Localization of sound in rooms

W. M. Hartmann[a)]
Department of Physics,[b)] Michigan State University, East Lansing, Michigan 48824 and Institut de Recherche et Coordination Acoustique/Musique, 31 rue Saint Merri, 75004, Paris, France

(Received 14 March 1983; accepted for publication 13 July 1983)

This paper is concerned with the localization of sources of sounds by human listeners in rooms. It presents the results of source-identification experiments designed to determine whether the ability to localize sound in a room depends upon the room acoustics, and how it depends upon the nature of the source signal. The experiments indicate that the localization of impulsive sounds, with strong attack transients, is *independent* of the room reverberation time, though it may depend upon the room geometry. For sounds without attack transients, localization improves monotonically with the spectral density of the source. Localization of continuous broadband noise *does* depend upon room reverberation time, and we propose the concept of direct signal to reverberant noise ratio to study that effect. Source identification experiments reveal certain localization biases, invisible to minimum-audible-angle experiments, and of uncertain origin. Appendices to this paper develop the statistics of the source-identification paradigm and show how they relate to the minimum audible angle.

PACS numbers: 43.66.Qp, 43.55.Br [JH]

INTRODUCTION

The ability to localize sounds in enclosed environments is an important function of the auditory system of humans and other animals. If, for example, one is sharing a dark cave with a sabre-toothed tiger, the ability to obtain information about the location of the tiger by listening to its growl is of considerable value. Nonetheless there are very few psychoacoustic experimental studies of our abilities to localize sounds in rooms. Instead, psychoacousticians have studied the localization of sound in free field, out of doors (Stevens and Newman, 1936), or in anechoic rooms (Mills, 1958), or have studied the lateralization of sound, the sense of sidedness that one can obtain from sounds delivered by headphones. (See Durlach and Colburn, 1978, for a review.)

From previous work one knows that localization is based upon interaural differences in intensity and spectrum, and upon interaural differences in arrival time of features of the direct sound waveform. The reflected sounds, which are potential sources of confusion in rooms, are presumed to be eliminated from the localization processing because they arrive later than the direct sound. The process which excludes room reflections from the brain's computation of location has come to be known as the precedence effect. (Wallach *et al.*, 1949) Though the precedence effect is specifically required to deal with the problem of localization in rooms, ironically, everything that we know about the effect has come from headphone experiments (e.g., Zurek, 1980) or from paired-loudspeaker experiments (e.g., Haas, 1951; Franssen, 1961).

The present paper is an experimental study of azimuthal localization of sound in rooms, with particular attention paid to the effects which different wall absorption and different room geometry have upon localization. The acoustical properties of the listening environment were controlled by performing all the experiments in a variable-acoustics concert hall, the Espace de Projection (ESPRO) at the Institut de Recherche et Coordination Acoustique/Musique (IRCAM) in Paris.

The ESPRO, shown in Fig. 1 is a medium-sized concert hall with a maximum volume of 4300 m^3 and an audience capacity of about 400. It is highly regarded as a recording environment for ensembles.

For the scientist, the ESPRO has a number of attractive features. The geometrical acoustics can be made simple because the ESPRO can be completely emptied leaving a nearly perfect rectangular box, with, if one so chooses, almost no diffusing surfaces. The wall acoustical properties and the geometry can be varied. The four walls and the ceiling consist of metal frames, which contain a total of 513 prisms with one absorbing surface, one specularly reflecting surface and one diffusing surface. The different surfaces are brought into play by rotating the prisms, in groups of three, by motors controlled from a control room. There are, in addition, 66 panels at eye level which can be rotated by hand into one of two positions, absorbing and specularly reflecting. In all,

