Speech intelligibility studies in classrooms

Bradley, J. S.

Publisher's version / Version de l'éditeur:

Journal of the Acoustical Society of America, 80, 3, pp. 846-854, 1986-09

National Research Council Canada
Conseil national de recherches Canada

Institute for Research in Construction
Institut de recherche en construction

Speech Intelligibility Studies in Classrooms

by J.S. Bradley

Reprinted from
Journal of the Acoustical Society of America
Vol. 80, No. 3, September 1986
p. 846-854
(IRC Paper No. 1417)

Price $2.00

NRCC 26640

Canada

RÉSUMÉ

Des essais d'intelligibilité de la parole et des mesures acoustiques ont été effectués dans dix classes occupées. On a ainsi mesuré en bandes d'octave les niveaux de bruit ambiant, les temps d'affaiblissement initial et les temps de réverbération, ainsi que divers rapports son initial-son subséquent et le temps du centre. Divers rapports son utile-son nuisible par bande d'octave et l'indice de transmission du son ont été calculés. On a étudié les relations entre ces mesures pour préciser celles qui convenaient le mieux dans les classes, et on a identifié les mesures qui permettent le mieux de prévoir les pourcentages d'intelligibilité de la parole. Ces résultats ont servi à déterminer les conditions acoustiques optimales pour les classes, soit en termes de rapports son utile-son nuisible de 50 ms, soit grâce à des combinaisons du temps de réverbération et du niveau de bruit ambiant.

Speech intelligibility studies in classrooms

J. S. Bradley
Institute for Research in Construction, National Research Council of Canada, Ottawa, Canada K1A 0R6

(Received 23 April 1985; accepted for publication 29 April 1986)

Speech intelligibility tests and acoustical measurements were made in ten occupied classrooms. Octave-band measurements of background noise levels, early decay times, and reverberation times, as well as various early/late sound ratios, and the center time were obtained. Various octave-band useful/detrimental ratios were calculated along with the speech transmission index. The interrelationships of these measures were considered to evaluate which were most appropriate in classrooms, and the best predictors of speech intelligibility scores were identified. From these results ideal design goals for acoustical conditions for classrooms were determined either in terms of the 50-ms useful/detrimental ratios or from combinations of the reverberation time and background noise level.

PACS numbers: 43.55.Hy, 43.71.Gv

INTRODUCTION

Classrooms are an obvious example of rooms where a very high level of acoustical quality is required. Conventionally a teacher talks to a group of students who are intended to hear everything that the teacher says. A number of studies have indicated that excessive background noise and reverberation can influence not only speech intelligibility, but also the achievements and behavior of the students. One must therefore first decide what criteria high acoustical standards are to be based upon. Many references contain suggested maximum background noise levels and reverberation times for classrooms. The rationale behind the recommendations is often obscure, and there is not complete agreement among the various sources. Recently, a number of newer acoustical quantities have been proposed that are intended to better relate to measures of speech intelligibility, but these remain largely untested, particularly in rooms similar to classrooms. The purpose of the present work was to evaluate a number of acoustical measures as predictors of speech intelligibility scores, and to determine preferred acoustical design criteria for classrooms.

Acoustical design criteria for classrooms must include limits on interfering background noise as well as the acoustical properties of the room, which have conventionally been considered in terms of reverberation time. Several studies have considered the influence of intruding noise on students and have reported decreased students' achievement for more severe noise conditions. Bronzaft[1] found that word knowledge and reading comprehension scores were 3 to 4 months retarded for second, third, and fourth grade students in classrooms exposed to overhead train noise.

Cohen *et al.*[2] reported further detrimental effects of noise on students in classrooms. Lukas *et al.*,[3] in a study of the effects of road traffic noise, found reduced performance on reading and math tests; the reading test scores of the sixth grade students exposed to noise were 0.7 years behind comparable students in quieter schools. Only the study by Lukas *et al.* included reasonably complete noise measurements, and all three studies represented more extreme noise conditions.

Several other studies have considered in detail the requisites for high speech intelligibility for students of various ages. Nábělek and Pickett[4] referenced several studies in which speech intelligibility scores increased continuously as reverberation time was decreased to zero seconds. Both high noise levels and the use of hearing aids aggravated the effects of reverberation. In a more recent paper, Nábělek and Robinson[5] again showed that speech intelligibility scores increased continuously as reverberation time was decreased to zero seconds, and also that the detrimental effects of reverberation depended on the age of the subject. Thus both 10-year-olds and older subjects of 64 and 72 years were more negatively influenced by reverberation time than were young adults. Nábělek and Robinson also found that 10-year-olds required an extra 5 dBA of speech signal level to produce equivalent scores to young adults, and Elliott[6] reported that 7-year-olds required an extra 5 dBA of signal-to-noise ratio to produce equivalent scores to those of young adults. From these studies it is clear that optimum conditions based on tests with young adults (college students) are not stringent enough for most other subjects. In particular younger students, older listeners, and hearing impaired subjects require lower background noise levels and reverberation times to achieve the same optimum results.

Houtgast[7] obtained both speech intelligibility scores and estimates of the speech and background levels in classrooms. He related aggregate speech intelligibility scores to A-weighted signal-to-noise ratios and concluded that a 15-dBA signal-to-noise ratio eliminated the detrimental effects of interfering noise. Unfortunately no measurements of the acoustical properties of the rooms, such as reverberation times, were made.

Sargent *et al.*[8] carried out a survey of the responses of teachers in British schools exposed to road traffic noise and aircraft noise. Their results show various degrees of "bother" with road traffic noise, increasing with noise level from a threshold of about 50-dBA L_{eq} outdoor noise level.

Although decreased achievement scores are a more severe detrimental effect of adverse acoustical conditions, they are probably not the best measure to use to determine preferred conditions for classrooms because they are not a very sensitive measure and are strongly influenced by many non-acoustical factors. It seems better to base ideal acoustical conditions on speech intelligibility tests, which will reflect only acoustical conditions. Further, when speech intelligibility is near perfect it is very unlikely that acoustical conditions would be detrimental to academic achievement.

The present study consisted of both speech intelligibility tests and acoustical measurements in occupied classrooms. From the acoustical measurements a number of acoustical measures were calculated, and tested as predictors of the speech intelligibility scores. Comparisons among the various acoustical measures were also used to investigate their relative merits and their appropriateness in rooms similar to classrooms. The results were then used to estimate preferred acoustical conditions corresponding to near perfect speech intelligibility. The conclusions are applicable to other rooms of similar size intended for speech.

