THE LOCALIZATION OF PHANTOM IMAGES IN AN OMNIDIRECTIONAL STEREOPHONIC LOUDSPEAKER SYSTEM

2371 (B-15)

David Moulton
Michael Ferralli
Phase Coherent Audio, Inc.
Girard, Pennsylvania
Steven Hebrock
Mark Pezzo
Audio-Technica, U.S., Inc.
Stow, Ohio

**Presented at
the 81st Convention
1986 November 12-16
Los Angeles, California**

AN AUDIO ENGINEERING SOCIETY PREPRINT

THE LOCALIZATION OF PHANTOM IMAGES IN AN OMNIDIRECTIONAL STEREOPHONIC LOUDSPEAKER SYSTEM

David Moulton
Micheal Ferralli
Phase Coherent Audio, Inc.
Girard, PA, USA

Steven Hebrock
Mark Pezzo
Audio-Technica, U.S., Inc.
Stow, OH U.S.A.

While testing a 360 degree dispersion stereophonic loudspeaker system, anomalies in terms of "depth" and "size" in the perception of phantom images were noticed. This has led to an investigation of phantom images in both stereophonic and monophonic music reproduction. Differences in the perception of high and low frequencies are considered, and a hypothesis regarding the relationship between the perception of reverberant fields of the source room and playback rooms is presented.

INTRODUCTION

In December, 1985, two of the authors completed construction of prototype high-frequency acoustic lens systems with 360° horizontal dispersion and 90° vertical dispersion. When these prototypes were integrated into a full-range system and auditioned, characteristics of the stereophonic sound field were noticed that were unique in the authors' experience. This has lead the authors to the investigation that is the basis of this paper. A specific physical and theoretical description of the lens system is described in another paper by two of the authors (Ferralli and Moulton)[1] and will not be repeated here.

Our physical experiments have been directed toward evaluation and understanding of the behavior of the prototypes, with an eye toward product development. As a result, our investigation of localization in regard to our omnidirectional prototypes has been generally straightforward and pragmatic. We have sought general confirmation by both experienced and inexperienced listeners that the effects that we have perceived are in fact generally perceivable and predictable. We have

further sought to determine the impact of environment upon this localization behavior in the hope of increasing our ability to successfully predict the character of that behavior. We have done this by extended listening to music and special research recordings in a variety of reverberant listening rooms, including:

- a comparatively live residential living room (15'x20'x9');
- a residential study, much less live, (15'x14'x9'), with extensive carpeting, floor-to-ceiling drapes, bookcases, diffuse brick hearth structure;
- a large recording studio (40'x50'x25') with an RT_{60} of 1.8 sec. @ 500 Hz.;
- a studio control room (19'x12'x10') with significant lateral absorbtion, but a large window behind the speakers and a diffuse wall (30% quadratic residue diffusor) behind the listeners. In this case, switching between the house monitor system (triamplified and equalized) and a popular brand of small near-field monitors placed on the same lateral plane as the prototypes was possible, and the prototype system was used as part of the overall monitor system for remix of a multitrack rock music project, in addition to other evaluations.

Speaker placement varied as a function of experimental curiosity and production needs. Generally, placement was away from room boundaries and was approximately symmetrical in relation to room boundaries. Speaker placement close to room boundaries was tried in several instances. Further, in the large recording studio, the room boundaries were manipulated by asymmetrical placement of an 8'x8' polycylindrical diffusor.

In addition, extended listening was done in a medium-sized anechoic chamber (24'x20'x16'), where auditions of music and the research tape (see below) were made and compared to the reference system using conventional tweeters. Swept sine-wave amplitude response tests and polar response tests were also conducted in this facility.

Listeners include the authors, numerous broadcast and recording engineers from around the country, audio equipment manufacturers sales and engineering staff personnel, numerous musicians with and without audio training, numerous students involved in the Sound Recording Technology program that one of the authors teaches, and a small number of lay individuals and children. Not less than 50 individuals have auditioned these prototypes, in the various rooms described above.

Listening trials have been conducted on 23 separate days, evenly spread between December, 1985 and the time of writing. System configuration has varied quite widely, but the lens elements and their transducers have been essentially unmodified except for periodic realignment. Approximately 600 person/hours of auditioning have occurred

TEST EQUIPMENT AND METHODS

The test prototypes are high-frequency elements wired in place of the conventional tweeters in a pair of conventional small 3-way studio monitors, utilizing the stock crossover. The prototype elements were placed on top of the studio monitor cabinets. Wiring was exposed so that it was possible to easily reconnect the stock tweeters for comparison purposes. Differences in response have been compensated for in some cases by corrective equalization, in others by simple crossover adjustment.

Occasionally, different low- and mid-frequency systems were used, including several different consumer-oriented systems, sometimes using the stock cross-over, other times using a simple proprietary crossover network. In one series of auditions, the prototype high-frequency elements were biamplified in combination with mid-frequency elements and a central woofer element.

In all of these configurations, the characteristics of behavior that we originally noticed remained, and we believe that they arise principally as a function of the dispersion of the prototype elements.

The system has been auditioned in two primary configurations: (a) with low-frequency and mid-frequency drivers facing the listener, and (b) with low-frequency and mid-frequency drivers facing away from the listener (the latter was never done in the control room setting).

A research tape (PCM 16-bit recording) was used for examining features of localization: the recording includes a set of amplitude-differential (ΔL) transient impulses were recorded to examine the monophonic characteristics of the phantom image localization, a stereophonic recording of an impulse and of pink noise sounding in a medium sized (50,000 cubic ft.) reverberant room (the recording studio referred to above), and a time-differential recording (ΔT) of impulses and pink noise.