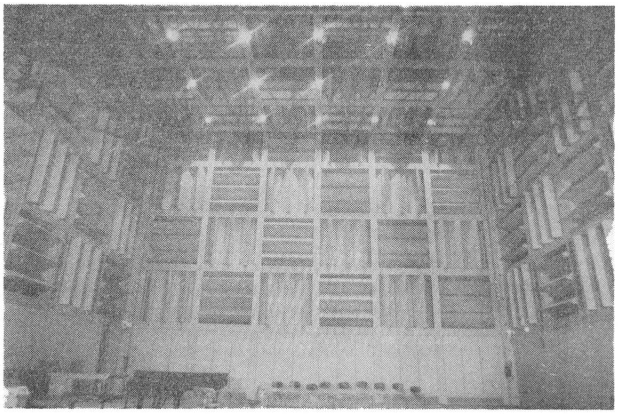

FIG. 1. View of the east end of the Espace de Projection at IRCAM. The ceiling is partly lowered and all prisms are in a diffusing condition, a condition which was not used in our work.

[a)] Work partially supported by the NIH, grant NS 17917.
[b)] Permanent address.

nearly 60% of the total area of the six interior surfaces (75% of wall and ceiling areas) can be varied. As a result the total room absorption in the intermediate frequency range can be varied by as much as a factor of 5. The ceiling of the ESPRO is formed from three 20-ton prism frames, which can be raised to 11.5 m or lowered to 3.5 m. Acoustical properties of the ESPRO have been discussed by Peutz (1981) and by Peutz and Bernfeld (1980).

I. EXPERIMENT 1: LOCALIZATION OF A PULSED TONE

The purpose of experiment 1 was to determine the ability of subjects to localize a pulsed sine tone. We were particularly interested to learn whether the acoustical properties of the room could affect performance.

A. Task

A single 50-ms pulse, rectangularly gated, of a 500-Hz sine tone was sent to one of eight numbered loudspeakers. The subjects' task was to declare which loudspeaker had sounded.

B. Method

The ESPRO was completely cleared, leaving four bare walls and a bare parquet floor. Planks and grills were removed from the catwalks to render them acoustically transparent at frequencies of interest. The catwalks were then raised to the ceiling.

The coordinates of listener and speakers were chosen to be simple to reproduce and to describe, but nonspecial. The coordinates of a point X on the floor were chosen to be the length and width of the room divided by $\sqrt{2}$, as shown in Fig. 2. A square, 1 m by 1 m was centered on point X. The subject was free to move at will within the square. From point X a center line, shown dashed in Fig. 2, was drawn to the far corner. Eight loudspeakers were placed symmetrically about this line, and on a circle of radius 12 m, centered on point X. Adjacent loudspeakers were 4° apart, and the speakers were 1 m above the floor. A vertical plane containing the center line divides the room volume equally. Therefore, reverberant sound arrived equally from the left- and right-hand sides of the speaker center line. Early reflections, by contrast, were not symmetrically distributed.

The loudspeakers were Philips type 544 amplifier–speaker combinations, with the automatic amplifier shutdown defeated. These speakers are small, with 17-cm-diam woofers; they presented an angular size of less than 1° of arc. The speakers had been matched for on-axis frequency response in an anechoic room. Speaker drive levels were chosen by using a white noise input and equating the outputs, as measured with a B and K A-weighted sound level meter at the position of the listener.

The actual level of the pulsed sine tone stimulus was chosen to be a comfortable listening level. The level was measured by the following procedure: The (unity gain) gate amplifier was bypassed to provide a continuous tone, and the room, in its maximum absorbing condition, was allowed to fill with sound. The sound level meter was moved around within the subject's square meter. The maximum reading

