CLICK ANYWHERE on THIS PAGE to RETURN to DRIPPING WATER SOUND SOURCES at InspectApedia.com

Localization of sound in rooms, II: The effects of a single reflecting surface

Brad Rakerd
Department of Audiology and Speech Sciences, Michigan State University, East Lansing, Michigan 48824

W. M. Hartmann
Department of Physics, Michigan State University, East Lansing, Michigan 48824

(Received 5 November 1984; accepted for publication 24 April 1985)

Auditory localization was studied in a room bounded by a single acoustically reflective surface. The position of that surface was varied so as to simulate a floor, a ceiling, and left and right side walls. The surface was eliminated in one condition so that we could examine localization in free field for purposes of comparison. Using a source identification method we assessed the influences of these various room configurations on the localization of both slow-onset and impulsive sine tones of low frequency (500 Hz). We also measured the steady-state interaural-time-difference (ITD) and interaural-intensity-difference (IID) cues available to subjects in the different room configurations and compared these data with the perceptual judgments. Our results indicate the following: (1) A sound must include transients if the precedence effect is to operate as an aid to its localization in rooms. (2) Even if transients are present the precedence effect does not eliminate all influences of room reflections. (3) Due to the interference of reflections large interaural intensity differences may occur in a room and these have a considerable influence on localization; this is true even at low frequencies for which IID cues do not exist in a free field. (4) Listeners appear to have certain expectations about the reliability and plausibility of various directional cues and perceptually weight the cues accordingly; we suggest that this may explain, in part, the large variation in time-intensity trading ratios reported in the literature and also the differing reports regarding the importance of onsets for localization. (5) In this study we find that onset cues are of some importance to localization even in free field.

PACS numbers: 43.66.Pn, 43.66.Qp, 43.55.Br [RDS]

INTRODUCTION

This paper is a report of our second experimental study of auditory localization in rooms. The first study (Hartmann, 1983a; referenced below as paper I) was conducted in a variable-acoustics concert hall—the Espace de Projection (ESPRO) at the Institut de Recherche et Coordination Acoustique/Musique. The wall absorption and ceiling height of the ESPRO are both manipulable, and their effects on localization were examined for several different classes of sounds.

One notable finding in the ESPRO was that the azimuthal localization of impulsive tones significantly improved when the ceiling was lowered. Since the principal acoustical consequence of this manipulation was to reorder the arrival times of reflections from the room's boundaries, we interpreted this result as indicating a limitation of the precedence effect (Wallach *et al.*, 1949; Haas, 1951; Blauert, 1971; Zurek, 1980) as it operates in rooms.[1]

Our attention centered, particularly, on the *first* reflections from the ceiling and side walls. The azimuth of a ceiling reflection invariably agreed with the azimuth of a direct sound and, as a consequence, might be expected to have reinforced listeners' sense of location. By contrast, the azimuth of a side wall reflection disagreed with that of a direct sound and was a potential source of localization error. When the ceiling was lowered, its reflection reached a listener much earlier than the side wall reflections. We hypothesized that this gave the ceiling reflection an increased importance and accounted for the enhancement of localization accuracy.

It might be hypothesized, more generally, that any reflection will influence the localization of impulsive tones in a manner which is fully determined by its azimuth and by its arrival time relative to a direct sound. One purpose of the present study was to provide a direct test of the contribution of reflection azimuth. To do so, we needed to be able to manipulate the directions of acoustical reflections while preserving the essential features of a room. We achieved this by designing a room in which there was one and only one sound-reflecting surface. The position of that surface was systematically manipulated across conditions so that we could examine its impact on localization and the precedence effect.

A second finding reported in paper I was that listeners had great difficulty localizing low-frequency tones with slow onsets. Their highly inaccurate localization judgments were nevertheless made with some consistency, as demonstrated by the fact that the variability of those judgments was significantly less than chance. Apparently the listeners were able to devise some strategy for localizing such sounds, albeit an inappropriate one. As part of the present study, we conducted an experiment to determine how listeners attempt to cope with slow-onset tones. For various arrangements of our room, we compared their perceptions of the location of slow-onset tones with measurements of the corresponding interaural time and intensity cues.

0001-4966/85/080524-10$00.80 © 1985 Acoustical Society of America

I. EXPERIMENT 1: LOCALIZATION OF IMPULSIVE TONES

Experiment 1 examined the effects of various room reflection azimuths on the azimuthal localization of impulsive tones. As nearly as possible, the procedures for this experiment paralleled those of paper I.

A. Method

1. Subjects

Ten subjects participated in the experiment. All of them had normal hearing according to self-report. All but two subjects, the authors, were ignorant of the experimental hypothesis under test.

2. Sources and stimuli

Stimuli were delivered from eight loudspeaker sources, constructed and arranged as follows. Eight, 2-in. drivers were selected from 20 for best-matched frequency response. Each driver was mounted in a cylindrical container, which was then packed with acoustical insulation and sealed. A speaker hole in the front of the container was covered by a wooden disk, 0.75 in. (19 mm) thick, with a 0.75-in.-diam. hole cut in its center to emit the signal. The eight containers were arranged in an arc and labeled with the numbers *1* through *8*, ordered left to right.

A chair was placed 10 ft (3 m) away, facing the center of this arc. At that distance, there were 3° of angular separation between adjacent sources, 0.36° of which were taken up by source aperture. The levels of all sources were adjusted to 50 dBA for a continuous 500-Hz sine tone measured at the chair under anechoic conditions. The stimuli for this experiment were 500-Hz sine tone pulses of 50-ms duration, turned on and off at zero crossings.

3. Room conditions

Testing was done in an anechoic chamber, IAC 107840, which was 15 ft long, 11 ft wide, and 8 ft high (4.6×3.4×2.4 m). We created different room conditions by varying the position of an acoustically-reflective panel within this chamber. The panel was a 4×8 ft (1.2×2.4 m) sheet of particle board, heavily braced on one side, and painted with epoxy enamel. This panel weighed 95 lb (422 N). The panel's reflection coefficient at 500 Hz was estimated in two ways. The first was to bounce a train of 8-ms tone pulses from it and to compare reflected and incident signals. The second method was to measure the maximum change in the intensity of a continuous tone that resulted from introducing the panel into the anechoic chamber. Both methods indicated that the panel had a reflection coefficient of 100%.