In the monophonic amplitude-differential recording, the left channel remains constant in amplitude, while the right channel is decreased in amplitude. Two sets of test data were recorded: (a) amplitude decreasing in increments of .2 db to a maximum difference of -2.4 db, and (b) amplitude decreasing in increments of 1 db to a maximum difference of -11 db.

The stereophonic recording is of a single conventional small 3-way studio monitor in a large reverberant room (the recording studio listed above), situated on a stand 4' from the floor and 10' from the wall behind the speaker. The stereophonic recording was made with coincident xy bidirectional mics (a "Blumlein pair") placed 4' from the speaker. The speaker was moved laterally to the left in 6" increments to a maximum lateral offset of 4' (45 degrees). Two test signals were used: pink noise and an impulse as described above.

The recording of ΔT impulses and Pink Noise stimuli included delays ranging from .25 ms ΔT to 15 ms. ΔT. Also, sections of Pink Noise with the earlier or later channel intermittently muted was included to examine the comparative masking-by-precedence of the various systems.

Playback of the research recording has been done over a wide variety of speakers (in numerous environments) in addition to the test prototypes, and for a fairly wide variety of listeners. The general effect of the research tape played over conventional loudspeakers in a variety of spaces is familiar to the authors and quite predictable. It is also within conformance with what is known about the effect of ΔL and ΔT upon localization with two loudspeakers. The music auditioned has included both a general range of commercially available and successful pop and classical recordings (on CD) and several rock and classical recordings by one of the authors, wherein the recording and remix techniques used were known precisely.

Trials have been informal and anecdotal. No formal testing of subjects has been done, and no attempt has been made to isolate listeners from either other listeners or from the authors (and their predispositions in regard to the performance of these prototypes). Instead, trials have sought to elicit comments regarding (a) the perception of depth when listening to the prototypes in comparison to their previous experience with listening to loudspeakers, (b) the perception of a stereo soundfield when listening to the prototypes and how this perception differed from their experience with conventional loudspeakers, (c) their subjective assessment(s) of the prototype system, including what they felt its strengths and weaknesses were, again in comparison with their experience with conventional loudspeakers.

The listeners' access has been as flexible and accommodating as time and facilities would permit. This has included A/B trials, equalization, moving speakers to different positions in the room, changing program material, reversing channels, reversing system and/or element polarity, listening to monophonic program material, answering all questions that we could about the systems, and so forth. In short, our trials have been directly related to product evaluation and development, and formal psychoacoustic experimental methodology has not yet been implemented.

SUMMARY OF PROTOTYPE AUDITION RESULTS

In a single-speaker monophonic configuration, playing in a reverberant room, the prototype system sounds reasonably similar to conventional speakers, except that high-frequency spectra does not change with axial position. "Glare" reflections and apparent changes to direct sound at the listener, as reported by Lipshitz and Vanderkooy[2], have been noted. No attempt has been made to resolve the flat-power-response vs. flat-axial-response problem. Subjectively, the prototype system sounded pleasant and smooth.

In "conventional" monophony (2 speakers, monophonic program), the phantom image appears to be behind the lateral plane running through the two speakers when panned to Center and at each speaker when panned Left or Right (thus, panning from left to right results in a somewhat vee-shaped motion of localization of the phantom image). The strength of this effect varies with the nature of the room and speaker placement. A less reverberant room will yield a less pronounced effect. The apparent "depth-of-field" appears also to be a function of the distance from the speakers from the rear wall.

In "true" (stereo (2 speakers, stereophonic program, including reverberant information), the image(s) appear(s) to emanate from a field behind the lateral plane. The size and depth of this field appear to be dependant upon:

A. The closeness of the speakers to the rear wall. As the speakers are moved away from the wall the image field appears to deepen. It often appears to extend beyond the wall.

B. The amplitude of the program material. As amplitude is increased, the width and depth of the image field increase. Curiously, low amplitude levels have resulted in an impression of a group of miniature, doll-like performers, particularly in a small room. High amplitude levels and/or a large reverberant room result in a "larger-than-life" impression. This effect is less reliable and predictable than A. or C.

C. The closeness of the listener to the lateral plane running through the speakers. The front of the image field moves closer to the listener as he/she approaches the plane, and recedes as he/she moves back from the plane, as if the listener was walking toward a mirror and the image was his/her reflection.

Several other phenomena need to be mentioned:

A. The stereophonic localizations within the image field remain fairly constant as the listener moves laterally, and the relationships between localized sources shift approximately as they would if the actual sources were present (this appears to hold true, roughly speaking, even beyond the point where the listener has moved "outside" of the listening area between the speakers, so that the listener will clearly perceive stereophony, with phantom localization even when standing at a side wall, close to the lateral plane);

B. As the listener approaches the lateral plane between the speakers, he/she perceives that the phantom image(s) are moving toward the lateral plane as well as noted above. As the listener passes between the speakers, he/she has the sensation of "walking into" the image of the performers. As he/she passes beyond the lateral plane running through the speakers, the image field shifts to the far side of the plane and the listener now hears the sound coming from behind him/her, on the other side of the room. With material that is primarily "multitrack mono," as the listener approaches the lateral plane the monaural center images come to be perceived as "within the head" (as per Plenge[3]), roughly similar in effect to headphones;

C. Program material that has both stereophonic and monophonic components (i.e. most recordings) tends to exhibit different behavior for the stereophonic and monophonic components. Monophonic components appear as described in the first paragraph, above, so that material panned Left or Right appears to come from the loudspeakers and material panned to Center appears to come from behind the plane running through the loudspeakers, while stereophonic components appear to emanate from the image field, well behind the plane running through the speakers. This leads to curiosities: a stereophonic orchestral recording with outrigger highlight mics used for the string sections tends to lead to an ambiguous image for the strings, so that they appear both behind the plane and also at the speakers (this can vary as a function of gain-riding in the mix, which is often more obvious on these prototypes than on the stock units); a rock recording with comparatively dry rhythm guitars panned L and R and wet vocals, drums and horns will sound as if the performance is taking place in a semicircle with the ends of the semicircle at the speakers.

