



The Relationship between Audience Engagement and the ability to Perceive Pitch, Timbre, Azimuth and Envelopment of Multiple Sources

David Griesinger

Consultant, 221 Mt Auburn St. #504, Cambridge, MA 02138, USA

PACS: 43.55.Fw, 43.55.Mc, 43.66.Ba, 43.66.Hg, 43.66.Jh, 43.66.Qp

ABSTRACT

It is well known in psychology that sounds perceived as close to a listener engage our attention, while more distant sounds can be ignored. But the physical properties of sound that lead to engagement are poorly understood. The phenomenon is well known by directors of cinema and drama, who demand that theatres and sound systems have high presence, and not just adequate intelligibility. This is achieved by bringing the audience close to the performers, attenuating both early and late reflections, and using highly directional loudspeakers for cinema. When sounds are perceived as distant it is not just engagement that suffers. Precise localization of instruments in ensembles becomes impossible, and instrumental voices – distinct in a historic hall – become blended into mere harmonies. Historic concert venues tended to be small, with sound absorbing stages and audience areas. Sonic intimacy and clarity of musical lines was taken for granted. But the physical properties that make a sound seem close to a listener are poorly understood, and there is no current acoustic measure for specifying the type of clarity that provokes audience attention. An overly reverberant, “well blended” sound has become the design goal for modern chamber music and orchestral venues, even though a few great halls demonstrate that it is possible to have both clarity and reverberation at the same time over a wide range of seats. This paper describes the physics of sound that provides engagement, and how reflections interact to reduce it. An impulse response based measure for engagement is described, along with a means of determining the degree of engagement from live recordings of speech. Time and equipment permitting, the author will demonstrate that the right combination of direct sound and reflections can provide engagement, localization, reverberation, and envelopment at the same time. The improvement in sound quality is dramatic.

INTRODUCTION

This talk and its associated demonstrations are centered on the properties of sound that promote *engagement* – the focused attention of a listener. Engagement is related to the perception of distance to a sound source, a perception usually dominated by vision. But everyone can aurally sense the difference between “near” and “far” – although the perception is usually subconscious. In some art forms the phenomenon is well known: drama and film directors insist that performance venues be acoustically dry, with excellent speech clarity and intelligibility. Producers and listeners of popular music, and customers of electronically reproduced music of all genres, also expect – and get – recordings and sound systems that demand our attention. Ipod listening provides clarity in abundance.

Engagement is associated with clarity of sound, but currently there is no standard method to quantify the acoustic properties that promote it. Acoustic measurements such as “Clarity 80” or C80, were developed to quantify intelligibility, not engagement. C80 considers all reflections that arrive within 80ms of the direct sound to be beneficial. As we will see, this is not what engagement requires. Venues often have adequate intelligibility – particularly for music – but poor engagement. But since engagement is subconscious, and

reverberation is not, acoustic science has concentrated on sound decay – and not on what makes sound exciting. Acoustic engineers need methods to specify and verify the properties they are looking for. We desperately need measures for the kind of clarity that leads to engagement. The work in this talk attempts to fill this gap by looking at the physics and neurology of sound detection.

Specifically we seek to understand how our ears and brain extract precise information on the pitch, timbre, horizontal localization (azimuth), and distance of multiple sound sources at the same time. Such abilities would seem to be precluded by the structure of the ear, and the computational limits of human neurology.

We present the discovery that this information is encoded in the phases of upper harmonics of sounds with distinct pitches, and that this information is scrambled by reflections. Human neurology is acutely tuned to novelty, so reflections at the onsets of sounds are critically important. If the brain can detect and decode the phase information in the onset of a sound – before reflections obscure it – pitch, azimuth and timbre can be determined. The sound, although physically distant, is perceived as psychologically close. We will show mechanisms by which the ear and brain can detect pitch, timbre, azimuth and distance by analyzing the information

that arrives in a 100ms window after the onset of a particular sound event.

The significance of this discovery for these papers is that phase information is scrambled predictably and quantifiably by early reflections. Given a binaural impulse response or a recording of a live performance the degree of harmonic phase coherence in a 100ms window can be used to measure the degree of engagement at a particular seat.

We will also present some of the experiences and people that taught me to perceive and value engaging sound, and some of the reasons a few well known halls are highly valued. Together these experiences become a plea for hall designs that deliver excitement and clarity along with reverberation.

With luck we will also *demonstrate* the dramatic effects on clarity and engagement that result when the direct sound is clearly audible in the midst of a reverberant field. We will demonstrate the ease with which humans can detect the distance of a sound source from the phase coherence of high frequency harmonics. We will show how the perception of “near” and “far” is associated with the ability to localize sounds in a reverberant field, and how abruptly localization disappears when the D/R falls below a critical level. We have found that these effects can be demonstrated to a large audience if the room is not reverberant, and loudspeakers reproducing a reverberant field can be set up around the listeners.

Objective data obtained from analysis of binaural recordings made by the author in a variety of venues will be presented, and hopefully it will be possible to hear some of the differences the analysis reveals. In many cases the differences in engagement can be clearly heard – at least when the playback room is sufficiently non-reverberant.

“NEAR”, “FAR” AND LOCALIZATION

The perception of engagement and its opposite, muddiness, are related to the perception of “near” and “far”. For obvious reasons sounds perceived as close to us demand our attention. Sounds perceived as far can be ignored. Humans perceive near and far almost instantly on hearing a sound of any loudness, even if they hear it with only one ear – or in a single microphone channel. An extended process of elimination led the author to propose that a major cue for distance- or near and far – was the phase coherence of upper harmonics of pitched sounds. [1], [2]. More recent work on engagement – as distinct from distance – led to the realization that engagement was linked to the ability to reliably perceive azimuth, the horizontal localization of a sound source. For example, if the inner instruments in a string quartet could be reliably localized the sound was engaging. When (as is usually the case) the viola and second violin could not be localized the sound was perceived as muddy and not engaging. Engagement is usually a subconscious perception, and is difficult for subjects to identify. Localization experiments are easier, so I used localization as a proxy for engagement.

Direct sound, Reflections, and Localization

Accurate localization of a sound source can only occur when the brain is able to perceive the direct sound – the sound that travels directly from a source to a listener – as distinct from later reflections. Experiments by the author and with students from several universities discovered that the ability to localize sound in the presence of reverberation increased dramatically at frequencies above 700Hz, implying that localization in a hall is almost exclusively perceived through harmonics of tones, not through the fundamentals. Further experiments led to an impulse response based measure that

predicts the threshold for horizontal localization for male speech [3][4]. The measure simply counts the nerve firings that result from the onset of direct sound above 700Hz in a 100ms window, and compares that count with the number of nerve firings that arise from the reflections in the same 100ms window.

$$S = 20 - 10 * \log \int_{.005}^{\infty} p(t)^2 dt$$

LOC in dB =

$$S - 1.5 + 10 * \log \int_0^{.005} p(t)^2 dt - (1/D) * \int_0^{D-.005} POS(S + 10 * \log \int_{.005}^{\tau} p(t)^2 dt) d$$

In the equations above S is a constant that establishes a sound pressure at which nerve firings cease, assumed to be 20dB below the peak level of the sum of the direct and reverberant energy. $p(t)$ is an impulse response measured in the near-side ear of a binaural head. $p(t)$ is band limited to include only frequencies between 700Hz and 4000Hz. LOC is a measure of the ease of localization, where $LOC = 0$ is assumed to be the threshold, and $LOC = +3dB$ represents adequate perception for engagement and localization. POS means positive values only. D is the 100ms width of the window.

The first integral in LOC is the log of the sum of nerve firings from the direct sound, and second integral is the log of the sum of nerve firings from the reflections. The parameters in the equation (the choice of 20dB as the dynamic range of nerve firings, the window size D , and the fudge factor -1.5) were chosen to match the available localization data. The derivation and use of this equation is discussed in [3]. The author has tested it in a small hall and with models, and found it to accurately predict his own perception. Similar results have been obtained by professor Omoto at the University of Kyushu.

