Audio Engineering Society

Convention Paper

Presented at the 123rd Convention
2007 October 5–8 New York, NY, USA

Room Reflections Misunderstood?

Siegfried Linkwitz

Linkwitz Lab, Corte Madera, CA 94925, USA
sl@linkwitzlab.com

ABSTRACT

In a domestic living space a 2-channel monopolar and a dipolar loudspeaker system are compared for perceived differences in their reproduction of acoustic events. Both sound surprisingly similar and that is further enhanced by extending dipole behavior to frequencies above 1.4 kHz. The increased bandwidth of reflections is significant for spatial impression. Measured steady-state frequency response and measured reflection patterns differ for the two systems, while perceived sound reproduction is nearly identical in terms of timbre, phantom image placement and sound stage width. The perceived depth in the recording is greater for the dipole loudspeaker. Auditory pattern recognition and precedence effects appear to explain these observations. Implications upon the design of loudspeakers, room treatment and room equalization are discussed.

1. INTRODUCTION

Much of music listening over two loudspeakers takes place in ordinary living rooms with greatly different furnishings, decorations and loudspeaker placements that express the priorities, life style and taste of their owners. Acoustic room treatment is usually provided by the "normal stuff of life". The sound from both loudspeakers impinges upon each ear, left and right, as do the loudspeaker sounds that are reflected from the various surfaces and objects in the room. The received sound at the listening position is a combination of the direct sounds from each loudspeaker and the delayed and filtered reflected sounds. From this mixture of sound waves the ear-brain system creates, for example, an impression of a particular musical instrument, an orchestra in its acoustic space, or a popular human voice at close range. Furthermore, an impression of the spatial arrangement of various sound sources and of the acoustics of the recording venue can be generated, if the two loudspeakers and the listener are properly located in the room. This is a rather amazing phenomenon that has no precedence in the gradual evolution of natural hearing. For example, not sufficiently often have the sounds from two roaring lions been similar enough to locate as one lion somewhere between the two. Yet this is the type of illusion that a 2-channel stereophonic setup is asked to create in our mind.

It has been found that the effectiveness of the illusion depends upon minimizing the cues that give it away as sound coming from two locations. Cues are usually contributed by the on-axis and off-axis frequency response of the two loudspeakers, by cabinet resonance, stored energy, and by the generation of new spectral content due to non-linear amplitude response. The room may contribute misleading cues by selectively emphasizing or reducing certain low frequencies at the listening position, or by reflecting differently filtered versions of the loudspeaker's direct sound from different directions. The illusion of acoustic space and phantom images can be strengthened by predistorting the loudspeaker signals to cancel their crosstalk at the ears. This technique requires listening from a precise location. While impressive at first, the experience soon becomes tiring which indicates that something is unnatural. The brain is subconsciously engaged to process misleading information, such as the uncompensated reflected signals in the room.

It is widely assumed that the effect of the room will be minimized if the loudspeaker is highly directional and the majority of its output is aimed at the listener. Audio covers a 1000:1 range of wavelength, from 17 m at 20 Hz to 17 mm at 20 kHz. It makes building broadband directional loudspeakers difficult. Below 500 Hz a dipole becomes the only practical directional source that fits a typical living room. Planar dipole loudspeakers have been in use for a long time. They are usually respected for their natural sound reproduction capabilities. Their placement in the room tends to be critical because they are acoustically large sources over most of their frequency range. Acoustically small dipole loudspeakers, though, can be built by using conventional voice coil drivers [1]. It has been found that their room placement is less critical as long as it follows a minimum set of criteria.

If a wideband, uniformly directional loudspeaker leads to reduced interaction with the room, then an omni-directional loudspeaker should have maximum interaction. To obtain omni-directional behavior a source must be small compared to the radiated wavelengths. Thus it will also exhibit few baffle diffraction effects that might affect imaging. It was thought that a listening comparison between an omnidirectional, or monopolar, source and a dipolar, small acoustic source in a normal living room would show up differences in phantom image creation due to different acoustic illumination of the room by each pair of loudspeakers. An omni-directional, monopolar loudspeaker was designed and built to investigate these postulates when compared to an existing dipolar loudspeaker.

1.1. Loudspeaker Configurations

The monopolar source M is constructed as a 3-way loudspeaker and uses a 40 mm tweeter, a 110 mm midrange driver and a 200 mm effective diameter woofer. (Figure 1). Crossovers are at 1 kHz and 100 Hz with 4th order Linkwitz-Riley acoustic filter slopes. Each driver has its own power amplifier and electronic equalization. The woofer driver is in a separate sealed box and downward firing. The midrange driver is mounted at the end of a sealed pipe and upward firing. The pipe eliminates panel resonances due to its extreme stiffness. Internal resonances are attenuated so that re-radiation of the backwave through the cone is attenuated by 40 dB relative to the direct signal [2].

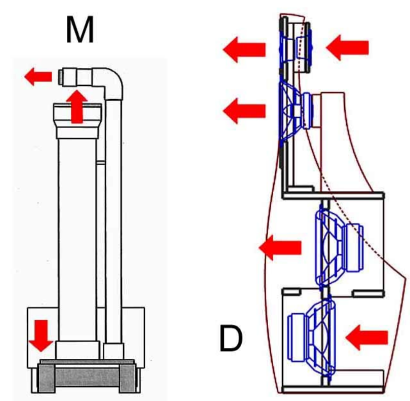

Figure 1: Acoustically small monopolar and dipolar sources M and D. The arrows indicate the direction of synchronous piston movement.

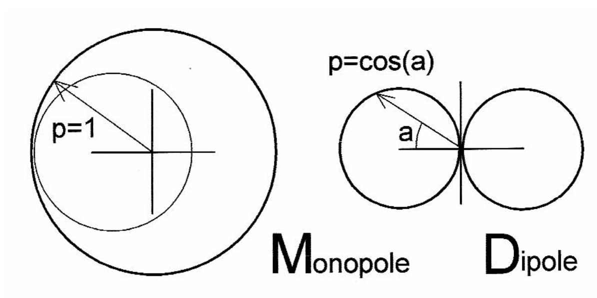

Figure 2: On-axis and off-axis frequency response level for monopolar and dipolar sources.

The tweeter is forward firing. It becomes increasingly directional above 3 kHz due to its piston diameter. (Figure 2). The dipolar loudspeaker D is a 3-way system with two back-to-back 25 mm tweeters, an open baffle mounted 165 mm midrange driver, and two 210 mm effective diameter woofer drivers in a ducted open baffle [3]. Crossovers are at 1.4 kHz and 120 Hz with 4^{th} order Linkwitz-Riley acoustic filter slopes. Each driver has its own power amplifier and electronic equalization. At higher frequencies, where midrange and tweeter drivers become directional due to their large piston diameter, dipole like behavior is preserved even when there is no longer direct interference between front and rear radiation from diffraction around the baffle edges. It requires a rear-facing tweeter, though.

Both loudspeakers have been equalized for a flat on-axis frequency response. It was measured under free-field (4π) conditions with the units outdoors and raised off the ground. The woofers were integrated into the flat higher frequency response by a ground plane measurement (2π) of each system. This accounts for the floor reinforcement in a typical room setup. The 4π to 2π transition was equalized by a 100-200 Hz, 6 dB/oct shelving highpass filter for D, and with a woofer level adjustment for M. Off-axis measurements in horizontal and vertical planes were taken to assure a consistent polar response over 360 degrees horizontally and +/-30 or more degrees vertically.

