The Detection of Reflections in Typical Rooms*

SEAN E. OLIVE AND FLOYD E. TOOLE

National Research Council, Division of Physics, Ottawa, Ont. K1A 0R6, Canada

Reflected sounds influence the timbre and spatial character of live and reproduced sounds. Most investigations of reflections have focused on the performance of live sounds in large halls. Current interest in the acoustical interactions of rooms, loudspeakers, and listeners requires further and possibly more relevant data than have been available. In this study, the effects of reflected sounds were examined as they would occur in stereophonic reproduction in rooms of domestic or control-room size. Thresholds were determined as a function of level relative to the direct sound, the angle of incidence, spectrum, the temporal form of the signal, and reverberation. The relationships between audible effects and measurements, such as energy–time curves and frequency response, are discussed.

0 INTRODUCTION

The importance of reflected sound on the perception of live and reproduced sound is unquestioned. Over the years, several investigators have looked into the various perceptual effects that reflections can have, the thresholds of their detection, and the limits of their tolerability. Most of the work has been motivated by the need to understand the architectural acoustics of performing spaces, concert halls, and the like, and by a quest to improve techniques for sound reinforcement.

More recently, concerns about reflections within rooms intended for stereophonic reproduction, such as studio control rooms and domestic listening rooms, have renewed interest in the subject. Information from the earlier studies has been extremely useful, but unfortunately it was not always directly applicable to the problem at hand.

The aspects in which the existing studies seem to be most limited are as follows:

1) The emphasis on large rooms has meant that there are limited data relating to the shorter time delays associated with reflections in rooms of domestic or control-room proportions.

2) The concern for good communication of speech and the acoustical embellishment of classical music has limited the range of program material used in the experiments. There have been suspicions that close-miked or synthesized instruments, especially those with strong percussive sounds, might be perceived differently.

3) The perceptual priorities that apply to live performances may differ from those in effect in other critical listening situations, such as control-room monitoring or stereo reproduction for critical audiophiles. For example, it is clear that most listeners in these situations expect more spatial discrimination and localization resolution than is typically available to listeners in concert halls. Similarly, spaciousness and timbral enhancement of the kind that are expected of a concert hall may be unwelcome if they are added by an environment for stereo listening.

4) Reflected sounds used in the studies have generally had the same sound spectrum as the direct sound. This is not a common situation in reality, given the frequency-dependent directivity of sound sources and selective sound absorption in room boundaries and furnishings.

5) Measurements of the sound levels used in quantifying the direct and reflected sounds have generally been based on spectral levels or frequency response. New time-domain measurements require examination to ensure that the visual representation corresponds to the auditory effects of the sounds.

This paper is an attempt to consolidate some of the data from earlier studies and to fill in certain missing information.

* Presented at the 85th Convention of the Audio Engineering Society, Los Angeles, CA, 1988 November 3–6; revised 1989 June 12.

1 BACKGROUND

The earlier of two similar sounds arriving from different directions tends to dominate the perceived localization. This effect, known for over 100 years, is generally known in the fields of acoustics and psychoacoustics as the *precedence effect* [1], [2], or the *law of the first wave front*, although the name *Haas effect* [3] has become increasingly common in the context of audio. It is an effect fundamental to our ability to accurately localize sound sources in acoustically complex environments.

The reflections arriving after the first, or direct, sound may be somewhat suppressed from the point of view of localization, but there are other perceptual effects related to these later sounds, and these are not necessarily diminished. For example, Haas noted changes in loudness, sound quality, "liveliness," and "body" [3, p. 150], and a "pleasant broadening of the primary sound source" [3, p. 159]. In the context of his major concern, speech reproduction, the loudness gain was a benefit, and the other modifications were of little concern so long as they did not reduce intelligibility.

In the context of concert halls for musical performance, it is clear that the addition of timbral richness and the creation of a generous spatial effect are, up to a point, highly desirable effects of reflected sounds. Indeed, so desirable are they that it has become an objective of concert-hall design to optimize the amounts of these effects for various kinds of music, even to the point of sacrificing some of the precision with which the originating sound sources can be localized.

In the context of stereophonic sound reproduction, things are not yet so well defined. The various interactions between the recording technique and the reproduction technique are not fully identified, nor are the expectations of listeners, in different contexts, to different types of music. Loudspeakers of certain directional properties in certain types of rooms may be more flattering to some styles of recording than to others. Professionals in control rooms may wish to be able to hear details that, in recreational listening, even they are less critical of.

In any event, there is good reason to clearly understand the thresholds at which reflected sounds within rooms begin to affect various aspects of reproduced sound. It is not sufficient to apply criteria developed in the context of communicating intelligible speech, or flattering a live orchestral performance.

2 EXPERIMENTS

2.1 Experimental Apparatus and Procedure

In order to manipulate the various experimental parameters independently, a technique of sound-field synthesis was used. All of the important direct and reflected signals were modified using analog and digital signal processing, and were reproduced by loudspeakers arranged within the appropriate listening environment.

The system used throughout the experiments is shown in Fig. 1. The flexibility is such that all of the operational modules can be adjusted or removed individually, and any of the signal inputs can be electrically added to each other and routed to any of the loudspeakers. The loudspeakers used in the experiments were small high-quality two-way units, with very flat axial responses, matched within 1 dB from 20 Hz to 20 kHz. They were located at a distance of 2 m, at positions around the listener to simulate a direct sound (0° horizontal, 0° vertical, relative to a position horizontally in front of the listener), a lateral wall reflection (65° H, 0° V), a vertical ceiling reflection (0° H, 60° V), and a reflection from the rear of the room (115° H, 0° V). The majority of the experiments were done with this listening configuration set up in an anechoic chamber, but certain experiments, as noted, were set up in a normal room (an IEC recommended listening room [4]) modified to provide two different reflected sound fields.

The "control" module in Fig. 1 is the listener's control box. With it, by means of a potentiometer, the listener was required to adjust the level of the test reflection until it was at the detection threshold, somewhere between the conditions of "just audible" and "just not audible." A nonlinear multiturn potentiometer was used so that the knob position would not be a reliable clue to the listener. To confirm what the target and background sounds were, the listener could, at will, switch to the signal without the test reflection, or to the signal with the test reflection at maximum level (10 dB above the level of the direct sound). No time limit was imposed on this adjustment.

This is a variation of the well-known "method of adjustment," and it is believed that the resulting

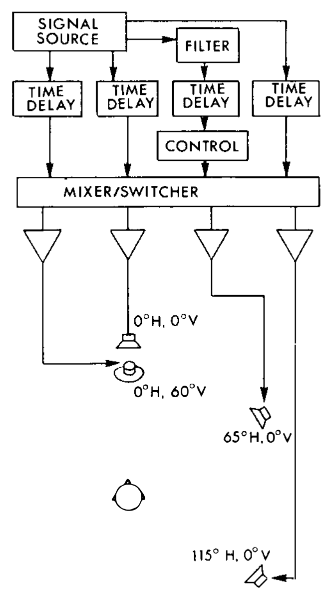

Fig. 1. System used to control various experimental sound-field parameters. The signal is passed through multiple digital delay units and processors to create the necessary delayed signals. These are then mixed together, amplified, and routed to loudspeakers arranged in an appropriate listening environment to simulate any combination of wall reflection, ceiling reflection, rear reflection, and direct or first-arrival sound. The listener is given a control box, allowing level adjustments to be made to the test reflection.

thresholds are close to the minimum attainable. The technique and basic apparatus had been used successfully in earlier experiments by these authors [5]. They also appear to be similar to that used by Seraphim [6].

In general, the *absolute threshold* was of interest. In this, listeners were instructed to respond to any audible change in the nature of the sound itself or of the sound field. In a few experiments, a second kind of threshold was determined. Called the *image-shift threshold*, this was the reflection level at which there was a just-discernible shift in the location or size of the principal auditory image.

Three subjects participated in these experiments. All had near-normal hearing thresholds and two had extensive prior experience in listening tests of various kinds.

The pulse and pink-noise test signals were generated electronically. The speech and castanet signals were taken from the European Broadcasting Union, subjective quality assessment material (EBU-SQAM), on Compact Disc no. 422 204-2. The female speech on track 49 was recorded using a cardioid microphone (AKG C414-ULS) at about 65 cm (26 in), in a drama studio with a midband reverberation time of 0.3 s. The castanets on track 27 were recorded using a B&K 4006 microphone at 1–2 m (40–80 in) in a 1000-m^3 studio with a reverberation time of 1.6 s [7].

Physical measurements of the sound fields, such as energy–time curves and waterfall diagrams (amplitude versus frequency versus time), were made using a Techron 12 time-delay spectrometry system. Other measurements were made with a computer-based measuring system incorporating an Amber 5500.

On the matter of the reproducibility of the experimental results it is sufficient to note that, with only isolated exceptions, the listeners throughout responded with standard deviations typically in the range of 1–3 dB.

2.2 Effects of Various Signal and Sound-Field Parameters

2.2.1 Effects of Signal Type and Angle of Incidence

Although other studies have investigated the effects of incident angle, they have tended to use music or speech signals which, although realistic, are not necessarily the sounds that most readily reveal the presence of delayed sounds. In an earlier study we found that brief impulses (10–40-μs rectangular pulses applied to a loudspeaker) and pink noise were similar to, if not better than, speech and music in revealing the presence of resonances and delayed sounds in detection threshold tests [5]. The following test was conducted as an orienting test to provide a logical extension to the work reported previously.