D. When the polycylindrical diffusor was put into place for trials in the recording studio (see Fig. 1), the monophonic Center image shifted clearly toward the diffusor, becoming offset in terms of both lateral and depth positions. Stereophonic images shifted similarly, with a less pronounced lateral off-set but a clear skewing of the depth of field, resulting in the perceived image of an ensemble that was closer to the listener on the side of the room where the diffusor was placed.

Interestingly, some of the most compelling music auditions were obtained with the speaker low-frequency and mid-frequency elements facing away from the listener. Localization remained excellent, dependent on reverberant surfaces, and appeared to be in an image field considerably behind the speakers and extending through the wall behind the speakers. It is our belief that localization behavior similar to the high-frequency prototypes will be noted with omnidirectional mid-frequency elements (currently under construction).

THEORY OF PHANTOM IMAGE BEHAVIOR

The electrically generated audio signal presented in common to two identical loudspeakers operating near each other in a reverberant space and in a phase-coherent and -locked manner, is a unique source unlike any other sound in nature. Such a sound may be characterized as having two actual sources with the normal psychoacoustic signature of a single source (i.e. the signal is phase-locked, has common spectra and formant, and identical or similar amplitudes). The precedence effect coupled with the inverse square law dictates that the sound shall be localized at the nearer speaker and that the farther speaker shall be imperceptible. In the case where the observer is equidistant from both loudspeakers, a "phantom image" is noted at the median plane, directly between the two speakers, and neither speaker itself is perceived, in conjunction with the paradoxical condition that there are two first arrivals of direct emission artifacts instead of the normal single first arrival of the direct sound.

It should be noted that monophonic signals, with ΔL derived from panning, are the primary sources for current multitrack recording. Comparatively little music today is recorded in "true" stereo. Stereo effect in "multitrack mono" recordings is generated by the addition of time-differential effects such as artificial reverb (which in its current manifestations generates quite effective simulations of acoustical reverberant spaces as well as the apparent positions of both the source[s] and the listener), interspeaker time delay of non-reverberant signals and simple amplitude-differential panning of non-reverberant signals. Often, "true stereo" recordings (such as drum overhead mic pairs) are mixed in as well.

"True stereo" recordings (i.e. "classical" music recordings) often employ monophonic "highlight" recordings (via multitrack recording techniques) to ease balance difficulties. Comparatively few released recordings today employ only two (or three) microphones placed in a single stereophonic configuration. Virtually all contemporary recordings, then, have both monophonic and stereophonic components, coexisting in reasonably convincing and pleasing fashion, and often enhancing each other substantially.

LOCALIZATION OF REAL SOURCES

It is axiomatic that we localize actual acoustic signals by interaural amplitude- and time-differences. Benade notes that this is true in free-field conditions and that in reverberant spaces the issue is more complex.[4] He points out that localization takes place in each ear as well as interaurally and that such localization may be more predominant than the traditionally noted free field case:

> "It is easy to show that listeners can localize a person or a musical instrument with good precision. It is interesting to find that they can do this much more confidently in a room than they can outdoors, and a little deliberate practice can improve their skill and their self-confidence a great deal. A person localizing a sound in a room processes temporal cues from the early reflections from the walls of the room; the pattern of the early reflections is a direct geometrical consequence of the position of both source and listener. An easily verified indication that this is actually going on is the following: If a listener is asked either to partly cover one ear with his hand or to plug it up completely, he will find that his localizing ability in a room is almost unimpaired, whereas in free space the same action will sharply reduce his ability to localize. This shows that in a room the signals can be processed on the basis of their temporal arrival patterns at each of the two ears separately, even when the detailed interaural amplitude and phase relationships are absent or totally disrupted."[5]

Benade goes on to describe a working hypothesis for the precedence effect with six rules, two of which are of significance here:

> "1. The human auditory system combines the information contained in a set of reduplicated sound sequences and hears them as though they were a single entity, provided . . . that most of them arrive within a time interval of about 40 ms following the arrival of the first member of the set. . . .
>
> "4. The apparent position of the source of the perceived composite sound coincides with the position of the source of the first-arriving member of the set, regardless of the the physical

> directions from which the later arrivals may be coming. . . .'
>
> "It is important to notice that these very strongly worded categorical statements all emphasize that there is an *accumulation of information* from the various members of the sequence. It is quite incorrect to assume that the precedence effect is some sort of masking phenomenon which, by blocking out the later arrivals of the signal, prevents the auditory system from becoming confused. Quite to the contrary, those arrivals that come within a reasonable time after the the first one actively contribute to our knowledge of the source."[6]

Therefore, according to Benade, the commonly accepted criteria for determining localization (ΔT and ΔL) are primarily valid in anechoic conditions.[7] At the same time, Roederer asserts that the time scale for neural processing at the basilar membrane is less than 50 ms.,[8] which coincides with the time base within which most primary early reflections occur (i.e. within the range of times normally considered to be subject to the precedence effect). This, in conjunction with Benade's thesis, implies that substantial neural processing leading to monaural localization at each ear separately is within the integration time range of the precedence effect. Roederer also indicates that neural pulse trains triggered by low-frequency signals contain detailed "state-of-phase" information.[9] Griesinger asserts that low-frequency difference information in stereophonic recordings is a primary "carrier" of spatial sensation and perception of auditory environment,[10] while Hartman notes that such lateral reflections also tend to "delocalize" real sources in the space[11] and Schroeder's studies regarding subjective preferences for concert halls indicate that listeners find lateral difference information (presumably mostly low-frequency) a desirable attribute.[12][13] These findings, taken together, indicate that the precedence effect, as a localization mechanism, may operate independently of low-frequency interaural analysis.