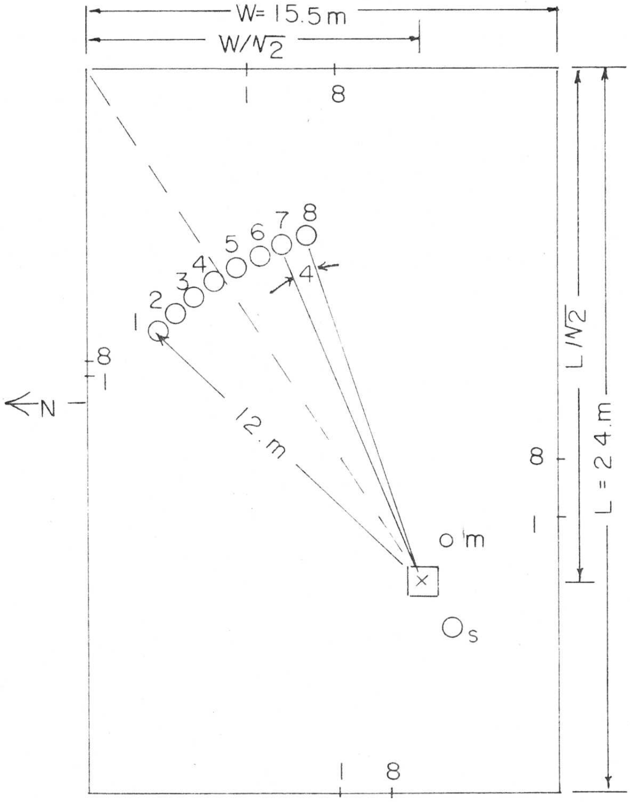

FIG. 2. Location of the eight target speakers, circles 1–8; the subject x; and talk-back microphone and speaker, m and s. Positions 1 and 8 on the walls indicate the origins of reflections, calculated by the image method for an ideal rectangular room. Reflections from all speakers appear, in numerical order, between positions 1 and 8.

obtained on the meter was 80 dB. Further details of the stimulus are described in Appendix A.

C. Subjects

Thirteen subjects participated in this experiment. The population included 11 males and 2 females, ranging in age from 8 to 42 years. The median age was 34. Most of the subjects were members of the technical and musical staffs at IRCAM and were experienced in listening to sounds and in manipulating them. Subjects averaged 12 years of formal musical study and regular performance. Subject 5 was left-handed, subjects 3 and 7 had been switched to right-handed in childhood. Two subjects, 9 and 12 were aware of some differential hearing loss at high frequencies.

D. Procedure

The experimenter was in the control room on the north side of the ESPRO, so that he could always observe the subject. The experimenter directed the pulsed sine to one of the eight loudspeakers and recorded the response of the subject, transmitted by microphone, "m" in Fig. 2. (The experimenter could talk to the subject through speaker "S.") Each experimental run consisted of 80 trials, ten presentations of the eight loudspeakers in random order. No feedback was given during the run, which lasted about 10 min. After the run the subject could examine the completed data form.

E. Room conditions

The localization experiment was performed in four room conditions as follows:

1. Absorbing room

The ceiling was at its maximum height of 11.5 m. All prisms were placed in the absorbing orientation. The log–log plot of reverberation time versus frequency reported by Peutz (1981) is mound shaped with RT-60 equal to 1 s at 500 and at 4000 Hz and equal to 1.3 s between 1000 and 2000 Hz.

2. Reflecting room

The ceiling was at its maximum height of 11.5 m. All prisms were placed in the specularly reflecting orientation. The reverberation time plot reported by Peutz is mound shaped with RT-60 equal to 4 s at 250 and 3000 Hz and equal to 5.5 s at 500 and 1000 Hz.

3. Low ceiling

All prisms were in the specularly reflecting orientation. The ceiling was lowered to a height of 3.65 m. According to reverberation time formulas of Sabine and Eyring the reverberation time for this configuration is exactly half that for the reflecting room condition above, e.g., 2.8 s at 500 Hz.

4. Mirror reversed

The ceiling and the prisms were placed identically to the absorbing room, but the positions of the listener and the eight loudspeakers were reflected in a vertical mirror plane passing through the north–south central axis of the room. The mirror-reversed condition is shown in Fig. 3.

