# RECENT EXPERIENCES WITH ELECTRONIC ACOUSTIC ENHANCEMENT IN CONCERT HALLS AND OPERA HOUSES

David Griesinger

Lexicon
3 Oak Park
Bedford, MA 01730
dg@lexicon.com
www.lares-lexicon.com

ABSTRACT

This paper gives a brief summary of acoustical theory based on human perception. It then uses this theory to discuss the design and performance data of electronic acoustic enhancement systems installed in a number of opera houses and concert halls. The installations include the Deutches Staatsoper in Berlin, the Hummingbird Center in Toronto, and the Adelaide Festival Center Theater in Adelaide, Australia. Solutions to the problems of maintaining optimum clarity of the singers while providing optimum envelopment for the orchestra are given.

Mörbisch

## INTRODUCTION

Electronic acoustic enhancement of spaces for music performance has frequently been long on promise and short on performance. The major problem has been uncontrolled acoustic feedback between the microphones and the loudspeakers in the enhancement system. This feedback induces an artificial metallic coloration into the system. Avoiding the coloration has meant operating the system at loop gains that provide little overall benefit. Enhancement systems that use multiple time variant reverberators have essentially solved the feedback problem, allowing the system designer to create the necessary acoustic fields without artificial coloration.

The technology is available – but how do we use it? We are faced with the problem of determining just what are the major acoustic difficulties of a particular space. Assuming we identify these correctly, how can microphones and loudspeakers be installed to solve them without breaking the budget? This paper will present a brief update on our research into the perception of musical acoustics, and show how this knowledge can be applied toward solving problems in real spaces.

## RECENT RESEARCH INTO PERCEPTION

In a previous paper (1) we wrote that acoustic descriptors can be divided into four categories: descriptors of localization, spaciousness, intelligibility, and reverberance. Since that time we have made some progress in understanding how humans perceive sound, and feel it is possible to revise this list.

We find that there are several key processes in sound perception, each working at a successively higher neural level. The low level processes, such as the separation of sound into different frequency bands, and the detection of localization through Interaural Intensity Differences (IIDs) and Interaural Time Differences (ITDs), occur early in the perception process, and in general have very short time constants. Higher level processes can take substantial amounts of time to complete. The time constants are vital to the way we perceive room acoustics.

## LOW LEVEL PROCESSES:

1. The analysis of incoming sound pressure into frequency bands. This analysis takes place on the basilar membrane and is fundamental to the hearing process. At the lowest level we hear sounds in separate frequency bands.
2. The detection of rapid increases in level in individual frequency bands. This “rising edge” detection occurs early in the neural process and is the first step in the detection of the starts of individual foreground sound units. In speech these individual sound units are called “phones”. In music they are called “notes”.
3. The detection of interaural time and level differences in each frequency band. These interaural time and level differences are gated with the “rising edge” data to determine the azimuth of the sound source. When the “rising edge” data is absent – when the sound is continuous – IID and ITD still determine azimuth if they are stable and consistent. For example, when a sound occupies several critical bands the IID and ITD should be the same for each band.
4. The determination of the average uncertainty in the IID and the ITD. In an acoustic environment these uncertainties are primarily due to fluctuations in the both the IID and the ITD. These fluctuations are caused by interference between the direct and reflected sound. Fluctuations in IID and ITD that occur during the rise time of a sound event broaden the source image. Fluctuations that occur later can be interpreted as room sound or as envelopment, depending primarily on when these fluctuations occur relative to the sound events.

Of these low level processes the most basic is the analysis of sound into frequency bands. In all the processes that follow this separation has already occurred. Thus when we speak later of localization or envelopment we are not assuming these perceptions to be independent of frequency. Localization can be sharp at high frequencies, and at the same time it can be poor at low frequencies. If a particular sound event includes both high and low frequencies the sharpness of localization in the different bands can be separately perceived – although in overall impression the most accurately localized bands will dominate. The same frequency selectivity applies to envelopment. The frequency dependence of intelligibility, localization, and envelopment is particularly important to musical acoustics.

## HIGHER LEVEL PROCESSES, IN APPROXIMATELY THE ORDER THAT THEY OCCUR:

1. The parsing of sounds into individual units, the phones and the notes. To perform this parsing the hearing process must find where one sound event ends and another begins. Thus detecting the ends of sound events is often as important as detecting the beginnings.
2. The determination of the direction and timbre of individual sound events.
3. The organization of groups of sound events into foreground streams. In speech the phones from a particular speaker are organized into phrases and sentences. In music the foreground streams consist of musical lines from individual instruments or sections.
4. When there are several speakers talking at the same time, we organize the sound units from each of them into separate streams. Likewise there can be several simultaneous foreground sound streams.
5. The stream formation process sorts individual sound events using all available clues, such as direction, timbre, and pitch. Thus the azimuth of a particular event can help assign it to a stream.
6. The formation of a "background stream" that contains the sounds perceived between elements of the foreground stream. The background stream contains room noise, reverberation, etc. While there can be several foreground streams, there is only one background stream.

