THE EVOLUTION OF NEWER AUDITORIUM ACOUSTICS MEASURES

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

ABSTRACT
This paper reviews the development of newer types of auditorium acoustics measures that go beyond the conventional reverberation time. Quantities have been developed for speech and music. Two different types of speech measures have been shown to correlate well with speech intelligibility scores. Research has led to a consensus that a small number of quantities explain almost all of the variance in subjective assessments of concert hall acoustics quality. These new quantities involve only simple energy summations over various time intervals or the decay rates of sound energy and can be readily measured in halls.

1. INTRODUCTION
It has long been recognized that reverberation time is inadequate as a single objective descriptor of auditorium acoustics quality. There are clearly other dimensions to acoustic quality than reverberance, and acoustic quality often varies from seat to seat in a manner not reflected by changes in reverberation times. This paper reviews the development of newer types of objective parameters for both speech and music that go beyond reverberation time.

Large halls are extremely complicated physical systems, and to understand them in detail is a substantial challenge. We can readily devise a variety of complex physical parameters to describe conditions in such spaces, and many authors have done this. Such parameters are not of much practical value unless they relate to subjective evaluations of the same acoustical conditions. It is essential to carry out controlled subjective tests to systematically validate objective parameters in terms of subjective judgments. Because of the vagaries of subjective judgments and the physical complexity of large auditoria, large complex studies are required to make substantial progress in this field.

In spite of the complex problem, the newer parameters that will be described in this paper generally involve quite simple concepts of energy summations over various time intervals as well as the rate of decay of this energy in large rooms. In addition, these measurements have usually been made in only octave or wider frequency bands.

From a review of research in this field, this paper attempts to show how the many new parameters that have been proposed can be reduced to a small group of perhaps four or
five quantities that cover the major aspects of auditorium acoustics quality. Finally, the importance of assessing halls in terms of these new objective parameters is stressed.

2. EARLY REFLECTIONS

To understand the concept of early reflections and many of the newer parameters, it is essential to appreciate the characteristics of an impulse response in an auditorium. Figure 1 illustrates such a pulse response showing the direct sound, a gap followed by several discrete early reflections, and then many overlapping reflections in the reverberant decay. Thus, early reflections are reflections that arrive very shortly after the direct sound, but before the many overlapping reflections in the reverberant decay. Frequently these early reflections represent a significant portion of the total sound energy in the pulse response.

After the pioneering work of Sabine to develop an initial understanding of reverberation, procedures were gradually developed for the design and evaluation of auditoria in terms of optimum reverberation times. It was only after studies revealed the particular importance of early reflections that progress beyond reverberation time commenced. Initial investigations first demonstrated that early reflections are usually not heard as separate events, but are integrated subjectively together with the direct sound by our hearing system. Thus, early reflections can make the direct sound seem louder and hence increase the loudness and intelligibility of speech in rooms. This is often referred to as the Haas Effect.

Haas carried out experiments using speech in simulated sound fields with a direct sound and a single reflection. His subjects identified when the added reflection seemed equally loud to the direct sound. Figure 2 illustrates how Haas' results showed that early reflections could be as much as 10 dB stronger than the direct sound or arrive as much as 50 ms after the direct sound and still be subjectively integrated together with it. Lochner and Burger produced similar curves, that are also shown in Figure 2, and included a small dependence on speech level. Haas' and Lochner and Burger's results in Figure 2 were for judgments of the added reflection sounding equally loud to the direct sound. Also shown is a curve of the threshold of detectability of an added early reflection by Schodder. Finally, there is a curve showing an expected inverse square law drop off of the level of a reflection relative to the level of the direct sound in a typical auditorium. By comparing the subjective test results to this inverse square law curve, it is seen that useful early reflections can be much stronger than would normally occur naturally in a room. This, of
course, explains why well-designed sound reinforcement systems can often produce very natural sounding increases in sound level with the illusion that there is no electroacoustic reinforcement.