I. MEASUREMENT PROCEDURES

Tests were performed in ten classrooms containing grade seven and eight students (12- to 13-year-olds) in Ottawa, Canada. No hearing test information was available on the students, but those students who knew they had impaired hearing were excluded from the analyses. The classrooms were chosen to give the widest possible range of acoustical conditions typical of normal classrooms. They included school buildings of various ages in both suburban and inner city settings, as well as one school adjacent to a freeway. Acoustical details of the rooms are given in the results section below.

Following the procedure of a previous study,[9] a Fairbanks rhyme test was used to obtain speech intelligibility scores, as it is readily performed by inexperienced listeners. Students were tested in their regular classroom during their regular class time. Thus the noises present both from outdoors and from other parts of the school were completely representative of normal conditions. The students were very quiet during the tests and there was no significant noise from within the classroom being tested.

The recorded speech material was played back at four different levels (varying according to local conditions) using a loudspeaker with directional characteristics similar to human speakers. The source was positioned at the front of the room where the teacher might frequently stand. With the students sitting in their normal seats, the student seating area was divided into four approximately equal areas and acoustical measurements were made at the center of each of these areas. The speech intelligibility scores of all students in each of these groups were averaged and compared with the acoustical measurements. Thus, with four measurement groups in each of ten classrooms for each of four different speech levels, there were a total of 160 sets of data obtained by administering 972 individual speech intelligibility tests. As there was an average of 24.3 students in each classroom, each of the four groups of students in each classroom consisted of approximately six students.

The acoustical measurements included both 1-min background noise recordings and pulse measurements in the occupied classrooms. Integrated octave-band background levels were determined in octaves from 125–8000 Hz immediately after the speech tests with no noise-producing activity within the test room. As described in a previous study,[10] a 0.38-caliber blank pistol, which was found to be suitably omnidirectional,[11] was used to obtain pulse responses in the room with the gun positioned at the same location as the speech source. From these a number of acoustical measures were calculated in octave bands. These included reverberation time (RT), early decay time (EDT), early/late sound ratios for 35-, 50-, and 80-ms early sound limits (C_{35}, C_{50}, C_{80}), and Schwerpunktzeit or center time (TS).

Lochner and Burger[12] proposed a useful/detrimental sound ratio (U_{95}), where the useful sound comprised the direct sound and early reflections, and the detrimental sound combined the later reflections and the background noise. Their procedure required a complicated summation of weighted early reflections to produce the useful early energy. Simpler measures using the unweighted early energy sums, and based on C_{35}, C_{50}, and C_{80} have been successfully used in previous work. These useful/detrimental ratios are referred to as U_{35}, U_{50}, and U_{80}, respectively. Finally, the Speech Transmission index (STI), based on a modulation transfer function concept,[13] was also calculated. Where the speech level at each listener was required, it was calculated from the known source level and the measured acoustical properties of the classroom. Measuring speech levels might have led to more precise relations with speech intelligibility scores (SI), but would not have reflected the real errors that would arise in designing classrooms where such quantities must be calculated.

II. RESULTS

A. Measured conditions in the classrooms

Table I presents the mean, minimum, maximum, and standard deviation for a number of the major variables measured in the ten classrooms. Rooms varied from 253 to 529 m^3 and source–receiver distances in the rooms varied from 2.9–8.0 m. All source–receiver distances exceeded the critical distances in that room. The RT values at 1 kHz varied from 0.39–1.20 s with a mean of 0.72 s, and background noise levels in the classrooms varied from 38–45 dBA. Table I includes similar information for three early/late ratios (C_{35}, C_{50}, C_{80}) and the center time (TS). The range of EDT values was very similar to that for RT values.

Comparisons among the speech intelligibility scores in each classroom can readily be made by determining the mean speech intelligibility score in each room for a particular speech source level 1 m from the source. This was done for source levels of 40 and 50 dBA by fitting curves to plots of measured speech intelligibility scores versus the corresponding speech source levels to eliminate the irregularities of particular tests. The resulting mean speech intelligibility scores are shown for each of the classrooms in Fig. 1. This proce-

TABLE. I. Details of measured data.

Quantity	Minimum	Mean	Maximum	Standard deviation
Room value, m^3	253	312	529	81.8
Source receiver distance, m	2.86	4.72	8.00	1.44
Background level, dBA	38.4	41.9	45.1	2.10
C_{35} (1 kHz), dB	− 5.82	0.82	4.97	2.47
C_{50} (1 kHz), dB	− 2.22	3.30	7.92	2.73
C_{80} (1 kHz), dB	0.61	7.08	13.38	3.41
TS, s	0.0241	0.0488	0.0981	0.0186
RT, s	0.39	0.72	1.20	0.26
EDT, s	0.38	0.71	1.24	0.26

dure is proposed as a simple method for comparing the overall degree of speech intelligibility in different rooms. The relative ranking of the classrooms varies with the speech level. Where the mean speech intelligibility scores are most decreased at the lower speech level, interfering background levels would be expected to be the cause. Where scores are lower for both speech levels, poor room acoustics conditions, such as excessive reverberation, would be expected. Figure 1 suggests that the measured classrooms include a reasonable range of acoustical problems.

B. Interrelation of physical measures

Simple predictions of early/late ratios and center time can be made assuming an ideal continuous exponential decay.[10] Figure 2 plots measured C_{50} values against measured RT values at 1 kHz, along with the predicted relationship assuming an ideal exponential decay. Measured values cluster quite closely about the prediction curve, and in this case the standard error about the curve is 1.09 dB. There is a small tendency for the measured values to lie above the curve as there is a mean difference between the curve and measured values of 0.48 dB. Table II provides further results for other measures in octave bands from 125–8000 Hz. For each case both the mean difference between the measured and predicted values, and the rms (root-mean-squared) deviation about the curve are given. These results show that one can predict many of the early/late ratio values with an rms error of just over 1.0 dB for the present classroom data. The errors are much larger in the lowest octave band (125 Hz), and are larger for C_{35} than for the other two early/late ratios. As the mean TS value at 1 kHz was 0.048 s, the associated rms error of 0.0059 s represents a 12% variation about the predicted curve.