1.2. Loudspeaker and Room Setup

The listening tests took place in the author's living room. It is a large room that is enclosed on three sides and extends towards a kitchen area and a hallway in the back. (Figure 3). No special acoustic treatment has been used and the listening area is acoustically fairly live with RT60 around 500 ms. Both loudspeaker systems are set up with >1 m distance from the side walls and the wall behind them. (Figure 4). Listening tests and measurements were performed at locations A and B, where A is at the apex of an equilateral triangle that is formed by the loudspeakers and the listener. The monopolar loudspeakers are closer to A to obtain a similar direct-to-reverberant sound level ratio as for the dipolar loudspeakers. Under ideally reverberant conditions the distance M-A should be 1/sqrt(3) = 58% of distance D-A. The subtended angle from the listener to each loudspeaker is identical for M and D to ensure similar head-related- transfer-functions.

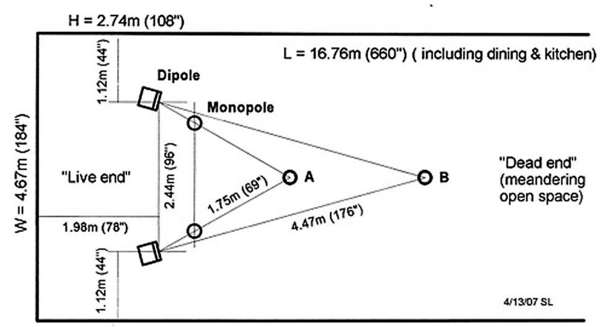

Figure 4: Layout of loudspeakers D and M for a listener position A at the apex of an equilateral triangle.

Figure 3: The living/listening room with loudspeakers M and D and listening positions A and B as seen from opposite directions.

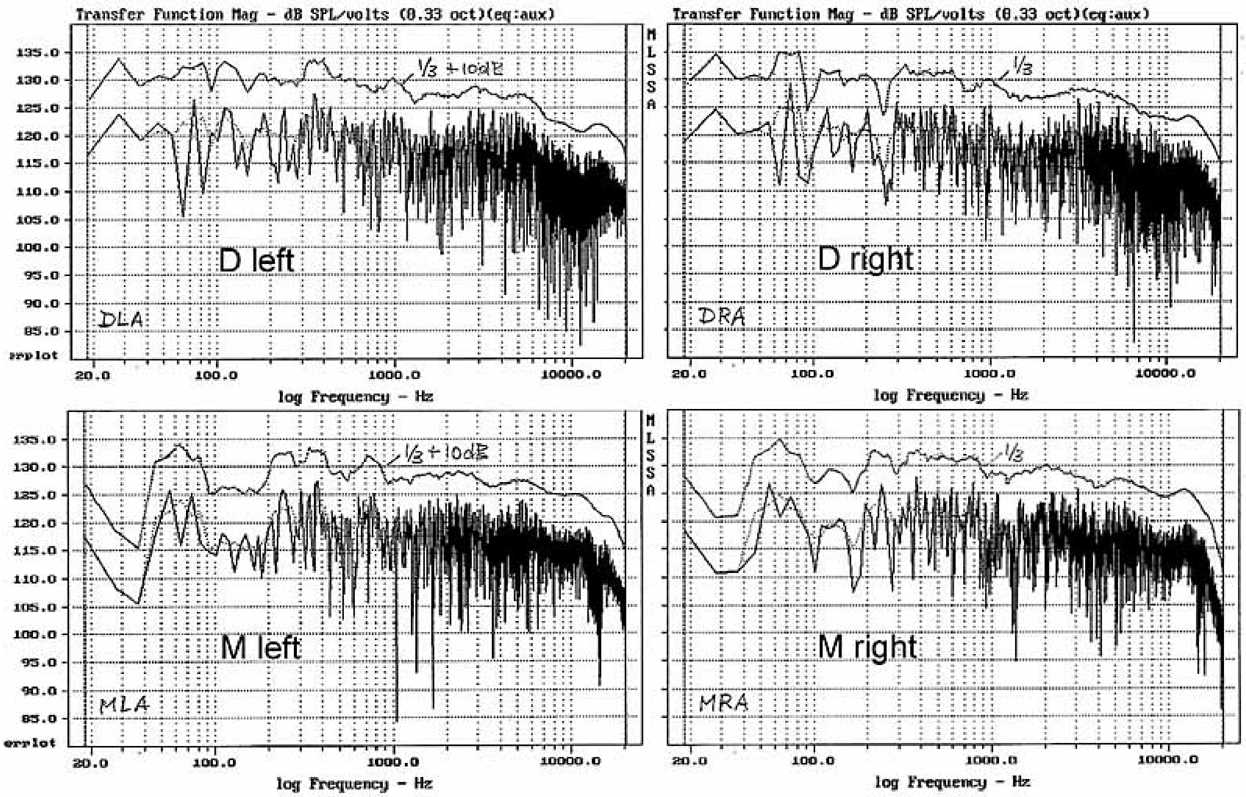

Figure 5: Frequency response in the room at location A for left and right loudspeakers D and M. Calculated from 200 ms long impulse response. The upper traces are 1/3rd octave smoothed and offset by 10 dB.

1.3. In-Room Frequency Response Measurements

Both loudspeaker systems exhibited a similar, flat on-axis frequency response when measured outdoors, but lose this similarity when measured with the same microphone from location A in the room. The frequency response graph shows the transformed data from the first 200 ms of a longer impulse response. This is barely a long enough rectangular time window to resolve low frequency room resonances, which have a bandwidth in the order of 3-5 Hz and to indicate their full magnitude. At high frequencies the measurement results fluctuate wildly because numerous room reflections are captured in the 200 ms time window. The response also rolls off above 1 kHz unlike the free-field measurement. In-room measurements are used by some as the basis for equalizing a loudspeaker to a flat response at the listening position. This works to some extent for very low frequencies, but it leads to an unnaturally bright top end. Thus great care must be taken to interpret the in-room response for its audible consequences. The visual differences between loudspeakers M and D would be significant if they were measured in a reflection-free environment. Surprisingly, and subjectively, in terms of timbre or perceived frequency response, both loudspeakers sound almost identical on program material. The main differences are in their spatial presentation.

When the initial listening tests were performed loudspeaker D did not have a rear-firing tweeter. Thus it was not behaving like a dipole above 1.4 kHz. The close similarity between D and M was slightly marred in the upper voice range, where there was a preference for the monopolar source M, particularly on female soprano. Attempts to equalize the frequency response of D did not give the desired result. The addition of a rear tweeter eventually shifted the preference back to D. Free-field frequency response measurements in the frontal hemisphere did not show any contribution from the rear tweeter for angles of up to 60 degrees off-axis. Thus any audible difference in the room was due to reflected rear radiation. (Figure 6). It had been thought previously that such room contribution to the response at the listening position could only be detrimental. Therefore loudspeaker D was originally designed without a rear tweeter.