We wish to emphasize that the formation of sound streams is a vital part of our sonic perception. We perceive localization, timbre, and reverberation much more strongly in a series of connected sound events than we do in individual, isolated sound events. For example, when we hear the reverberation from a loud chord that is followed by silence, the reverberation becomes a foreground sound event – we can apply the entire analysis power of our brains to it. When the reverberation is heard during the gaps between syllables of speech an entirely different neural process is involved in its perception. This perception is not of an event, but of a continuous stream of sound. If the reverberant level is strong and there are significant fluctuations in the IID and ITD during these gaps we will perceive significant envelopment. During and after the stream formation process several further actions occur:

7. The assignment of meaning to the various foreground streams.
8. The inference of source distance from the relative strength of the foreground and background sound streams.
9. The interpretation of the fluctuations in the IID and ITD as either a "room" impression, or as envelopment. This interpretation depends on the time delay between the end of the sound event and the reflected energy that produces the fluctuation.

Of the higher level processes, the parsing of sounds into individual events is by far the most critical. Speech comprehension drops very rapidly when noise or acoustic conditions prevent the reliable detection of the ends and beginnings of phones. This is why the modulation transfer function of an acoustic channel is a meaningful measure of speech intelligibility.

From the nature of the event detection process we can see that:

1. The effect of early lateral reflections on localization will depend on the rise time of the sound events used as a sound source.
2. Where sound events have rapid attacks the sharpness of the sound image (the apparent source width) is determined by the presence or absence of reflected energy that arrives during the rise time of the sound event. The gating of localization with the "rising edge" data gives a significant advantage to us as a species. The rise time of sounds is usually not corrupted by reflections. Speech phones can rise quite rapidly – in under 10ms. Thus speech can be accurately localized even in small spaces. The same is true of musical

sounds from many solo instruments. Perceptual experiments show that lateral reflected energy arriving later than about 10ms has little effect on the source width of such sounds.

3. Legato music for a large string section tends to have long rise times for individual notes. When we do a perceptual experiment using such music, we expect that the image will be broadened by lateral reflected energy with delay times of 50ms or more. The expected broadening is easy to confirm.
4. Intelligibility of either speech or music will be reduced if reflected energy reduces the ability of the hearing mechanism to detect the starts and ends of phones. Phones in rapid speech come as frequently as every 150ms. Normal speech can be somewhat slower. The gaps between phones are typically 50ms or greater. From this data we can immediately infer that reflected energy arriving between 50 and 150ms after the ends of a sound event will be particularly detrimental to intelligibility.
5. The perception of reverberation and envelopment will depend on the presence of gaps between phones or notes where reverberation can be heard, and on the ability of the hearing mechanism to separate the sound in these gaps from the foreground sound events. This separation process takes time – at least 100ms after the end of the sound event must elapse before the sensitivity to background sound is at a maximum.
6. Thus the perception of reverberation and envelopment depend on:
   a: the "transparency" of the musical material
   b: the strength of the reverberant sound at least 100ms after the ends of notes.

For speech, solo music, and thinly orchestrated music, lateral reflections arriving in various time ranges have the following properties:

0-10ms – these reflections make the sound event louder, change the timbre, broaden the source image, and/or cause image shifts.

10-50ms – these reflections cause a "room" impression that is not enveloping, but desirable if the reflections do not exceed the energy of the direct sound. Reflected energy in this time range also increases the loudness and affects the timbre of the sound event. We call the spatial impression created by these reflections "early spatial impression" or ESI.

50-150ms – these reflections produce some sensation of envelopment, but the primary effect of energy in this range is to reduce intelligibility.

150-400ms – these reflections contribute to the background sound stream.

The background stream is highly audible. It produces the major perception of "support" for a solo musician, and envelopment for an audience member. The strength of the envelopment perception depends on the absolute loudness of the reverberation. The louder the musician plays the stronger the envelopment perception will be. We call this envelopment perception "background spatial impression" or BSI.