A previous study of measurements in larger rooms[10] found much larger errors when attempting to predict measured values of early/late ratios from ideal exponential decay theory. While such simple predictions were inadequate in larger rooms, they are more satisfactory in smaller rooms such as classrooms. Barron and Lee[14] have recently proposed an improved method for predicting C_{80} values from the RT, room volume, and source–receiver distance. When his method [Barron's Eq. (7)] was tried with the present data, the rms deviation was reduced to 0.82 dB. Although this represents an improvement (rms error reduced by 0.2 dB), it is probably not of great practical importance for smaller rooms.

The EDT is usually a better correlate of subjective evaluations of rooms. For the present data it was very closely related to the conventional RT of the room. Only at lower

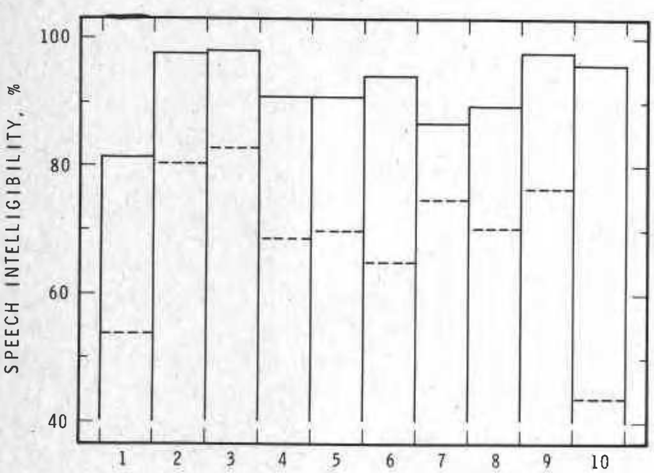

FIG. 1. Mean classroom speech intelligibility scores for speech source levels of 50 dBA —— and 40 dBA - - -.

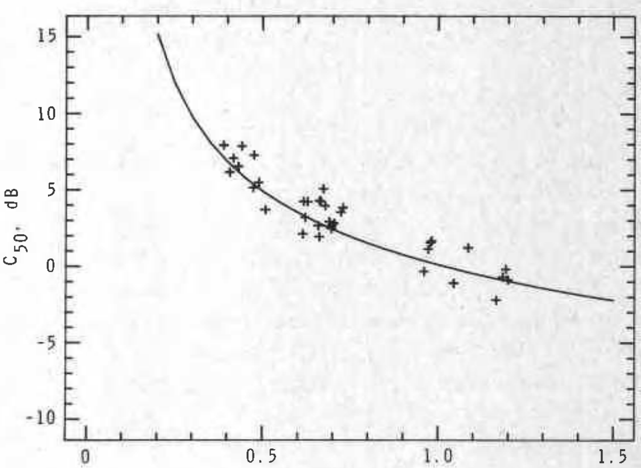

FIG. 2. Measured C_{50} values versus measured RT values at 1 kHz + + + and prediction from exponential theory ——.

TABLE II. Prediction errors for ideal exponential decay theory.

Measure	125	250	500	1000	2000	4000	8000
	Octave-band frequency, Hz						
	Mean differences						
C_{35}, dB	−4.25	−0.30	0.88	0.34	0.70	0.85	1.06
C_{50}, dB	−2.76	0.01	0.59	0.48	0.63	0.76	1.00
C_{80}, dB	−0.93	−0.22	0.48	0.47	0.62	0.67	0.83
TS, s	0.0136	0.00592	−0.00305	−0.00255	−0.00373	−0.00376	−0.00399
	rms differences						
C_{35}, dB	5.55	2.42	1.67	1.27	1.24	1.34	1.38
C_{50}, dB	4.43	2.16	1.37	1.09	1.02	1.20	1.27
C_{80}, dB	2.44	2.00	1.10	1.04	0.97	1.05	1.04
TS, s	0.0179	0.0107	0.00608	0.00593	0.00621	0.00627	0.00538

frequencies were there substantial differences between EDT and RT values, and these differences diminished with increasing frequency. Figure 3 plots measured EDT values against measured RT values at 125 Hz; the EDT values deviated from RT values by up to 50% in the worst case. At 500 Hz, differences were diminished to a maximum of 0.1 s and continued to decrease at higher frequencies. Thus for these smaller rooms it would be important to differentiate between RT and EDT values only at lower frequencies.

The STI values derived from modulation transfer functions have a quite different origin than do early/late ratios. Figure 4 compares measured STI values with measured C_{80} values. These STI values do not include the effects of background noise. Although the best-fit curve is not linear, the two measures are very closely related, and knowledge of one implies, to quite a high level of precision, knowledge of the other. This confirms earlier results for larger rooms,[10] and further demonstrates the strong relationships that exist between various newer acoustical measures.

C. Predictors of speech intelligibility

Various acoustical measures were considered as predictors of speech intelligibility scores by fitting third-order polynomials to each set of data and considering the resulting multiple correlation coefficients and associated standard errors (SE) as indicators of the prediction accuracy. Figure 5 shows the example of measured speech intelligibility scores versus measured steady state A-weighted signal-to-noise ratios [S/N(A)] and the best-fit curve. The mean trend suggests that speech intelligibility scores increase with increasing signal-to-noise ratios up to approximately + 15.0 dBA, where a plateau is reached. Similar ideal minimum signal-to-noise ratios were obtained in two other recent studies.[8,9] The multiple correlation coefficient for the third-order polynomial fit to the data of Fig. 5 was 0.805, with an SE of ± 9.94%. When the articulation index (AI) was used as a predictor, the resulting multiple correlation coefficient was 0.828 and the SE was ± 9.38%. Thus the S/N(A) ratio and the AI exhibited very similar prediction accuracies.

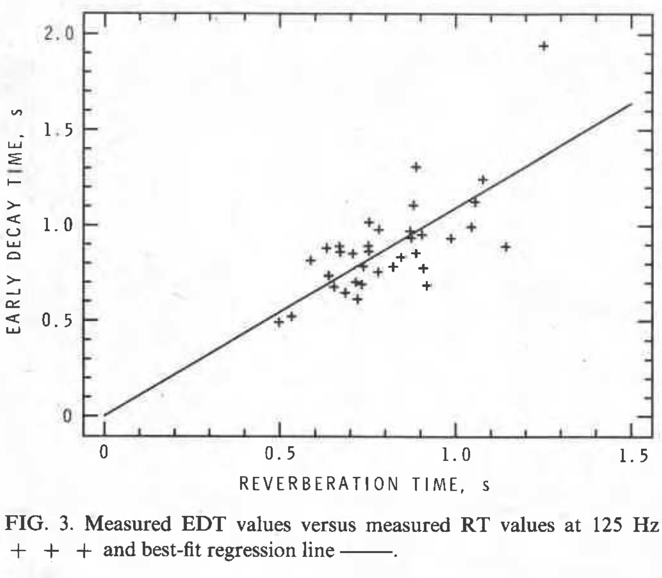

FIG. 3. Measured EDT values versus measured RT values at 125 Hz + + + and best-fit regression line ——.