Fig. 2 shows, for one listener, the absolute thresholds of simulated reflections arriving from a sidewall and from the ceiling, using pulses and pink noise. The experiment was conducted in an anechoic chamber so that the test reflection was the only significant delayed event. The differences caused by these two very different incident directions were very small—in practical terms, negligible. In contrast, the differences resulting from the use of pulses or noise were enormous. At short time delays, below about 10 ms, the continuous sound was more revealing, and at longer delays the discontinuous sound was much more revealing. These results are similar in form to those found in the earlier study where the delayed sound was broadband, as well as variously filtered [5, figs. 9, 10]. Further comment on directional effects may be found in the following section.

2.2.2 Individual Differences among Listeners

Figs. 3–5 show that the three listeners used in these experiments were in substantial agreement about the detectability of reflections from three very different directions, using pink noise. The differences among the three results are somewhat smaller than those found by Schubert, whose 16 subjects yielded thresholds that differed by as much as 12–17 dB over this delay range [8, fig. 8]. We found, as did Schubert, that prior experience at critical listening had little effect on the results. This may be because, in the context of the experiments, the task was such a focused activity. Even though the sound field may have been extremely complex, there was no doubt about the specific aspect of it that the listener was required to attend to.

The unimportant differences between thresholds determined for reflections incident from beside and above are in distinct contrast to the large difference between either of these and reflections arriving from the same direction as the direct sound. The masking effect of the direct sound is clearly much greater when the reflection arrives from the same direction. This effect

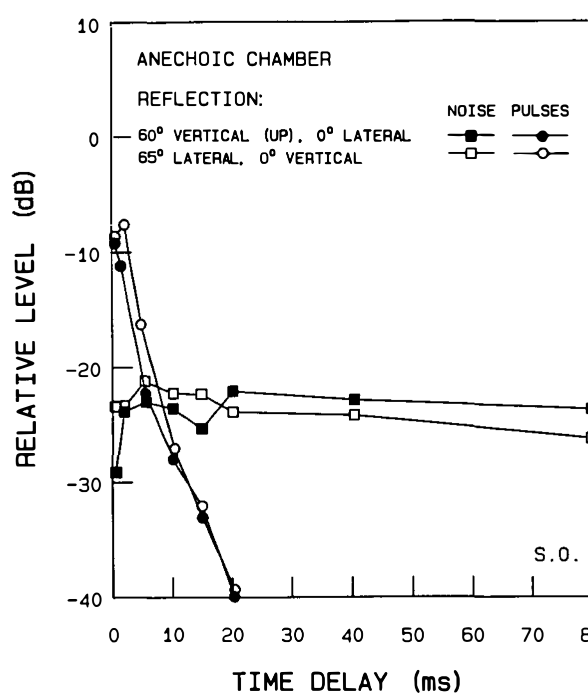

Fig. 2. Absolute thresholds for single vertical reflection (solid symbols) and single lateral reflection (open symbols) using two different types of signals: pink noise (squares) and pulses at 2/s (circles). Anechoic listening.

has been studied in more detail by Schubert [8] and Burgtorf and Oehlschlägel [9]. In practical terms it means that the direction of the reflection is an important factor in assessing the audible consequences. Reflections originating from directions that are close to the principal sound source can be up to 5–10 dB louder before they can be detected. It is interesting that, in this case at least, elevation had an effect similar in magnitude to that of azimuth. The audible effect was quite different, however, with the lateral reflection generating much more of a sensation of spaciousness at levels just above threshold, whereas the vertical reflection, on the median plane, was apparent more as an effect on timbre, with a slight amount of high-frequency spatial diffuseness. These observations are absolutely consistent with studies of spaciousness, and its relationship to interaural cross correlation [10]–[12].

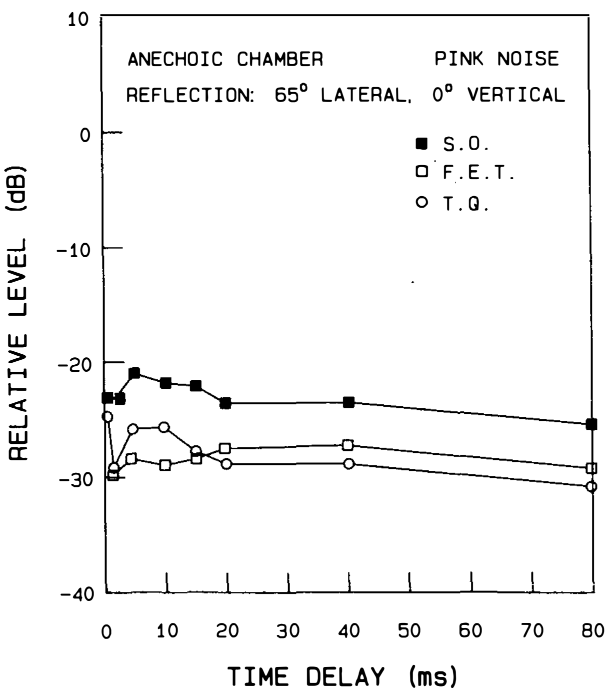

Fig. 3. Absolute thresholds for single lateral reflection shown for three listeners, using pink noise as the signal source. Anechoic listening.

2.2.3 Effect of Signal Type—A Synopsis

Drawing from the present and other works, Fig. 6 shows how the signal type affects the detectability of a single lateral reflection, auditioned in an anechoic environment. Data from Fig. 2 are included in this compilation, confirming the earlier statement that pulses and pink noise are good examples of discontinuous and continuous signal types, respectively. Anechoically recorded or close-miked speech yields thresholds that are intermediate between continuous sounds, like pink noise, or orchestral or virtually any other sound with reverberation. As found previously [5], the addition of even modest amounts of reverberation to impulsive or discontinuous sounds alters their behavior to one resembling naturally continuous sounds. This characteristic is apparent in the threshold curves for castanets, a highly impulsive sound in itself, but combined with some natural reverberation in this recording (see Sec. 2.1).

2.2.4 Adding Reverberation to the Signal

Listening in an anechoic chamber, the influence of adding artificial reverberation to a pulse signal was explored. Fig. 7 illustrates that, at delays below about 10 ms, the added reverberation has the effect of increasing the sensitivity to the reflection, while at delays

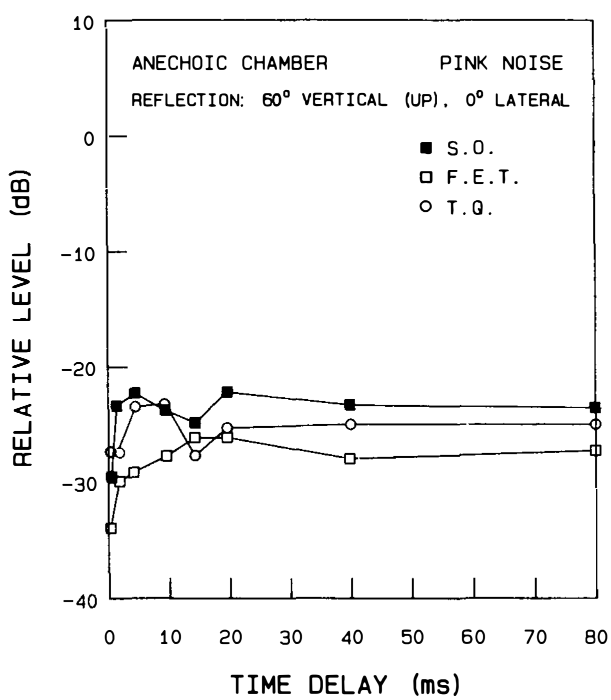

Fig. 4. As Fig. 3, but with test reflection arriving at listener from a vertical angle of incidence.

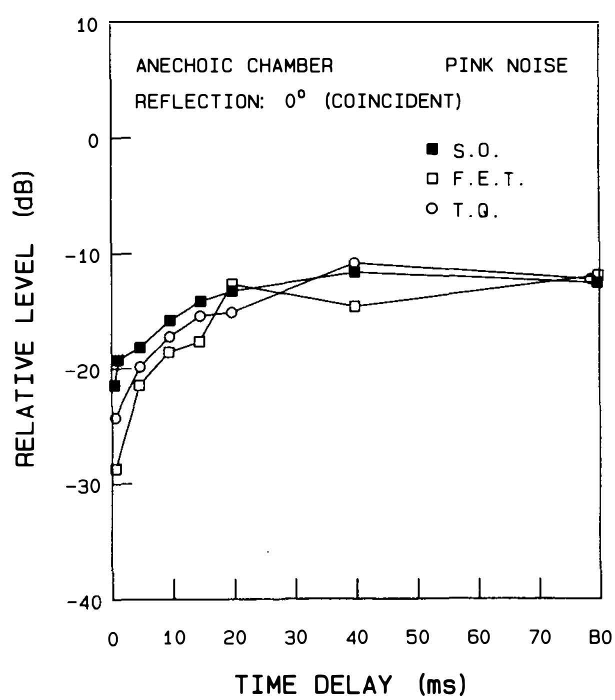

Fig. 5. As Fig. 3, but with test reflection arriving from same direction as direct sound.

above about 20 ms, it has the reverse effect, with the threshold being elevated by 20–30 dB or more.

It is now possible to speculate about the difference, in Fig. 6, between the speech data of Seraphim and those of this study. The Seraphim results were obtained using anechoically recorded speech and this study used close-miked speech recorded in a drama studio (see Sec. 2.1). In the data from this study, the lower threshold at short time delays and the higher threshold at long time delays are absolutely consistent with the results of Fig. 6. The agreement extends even to the proximity of the two speech curves in the delay range of 10–20 ms, where they are within about 3 dB of each other. Similar effects were seen in the audibility of delayed low-Q resonances in our earlier study [5, figs. 16 and 17].

