. (2002).

[5] H. J. M. Steeneken and T. Houtgast, “A physical method for measuring speech-transmission quality”, *J. Acoust. Soc. Am.* **67**(1) 318-326 (1980).

[6] M. R. Hodgson, "Experimental investigation of the acoustical characteristics of university classrooms", *J. Acoust. Soc. Am*. **106**(4) 1810-1819 (1999).

[7] *American National Standard Methods for Calculation of the Speech Intelligibility Index,* ANSI S3.5-1997.

[8] R. Wang, “Sound-intensity measurement of the sound-power outputs of classroom projectors and ventilation outlets”, Project report, University of British Columbia (2001).

[9] M. R. Hodgson and E.-M. Nosal, “Effect of noise and occupancy on optimum reverberation times for speech intelligibility in classrooms”, *J. Acoust. Soc. Am*. **111**(2) 931-939 (2002).

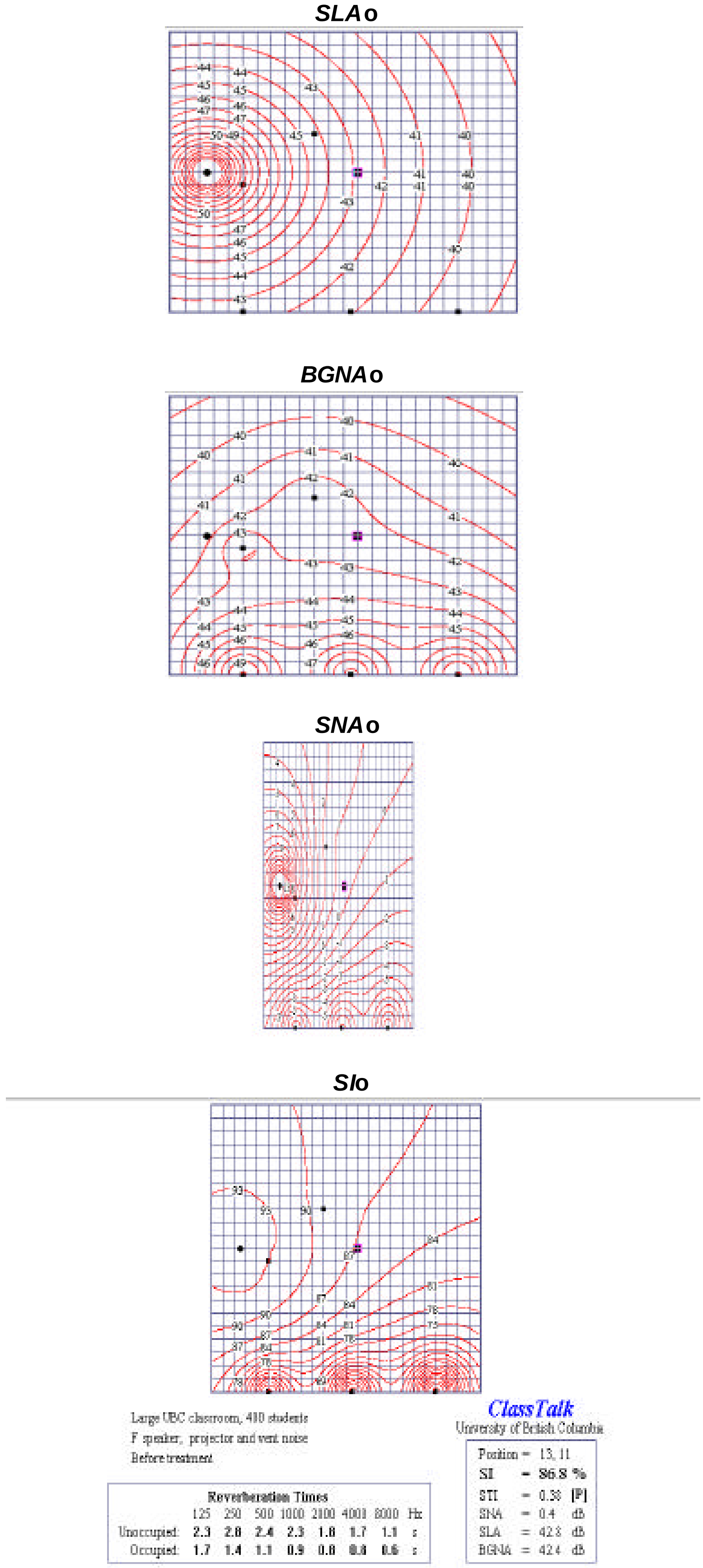

Figure 1. *ClassTalk* predicted sound-field and speech-quality descriptors for a large classroom before acoustical treatment. Classroom and treatment details are in the text. The bottom figure shows the full *ClassTalk* visual display.