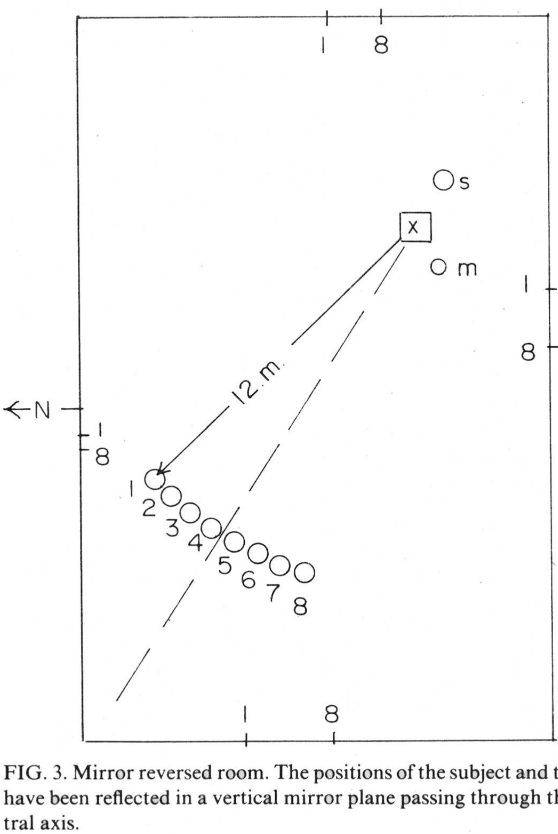

FIG. 3. Mirror reversed room. The positions of the subject and the speakers have been reflected in a vertical mirror plane passing through the N–S central axis.

The initial experiment compared performance in the reflecting room with performance in the absorbing room. Subjects 1–7 performed their first run in the reflecting room, their second run in the absorbing room and their third run in the room with the low ceiling. Subjects 8–13 performed their first run in the absorbing room, their second run in the reflecting room, and their third run in the room with the low ceiling. Because we suspected that performance might improve with time the fourth run was a repeat of the first. Data from the fourth run were used in two ways. They were added to the data set for averages across subjects, and they afforded an intrasubject test for improvement with experience.

F. Results

Data from individual runs were entered into a computer, which enabled us to compute numerous statistics. Definitions of the most important statistics used are given in Appendix B. Below we use the notation $\langle \cdots \rangle$ for quantities averaged across subjects.

1. Run rms error $\langle \overline{D} \rangle$

To quantify localization performance we computed the rms discrepancy between the azimuth of the loudspeaker and the subject's response, statistic D. The results for 13 subjects in three room conditions are shown in Fig. 4. The means of D across subjects, with their standard deviations are reflecting room $\langle \overline{D} \rangle = 3.3°$ (0.6) for 13 subjects and 20 runs, absorbing room $\langle \overline{D} \rangle = 3.4°$ (0.6) for 13 subjects and 19 runs, low ceiling $\langle \overline{D} \rangle = 2.8°$ (0.6) for 12 subjects and 12 runs. We regard D as the most meaningful single number to describe localization performance.

2. Run standard deviation $\langle \overline{s} \rangle$

The standard deviation $s(k)$ is the standard deviation of the subject's responses to a stimulus produced by speaker k.

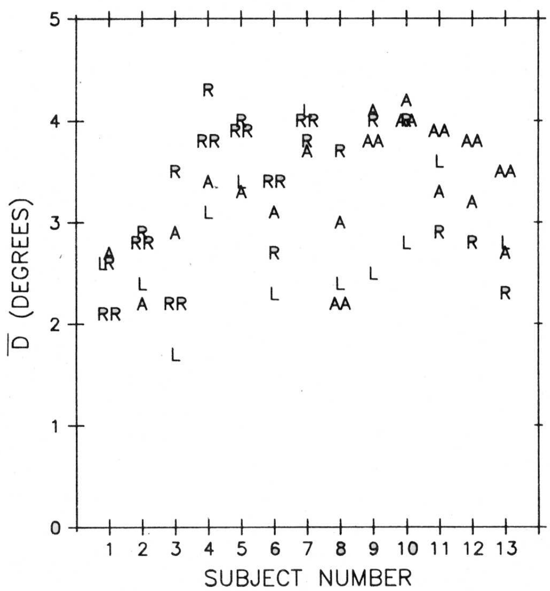

FIG. 4. The rms error, $\overline{D}$, for 13 subjects for localizing a 500 Hz, 50 ms sine tone in three room conditions; A absorbing, R reflecting, AA repeat absorbing, RR repeat reflecting, L low ceiling.