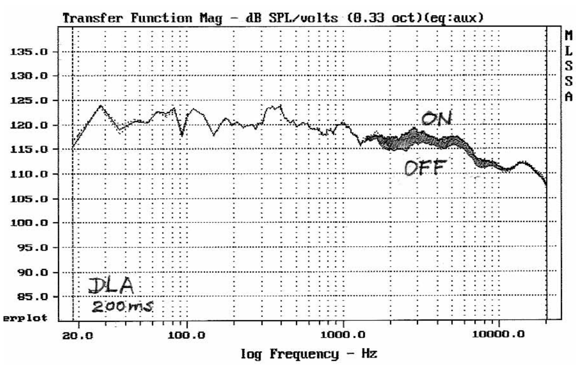

Figure 6: Effect of the rear tweeter on the 200 ms in-room frequency response of the left loudspeaker D.

2. ROOM REFLECTION OBSERVATIONS

In the following the differences and similarities in reflection patterns for the two loudspeakers M and D are investigated. Measurements and listening observations will be compared. For example, the author and others had observed that wide horizontal dispersion in the high frequency range of a loudspeaker is subjectively desirable and increases fidelity [4]. Conclusions about optimizing the loudspeaker-room-listener interaction will be drawn from measured data, informal listening tests and from published studies of sound perception for multiple sources with reflections.

2.1. Test Signal for Reflection Measurements

The polar response of loudspeakers is usually frequency dependent. The absorptive, diffusive and reflective properties of room surfaces can be frequency dependent as well. Thus reflections are best measured with a bandwidth limited test signal such as a shaped toneburst [5]. To resolve multiple reflections the duration of the burst must be short compared to the travel time difference between reflected bursts.

A 4-cycle Blackman windowed toneburst is well suited to typical room dimensions and for the frequencies of interest. Its spectrum is one octave wide at –9 dB. (Figures 7, 8). The received, reflected tonebursts can be displayed directly on an oscilloscope without further signal processing. The presentation lacks dynamic range, though. Calculating the envelope of the toneburst, i.e. the magnitude of the analytic signal, and displaying it on a logarithmic amplitude scale leads to a more

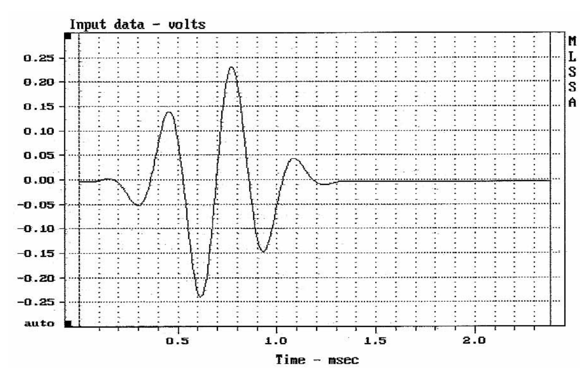

Figure 7: Blackman windowed 4-cycle tone burst of 1.33 ms duration at 3 kHz.

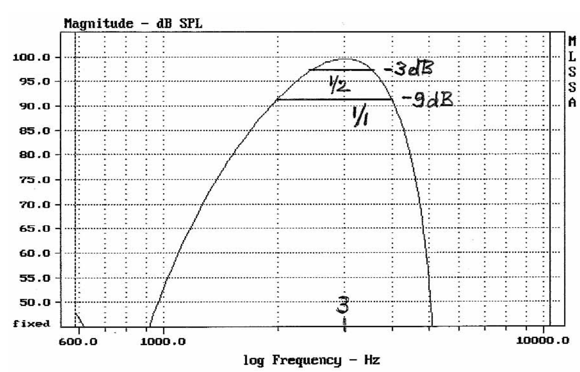

Figure 8: Spectral content of a 4-cycle Blackman windowed toneburst with one octave width at –9 dB.

informative presentation. Processing artifacts though show up at low signal-to-noise ratios. Alternatively the full-wave rectified signal can give an unambiguous display but only over a limited linear amplitude range. (Figure 9).

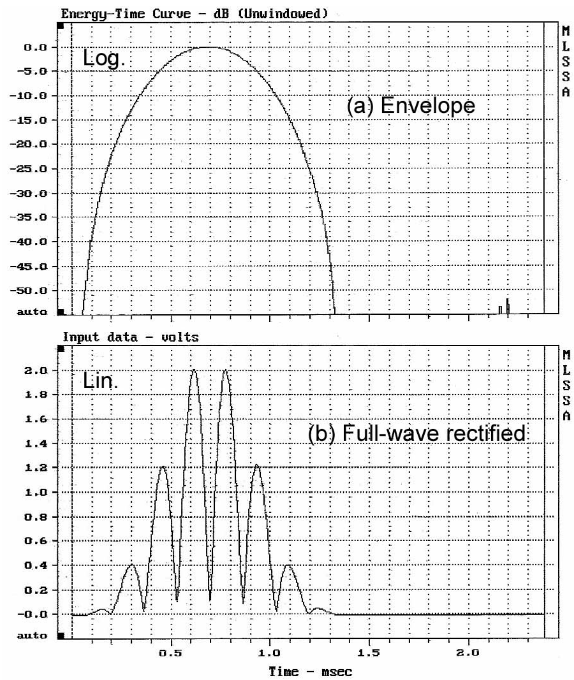

Figure 9: Display of the received toneburst by its envelope (a), and as full-wave rectified burst (b).

2.2. Room Corner Model

A loudspeaker that is placed near the corner of a room where floor, sidewall and rear wall intersect will have a number of possible reflections. (Figure 10). Of concern to some audiophiles are primarily the floor F and sidewall S reflections, which usually arrive at the listener soon after the direct signal. The rear wall reflection R can be delayed and reduced in magnitude by moving the loudspeaker out into the room. The double reflections between side wall and floor S+F, rear wall and floor R+F, and rear and side wall R+S, are not necessarily negligible, nor is the triple reflection R+F+S. Table 1 gives the strength and delay of the various reflected signals relative to the direct signal. The values were calculated for the left loudspeaker D and listening position A. (Figure 4). While in reality D sees a fairly uncluttered environment the actual surfaces are not perfectly reflecting as assumed in the Table. Also there will be additional reflections from the ceiling and opposite walls leading to an infinite series of reflections. All reflections decay over time because sound energy is absorbed and diffused with every surface encounter and path through the air.

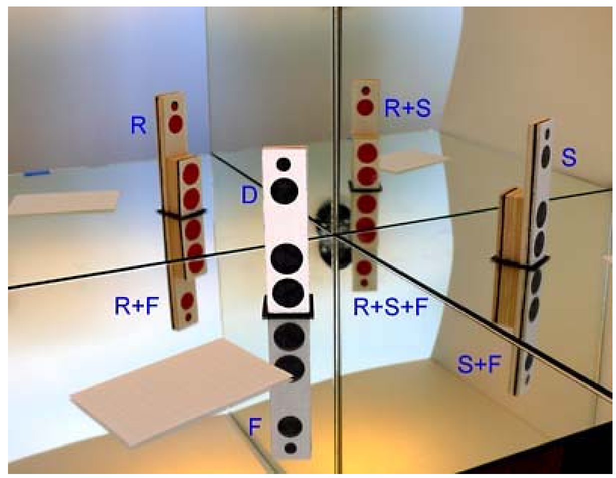

Figure 10: Loudspeaker in front of a perfectly reflecting room corner.