In what will be called the *ceiling* condition, the panel was suspended 3 ft (0.9 m) above the loudspeaker sources. This corresponded to a distance of approximately 2.5 ft (0.8 m) above the subjects' ears. (The precise distance varied slightly depending on a subject's height.) Reflection delay times, and the angle of reflections, for this and the other room conditions are reported in Table I. The delays ranged from 0.6 to 2.3 ms, and averaged 1.4 ms for all conditions combined. In the *floor* condition of the experiment, the panel was situated 3 ft beneath the sources and, typically, 3.5 ft (1.1 m) beneath the subjects' ears. Two side-wall conditions, *left-wall* and *right-wall*, respectively, were produced by positioning the panel 3 ft to either side of the speaker midline, as viewed from the subject's position. For comparison, testing was also done in an *empty-room* condition in which the panel was removed, leaving only the absorbing surfaces of the anechoic chamber.

TABLE I. Delay of reflections and (angle of reflections), both relative to the direct sound. The delay times are in milliseconds, the angles in degrees.

	Room condition				
Source	Empty	Floor	Ceiling	Left wall	Right wall
1	···	1.5 (31)	1.2 (32)	0.6 (12)	2.3 (49)
2	···	1.5 (31)	1.2 (32)	0.9 (18)	2.1 (44)
3	···	1.5 (31)	1.2 (32)	1.1 (23)	1.8 (39)
4	···	1.5 (31)	1.2 (32)	1.4 (29)	1.6 (33)
5	···	1.5 (31)	1.2 (32)	1.6 (33)	1.4 (29)
6	···	1.5 (31)	1.2 (32)	1.8 (39)	1.1 (23)
7	···	1.5 (31)	1.2 (32)	2.1 (44)	0.9 (18)
8	···	1.5 (31)	1.2 (32)	2.3 (49)	0.6 (12)

4. Perceptual task

Subjects were tested, individually, with a source-identification method, as described in paper I. Accordingly, they made a forced-choice decision as to which of the eight numbered speakers was the most likely source of the sound presented on each experimental trial. Subjects reported their choices to the experimenter through an intercom; they received no feedback. So as to maintain the angular separation of sources, we asked subjects to remain seated in the chair throughout testing. Otherwise, the subjects were free to do whatever they thought most useful to perform the task.

The order of testing in the five room conditions was randomly varied across subjects. In each condition, the testing comprised two blocks of 80 trials, for 160 trials in all. Within a block, there were ten stimulus presentations from each of the eight sources. The order of these presentations was randomized with the constraint that no source was presented more than twice in succession.

B. Results and discussion

A comparison of subjects' responses with the known source locations permitted computation of the following statistics, all of which are expressed in degrees (see paper I for a detailed description of the calculations): (1) Root-mean-square (rms) error, symbolized as D, which indicates the overall accuracy of localization. (2) Signed error, E, which indicates any systematic response biases to the left (negative E) or right (positive E). (3) Response standard deviation, s, which indicates the variability of subjects' choices. The averaging of these statistics over all trials for a given source k, for a single subject, is indicated by parentheses around the source number, e.g., $E(k)$. Averaging over all sources for a given subject is indicated by a bar above the corresponding symbol, e.g., $\overline{E}$. Finally, averaging over subjects is indicated by angular brackets around a symbol, e.g., $\langle \overline{E} \rangle$.[2] In Table II, we report the values of $\langle \overline{D} \rangle$, $\langle \overline{E} \rangle$, and $\langle \overline{s} \rangle$ which resulted

TABLE II. Results for the localization of impulsive tones (experiment 1), depending on room condition. $\langle \bar{D} \rangle$ is rms error, $\langle \bar{E} \rangle$ is signed error, $\langle \bar{s} \rangle$ is response standard deviation.

Room condition	$\langle \bar{D} \rangle$ deg	$\langle \bar{E} \rangle$ deg	$\langle \bar{s} \rangle$ deg
Empty room	3.0	−0.4	2.0
Floor	2.9	0.3	1.9
Ceiling	3.4	0.1	1.9
Left wall	4.4	0.1	2.3
Right wall	4.3	0.5	2.4

from localizing impulsive tones in each condition of this experiment.

We also determined the mean location judgment, statistic $\langle R(k) \rangle$, for each loudspeaker source. Values of $\langle R(k) \rangle$ are on a scale defined by the source numbers, 1 through 8. Figure 1 shows $\langle R(k) \rangle$ for each experimental condition. Perfect performance would be indicated by points lying along the 45-deg diagonal line, which equates true and perceived locations. Points above the diagonal indicate a response bias to the right of a source, points below the diagonal indicate a bias to the left.

1. Comparison with the ESPRO study

So that the results of the present study can be compared with those of paper I, $\langle \bar{D} \rangle$, $\langle \bar{E} \rangle$, and $\langle \bar{s} \rangle$ for three ESPRO conditions are given in Table III. Tables II and III show that there was substantial agreement in the outcomes of the two studies. In both, subjects' judgments of the location of impulsive tones: (a) were accurate to within an error of approximately 3.5° (statistic $\langle \bar{D} \rangle$); (b) systematically deviated to the left or right of sources by less than 1° ($\langle \bar{E} \rangle$); and (c) had a standard deviation of 2.5° or less ($\langle \bar{s} \rangle$). This close agreement between the two studies is notable given that the two testing environments were of very different scales. The ESPRO had a maximum volume of 4300 m^3, our anechoic chamber a volume of only 38 m^3.

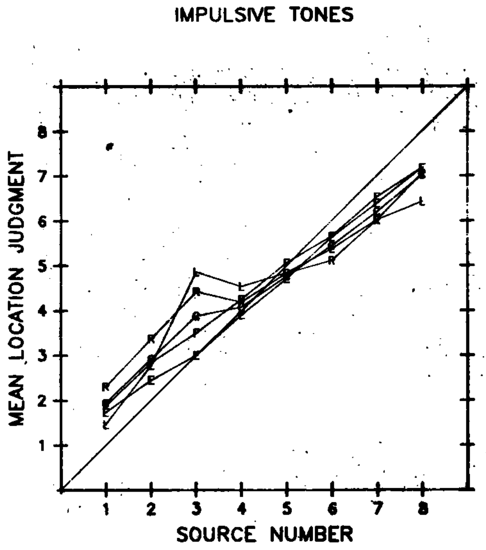

FIG. 1. Impulsive tones: Mean location judgment $\langle R(k) \rangle$ for each source depending on room condition. Symbol E = empty room condition, F = floor, C = ceiling, L = left wall, R = right wall.

TABLE III. Results for the localization of impulsive tones in several conditions of an experiment in paper I.