MEASURING ENGAGEMENT AND LOCALIZATION WITH LIVE MUSIC - THE IMPORTANCE OF PHASE.

The equation for LOC presented above requires binaural impulse responses from fully occupied halls and stages. These are extremely difficult to obtain. The author has struggled for some time to find a way to measure both localization and engagement from binaural recordings of live music. It ought to be easy to do – if you can reliably hear something, you can measure it. You just need to know how!

In the process of trying to answer this question, the author came to realize that the reason distance, engagement, and localization are related is that they all arise from the same stream of information: *the phase relationships of harmonics at the frequencies of speech formants.*

Perplexing Phenomena of Hearing

Human hearing uses several ways of processing sound. The basilar membrane is known to be frequency selective, and respond more or less logarithmically to sound pressure. With the help of sonograms much has been learned about speech perception. But these two properties of hearing are inadequate to explain our extraordinary ability to perceive the complexities of music – and our ability to separate sounds from several simultaneous sources.

For example, the frequency selectivity of the basilar membrane is approximately 1/3 octave (~25% or 4 semitones), but musicians routinely hear pitch differences of a quarter of a semitone (~1.5%). Clearly there are additional frequency selective mechanisms in the human ear.

Additional unexplained phenomena are numerous. The fundamentals of musical instruments common in Western music lie between 60Hz and 800Hz, as do the fundamentals of human voices. But the sensitivity of human hearing is greatest between 500Hz and 4000Hz, as can be seen from the IEC equal loudness curves. In addition, analysis of frequencies above 1kHz would seem to be hindered by the maximum nerve firing rate of about 1kHz. Even more perplexing, a typical basilar membrane filter above 2kHz has three or more harmonics from each voice or instrument within its bandwidth. And yet if we listen to a band limited signal at these frequencies we can easily hear the pitches of each instrument. How can we possibly separate them? Why has evolution placed such emphasis on a frequency range that is difficult to analyze directly, and where several sources seem to irretrievably mix? In a good hall, and in nearly any recording, I can detect the azimuth, pitch, and timbre of three or more musicians at the same time, even in a concert where musicians such as a string quartet subtend an angle of +/-5 degrees or less! (The ITDs and ILDs at low frequencies are miniscule.) Many concert halls prevent me from hearing the inner voices of a quartet, although I can hear the harmonies.

As a further example, the hair cells in the basilar membrane respond mainly to negative pressure – they approximate half-wave rectifiers, which are strongly non-linear devices. How can we claim to hear distortion at levels below 0.1% ?

Why do so many creatures – certainly all mammals – communicate with sounds that have a defined pitch? Is it possible that pitched sounds have special importance to the separation and analysis of sound?

Answers to these perplexing properties of hearing become clear with two basic realizations:

1. *The phase relationships of harmonics from a complex tone contain more information about the sound source than the fundamentals.*
2. *And these phase relationships are scrambled by early reflections.*

For example: my speaking voice has a fundamental of 125Hz. The sound is created by pulses of air when the vocal chords open. All the harmonics arise from this pulse of air, which means that exactly once in a fundamental period all the harmonics are in phase.

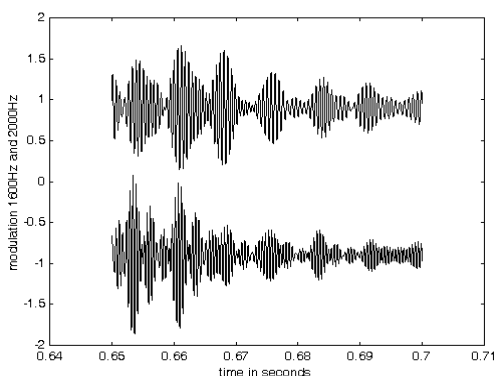


Figure 1: Top trace: The motion of the basilar membrane at a region tuned to 1600Hz when excited by a segment of the word “two”. Bottom trace: The motion of a 2000Hz portion of the membrane with the same excitation. The modulation is different because there are more harmonics in the higher frequency band. In both bands there is a strong (20dB) amplitude modulation of the carrier, and the modulation is largely synchronous between the two bands.

A typical basilar membrane filter at 2000Hz contains at least four of these harmonics. The pressure on the membrane is a maximum when these harmonics are in phase, and reduces as they drift out of phase. The result is a strong amplitude modulation in that band at the fundamental frequency of the source. When this modulation is below a critical level, or noise-like, the sound is perceived as distant and not engaging.

Amplitude Modulation

The motion of the basilar membrane above 1000Hz as shown in figure 1 appears to be that of an amplitude modulated carrier. Demodulation of an AM radio carrier is achieved with a diode – a half-wave rectifier – followed by a low pass filter. Although the diode is non-linear, radio demodulation recovers linear signals, meaning that sounds in the radio from several speakers or instruments are not distorted or mixed together. A similar process occurs when the basilar membrane decodes the modulation induced by the phase relationships of harmonics. Harmonics from several instruments can occupy the same basilar region, and yet the modulations due to each instrument can be separately detected.

Both in an AM radio and in the basilar membrane the demodulation acts as a type of sampling, and alias frequencies are detected along with the frequencies of interest. In AM radio the aliases are at high frequencies, and can be easily filtered away. The situation in the basilar membrane is more complicated – but can still work successfully. This issue is discussed in [3].

Figure 2 shows a model of the basilar membrane which includes a quasi-linear automatic gain control circuit (AGC), rather than a more conventional logarithmic detector. The need for an AGC is discussed in [3], but in other ways the model is fairly standard. The major difference between the model in figure 2 and a standard model is that the modulations in the detected signal are not filtered away. They hold the information we are seeking.

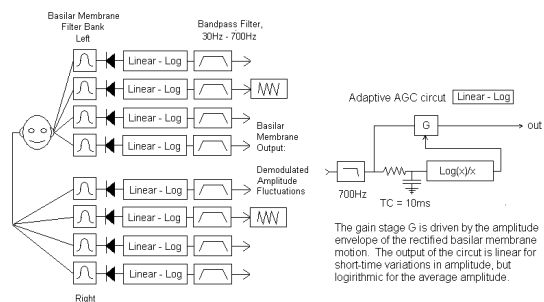


Figure 2: A basilar membrane model based on the detection of amplitude modulation. This model is commonly used in hearing research – but the modulation detected in each band is normally not considered important.

There is one output from figure 2 for each (overlapping) frequency region of the membrane. We have converted a single signal – the sound pressure at the eardrum - into a large number of neural streams, each containing the modulations

present in the motion of basilar membrane in a particular critical band.

How can we analyze these modulations? If we were using a numeric computer some form of autocorrelation might give us an answer. But autocorrelation is complex – you multiply two signals together – and the number of multiplications is the square of the number of delays. If you wish to analyze modulation frequencies up to 1000Hz in a 100ms window more than 40,000 multiplies and adds are needed

I propose that an analyzer based on neural delay lines and comb filters is adequate to accomplish what we need. Comb filters are capable of separating different sound sources into independent neural streams based on the fundamental pitch of the source, and they have high pitch acuity. Comb filters have interesting artifacts – but the artifacts have properties that are commonly perceived in music. A comb filter with 100 sum frequencies in a 100ms window requires no multiplies, and only 2000 additions. The number of or taps (dendrites) needed is independent of the delay of each neuron – which means in this model that the number of arithmetic operations is independent of the sample rate.