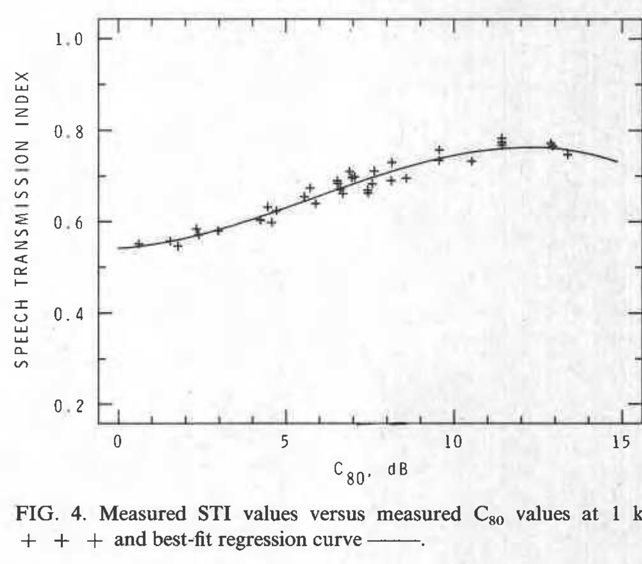

FIG. 4. Measured STI values versus measured C_{80} values at 1 kHz + + + and best-fit regression curve ——.

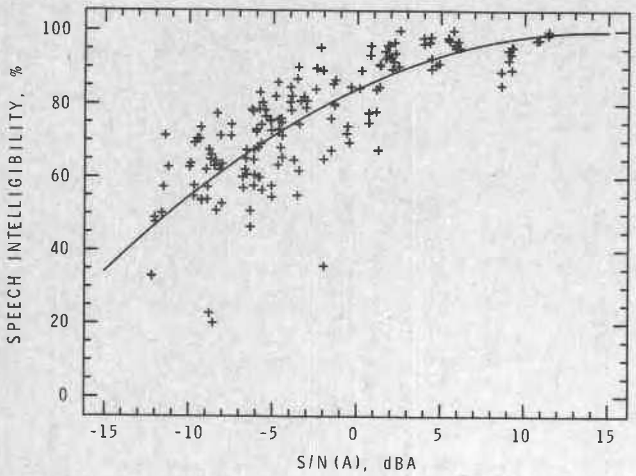

FIG. 5. Measured SI values versus measured S/N(A) values + + + and best-fit regression curve ——.

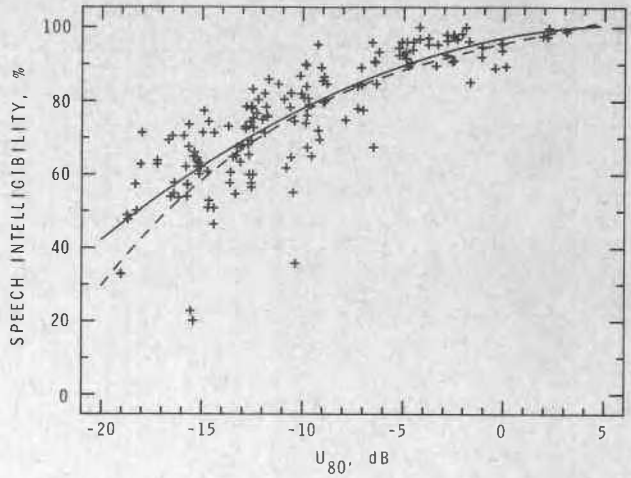

FIG. 6. Measured SI values versus measured U_{80} values at 1 kHz + + + best-fit regression curve to this data —— and best-fit regression curve to similar data from Ref. 9 - - -.

In previous work[9] the 80-ms useful/detrimental ratio (U_{80}) was one of the better predictors of speech intelligibility. Figure 6 plots speech intelligibility scores versus U_{80} values. The solid line is the best-fit curve to the present data, and the dashed line is the best-fit curve from the data of the previous research. The two curves are in very close agreement and the mean trends of both sets of data are very similar. However, compared to Fig. 4 of Ref. 9, the present data exhibit greater scatter about the best-fit curve.

Table III gives the multiple correlation coefficients and associated standard errors for third-order polynomial fits to each measure in each octave band. The multiple correlation coefficients are generally similar in magnitude, and almost all of them are greater than 0.8. The standard deviations show that the best predictors predict speech intelligibility within an SE of close to $\pm 9.0\%$. There are only small differences between the different useful/detrimental ratios, and only in the 125-Hz octave band do the errors increase a great deal. Although the Lochner and Burger useful/detrimental ratio values (U_{95}) are more complicated to calculate, they were of essentially the same prediction accuracy as the less complex U_{50} values. In comparison, the third-order fit between speech intelligibility and STI values, including the effects of background noise, produced a multiple correlation coefficient of 0.844 and a standard error of $\pm 9.0\%$. Thus it was of similar prediction accuracy to the other measures in Table III.

Although many of the present results in classrooms are quite similar to the previous tests in larger rooms,[9] there are some differences. In the larger rooms, the 80-ms useful/detrimental ratio was a slightly better predictor, while in the smaller rooms of this study U_{50} was slightly better. In the present study speech intelligibility scores were not significantly related to any of the octave-band early/late ratios or to any of the decay time measurements. The best predictions in the present study had associated standard errors of $\pm 9.0\%$ or larger, while in the larger room study, standard errors were approximately 2.0% smaller. The increased scatter in the present results is largely due to the difficulty of precisely representing the background noise levels that existed during the classroom tests. Because the tests were carried out during school hours there were many intermittent noises from other classrooms and from outside the school. Noise measurements after the speech intelligibility tests were completed did not always precisely represent the noise levels that existed during the speech tests. In the previous large room study, conditions were more controlled and intruding noise was a relatively minor problem. There may also be some increased scatter in these results because the subjects were younger than the adults of the previous study, and on average there were six subjects in each group compared to the nine of the previous study. Although the number of subjects in each group was smaller in this study, each group covered a much larger portion of the room and so acoustical conditions within each group could have been more varied.