Room condition	$\langle \bar{D} \rangle$ deg	$\langle \bar{E} \rangle$ deg	$\langle \bar{s} \rangle$ deg
Absorbing walls, high ceiling	3.4	−0.9	2.4
Reflecting walls, high ceiling	3.3	−0.6	2.5
Reflecting walls, low ceiling	2.8	−0.7	2.2

2. Limitations of the precedence effect

In paper I we reported that changing the geometry of a room significantly affected listeners' ability to localize impulsive tones, and we took the result to suggest an imperfect precedence effect, as it operates in rooms. The present experiment provides stronger evidence of such an imperfection. The nature of the acoustical reflection of an impulsive tone was systematically manipulated here and we observed, as a consequence, a significant influence on the overall accuracy of localization. Statistic $\langle \bar{D} \rangle$ was significantly different across the five room conditions tested [$F(4,9) = 14.881$, $p < 0.001$].

3. Expectations and results regarding overall accuracy

Consistent with the results of paper I, we earlier proposed that acoustical reflections will either negatively or positively affect localization accuracy depending on whether they agree or disagree with a direct sound in terms of their azimuth. In the present experiment, the side-wall conditions presented cases of disagreement and the floor and ceiling conditions cases of agreement. We therefore predicted that subjects' overall accuracy would be worst with the side walls and best with the floor and ceiling. The first of these predictions was clearly borne out: values of $\langle \bar{D} \rangle$ for the left and right-wall conditions exceeded those of the other conditions by a degree or more (Table II), a highly significant difference ($t' = 7.58, p < 0.001$). The second prediction failed. Subjects' accuracy in the vertical reflection conditions (floor, ceiling) did not differ significantly from accuracy in the empty room.

4. Expectations and results regarding response biases

The hypothesis that the azimuth of a reflection will determine the nature of its influence also leads to predictions regarding response biases, as indexed by statistics $\langle \bar{E} \rangle$ and $\langle R(k) \rangle$. One expects that when the azimuth of a reflection competes with the azimuth of a direct sound subjects' responses will be biased in the direction of the reflection. Therefore, $\langle \bar{E} \rangle$ should be oppositely signed in the left and right-wall conditions, and, more generally, there should be a bias in the direction of a side wall for each of the sources. The results were not consistent with these expectations. Statistic $\langle \bar{E} \rangle$ did not differ significantly across the left- and right-wall conditions and was, in any case, positively signed for both. Moreover, subjects' responses were not uniformly biased in the direction of a side wall. This can be seen in Fig. 1, which shows the mean location judgments for each condition. The points for the left wall do not lie exclusively below the diagonal line as they should if all responses were biased to the left, and the points for the right wall do not lie exclusively above the line.

C. Alternative accounts of these results

It appears that the straightforward hypothesis that a reflection's azimuth will determine its impact on the azimuthal localization of impulsive tones is insufficient to account for our results. Therefore that hypothesis is rejected and, in this section, some alternatives are considered.

1. End effects, visual effects

At least a portion of subjects' response biases can be explained in terms of small end effects and somewhat more pronounced visual effects. Because the subjects' responses are constrained to sources 1 through 8 the source identification method includes a bias at the ends of the range. $E(k=1)$ is guaranteed to be non-negative, $E(k=8)$ must be nonpositive. Because $\langle \overline{D} \rangle$ is generally slightly greater than the angular spacing between the sources we expect that the bias is present for sources 1 and 8, but only weakly present for sources 2 and 7. One measure of the importance of the end effect is obtained by examining $\langle D(k) \rangle$, the rms error, as a function of source position. One expects that the end effects would lead to values of $\langle D(k) \rangle$ which are smaller for $k=1$ and $k=8$ than for other sources. Examination of $\langle D(k) \rangle$, for the five room conditions, shows exactly the opposite effect. The smallest values of $\langle D(k) \rangle$ occur for sources in the center of the range, and $\langle D(k) \rangle$ increases as k departs from the center. We attribute this result to the effects of visual bias. It is likely that the gaze of subjects, prior to receiving a tone, was in the forward direction, between sources 4 and 5. The effect of directed gaze is to bias the auditory judgments in the direction of the gaze. A bias of 2° was found by Weerts and Thurlow (1971) and by Hartmann (1983b). For the present experiment, therefore, a bias towards the center of the range of sources, which we attribute to directed gaze, is actually of greater significance than bias associated with end effects in the method.

The end effects and the visual effects exert their influences across all of our experimental conditions. *Differences* among conditions must have some other origin.

2. The temporal extent of the precedence effect—An initial consideration

Recent models of the precedence effect (Blauert, 1982; Lindemann, 1983) propose a neural inhibition process which prevents the processing of binaural difference following an onset. There are indications that this inhibition is quite general, affecting sensitivity to both interaural time and intensity, and affecting binaural discrimination judgments as well as localization judgments (Zurek, 1980; Gaskell, 1983). More important from our standpoint, however, are findings which suggest that there is some release from this binaural inhibition after approximately 10 ms and almost complete release within 50 ms (Zurek, 1980). Since our impulsive tones were 50 ms long, binaural processing should have recovered at some point during their course. In order to describe the expected consequences of this recovery for localization accuracy and response bias we need to know the nature of the ongoing sound field in each of our room conditions and also the perceptual responses associated with it. These are the topics of Secs. II and III. We therefore defer, until the end of Sec. III, further discussion of the temporal extent of the precedence effect as it applies to impulsive tones.

II. EXPERIMENT 2: LOCALIZATION OF TONES WITH A SLOW ONSET

Experiment 2 examined the effects of a single reflecting surface on localization of slow-onset tones of low frequency.

A. Method

The experimental protocol was the same as that of experiment 1, with three exceptions. (1) In this instance the 500-Hz sine tone was turned on with a 7-s onset and left on to deprive listeners of transient cues. (2) There were eight subjects instead of ten. All eight of those subjects also participated in experiment 1.[3] (3) In each condition, subjects judged only one block of trials—ten presentations per source.

B. Results and discussion

Table IV summarizes the results for several conditions of experiment 2. These data clearly slow that the localization of slow-onset tones was greatly affected by our manipulation of room reflections. Subjects' accuracy [$\langle \overline{D} \rangle$: $F(4,28\rangle = 38.86$, $p\langle 0.001$], systematic response biases [$\langle \overline{E} \rangle$:$F(4,28) = 19.78$, $p\langle 0.001$], and overall consistency [$\langle \overline{s} \rangle$: $F(4,28) = 8.43$, $p\langle 0.001$] were all significantly different across the five room conditions.