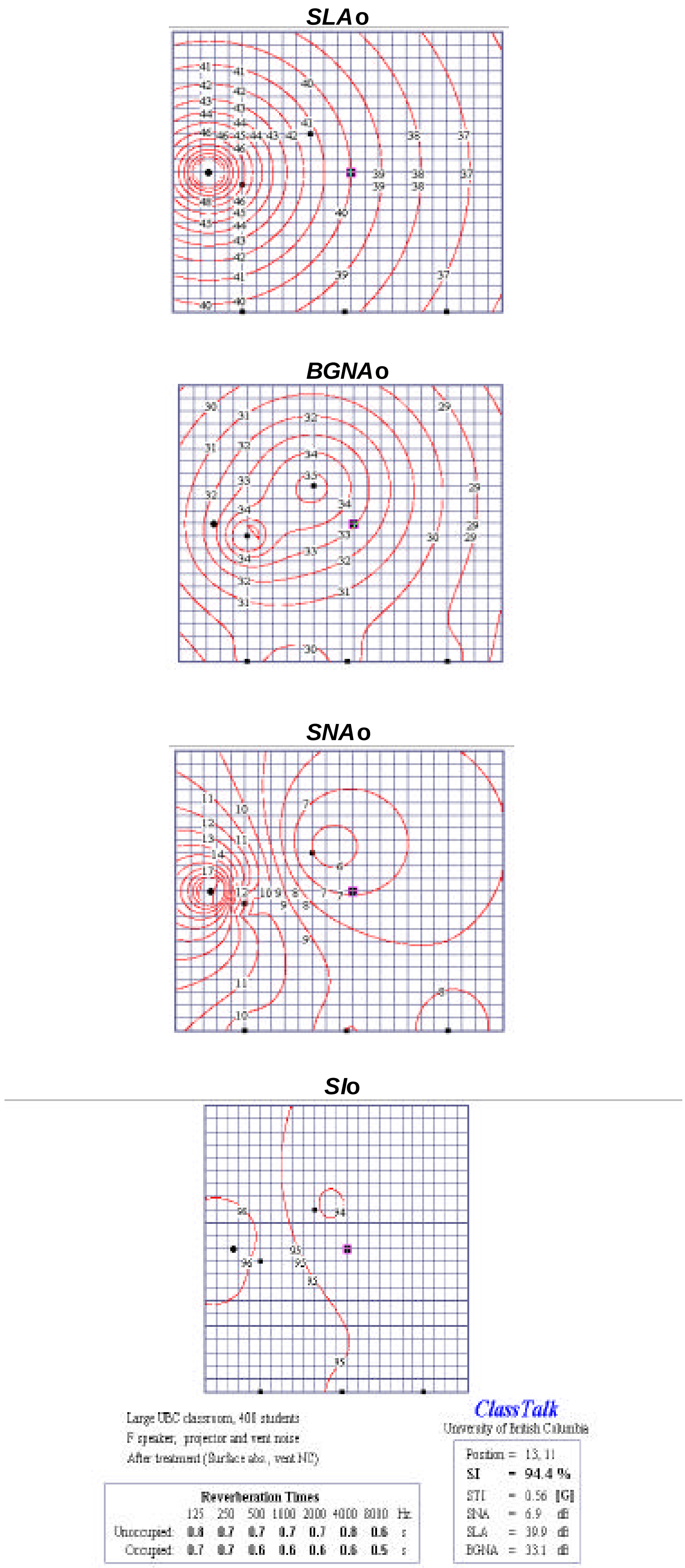

Figure 2. *ClassTalk* predicted sound-field and speech-quality descriptors for a large classroom after acoustical treatment. Classroom and treatment details are in the text. The bottom figure shows the full *ClassTalk* visual display.