FURTHER INVESTIGATION INTO THE LOUDNESS OF RUNNING REVERBERATION

D. Griesinger

Lexicon, 100 Beaver Street, Waltham, MA 02154, dg@lexicon.com

1. INTRODUCTION

Many acousticians associate "spaciousness" and "spatial impression" with early reflected energy. (See Barron, Ando, Beranek, etc.) While such association is certainly correct in some cases, recent work (Morimoto and Plosselt, Gold, Soulodre, Gardner, Kahle, Griesinger) shows that spaciousness and envelopment are determined (mostly or in part) by sound energy arriving later than 100ms. (For example, Gold finds that early lateral reflections in a time range from 25-40ms added to a 2.0 second RT reverberant field REDUCE spatial impression.) In (8) we show that lateral energy causes source broadening for frequencies below 200 - 300Hz, and a feeling of being surrounded by an acoustic field for higher frequencies. This second sensation can be easily called spatial impression, spaciousness, or envelopment, but the arrival time is not critical for many types of music. When the music has sharp attacks the first 50ms can be separately distinguished, and the acoustic impression of this segment alone can be heard. However, in a concert hall the majority of the spaciousness and envelopment perceived comes from energy arriving later than this. For a soloist on stage, Gade has shown that any first order reflection arriving before about 75ms is inaudible, at least if it arrives in the medial plane. Thus musician self support must involve times later than this. In (2) we introduce the concept of reverberant loudness as a way of understanding these conflicting observations. reverberant loudness depends on the ratio of the loudness of an event stream which contains the musical notes - the direct stream - and the stream which represents the background - noise and reverberation. To understand reverberant loudness in halls, we must understand how the ear separates incoming sound into these two streams, and this is particularly difficult when the majority of the sound is reflected. The results of a series of experiments into these issues have, as usual, been surprising.

2. RUNNING REVERBERATION WHEN THE DIRECT SOUND IS HIGH

The work in (1) and (2) was derived from a study of soloist self support on concert stages. Thus we were at least originally working with high direct to reverberant ratios; 18dB or more for large concert halls, and at least 10dB for small spaces of equivalent reverberant loudness. Under these conditions we find that running reverberance (RR), the loudness of the reverberation while the music is playing, is the most important predictor of preference. Untrained subjects can reliably make comparisons of the loudness of different reverberant profiles. The ease and comfort with which these comparisons are made lead us to believe that reverberant loudness (a measure of the direct to reverberant ratio) is a built-in property of human hearing (1), (2). It is associated with a means of determining the relative distance of sound sources in an environment which contains reverberation and noise. To determine the direct to reverberant ratio the brain must first separate sound events into at least two streams, and the sorting of sound events into streams takes time - typically in our measurements on the order of 160ms. The time depends on the individual listener.

The data in (1) show that for short reverb times the reverb loudness depends on the pre-delay. We have now been able to study the masking of reverberation by music as a function of frequency and reverberant level.