For legato strings and continuous thickly orchestrated music, the beginnings and endings of notes are not easily detected, and the effect of lateral reflections becomes much less dependent on the delay time. Reflections in the following ranges have the effect of:

0-10ms – these reflections affect loudness and timbre. They can also widen the sound image and shift the azimuth, but they do not affect envelopment or "room" impression.

10-50ms affect loudness and timbre, and broaden the sound image. They also contribute to a form of envelopment we call "continuous spatial impression" or CSI. CSI is less audible than BSI, and depends on the direct to reverberant ratio, not on the absolute level of the reverberation.

50-150ms – these reflections affect the musical intelligibility, can broaden the source width, and contribute to CSI

150-400ms – these reflections contribute to CSI. With very legato sources these reflections can also affect source width.

Notice for continuous music nearly all lateral reflections affect envelopment. In fact, it is the ratio between the total medial and the total lateral energy that will determine the amount of envelopment. This is particularly true at low frequencies. At least in opera houses the reflected energy tends to be medial at low frequencies, since it comes primarily from the front. Low frequency envelopment tends to be low in such spaces, even for continuous music.

With the development of this perceptual theory, we can narrow our list of acoustic properties. For perception the vital sound properties are intelligibility, localization, "room impression" (ESI) and envelopment (BSI).

The importance of each of these perceptions may depend on the type of sound and on personal preference. For speech everyone agrees that intelligibility is of primary importance. For music many if not most listeners believe envelopment becomes much more important, and a substantial degradation of intelligibility and localization is acceptable to achieve adequate envelopment. A lack of "room impression", ESI, is perceived as a lack of distance and blending between the listener and the sound source. This distance or blending is nice to have – but in our experience it is much less important than intelligibility and envelopment. The distinction between ESI and BSI is important, since traditionally it is a lack of early reflections that is blamed for most acoustic problems, and yet it may be possible to augment late reverberation much less expensively than early reflections.

## APPLICATION OF THE THEORY TO ACOUSTIC ENHANCEMENT

Modern halls come in all shapes and sizes. In spite of an enormous range of audience capacity and internal volume these halls are expected to sound good with a wide range of sound sources. The problems of designing a hall with good acoustics for speech are well known and will not be extensively discussed here. For speech one wishes to minimize the reflected energy that arrives 50ms or more after the direct sound. Reflected energy or energy from a reinforcement system that arrives earlier than 50ms can be helpful as long as it is primarily medial.

It is well known that if a hall is to be used for music it should have a longer reverberation time than a hall designed for speech, and it is widely believed that a two second reverberation time is optimal for music. Unfortunately the acoustic properties of a hall depend both on the reverberation time and the volume of the hall. A small hall with a two second reverberation time sounds very different from a large hall with a two second reverberation time.

One reason is that in natural acoustics the reverberation time, the reverberant level, and the distance at from the sound source for a given direct to reverberant ratio are all linked. You can not alter one without altering all of them. In general a small hall designed for a two second reverberation time will have much too high a reverberant level for most purposes. The critical distance – the distance where the direct sound and the reflected energy are equal – is too small.

Another difference is that in large halls the early reflections in the time range of 50-150ms can be lower in energy than in small halls of the same reverberation time. In many large halls the reverberation decay is not immediately exponential. There may be a few early reflections from the stage house, but then there is a little less energy than one might expect before exponential decay begins. This characteristic is particularly noticeable in a hall without a stage house, such as the Concertgebouw in Amsterdam. The result is a sound that has both high intelligibility and an strong sense of envelopment. A small hall can not achieve this sound. If we make it reverberant enough to supply envelopment, the energy in the 50-150ms range is too high. Intelligibility, localization, and timbre are all compromised.

With electronic enhancement the reverberation time and the critical distance do not have to be linked. The energy in the time range of 150ms and beyond can be altered without excessive energy in the early field. When using multiple time variant reverberation systems it is not necessary to use a large number of microphones, and we recommend that two to four microphones be installed as close as possible to the sound source consistent with uniform coverage. With such an array the feedback can be minimized, and the designer has much more control over the reverberant level.

All enhancement systems are not equal in this regard. Some current enhancement systems place the pickup microphones at distances greater than the reverberation radius from the sound source. With these systems the enhancement system acts as negative absorption. Any increase in level from the enhancement system must increase the reverberation time through acoustic feedback.