A tenable hypothesis regarding the principles of localization of real acoustic sources in a free space is that given that long wavelength phase detection can clearly be detected interaurally and that short wavelength phase data is not detected at the basilar membrane, quasi-statistical summation of early reflections and direct sound (with little or no information regarding time offset between direct and reflected artifacts) is the basis of neural information about the nature of the sound source sent from ear to brain. In a non-reverberant space, only direct information is available and much of the auditory information upon which this summation depends is not present. Therefore, for non-reverberant conditions interaural comparison of low-frequency phase components and overall amplitude may be the predominant or the only data available.

In a reverberant space, a richer data source is available: each ear takes in the direct sound and early reflections from all walls (what we shall refer to as the "7-Path Model" ["5-Path Model" in the illustrations where floor and ceiling have been deleted for graphical simplicity, see Fig. 2); it is reasonable to hypothesize that these are identified as a set by their phase-locked character and that (a) vector information regarding the set is derived from pinna cues at each ear and interaurally [14][15][16][17][18] and that (b) this set is analyzed by a complex algorithm arising from the patterns of excitation and inhibition of the inner and outer rows of cilia on the basilar membrane, and that the resulting construct is "sharpened" by neural processing in the auditory nerve en route to the brain, possibly including neural feedback.[19][20][21][22][23] Bloom has shown that high-frequency localization takes place at each ear as a function of comb filtering caused by pinna reflections that yield a unique spectral signature to a given direction of arrival of sound at the ear.[24] If the reverberant space is "known" (i.e. the dimensions and character of the space have already been assimilated by prior stimuli), then a unique source location can be derived from any real set of early reflections.

From all this, a strong case exists for the notion that localization of real sources in a reverberant space is derived from a multiplicity of data and neural templates. A single-mechanism localization theory does not appear to be consistent with known facts. We hypothesize further that "knowledge of space" is derived primarily from interaural low-frequency analysis and "knowledge of source" is derived primarily from monaural high-frequency analysis of direct and early reflection versions of the source. This is related to the "duplex theory" (Stevens and Newman [1936]) as cited by Yost[25].

LOCALIZATION OF PHANTOM IMAGES

For the purposes of this paper, we will categorize phantom image signals as monophonic (meaning that the signal emitted from both speakers is identical) or stereophonic (meaning that the signal emitted from both speakers contains amplitude- and time-differential information regarding the recorded source and may also carry reverberant cues derived from the recording room as well).

The monophonic phantom image emitted by conventional loudspeakers is perceived as originating from a point in space directly between those loudspeakers. When this effect is tested in an anechoic chamber, the phantom image proves to be ambiguous in terms of depth, although the lateral discrimination is quite precise[26]. (Plenge notes contrasting sensations of "sound-in-head" and "sound-out-of-head" which he attributes to richness of detail of source stimulus.[27] We believe that this is related to the anechoic sensation we experienced.) In a reverberant space and with conventional loudspeakers, the image appears on the lateral plane so reliably that it has become axiomatic, and sensations of depth have been attributed (in advertising) to extreme precision of phase response in the loudspeaker and between matched loudspeakers and special attributes of particular stereophonic microphone configurations.

A comparison of the 7-Path Models (see Figs. 3 and 4) for a symmetrically placed real source and a centered monophonic phantom image generated by directional speakers yields the insight that the vector data gleaned from the two speakers is incompatible with that of the real source. Common sense suggests that the two first arrivals (i.e. direct) from the speakers are fused into a unique vector based on ΔL as the determinant for the perceived angle of arrival of the phantom source. In fact, the ear has no reason to infer the existence of multiple sources of sound. Localization of the phantom is more likely based on a "best-fit" interpretation of the neural templates for localization, in which the actual arrivals of artifacts (amplitudes, times and directions) infer a localization of source whose direct artifact was not detected.

In the case of conventional directional speakers, no high-frequency artifacts are perceived from the front of the room. Only lateral and rear artifacts are present. The ΔL of the left and right artifacts are essentially at unity, leading the auditory system to infer that localization "must be" along the median plane. Further, time and amplitude relationships between the lateral artifacts imply the existence of a source at or slightly behind the lateral plane through the speakers.

The 7-Path Model derived from a monophonic signal (Fig. 5) from omnidirectional speakers includes artifacts reflected from the front wall. The angle of arrival of these artifacts are quite close to the hypothetical angle of arrival of a phantom image on the median plane. We hypothesize that it is these front-wall reflection artifacts that are a primary element in inducing the perception of depth in the monophonic phantom image. Experimental observations support this notion. Depth is least pronounced, most ambiguous with conventional speakers. Omnidirectional tweeters induce a sense of depth, particularly at the median plane. When mid-frequency and low-frequency elements are faced away from the listener, sense of depth becomes much more pronounced, and even Left and Right ΔL (up to -11 dB) and ΔT (up to the point of loss of fusion, ca. 10 msec.) are perceived as originating from well behind the speakers. This progression correlates well with relative amplitude of front-wall reflections. In the anechoic chamber, we expected to find no difference between the conventional and omnidirectional systems. However, the omnidirectional system yielded an increased sense of depth that was easily noted by all listeners, even if the effect was less pronounced than in a reverberant space. We hypothesize that this is due to imperfections (i.e. the presence of reflections, however slight) in the anechoic chamber. This remains to be tested.