A result reported in paper I was that subjects found it extremely difficult to localize slow-onset tones in a room. That result was replicated in the present experiment. For three of the room conditions (ceiling, left wall, right wall) localization errors were as large or larger than the value that would result from random guessing, $\overline{D} = 9.7°$. The localization responses were by no means random, however. In each condition, the standard deviation of those responses, though larger than the corresponding value for impulsive tones (Table II), was still significantly smaller than would have resulted from guessing, $\overline{s} = 6.5°$. Here, as in the ESPRO, subjects developed consistent (but inappropriate) strategies for localizing slow-onset tones on the basis of the acoustical cues available. Sec. III of this paper describes our measurements of those cues and relates the measurements to the perceptual

TABLE IV. Results for the localization of slow-onset tones (experiment 2), depending on room condition.

Room condition	$\langle \overline{D} \rangle$ deg	$\langle \overline{E} \rangle$ deg	$\langle \overline{s} \rangle$ deg
Empty room	3.5	0.4	2.6
Floor	4.8	1.0	2.5
Ceiling	9.7	1.2	3.6
Left wall	9.7	− 3.6	4.4
Right wall	10.4	5.0	3.4

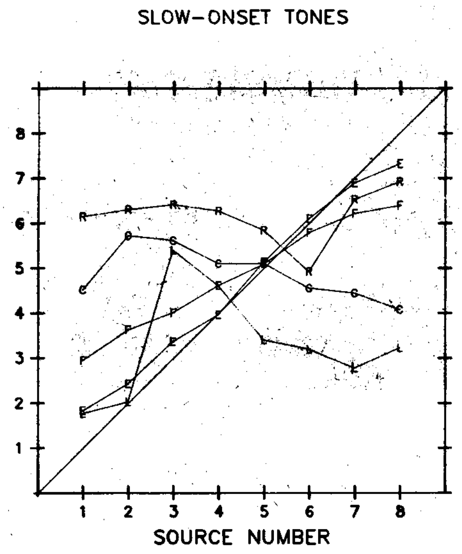

FIG. 2. Slow-onset tones: Mean location judgment $\langle R(k)\rangle$ for each source depending on room condition. Symbols are as in Fig. 1.

judgments. Here, we will be concerned with the judgments themselves.

1. The empty room

Figure 2 displays the mean location judgments, $\langle R(k)\rangle$, for each room condition. The values for the empty-room approach ideal performance—i.e., they lie very nearly atop the diagonal line equating true and perceived locations. This finding is of some interest given the nature of our stimuli. The rate of onset of the tones (7-s ramps) was sufficiently slow to deprive listeners of any transient onset cues. They were likewise deprived of transient offset cues because the tones remained on until the localization judgments had been made. Hence, the empty-room condition provides a test of the accuracy of azimuthal localization that can be achieved in free field when one relies solely upon the ongoing cues of interaural time and intensity difference. In this case, the standard deviation of all responses was $\langle \bar{s}\rangle = 2.6°$.[4] Appendix C of paper I shows that statistic $\bar{s}$ should be only about 5% greater than the minimum audible angle statistic introduced by Mills (1958). Our value of 2.6° is, however, larger than would be expected given the minimum audible angle of 1° measured by Mills. Part of the difference may be attributable to onset transient information in the Mills experiment, where onsets and offsets were only 70 ms in duration. It has been shown, at least for lateralization, that judgments are affected by onsets as long as 200 ms (Kunov and Abel, 1981).

2. Floor, side walls, and ceiling

Figure 2 shows that in the floor condition subjects exhibited a general bias to choose central locations. Source 1, for example, was heard to be positioned near location 3, and source 8 near location 6. With the wall on the left, there was much more dramatic deviation from the true location of sources: sources 5 through 8 were displaced far to their left, as demonstrated by the fact that their corresponding points lie well below the diagonal line, and source 3 was displaced far to its right.

In our testing room the right- and left-wall conditions were mirror images of each other, as reflected in the vertical median plane. Therefore, to compare the localization results for these two conditions it is appropriate to imagine reflecting the room and the perceived locations in this plane such that source 8 (and location 8) would correspond to what had been source 1 (and location 1), source 7 (location 7) would correspond to source 2 (location 2), and so on. Figure 2 shows that the right-wall function is, in essence, just such a reflection of the left-wall function. Hence, subjects appear to have followed similar strategies when localizing in these two complementary conditions. Lastly, we note that with the ceiling the perceived locations of all sources clustered toward the center of the room and the ordering of all but source 1 was *reversed*. The more the true location of a source was to the right, the more it was heard to lie to the left, as evidenced by the downward slope of the ceiling function.

3. Adequacy of the method

From Fig. 2 it is evident that the presence of a reflection can result in a large discrepancy between a subject's response and the azimuth of the source. There is no *a priori* reason to suppose that the source identification method, in which the allowed responses are limited to the sources themselves, provides an adequate range for responses. Evidence that the range is actually adequate appears in our data as follows:

a. Intrasubject variation. The largest value of the standard deviation, $\langle \bar{s}\rangle$ (left wall condition in Table IV) is 4.4°, corresponding to about 1.5 interspeaker spacings. We conclude that individuals can reproducibly associate a sound with a location near one of the allowed responses.

b. Intersubject variation. We computed the variance of the mean responses for a given source across all subjects, $\langle [R(k) - \langle \bar{R}(k)\rangle]^2\rangle$. The average of this variance over all sources provided a measure of the intersubject variation. The largest variation occurred in the right-wall condition, for which the standard deviation was 4.7°. We conclude that *different* subjects tend to associate a sound from a given source with the *same* allowed response.

c. Number of extreme responses. We expect that if a subject perceives an azimuth outside the range of allowed responses then he will give one of the extreme values, 1 or 8, as a response. A large number of extreme responses would be evidence that the range of allowed responses is inadequate. The largest numbers of extreme responses in our data occurred for left-wall and right-wall conditions, in which responses were strongly biased left and right. For the left-wall condition the largest number of responses was for location 1, but this number was only 14% larger than the number for location 2. For the right-wall condition the largest number of responses was for location 8, but this number was only 4% larger than the number for location 7.

In sum, although conditions in which responses deviate appreciably from the source locations would seem to push the source identification method beyond its limits of applicability, the data suggest that the method is still valid in these conditions.

III. INTERAURAL TIME AND INTENSITY DIFFERENCES

The slow-onset tones of experiment 2 were devoid of any transient cues. We therefore measured the ongoing interaural time difference (ITD) and the ongoing interaural intensity difference (IID) for each source in each room condition. These data are also relevant for a discussion of the precedence effect as it applies to impulsive tones.