Speaker D - left	Distance m	rel. Delay ms	Strength dB
Direct	2.4	0.0	0.0
Side	4.0	5.0	-4.2
Rear	5.9	11.5	-7.7
R+S	6.7	14.0	-8.8
Floor	3.2	2.6	-2.5
S+F	4.5	6.8	-5.3
R+F	6.3	12.7	-8.3
R+S+F	7.0	15.1	-9.2

Tweeter height = Listener height = 1.1 m
Ideal dipole polar response

Table 1 Reflected signal timing and strength

2.3. Rear Tweeter Measurements and Audibility

The rear tweeter of loudspeaker D only contributes to the measured response at A via room reflections. Its output is in opposite phase to that of the front tweeter, which causes cancellation at 90 degrees off-axis angle to preserve dipolar behavior. This is the normal ON configuration. (Figure 11). With the rear tweeter turned OFF the magnitude and density of the reflections in the displayed 50 ms time window is reduced.

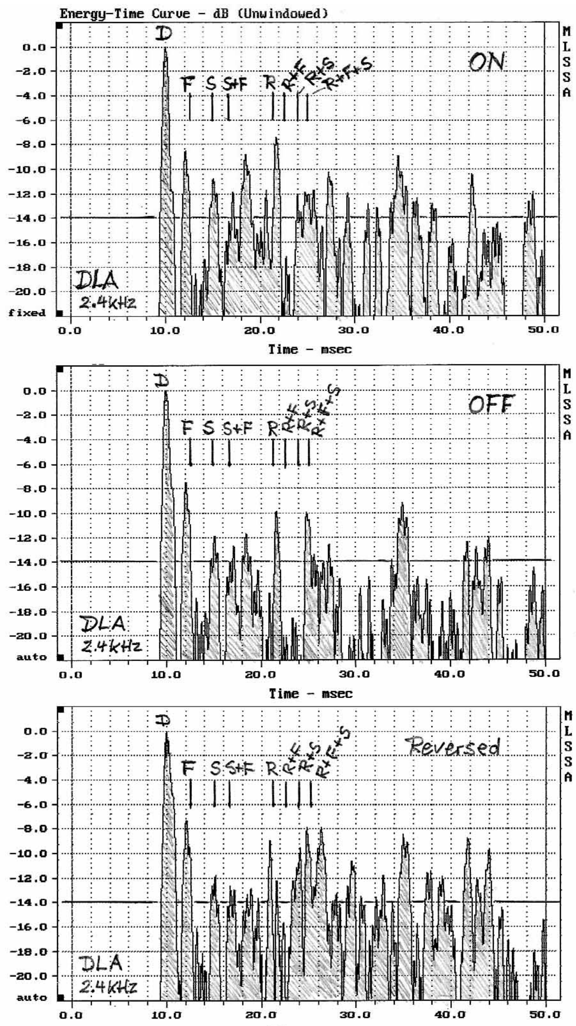

Figure 11: Envelope of the initial 50 ms time response to a 2.4 kHz burst from left loudspeaker D.

With “Reversed” or monopolar polarity the rear tweeter adds to the front tweeter output at 90 degrees and other off-axis angles. Consequently the room reflection pattern observed at D becomes even denser than for the normal ON condition. There is some correlation between various reflection amplitudes and their timing to the values in Table 1. Many more reflections are observed in the furnished room than those from the simple corner model. Some reflections are as large as 40% (-8 dB) of the direct signal level.

It would seem obvious from Figure 11 that the addition of a rear tweeter in whatever polarity should have detrimental effects upon the accuracy of reproduced sound because it adds reflected sound. The room participates more strongly in what is measured and perceived at A. As had been observed, though, the addition of a rear tweeter improved the perceived accuracy of loudspeaker D in the voice range. The similarity between loudspeakers M and D was increased. When the rear tweeter polarity is monopolar, phantom imaging deteriorates.

Surprisingly, the improvement of D extended beyond the voice range. Perceived high frequency energy is increased and the combined tweeter level had to be reduced by about 1 dB relative to a forward-firing tweeter only. The level setting is critical. A +/-0.25 dB change has a significant effect when assessed by long-term listening and conclusions are more reliable than from an instant A/B comparison. The integration of the rear tweeter also pointed to a now desirable frequency response correction of about 0.5 dB over an octave around and below 400 Hz. Much of this “tuning” was done in collaboration with Don Barringer [6]. His loudspeakers D are set up in a smaller room. These same adjustments to duplicate loudspeakers D at other locations and in similar setups gave similar perceptual results according to their owners [8]. This would seem surprising since the specific room acoustics are bound to be different. There must be an overriding commonality then, which makes such re-occurring similarity of perception possible.

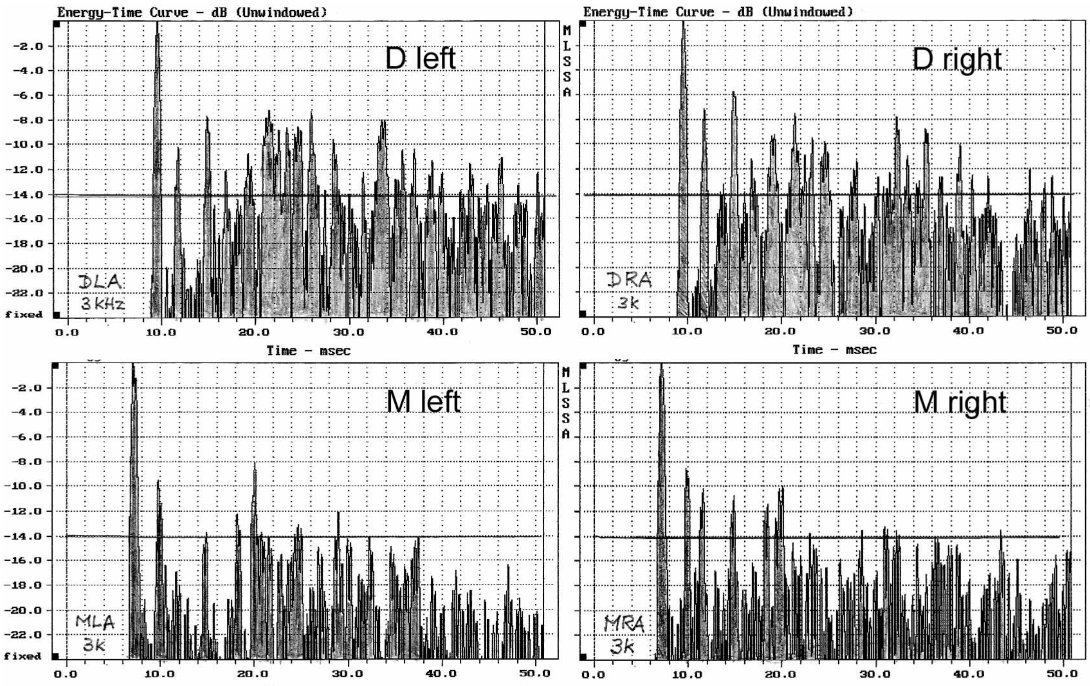

Figure 12: Envelope of direct and reflected 3 kHz toneburst from left and right loudspeakers D and M at microphone position A during the first 50 ms.