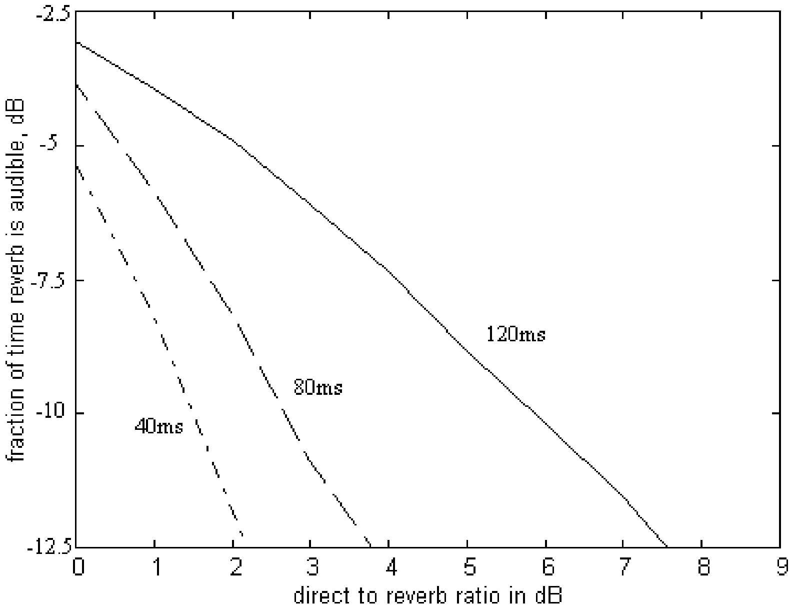

Figure 1 shows the fraction of the time reverberation in the 500Hz to 2kHz bands is audible as a function of direct to reverberant ratio for a 0.5s RT decay curve and three different pre-delays. The source is anechoic solo recorder music. If we look at a horizontal line of constant masking, we see that equivalent masking as a function of predelay depends on the level of the reverb. The lower the level, the greater the dependence on predelay. These curves match the data we obtained in (1) reasonably well.

RR is best understood as a ratio - the relative loudness of two different event streams; the stream associated with the direct sound (or the part of the sound which carries primary information) and the stream associated with the background. To make a measure we must decide how to measure both streams. Fortunately for solo self support the direct sound is dominant, and its level is a obvious choice for the level of the primary stream. From psychoacoustics we know loundess of a sound event can be approximated by a 160-200ms integration of the sound energy. We propose a 160ms integration as a measure for the direct stream. Experiments in (1), (2) show that for reverberant levels associated with self support on concert stages the level of the reverberant stream can be estimated by the energy in a 160ms window starting at 160ms. We define RR160 as a measure for RR on stages. It is the ratio of the energy in the second 160ms of the sound divided by the energy in the first 160ms.

$$RR = RR160 = 10 * \log 10\left(\frac{\int_{160ms}^{320ms} p(t)^2\, dt}{\int_{0}^{160ms} p(t)^2\, dt}\right)$$

We can convert RR160 to an equivalent slope of a decay curve, and in this case we get a measure which is quite similar to one first proposed by Schroeder and further developed by Kahle, who calls it EDT160. RR160 does not fit the data generated from 0.5sec RT decay curves with various pre-delays. It appears that this data is better matched when the effects of musical masking are included.

3. STAGE SUPPORT AND EDT

In room acoustics the time it takes for the sound to decay 10 decibels (multiplied by six to be comparable in magnitude to the reverberation time) is usually used to describe reverberant level. This measure is called the Early Decay Time, or EDT. EDT is measured from the Schroeder integrated impulse response, and thus is appropriate for music which contains mostly notes of long duration. We (and Kahle) prefer to measure EDT by determining the slope of a line drawn from the peak of the Schroeder integral to a point 10dB lower. This method is consistent with the original meaning of EDT as proposed by Jordan, and correctly accounts for the loudness of the direct sound. I will refer to it in this paper by EDT(-10). The method most widely used does a linear regression on all the points between the peak and -10dB, which underestimates the direct sound. I will call this EDTR. The difference between these measures in practice is small. In the domain of solo music and stage acoustics EDT fails badly as a measure of RR. Reverberant profiles with identical self support and RR can be found which will have very different EDT's.

4. RR FOR ORCHESTRAL MUSIC

Orchestral music is much more effective at masking reverberation. One effect of this is to make the reverberant loudness less dependent on individual physiology. Another effect is that much lower direct to reverberant ratios must be used before RR is audible at all, let alone comparable in loudness to reverb from solo music or speech. The steepness with which audibility decreases as direct to reverberant ratio increases is quite remarkable, and it makes the preferred reverberant level critical. The preferred level also depends strongly on decay time and predelay. To make matters more complicated, listeners are often well beyond the critical distance. The sound they hear is largely reverberant, and we must confront the question of how to measure the level of the primary stream in the absence of a dominant direct sound. Our choice will influence all measures which depend on time or level - of primary importance are clarity and EDT. Kahle makes the observation (private communication) that the sound in the Grosse Muiskverreinsalle in Vienna changes very little when you stand behind a pillar and the direct sound is blocked. From the point of view of perception we are primarily concerned with identifying the start time and end times of musical events (notes). The choice of start time is absolutely critical - measures such as EDT or Clarity become meaningless if the times chosen are based on aspects of the sound which are not audible or not important. We studied this problem with various artificial reverb profiles using speech or string quartet music as sources. In our limited experience there is no single measure for clarity which works in all cases. The cases roughly divide into groups. To identify the groups we first smooth the squared impulse response by averaging all points inside a 160ms window. We start with the window beginning at -160ms, so when the direct sound is high the peak value of the loudness curve is at 160ms, not at zero. The resulting curve represents the apparent loudness of a 160ms note as a function of time, and the level grows while the note is held. We choose as the apparent Start Time of the note (ST) a point 3dB before and below the peak level of this sliding integration. When the direct sound has more energy than the total reverberation and predelay is short, ST will be zero. The level of a point about 20ms into the integrated impulse response (the integrated energy from 0ms to 20ms) is a measure of the strength of the direct sound. If ST > ~20ms or if the peak of the curve is at 180ms or more, the peak value of the curve (which does not contain the direct sound) is an approximate measure of the strength of the reverberant energy. If the direct/reverb ratio measured this way is less than -10dB, then the direct is not particularly audible in music or speech. Measures for clarity, EDT, etc. should use ST as the start point and not the beginning of the direct sound. The equation to the left shows the method for Clarity.