Note that it is computed with respect to the subject's mean response, not with respect to the correct answer. Statistic $\overline{s}$ is the average of $s(k)$ across source speakers. It is guaranteed to be less than or equal to $\overline{D}$. The means of $\overline{s}$ across subjects with standard deviations are reflecting room $\langle \overline{s} \rangle = 2.5°$ (0.5), absorbing room $\langle \overline{s} \rangle = 2.4°$ (0.4), low ceiling $\langle \overline{s} \rangle = 2.2°$ (0.6).

3. Mean error $\langle \overline{E} \rangle$

The mean error $\overline{E}$ is the average discrepancy between the subject's response and the correct loudspeaker azimuth. It can be positive or negative, indicating rightward or leftward biases, respectively. In a symmetrical experimental situation such as ours its expected value is zero. Averaged across subjects the mean errors, with standard deviations are reflecting room $\langle \overline{E} \rangle = -0.59°$ (1.13), absorbing room $\langle \overline{E} \rangle = -0.93°$ (1.21), low ceiling $\langle \overline{E} \rangle = -0.73°$ (0.89).

G. Discussion of experiment 1

1. The effect of room absorption

The rms error, $\langle \overline{D} \rangle$ for the reflecting room, is virtually identical to that for the absorbing room. To compare these two conditions further we performed a *t*-test. We considered the two groups of listeners, subjects 1–7 and subjects 8–13 and two pairs of runs 1st–2nd and 4th–2nd. The *t* scores for the differences of *D* values between reflecting and absorbing conditions were less than 1.0 for all of the four possible comparisons. This means that localization accuracy cannot be shown to depend upon the wall absorption at any reasonable level of confidence.

This result *does not* mean that localization is not affected by early reflections. The intensity of the early reflections in the two extreme room conditions differs by only a factor of 5, i.e., 7 dB. It is quite possible that early reflections play a role in localization but that the processing of early reflections is insensitive to a 7-dB change in reflection strength. Our result does suggest that localization is independent of reverberation over the entire range of practical reverberation times in concert halls.

2. The effect of room geometry

Although we were not able to change the intensities of early reflections by more than 7 dB we were able to reorder these early reflections dramatically by lowering the ceiling. Altering the room geometry in this way enabled us to test the operation of the precedence effect. We reasoned that if localization performance could be shown to depend upon the ordering of early reflections then that would indicate that the precedence effect, as it operates in rooms, is not absolute.

The results of the experiment with low ceiling clearly show that the ability to localize a tone is affected by reordering the reflections. The error $\langle \overline{D} \rangle$ was reduced by half a degree when the ceiling was lowered. We tested the hypothesis that $\overline{D}$ for the low ceiling is smaller than the average $\overline{D}$ for the high ceiling conditions. The *t* score was 3.0, indicating that the hypothesis is acceptable at the 0.01 confidence level; it is almost acceptable at the 0.005 level. We conclude that it is easier to localize a source with a low ceiling than with a high ceiling. This conclusion is in agreement with the informal remarks of subjects who participated in the experiment.

Our interpretation of this result is based upon the assumption that very early reflections are more likely to be confused with the direct sound and hence are more likely to result in a small breakdown in the precedence effect. We therefore consider the reflection delay times (as shown in Table I). In all our geometries the first reflection comes from the floor. The azimuth of this reflection agrees with the azimuth of the direct sound and may reinforce the perception of this azimuth. When the ceiling is *high* the next reflections come from the side walls. The azimuths of these reflections do not agree with the direct sound. The reflection from the high ceiling, which, like the floor, has the same azimuth as the direct sound, is greatly delayed. But when the ceiling is *low* the reflection from the ceiling precedes all the reflections from the side walls by a sizeable margin. Thus the low ceiling condition provides a second early reflection which agrees in azimuth with the direct sound and leads, we propose, to improved localization.