It appears that even a small amount of room-reflected sound has an immediate effect on the absolute thresholds, and the effect tends to be constant above a reverberation time of about 0.3 s. The relationship in the range of reverberation times between zero and 0.3 ms was not explored, as it becomes a moot point whether there even *is* a reverberant sound field in such acoustically "dead" rooms.

2.2.5 Effect of a Single Room Reflection on the Audibility of a Single Reflection in the Signal

The audibility of an early reflection in a signal, a recording, for example, in the presence of reflections within the listening environment is a matter of concern for critical listening. This experiment looks at the detection thresholds for a single repetition (a simulated reflection) added to the signal, as a function of the time delay of a single reflection within the listening room. The room reflection was presented from above, and from the front side, and the listeners determined thresholds for both pink noise and pulses. The added room reflection was at a level 3 dB below that of the direct sound. This is a worst-case situation, since the level is higher than that which would be encountered in real environments.

The results shown in Figs. 8 and 9 indicate that there was no significant change in the threshold of audibility of the reflection in two different signals for any delay or for either direction of reflection within the listening room. Furthermore, the levels observed here were closely similar to the threshold levels observed for no room reflection at all; compare Fig. 8 with this listener's data in Fig. 5.

This finding must be qualified in that it does not comment on the audible effects at levels above threshold. For example, it is possible (albeit untested) that a reflection within the listening room could have audible spatial or timbral effects similar to those of a reflection within the recorded signal. This could lead to a confusion of the origins of the audible effect during critical listening.

What this experiment *does* show is that, given a comparison of two recorded sounds, differing in the amounts of a delayed sound component, even a strong reflection in an otherwise acoustically "dead" listening room will not prevent the difference from being audible. That this should be so is almost intuitive, since the delayed component is part of the signal itself, and is thus included in the direct and all subsequent reflections of the sound.

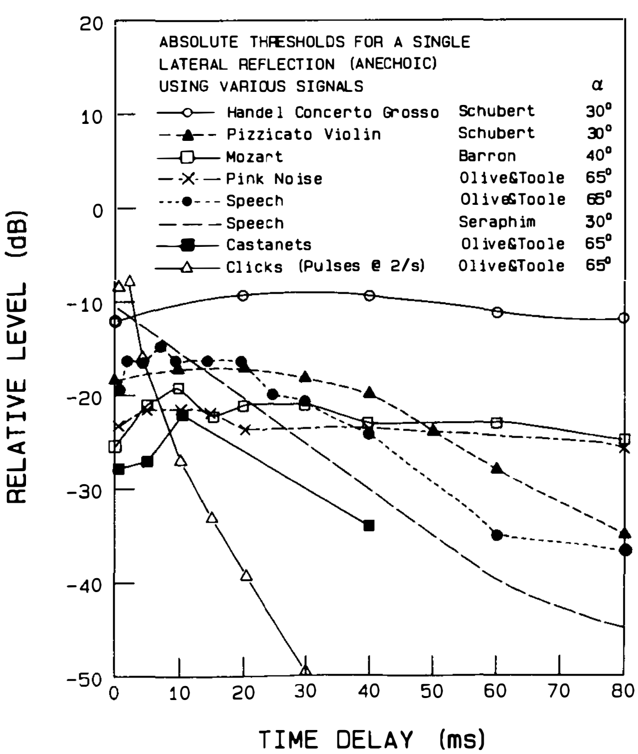

Fig. 6. Comparison of data from this study and from Schubert [8], Barron [10], and Seraphim [6] showing absolute thresholds for single lateral reflection using a variety of natural and artificial signals. All studies used anechoic listening conditions, with slight variations in the horizontal angle of incidence α selected for lateral test reflection.

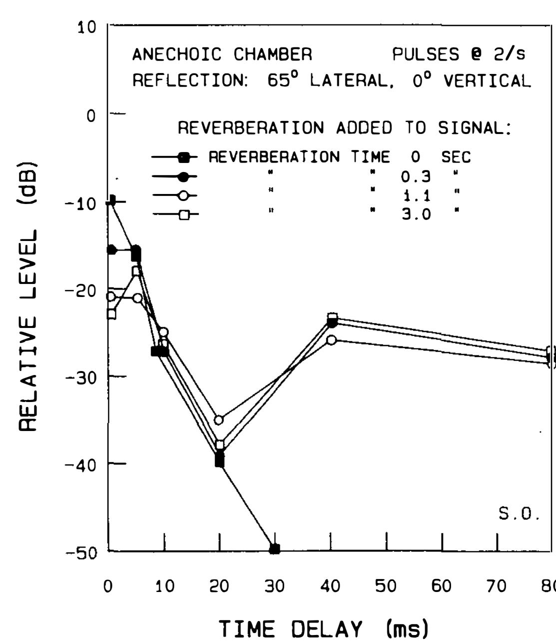

Fig. 7. Absolute threshold for lateral reflection, using pulses at 2/s that are processed with increasing amounts of reverberation using a Yamaha DSP-1. Program—chamber; initial time gap—5 ms, R_{t60} = 0–3.0 s. Anechoic listening.

2.2.6 Influence of Room Reflections and Reverberation

Real listening rooms are not anechoic, nor is there likely to be only one significant reflection. The next logical step in this progression of experiments is to examine the audibility of a "target" reflection in the presence of increasing amounts of room reflections. We have chosen to use a lateral reflection as a test signal. This can be interpreted as a single component of a room sound field, or as a cross-channel delayed component in a stereo or surround-sound system. The question here is: does the increasing complexity of room reflections have an effect on the audibility of an individual reflection within the room itself, or within a multichannel recording reproduced within the room?

To shed further light on the situation, a new threshold measurement was added. In this experiment, listeners were required to identify the now familiar absolute threshold (where *any* change is registered) as well as an image-shift threshold, where the emphasis is on changes in the position and size of the main auditory image.

When the sound level of the test reflection is increased above threshold a, variety of audible effects can arise, depending on the signal and the time delay. For this signal, speech, the effects were as follows. At reflection levels well above the threshold, the sensation was one of image spreading (from the direct-sound source toward the lateral reflection) at time delays less than about 10 ms, spaciousness and image spreading at time delays between about 10 and 40 ms, and spaciousness and an identifiable echo at longer time delays. As the reflection level was reduced, there came a point where the image spreading was no longer significant and the reflection was not separately identifiable. This was the condition required for the image-shift threshold. At this threshold there were still other artifacts, but they did not affect the location or apparent size of the main auditory image. The principal side effects betraying the presence of the low-level reflection were a slight sense of spaciousness and occasional high-frequency sibilant "splashes" localized at the origin of the lateral reflection.

It was observed by listeners that, in general, they were responding to directional and spatial effects, rather than to changes in timbre or sound coloration.

The experiment was conducted in the anechoic chamber, and then moved to the IEC recommended listening room [4] [13, app.]. The physical equipment, the arrangement, and all signal conditions were precisely duplicated in the new environment. The IEC room was employed first in its most "live" form, with the movable drapes compressed into the corners [14, fig. 14]. The midfrequency reverberation time was about 0.4 s. Next, by careful positioning of the drapes and by adding a few patches of sound-absorbing materials, the early reflections from adjacent room boundaries were reduced in amplitude, and the overall reverberation was subdued. Fig. 10 shows energy–time curve (ETC) measurements of the three conditions employed in these experiments. Fig. 10(a) reveals the complex array of reflections typical of an unprepared domestic living room. Fig. 10(b) shows a much tidier display, with all reflections at least 15 dB below the direct sound, and a relatively smooth and dense reverberant tail. We call this acoustically treated condition the relatively reflection-free

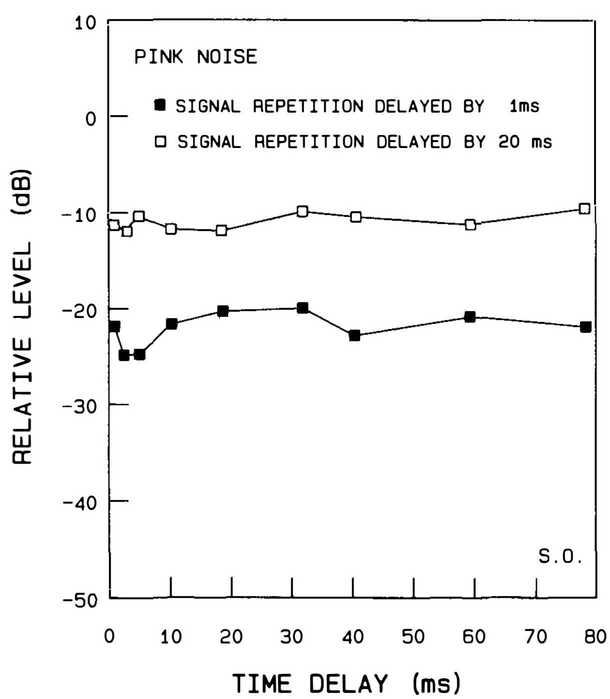

Fig. 8. Thresholds for single repetition in signal path (simulating a reflection in a single channel) in the presence of a vertically incident room reflection (45° V, 0°H) at different time delays. The simulated ceiling reflection is at −3 dB relative to the direct sound. Thresholds were determined for two repetition delays, 1 ms (e.g., reflection from a lectern or music stand) and 20 ms (e.g., ceiling or wall reflection). The signal was continuous pink noise.