2.4. Room Reflections at the Listening Position

The two loudspeakers M and D had captured the author's attention by their unexpected likeness of sound reproduction, which did not seem plausible. Under free-field conditions essentially the same on-axis frequency response was measured for both loudspeakers, but their off-axis response is very different. Therefore the room reflection pattern should also be different and with it the perceived sound.

The burst response envelope measurements at A indicate a loudspeaker placement asymmetry with respect to the room reflections. (Figure 12). In practice there is a trend for the phantom soundstage to be shifted slightly to the right, but this is highly program material dependent. Recordings are not always spatially centered. The relative level of reflections is lower for M. That loudspeaker is positioned more closely to the microphone, which increases the direct-to-reflected amplitude ratio. If the loudspeaker were truly omni-directional and the room ideally reverberant to the 3 kHz burst excitation, then the reflection density should be higher for M than it is for D.

The power spectrum of the 50 ms burst response fluctuates around the power spectrum of the source burst for both D and M. (Figure 13). The reflections do not change the overall shape of the spectrum and must be fairly close copies of the direct sound for this to happen. Distortion and ambient noise limit the measurement dynamic range, particularly for the tweeter of M.

As a side note, the onset of distortion is readily heard with a burst signal.

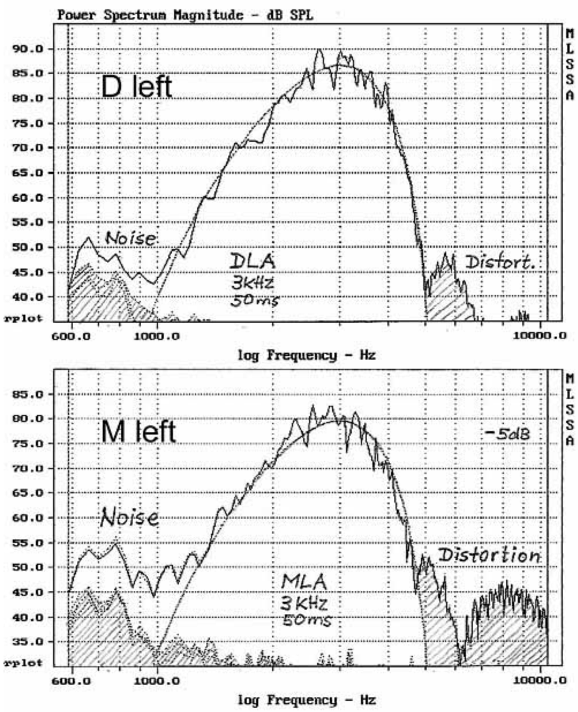

Figure 13: Power spectrum of 3 kHz burst from left loudspeakers in Figure 12.

2.5. Reflections at Greater Distance

A comprehensive picture of how the reflections decay over time is obtained by increasing the observation window to 400 ms and by moving the microphone farther out into the room to position B. At this location the direct signal has decreased in inverse proportion to the relative change of distance. The reflected signal amplitudes become more dominant. (Figure 14). Their rate of decay can be estimated by drawing a trend line, for example at 3 kHz and at 800 Hz burst frequency. There are however too few reflections in this acoustically small living room to obtain reverberation time numbers that can be repeated with low uncertainty for different locations in it. Reverberation time is a useful measure for large spaces like concert halls or churches, but possibly not for a listening room.

It can be seen that at 3 kHz loudspeaker M generates less reflected energy than D. This is a consequence of the increasing directivity of the forward pointing tweeter with its relatively large diameter. At 800 Hz, though, in the upward pointing mid-frequency driver's range, M produces a stronger series of reflections than D.

When listening to program material from location B the phantom imaging precision is greatly reduced for both loudspeakers, but with D suffering less. The similarity of timbre between the two loudspeaker types however remains strong. All of the author's listening to background music or to news is from place B. For full involvement the presentation at A is far more convincing

2.6. Reflections at Different Frequencies

A measurement of the reflection patterns at A when carried out over a wider range of toneburst frequencies shows a graininess in the overall time response envelope. (Figure 15) The peaks shift in time and they increase in width as frequency is lowered. This is a result of the increase in burst cycle duration and the changing interference between individual reflections, which add and subtract from each other. The graph of the half-wave-rectified microphone output signal shows both the individual burst cycles and to some extent also their envelope. At the lowest frequencies and when the frequency is properly tuned, a long and smoothly decaying tail of the initial burst indicates a room resonance.

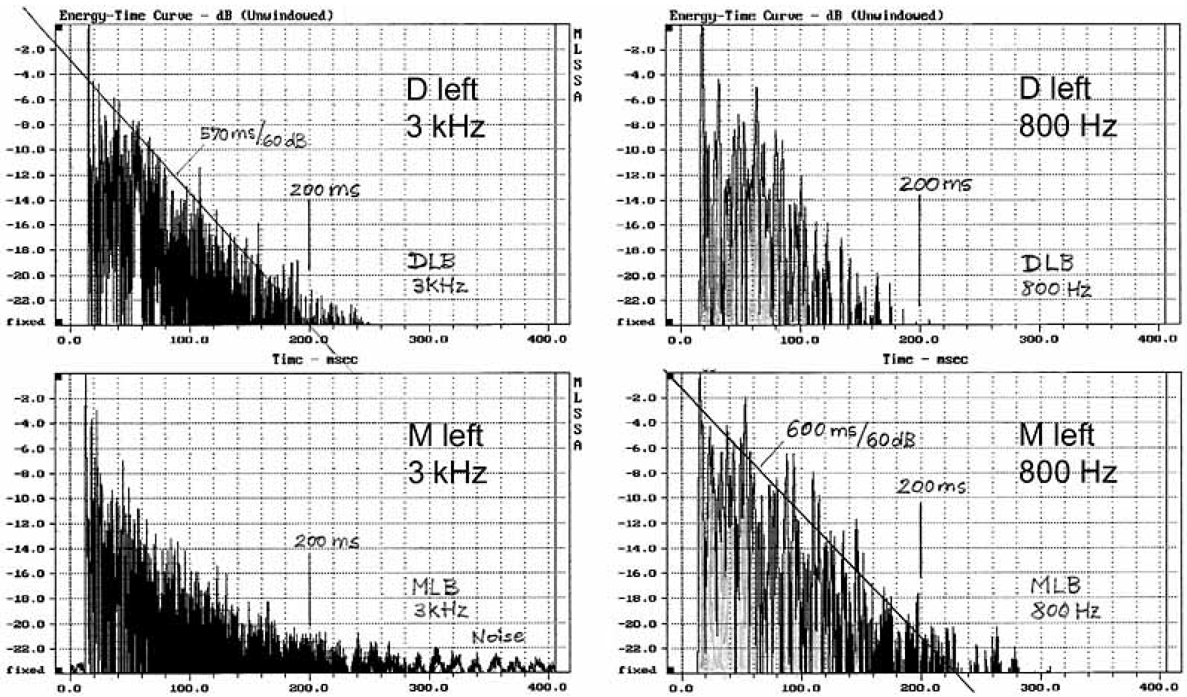

Figure 14: Burst response envelope for increased microphone distance at position B and in a 400 ms time window.