The stereophonic case is more complicated, of course, and at the same time informative. If we use two omnidirectional microphones in a reverberant space, each microphone will detect its own unique 7-Path Model for the source and room. Angle-of-arrival information is not detected. When such data is emitted from loudspeakers, vector information *about the recording room* may only be derived from ΔT and ΔL of the *two* loudspeakers, essentially by interaural comparison. For this to happen, we need (in theory) anechoic conditions, and in fact LEDE®[28] and other[29] control room design principles are based on the notion that the 7-Path Model of the playback room tends to obscure the ΔT and ΔL of the recording room and that such paths should be suppressed until the ΔT/ΔL of the recording room have been assimilated by the ear(s). However, the overwhelming majority of listening to music via loudspeakers is done in reverberant spaces and the LEDE® environment is at present a "special-case" recording production tool.

In the case of the omnidirectional speakers, each speaker emits common temporal data (the 7-Path temporal signature detected by the microphone) in all directions, so that the ear receives 7 versions (Fig. 6) of the same set of temporal arrival patterns carried on the reflective as well as direct paths of the playback room's 7-Path Model. Thus, we have a condition where ΔT/ΔL information from the recording room (superimposed on the vector information of the playback room) is carried on the two 7-Path Models from speakers to listener (ca. 100 early reflection versions of the ΔT/ΔL data acquired at each microphone, each version with the same temporal pattern). It is our hypothesis that this results in multiple neural localization templates that the auditory mechanism resolves, as best it can and with considerable effectiveness, into a single perceptual construct[30] and that this perceptual construct is representative of neither the recording space or the playback space, but rather a fusion of the two. By direct observation, this construct appears to be a space (the recording room) within or around (depending upon amplitude) a space (the playback room), based upon auditory assimilation of the ΔT of both spaces

At present, we hypothesize that the use of omnidirectional speakers away from the boundaries of the playback room (placed, in fact, in a manner similar to the way live musicians might be placed in the same space) represents the opposite side of the coin of LEDE® theory and is similarly valid: where LEDE® theory attempts to suppress he influence of the playback space in order to gain access to the $\Delta L/\Delta T$ information from the recording room (by placing speakers at boundaries, rendering surfaces either non-reflective or as highly diffuse as possible), omnidirectional/reverberant theory would accomplish the same thing by transmitting as many versions as possible of the $\Delta L/\Delta T$ information from the recording room, via all available playback room paths. There is clear evidence that the auditory system can and does make use of this multiplicity of data.

CONCLUSION.

At this time, our test results remain preliminary. Although all listeners (perhaps 50 in all) have confirmed the sensations that we have noticed, strictly controlled testing needs to be done. The strength of the effect(s) that we have noticed is of sufficient magnitude that we feel justified in predicting that our tests and results will prove to be replicable

In our examination of the literature regarding localization studies, we have found no reference to concern about loudspeaker directionality in experimental methodology. Toole notes, however, that different loudspeaker directivities may render other comparative data invalid.[31] While our omnidirectional elements are clearly, at this time, a "special case" in the world of loudspeakers, we believe that they may reveal that conventional (directional) loudspeakers as test sources may present a "special case" insofar as the human auditory localization mechanism is concerned

Finally, on a more subjective note, both Benade[32] and Plenge[33] note a human preference for a rich diet of early reflections in their musical perceptions (via speakers, headphones or in live performance), and the plethora of time-domain signal-processing devices reaching the production-studio effects racks these days speaks eloquently if indirectly in support of this preference. Our own observations of this prototype system support this notion: we find ourselves drawn to prefer listening to music via the prototypes to conventional speakers and we find we have developed a profound affection for the musical nature of the sounds that we perceive in the stereophonic omnidirectional playback system: the *musical* essence of the sound seems more palpable, more enduring and more directly accessible than we have experienced with other loudspeaker systems.

ACKNOWLEDGEMENTS

The authors wish to thank Audio-Technica, U.S., of Stow, Ohio for the use of their research facilities and anechoic chamber, and The State University of New York, College at Fredonia, Fredonia, NY for the use of their recording studio facilities.

REFERENCES:

1 Ferralli, Michael W. & Moulton, David, 360 Degree Frequency Invariant Acoustic Transduction System, Audio Engineering Society 81st Conv., L.A., Nov. 1986.

2 Lipshitz, Stanley P. & Vanderkooy, John, Experiments in Direct/Reverberant Ratio Modification, Audio Engineering Society 79th Conv., NYC, Oct. 1985.

3 Plenge, G., On the differences between localization and lateralization., Journal of the Acoustical Society of America 56:3, 1974, p. 944-951.

4 Benade, A.H. From Instrument to Ear in a Room: Direct or via Recording, Journal of the Audio Engineering Society 33:4, April, 1985, pp. 218-233.

5 Benade, A.H., p. 223.

6 Benade, A.H., pp. 224-5.

7 Benade, A.H., p. 222.

8 Roederer, Juan G., Introduction to the Physics and Psychophysics of Music, 2nd Ed., Springer Verlag , NYC, 1973, 1979, pp. 5-7.

9 Roederer, Juan G., pp. 48-50.

10 Griesinger, David, Spaciousness and Localization in Listening Rooms and Their Effects on the Recording Technique, Journal of the Audio Engineering Society 34:4, 1986, pp. 255-268.

11 Hartman, W.M., Localization of Sound in Rooms, Journal of the Acoustical Society of America 74:5, Nov. 1983, p. 1388.

12 Schroeder, Gottlob & Siebrasse, Comparative study of European concert halls: correlation of subjective preference with geometric

and acoustic parameters., Journal of the Acoustical Society of America 56:4, 1974, p. 1195 ff.

13 Schroeder, Manfred R., Progress in Architectural Acoustics and Artificial Reverberation: Concert Hall Acoustics and Number Theory, Journal of the Audio Engineering Society 32:4, April, 1984, p. 197.

14 Bloom, P. Jeffrey, Creating Source Elevation Illusions by Spectral Manipulation, Journal of the Audio Engineering Society 25:9, 1977, pp.560-565.