Studies with multiple early reflections indicated that their combined effect was related to the sum of their energies. Accordingly, Thiele developed a parameter that is the ratio of the sum of the early energy to the total energy in a pulse response. (This is later defined in Figure 5.) This parameter was called "Deutlichkeit" in German, and would translate to distinctness or definition in English. Thus, the greater the combined energy in the direct and early reflections, the greater the definition of the sound, due to the integration of the early reflections with the direct sound by our hearing system. Thiele made measurements in a number of rooms and generally showed that this was a useful new parameter nearly 40 years ago.

3. SPEECH INTELLIGIBILITY

The work of Lochner and Burger concentrated on the influence of early reflections on speech in rooms. They developed a measure that combined early reflections by a weighting scheme derived from extensive subjective tests. Figure 3 shows their results in terms of the fraction of an early reflection that would be judged useful to improved speech intelligibility. According to this procedure, the importance of a particular reflection depends on both its amplitude and arrival time relative to the direct sound. Using this weighting procedure, they summed up the effect of direct and early reflections as the combined useful sound. They then developed a useful/detrimental sound ratio as the ratio of this useful early sound energy to the detrimental energy formed by the sum of the later arriving speech energy and the ambient noise energy. Approximately 30 years ago, they showed these useful/detrimental sound ratios to be well correlated with speech intelligibility test results.

Due to the complexity of implementing their weighting procedure, few other studies have followed this approach. More recently, work by Bradley compared their procedure and others with speech intelligibility scores obtained in a range of rooms. This study found that a simplification of the Lochner and Burger concept, where the useful energy was a simple unweighted summation of the direct and early energy, was well correlated with speech intelligibility test scores. Thus, a simplified useful/detrimental sound ratio was derived as follows:

\[ U = 10 \cdot \log \left( \frac{\text{Early}}{\text{Late} + \text{Ambient}} \right), \text{dB} \]

where the early energy was summed over either 50 or 80 ms.

This measure combines both the room acoustics and signal/noise aspects of evaluating conditions for speech into
a single quantity, and includes the influence of subjectively important early reflections. (The conventional approach would be to separately measure reverberation times and background noise levels.)

More recently, another speech measure has been produced that again combines both aspects into a single quantity. Work in The Netherlands by Steeneken and Houtgast\textsuperscript{12} produced the speech transmission index, STI, and a simplification of it, the rapid speech transmission index, RASTI. It was reasoned that speech intelligibility in rooms is degraded because the acoustical properties of the room and existing ambient noise diminish the natural amplitude modulations of speech. Accordingly, the STI measure assesses modulation transfer functions for the 96 combinations of 6 speech frequency bands and 16 modulation frequency bands. From this matrix of values, the single STI value is derived as a number between 0 and 1.0 similar to the articulation index. Houtgast and Steeneken have published many papers on their work and have shown the STI measure to be well correlated with speech intelligibility scores for a wide range of conditions.

Bradley\textsuperscript{10} calculated STI values from impulse responses by a procedure proposed by Schroeder\textsuperscript{13}. It was found that STI values and useful/detrimental sound ratios were quite strongly correlated with each other (see Figure 4). Thus, in spite of their quite different derivation, the two quantities are apparently assessing essentially the same physical aspects in the rooms that were studied.

4. MEASURES OF MUSICAL CLARITY

Thiele's Deutlichkeit parameter (see Figure 5) was the first of a number of parameters based on ratios of energy sums from various parts of an impulse response. While such quantities have been developed as measures of speech intelligibility, they have also been used to assess conditions for music. They are usually thought to relate to clarity, definition or distinctness in the sound quality. Four of these quantities are compared in Figure 5, and these are: Thiele's Deutlichkeit, Schultz's running liveness, Reichardt's clarity measure, and Cremer's centre time. This figure includes the defining equations of each measure as well as a pictorial representation of the intervals over which energies are integrated in calculating each quantity.

While Thiele's Deutlichkeit is an early/total sound ratio, Schroeder\textsuperscript{14} somewhat later published results of measurements using an early/late arriving sound energy ratio. At about the same time, Schult\textsuperscript{15} proposed his running liveness measure that is the ratio of the late/early
arriving sound. Although the ratio is inverted, because it uses the same 50 ms early time interval, it would be exactly related to Deutlichkeit.