A. Acoustical measurements

1. Method

In all conditions except the floor, we measured the interaural differences by means of binaural microphones (JVC HM-200E), mounted at the ears of a dummy head. To measure the ITD, the outputs of these microphones were amplified and displayed on a two-channel oscilloscope. We determined the number of microseconds of disparity (right ear minus left ear) between zero crossings of a sustained waveform. The IID was measured by comparing signal levels on a spectrum analyzer (HP-3580A) with a bandwidth of 30 Hz. In the floor condition, unlike the other conditions, we found that the presence of a body had significant impact upon our acoustical measurements. Therefore, one of the experimenters wore the microphones and sat quietly while the floor measurements were made.

2. Perceptual test of the method

Our binaural microphone system was of entertainment grade and we had reservations about the quality of the measurements made with it. Therefore, we tested the system in a perceptual experiment. One subject (the second author) wore the binaural microphones and sat in the chamber for a session of testing in the empty-room condition. He made no responses at this time. Instead, he sat quietly while outputs from the microphones were recorded on separate channels of a two-channel tape recorder (Revox A700). One block (80 trials) of slow-onset tones were recorded. These tones had the usual 7-s onset and a total duration of 15 s. Later the recording was played back to the subject through headphones while he sat in the chamber, and he made localization judgments of the usual sort. The value of his $\overline{D}$ obtained in this way was only 0.1° larger than when localizing actual sources in experiment 2. The value of $\overline{E}$ was shifted by less than 0.5°.[5] We concluded that the binaural microphone recordings were accurate enough to support appropriate localization judgments and were adequate for measurements.

B. Comparison of the acoustical measurements with the slow-onset tone perceptual judgments

1. The figure

Figure 3 displays both our acoustical measurements and the localization judgments from experiment 2. The left axis of the figure indexes the measured ITD in microseconds. Positive numbers indicate instances in which the signal reached a subject's right ear sooner than the left. These points should therefore cue a location somewhere to the subject's right. Negative numbers should, of course, have the opposite effect. For a 500-Hz tone, the extreme values on the graph, − 1000 and + 1000 μs, are in fact equivalent. Both correspond to a 180-deg phase difference at the two ears.

The IIDs are indexed at the right of the figure. The numbers on this scale have been chosen assuming a time-intensity trading ratio of 5 μs/dB, a choice which we will consider for the moment to be arbitrary. Positive values on this scale should cue locations to a subject's right, as they correspond to instances in which the signal at the right ear

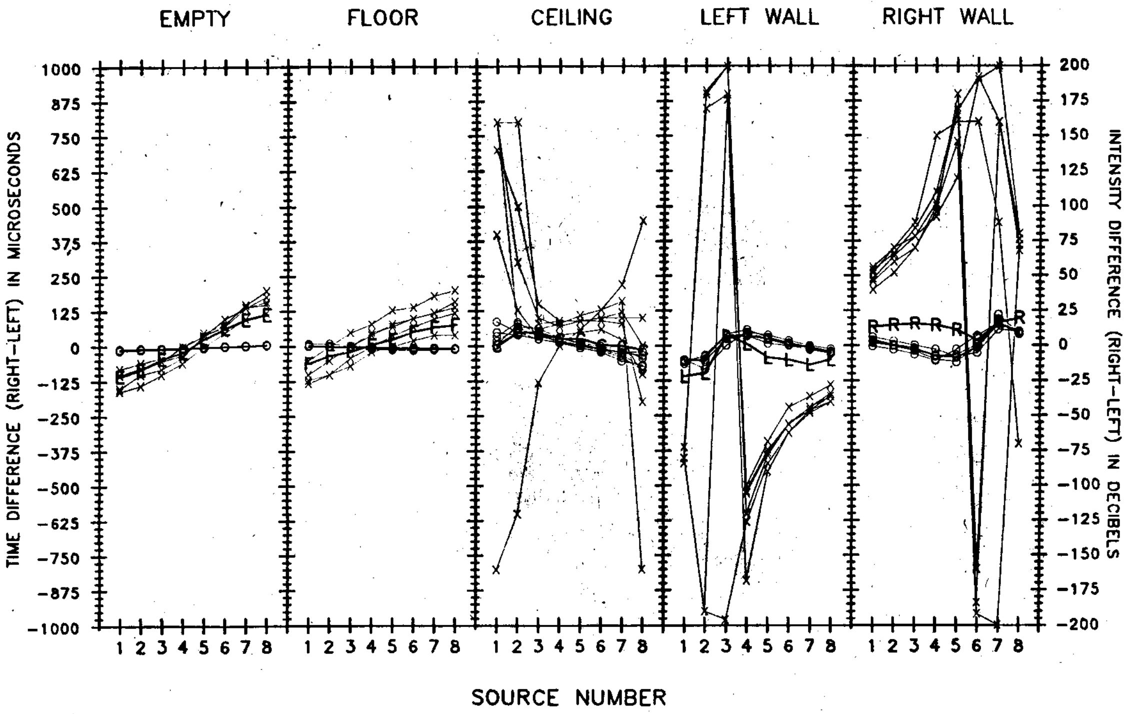

FIG. 3. Interaural time (symbol ×) and intensity (symbol O) differences as measured at six different points approximating subjects' head positions. Mean location judgments for slow-onset tones (Fig. 2) have been replotted here, on this scale, for comparison. Each panel represents a different room condition.

was more intense than the signal at the left.

We have replotted the localization functions from Fig. 2 on the ITD scale. To do so, we used the formula

$$\mathrm{ITD} = 3r \sin\langle\theta\rangle / c,$$

where r is the head radius (8.75 cm), c is the speed of sound, and $\langle\theta\rangle$ is the mean localization judgment from experiment 2 expressed as an angle with respect to the midline. Results presented in a review by Durlach and Colburn (1978) suggest that this formula should work very well at 500 Hz (see their Fig. 3).

The room conditions are represented in separate panels of the figure. In each panel, we have plotted the appropriate localization function and then overlaid it with a family of six functions representing ITD measurements (symbol ×) and six functions representing IID measurements (symbol ○). Each of the ITD and IID functions corresponds to measurements made from a slightly different point in the room, where those points were chosen to approximate different head positions of subjects seated in the testing chair.[6] At times the IID functions were so similar as to lie atop one another. In the empty-room, for example, six functions are perfectly superimposed.

2. Observations without regard for the trading ratio

Some comparisons between the localization judgments and the physical measurements do not depend upon our choice of the trading ratio between time and intensity. This pertains, for example, to the empty-room (panel 1) in which the IIDs were negligible, as expected for 500-Hz tones in free field. What varied in the empty-room were ITD values, and the figure shows that these varied in a way which was in close agreement with the perceptual judgments. This strongly suggests that in the empty-room localization was determined by the interaural time cue. The small discrepancy between localization judgments and the ITD functions for sources 1 and 2, and sources 7 and 8, might be attributed to end effects in the source-identification method.