## EXAMPLES

### Deutches Staatsoper, Berlin

The Berlin Staatsoper is typical of a great number of European opera houses. It has a horseshoe plan with four rings, and a seating capacity of about 1500. The stage house and proscenium are small by modern standards, which gives the house an intimate character very well suited to many types of opera, particularly Mozart. The occupied reverberation time is below 1 second at most frequencies, but speech intelligibility, a vital component of the dramatic connection between an actor and the audience, is very good throughout the hall. Although the acoustics were excellent for drama, they lacked envelopment for the orchestra, particularly for the music of Strauss and Wagner. The lack of envelopment was blamed (as usual) on a lack of early reflections, although attempts to improve the situation with reflectors had failed. Like many other opera houses the Staatsoper had minimal funds available for any improvements to the hall acoustics, which had been judged “good enough” for many years.

Through the efforts of Albrecht Krieger, the tonmeister, and Daniel Barenboim, the music director, a Lexicon Lares system was temporarily installed in 1996 for a series of performances of Wanger’s “Das Rheingold” and “Die Walkure”. Installing the system in an historic building presents particular challenges. The theater had already installed a set of 10 loudspeakers in the back wall of each ring. Eight additional loudspeakers were installed in a circle around the domed ceiling, and a pair of subwoofers were installed over the proscenium. The ceiling speakers and the speakers in the rings were driven by separate Lares frames, so each could be balanced separately under computer control. The Lares systems were driven by two hypercardioid microphones installed high over the orchestra pit.

The ring system and the ceiling system were separately equalized for flat feedback transfer between the loudspeakers and the pickup microphones using the built-in calibration programs in the Lares system. The overall balance was then adjusted for uniform coverage through the hall. It turned out to be possible to achieve a +- 1.5dB uniformity.

During rehearsals for “Das Rheingold” we learned that the adjustment of the system was critical, and that the ability to separately adjust the performance of the system at different frequency bands was vital. We found that the intelligibility of the singers must be preserved at all times. At first, this meant reducing the system level to the point where there was little improvement for the orchestra. However, we realized that the frequencies that convey the most information in speech and singing lie between about 700Hz and 4000Hz, and the majority of the orchestral energy lies in the fundamentals of the musical tones. These fundamentals lie chiefly below 500Hz. Thus in theory it is possible to increase the envelopment for the orchestra without compromising the acting. The system needs to be frequency dependent. With the help of the artistic staff, including the music director, we found an equalization that did the job – about a 6dB reduction in reverberant level above 500Hz.

With the equalization the orchestral sound was greatly improved. All the instruments became richer, and the sound spread out from the pit and surrounded the audience. The measured reverberation time rose to 1.7 seconds at 500Hz, somewhat less above and somewhat more below. Compared to the original house these changes are enormous – but to an untrained listener the sound was completely normal. The following performances brought critical praise – particularly for the orchestral sound, and no complaints.

A permanent installation was completed in March of 1997, and has been in continuous operation on every performance since that time. In the permanent system some of the loudspeakers in the rings were replaced with cardioid loudspeakers above the door frames, a solution that solved some problems with hot spots directly in front of the earlier speaker positions. Two independent subwoofers were installed at opposite sides of the ceiling dome. The system was installed entirely by the house staff, using available equipment were ever possible. In spite of (perhaps because of) the low budget approach the system fits the hall well. The system is particularly beneficial for ballet, where the reverberation time is raised to 2.0 seconds, and less equalization is used. Critical reception continues to be excellent.

The relative simplicity and low cost of this system depend on several factors unique to the Staatsoper. One is the skill and dedication of the resident sound staff under Albrecht Krieger. Another is the excellent speech intelligibility everywhere in the hall. Thus no conventional acoustic modification was needed to achieve this vital goal. The relatively low level of early lateral reflections in this case contributed to a sense of intimacy and connection between the performers and the audience. The small proscenium opening and the very high use of the theater also contribute. Because there is always a new production in the wings, the upper stage house is always full of absorbent curtains and sets. The natural reverberation time of the stage house is thus usually quite low. An actor can move from far downstage to far upstage with only a moderate change in the reverberant quality of the voice. This simplifies the pickup problem. A single pair of microphones can successfully capture both the actors and the orchestra.

## Hummingbird Centre, Toronto

The Hummingbird Centre – formerly the O'Keefe Centre – is the home of the Royal Canadian Opera company, and the site of many music and ballet performances. It is a very large hall, with 3200 seats. In spite of the large hall volume the reverberation time is low, about 1.2 seconds. The unaugmented sound in the stalls is lacking in envelopment and reverberation, while the sound in the balconies is weak and muddled.