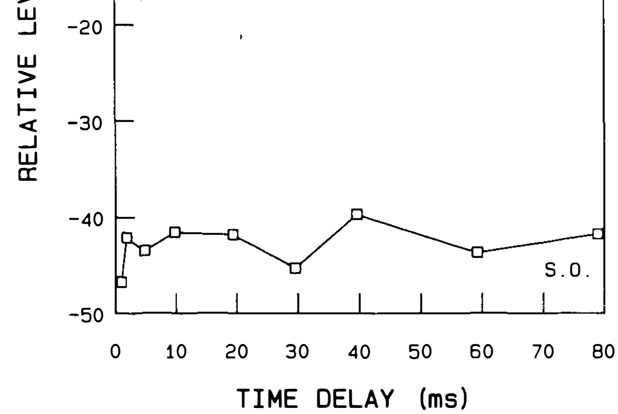

Fig. 9. As Fig. 8, but with simulated wall reflection at −3 dB incident from 65°H 0°V. The signal was pulses at 2/s.

(RRF) room. Fig. 10(c) shows the comparable measurement in the anechoic chamber.

Figs. 11 and 12 show the results for the absolute and the image-shift thresholds, respectively. It is no surprise that the lowest thresholds were obtained mainly in the anechoic condition. However, up to a reflection time delay of about 30 ms, the thresholds changed relatively little between anechoic and RRF conditions, and went up no more than about 6 dB in the moderately reverberant IEC room. This somewhat remarkable finding means that, with speech as the signal, the audibility

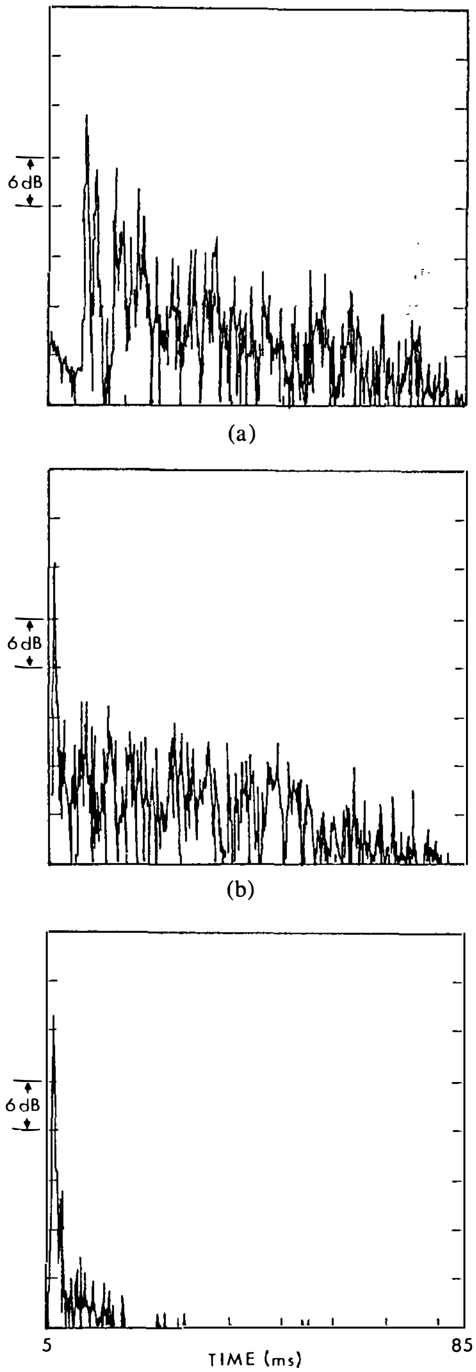

Fig. 10. Energy–time curves showing sound field generated by direct-sound (0° H, 0° V) loudspeaker at the listener's head location. (a) In "normal" (IEC) listening room. The large peaks following the initial arrival indicate, in order, floor, ceiling, and sidewall reflections. (b) As (a), but with absorptive material strategically placed to attenuate early energetic reflections and to reduce reverberation (RRF room). (c) In anechoic chamber used for these experiments. Note that onset of direct sound is displayed about 5 ms later in (a) than in (b) and (c).

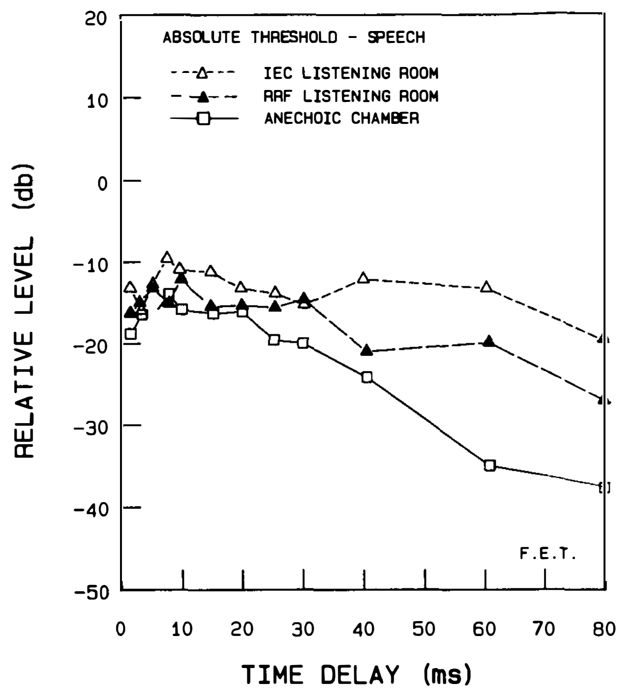

Fig. 11. Absolute thresholds for single lateral reflection (65° H, 0° V), using speech, measured in three different rooms that range from total absence of reverberation (anechoic) to room with some reverberation (RRF) and a normal room with relatively strong early reflections and moderate reverberation (IEC).

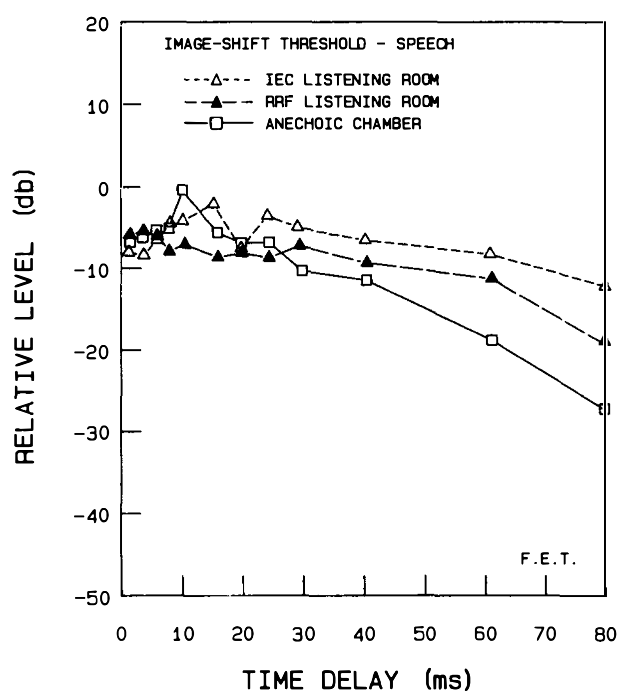

Fig. 12. Using the exact experimental conditions as in Fig. 11, here image-shift threshold is shown, indicating the minimum level at which the lateral reflection is perceived to cause a spreading or shift in localization of the auditory event associated with direct sound.

of a specific lateral reflection at threshold is similar, whether it is auditioned in splendid isolation or in an anechoic room, or whether it is one reflection among the many within a normal room.

In contrast, above about 30 ms, the pattern of change was quite clear, with the thresholds for the delayed reflection rising sharply with each move to a more reflective listening space.

This elevation of thresholds at longer time delays is very likely related to the presence of a significant reflected sound field in the listening room. The continuous-speech test signal built up a more or less continuous background of reflected sound in the listening room in both of its configurations, and it is logical that this should assist in the masking of substantially delayed sounds. Seraphim shared this belief [6, figs. 10 and 11].

The image-shift thresholds were consistently well above the absolute thresholds. Surprisingly, judgments of this more exotic threshold criterion were made with standard deviations that were indistinguishable from those for the absolute thresholds. It appears to be a "durable" phenomenon, even in the acoustically hostile environment of a normal listening room.

The difference between the two thresholds depended on the listening environment. In the anechoic chamber the image-shift threshold was an average of 12.3 dB above the absolute threshold. In the RRF room this difference was 8 dB, and in the normal room it was 7 dB. In other words, introducing progressively increasing amounts of listening-room reflections caused the absolute threshold for a lateral reflection to rise more rapidly than the threshold for its influence on the position and size of the primary auditory image. However, the larger change occurred in the transition from the anechoic chamber to the RRF room.

In this context, it is relevant that, from his anechoic experiments, Seraphim speculated that the threshold for the contribution of a reflection to "room effect," or spaciousness, would be about 6 dB above the absolute threshold [6]. The present results, and the discussion of the auditory effects just below the image-shift threshold earlier in this section, support this speculation.

Combining the last two points, it can now be suggested that increasing amounts of listening-room reflections and reverberation will inhibit the audibility of spaciousness generated by a specific reflection, in the recording or in the room, at levels 7–12 dB below those at which the image location is itself influenced. A major transition occurs, however, when even controlled room reflections are introduced. This is consistent with earlier findings on the effects of adding reverberation to the recorded signal, where even small amounts had a substantial effect on threshold levels (see Sec. 2.2.4). In a very different kind of experiment, Haas found that the disturbance of speech by an early reflection was substantially reduced by going from essentially anechoic listening conditions to an auditorium with a reverberation time of 0.8 s, but was only slightly reduced in a further doubling of the reverberation time to 1.6 s [3, fig. 15].