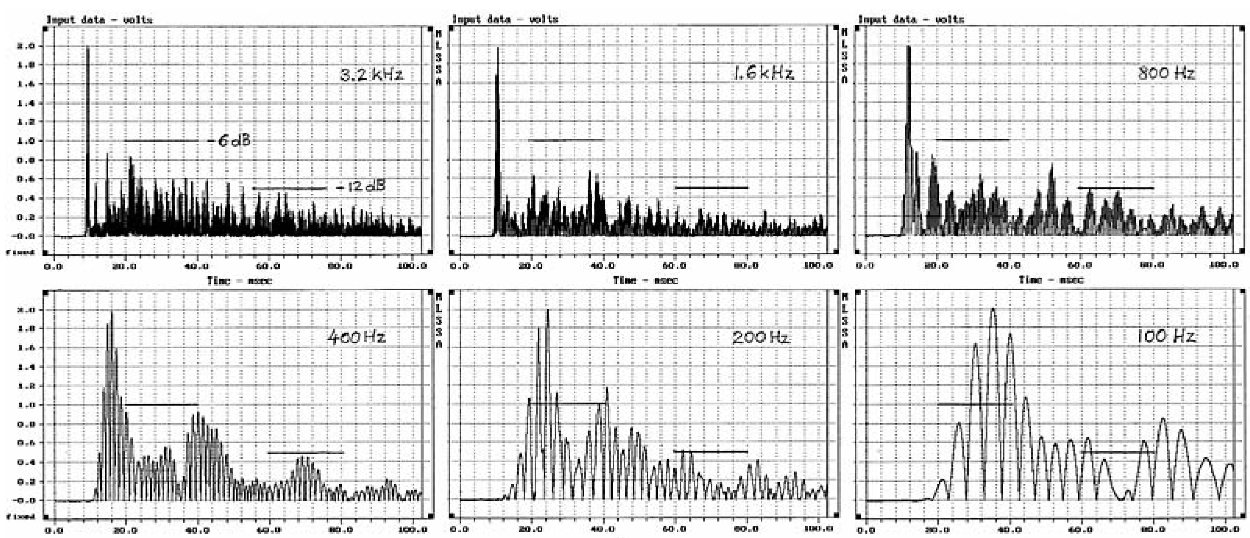

Figure 15: Burst response of loudspeaker D at microphone position A. The full-wave rectified burst response is shown on a linear amplitude scale in a 100 ms time window

2.7. Perceptual Comparison between Loudspeakers D and M

Both loudspeaker types render program material in almost identical timbre. It would be difficult to describe differences other than in the bass region. Here the sealed box woofer of M excites room resonance modes differently from the open baffle woofer of D and exhibits some unevenness, though less so than was expected.

Furthermore, M retains much of the clarity and transparency of the dipole D. This came as a real surprise, because the individual drive elements were of lower cost and higher non-linear distortion than those in D. It led the author to believe that there exists a distortion threshold that is "good enough" and which is higher than was assumed. The absence of secondary radiation from the enclosure surfaces and from behind the cones certainly should help to give M a box-less sound. At very high volume levels the small drivers in M will distort and the loudspeaker loses the effortlessness that characterizes D at the same output level.

It was expected that M would image precisely because it is almost an acoustic point source. This is indeed what was observed. Lateral phantom image placement is pinpoint like. The sound stage is tall, almost like for D, but the depth and layering of it is much less pronounced. Here is the domain where the dipolar system gives a much stronger impression of the recording venue space and its depth in addition to establishing a realistic sounding phantom image placement.

Some audiophiles have claimed that the perceived sound stage depth corresponds to the distance from the loudspeaker to the wall behind it. That claim is mistaken. The wall behind the loudspeakers as well as the loudspeakers themselves completely disappear on many recordings of live events where apparently the venue acoustics are sufficiently embedded in the sound to recreate the sense of space. A dipole illuminates that wall more strongly than a conventional box type loudspeaker and this might contribute to a stronger sense of depth and openness. A monopolar or omni-directional loudspeaker produces a similar effect.

It is safe to state that reflections affect what is perceived by the listener. The question then becomes:

- How could two loudspeakers so different as D and M sound so similar in a reflecting environment?

- How can two loudspeakers create a spatial impression of a recording studio, a concert hall, a church or even the outdoors inside a usually much smaller listening room that has its own acoustics?

The answer is not obvious from the in-room measurement results, but there must be a clue in the beneficial rather than detrimental effect of the rear tweeter and the associated change in room reflections. With the addition of a rear tweeter the rear radiation pattern of D becomes consistent over the full frequency spectrum and closely resembles the frontal radiation pattern. Similarly the radiation pattern from M is consistent over most of that loudspeaker's frequency range and only becomes forward pointing at the highest frequencies. The reflections generated by both loudspeakers then are essentially delayed and attenuated copies of the direct sound provided that the room surfaces are broadband reflective.

.

3. HYPOTHESIS

Previous observations from having designed different types of loudspeakers and having listened to them in a variety of rooms, and when combined with the most recent findings about reflections, have led us to a novel hypothesis about an optimum loudspeaker-room-listener interaction.

- The two loudspeakers and the listener should be set up symmetrically with respect to the room boundaries and with the listener at the apex of a symmetrical triangle.

- Reflections generated by the two loudspeakers should be delayed copies of the direct sound to the listener. The delay should be greater than 6 ms. The high frequency content of the reflections should not be intentionally attenuated.

- Under these conditions the direct sound from the loudspeakers dominates perceptually. The room interferes minimally with the spatial, temporal and timbral cues embedded in the direct sound and with the creation of a phantom sound stage between and behind the loudspeakers.

- Under these conditions the cognitive faculty of the brain is better able to separate the static listening room acoustics from the acoustics embedded in the recording which are presented dynamically by the two loudspeakers.

The hypothesis points to a number of requirements that are rarely fulfilled by current practices. For example and most importantly, the polar response of the loudspeaker must not change over its whole frequency range. Only then can the room reflections be copies of the direct sound. It means that the loudspeaker has to behave either as an omni-directional source, a dipole or a cardioid. These are the only radiation patterns that are practically realizable even for lower frequencies and their long wavelengths. For the reflections to be delayed relative to the direct sound the loudspeakers must be placed at least 1 m away from adjacent surfaces. Still, the first arriving floor reflection will have less than 6 ms delay. This does not appear to cause a problem. After all, almost every source that we hear has a floor reflection and we seem to use this information primarily to determine the height of the source above the floor.

The requirement for full spectral content of the reflections rules out the use of frequency dependent absorbers on the room surfaces. The various commercially available foam or fiberglass panels absorb predominantly higher frequencies only and would color the room reflected sound dynamically to where it no longer can be cognitively separated from the direct sound. This then argues for relatively live room acoustics that are determined by the "normal stuff of life" with which the room is filled and decorated, acoustics with which we are intimately familiar as normal.

At very low frequencies, below about 150 Hz, the room response may not be separable from the direct loudspeaker signal at specific modal resonance frequencies. A dipole bass loudspeaker, though, appears to suffer perceptibly less than a monopole source because it excites fewer modes due to its directionality [7].