The use of a 50 ms early time interval seems to be a result of earlier studies concerning speech. Reichardt developed a clarity measure for music with an 80 ms early time interval from subjective tests in simulated sound fields. His clarity measure is a simple early/late arriving sound energy ratio and is included in Figure 5. Because the early time interval is different than the earlier measures, there will not be an exact relationship among these quantities, but measurements in halls have shown that C80 values are typically 2 to 3 dB greater than C50 values (where C80 and C50 are early/late arriving sound energy ratios with 80 and 50 ms early time intervals respectively).

Clearly, many different early time intervals could be proposed or a weighting scheme similar to that of Lochner and Burger could be used to blur the boundary between early and late energy. Cremer proposed that this problem could be solved by using the centre time as a clarity related measure.

Measurements in halls indicate that these quantities are usually highly correlated with each other and it is difficult to discriminate among them. Figure 6 plots hall mean values of Schultz's running liveness, R, versus Reichardt's clarity measure, C80, showing how these two quantities are related to each other. Figure 7 plots values of the centre time versus C80 values. Because the centre time is a linear quantity and C80 is a logarithmic quantity, the relationship is not linear, but the two measures are clearly closely related.

Eysholdt considered this same question by comparing correlations between subjective judgments using music signals and early/late ratios with varied early time interval. His results, shown in Figure 8, show a broad maximum in the range from 50 to 180 ms early time intervals. Again, it is not possible to find large differences among the various parameters. They all seem to assess the balance between clarity and reverberance in a more or less similar manner.

5. LATERAL REFLECTIONS

After it was established that early reflections are subjectively important, the direction of arrival of these reflections was next found to be important. Marshall and West both suggested that strong early lateral reflections were particularly important. Pioneering work by Barron showed that the sense of spatial impression or envelopment was determined by having adequate strong early lateral reflections. Figure 9 taken from Barron's work shows the region of delays and amplitudes of lateral reflections that lead to spatial impression. It is seen that early lateral reflections as much as 20 dB below the level of the direct sound contribute to perceived spatial impression. Very early reflections, arriving within less than 10 ms after the direct sound, lead to colouration of the perceived sound.
Longer delayed reflections, that fall to the right of the shaded area on Figure 9, are detected as discrete echoes and lead to disturbance rather than ideal spatial impression. If the amplitude of early reflections is too great, then there is a perceived image shift.

As shown in Figure 9, Barron also found that the degree of spatial impression was relatively independent of time delay. He also showed that spatial impression was greatest for reflections arriving at listeners' ears in the horizontal plane at 90 degrees to straight ahead (see Figure 10).

Ando\textsuperscript{23} has published many papers on subjective preference tests using simulated concert hall sound fields. His work also includes the effect of angle of incidence of early lateral reflections, and his results are reproduced here in Figure 11. Again, early reflections from straight ahead have the least effect, but for one of Ando's music motifs, preference was less at 90 degrees than at lower angles, leading Ando to conclude that there is a preferred angle of about 55 degrees to straight ahead. This conclusion would conflict with Barron's results in Figure 10 and depends on only one data point in Figure 11.

Barron's work led to the derivation of the lateral energy fraction, LF, as an objective parameter related to subjective spatial impression. It is simply the ratio of the early arriving energy from lateral directions to the total early arriving energy from all directions. The early time interval is taken as 80 ms. Ando has used the inter-aural cross correlation coefficient, IACC, as an objective correlate of spatial impression. Ando uses the maximum of the short-term cross correlation of the A-weighted signals obtained from a dummy head. Both the LF and the IACC have been shown to relate to subjective judgments of spatial impression and some authors\textsuperscript{22} have suggested that the two quantities are related to each other. However, measurements in real rooms\textsuperscript{24} shown in Figure 12 show no correlation at all between the two measures.