Reflection thresholds are also affected by the type of signal. However, the effect is much smaller in the presence of room reflections than it is in an echo-free environment. Fig. 13 shows data from earlier work [5, fig. 21], drawn with a revised time scale, showing thresholds for different signals auditioned in a normally reverberant room. In anechoic listening (Fig. 6), comparable signals yielded thresholds with greater differences, especially for impulsive sounds at longer time delays.

As we found in previous work [5], an important factor in the sounds arriving at the listeners' ears appears to be continuity. If the sounds are relatively continuous, the detectability of delayed resonances or of delayed broadband sounds is relatively independent of the amount of the delay. The thresholds are also relatively high. This continuity can be inherent in the structure of the signal (such as pink noise or complex orchestrations), or it can result from reverberation synthetically added to or naturally incorporated in a recorded signal, or it can be contributed by reflections and reverberation in the listening environment. In order for this not to be so, both the recording and the listening conditions must be very nearly anechoic.

2.2.7 Effect of a Single Strong Reflection in the Presence of Differing Amounts of Natural Room Reflections

If the direct sound can inhibit the perception of some potentially distracting early room reflections, it is worth considering whether judiciously timed additional strong reflections might not be able to accomplish more of

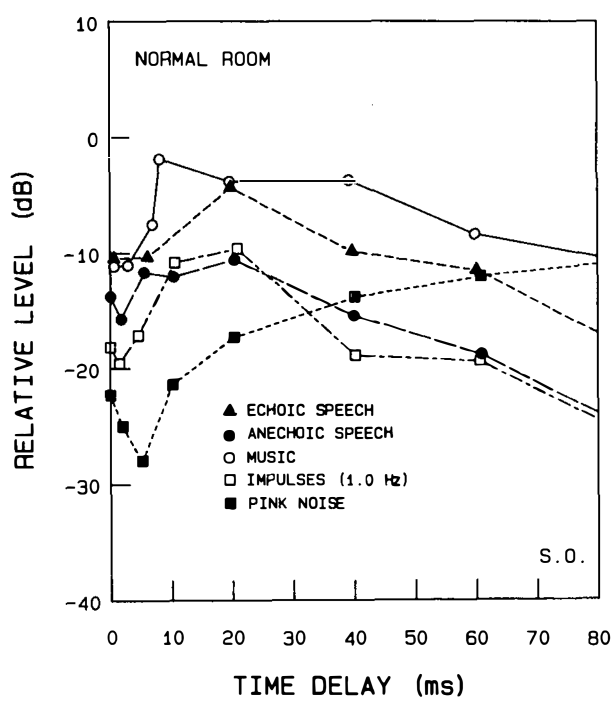

Fig. 13. Results taken from a previous study of detection of resonances [5], showing absolute thresholds for delayed broadband repetition added to signal and reproduced through single loudspeaker in the IEC room, using a variety of different signals.

the same. Based mainly, it appears, on certain data from Seraphim [6, fig. 8], which appeared in a well-known book by Kuttruff [15, fig. VII.4], the idea has developed that, if there is a strong reflection following a reflection-free interval, certain desirable aftereffects of the initial, direct sound are extended in time. Davis has called the deliberate strong reflection a Haas "kicker" [16, p. 221] and the extension of the Haas zone has been called the *Kuttruff effect* [16, p. 363], probably because Kuttruff was careless about clearly attributing his figures to the original sources, in this case, Seraphim. Whether one discusses it in terms of Haas effect [16] or precedence effect [17], the perceptual process involved is a binaural variation of forward temporal masking, wherein a given sound can inhibit the perception, in certain respects, of later sounds having spectral similarities [18].

However, there are some problems with using that particular example from Seraphim's work as a basis for further hypothesis in the realm of room acoustics. The data were simply not acquired in realistic conditions. The widely known figure applies to the case of both the direct and the strongly reflected sounds arriving from the same direction, in an anechoic listening environment. Such circumstances are unrealistic, as are the constant-amplitude delayed sounds extending to 70 ms. This experiment was clearly designed to test a principle, not necessarily to be realistic.

Elsewhere in the same paper, however, Seraphim shows data that *are* more typical of reality. His fig. 6 shows the effect of a reflection arriving from a different direction and, equally important, he examines the effects of strong reflections at different sound levels. It was found that, under these conditions, the absolute thresholds of the test reflections were not constant in level, but decline monotonically with increasing delay. See the Seraphim (speech) curve shown in our Fig. 6, which falls nicely into context with the other data. When Seraphim added a strong reflection at 30 ms, it was found that the threshold curve of the test reflections changed. If the strong reflection was at the same sound level as the direct sound, the thresholds followed the monotonic decline shown here, but at 30 ms it was elevated by about 15 dB, whereupon it continued to decline at the original slope. There was a threshold elevation of +15 dB following the strong reflection. When the level of the strong reflection was reduced to a more realistic 10 dB below the direct sound, the shift in the threshold curve was reduced to about +4 dB.

In the present experiment, a strong reflection from the side rear (115° H, 0° V) was added at a delay of 20 ms and at a level 4 dB below that of the direct sound. As before, the experiment was conducted in the anechoic chamber and both versions of the IEC room. Fig. 14 shows an ETC for the anechoic condition.

The results are shown in Fig. 15. The display is in the form of threshold shifts, so that the effects can be seen clearly. For the absolute threshold, in the anechoic environment, the threshold shift rose to about +8 dB following the −4-dB reflection, which is appropriately between the Seraphim elevations of +15 dB for a 0-dB reflection and +4 dB for a −10-dB reflection. Given the many opportunities for error, the agreement is very good.

However, the real interest is in what happens in the more realistic listening spaces. Fig. 15 also reveals that the forward-masking effect of the strong reflection is barely noticeable in the presence of controlled amounts of room-reflected sound (RRF room) and virtually nonexistent in the normal room reflections (IEC room). Based on previous observations that, in the presence of room reflections, all thresholds tend to be relatively independent of time delay, it is consistent that the effect of adding a discrete strong reflection should be substantially diluted.

This is not to say that adding the delayed component had *no* effect in the two listening rooms. On the contrary, all of the absolute thresholds, at all delays, were slightly elevated. The explanation for this appears to be that, since the masking effect is being provided substantially by the background of room reflections, the addition of the strong reflection raises that background level by about 1.5 dB (the sum of uncorrelated sound fields: the reflected sound field due to the direct sound at 0 dB and that created by the delayed sound at −4 dB). A horizontal line at a level of +1.5 dB would approximate an average curve for the shifts in absolute thresholds in both versions of the listening room.

The image-shift threshold is considerably higher to begin with and, being much more a time-domain phenomenon (it is the interaural signal differences that are of prime concern here), one would not expect it to be influenced by all of the same factors as the absolute threshold. Accordingly, it is not surprising to find that there is little in the way of a consistent effect in the time interval before the strong reflection, and a general elevation of thresholds after the event. The elevation was moderate (no more than about 6 dB) in the anechoic chamber and much less in both listening rooms.

In summary, it seems clear that the addition of a single strong reflection to the sound field of a normal

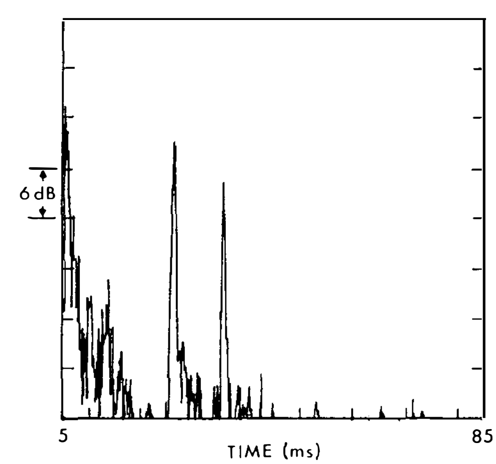

Fig. 14. ETC measured at the listener's head location in anechoic chamber, showing direct sound followed by −4-dB side–rear reflection at 20 ms and −9-dB lateral reflection at 30 ms.

or a relatively reflection-free room has little effect on the audibility of other reflections, arriving before or after the event. Similarly, the effect of the added strong reflection on the image-shift potential of other reflections in the room is very small. On the other hand, it is equally clear that the strong reflection itself is audibly evident in a variety of ways, as it is well above the absolute threshold and slightly above the image-shift threshold, even in the reflective rooms. Other workers have arrived at similar conclusions using quite different techniques [17], [19].

In practical listening environments, therefore, a deliberate strong reflection at 20-ms delay, at a very high level (−4 dB), yielded negligible extension of the masking and image-displacement inhibitions of the direct speech sound. It is possible that essentially anechoic strongly impulsive sounds might trigger more of these precedence (Haas) effect attributes, but prior experiments suggest that the vast majority of sounds would be sufficiently continuous, or contain enough reflections from the recording environment, to be even less demonstrative than the speech signal used here.

2.2.8 Effect of Spectrum on the Audibility of Reflections

Sounds reflected from room boundaries and other surfaces tend to be spectrally modified. The sound absorption of drapes and upholstery, as well as the common fiberglass and acoustic-foam sound-absorbing devices used in dedicated listening rooms, all remove high-frequency energy from reflections. This, combined with the normal off-axis decline in high-frequency response of loudspeakers, means that typical room reflections will have reduced high-frequency content, compared with the direct sound.

All of the experiments reported thus far, including virtually all of the investigations by other workers, have employed reflections having spectra that are essentially identical to the direct sound. There is a clear need to investigate the effect of removing the high frequencies from reflected sounds.