The live-end, dead-end room arrangement practice, with loudspeakers at the dead-end and the listener at the live-end, is actually reversed from what the hypothesis requires. The absorbing end of the room is usually not broadband attenuating and increasingly reflects the lower portion of the spectrum. The live-end is often made diffusive which in reality only affects the higher frequencies. Thus the room assumes a peculiar acoustic character through its reflection pattern which cognitively would be difficult to separate from the sonic cues of the recording in the direct loudspeaker signals. Furthermore, the vast majority of box loudspeakers exhibits omni-directional radiation behavior up to a few hundred Hz and then becomes increasingly forward firing with increasing frequency. This consistency in frequency dependent polar response over many different consumer and professional products is the primary reason why box loudspeakers sound different from all forms of open baffle loudspeakers with their inherent dipolar radiation. The differences typically increase the larger the size of the box loudspeaker. Both types of loudspeakers impart their own generic signature to sound reproduction in a room.

4. THE PRECEDENCE EFFECT

The above hypothesis is based on extensive experiments and observations that were confirmed by several collaborators [8]. It should be further investigated for its range of validity. Many papers have been published about sound measurements in large performance spaces and their correlation to sound perception in those environments. It has not been shown which of those measurements apply to the acoustically small space of a living room. Here the issue is accurate sound reproduction via two loudspeakers and not sound production. A concert hall becomes part of the sound creation process together with the instruments of the orchestra. The listening room should not interfere with what has been recorded. Reverberation of sound due to very large numbers of reflections is exceedingly important in the concert hall. It is questionable whether this measure is even applicable to listening rooms and their relatively small number of reflections and high absorption. Low frequency room modes on the other hand are not an issue in concert halls but often are in listening rooms. The playback problem is complicated by the fact that real listening rooms can vary greatly in shape and reflective characteristics and present many variables to investigate. Studies are therefore done selectively. Thresholds of audibility for a single reflection from a single source in an otherwise anechoic environment have been studied extensively [9]. For this case of a single source and single or multiple reflections the acuity for localizing the single source suffers. This does not necessarily translate to localization in two-channel stereo in a reflective environment, which behaves as predicted by uncomplicated ITD calculation [10]. Furthermore, it had been found that loudspeakers with wide horizontal dispersion were preferred in listening tests [4]. What is known scientifically about loudspeakers and rooms for sound reproduction has been reviewed by F. E. Toole [11].

Common and underlying to all these investigations is some form of the extensively studied precedence effect. The effect must also be at play when listening to loudspeakers D and M, because the commonality in their direct sound is so dominant and the different room reflection patterns have little audible effect.

A comprehensive review of the precedence effect cannot be given here. It is instructive, though, to quote W. M. Hartmann from "Listening in a room and the Precedence Effect" [12]

- "The precedence effect makes its appearance in several guises: as a localization phenomenon, as the Haas effect, and as de-reverberation and de-coloration."

- "When the precedence effect operates, the combination of direct and reflected sounds is heard as a single entity, and the perceived location of the entity corresponds to the direction of the direct sound. … Reflections add a sense of "space filling" and loudness to the sound as a whole, but the reflections are fused with the direct sound."

- "The integration of a direct sound with a reflected sound, according to Haas, was neatly described by Green [13]: If one stands in a room 1m from a reflecting wall and creates an impulsive sound, by clicking two rocks together, there is a reflection from the wall that arrives 6 ms after the direct sound. One never hears that kind of reflection. On the other hand, if a listener wears headphones and hears two clicks in one ear separated by 6 ms, the listener immediately hears two well-separated clicks. The suppression of the reflection that takes place in the room, but not with headphones, can be called the precedence effect."

- "A third viewpoint concerns an effect that might be called "de-reverberation". Unlike the localization precedence effect, there is no standard defining experiment for de-reverberation, but the idea is simply that we are not normally much aware of reverberated sound, even though the energy in the reverberated sound may be several times larger than the energy in the direct sound."

Studies of the so-called cocktail party problem, i.e. how we recognize what one person is saying when others are speaking at the same time, might give further clues to what we hear from two loudspeakers in a room. In particular the following quote from William A. Yost, "The Cocktail Party Problem: 40 Years Later", has relevance to the full-spectrum reflection requirement in the hypothesis above [14].

- "As the sound from reflective sources becomes de-correlated (in Blauert's terms incoherent) from that of the original source, the listener perceives both sources as the precedence effect breaks down. Divenyi (1992) has studied the precedence effect under conditions in which the original and reflected sources were spectrally incoherent. When the source and its echo are spectrally different, the location of both the original source and the echo is perceived; however, the location of neither is as accurately determined as when each is presented as a separate source."

The requirement to place speakers at least 1m away from reflecting surfaces (measured from the tweeter) had been found empirically for box and panel type loudspeakers. All reflections are then delayed at least 6 ms, except for the floor reflection. To quote from Brian C. J. Moore's chapter on the precedence effect [15]:

- "If the interval between the arrival of the two sounds is 1 ms or less, the precedence effect does not operate; some average or compromise location is heard. This is called summing localization."

Thus the precedence effect will be operative for the loudspeaker reflections and place them in the time constant range of the Haas effect, if the loudspeakers are appropriately set up in the room. Reflections would have to be delayed in the order of 50 ms to become audible as separate echoes.

5. SPACES SPEAK

"Spaces speak. Are you listening?" is the title of a book by Barry Blesser and Linda-Ruth Salter [16]. We experience spaces not only by seeing, but also by listening and easily remember the sound of an empty house or a concert hall. We quickly learn the sound of a new space upon entering it and hearing or making noise.

When comparing loudspeakers D and M on the same recording and switching from one to the other, it took some time to perceive a changed spatial impression. The impression developed in parts as different instruments in the recording illuminated their venue space. Thus it took a few seconds to minutes to register a change, depending upon the rate with which informational cues were delivered by the recording. Once the spatial impression was established, it was readily maintained by further cues and only slowly lost. It helped to close the eyes. When visual bias is removed, the listening room and the objects in it are largely de-correlated from experiencing the recording venue space. Seeing the loudspeakers and knowing that the sound is coming from them is inconsistent with the perceived disappearance of those loudspeakers. It is remarkable that the illusion of listening into a foreign space can be generated inside the confines of a familiar living room.

For 2-channel stereophonic sound perception the loudspeakers, room and listener can form a living system, a symbiosis that constantly adapts to the cues in the vibrations that emanate from the loudspeakers.

- "Listening is more than hearing; it is more than sensing, detecting, and discriminating sounds. Listening is the act of making sense out of an aural experience by incorporating all that has been remembered from previous experiences." [16]

The perception of source location is generated by timing and level differences between the signals at the two ears. Two loudspeakers can recreate source locations in the horizontal plane within some limits of accuracy. The perception of distance of a source within a space, however, depends largely upon the degree to which reverberation has changed the source signal's onset, decay and envelope at the receiving location [17]. The two ear signals do not need to be different from each other to hear distance and space. With an appropriate microphone technique and in a suitable recording venue, space information is automatically embedded in the two channels. Even a single channel contains sufficient cues to create the impression of space, as observed when listening to old monaural recordings over a single loudspeaker. The precedence effect, in conjunction with appropriate loudspeakers and their setup, helps to minimize room effects, which then allows optimum recognition of spatial cues.

A simple experiment may serve to illustrate the perception of space.

- Listen to a CD through the 2-channel loudspeaker setup in your room.
- Record the CD playback with small omni-directional microphones on the sides of your head near the pinnae.
- Play back the recording over the two loudspeakers and compare it to the initial CD reproduction.