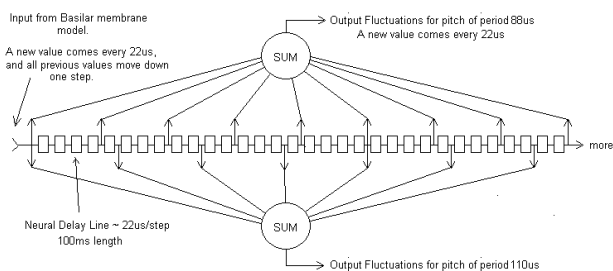


Figure 3: A comb filter analyzer showing two tap periods, one a period of four neural delay units, and one of five neural delay units. In human hearing such a delay line would be 100ms long, and be equipped with perhaps as many as 100 tap sums, one for each frequency of interest. There is one analysis circuit for each overlapping critical band. I have chosen a sample rate of 44.1kHz for convenience, which gives a neural delay of 22us.

Figure 3 shows the analyzer that follows the basilar membrane circuit in the author’s model. The analyzer is driven by the amplitude modulations created by the phase coherence of harmonics in a particular critical band. When the fundamental frequency of a modulation corresponds to the period of one of the tap sums, the modulations from that source are transferred to the tap sum output, which becomes a neural data stream specific to that fundamental. The analysis circuit separates the modulations created by different sound sources into independent neural streams, each identified by the fundamental frequency of the source.

If we use a 100ms delay window and plot the outputs of the tap sums as a function of their frequency, we see that the analyzer has a frequency selectivity similar to that of a trained musician – about 1%, or 1/6th of a semitone.

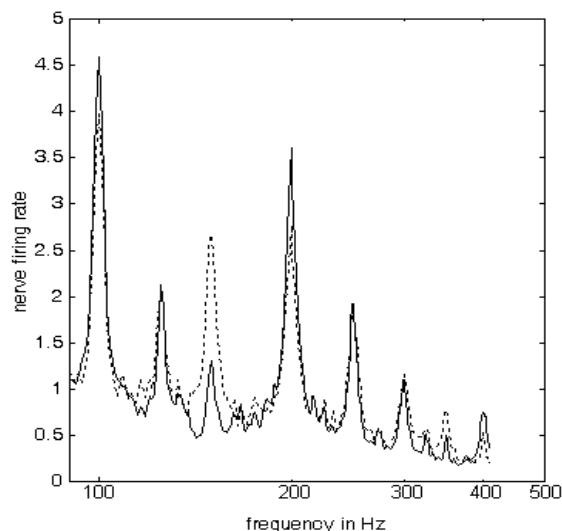


Figure 4: The output of the analysis circuit of figure 3 after averaging the tap sums of six 1/3 octave bands from 700Hz to 2500Hz. Solid line: The modulations created by the harmonics of pitches in a major triad – 200Hz, 250Hz, and 300Hz. Dotted line: The modulations created by harmonics of the pitches from the first inversion of this triad – 1500Hz, 200Hz, and 250Hz. Note the patterns are almost identical, and in both cases there is a strong output at the root frequency (200Hz) and its subharmonic at 100Hz.

Figure 4 shows one of the principle artifacts – and musical advantages – of the comb filter used as an analyzer. The advantage is that the comb filter inherently repeats triadic patterns regardless of inversions or octave, and produces similar output patterns for melodies or harmonies in any key.

The reason for this advantage – and a possible disadvantage – is that the tap sums are equally sensitive to the frequency corresponding to their period and to harmonics of that frequency. In practice this means that there is an output on a tap sum which is one octave or more below the input frequency. The subharmonic is not perceived, which suggests that the perception is inhibited because of the lack of output from a region of the basilar membrane sensitive to this fundamental frequency (in this case 100Hz).

The comb filter analyser is composed of simple neural elements: nerve cells that delay their output slightly when excited by an input signal, and nerve cells that sum the pulses present at their many inputs. The result is strong rate modulations at one or more of the summing neurons, effectively separating each fundamental pitch into an independent neural stream.

Not only is the fundamental frequency of each pitch at the input determined to high accuracy, once the pitches are separated the amplitude of the modulations at each pitch can be compared across critical bands to determine the timbre of each source independently.

The modulations can be further compared between the two ears to determine the interaural level difference (ILD) and the interaural time delay (ITD). The ILD of the modulations is a strong function of head shadowing, because the harmonics which create the modulations are at high frequencies, where head shadowing is large. This explains our abilities to localize to high accuracy, even when several sources subtend small angles.

Simple experiments by the author have shown that humans can easily localize sounds that have identical ITD at the onset

of the sound, and identical ILDs, but differ in the ITD of the modulations in the body of the sound, even if the bandwidth of the signal is limited to frequencies above 2000Hz. A demonstration of this ability using pink noise is on the author's web-site.

WHY THE HEARING MODEL IS USEFUL

The hearing model presented here need not be entirely accurate to be useful to the study of acoustics. The most important aspect of the model is that it demonstrates that many of the perplexing properties of human hearing can be explained by the presence of information in harmonics above 700Hz, that this information can be extracted with simple neural circuits, and that this information is lost when there are too many reflections.

Our model detects and analyses modulations present in the motion of many overlapping regions (critical bands) on the basilar membrane. Although the detection process is non-linear, as in AM radio the modulations themselves are (or can be) detected linearly. The analysis process creates perhaps as many as one hundred separate neural streams from each critical band. But most of these streams consist of low amplitude noise. A few of the outputs will have high amplitude coherent modulations, each corresponding to a particular source fundamental. The frequency selectivity is very high – enabling the pitch to be determined with accuracy. The brain can analyse the outputs from a single pitch across critical bands to determine timbre, and between ears to determine azimuth.

The length of delay line in the analyser (~100ms) was chosen to match our data on source localization. As the length of the delay line increases the pitch acuity increases – at the cost of reduced sensitivity and acuity to sounds (like speech) that vary rapidly in pitch. Tests of the model have shown 100ms to be a good compromise. As we will see, the model easily detects the pitch-glides in speech, and musical pitches are determined with the accuracy of a trained musician. The comb filter analyser is fast. Useful pitch and azimuth discrimination is available within 20ms of the onset of a sound, enabling a rapid response to threat.

But the most important point for these papers is that the fine perception of pitch, timbre, and azimuth all depend on phase coherence of upper harmonics, and that the acuity of all these perceptions is reduced when coherence is lost. When coherence is lost the brain must revert to other means of detecting pitch, timbre, and azimuth. When the coherence falls below a critical level a sound source is perceived as distant – and not engaging.

The degree of coherence in harmonics is a physical property. The model presented above can be used to measure coherence, and this measure can be useful in designing halls and opera houses.

THE EFFECTS OF REFLECTIONS ON HARMONIC COHERENCE

The discrimination of pitch

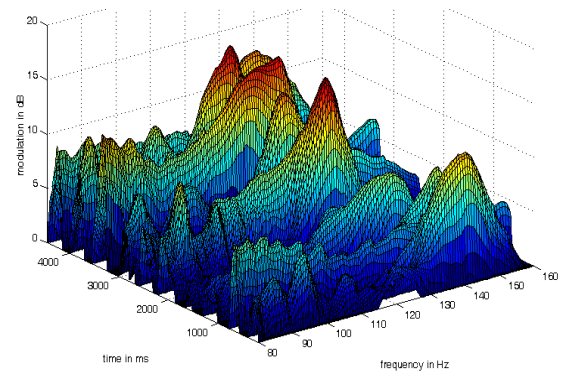


Figure 5: The syllables “one” to “ten” in the 1.6kHz to 5kHz bands. Note that the voiced pitches of each syllable are clearly seen. Since the frequencies are not constant the peaks are broadened – but the frequency grid is 0.5%, so you can see that the discrimination is not shabby.