$$C(x) = \frac{\int_{ST}^{x+ST} p(t)^2\,dt}{\int_{x+ST}^{\infty} p(t)^2\,dt}$$

If the direct/reverb ratio is higher than -5dB and ST < 50ms then the direct and the reverb will fuse, and measures should start with the direct sound; i.e. ST should be set to zero. If ST > 60ms, and the direct/reverb is > -10dB for speech, and >-5dB for string music, then an echo condition probably exists, and clarity will not be easily measured by simple integrations. These seats should be flagged. If the direct/reverb ratio is > -2dB, then current measures for Clarity, EDT, and reverberance will probably be OK. In this case measures should start with the direct sound and ST should be set to 0. In this case if ST > 100ms an echo condtion may exist.

5. RR WHEN THE DIRECT/REVERB RATIO IS > -2DB

We have done several reverberant level matching experiments on reverberant profiles where the direct/reverb ratio is -2dB or greater. The object is to find two different reverberant profiles which sound just as reverberant on orchestral music. It is often the case that such profiles are indistinguishable until the music stops. For example, figure 2 shows the Schroeder integrals from two such pairs. It is clear that in both cases the Schroeder integrals are equal at their peaks and at a point 380ms delayed.

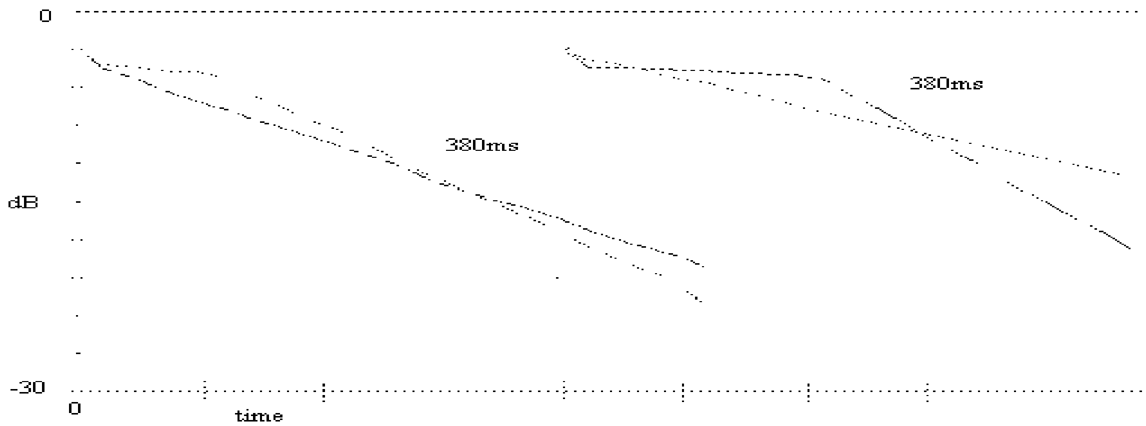

Figure 2: Two pairs of equally loud reverberation profiles - Schroeder integrals cross at ~380ms but not at -10dB