This argument, while compelling in its simplicity, was not *a priori* an obvious one. Reflections from the floor and ceiling actually provide less binaurally differentiated input than reflections from the side walls. One might have expected poorer localization performance in the case of the low ceiling. Indeed, Benade (1976) has proposed that lateral reflections play a positive role in the localization of sound. We note here only that our data are more economically explained if one assumes that *reflections from floor and ceiling reinforce the sense of localization while reflections from the side provide only confusion.*

Recent works in architectural acoustics stress the importance of early lateral reflections. (Schroeder *et al.*, 1974; Baron and Marshall, 1981). Preference judgments in simulated concert halls suggest that early reflections from the sides are much preferred to early reflections from above because the former provide a "stereophonic" effect. (Ando,

TABLE I. Early reflections: Angles and delay times (both with respect to the direct sound).

	Side walls			
	South		North	
Speaker	degrees	ms.	degrees	ms.
1	11.7	10.6	111.4	21.9
2	14.8	12.8	105.6	21.2
3	18.5	15.3	99.1	20.3
4	22.2	17.7	92.9	19.4
5	26.4	20.2	86.2	18.3
6	30.4	22.6	80.1	17.3
7	34.4	24.8	74.2	16.3
8	38.9	27.2	67.6	15.1
	Ceiling			
	High (11.5 m)		Low (3.65 m)	
all	63.2	33.4	24.4	2.3
	Floor			
all	9.3	0.9		

1977; Schroeder, 1978, 1980). It is argued for example, that acoustical clouds fail to provide a desired sense of intimacy because they produce reflections from the wrong direction. Instead, narrow halls with high ceilings are favored, supposedly because the first reflections come from the sides, thus resulting in the desired stereophonic effect.

It is not clear how one should interpret this concert-hall stereophony so far as localization is concerned. Stereophonic reproduction of sound by loudspeakers can sometimes produce a realistic sense of localization of the recorded source. On the other hand, stereophony provides confusion in that the source is not localized at its true azimuth, namely, at one or the other loudspeaker. Our experimental results and explanation suggest that the favored early lateral reflections in good concert halls actually delocalize the source. The preference for early lateral reflections most likely has nothing to do with localizability of sources but rather with the sense of surround which comes simply from interaural incoherence.

3. Average shifts

From the mean error, statistic $\overline{E}$, one can learn whether localization judgments are biased in some direction. Averaged across subjects, $\langle \overline{E} \rangle$ is negative for all three room conditions, i.e., incorrect judgments tend to fall to the left of the correct azimuth more than to the right.

The average shift of $-0.75°$ is smaller than the width of the distribution, a standard deviation of 1.2°. However, the leftward shifts are rather consistent. Of 12 subjects who performed four runs, seven of them exhibited leftward shifts exclusively, and one exhibited rightward shifts exclusively. The data for $\overline{E}$ are shown in Fig. 5.

The hypothesis that there is a leftward bias over the 50 runs shown in Fig. 5 results in a t score of 4.76, strong evidence in favor of the hypothesis. This bias was not expected given that, apart from the location of the speakers in the room, the experiment was methodologically symmetrical.

The bias can be seen in Fig. 6 showing $\langle E(k) \rangle$, the error as a function of source number, averaged over subjects and runs. For source $k = 1$ the error must logically be non-negative; for $k = 8$ it must be nonpositive. This end effect is probably present also in the data for sources two and seven; it is probably not present for sources three and six because errors of 8° are very uncommon. To understand the graph in Fig. 6 completely would seem to require end effects, some degree of central tendency, and a constant bias of approximately $-0.75°$ independent of k.

The bias may have an acoustical origin, essentially caused by the asymmetrical arrangement of the speakers in the room and the consequent asymmetry of the first reflections from the walls. The origins of early reflections, determined by the image method, are shown for sources 1 and 8 on the walls in Fig. 2. Reflections from the other sources originate from points in between points 1 and 8 in Fig. 2. Reflection delay times and angles, with respect to the direct sound are given in Table I. The first reflection for sources 1–4 comes from the north wall; the first reflection for sources 5–8 comes from the south wall.