The second panel of Fig. 3 shows that in the floor condition the ITD functions again varied across sources in a way which was consistent with the localization judgments. The IID functions varied in a way which was opposed to the localization judgments. Compared to the empty-room, both the perceptual function and the ITD functions are somewhat flattened.

The results for the ceiling, left-, and right-wall (shown in the three rightmost panels) contrast sharply with those for the empty-room and floor. First, in these three conditions the values of the ITD depended strongly upon the point of measurement. Small variations in the position of the dummy head produced large variations in the time cue. Second, with the ceiling and side walls we observed a number of extreme values of the ITD. These interaural differences often exceeded several hundred microseconds. Third, in these conditions there was a significant dependence of interaural intensity on source location. In contrast to our measurements of time, these measurements of intensity remained rather stable across the different positions of the dummy head.[7]

In the ceiling, left-, and right-wall conditions the overall slopes of the IID functions agreed with the slopes of the localization judgments. For example, with the ceiling, both the IID and the judgments showed a downward slope for sources 2 through 8. There is no such agreement between the time functions and the judgment data. We take this to be a strong indication that interaural intensity differences were the major determinant of subjects' localization judgments in the ceiling, left-, and right-wall conditions.

The conclusion that interaural intensity differences in low-frequency sine tones can determine localization is not altogether surprising. The data of Yost (1981) indicate that lateralization judgments depend upon IIDs in a way which is independent of frequency from 200 Hz to 5000 Hz. The dependence of localization on IID at low frequency has not been noted in the past because the experiments have been done in free field where such differences are not present. In a room, however, the interference between reflections and the direct sound results in sizable interaural intensity differences and these *do* affect the localization decision.

3. Time-intensity trade and the plausibility hypothesis

Although the measured IID appears to account, in a general way, for localization judgments in the left- and right-wall conditions, the agreement between the IID functions and the localization data is not perfect. We propose that this discrepancy is the result of a competition between the IID and the ITD. Calculations of this competition require a trading ratio. The scales of Fig. 3 were drawn with a trading ratio of 5 μs/dB. This value is smaller than most values reported in the literature, though it is similar to values found by Shaxby and Gage (1932) and by Moushegian and Jeffress (1959) for low-frequency tones such as ours. However, the numerical value of the trading ratio is, in our opinion, of little consequence because we reject the idea that the trading ratio can be described by a single number for our experiment. To deal with the time-intensity tradeoffs we suggest a plausibility hypothesis, as described below.

The plausibility hypothesis states that interaural time cues are weighted by listeners according to their plausibility. In our experiment the sound sources were always visible to the listeners and the visual image provided grounds for assessing plausibility of the ITD cues. The hypothesis assumes that listeners do not know how to compensate for room effects, even though they can see the room surfaces. Instead, listeners assess plausibility as though all sounds were direct.

We illustrate the operation of the hypothesis by considering the data for the left-wall condition in Fig. 3. The extreme sources, numbers 1 and 8, produce ITDs of $-140\,\mu$s and $+140\,\mu$s for the direct sound. ITDs outside this range are *to some extent* implausible. At a frequency of 500 Hz a source at 90° azimuth corresponds to an ITD of $\pm 760\,\mu$s. ITD cues outside this range are *completely* implausible and, we propose, are excluded from the localization computation. For sources 2 and 3 in the left-wall condition all of the ITD cues are in this completely implausible range. Therefore, localization should be entirely determined by IID cues, in agreement with Fig. 3. For sources 4, 5, 6, 7, and 8 the ITD cues become increasingly plausible. We propose that the contribution of the ITD cue to localization is given by a function, which increases with increasing plausibility, multiplied by the ITD itself.

With this hypothesis we can account for all the data in the left-wall condition in the following way. Let P be the perceived azimuth for a given source, expressed in units of μs, as in Fig. 3. The perceived azimuth is a linear combination of interaural intensity and time differences,

$$P = a(\mathrm{IID}) + b(\mathrm{ITD}),$$

where a and b are weighting factors. The time-intensity trading ratio is then a/b. For source 3 in the left-wall condition the ITD cue is completely implausible (Fig. 3 shows that the interaural phase difference is 180°) so that $b = 0$ for this source. The observed values of P and IID then determine $a = 5\ \mu\mathrm{s/dB}$, which is the scaling factor used in plotting the IID measurements in Fig. 3. Using this value of a, and the observed values of P and ITD for every other source, determines the trading ratio for each source. The trading ratios for sources 4, 5, 6, 7, and 8 are 61, 29, 23, 17, and 22 μs/dB, respectively. These ratios have a decreasing pattern. From the increasing plausibility of the ITD cue, as shown in Fig. 3, one would expect this decreasing pattern to be monotonic, and it almost is. The trading ratio for source 1 is 27 μs/dB. That value is slightly less than the ratio for source 5, as would be expected, because the ITDs for source 1 are slightly more plausible than those for source 5. An entirely similar analysis accounts for the data in the right-wall condition, with mirror reflection about the midline, i.e., source 1 exchanged with 8, 2 exchanged with 7, etc., and all functions reversed in sign.

C. Comparison of the acoustical measurements with the impulsive tone perceptual judgments: The precedence effect reconsidered

As noted in Sec. I, the impulsive tones of experiment 1 were sufficiently long that listeners very likely experienced some release from the binaural inhibition triggered by pulse onsets (Zurek, 1980). This may have contributed to the response biases shown in Fig. 1. Because experiment 1 included only a single early reflection, and no reverberation, we know that the sound field present during release was precisely the sound field heard with slow-onset tones and measured as described above. We therefore consider the possibility that the localization judgments in experiment 1 contain the ghostly image of the localization biases seen in experiment 2.

A first comparison of the two experiments makes this idea seem unlikely. Whereas right- and left-wall conditions resulted in large biases in $\langle \overline{E} \rangle$ for experiment 2, there was very little bias in $\langle \overline{E} \rangle$ for experiment 1. However, a more detailed examination reveals some similarities among the results of the two experiments. The following features are common to both: (1) Introducing either the floor or the ceiling flattened the function $\langle R(k) \rangle$ compared to the empty room (see Figs. 1 and 2). (2) In the right-wall condition sources 1, 2, and 3 were pulled to the right. (3) In the left-wall condition source 8 was pulled to the left. (4) The largest values of $\langle \overline{D} \rangle$ occurred for ceiling and side wall conditions. We conclude that the ghost of the slow-onset data is probably present in the impulsive-tone data.