It has not been possible so far to find a simple measure which predicts the outcome of all these experiments. The best match comes directly from figure 2. It is a measure similar to EDT(-10), but based on time and not on level. We use the Schroeder integral as an input. The best match to our data occurs by comparing the level of the peak of the Schroeder integral to a point about 380ms delayed. We call this EDT380. The slope of a line connecting these points is converted to a reverberation time: For example here is EDT380. If S(0) is the value of the Schroeder integral at the peak, and S(380) is the value at 380ms, then

$$EDT380 = \frac{60 * 380ms}{(S(0) - S(380)) * 1000ms / \sec}$$

EDT380 can be extended to high reverberant levels by substituting S(ST) and S(380+ST), where ST is defined as above for Clarity. The first equally reverberant pair in figure 2 is a pair with one 2.2s RT exponential decay and one 1.8s decay with predelay. For this case it is clear that the point where the Schroeder integrals intersect is both at 380ms and at -10dB. For this pair EDT(-10) is just as good a measure as EDT380. The second pair shows that this close agreement is accidental. Here we have chosen a 3.8s RT and have paired it with a 1.5s RT with predelay. The curves still cross at 380ms, but this is at -6dB and not -10dB. EDT(-10) for these curves is 2.8s and 3.5s, although they sound equally reverberant. A third pair (not shown) using a direct/reverberant ratio of +4dB and RTs of 2.2s and 1.7sec with predelay give EDT(-10) values of 1.4 seconds and 1.6seconds, but EDT380 values of 1.6 and 1.5seconds. The optimum measure for this reverb level is EDT320. When the direct sound is clearly audible the reverb loudness is largely determined by the masking of reverberation by music. Figure 3 shows the masking of reverberation by three different types of music. Notice that for orchestral music the slope is very steep - a 1dB change in reverb level causes at least a 4dB change in masking. Only reverberation levels higher than 3dB direct to reverberant ratio are likely to be audible at all.

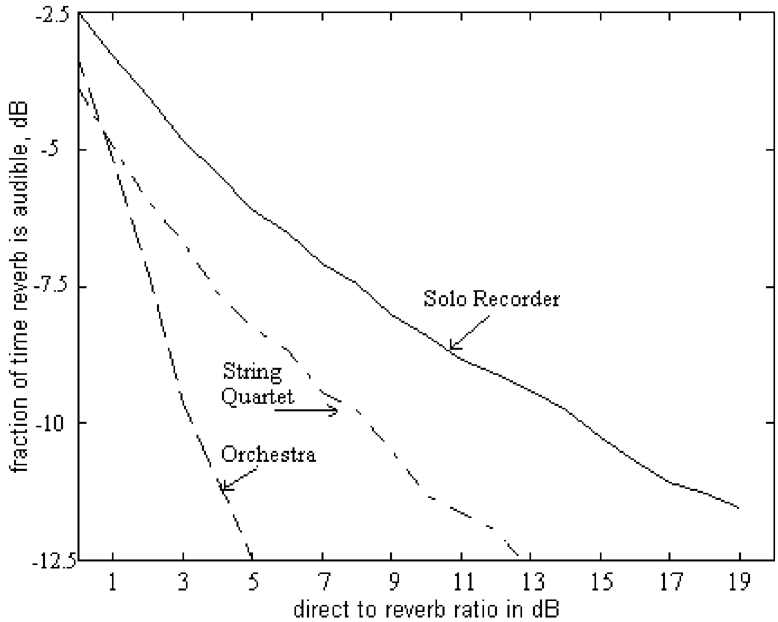

Figure 3: Fraction of the time reverb is audible for a 2.0s RT, zero predelay, for three different types of music.

Figure 4 shows the effect of changing the reverb time and predelay. Note that the higher the reverb time and the longer the predelay the lower the reverberant level needs to be to sound equally loud. Notice also however that as the direct to reverberant ratio decreases there is less dependence on both reverb time and predelay. These curves do not depend on physiology - although a time delay of 120ms before the detection of reverb is assumed. They are dependent on the musical composition and orchestration.