Fig. 16 shows the results of tests done using signals ranging from continuous pink noise through to pulses

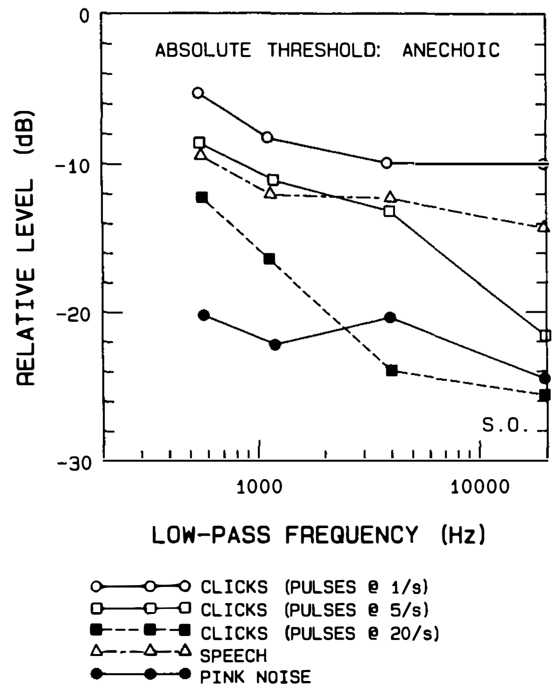

Fig. 16. Absolute thresholds for lateral reflection (50° H, 0° V) that has been low-pass filtered at 20 kHz, 4.4 kHz, 1.2 kHz, and 600 Hz. Reflections were all delayed by 4 ms and signals ranged from discontinuous (pulses at 1/s) to continuous (pink noise).

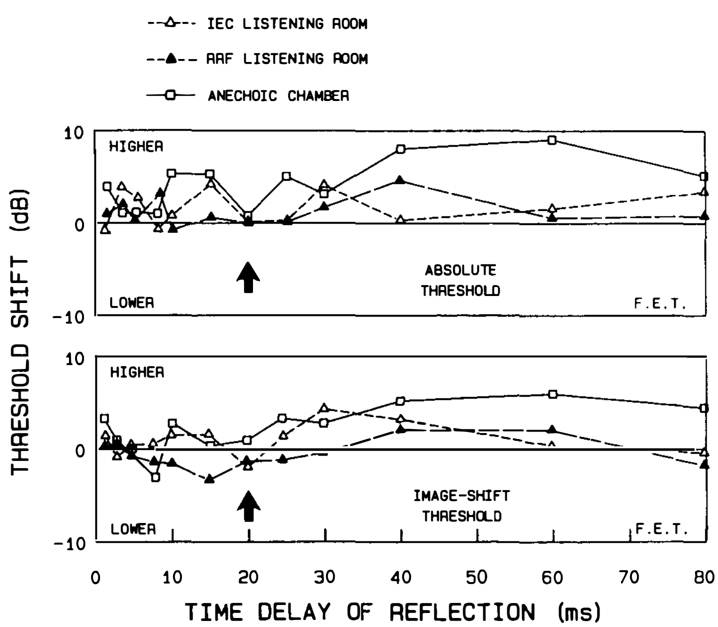

Fig. 15. Shift in absolute (top graph) and image-shift thresholds (bottom graph) of lateral reflection (65° H, 0° V) as a result of adding a strong (−4-dB) side–rear reflection (115° H, 0° V) at a delay of 20 ms. Positive shifts in threshold indicate that lateral reflection is less detectable (i.e., rear reflection produces masking), while negative shifts mean that lateral reflection is more easily detected. Threshold shifts were determined, using speech, in three types of listening rooms: anechoic, RRF, and normal (IEC).

at 1/s. For these two sounds and for speech, the effect of some dramatic spectral limiting (down to 500 Hz low-pass) on the absolute thresholds remained very close to their broadband levels. For the 5/s and 20/s pulses there was a spectral effect, however, in the form of an elevation of the threshold as the high frequencies were removed. This same effect was noted for 10/s pulses in our earlier study [5, fig. 19]. All of this listening was done under anechoic conditions (this study), or through headphones [5]. The reduction in echo disturbance when high frequencies are removed from signals has been noted by other workers, including Haas [3].

The observation of a relatively constant absolute threshold for some of the sounds should not be construed as suggesting that there was no change in the perception. It was obvious to the listeners that their judgments were based on different kinds of sound-field modifications for the different reflected sounds, although they resulted in similar detection thresholds. This and other variables need further study.

2.2.9 Effect of Spectrum on the Measurement of Reflection Levels

The use of ETC measurements to evaluate the reflected sound field of rooms is very attractive. With the right equipment, it is easy to do and it yields a plot of sound events versus time that has a strong intuitive appeal. Unfortunately, depending on how the ETC is implemented, it can be misleading because the indicated amplitude of the sound event may not be a correct portrayal of spectrum levels at all frequencies.

A number of users have noted measurement anomalies in ETC implementations incorporating Hamming windows that spectrally weight the signal in the frequency domain. Such windows suppress the low- and high-frequency contributions to the ETC calculation. Since this is done on a linear frequency scale, the result is a severe reduction in the weight given to low, and even lower middle, frequencies. The errors that result from this very common scheme have been examined in detail by Vanderkooy and Lipshitz [20]. In the following, we will show how these errors relate to the audibility of reflected sounds.

Fig. 17 shows that image-shift thresholds for speech sounds are similar to the corresponding absolute thresholds in that they are relatively unaffected by the removal of high frequencies from the lateral reflection. At least this is the case in the conventional method of making such measurements by means of what is known as spectrum level: the amplitude of the signal over its effective bandwidth or, more simply, its frequency response. Also shown in the figure are the corresponding peak ETC levels, relative to the level of the direct sound. The ETC incorporated Hamming windowing, the default condition of the analyzer used. The measurement suggests that listeners are much (20 dB) more sensitive to the presence of a reflection if the high frequencies have been removed. In fact, the sensitivity to the remaining portion of the spectrum is almost the same as it is to the broadband reflection.

The practical consequence of this observation is that using the peak level of these ETC spikes as a measure of the audible importance of room reflections can lead to a serious underestimation of the importance of reflected sounds that contain less high-frequency energy than the direct sound. Because of loudspeaker directivity and sound absorption by reflecting surfaces, this is likely to be a common situation.

Figs. 18 and 19 show sets of data taken at the threshold conditions of a 30-ms-delayed test reflection, following a 20-ms-delayed strong reflection, in an anechoic chamber. In Fig. 18 the test reflection is broadband, and in Fig. 19 it was 500-Hz low-pass filtered. Each

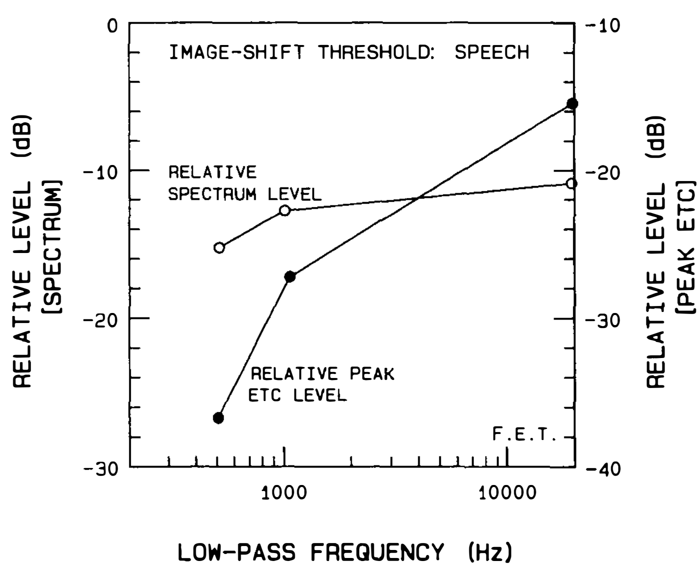

Fig. 17. Image-shift thresholds for 65° lateral reflection that has been low-pass filtered at 20 kHz, 1 kHz, and 500 Hz, all delayed by 30 ms. Measured level of reflection relative to broadband direct sound is shown in two ways. Scale on left shows relative spectrum level of reflected sound at threshold, in the manner used throughout this paper. Scale on right shows corresponding relative levels as indicated by the appropriate peaks in an ETC display. Anechoic listening.

figure shows the ETC for the threshold conditions, the waterfall diagrams including the direct sound and both reflections, and the frequency response of the test reflection extracted from the waterfall diagram at the time "slice" closest to 30 ms. Fig. 20 shows a waterfall diagram for the isolated direct sound and the appropriate time-slice frequency responses for the broadband direct sound and for the strong broadband reflection at 20 ms.