D. Compound predictors

Partial correlation analyses were used to identify other variables that could be combined with those discussed in the section above to form improved compound predictors. Using this technique the variance associated with particular

TABLE III. Prediction accuracies of third-order polynomials.

Quantity	Octave-band frequency, Hz 125	250	500	1000	2000	4000	8000
	Multiple correlation coefficients ($N = 160$)						
U_{35}	0.725	0.832	0.838	0.841	0.835	0.835	0.838
U_{50}	0.758	0.844	0.835	0.842	0.836	0.835	0.838
U_{80}	0.823	0.829	0.826	0.832	0.835	0.831	0.835
U_{95}	0.725	0.829	0.837	0.843	0.837	0.833	0.837
	Standard errors, percent speech intelligibility						
U_{35}	11.5	9.3	9.2	9.1	9.2	9.2	9.1
U_{50}	10.9	9.0	9.2	9.1	9.2	9.2	9.1
U_{80}	9.5	9.4	9.5	9.3	9.2	9.3	9.2
U_{95}	11.5	9.4	9.2	9.0	9.2	9.3	9.2

major variables is first partialled out, and the remaining variance in speech intelligibility scores is then related to the variance in other possible predictors. Three different sets of analyses were performed partialling out the effects of the third-order polynomial predictors associated with S/N(A), U_{80}, and U_{50}. In all three cases, early/late ratios in the 250-Hz octave band exhibited strong positive partial correlations with speech intelligibility scores. In addition, there were strong negative partial correlations with the 250-Hz RT, and somewhat stronger negative partial correlations with the 125-Hz EDT values. For each of these measures the partial correlations in other octave bands were smaller in magnitude and sometimes not statistically significant.

Multiple regression analyses were then performed regressing speech intelligibility scores onto combinations of third-order polynomial predictors and either the 250-Hz C_{50} values or the 125-Hz EDT values. The new compound predictors produced increased multiple correlation coefficients and decreased standard errors by about 0.5%. Similar analyses were carried out on the data from the previous study of large rooms but these suggested that the 125-Hz C_{35} value was the most important additional predictor. The importance of these additional predictors cannot be completely determined from the existing data, but results from both sets of data do suggest that lower frequency early/late ratios influence speech intelligibility, and should be included in better acoustical measures. The precise nature of the necessary combination of early/late ratio values from different octaves might best be determined using synthesized sound fields in laboratory studies.

Further compound predictors were produced with combinations of either S/N(A) or AI with one of the decay time measures (RT or EDT). Combinations of AI values and either RT or EDT values at 1 kHz were no better as predictors of speech intelligibility scores than the AI values alone. Combinations of S/N(A) values with either RT or EDT values at 1 kHz produced small improvements in prediction accuracy. In both cases the standard error decreased by approximately 0.4%. The resulting regression equation for combinations of S/N(A) and the 1-kHz RT values was

$$SI = 2.26\,S/N(A) - 0.0888\,S/N(A)^2 - 13.9\,RT + 95.0. \quad (1)$$

Figure 7 is based on this same regression equation and plots speech intelligibility score (SI) versus S/N(A) for 1-kHz RT values of 0.5, 1.0, and 1.5 s. As the associated standard error was $\pm 9.6\%$, the above equation and Fig. 7 represent a practical method of estimating SI values from readily available acoustical measures. They will also be used in the following section as one approach to determining preferred design goals for acoustical conditions in classrooms.

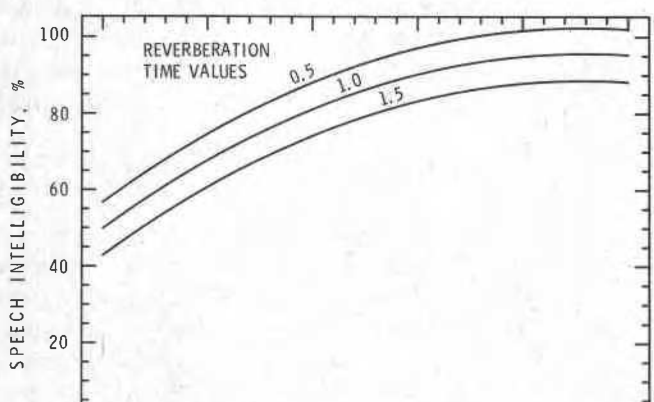

FIG. 7. Best-fit curves from multiple regression analysis of measured values of SI vs S/N(A) for RT values of 0.5 (top), 1.0 (middle), and 1.5 s (bottom).

III. DETERMINING PREFERRED DESIGN GOALS FOR ACOUSTICAL CONDITIONS IN CLASSROOMS

To specify preferred acoustical conditions for classrooms one must specify both maximum permissible background noise levels and preferred room acoustics conditions (which have usually been specified in terms of an optimum reverberation time). One can find a range of recommended values in various references. Beranek[15] suggests NC 30–40 (Noise Criterion) for classrooms, Stumpf[16] and Sharland[17] recommend NC 30–35, Parkin and Humphreys[18] give an octave spectrum that approximates NC 25, while Burns[19] gives NR 20–30 (Noise Rating) as ideal. Analysis of the present classroom data indicated that A-weighted background levels were on average 5.6 dB greater than NC values. Further, NR values can be taken as approximately equal to NC values. If we take the low end of the various recommended ranges to be more ideal, and convert them to approximate A-weighted levels, we obtain recommended maximum background noise conditions for classrooms ranging from 26–36 dBA. Preferred reverberation times for speech are frequently based on Knudsen and Harris's curve,[20] which recommends a value of 0.7 s for a room of 300 m^3.

A. Combinations of RT and S/N(A)

The simplest approach to determining ideal conditions from the present results is to use Eq. (1) as illustrated in Fig. 7. As a criterion for excellent acoustics, it is again[9] argued that this corresponds to the point where the best-fit curve approaches most closely to 100% speech intelligibility for these results using the rhyme test. Various combinations of S/N(A) and 1000-Hz RT in Eq. (1) would lead to an expected speech intelligibility of 100%. Such combinations were determined and are plotted as the dash–dot curve in Fig. 8. This curve indicates that reverberation times less than 0.7 s are necessary for 100% SI. As reverberation time decreases below this, some increase in background levels can be tolerated while still expecting 100% SI. For example, this curve suggests that with 0.4-s RT, an S/N(A) of 6 dBA could be accepted.