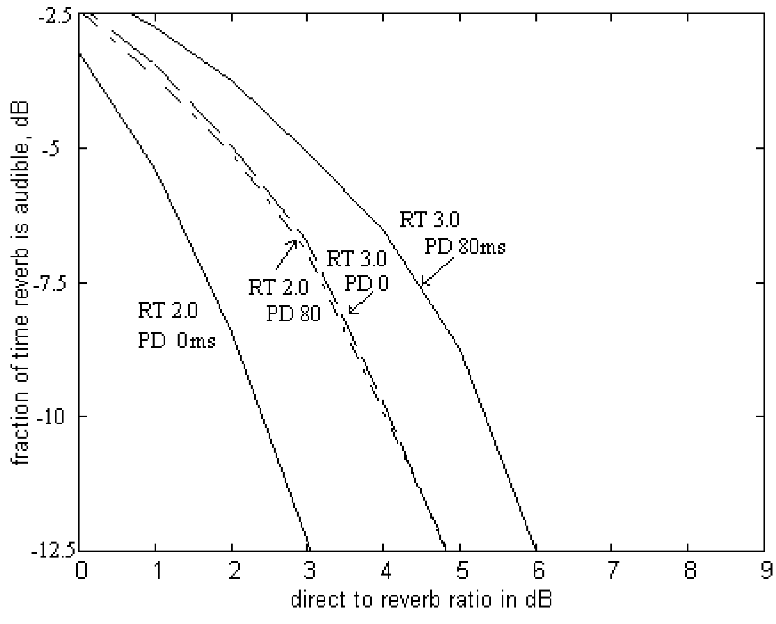

Figure 4: fraction of the time reverb is audible for orchestral music with three different reverb profiles

6. RR WHEN DIRECT/REVERB RATIO IS LESS THAN -2dB

We have a few occupied hall measurements in Boston Symphony Hall. The values of C40 (as measured conventionally with ST=0) are unfortunately corrupted by the directionality of the balloons used as sound sources, but by averaging several we can conclude that at least in the 1000Hz frequency range C(40) is about -2dB at row S with the source on the extended stage. In the front of the first balcony C(40) is in the range of -6 to -8dB. Thus in this hall a great number of the seats are quite far into the reverberant field. Does the concept of reverberant loudness - based as it is on the loudness of a background stream compared to the loudness of a foreground stream - make sense when the background stream has more energy than the foreground? As the direct sound becomes weaker in our occupied hall measurements the Schroeder integral becomes simply a straight line - the slope of which is given by the reverberation time. EDT(-10) is equal to EDT380 and all are equal to RT. There is no change as you move further back in the hall. Does this mean that all seats further back than row W in Symphony Hall Boston will sound equally reverberant?

Kahle has studied reverberance by looking for correlations between subjective impressions of reverberance during concerts to measurements made at the same seats in the unoccupied hall. One of his best measures (EDT160) is nearly identical to RR160 as defined above for stage acoustics. (The best match is for the 2kHz octave band.) However slightly better correlation occurs with a measure which is the ratio between the energy arriving in the first 40ms divided by the energy arriving later than 160ms. The success of this measure seems to indicate that loudness of the direct sound may be at least partially determined by the loudness of the first 40ms of a note. If Kahle is right, then the seat in the balcony - with more than 4dB less energy in the first 40ms - should sound substantially more reverberant. In this case EDT would fail as a measure and Kahle's would succeed. Does this seat in fact sound much more reverberant? In my experience it does not. We decided to perform some experiments.

7. THE APPARENT WIDTH OF REVERBERANT SOUND

An artificial reverberator was adjusted for a decay profile very similar to Boston Symphony - about a 1.6s decay slope for the first 150ms or so, and a 1.8s slope after that. Monaural anechoic speech and stereo string quartet were the primary sources. The reverberator was stereo - which means the left and right outputs are uncorrelated. Frequency contouring in the reverberator was turned off, so all frequencies had the same profile. Direct sound was mixed together with the reverb and presented to earphones in two ways - one way using a spherical head model so the reverberation would appear to come from +-90 degrees on either side of the listener, and one without the head model. The listener was asked to determine the apparent width of the reverberant sound field as the level of the direct sound was varied. Both methods gave the same variation with level, although they sounded different.