15 Hebrank & Wright, Sound localization on the median plane., Journal of the Acoustical Society of America 56:3, 1974, p. 935.

16 Lambert, Robert M., Dynamic theory of sound source localization, Journal of the Acoustical Society of America 56:1, 1974, p. 165 ff..

17 Searle, Braida, Cuddy & Davis, Binaural Pinna Disparity: another localization cue. Journal of the Acoustical Society of America 57:2, 1975, p.448 ff.

18 Wright, Hebrank & Wilson, Pinna reflections as cues for localization. Journal of the Acoustical Society of America 56:3, 1974, p. 957.

19 Dallos, Peter, The Auditory Periphery, Academic Press, NYC, 1973.

20 Robinson, Charles & Pollack, Irwin, Interaction Between Forward and Backward Masking, a measure of the integrating time of the auditory system., Journal of the Acoustical Society of America 53:5, 1973, pp. 1313-1316.

21 Roederer, Juan G., 44-56.

22 Yost, W.A., Man as Mammal: Psychoacoustics, from "Comparative Studies of Hearing in Vertebrates", ed. Popper & Fay, Springer Verlag, NYC 1980, ©1978, pp. 399-420.

23 Von Bekesy, Georg, Auditory Backward Inhibition in Concert Halls: The Parallelism Between Vision and Hearing Suggests New Features in Room Acoustics, Journal of the Audio Engineering Society 27:10, 1979, pp 780-787

24 Bloom, P Jeffrey, p. 560, 562.

25 Yost, W.A., p 413.

26 Blauert, Jens (tr. Allen), Spatial Hearing, MIT Press, Cambridge, MA, 1974, 1983, pp 203-222

27 Plenge, G., p. 947

28 Davis, Don & Davis Chips, The LEDE® Concept for the Control of Acoustic and Psychoacoustic Parameters in Recording Control Rooms Journal of the Audio Engineering Society 28:9, Sept. 1986, pp 585-595.

29 Jones, Martens, Kendall, Optimizing Control Rooms for Stereo Imagery, Draft for Acoustical Society of America presentation. Evanston, IL, 1985.

30 Roederer, Juan G., pp. 59-60, 175-183

31 Toole, Floyd E., Loudspeaker Measurements and Their Relationship to Listener Preferences: Part I, Journal of the Audio Engineering Society 34:4, April, 1986, p. 230.

32 Benade, A.H., p. 226-227.

33 Plenge, G., p. 944.

Figure 1: Audition in recording studio, observe phantom migration with diffusor.

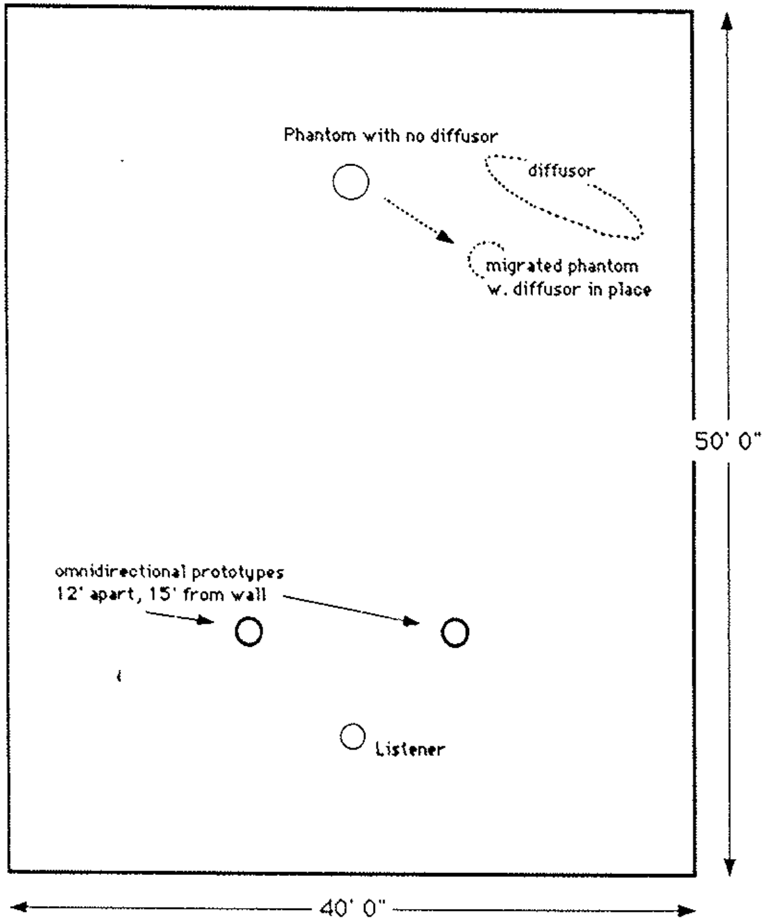


FIGURE 2: (A) 7-PATH REVERBERANT MODEL
(B) 5-PATH MODEL (FLOOR AND CEILING PATHS DELETED)