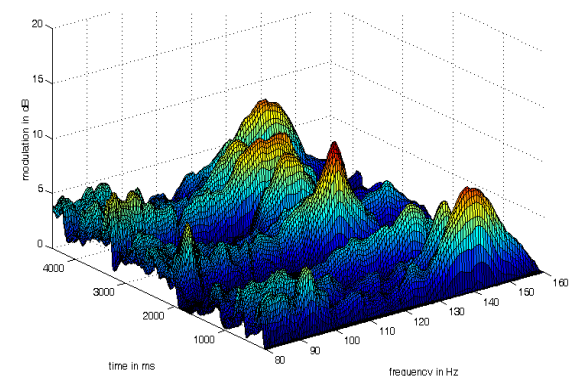


Figure 6: The same syllables in the presence of reverberation. The reverberation used was composed of an exponentially decaying, spatially diffuse, binaural white noise. The noise had a reverberation time (RT) of 2 seconds, and a direct to reverberant ratio (D/R) of -10dB. Although the peak amplitude of the modulations is reduced, most of the pitch-glides are still visible. The sound is clear, close, and reverberant.

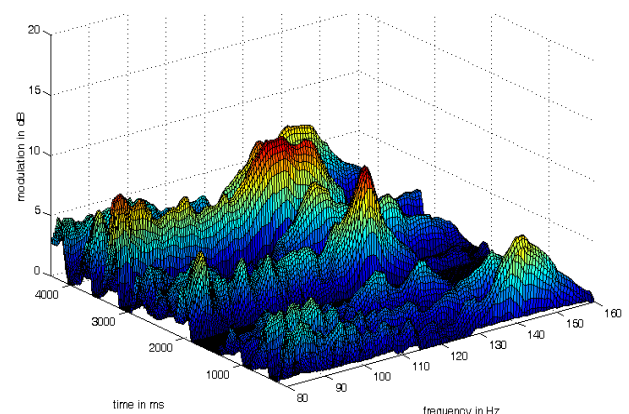


Figure 7: The same as figure 6, but with a reverberation time of 1 second, and a D/R of -10dB. The shorter reverberation time puts more energy into the 100ms window, reducing the phase coherence at the beginning of each sound. Notice that many of the pitch-glides and some of the syllables are no longer visible. The sound is intelligible, but muddy and distant.

The discrimination of horizontal direction (ILD)

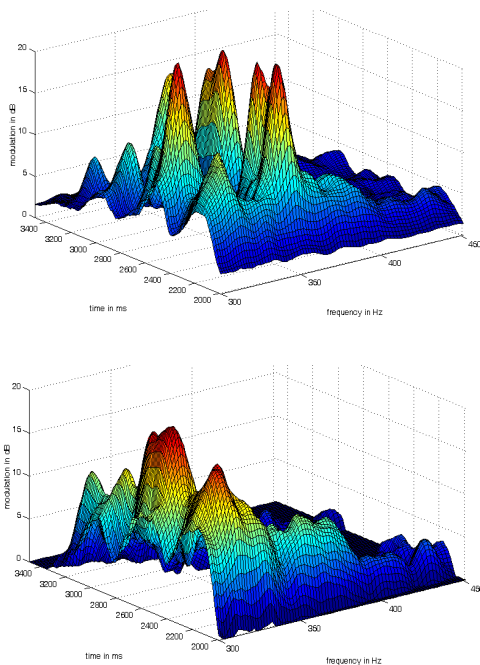


Figure 8: The modulations from two violins playing a semi-tone apart in pitch, binaurally recorded at ± 15 degrees azimuth. The top picture is the left ear, the bottom picture is the right ear. Note the higher pitched violin (which was on the left) is hardly visible in the right ear. There is a large difference in the ILD of the modulations.

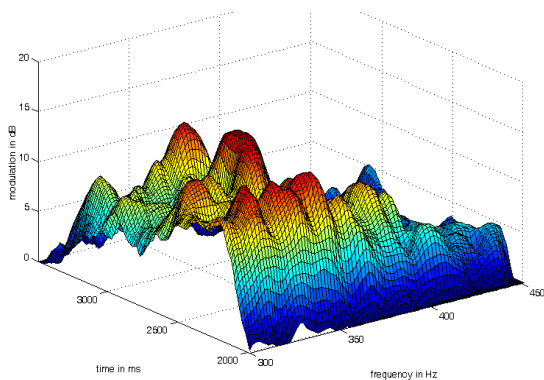


Figure 9: The same picture as the top of figure 8, but with the 1 second RT of figure 7. Note the difference in ILD is far less. The pitch of the higher frequency violin can still be determined, but the two violins are perceived as both coming from the centre. The azimuth information is lost.

Timbre – comparing modulations across critical bands

Once sources have been separated by pitch, we can compare the modulation amplitudes at a particular frequency across each 1/3 octave band, from (perhaps) 500Hz to 5000Hz. The result is a map of the timbre of that particular note – that is, which groups of harmonics or formant bands are most prominent. This allows us to distinguish a violin from a viola, or an oboe from a clarinet. This aspect is discussed in more detail in the first ICA2020 preprint.

SUMMARY THE HEARING MODEL

We postulate that the human ear has evolved not only to analyze the average amplitude of the motion of the basilar membrane, but also fluctuations or modulations in the amplitude of the basilar membrane motion when the membrane is excited by harmonics above 1000Hz. These modulations are at low frequencies, and easily analyzed by neural circuits. As long as the phases of the harmonics that create the modulations are not altered by reflections, the modulations from several sources can be separated by frequency and separately analyzed for pitch, timbre, azimuth, and distance.

The modulations – especially when separated – carry more information about the sound sources than the fundamental frequencies, and allow precise determination of pitch, timbre, and azimuth.

The phases of the harmonics that carry this information are scrambled when the direct sound from the source is combined with reflections from any direction. However if the amplitude of the sum of all reflections in a 100ms window starting at the onset of a sound is at least 3dB less than the amplitude of the direct sound in that same window the brain is able to perceive the direct sound separately from the reverberation, and timbre and azimuth can be perceived. The sound is likely to be perceived as psychologically close, and engaging.

Reflections from any direction – particularly early reflections – scramble these modulations and create a sense of distance and disengagement. But they are only detrimental to music if they are too early, and too strong. The model presented above makes it possible to visualize the degree to which timbre and pitch can be discerned from a binaural recording of live music in occupied venues.

I have been attending operas and concerts all over the world, and recording many of them with tiny microphones stuck to my eyeglasses above my pinna. For the last three years I have been doing this with probe microphones which sit gently on my eardrums. The probe microphone system, when headphones are equalized with the same microphones, is capable of astounding realism, at least for me and about 50% of the people who listen to it. [10].

FORMATIVE EXPERIENCES WITH OPERA

The author has had the good fortune to work with sound and reverberation in many capacities. My earliest experiences came as a sound engineer of classical music, work I still enjoy. I recognized the critical lack of natural-sounding artificial reverberation, and started a career designing these devices. This led to extensive work in recording, sound playback, and electronic architecture - in spaces as large as Chicago's Grant Park, and as small as automobiles.

Electronic manipulation of acoustics has enormous advantages as a research tool. It is possible to quietly wander through a hall during a rehearsal or a performance, and vary the acoustics with a remote control. Suddenly valid A/B comparisons can be made – and in the company of a skilled conductor or director the real – as opposed to imagined – effects of small modifications can be heard, and remembered.

Coming from the field of sound recording, and as a designer of reverberation equipment, I was all in favour of more reverberation than some halls provided. When I had the chance to work in opera houses I was rapidly disabused of this notion. My first settings for the reverberation enhancement in the Berlin Staatsoper were deemed way too strong by Mr. Barenboim. He insisted that the clarity of the singers be in no way reduced by reverberation.



Figure 10: Deutsches Staatsoper, Berlin

The solution was to increase the reverberation time and the reverberation level at frequencies below 500Hz – but to reduce the level and RT at upper frequencies. The result was that the extremely high engagement of the Berlin Staatsoper was retained, while the orchestra, and the fundamental frequencies of the singers, gained the richness they had lacked. This opera house is still my favourite in Europe. Critics were very happy with the change – but oblivious to its electronic origin. “Barenboim has managed to get the Staatscapella to sound like a real orchestra” was the comment.