As we changed the direct sound the apparent width of the reverb varied, but not the way we expected. Pure reverb was perceived as coming from about +-60 degrees in front of the listener - even when no head model was used! As direct sound was added the reverb became wider. Spaciousness increased to a maximum when the direct/reverb ratio was equal to 0dB, and decreased on either side of this maximum. The apparent frontal localization for reverberation is probably an artifact of the experiment. It is likely that the earphones used were equalized for a more frontal response pattern than a diffuse sound pattern. That the experiment has this artifact is actually useful, because reverberation in the rear of a hall usually does seem to come from the front. As the direct energy increases this frontal cue becomes unrecognizable, and the apparent width increases. We then tried various profiles (by adding predelay and a few reflections in the 50ms to 100ms range.) The results appear to depend on the instructions given to the listener. In (1) and (2) the suggestion is made that the fundamental purpose of the neural mechanism for hearing running reverberation is determining distance. Thus distance and reverberance are closely linked. There is no question that as you decrease the level of the direct sound in the experiment the apparent distance increases. However when the direct sound becomes less than -10dB it is no longer particularly audible, and the sound can only be described as "distant". If you ask the subject to tell you about distance they will tell you about the ratio of the sound in the first 40ms to the total energy. However, if you ask them how spacious or enveloping they find the sound you get a different answer. Surprisingly, direct to total reverberant ratios of 0dB to -2dB give maximum envelopment, and the 1.8sec reverberation with no direct sound at all is not maximally spacious. It appears to come from the front more than from any other direction.

As direct sound is added, the reverberation becomes wider, reaching a maximum width when C40 ~=0. This width is independent of predelay, although the apparent loudness of the reverberance is not. The direct sound increases the loudness of the reverberance by providing a reference with which any predelay in the reverberation can be perceived. These results suggest the two seats in Boston would sound different in apparent distance, but similar in reverberance. If there was a difference, row S might sound more spacious than the balcony. Measures of these curves are particularly dependent on the choice of ST. Table 1 gives the measurements both as they normally are made, and with ST as defined above. The first entry in each column is the old measure, and the second is with ST. The results are dramatically different. The trials are in order of increasing reverberance. Note that without the direct sound or without predelay EDT = EDT380 ~= RT. This is as we expect with a reasonably exponential decay. However adding predelay and then some direct sound triggers the measuring system before the onset of the main reverberant energy, and EDT and EDT380 values increase dramatically in the normal measures. Trial #4 and #5 are quite close in sound. In spite of the large difference in C40 the apparent distance of trials 3 and 4 is nearly identical. The second set of clarity values are also calculated using ST. If these curves represented real seats in a hall and we used the old method of calculating EDT and clarity, averages of both over the whole hall would make the hall appear much more reverberant than it actually is. Notice to get reasonable values for Clarity one must use the new C40 and C80 when the direct/reverb is -5dB or less, and the old when it is higher.

Table 1:

trial	EDT(-10)		EDT380		C40		C80		RT(5-35)	Predelay	ST	
1	1.7	1.7	1.7	1.7	-3dB	-3dB	0dB	0dB	1.8	-	0	No direct sound
2	2.2	1.8	2.1	1.7	-8dB	-4dB	-7dB	-.3dB	1.8	100ms	100ms	Direct -8dB
3	1.6	1.6	1.6	1.7	-1dB	-1dB	+1dB	+1dB	1.8	10ms	0	Direct -5dB
4	2.1	1.9	2.0	1.8	-5dB	-9dB	-3.6dB	-1.8dB	1.8	100ms	60ms	Direct -5dB
5	2.0	1.9	2.0	1.8	-2dB	-10dB	-1.0dB	-2.3dB	1.8	100ms	40ms	Direct -2dB