The amplitude of the test reflection, relative to the direct sound, as estimated from the ETC, would be about −9 dB for the broadband case and about −30 dB for the 500-Hz low-pass filtered case. From the frequency responses, the comparable levels would be assessed at about −7 and −5 dB, respectively. The latter are satisfactorily close to the true free-field spec-

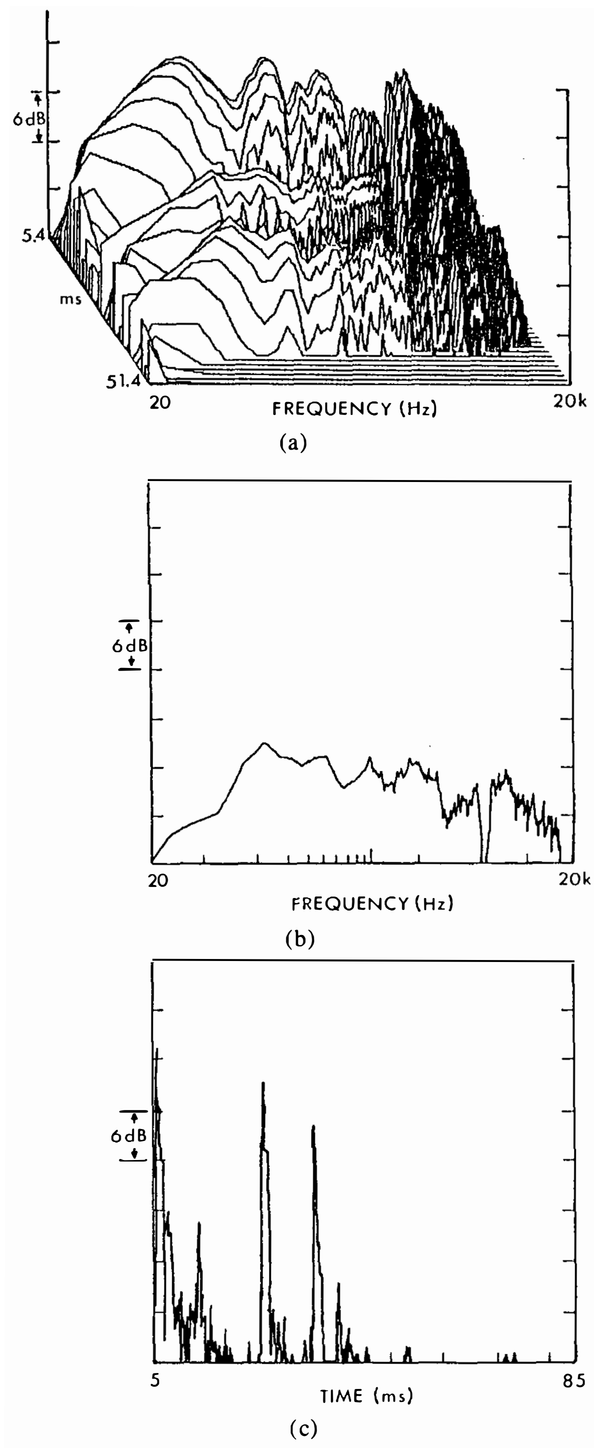

Fig. 18. Measurements made at image-shift threshold conditions for lateral reflection at 30 ms, following a strong side-rear reflection at 20 ms. All three sounds were broadband, and signal was speech. Measurements were made using a time-delay spectrometry system (Techron 12). (a) Three sound arrivals as "waterfall" of amplitude as a function of both frequency and time. (b) "Slice" from waterfall display at time (30 ms) of broadband lateral reflection, after adjustment to threshold level. (c) ETC of complete sound field, with three largest peaks corresponding to direct sound, rear reflection, and broadband lateral reflection.

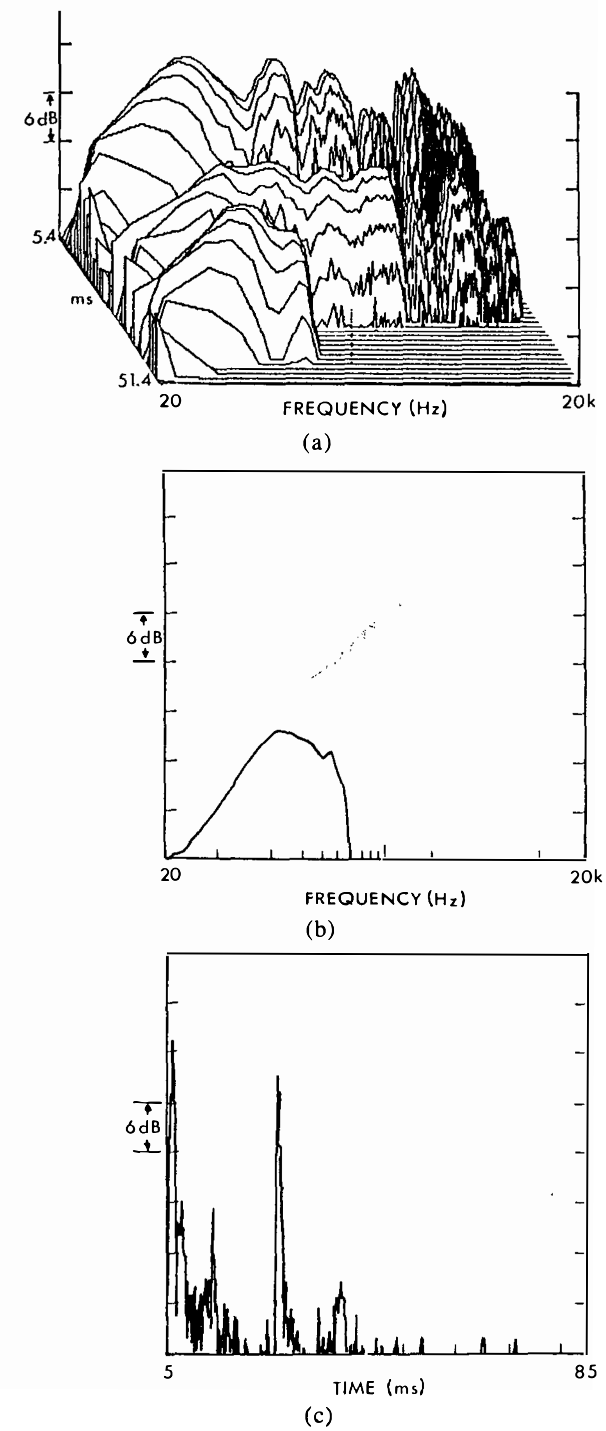

Fig. 19. As Fig. 18, but with lateral test reflection low-pass filtered at 500 Hz and level adjusted to new threshold condition.

trum levels of −5 and −4.5 dB. The small discrepancies most likely result from slight time error and smearing in the waterfall slices, and the ragged frequency responses created by reflections from the listener's chair over which the microphone was located. (In retrospect, the chair should have been removed because it adds nothing to the authenticity of the measurement. Reflections from the listener's own body would, at his or her ears, cause different acoustical interference of similar magnitude.)

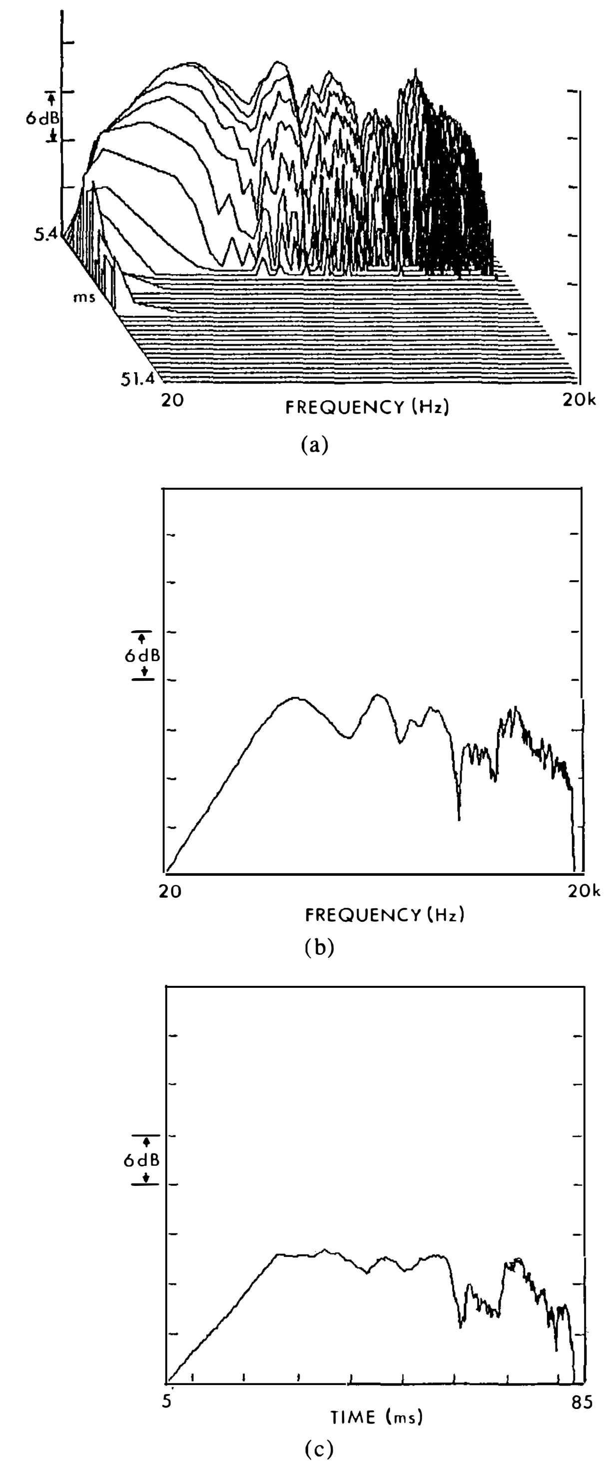

Fig. 20. Data related to Figs. 18 and 19. (a) Direct sound component in isolation. (b) "Slice" in time from waterfall graph of direct sound. (c) 20-ms "slice" in time of strong rear-side reflection.

When there is a substantial difference in spectrum between the sounds being compared, the disparity between the ETC measurement and the spectrum measurement can be very large.

On the other hand, if the spectra of the sounds being compared are similar, the agreement can be excellent. For example, the direct sound and the strong reflection at 20 ms were both broadband. The level of this reflection relative to the direct sound is about −4 dB according to the ETC, and about −6 dB according to a comparison of frequency responses. The true spectrum level was −4 dB.

As in all measurements, it is important to know what one is measuring and what one is looking for in the measurement. In these cases it is clear that if an ETC is to display data that are directly related to the audibility of sound events with differing spectra, the frequency weighting cannot be ignored. We have illustrated one technique for obtaining useful data. There are others.