Peter Lockhart, the assistant conductor of the Amsterdam Muziektheater, and I were in the Muziektheater stalls with a remote control during a rehearsal of “Siegfried”. We raised the strength of the reverberation above 1000Hz by 1dB. “Stop – that’s too much” said Peter. “Why”, I asked. “Because the singer just moved away from us by 3 meters” he said. In time I was able to hear this myself. It is difficult to do, because the visual impression of distance is so dominant. The effect of reflections and reverberation on sonic distance seems to be all or nothing. Either the singers are clear – and engaging – or they seem far, and less interesting. But you have to close your eyes to hear the difference as in increase in distance.

Hartmut Haenchen, the music director in Amsterdam, was conducting in the pit when I chose to raise the reverberation level by 1/2dB. He immediately waved to me from the podium, and told me to put it back. He had easily heard the increase, even standing right in front of the pit orchestra. I had similar experiences with Michael Schönwandt in Copenhagen.

All these experiences convinced me of the vital importance of immediately hearing the results of acoustic adjustments. I am convinced that without the ability to rapidly compare the small changes we were making we would not arrive at the settings we eventually used. We would probably have used more reverberation, and compromised engagement.

DRAMA EXPERIMENTS IN COPENHAGEN

Due to the success of the enhancement system in the old Royal Theatre in Copenhagen, Steve Barbar and I were asked to improve the speech intelligibility in a drama theatre – the “new stage” across the street. We had previously installed 64 loudspeakers in this theatre for use in opera, but the system was not intended for speech. We used a pair of line-array microphones to pick up the sound from the actors on stage, and I designed a sophisticated electronic gating circuit to remove as much as possible of the reflections and reverbera-

tion from the direct sound. The resulting signal was distributed through the speakers with appropriate sound delays.

The New Stage is a shoebox. Not the best shape for a drama theatre, as the average listener is too far from the actors for good engagement.

I turned the system on and off every 10 minutes during a live performance of Chekov in Danish with a full audience. Five of the major drama directors in Copenhagen were in the audience. At the intermission the directors were unanimous. “The system works – the actors are louder and more intelligible” they said. “We don’t like the system – turn it off.” “All right”, I said, “tell me why.”

Finally one said “The system makes the actors seem further away. I would rather the actors be unintelligible than sound further away. If the audience can’t understand them, the audience will listen more intently. This is just what I want.” The other directors agreed. They decided the solution to the intelligibility problem was better training of the actors.

RECENT OPERA HOUSE EXPERIENCES

Much of the work described in these papers started when I was working in Moscow at a new opera theatre built next to the old Bolshoi. The new theatre was intended as a replacement for the old theatre which was scheduled to be rebuilt. The new theatre was modelled after the Semperoper in Dresden, which was rebuilt in 1983 after being destroyed in the second world war. The redesigned Semperoper eliminated the layers of fabric and other absorptive surfaces that were – and are – typical of European opera houses. I measured the fully occupied reverberation time at 1.6 seconds at 1000Hz – quite long for a venue with only about 1200 seats. I have a recording of “Arabella” from the front of the first balcony. The singers seem far away, and the balance between singers and orchestra is poor. The sound is reverberant, and the orchestra sounds good. But the singers are not engaging.

The new Bolshoi is smaller, and more reverberant than the Semperoper. In addition there are strong focused early reflections from the curved side walls into the stalls. Singers sounded even further away than in the Semperoper. I had been asked to add reverberation to the opera house – which was not what it needed.

With the help of a scalper I attended an opera performance and a ballet in the old Bolshoi theatre. Both were magnificent. The hall has 3000 seats – more real seats than La Scala – and an occupied reverberation time of 1.2 seconds. The singers have enormous emotional power throughout the hall, and balance between them and the orchestra is overwhelmingly in favour of the singers. The ballet was surprisingly reverberant. The orchestra was playing the reverberation, and the very wide pit made the sound enveloping. This is what an opera house should be!

Two recently completed opera houses are discussed in the second of the three preprints prepared for the IOA 2010 conference.

EXPERIENCES IN LARGE CONCERT HALLS

In 2004 Leo Beranek asked me to join him for a talk at the 50th anniversary of the ASA. He wanted me to talk about concert halls – and I thought I knew too little about the subject. So I compared my recordings of violin concertos from three well known halls – Boston Symphony Hall, Avery Fisher Hall in New York, and the Kennedy Centre in Washington DC [12].

All these halls have similar cubic volume, rectangular shape, and similar numbers of seats. But they sound very different, with Boston clearly the best. I concluded at the time that the major differences were in the stage houses – but I did not test seats more than half-way back, and I was not convinced the stage house was the only reason they sounded different.

A few years later I was asked to write a short mathematical article on concert acoustics for the IEEE [11]. Once again I felt I knew too little about the subject, so I did some experiments. I modelled the Boston Hall and the Concertgebouw in Amsterdam, and did binaural convolutions of the result with my own HRTFs. The models sounded very good, and plausibly like the real halls. But the two halls sounded quite different.

I swapped the late reverberation from one hall with the late reverberation of the other – and found it made no difference. The shape of the build-up of reverberation in the two halls was similar - but there was an additional time delay in Amsterdam of about 10ms. When I shortened this time delay to match the delay of the Boston hall the two models sounded identical. Both halls create good engagement over a wide range of seats.

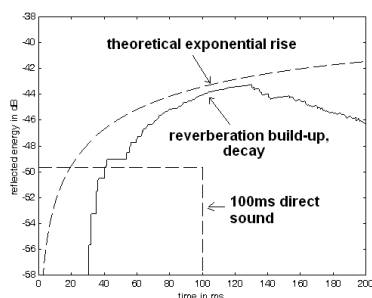


Figure 11: Reverberation build-up and decay from a 100ms excitation in a model of the Amsterdam Concertgebouw. The seat was chosen so that the D/R would be -10dB for a continuous excitation. Note there is more than 35ms time delay before the reflected sound pressure equals the direct sound pressure, enough time to perceive the direct sound as separate from the reverberation. Good azimuth perception and high engagement results. The value of *LOC* is +6dB.

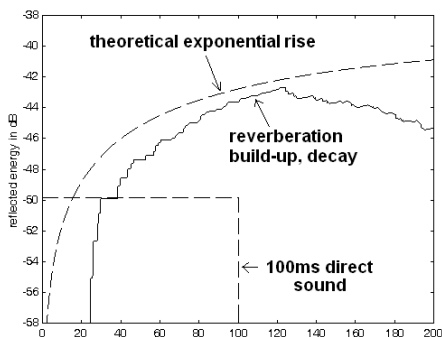


Figure 12: Reverberation build-up and decay from a 100ms excitation in a model of Boston Symphony Hall. The seat was chosen so that the D/R would be -10dB for a continuous excitation. Note that the initial delay is less than in Amsterdam. The sound is slightly less clear, but still engaging. The value of *LOC* is +4.2dB.

Clearly the rate at which reverberant energy builds up, and the level of this reverberation compared to the direct sound is

an extremely important – and neglected – aspect of hall design.

As a further experiment, I took the model of the Boston hall and reduced its dimensions by a factor of two. The sound went from clear and reverberant to muddy and unpleasant. If too many reflections come too soon, the result is disastrous. The value of *LOC* is ~0dB.

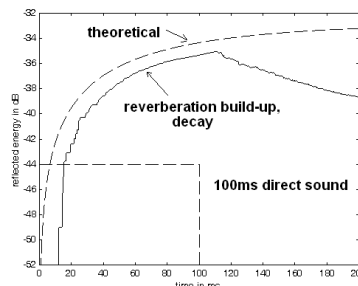


Figure 13: Reverberation build-up and decay from a 100ms excitation in a model of the Boston Symphony Hall with the dimensions all reduced by a factor of two. This hall has a reverberation time of under one second, but the sound is muddy and there is no engagement. Since the hall is smaller than the original, and the reverberation builds up more quickly, there is far more reverberant energy in the 100ms window that follows the onset of the direct sound. The direct sound is completely masked by the reflections.