Two situations are of special interest, source 3 in the left-wall condition and its mirror image, source 6 in the right-wall condition. These cases are unique in that the discrete reflection came from one direction whereas the ongoing sound field, as deduced from the localizations of the slow-onset tones, clearly suggested a source in the *opposite* direction. In these cases the localization of the impulsive tones *followed the ongoing sound field*, as can be seen in Fig. 1. This observation provides rather firm evidence that many of the failures of the precedence effect with impulsive tones were caused by a recovered sensitivity to the ongoing sound field.

But although the ghost of the slow-onset data is present in the impulsive-tone data, it is only a nebulous ghost. Its influence is sufficiently small that it was not seen in the overall statistic $\langle \overline{E} \rangle$. We conclude, therefore, that in our experiment subjects did experience some recovery of binaural sensitivity, but that the effect was much smaller than the complete recovery seen by Zurek (1980). We interpret the smallness of the effect upon recovery in this way. First we note that in Zurek's experiment a release from binaural inhibition was to the subject's advantage; in the case of localization in a room the release is clearly to the subject's disadvantage. We presume that the ongoing cues available to the subject upon release from inhibition are evaluated according to their plausibility. With large probability these ongoing cues in a room are implausible and are discounted. We conclude, therefore, that the precedence effect consists of two processes: binaural inhibition, and the process implied by the plausibility hypothesis.

IV. GENERAL DISCUSSION

A. The importance of onsets in free field

Although our work has been primarily concerned with the effects of reflections, the empty-room data permit us to comment on the importance of onsets for localization performance in free field. In the impulsive-tones experiment subjects received only a single brief exposure to a sound. In the slow-onsets experiment they could listen to a tone indefinitely before making a decision. Despite a greatly increased exposure to the tone in the slow-onsets experiment subjects' performance, as measured by both $\langle \overline{D} \rangle$ and $\langle \overline{s} \rangle$, was about half a degree poorer than performance in the impulsive-tones experiment. We conclude that the onset transient in the impulsive tones provided our listeners with important localization information.

The additional information provided by abrupt onset transients can be described in several ways. (1) The abrupt onset provides an envelope cue which is absent from the slow onset tone. (2) The abrupt onset provides a signal which is spectrally broadened somewhat and excites more neurons, if only briefly, than does a slow onset tone. Information therefore arrives through more channels for the abrupt onset. (3) The broadened spectrum of the abrupt onset tone can be modified by direction-dependent filtering by the pinna (Blauert, 1969/70; Wightman *et al.*, 1983).

Because it appears that onsets provide localization information, there is a methodological difficulty when the results of experiments using tones with abrupt onsets are analyzed in terms of models which involve only ongoing interaural time and intensity differences. It is clear that the *correct* stimuli to use to test models involving only ongoing cues are stimuli with slow onsets, though it is not clear that the onsets need to be as long as ours (7 s).

B. The plausibility hypothesis

The plausibility hypothesis introduced in Sec. III is not, we believe, a totally new idea. Its antecedents appear in other binaural studies where distinctions are made between compact and diffuse source images (Durlach and Colburn, 1978) or where multiple images are perceived (Hafter and Jeffress, 1968; Hafter and Carrier, 1971). In the context of a localization experiment in which the listener expects a compact image, anomalous ITD cues acquire the status of implausible and, according to the hypothesis, are discounted. The plausibility of cues may be of especial importance in an experiment such as ours in which the sources are visible.

The plausibility hypothesis suggests that in some circumstances it may be quite pointless to look for a single time-intensity trading ratio, as the trade should vary with the degree of plausibility. This may explain, in part, the disturbingly large range of trading ratios reported in the literature, from 2 μs/dB for low-frequency tones to 300 μs/dB for clicks.

The plausibility hypothesis may be used to explain another difficulty in the theory of localization, one regarding the relative importance of onset (envelope) cues and ongoing interaural differences. Studies in free field suggest that for low-frequency tones the onset cue may not be very important because localization performance does not change very much when it is removed. Perrott (1969) found no change when the onset was removed; our own data above show a change of only half a degree. Studies of localization in rooms, however, suggest that the onset cue is of overwhelming importance, and that ongoing cues matter little. A demonstration of this fact was given by Franssen (1961) in a two-speaker experiment within a room. He showed that if a tone is turned on abruptly in one speaker and then slowly faded out while the tone in the second speaker is slowly faded in, then the sound image remains at the first speaker. The plausibility hypothesis copes with this contradiction by noting that in free-field experiments the ongoing cues are plausible. By contrast, in a room the ongoing cues are, with high probability, implausible and are discounted by the listener in localization judgments. We note that this explanation predicts that in an anechoic room the Franssen illusion should fail.

The plausibility hypothesis as introduced here is, however, not totally satisfactory because it regards only the ITD as subject to a plausibility evaluation. No mention was made of the fact that, at 500 Hz, an IID of 10 dB (source 4, left-wall condition) is itself rather implausible. A more symmetrical and more satisfying hypothesis would subject both ITD and IID to a plausibility evaluation. Further work may clarify this matter. At this time we note only that the asymmetrical hypothesis seems to account rather well for our data, possibly because the IID cues were rather stable across the several positions that a subject might sample with a head turn, while ITD cues varied greatly across these positions (Fig. 3).

V. SUMMARY

Using the source identification method we have studied localization in a room which provided a single early reflection. Sources were impulsive tones and slow-onset tones. Our experiments led to the following conclusions:

(1) We observed an imperfect precedence effect, which did not absolutely exclude reflections from localization judgments.

(2) Lateral reflections were particularly detrimental to the azimuthal localization of impulsive tones.

(3) The precedence effect failures appear to be failures on a long time scale, suggesting recovery from binaural inhibition. Accordingly, the impulsive tone data included a ghost of the response biases seen with slow-onset tones.

(4) It is almost impossible to localize a tone in a room in the absence of transients because the ongoing sound field provides no reliable information concerning the source location.

(5) In a room, there are often large intensity differences due to the interference between a source and its reflection. These differences greatly influence localization decisions, even at low frequencies.

(6) Our results for the localization of slow-onset tones suggest the plausibility hypothesis. According to this hypothesis the weight given to the interaural time difference cue in a computation of localization depends upon its plausibility.

(7) The plausibility hypothesis is offered as an explanation for the smallness of the effect of the ongoing sound field on localization of impulsive tones. The plausibility hypothesis may also account for part of the large variation in the time-intensity trading ratio reported in the literature, and for the discrepancies in the reported relative importance of onset cues and ongoing cues for localization.

(8) In free field, our listeners localized impulsive tones somewhat more accurately and less variably than they localized slow-onset tones. Therefore, onset cues appear to be of some importance even in free field.