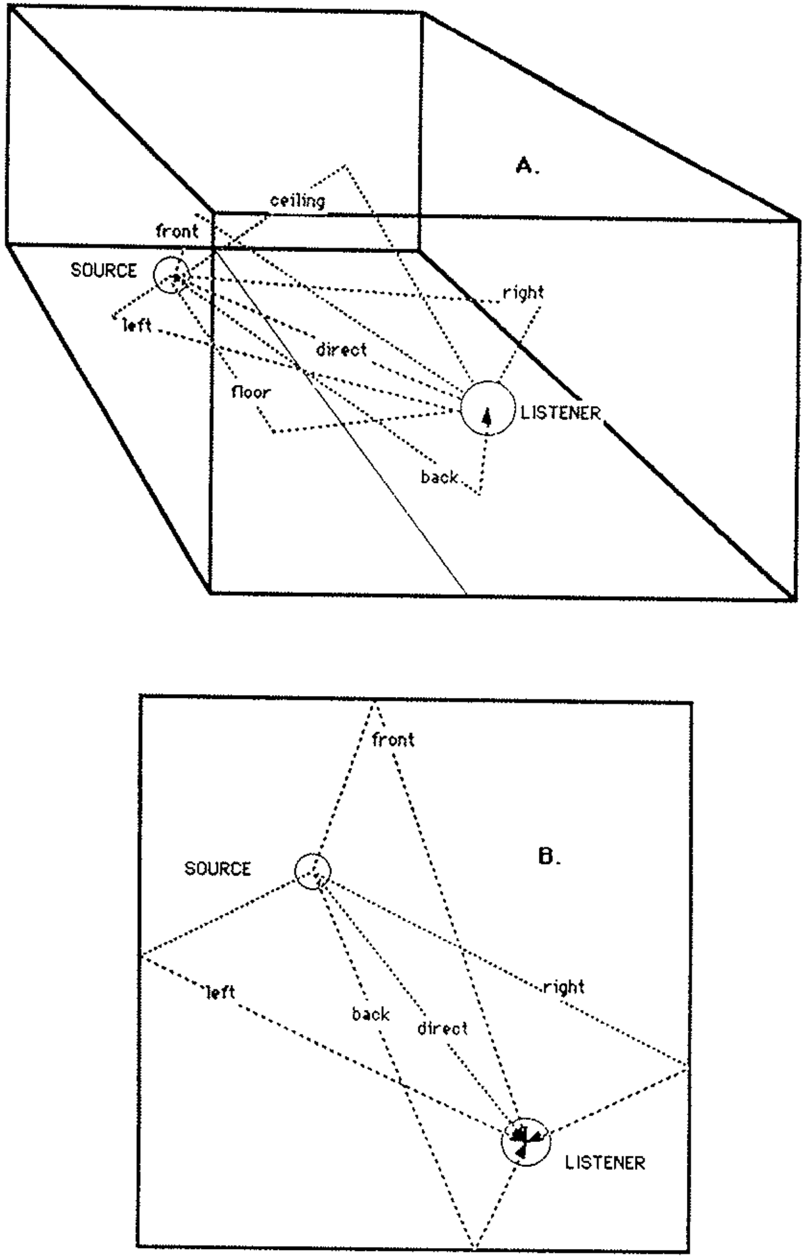
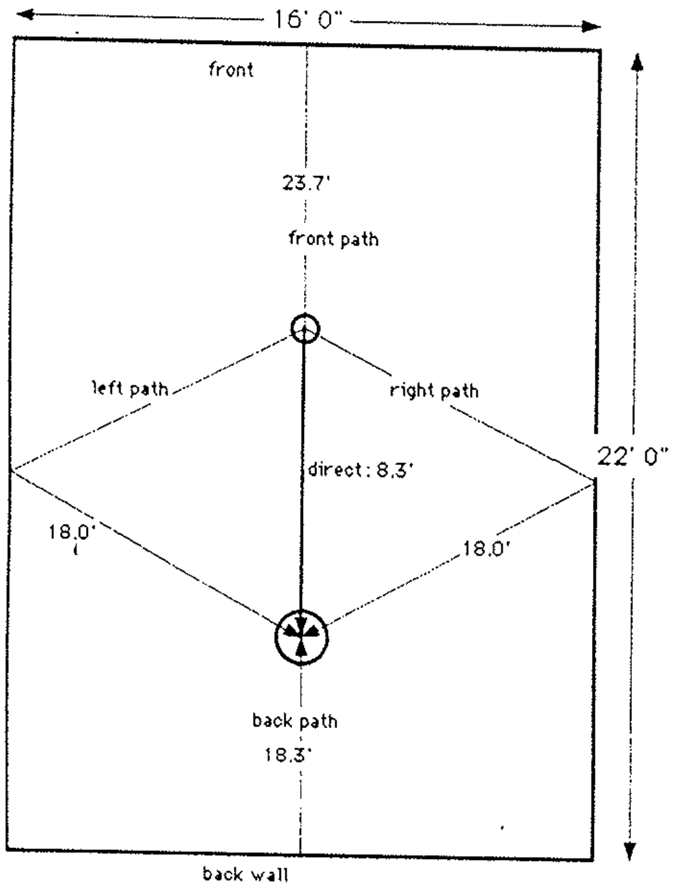


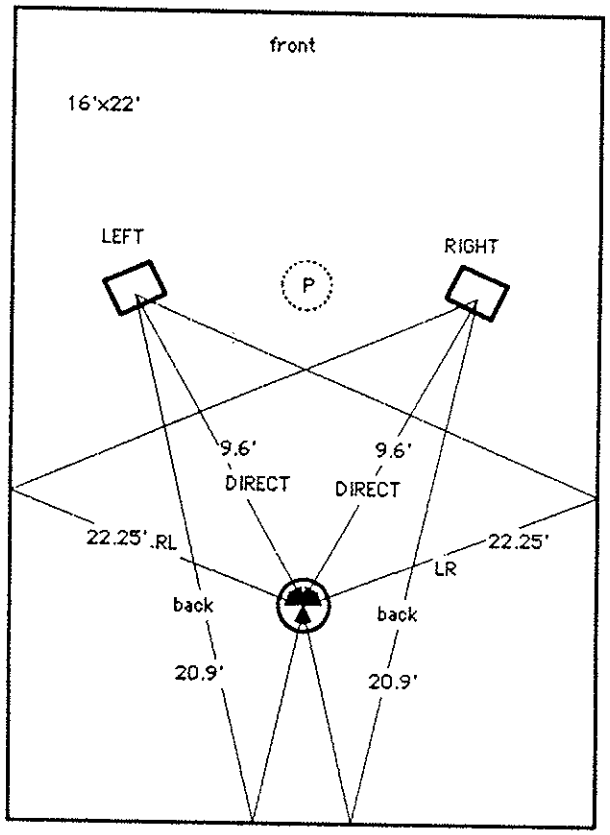
Figure 3: 5-Path Model (floor and ceiling excluded) for symmetrically placed real source in 16x22 room.