As we will see later in this paper, the shape of the stage and the delay of the first lateral reflections are only part of the reason that Boston and Amsterdam are held in such high regard. I now believe an essential element of their success is the frequency dependence of the scattering elements on the walls and ceiling. These features increase the direct to reverberant ratio at high frequencies by directing high frequency reverberant energy downward to the front of the hall, where it is absorbed before it can travel to the back of the hall. The result is similar to the electronic enhancement in the Berlin Staatsoper. Reverberation and engagement can co-exist.

EXPERIENCES IN SMALL CONCERT HALLS

A little more than a year ago I had the opportunity to work in a 300 seat chamber music hall that had problems with muddiness and lack of engagement. If you sat close enough to the stage to hear the sound clearly the sound was too loud. If you sat further back it was difficult to localize and separate instruments. The occupied reverberation time was about one second – so by conventional thinking the sound should have been clear. But it was not. In the opinion of several of my acoustician friends the lack of clarity could be solved by adding a small shell behind the musicians. My work with models suggested the opposite. I felt the problem was that the first reflections came too soon, and were too strong. The solution would be to add absorption to the back wall and side walls of the stage.

We tried both solutions. The small shell made the muddiness worse throughout the hall. Adding a small amount of absorption – just panels at the bottom of the back wall - made an enormous improvement. The dean of the music school that owned the hall said he had never heard what a difference a seemingly small acoustic modification could make to the power of the music. The piano sounded like a newer and far superior instrument. We plan to install absorptive panels at the back of this stage. [See figure 14.]

I used my probe microphones to record the sound from a string quartet in two seats in London's renowned Wigmore hall. The hall is a long, narrow shoebox. I was convinced I could localize at least some of the instruments in the quartet from a seat half-way back into the hall, and could occasionally localize them from a seat two-thirds of the way back. But when I listened later without the visual image I was surprised by the poor localization, and due to a strong prompt reflection from the side wall half of the time the sound was actually coming from my right side. Engagement in the recording was poor. I sat much closer in a previous visit to the hall – and loved the sound. This hall – like many others – has a deservedly good reputation, but only for the best seats.



Figure 14: The stage of a small concert hall that lacks clarity and engagement.

A medium sized (1000 seats) shoebox hall

The problems that accompany a long reverberation time and prompt early reflections do not go away as a hall gets larger. I recently attended several concerts in a new medium sized shoebox hall, and recorded a string quartet in two different seats in the stalls. Once again the reverberation time was specified as at least 1.7 seconds. A seat in row F provided clear localization and engagement for all the instruments in the quartet. The sound was beautiful and exciting. In row K the clarity was lost – all the instruments blended into the centre of the sonic image, and the inner voices were often inaudible. The music was nice, beautiful in a way, but it lacked the excitement of row F. Row K was less than half-way back into the hall – which does not bode well for the rest of the seats.

I played these recordings for one of the designers of the hall and one of the musicians. The musician strongly preferred row F. The designer clearly heard the difference – but claimed it was unimportant. “The hall looks and sounds like a concert hall. If a listener wants a clear sound he can sit in row F. Otherwise he can choose row K.” But the sound in row K was less reverberant than row F – why would anyone choose it? You could call the sound in row K “well blended” but I prefer Beranek's description: “you can sell them to tourists.”

REVERBERATION AND ENVELOPMENT

Reverberation also plays a vital role in live performances – and the properties of halls that provide reverberation seem to conflict with the properties that provide engagement. The loudness of music in a hall also plays a role. But great halls exist that successfully provide loudness, engagement, and reverberation at the same time over a wide range of seats. We will discuss here the reasons these halls work better than others, with the aim of providing methods that can be used to increase the number of engaging seats in new and existing halls and opera houses – and to improve the audibility of

reverberation when it is lacking. We find that engagement and reverberation are not opposites of each other. Both require the perception of the direct sound to be optimally heard.

Reverberation in recorded music

Reverberation is technically the sum of all the sound that does not travel directly to a listener. The most common measure of reverberation is the reverberation time (RT), the time it takes for sound to decay 60dB. But the perception of reverberation is more complicated than can be expressed with a single number. Recording engineers of both classical and popular music use reverberation as one of the essential components of a good recording, and carefully add it to sound mixes using a variety of commercial digital equipment, or with special purpose microphones in recording venues.

In all such recordings it is the *level* of the reverberation relative to other elements of the mix that is the most important parameter, not the reverberation time. I have measured the amount of reverberation in many classical mixes, and have made experiments where good acousticians add reverberation to a mix, and then measure the amount used. In all cases the answer is the same. In classical mixes the total energy in early reflections and late reverberation is between minus 4dB and minus 6dB of the total energy in the direct sounds. This means that in recordings – which in some sense represent an ideal representation of a performance – the D/R is between +4 and +6dB. This level of reverberation can be considered ideal because recordings can be A/B compared to each other, and customers can choose which ones to play, and which to leave to languish. Engineers – aided by some very critical conductors in the playback room – have learned what kind of sound does the music the most justice.

This is the range of D/R that was explored by Barron [5,6] and others in their studies of spatial impression. The author knows of *NO* successful classical or popular music recording where the D/R is less than -3dB. Very few seats in a concert hall have D/R ratios this high. Recording engineers add reverberation – or arrange their microphones to record reverberation – at levels just strong enough for it to be frequently, if not continuously, audible while the music is playing. There is no point of reverberation if you cannot hear it, and more than enough reverberation muddies the recording.

Recordings have become the norm for music listening, and opera performances such as the New York Metropolitan Opera HD broadcasts are seen by far more people than the live events. The sound of the MET broadcasts in most theatres is harsh, direct, nearly devoid of reverberation, and highly engaging. (Movie music in the same theatres is more reverberant than the operas – but movie dialog is always dry.) The HD opera sound is not beautiful, but the dramatic experience is very powerful. The video image brings you close – sometimes too close – to the performers, and the sound makes them seem to shout in your face. The result can be overwhelming. The performance of “Salome” with the Finnish soprano Mattila was blood-curdling to this author. It was emotionally far beyond what I would have experienced from a balcony at the MET.

I also saw “Salome” in the State Opera House in Vienna. The sound was far superior to the broadcast in timbre, and also nearly devoid of reverberation. The Vienna Philharmonic can play very loud in that house! The result was highly engaging. In Vienna the visual distance was greater than the HD image – but it was still a powerful performance. Like it or not, audiences have come to expect, or will come to expect, a similar experience to the HD broadcast when they come to a live event. They will get it in the the Vienna Opera, or the Staat-

soper Berlin. I can't imagine seeing or hearing "Salome" in an opera house like the Paris Bastille.

A FEW EXAMPLES

Although a small percentage of shoebox concert halls with a reverberation time of about 2 seconds have a good reputation, the success of a hall (of any shape) with the same reverberation time is not guaranteed. The opposite is proved by halls all over the world. We can glean some of the reasons some halls work better than others by looking at a few examples.

In [12] the author examines three shoebox halls of similar size and shape. A major difference is the design of the stage house. The stage of New York's Avery Fisher hall is deep and low ceilinged, with no absorption besides the orchestra on the floor. There are multiple prompt internal reflections which add to the direct sound of the instruments, particularly those in the back of the orchestra. These instruments sound muddy and far away, although instruments in the front row, such as a violin soloist, have some engagement. But the engagement is lost as you move back in the hall. In the front of the first balcony the sound is muddy, not localizable, and not reverberant. It is simply unclear. The sound from the rear of the stage lacks clarity because there are too many reflections in the stage house. The sound more than half-way back in the hall is not engaging because there are too many reflections above 700Hz in the first 100ms.