A Lares system was installed in the spring of 1998. The system uses four Lares mainframes, and four B&K cardioid microphones as pick-ups. The 312 loudspeakers are hidden in the procenium, in the diffusing elements along the side walls, in the ceiling of the hall, and under the balconies. The hall is divided into four separate time delay zones for the reverberation and direct sound reinforcement.

In the Hummingbird there are two problems. The envelopment needs to be augmented in the stalls and under the balconies. But at the same time the loudness and intelligibility needs to be raised throughout the house. Here is where the use of close directional microphones is helpful. By using these microphones it is possible to reinforce the direct sound without increasing the reverberant level in the hall through feedback.

The Lares software allows a direct sound reinforcement with some feedback reduction to be mixed in with the later reverberation. This feature was used in the proscenium and in the later time zones in the Hummingbird with good effect.

The equalization we found useful in the Staatsoper turned out to be equally useful in the Hummingbird – and in several other installations. Since this equalization is based on the properties of human perception, it is likely to be generally applicable.

Once again critical responses to the hall has been very favorable. The system is in use for all music performances (except for modern musicals and operas that include electronic amplification of all the instruments and voices.)

## The Circle Theater, Indianapolis

The Circle Theater was originally a Vaudeville house, converted to a concert hall for the Indianapolis Symphony. It seats about 1800. The natural reverberation time is low. A Lares system was installed with Jaffe Holden Scarborough Acoustics. It consists of three Lares frames, one dedicated to early reflections in the front of the hall, and two to the overall late reverberation. The reverberation under the balconies is controlled by one system, and in the stalls and over the balconies by another.

The primary purpose of the system is to augment the later reverberation, thus increasing the musical support and envelopment. During the final adjustment I spent some time with Paul Scarborough listening to the orchestra in many seats in the hall, while turning on and off the augmentation of early lateral reflections. These reflections are chiefly responsible for increasing a sense of distance and blend in the early sound of the orchestra. We both concluded that the augmentation was worthwhile. However for me the effect was not essential for the enjoyment of the music. Without the early reflection augmentation the sound of the orchestra was a bit too direct or "in the face." With the early reflections the sound was a bit better blended and more gentle. However without it doubt it was the later reverberant energy supplied by the system that was essential. Sound throughout the hall was richer, more enveloping, and better balanced with the system on. Once again the system has achieved critical acclaim. The major orchestral critic in Indianapolis immediately noticed the system and wrote a glowing review of both the orchestra and the hall.

### Adelaide Festival Center Theatre (AFCT) – Adelaide, Australia

The AFCT is an example of what can be done with a large hall when there is adequate funding. The hall has 2200 seats, and an occupied natural reverberation time of 1.2 seconds. As in many opera houses the stage house was too reverberant and the audience house was too dry. About 350 sq. meters of absorption was added to the stage house to control the reverberation there, and 250 sq. meters of absorption was added to the stalls to correct low frequency problems. Additional absorption was added to the rear walls of the rings to reduce focusing of sound back to the stage. Carpets were removed and replaced with wooden flooring. The result was an articulate room with low natural reverberation time.

The Lares system consisted of five Lares frames, two for early energy throughout the hall, two for late energy, and one for reverberation to the stage. Six B&K cardioid microphones were used, and 288 loudspeakers. The adjustment and calibration of this system was performed by the local staff with the help of my colleague, Steve Barbar. The flexibility of adjustment is very high, putting a large burden on the ears and the knowledge of the personnel making the final selection of settings.

The system has been in use for an entire performance series of the Wagner "ring", and for several purely orchestral performances. The control system allows easy selection of settings of the system for opera, ballet, or orchestra. The critical response to the system has been very good. Glowing reviews appeared in newspapers in both Australia and London, with several reviewers commenting that the acoustics were better than any venue in London.

## CONCLUSIONS

Electronic enhancement of spaces for musical performance has finally come of age. Multichannel time variant reverberation technology has made it possible to design practical systems that produce very significant improvements with no artificial artifacts. The success of these systems depends, as always, on careful consideration of the real acoustic needs of the hall in question, and the resulting design and installation of the enhancement system. Through the use of frequency dependence, and by controlling the undesirable reflections in the range of 50 to 150ms, it is possible to create acoustics that combine both high intelligibility and high envelopment. The result is musically very effective, particularly for opera.

REFERENCES

1. Griesinger, D. “The Psychoacoustics of Apparent Source Width, Spaciousness and Envelopment in Performance Spaces” Acta Acustica Vol. 83 (1997) 721-731
2. Griesinger, D. “Design and Performance of Multichannel Time Variant Reverberation Systems” Proceedings of the Active 95 conference, Newport Beach CA 1995 p1203-1212