9. RR AND EDT IN PRACTICE

We have tried EDT(-10), EDT(380), and EDTR on data sets from a number of occupied halls, including Boston Symphony Hall, Portland Maine Civic Auditorium, and Kresge Auditorium at MIT. In nearly every case EDT(-10) and EDT380 agree closely. Where EDT380 and RT disagree, the direct sound is high, and EDT380 and EDT(-10) predict the subjective impression of the seat better than RT. Portland has an average EDT380 in the range of 2.2 seconds, Boston has EDT380 in the 1.7s range, and Kresge in the 1.5s range. These values correspond well to the perceived reverberance. The sound is Portland is highly enveloping and reverberant, even in seats as close to the orchestra as row J on the floor. It is also clear. Boston is at the low end of optimal reverberance for my taste, and Kresge is too low, even for string quartet. The data sets include few seats where the direct sound is very low, so the good results could perhaps be expected. Under these conditions it appears that for practical purposes both EDT(-10) and EDT380 work pretty well as measures of hall reverberance. We have also compared EDT(-10) and EDT measured with linear regression, EDTR. When the direct/reverberant ratio is +2dB or less EDT(-10) and EDTR are nearly always identical. As mentioned before when direct/reverb is +4 they differ by 0.2 seconds, but then neither is a particularly good measure of the reverberance. In this case EDT320 is the best, and EDT380 is next best. It appears that EDT(-10) and EDTR agree except when both are broken. An observation from this work is that when the direct sound is strong - equal to or greater than the total reverberation - reverberant loudness is one of the major factors in preference. However as we move back in the hall the reverberant loudness (as measured by EDT380) becomes constant and is determined primarily by the reverberation time. For these seats other factors will determine preference, most likely intelligibility or clarity. Measuring reverberance, clarity, and intelligibility in highly reverberant fields is much trickier than is currently appreciated. When the direct sound is strong or when there is high musical intelligibility due to the shape of the reverberant profile there is a preference by listeners for EDT380 values of 1.9 seconds or more for orchestral music. These values for EDT are longer than many authorities currently recommend. The masking data indicate that skimping on EDT just a little can have disastrous consequences in reverberant audibility.

10. CONCLUSIONS

For halls with reverberation times in the range of 1.8 to 2.2 seconds and direct/reverberant ratios of +2dB to -2dB EDT(-10) appears to be an adequate measure of hall reverberance. EDT380, or perhaps EDT350, works over a much wider range of reverberation time and level. For reverb levels typical of stage acoustics RR160 is a better choice. Spaciousness appears to be a maximum when the direct energy and the reverberant energy are about equal. Where the direct sound is low (for many seats in large halls) conventional methods of calculating Clarity and EDT need to be modified to accommodate the apparent starting time of musical notes.

11. ACKNOWLEDGMENTS

Many people have contributed to the ideas in this paper. I would like to particularly thank Bill Gardner, Dick Campbell, and Eckhard Kahle for their ideas, data, and comments through conversations and email. I also want to thank Lexicon for continuing to support this work.

12. REFERENCES

[1] W GARDNER ‘Reverberation Level Matching Experiments’ Proceedings of the Sabine Memorial Conference, MIT June 5-7 1994 p263 Available from the Acoustical Society of America, 500 Sunnyside Blvd. Woodbury, NY 11797 This reference is highly recommended

[2] D GRIESINGER ‘Subjective loudness of running reverberation in halls and stages’ ibid p89

[3] E KAHLE J P JULLIEN ‘Some new considerations on the subjective impression of reverberance and its correlation with objective criteria’ ibid p239

[4] M A GOLD ‘Subjective evaluation of spatial impression: the importance of lateralization’ ibid p97

[5] G A SOULODRE J S BRADLEY ‘The influence of late arriving energy on concert hall spatial impression’ ibid p101

[6] M MORIMOTO K IIDA K SAKAGAMI A H MARSHALL ‘Physical measure for auditory source width (ASW): part 1 and 2’ ibid p109

[7] D GRIESINGER ‘Progress in Electronically Variable Acoustics’ ibid p57

[8] D GRIESINGER ‘Quantifying Musical Acoustics through Audibility’ Knudsen Memorial Lecture, Denver ASA meeting, 1993. Copies available from the author

[9] Y ANDO ‘Subjective preference in relation to objective parameters of music sound fields with a single echo’ JASA vol 62 p1436-1441 (1977)

[10] M BARRON ‘Spatial Impression due to Early Lateral Reflections in Concert Halls: The Derivation of a Physical Measure’ J. Sound and Vibration (1981) 77(2), 211, 232

[11] M MORIMOTO, C POSSELT ‘Contribution of Reverberation to Auditory Spaciousness in Concert Halls’ J. Acoust. Soc. Jpn (E)10, 2 (1989)

[12] L BERANEK ‘Concert Hall Acoustics - 1992’ J. Acoust. Soc. Am., Vol 92, No. 1, July 1992

[13] A C GADE ‘Investigations of musicians’ room acoustic conditions in concert halls’ Acoustica 69 p193-203 p249-262