MONAURAL REAL SOURCE, SYMMETRICAL.

Artifact	dist.	atten.	angle	Δt
Direct	8.3'	0 dB	0 deg.	0 ms
L,R	18'	-6.7	27 L,R	9
front	23.7	-9.1	0	14
back	18.3	-6.9	180	9

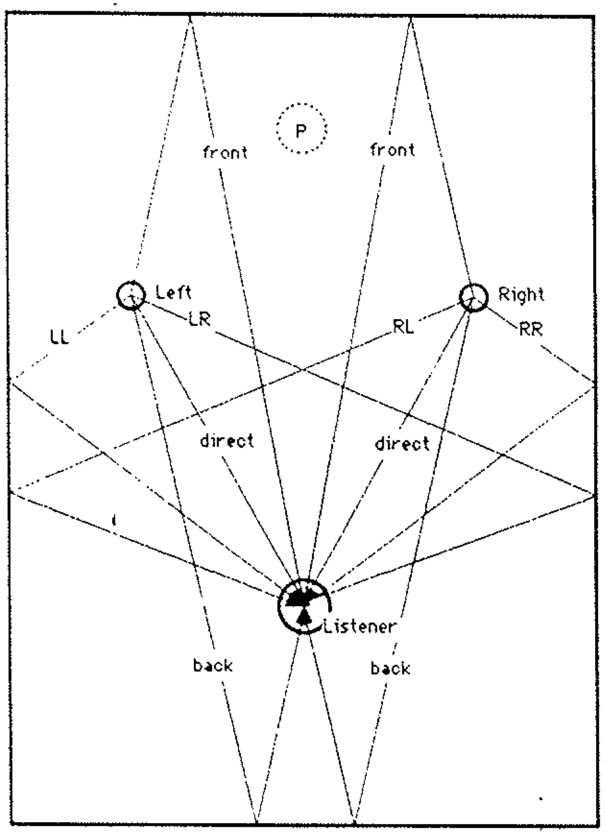
Figure 4: 5-Path model (floor and ceiling excluded) for symmetrically placed monophonic phantom image from conventional (directional) speakers in 16x22 room.



MONOPHONIC PHANTOM IMAGE, SYMMETRICAL, CONVENTIONAL SPEAKERS

Artifact	dist.	atten/direct	angle	Δt
direct	9.6'	0 dB	29 L,R	0 ms
LR, RL	22.25'	-7.3	68 L,R	11 ms
back	20.9'	-6.8	167 L,R	10 ms

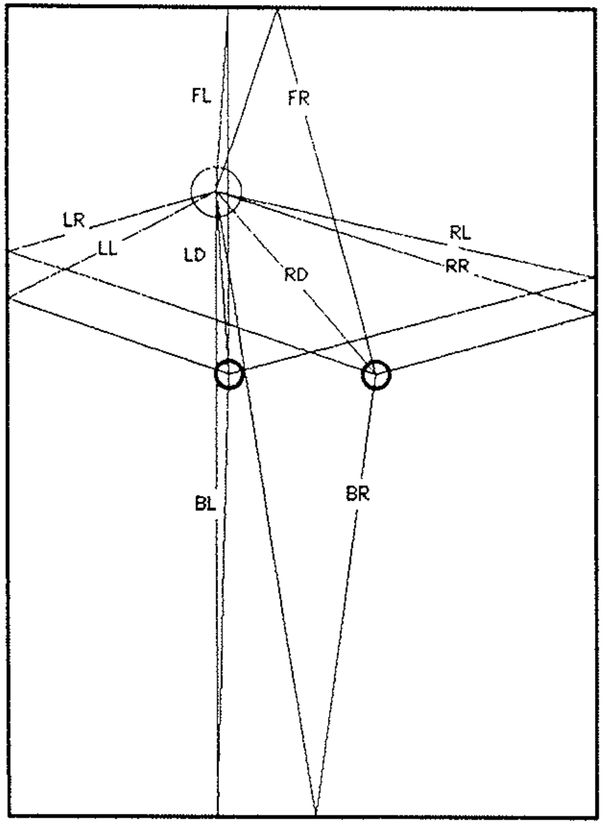
Figure 5: 5-Path model (floor and ceiling excluded) for symmetrically placed monophonic phantom image via omnidirectional speakers in 16x22 room.



MONOPHONIC SOURCE, OMNI SPEAKERS

ARTIFACT	Dist.	Atten	angle	Δt
Direct	9.6'	0 dB	29 L,R	0 ms
LL, RR	14.1'	-3.3	37 L,R	4
LR, RL	22.25'	-7.3	70 L,R	11
Front	24.2	-8	11 L,R	13
Back	20.8	-6.7	167 L,R	10

Figure 6: 5-Path models (floor and ceiling excluded) for 'spaced omni" stereophonic recording in 16x22 room(!). "Atten' and "ΔT" is data that will be emitted playback speakers.



LEFT	DIST	ATTEN.	ΔT
LD	5'	0 dB	0 ms.
LF	15'	-9.5	9
LL	12.75'	-8.1	7
RL	20.9'	-12.4	14
BL	27'	-14.6	19

RIGHT	DIST	ATTEN	ΔT
RD	6.6'	0 dB	0 ms
RF	15.6'	-7.5	8
LR	16.4'	-7.9	9
RR	17.1'	-8.3	9
BR	29.25'	-12.9	20