These S/N(A) can be used to determine maximum permissible background noise levels using accepted normal speech levels, and typical acoustical conditions for a classroom sized room. Pearsons *et al.*[21] determined mean speech source levels for males, females, and children as a function of the level of vocal effort. For "normal" vocal effort, females

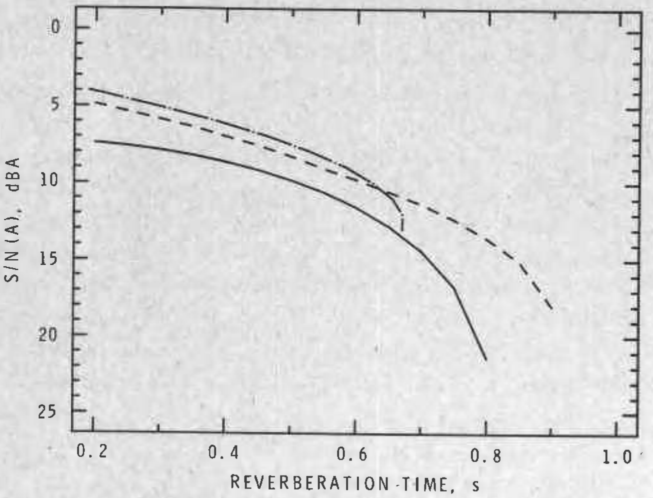

FIG. 8. Equal speech intelligibility contours: U_{50} equal to + 1.0 dB ——, 100% SI values from Fig. 7 – · –, and STI equal to 0.6 - - -.

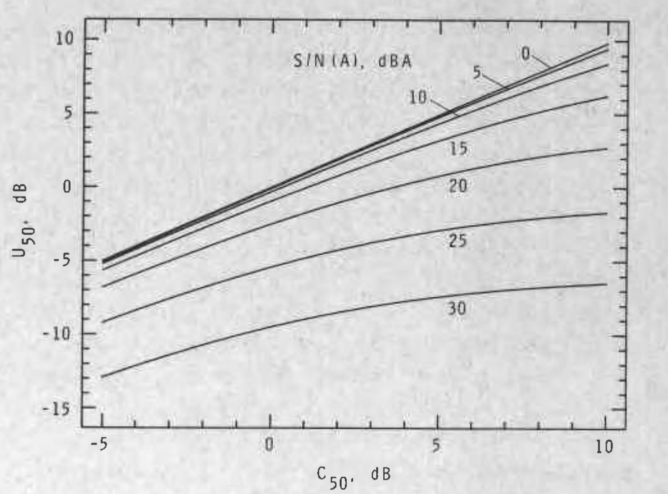

FIG. 9. Calculated 1-kHz U_{50} values versus 1-kHz C_{50} for different S/N(A) values.

had the lowest mean speech source level of 55 dBA with a standard deviation of 4 dBA. Thus 51 dBA (i.e., one standard deviation less than the mean) represents a conservative speech source level, below which the levels of few speakers would fall. For rooms of 300 m^3 a conservative estimate of the reduction in levels from the source to positions in the reverberant field is -2 dBA.[9] Thus if one selects an optimum S/N(A) of 15 dBA, as suggested by the data in Fig. 5, one arrives at a recommended maximum background level of $55-4-2-15 = 34$ dBA. Similarly, S/N(A) values from Fig. 8 could be converted to maximum preferred background noise levels.

The influence of RT is predicted by the -13.9 RT term in Eq. (1). This indicates that SI values would decrease by 13.9% for a 1.0-s increase in RT. It is of interest to compare this rate of change of SI scores versus RT with other results in the literature. Nábělek and Robinson[5] presented word identification scores as a function of RT for various age groups including 10- and 27-year-olds. Elliott's results[6] suggest that the 12- and 13-year-olds of the present study would perform approximately intermediate to the 10- and 27-year-olds of Nábělek's research. When the results of the 10- and 27-year-olds were averaged to obtain closer agreement with the 12- and 13-year-olds of the present study, the mean slope was -12.6% per second of RT. For the 27-year-olds the slope was less and for both younger and older subjects there were stronger effects of RT. Finitzo-Hieber and Tillman's[22] results for younger children also showed a larger effect of increased RT, -19.2% per second of RT. In an earlier paper Nábělek and Pickett[4] quote results from a number of studies in terms of percent words correct versus RT. For conditions of negligible background noise, the average mean slope was -12.3% per second of RT; for conditions with background noise, the average mean slope was approximately -21% per second of RT. All of these results suggest continuously improved speech intelligibility as RT is decreased to 0 s, and do not find nonzero optimum RT values. The detrimental influence of RT depends on both the age of the subjects and also the amount of background noise that is present. For similar conditions and age of subjects the present results agree with other published work and suggest a decrease of between 12% and 14% per second of RT.

B. Derivation from useful/detrimental ratios

In previous work,[9] optimum conditions for speech were derived from U_{80} values, which were found to effectively combine the detrimental effects of background noise and late reflections with the beneficial effects of the direct sound and early reflections. The present work suggests that U_{50} is better suited for predicting speech intelligibility in classrooms. From the regression analyses the best-fit curve relating measured SI values to U_{50} values at 1 kHz was:

$$SI = 1.027 \cdot U_{50} - 0.838 \cdot U_{50}^2 + 99.42. \qquad (2)$$

Using this mean trend, 100% SI is reached for a U_{50} of approximately + 1.0 dB. A U_{50} of + 1.0 dB can be achieved by various combinations of S/N(A) and C_{50}. This is illustrated in Fig. 9, which plots U_{50} versus C_{50} for a range of values of S/N(A). In the present results in classrooms (see Fig. 2), C_{50} values were quite closely related to RT values. In fact, at 1 kHz one can predict C_{50} values from RT values with a standard error of slightly less than ± 1.0 dB. The best-fit curve to the measures of C_{50} versus RT at 1 kHz was given by:

$$C_{50} = -20.83\ \mathrm{RT} + 7.020\ \mathrm{RT}^2 + 14.204 \pm 0.98\ \mathrm{dB}. \qquad (3)$$

Using Eqs. (2) and (3) with Fig. 9, one can determine combinations of RT and S/N(A) that should combine to produce a U_{50} of + 1.0 dB and hence correspond to 100% SI. Such calculations were performed in an iterative manner until a number of points describing the $U_{50} = +1.0$ contour were obtained. This contour is plotted in Fig. 8 for comparison with the contour derived from the consideration of S/N(A) and RT values. The result is a second contour that is quite similar to the first contour, and that differs most from the other contour at larger RT values in the range of 0.7–0.8 s. The two methods seem to combine background noise and reverberation effects in a similar form.