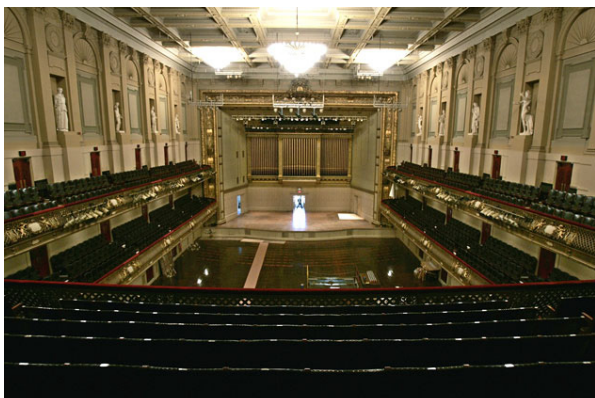


Figure 15: Boston Symphony Hall. The stage house is high, wide, and shallow, with sloping side walls and ceiling. Reflections from these surfaces are directed into the hall, and multiple reflections do not occur within the stage house. Instruments in the rear of the orchestra have equal clarity as instruments in front. Notice the coffers on the ceiling, and the niches along the side walls.

The stage in Boston gives the orchestra both clarity and power. Instruments in the rear of the orchestra are heard with equal clarity as those in front. The coffered ceiling and the niches on the side send frequencies above 1000Hz back to the front of the hall, effectively increasing the D/R ratio for seats in the rear. The hall is engaging over a wide range of seats. With an occupied reverberation time of only 1.8 seconds, the hall is perceived as both reverberant and enveloping.

The Amsterdam Concertgebouw is square in plan, and there is no stage house. The average distance from the orchestra to a listener is smaller than it is in Boston. There are no reflections from the wall behind the orchestra, as they are absorbed by the audience and the organ. The ceiling is coffered, as in Boston, and the reflections from the side walls arrive later than they do in Boston. All these factors combine to give the hall unusual clarity. The reverberation time is longer than in Boston, and the late reverberation is strong, as there are a great many surfaces that reflect the sound upward above the

audience, where it can take its time to get back down. The high late reverberation level, combined with the clarity of the direct sound, give a rich sense of envelopment throughout the hall.

Disney Hall, Los Angeles

Disney Hall is a vineyard hall, not a shoebox. There is no stage house, but reflections from the rear of the orchestra are directed into the stalls by the wall behind the orchestra. This adds a prompt, strong early reflection to the direct sound. This reflection is not sufficient to eliminate engagement, but it is a major component of the sum of early reflections in the first 100ms. The ceiling is devoid of frequency-dependent scattering, and reflects sound from the orchestra down into the audience, where it adds to the prompt reflection from stage wall. The sum is sufficient to scramble the phases of the direct sound in the first 100ms. All these reflections are absorbed by the orchestra and audience, so very little is left over to contribute to late reverberation. The result is very strange. Even in the middle of the stalls the orchestra seems far away. At the same time late reverberation is almost inaudible. It is unusual that a hall with a two second reverberation time should sound dry – but the late reverberation level is too low, and the direct sound is not separately perceived.

I heard a performance of "Le Sacre du Printemps" in Disney Hall from a seat in the middle of the stalls. The sound was distant, relatively quiet, and might be best described as "nice". I was surprised by the sense of distance. I expected at least some engagement in that seat. The next week I was in Berlin, tuning the Staatsoper. As luck would have it, after the tuning the Staatscapella performed "Printemps" with the Berlin Staatsoper Ballet. I happened to record both the performance in Disney and the performance in the Staatsoper with the same equipment. The Staatsoper was 10dB louder than Disney Hall. The sound from the centre of the first balcony in the Staatsoper was anything but "nice". It was wild, orgasmic, gut wrenching. This is the music that started a riot in Paris when it was heard in the dry acoustics of the Theatre des Champs-Elysees. No riot was started by the performance in Disney. The audience politely applauded.

MAIN POINTS

The ability to distinctly hear the *Direct Sound* – as measured by *LOC* or through the analysis of a binaural recording – is a vital component of the sound quality in a great hall.

Hall shape does not scale. Both D/R and the rate of build-up change as the hall size scales – but human hearing (and the properties of music) do not change.

A hall shape that provides good localization in a high percentage of 2000 seats will produce a much lower percentage of great seats if it is scaled to 1000 seats. We need to bring the average seating distance closer to the musicians if a small hall is to be both reverberant and engaging. We also need to reduce the reverberation time.

Frequency-dependent diffusing elements are necessary, and they do not scale. The audibility of direct sound, and thus the perceptions of both localization and engagement, depends on frequencies above 700Hz. Diffusing elements that direct high frequencies down to the front of a hall will improve the audibility of direct sound by raising the D/R in the rear of the hall. (The absorption only occurs in occupied halls – so the effect will not be detected in unoccupied measurements!) Measurements occupied in Boston Symphony Hall (BSH) above 1000Hz show a clear double slope that is

not visible at 500Hz. Although BSH is a large shoebox, the hall has high engagement in at least 70% of the seats.

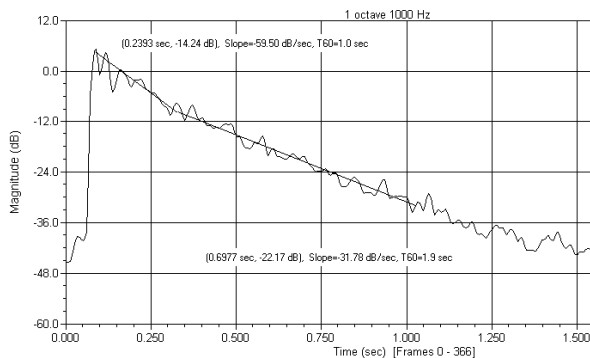


Figure 16: 50ms window-integrated impulse response of Boston Symphony Hall with occupied hall and stage, 1000Hz octave band. The source was in the middle of the violin section, the receiver was in the front of the first balcony – nearly 100ft from the source. Note the clear double slope. The RT for the first 10dB of decay is 1.0 seconds. The RT of the later decay is 1.9 seconds. The side wall and ceiling reflections have been significantly attenuated at this frequency. This is Leo Beranek’s favorite seat. It provides excellent localization, engagement, and envelopment. The double slope is not seen in the 500Hz octave band. The direct sound has been overwhelmed by reflections and reverberation – as one would expect at so great a distance.

The most important factor that contributes to both engagement and to the beneficial perception of reverberation is the size and shape of the hall.

The author has had the experience of hearing a fine string quartet from a distance of only two meters. The clarity was fantastic. Shortly thereafter I heard another fine quartet from the middle of the stalls in the auditorium at the Metropolitan Museum of Art in New York City. The sound was reasonably clear but too soft. There may be an ideal distance from which to hear various kinds of music. Perhaps the spaces in which the music was first performed are a guide. But one should be cautious about the current condition of these spaces. Many were far less reverberant in the past, filled with fabric since removed, and richly dressed audiences.

Modern spaces need to be larger, but it is possible to build large venues which bring the audience closer to the musicians. The Concertgebouw in Amsterdam does this for a large orchestra. Asbjørn Krokstad, Norway’s best known acoustician and a noted conductor, gave a provocative lecture in Oslo about why current concert halls are not attracting younger audience members. He suggested that halls need to be *engaging*, not just nice. I was very excited – he had given me the word to describe the perception I had been attempting to communicate. At the end of the lecture he showed a picture of the Teatro Colón in Buenos Aires, Argentina. “Is this the concert hall of the future?” he asked.



Figure 9: Teatro Colón in Buenos Aires, Argentina. This hall is not a shoebox, but a large opera theatre with four tiers of balconies and 2,487 seats. It is renowned as a concert hall, as is shown in the publicity picture above. In this hall the average listener is close to the orchestra. The cubic volume needed for good late reverberation is provided by a high ceiling, which is also high enough that the ceiling reflection into the stalls is relatively weak. Beranek writes “as a concert hall it is surprisingly satisfactory.” He has never heard a conductor who did not say that the Teatro Colón is one of the best halls in the world to conduct in, and to listen in. Orchestras love playing there. Why has it not been widely copied?