3 SUMMARY

Sounds reflected within a room have a range of effects on the perception of sounds originating either from live sound sources or from loudspeakers in a stereophonic system. The first sound to arrive, the direct sound, has a dominant effect on the localization of the auditory event, even when the later arrivals are at the same or even a higher sound level. However, this does not mean that the later arrivals have no effect. In fact, depending on the specific demands of the listening situation, reflections may be problematic at levels much below that of the direct sound.

In this paper these effects are classified, and the thresholds quantified, for different modifications to the auditory illusion created by the direct sound alone. In this we have drawn from past studies, as well as generating some new data to fill gaps in the published material.

Finally, we examine a commonly used technique for measuring the amplitudes of reflected sounds and discuss some of the requirements for measurements that are unambiguously related to the subjective effects of those sounds.

The major conclusions from this study may be summarized as follows.

1) There are many fundamental similarities between the detection of broadband reflections and the detection of delayed resonances, in the earlier study by these authors.

2) Delayed sounds arriving from the same direction as the direct sound are often less audible than delayed sounds from most other directions.

3) Discontinuous (impulsive) and continuous sounds exhibit fundamentally different absolute thresholds as a function of reflection delay, but only when there is little or no reflected sound in either the recording or the listening room.

4) Reflected sounds and reverberation, whether they

are in the recording or in the listening room, tend, through multiple repetitions, to add continuity to otherwise discontinuous sounds. The absolute thresholds for most sounds under these conditions tend to be similar: relatively high and relatively independent of reflection delay.

5) An early reflection in a monophonic recorded signal is not significantly masked by a single lateral room reflection at delays ranging from shorter to much longer than that of the signal reflection.

6) Natural reflections and reverberation in a listening room have relatively little effect on the absolute threshold of an individual lateral sound delayed by less than about 30 ms. This delayed sound could be a room reflection or a sound component in a stereo or surround-sound signal. At longer delays the thresholds were progressively elevated as a result of increasing amounts of room-reflected sound. This was true for speech, a moderately discontinuous sound. For more continuous sounds, or sounds with reverberation, the thresholds should be close to the elevated level under all listening conditions.

7) It has been suggested that a deliberate strong reflection following a relatively reflection-free interval could have audibly beneficial effects by renewing the precedence effect. In anechoic listening it was found that, at a sound level where the strong reflection showed slight enhancement of the precedence effect caused by the direct sound, the strong reflection was itself responsible for significant modifications to the auditory illusion. In normal or reflection-reduced listening rooms, the beneficial effects were negligible, and the audible detractions were still present.

8) Reflected sounds rarely have the same spectrum as the direct sound; normally the high-frequency content is reduced. Reducing the high-frequency content of reflected sounds had a relatively minor effect on the absolute thresholds of several common sounds, as assessed by measurements of spectrum level. However, ETC measurements incorporating frequency-weighting windows, such as Hamming, can lead to incorrect estimates of the audibility of reflections. Hamming-windowed ETC measurements remain reliable measures of the relative levels of sound with similar, and limited, bandwidth and as a guide to the effectiveness of removing high frequencies from reflections. Other windows may be more generally useful. Along with the traditional impulse response, the ETC is also an important aid in locating the times of potentially audible reflections.

The results of this study have been reassuringly consistent with previous work in this area. While certain aspects of the findings are well defined, there are others that require more data, and these remain under examination.

4 ACKNOWLEDGMENT

This work was supported by Audio Products International, Scarborough, Ontario.

5 REFERENCES

[1] H. Wallach, E. B. Newman, and M. R. Rosenzweig, "The Precedence Effect in Sound Localization," *Am. J. Psychol.*, vol. 62, pp. 315–336 (1949); reprinted in *J. Audio Eng. Soc.*, vol. 21, pp. 817–826 (1973 Dec.).

[2] P. M. Zurek, "The Precedence Effect," in *Directional Hearing*, W. A. Yost and G. Gourevitch, Eds. (Springer, New York, 1987), chap. 4.

[3] H. Haas, "The Influence of a Single Echo on the Audibility of Speech," *J. Audio Eng. Soc.*, vol. 20, pp. 145–159 (1972 Mar.), reprint of a translation by K. P. R. Ehrenberg of Haas's Ph.D. dissertation of 1949.

[4] "Listening Tests on Loudspeakers," in *Sound System Equipment*, International Electrotechnical Commission, Publ. 268-13, pt. 13; 1985.

[5] F. E. Toole and S. E. Olive, "The Modification of Timbre by Resonances: Perception and Measurement," *J. Audio Eng. Soc.*, vol. 36, pp. 122–142 (1988 Mar.).

[6] H.-P. Seraphim, "Über die Wahrnehmbarkeit mehrerer Rückwürfe von Sprachschall," *Acustica*, vol. 11, pp. 80–91 (1961).

[7] "User's Handbook for the EBU-SQAM Compact Disc," European Broadcasting Union, Tech. 3253-E, Brussels, 1988 Apr.

[8] P. Schubert, "Die Wahrnehmbarkeit von Rückwürfen bei Musik," *Z. Hochfrequenztech. Electroakust.*, vol. 78, pp. 230–245 (1969).

[9] W. Burgtorf and H. K. Oehlschlägel, "Untersuchungen über die richtungsabhängige Wahrnehmbarkeit verzögerter Schallsignale," *Acustica*, vol. 14, pp. 255–266 (1964).

[10] M. Barron, "The Subjective Effects of First Reflections in Concert Halls—The Need for Lateral Reflections," *J. Sound Vib.*, vol. 15, pp. 475–494 (1971).

[11] Y. Ando, *Concert Hall Acoustics* (Springer, New York, 1985).

[12] J. Blauert, *Spatial Hearing* (M.I.T. Press, Cambridge, MA, 1983).

[13] F. E. Toole, "Listening Tests—Turning Opinion into Fact," *J. Audio Eng. Soc.*, vol. 30, pp. 431–445 (1982 June).

[14] F. E. Toole, "Loudspeaker Measurements and Their Relationship to Listener Preferences: Part 2," *J. Audio Eng. Soc.*, vol. 34, pp. 323–348 (1986 May).

[15] H. Kuttruff, *Room Acoustics*, 2nd ed. (Applied Science Publ., London, 1979).

[16] D. Davis and C. Davis, *Sound System Engineering*, 2nd ed. (Howard Sams, Indianapolis, IN, 1987).

[17] J. Wrightson, "Psychoacoustic Considerations in the Design of Studio Control Rooms," *J. Audio Eng. Soc.*, vol. 34, pp. 789–795 (1986 Oct.).

[18] P. Divenyi and J. Blauert, "On Creating a Precedent for Binaural Patterns: When Is an Echo an Echo?"

in *Auditory Processing of Complex Sounds*, W. A. Yost and C. S. Watson, Eds. (Lawrence Erlbaum, Hillsdale, NJ, 1987).

[19] J. Wrightson and R. Berger, "Influence of Rear-Wall Reflection Patterns in Live-End-Dead-End-Type Recording Studio Control Rooms," *J. Audio Eng. Soc.*, vol. 34, pp. 796–803 (1986 Oct.).

[20] J. Vanderkooy and S. P. Lipshitz, "Uses and Abuses of the Energy–Time Curve," presented at the 85th Convention of the Audio Engineering Society, *J. Audio Eng. Soc. (Abstracts)*, vol. 37, pp. 399–400 (1989 May).

THE AUTHORS

S. E. Olive

F. E. Toole

Sean E. Olive was born in Brockville, Ontario, Canada, in 1959. He received a Bachelor of Music degree in 1982 from the University of Toronto and a Master of Music degree in sound recording in 1986 from McGill University, Montreal, Quebec, Canada.

He taught sound recording at McGill University and has engineered and produced numerous recordings of live concert performances and studio sessions. Since 1985 May he has been guest researcher at the National Research Council in Ottawa, conducting research in psychoacoustics related to the perception and measurement of loudspeaker and listening-room performance. He has coauthored, with Floyd Toole, papers on the perception and measurement of resonances in loudspeakers and reflections found in listening rooms. He is currently investigating the front end of the recording–reproduction chain, evaluating popular microphones used in commercial recordings.

Mr. Olive is a member of the Audio Engineering Society.

•

Floyd E. Toole was educated in electrical engineering at the University of New Brunswick before being diverted into psychoacoustics for his Ph.D. studies at the Imperial College of Science and Technology, University of London, England. After graduation in 1965 he joined the acoustics group in the Physics Division of the National Research Council in Ottawa, Ontario, Canada. Since then he has been involved in research activities in the areas of noise control, architectural acoustics, and psychoacoustics, mostly as they apply to the problems of sound reproduction through loudspeakers in rooms.

He is probably best known for his work in the subjective and objective evaluation of loudspeakers, where he developed detailed procedures for subjective measurements of sound and stereo quality and identified the technical measurements of loudspeaker performance that were most directly related to these subjective assessments. For this he received the AES Publications Award in 1988. His recent research has been focused on the audibility and measurement of resonances and their effect on the perception of timbre. His current investigations are concentrating on reflections in listening rooms and loudspeaker–room interactions.

Dr. Toole has published several papers in the journals of the ASA and AES, chapters in two recent books on audio engineering, and dozens of articles in various audio magazines. He is also familiar as a speaker at many of the AES sections. He is a member of the ASA, a Fellow of the AES, and chairman of the newly formed Working Group on Listening Tests of the AES Standards Committee. For many years he has been active in IEC standards work on loudspeaker and headphone measurements and listening tests.