C. Comparison with STI

Houtgast[8] derived a similar contour based on the STI concept. His STI = 0.6 contour was said to separate "fair" conditions from "good" conditions. This contour is also plotted in Fig. 8, and it too is quite similar to the $U_{50} = +1.0$ contour. The STI = 0.6 contour is approximately 2 dB higher, except at larger RT values, where it differs by a greater amount. Thus the three contours suggest a somewhat similar trade-off between S/N(A) and RT values for preferred classroom acoustics with some disagreement for cases of lower background noise and larger RT values. For example, at S/N(A) values in the range of 15–20 dBA, the three contours specify optimum maximum RT values varying between about 0.7 and 0.9 s.

D. Combinations of RT and background levels

The results in Fig. 8 suggest that improved speech intelligibility should occur as RT is decreased to 0 s. This is because the beneficial effects of reverberation have been ignored, and anechoic conditions are certainly not optimum for a classroom. Reflected sound is needed for satisfactory speech intelligibility when the person speaking is not looking directly at a particular listener. Reverberation is also related to beneficial increases in speech levels in rooms. We really only want the beneficial increases caused by early reflections and not the addition of later reflections, which can only degrade conditions for speech. Unfortunately, in real rooms the provision of strong early reflections will lead to some later reflections and hence to a preferred condition with a nonzero reverberation time. Assuming the useful/detrimental ratio concept optimally combines the signal/noise and room acoustics factors, it is possible to estimate an optimum trade-off in terms of conventional acoustical quantities.

Calculations similar to those performed to obtain the $U_{50} = +1.0$ contour in Fig. 8 were carried out, after first calculating the reverberant field levels that would exist in a 300-m^3 room for each reverberation time, source, and background level. In this way contours for $U_{50} = +3$, $+1$, -1, and -3 dB were calculated and are plotted in Fig. 10. These can be regarded as equal speech intelligibility contours for this size of room. Values of $U_{50} \geqslant 3.0$ have been labeled "excellent," -1.0 to $+1.0$ "good," and $\leqslant -1.0$ "less satisfactory." All contours have maxima at RT values greater than zero. Thus there is an "optimum" RT in that at these maxima greater background noise can be tolerated for the same speech intelligibility. The combination of Knudsen and Harris's recommended RT of 0.7 s and the 36 dBA (NC 30) recommended by a number of sources comes close to the $U_{50} = +1.0$ contour and is seen to be acceptable using Fig. 10. However, the present results indicate that a smaller RT between 0.4 and 0.5 s is optimum in that more background noise can be tolerated in this range of RT values.

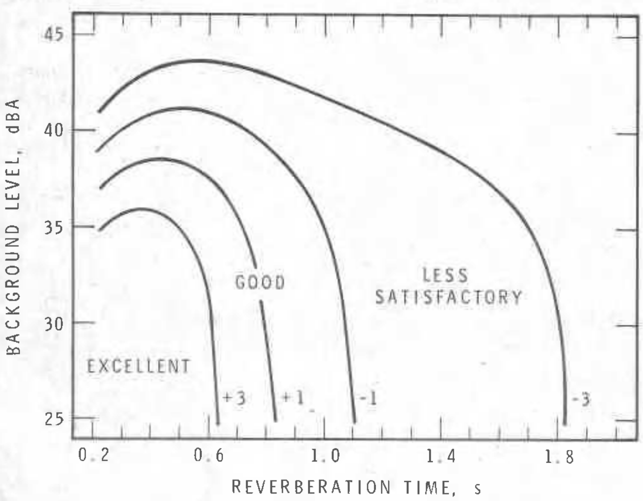

FIG. 10. Equal speech intelligibility contours in a 300-m^3 room based on calculated U_{50} values.

Plots similar to Fig. 10 can be produced for rooms of larger volumes. The curves would follow the same form and the scale on the vertical axis would vary as ten times the logarithm (to the base 10) of the room volume if one considers only positions in the reverberant field. Further adjustments should also be made to account for expected higher voice levels in larger rooms. Correcting only for volume, an acceptable background level of 30 dBA would become 20 dBA in a room of 3000 m^3. However, if one assumed that a "raised" voice level would always be used in this larger room, then from the data of Pearsons *et al.*[21] the voice level would be raised by an average of 7.0 dBA. Thus the 30-dBA background level for a "normal" voice level in a 300 m^3 room would be equivalent to a 27-dBA background level in a 3000-m^3 room with a "raised" voice level.

From the work of Nábělek and Robinson,[5] Elliott,[6] and others[23] we know that these conditions will not be optimum for all subjects. Although young adults may perform slightly better than the students of the present study, younger students and older adults will require better conditions than those found optimum for the 12- and 13-year-olds of the present study. Nábělek and Robinson[5] suggested that seven- and ten-year-old students would require a 5-dBA larger signal level. Elliott[6] suggested younger children require a 5-dB greater signal/noise. Nábělek and Robinson[5] and Neuman and Hochberg[23] showed that younger children were more strongly affected by reverberation. Thus, in rooms designed for younger children, the RT should fall between 0.4 and 0.5 s and the maximum background level should be 5 dBA lower than the values of Fig. 10. Where rooms are designed for general audiences including adults of various ages, again more stringent conditions are required. Thus overall optimum conditions for classrooms or other rooms of similar size intended for speech could comprise an RT of 0.4–0.5 s and a maximum background level of 30 dBA. All of the ten classrooms in this study had background noise considerably in excess of this optimum value, and only two rooms had RT values close to this optimum range.

IV. CONCLUSIONS

Although a number of published studies have demonstrated the importance of several newer acoustical measures such as early/late ratios, early decay times, and center time, in the smaller rooms that are representative of normal classrooms, the importance of such measures seems to be diminished, because many of these quantities could be predicted within an error of about ± 1.0 dB from the measured reverberation times. Such predictions could be made on an analytical basis to almost the same accuracy by assuming that decays were ideally continuous and exponential in form.