6. MULTI-DIMENSIONAL SUBJECTIVE STUDIES

Focussed laboratory studies can explore the importance of particular aspects of early reflections, but broad comprehensive studies are necessary to appreciate the overall picture. Two such major studies were undertaken in the early 1970's by research groups in Berlin and Göttingen in Germany. These studies looked at the overall multidimensional subjective evaluation of concert hall acoustics in combination with a large number of objective acoustical parameters. Figure 13 summarizes some details of these studies.

The Berlin group\textsuperscript{25,26} used dummy head recordings of a live orchestra in six different halls as their source of test sounds.
These recordings were played back to subjects via headphones. Subjects judged the sound fields by responding to 19 bipolar response scales. The overall strength or relative sound level, $G$, of the sound fields was found to correlate best with the major subjective response factor. The centre time, $TS$, and the slope of the early decay time versus frequency, $EDT(f)$, were also found to be important correlates of subjective judgments.

In the Göttlingen study, 27, 28, 25 halls were included by playing anechoic music recordings into the halls via two loudspeakers. The sounds were again recorded using a dummy head, but were played back to subjects using loudspeakers in an anechoic room. Cross talk filters were developed so that the sound recorded at one ear of the dummy head was only heard at one ear of the listener in the tests. Subjective judgments were of the overall preference of pairs of sound fields. Because the overall level or loudness was assumed to be of dominant importance, test sounds were adjusted so that they were all heard at the same loudness. Thus, there were no observed effects of the loudness or strength of the sounds. Deutlichkeit, $D$, reverberation time, $RT$, (or early decay time, $EDT$), and the inter-aural cross correlation coefficient, $IACC$, were found to be the strongest correlates of principal subjective dimensions. Because $RT$ and $EDT$ values were strongly correlated with each other for these halls, either was said to be an equivalent predictor of subjective preference.

Although the importance of $D$ in the Göttlingen study might seem to parallel the importance of $TS$ in the Berlin study, the sign of the correlation was reversed between the two groups. Thus, the subjects in the Berlin study preferred more clarity, but the Göttlingen subjects preferred less clarity. Subsequent investigations have attempted to explain this difference as due to a less reverberant character of the recorded sounds when the source was two loudspeakers and not a real orchestra. There can be many other criticisms of the details of these studies which just reflect how difficult it is to carry out this type of research as an ideal controlled experiment. Together they have given a much better picture of the overall problem. They have revealed that a relatively small number of subjective dimensions explain almost all of the variance in subjective assessments of concert hall sound and that only a relatively small number of objective parameters are necessary to describe acoustical conditions in concert halls.

A more recent study has been published by Barron 29 based on subjective assessments of conditions in 11 British concert halls during regular performances. His subjects filled in a short questionnaire after each concert and the responses were correlated with objective measurements at the same locations in these halls. Figure 14 summarizes how subjects' responses were inter-related, including the correlation coefficients between pairs of subjective responses. When all subjects were included in the analyses,
overall impression was found to be determined by perceived reverberance, envelopment, and intimacy. Perceived loudness appeared to influence overall impression via its effect on intimacy.

Barron found that his subjects could be grouped into two groups. One group seemed to prefer reverberance while the other preferred intimacy. Figure 15 shows that for the first group, judgments of reverberance and envelopment seemed to influence overall impression. For the second group, intimacy and envelopment judgments were most strongly correlated with overall impression.

Barron concluded that overall sound level or strength, early decay time, and lateral energy fraction values correlated with the principal dimensions of his subjects' responses. He also found that the perceived bass balance was related to either the ratio of low frequency to mid frequency EDT or TS values. His results were in reasonable agreement with those of the earlier German studies and his survey was an improvement over the German studies because of the more realistic presentation of the sounds. Thus, his subjects' evaluations of intimacy were influenced by both acoustical and non-acoustical factors present in actual halls. This complete concert experience is not entirely reproduced in laboratory studies. However, his survey was not as controlled an experiment as the German studies. His expert listeners were well aware that they were in a hall with a possibly good or bad reputation and their judgments could have been influenced accordingly.