ACKNOWLEDGMENTS

Financial support for this research was provided by NIH grant NS17917 and by Michigan State University. We wish to thank the Michigan State University Department of Audiology and Speech Sciences for the use of their anechoic chamber.

---

[1] The precedence effect is a perceptual process which reduces the potentially disruptive influence of acoustical reflections. (See Gardner, 1968, for a historical review.)

[2] As in paper I, statistics $\langle \overline{D} \rangle$ and $\langle \overline{s} \rangle$ were averaged over subjects by computing the root-mean-square values of $\overline{D}$ and $\overline{s}$, respectively. Appendix B of paper I does not make this clear.

[3] For the eight subjects who judged both impulsive tones and slow-onset tones, the various conditions of experiments 1 and 2 were randomly inter-

persed, with a different random order for each subject.

[4] Standard deviation ranged from $\bar{s} = 1.5$ to 3.8 deg across subjects, and $\langle s(k) \rangle = 1.8$ to 3.1 deg across source locations. It did not increase systematically with increasing distance from the midline, probably because the maximum azimuth for any source was only 10.5 deg.

[5] In a second experiment both authors listened to the recorded signals while looking at numbers 1 through 8, separated by 3 degrees, drawn on a piece of paper taped on a wall outside the anechoic room. The rms error was $\langle \bar{D} \rangle = 2.88$ deg, only 0.31 deg greater than measured for these subjects in experiment 2. For recorded impulsive tones the rms error, for the authors looking at numbers, was $\langle \bar{D} \rangle = 3.27$ deg, which is 0.83 deg larger than in experiment 1.

[6] One set of measurements was made at the nominal listener's position, X, which was 10 ft from the loudspeakers, perfectly centered along their midline, and 46 in. above the wire grid floor of the chamber. (The average ear height of seated subjects was 46 in.) The five other positions were 3 in. above and beneath X, 3 in. to the left and right of X, and 3 in. in front of X. These positions span the range of ear heights for different subjects and correspond approximately to head positions which are accessible without movements of the trunk of the body.

[7] The considerable difference between the floor and ceiling conditions was initially surprising because these two conditions are identical so far as azimuth is concerned. Additional measurements showed that the difference was caused by the fact that the ceiling was 30 in. above the ears of the subject or of the dummy head, whereas the floor was 38 in. below the ears. When the ears were lowered to be 38 in. below the ceiling, the dummy head measurements of IID and ITD in the ceiling condition resembled those measured in the usual floor condition.

Blauert, J. (**1969/70**). "Sound localization in the median plane," Acustica **22**, 205–213.

Blauert, J. (**1971**). "Localization and the law of the first wave front in the median plane," J. Acoust. Soc. Am. **50**, 466–470.

Blauert, H. (**1982**). "Binaural localization: Multiple images and applications in rooms and electroacoustics," in *Localization of Sound: Theory and Applications*, edited by R. W. Gatehouse (Amphora, Groton, CT), pp. 65–84.

Durlach, N. I., and Colburn, H. S. (**1978**). "Binaural phenomena," in *Handbook of Perception*, Vol. 4, edited by E. Carterette (Academic, New York), pp. 365–466.

Franssen, N. V. (**1961**). "Eigenschaften das naturlichen Richtungshorens und ihre Anwendung auf die Stereophonie," in *Proceedings Third International Congress on Acoustics* (Elsevier, Amsterdam).

Gardner, M. B. (**1968**). "Historical background of the Haas and/or precedence effect," J. Acoust. Soc. Am. **43**, 1243–1248.

Gaskell, H. (**1983**). "The precedence effect," Hear Res. **11**, 277–303.

Haas, H. (**1951**). "Uber den Einfluss des Einfachechos auf die Horsamkeit von Sprache," Acustica **1**, 49–58: translated (**1972**). "The influence of a single echo on the audibility of speech," J. Audio. Eng. Soc. **20**, 146–159.

Hafter, E. R., and Carrier, S. C. (**1971**). "Binaural interaction in low-frequency stimuli: The inability to trade time and intensity completely," J. Acoust. Soc. Am. **51**, 1852–1862.

Hafter, E. R., and Jeffress, L. A. (**1968**). "Two-image lateralization of tones and clicks," J. Acoust. Soc. Am. **44**, 563–569.

Harris, G. C. (**1960**). "Binaural interactions of impulsive stimuli and pure tones," J. Acoust. Soc. Am. **32**, 685–692.

Hartmann, W. M. (**1983a**). "Localization of sound in rooms," J. Acoust. Soc. Am. **74**, 1380–1391.

Hartmann, W. M. (**1983b**). "Localization of sound in rooms—The effect of visual fixation," Proc. 11th ICA **3**, 139–142.

Kunov, H., and Abel, S. M. (**1981**). "Effects of rise/decay time on the lateralization of interaurally delayed 1-kHz tones," J. Acoust. Soc. Am. **69**, 769–773.

Lindemann, W. (**1983**). "The extension of binaural cross-correlation modeling by a mechanism of lateral inhibition," J. Acoust. Soc. Am. Suppl. 1 **74**, S85.

Mills, A. W. (**1958**). "On the minimum audible angle," J. Acoust. Soc. Am. **30**, 237–246.

Moushegian, C., and Jeffress, L. A. (**1959**). "Role of interaural time and intensity differences in the lateralization of low-frequency tones," J. Acoust. Soc. Am. **31**, 1441–1445.

Perrott, D. R. (**1969**). "Role of signal onset in sound localization," J. Acoust. Soc. Am. **45**, 436–445.

Shaxby, J. H., and Gage, F. H. (**1932**). "Studies in the localization of sound," Med. Res. Council (British) Spec. Rep., Ser. No. 116, 1–32.

Wallach, H., Newman, E. B., and Rosenzweig, M. R. (**1949**). "The precedence effect in sound localization," Am. J. Psychol. **62**, 315–336.

Weerts, T. C., and Thurlow, W. R. (**1971**). "The effect of eye position and expectation on sound localization," Percep. Psychophys. **9**, 35–39.

Wightman, F. L., Kistler, D., and Barker, N. (**1983**). "Localization in man: Importance of outer-ear structures," J. Acoust. Soc. Am. Suppl. 1 **73**, S1.

Yost, W. A. (**1981**). "Lateral position of sinusoids presented with interaural intensive and temporal differences," J. Acoust. Soc. Am. **70**, 397–409.

Zurek, P. M. (**1980**). "The precedence effect and its possible role in the avoidance of interaural ambiguities," J. Acoust. Soc. Am. **67**, 952–964.