The 1-kHz U_{35}, U_{50}, and U_{95} values as well as the STI values were the most accurate predictors of speech intelligibility scores and were of essentially equivalent prediction accuracy. As both the Lochner and Burger type of useful/detrimental ratios (U_{95}) and the STI values are considerably more difficult to calculate, there seems to be no particular advantage to using them. The C_{35} values, and hence U_{35} values, were less accurately predictable from RT values than were U_{50} values. Much previous work has been based on measures incorporating a 50-ms early sound limit, and it would appear to be the preferred measure for evaluating classrooms and other similar sized rooms intended for speech. Because of the good approximate relationships between the newer measures and reverberation times, speech intelligibility scores can be predicted with only a little less accuracy from steady state signal-to-noise ratios and reverberation times for these smaller rooms, where they cannot be directly measured. The 1.0-kHz U_{50} values are a better and more universal measure for predicting speech conditions in such rooms, and an optimum of + 1.0 dB is required for very good speech intelligibility.

Preferred conditions for classrooms derived in this work must be further restricted to render them suitable for both much younger and much older audiences. Optimum reverberation times for classrooms were estimated to be in the range from 0.4–0.5 s, which is shorter than many standard references suggest. To accommodate all age groups of normal hearing listeners, background levels of approximately 30 dBA are required. The actual trade-off between background noise levels and reverberation times for optimum conditions took similar forms using three different calculation approaches.

The preferred variation of early/late ratios such as C_{50} with frequency remains to be determined, but low-frequency early/late values further influenced speech intelligibility scores. No attempt was made to consider the more stringent needs of hearing impaired subjects.

Further analyses of these data and that of a previous study are now available.[24]

ACKNOWLEDGMENTS

The helpful cooperation of the Ottawa Board of Education, and of various teachers and school officials, is gratefully acknowledged. This paper is a contribution from the Institute for Research in Construction, National Research Council of Canada.

[1] A. I. Bronzaft and D. P. McCarthy, "The Effect of Elevated Train Noise on Reading Ability," Environ. Behav. **7** (4), 517–527 (1975).

[2] S. Cohen, G. W. Evans, D. S. Krantz, D. S. Kelly, and S. Kelly, "Aircraft Noise and Children: Longitudinal and Cross-Sectional Evidence on Adaptation to Noise and the Effectiveness of Noise Abatement," J. Pers. Soc. Psychol. **40**, 331–345 (1981).

[3] J. S. Lukas, R. B. DuPree, and J. W. Swing, "Effects of Noise on Academic Achievements and Classroom Behaviour," State of California Report FHWA/CA/DOHS-81/01, Berkeley (September 1981).

[4] A. K. Nábělek and J. M. Pickett, "Reception of Consonants in a Classroom as Affected by Monaural and Binaural Listening, Noise, Reverberation, and Hearing Aids," J. Acoust. Soc. Am. **56**(2), 628–639 (1974).

[5] A. K. Nábělek and P. K. Robinson, "Monaural and Binaural Speech Perception in Reverberation for Listeners of Various Ages," J. Acoust. Soc. Am. **71**, 1242–1248 (1982).

[6] L. L. Elliott, "Effects of Noise on Perception of Speech by Children and Certain Handicapped Individuals," J. Sound Vib. **16**, 10–14 (1982).

[7] T. Houtgast, "The Effect of Ambient Noise on Speech Intelligibility in Classrooms," Appl. Acoust. **14**, 15–25 (1981).

[8] J. W. Sargent, M. I. Gidman, M. A. Humphreys, and W. A. Utley, "The Disturbance Caused to School Teachers by Noise," J. Sound Vib. **70**, 557–572 (1980).

[9] J. S. Bradley, "Predictors of Speech Intelligibility in Rooms," J. Acoust. Soc. Am. **80**, 837–845 (1986).

[10] J. S. Bradley, "Auditorium Acoustics Measures from Pistol Shots," J. Acoust. Soc. Am. **80**, 199–205 (1986).

[11] M. J. R. Lamothe and J. S. Bradley, "Acoustical Characteristics of Guns as Impulse Sources," Can. Acoust. **13**, 16–24 (1985).

[12] J. P. A. Lochner and J. F. Burger, "The Influence of Reflections on Auditorium Acoustics," J. Sound Vib. **1**, 426–454 (1964).

[13] H. J. M. Steeneken and T. Houtgast, "A Physical Method for Measuring Speech-Transmission Quality," J. Acoust. Soc. Am. **67**, 318–326 (1980).

[14] M. Barron and L.-J. Lee, "Energy Relations in Concert Auditoria," J. Acoust. Soc. Am., submitted for publication.

[15] L. L. Beranek, *Noise and Vibration Control* (McGraw-Hill, New York, 1971), p. 585.

[16] F. B. Stumpf, *Analytical Acoustics* (Ann Arbor Science, Ann Arbor, MI, 1980), p. 217.

[17] I. Sharland, *Woods Practical Guide to Noise Control* (Woods of Colchester, Great Britain, 1972), p. 33.

[18] P. H. Parkin and H. R. Humphreys, *Acoustics, Noise and Buildings* (Faber, London, 1969), p. 295.

[19] W. Burns, *Noise and Man* (John Murray, London, 1968), p. 118.

[20] V. O. Knudsen and C. M. Harris, *Acoustical Designing in Architecture* (Wiley, New York, 1965), p. 194.

[21] K. S. Pearsons, R. L. Bennett, and S. Fidell, "Speech Levels in Various Noise Levels," Bolt Beranek and Newman Inc., Report to U.S. EPA, Pg-270-053, Canoga Park, CA (May 1977).

[22] T. Finitzo-Hieber and T. W. Tillman, "Room Acoustics Effect on Monosyllabic Word Discrimination Ability for Normal and Hearing-Impaired Children," J. Speech Hear. Res. **21**, 440–458 (1978).

[23] A. C. Neuman and I. Hochberg, "Children's Perception of Speech in Reverberation," J. Acoust. Soc. Am. **73**, 2145–2149 (1983).

[24] J. S. Bradley, "Uniform Derivation of Optimum Conditions for Speech in Rooms," Building Research Note 239, National Research Council of Canada, Ottawa, Canada (1985).