Note that your room's contribution to the overall sound has become clearly audible because the room response is now imbedded in the direct signal from the loudspeakers. The reflections from your room are again fused with the direct signal and it dominates perception. Thus you hear a recording of your loudspeakers' sound in your room. The experience is similar to how you would hear a recording of a person speaking in your room, which is not how you actually hear the person or the loudspeakers in the live situation where their sound is fused with the familiar room response.

It should also be noted in this experiment that when you listen to a CD with well-defined phantom images and sound stage between the two loudspeakers and then turn your head slowly clockwise, the phantom images and sound stage stay firmly centered between the loudspeakers. Even the sound timbre remains unchanged, just as with real sources and space. The frequency responses for left and right ear, though, change drastically with the rotation, as do the microphone signals. This is an example of the processing power of the ear-brain cognitive system, which compensates for the head movement. Playback of the in-room recording, however, shifts the phantom images to the left loudspeaker and collapses the soundstage.

The recording technique used here is not binaural, where the microphones would have been placed at the ear canal entrance. The microphone response then includes the effects of the pinna. If such recording is played back over loudspeakers, the sound passes the pinna a second time and this is heard as a coloration. With the microphones outside the pinna the head introduces amplitude, phase and delay differences to the left and right microphone signals that have some resemblance to normal hearing processes. It had been found on many occasions that recordings could be made using this head-related microphone technique, which sounded realistic over two loudspeakers. Similarly, sphere microphones are sometimes used.

A buyer of new loudspeakers often wants to hear them in his own room. Differences between the loudspeaker's performance in the show room and the customer's room are usually blamed on the room, not on the loudspeakers. Only if a loudspeaker illuminates a room evenly at all frequencies and is set up appropriately, will it sound similar in different rooms.

In the evolution of natural hearing forest and savanna were the spaces for survival. To detect the proximity and direction of sound sources was essential for recognizing the largest threat. Forests are highly reverberant, the savanna is acoustically dead by comparison. Hearing evolved by learning how these environments change sounds. To recover the direct sound from a mixture of sounds, and therefore to know the true direction and distance of a potential thread, helped to survive. Savanna and forest provided an acoustic background. It could be ignored as long as there was no change from familiar sound patterns. A listening room is the modern equivalent to forest and savanna. We still use the now hardwired portions of the hearing process but adapt them to the new situation. We still can ignore the static background, in this case the room and the fixed loudspeakers, and automatically focus our attention on the direct sound, even when it creates an illusion.

6. SUMMARY

The investigation has shown that a pair of omni-directional and a pair of dipolar loudspeakers will sound essentially identical in a reflective listening room, if their free-field measured on-axis frequency response is the same. The different room reflection patterns due to the different polar responses become perceptually fused with the direct sound of the loudspeakers, if the reflections are sufficiently delayed and if their spectral content is coherent with the direct sound. To minimize skewing or destroying the position of phantom images a left to right symmetry of the loudspeaker and listener triangle with regards to the reflecting room surfaces is necessary. This is particularly important for omni-directional loudspeakers.

It is very difficult to predict the subtleties of what is perceived by a listener from analyzing in-room steady-state frequency response or reflection measurements. The response graphs are difficult to interpret and can be misleading. This has implications about the efficacy of electronic room response equalization based on such measurements for more than the lowest frequencies. Aspects of the precedence effect must be considered when trying to use in-room measurement data for realistic sound improvements.

The room response observations came as a surprise because loudspeaker M was built for an investigation of baffle edge diffraction and non-linear distortion in comparison to D. The ensuing conclusions about polar response and loudspeaker setup must be added to previously published loudspeaker requirements [18]. Together they appear to complete the set of necessary parameters for accurate 2-channel sound reproduction.

Hopefully this work stimulates further investigation of the loudspeaker-room-listener interplay, especially since it points to the limitations of common practices in loudspeaker design and room setup. It could also lead to recording techniques that consistently capture the instruments and the spatial context in which they are being used. Monitoring during the recording process could become more accurate and would match how the audience will hear the recording under optimal conditions.

Two-channel playback in a normal living space can provide an experience that is fully satisfying as loudspeakers and room disappear and the illusion of being transported to a different place and moment in time takes over.

7. REFERENCES

[1] Siegfried Linkwitz, "Development of a Compact Dipole Loudspeaker", presented at the AES 93rd Convention, San Francisco, USA, 1992 October 1-4, (Preprint 3431)

[2] http://www.linkwitzlab.com/Pluto/intro.htm

[3] http://www.linkwitzlab.com/orion_challenge.htm

[4] F. E. Toole, "Loudspeaker Measurements and Their Relationship to Listener Preferences", J. Audio Eng. Soc., vol. 34, part 1, pp. 227-235 (1986 April); part 2, pp. 323-348 (1986 May)

[5] Siegfried Linkwitz, "Shaped Toneburst Testing", J. Audio Eng. Soc., vol. 28, pp. 250-258 (1980 April)

[6] Numerous private conversations and exchanges of listening observations with Don Barringer, a professional musician and retired recording engineer for the US Marine Band, Arlington, VA.

[7] Siegfried Linkwitz, "Investigation of Sound Quality Differences between Monopolar and Dipolar Woofers in Small Rooms", presented at the AES 105th Convention, San Francisco, USA, 1998 September 26-29, (Preprint 4786)

[8] Private reports by builders of loudspeakers D and M about listening observations in their home environment, USA. Comparison of experiments with Brian Elliott, retired electronics engineer, acoustician and loudspeaker designer, Palo Alto, CA.

[9] S. E. Olive and F. E. Toole, "The Detection of Reflections in Typical Rooms", J. Audio Eng. Soc., vol. 37, pp. 539-553 (1989 July/August)

[10] Eric Benjamin, "An Experimental Verification of Localization in Two-Channel Stereo", presented at the AES 121st Convention, San Francisco, USA, 2006 October 5-8, (Preprint 6968)

[11] Floyd E. Toole, "Loudspeakers and Rooms for Sound Reproduction – A Scientific Review", J. Audio Eng. Soc., vol. 54, pp. 451-476 (2006 June)

[12] William M. Hartmann, "Listening in a Room and the Precedence Effect", Chapter 10 in "Binaural and Spatial Hearing in Real and Virtual Environments" R. Gilkey and T. Anderson (Eds.), Lawrence Erlbaum Associates, Hillsdale, NJ, 1997

[13] D. M. Green, "An Introduction to Hearing", Lawrence Erlbaum Associates, Hillsdale, NJ, 1976

[14] William A. Yost, "The Cocktail Party Problem: Forty Years Later", Chapter 17 in "Binaural and Spatial Hearing in Real and Virtual Environments" R. Gilkey and T. Anderson (Eds.), Lawrence Erlbaum Associates, Hillsdale, NJ, 1997

[15] Brian C. J. Moore, "Psychology of Hearing", Academic Press, 1997

[16] Barry Blesser, Linda-Ruth Salter, "Spaces Speak, Are You Listening? – Experiencing Aural Architecture", MIT Press, 2007

[17] David Griesinger, "Pitch Coherence as a Measure of Apparent Distance in Performance Spaces and Muddiness in Sound Recordings", Institute of Acoustics Conference, Copenhagen (2006 May), http://world.std.com/~griesngr/

[18] Siegfried Linkwitz, "Which loudspeaker parameters are important to create the illusion of a live performance in the living room?", presented at the AES 113th Convention, Los Angeles, USA, 2002 October 5-8, (Preprint 5637)