7. A CURRENT CONSENSUS

Barron's British results combined with the earlier German studies lead to a reasonable consensus as to which objective parameters are most important. All studies agree that the overall sound level or strength, is important. A measure of clarity or definition is found to be important in each of these studies as is a measure of reverberance. Both Barron and the Berlin group found that the variation of early decay time with frequency influences the timbre of the perceived sound.

One can thus draw up a list of subjectively important acoustical parameters for evaluating conditions in concert halls and auditoria. These are listed in Figure 16. There are five different categories of parameters given but there is clearly some overlap between measures of clarity and measures of reverberance. At this point, it seems desirable...
to include both and to assume that they relate to different but related aspects of subjective impressions. Although reverberation time is not included, it is certainly an important quantity related to the physical parameters of the room and the statistical behaviour of sound in the room.

8. THE VALUE OF OBJECTIVE MEASUREMENTS

Thorough objective assessments are essential to evaluations of acoustical conditions in auditoria. Informal individual subjective assessments are not reliable and there is much evidence that such evaluations can be more influenced by obvious visual changes than by the often quite subtle acoustical changes. Small adjustments to halls may initially be thought to produce marked improvements to halls, but careful measurements usually reveal that small physical changes produce only small acoustical effects. In such cases, the initial perception of improvements usually soon leads to a sense of disappointment as the initial euphoria diminishes.

The newer parameters are expected to vary from seat to seat within a hall so that measurements must be made for a number of source-receiver combinations. This spatial sampling problem\(^3^0\), combined with the often short time available in halls, demands a measurement system that is not only comprehensive and precise, but also fast and efficient.

We have gradually developed improved measurement techniques to evaluate these newer parameters in auditoria. Recently, we have been using our RAMSoft measurement program\(^3^1^,^3^2\) that runs on an IBM PC computer interfaced to a Norwegian Electronics real time analyser. The program calculates 12 different acoustical parameters in a matter of seconds at each position in an auditorium. This system was a development from an earlier system\(^3^3\), and has been shown to produce results in close agreement with the quite different measurement systems of Gade\(^3^4\), and Barron\(^3^5\). Our measurements have been used to evaluate conditions in particular halls with acoustical problems and have been most useful in evaluating changes to halls. We are also accumulating data from a range of different halls\(^3^6\) so that we can learn what conditions exist in better halls and how the various acoustical parameters relate to the geometry and materials of the halls.

9. CONCLUSIONS

There is a reasonable consensus that a small number of objective parameters are required to explain the major portion of subjective assessments of auditorium acoustics quality. A list of five different parameters can be produced,
and all are based on quite simple energy summations or decay rates of sound energy. These five parameters have been shown to be correlated with different but perhaps partially overlapping subjective attributes of the acoustical quality of auditoria.

Efficient procedures exist to routinely obtain values of these parameters in halls, and studies in several countries have reported such measurements. These measurements can provide a clear objective assessment of acoustical conditions in an auditorium, and are certainly vastly superior to informal individual subjective judgments. By the systematic use of these newer parameters, we can gradually learn to better interpret their values and to make auditorium acoustics a more quantitative science. The addition of these newer quantities to more conventional considerations such as reverberation times and background noise levels, should help to make future consulting in this area a more robust and reliable business.

REFERENCES


ACOUSTICAL INTERFACE™
SYSTEM precision acoustical measurements
with your FFT, scope or meter

PS9200 POWER SUPPLY
- Dual Channel
- 9V “Radio” Battery
- Portable
- 50 Hours Operation
- Low Noise
- LED Status Indicator

7000 SERIES MICROPHONES
- Type 1 Performance
- ¼, ½ and 1 Inch Models

4000 SERIES PREAMPLIFIERS
- 2Hz to 200kHz ± 0.5db
- Removable Cable
- PS9200 and 7000 Series Compatible

NEW LOW COST
PRECISION MEASUREMENTS

- SINGLE CHANNEL SYSTEM UNDER $1,200
- DUAL CHANNEL SYSTEM UNDER $2,000
(¼ or 1 inch microphones)

ACO Pacific, Inc.
2604 Read Avenue
Belmont, CA 94002
(415) 595-8588

ACOUSTICS BEGINS WITH ACO