Jordan Hall at New England Conservatory

Boston is blessed not only with one of the three halls rated “excellent” in Bernack’s surveys, but with at least one of the finest chamber music halls that I know. Jordan Hall at New England conservatory is semi-circular in shape with a single balcony. The balconies are spaced relatively high above the parquet, giving ample space for reverberation from the high ceilings to reach the audience members sitting below.

If you are a chamber musician and can attract a large audience, Jordan Hall is your Mecca. Jordan is intimate. The average seating distance is close enough that the direct sound is strong and engaging in almost every seat, and yet the reverberation is almost always audible and rich. Why build a shoebox, when this shape is so successful acoustically?

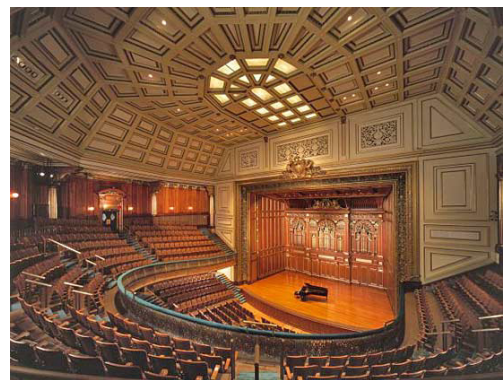


Figure 10: Jordan Hall at the New England Conservatory, Boston. (1020 seats) The hall is semi-circular in shape, with a single large balcony. This arrangement shortens the average seating distance compared to a shoebox hall. The high ceiling and ample volume above the second balcony provides plenty of resonance. The hall is in near constant use – and expensive to rent! It is known all over New England, and through the radio show “From the Top” it has become known throughout the United States.

SMALL HALLS

The smaller the hall the more difficult it is to combine resonance and engagement at the same time. In a small hall reverberation – whether in the form of early reflections or late reflections – builds up very quickly with time. As described earlier, the brain needs time to separate the direct sound from the reflections that follow. The time needed is dictated by human physiology, and not by the size of the hall. Human physiology also dictates that the sense of reverberance and envelopment that audience and musicians desire arises from reflections that arrive at least 100ms after the direct sound. In small halls the reverberation time is by necessity lower than in large halls, and the sound has decayed substantially before it can be heard as reverberance.

Since engagement is subconscious, and reverberance is not, acousticians advise that small halls be made as reflective as possible. This increases the reverberation time, and thus the resonance. But removing absorption will always raise the strength of the early reflections, and raise the total reverberant energy. The result will be even poorer clarity and engagement. The solution to this conundrum has been around a long time: add high frequency absorption to the stage! By absorbing some of the high frequencies that do not travel directly to the audience, the D/R and the engagement can be increased with little effect on the reverberation time.

How many modern recital halls surround the musicians with thick curtains, and hang a curtain in front of the proscenium? Who remembers the good old days when Carnegie Hall in New York had similar adornments? How many people with long memories wish the fabric would return?

The curtains behind the stage and in front of the proscenium absorb sound energy that would otherwise overwhelm the direct sound. The direct to reverberant ratio above 700Hz can be increased 3dB or more.

FREQUENCY DEPENDENT CANOPIES

Tanglewood Music Shed, Lenox Massachusetts

The canopy over the orchestra in the Tanglewood Music Shed consists of open and closed sections of equilateral triangles of variable size. The canopy acts as a filter, directing high frequencies down into the orchestra and the first few rows of the audience, and letting the low frequencies into the upper reaches of the hall, where they have ample time to bounce around before coming back down. The high frequencies absorbed by the orchestra and audience do not contribute to late reverberation, thus raising the D/R above 1000Hz in the middle and rear of the hall.

The addition of the canopy to the Tanglewood Music Shed successfully changed the sound from impossibly muddy to clear and engaging for a wide range of seats. Such semi-open canopies (clouds) are relatively common in halls, but the people who design them usually do not think of them as frequency filters for reducing high frequency reverberation.

ELECTRONIC ARCHITECTURE

In small and medium sized halls – and in most traditional opera houses – the only practical way to achieve the ideal balance between clarity and reverberation is to gently increase the late reverberation through the careful use of electronics. The success of some of these systems has been demonstrated in halls and opera houses around the world. The author's recent algorithms increase late reverberation transparently, with no effect on clarity.

But not all electronic systems work well, and the idea of electronics in classical music halls is often resisted. There are two essential requirements for a successful installation. The first is that the hall must already have excellent clarity and engagement, and the electronics should provide mostly late energy. With careful adjustment electronic enhancement can successfully augment the direct sound in regions where the direct sound has become too weak to be audible, but in most small halls there are too many early reflections. Lack of engagement must be corrected before the electronics are used. Otherwise electronic reflections just add to the mess.

Adding absorption to the stage and side walls of a small hall will improve engagement, but reduce the reverberation time. The audience will hear greater clarity, and the remaining reverberation will be more audible. But the performers, who

rely on late reverberation to judge their loudness and balance, will not be happy. We have found that a minimal enhancement system can add just enough late energy to restore or slightly increase the reverberance on stage and in the hall. Everyone is delighted.

The other requirement for a successful system is that the microphones receive primarily direct sound. Some electronic enhancement systems work by picking up sound in multiple positions in the hall, amplifying and delaying it a bit, and reproducing it somewhere else. The reverberation time goes up – but the sound being amplified is already muddy, and the amplified reverberation contributes to the mud. It does not sound pleasant or natural.

ACKNOWLEDGEMENTS

The experiences described in this paper would not have been possible without the participation, encouragement, and support of friends and colleagues too numerous to name. But I would especially like to thank Steve Barbar of Lares Associates for his passionate commitment to electronic acoustics, fine ears, and his unlimited appetite for hard work. I also thank Leo Beranek for his continually generous encouragement and support.

REFERENCES

- 1 D.H. Griesinger, "[Pitch Coherence as a Measure of Apparent Distance and Sound Quality in Performance Spaces](#)" Preprint for the conference of the British Institute of Acoustics in May, 2006. Available on the author's web site: www.davidgriesinger.com
- 2 D.H. Griesinger, "[The Relationship Between Audience Engagement and Our Ability to Perceive the Pitch, Timbre, Azimuth and Envelopment of Multiple Sources](#)" A powerpoint presentation given to AES local sections in Boston and Washington DC, June 2010
- 3 D.H. Griesinger, "[The importance of the direct to reverberant ratio in the perception of distance, localization, clarity, and envelopment](#)" A power point presentation with audio examples given at the Portland meeting of the Acoustical Society of America, May 2009.
- 4 D.H. Griesinger, "[The importance of the direct to reverberant ratio in the perception of distance, localization, clarity, and envelopment](#)" A preprint for a presentation at the 126th convention of the Audio Engineering Society, May 7-10 2009. Available from the Audio Engineering Society.
- 5 M. Barron: "Auditorium Acoustics and Architectural Design" E&FN SPON 1993.
- 6 D.H. Griesinger, "How Loud is my Reverberation" - on the author's web page.
- 7 D.H. Griesinger, "[The psychoacoustics of apparent source width, spaciousness & envelopment in performance spaces](#)" Acta Acustica Vol. 83 (1997) 721-731. (this paper is on the the author's web page)
- 8 S. Olive and F. Toole, "The Detection of Reflections in Typical Rooms" Preprint 2719 from the Audio Engineering Society Nov. 1988. Available from the AES.
- 9 F. Toole "Sound Reproduction, Loudspeakers and Rooms" Book available from Amazon.com
- 10 D.H. Griesinger "[Frequency response adaptation in binaural hearing](#)" A powerpoint presentation given at the May 2009 meeting of the ASA in Portland OR. Available on the author's web-page.
- 11 D.H. Griesinger, "Concert Hall Acoustics and Audience Perception" IEEE Signal Processing Magazine [1] March 2007
- 12 D.H. Griesinger, [Slides for the Acoustical Society Workshop with Leo Beranek, June 2004](#)