A priori one does not expect these reflections to influence localization judgments. The shortest first-reflection delay (source 1 from the north wall) is 11 ms; the precedence effect should easily eliminate it from localization processing. But, as we have noted above, the precedence effect is not absolute. Apparently early reflections can lead to a minor breakdown in its operation. The earliest reflections from the north wall could lead to a small localization bias such as we observed. However, the first reflections from the south wall present a problem for this conjecture. One would have to argue further that the north wall reflections emerge from origins close to the original sources and are more easily con-

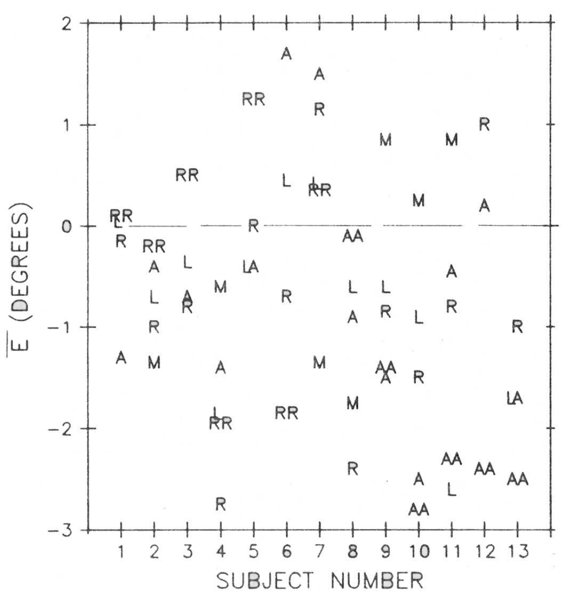

FIG. 5. Signed error $\overline{E}$ for 13 subjects for localizing a 500-Hz, 50-ms sine tone in three room conditions. Symbols are the same as for Fig. 4, and symbol "M" is for the mirror reversed room.

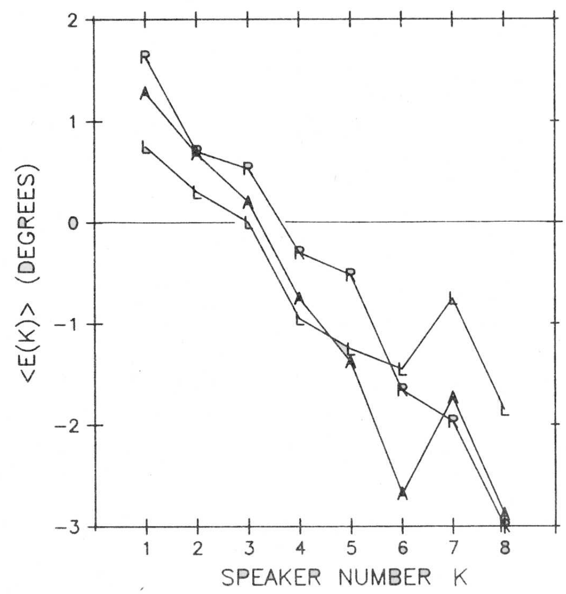

FIG. 6. Signed error $\langle E(k) \rangle$ averaged over 13 subjects given a 500-Hz, 50-ms sine tone in loudspeaker k for three room conditions: A absorbing, R reflecting, L low ceiling.

fused with the sources than are the south wall reflections with origins far from the sources.

Alternatively, the bias may have a more psychological origin, indicating some, as yet unknown, leftward bias for localization in the central processor. It may be significant that the localization data of Sandel *et al.* (1955), taken in an anechoic room, show a leftward bias at 500 Hz almost identical to ours. Possibly related is the finding of Jonquet and Pignon (1977) that localization accuracy is somewhat better in the left half-space than in the right.

To try to distinguish between these two possibilities we performed the localization experiment with the mirror reversed room. In this condition the acoustical situation was identical to the original absorbing room condition, so too was the visual aspect. The only difference was the "handedness" from the point of view of the listener. The listeners in the experiment were seven subjects who all showed a consistent bias in the previous four runs, subjects 2, 4, 7, 8, 9, 10, and 11. The results of the experiment, shown by symbols "M" in Fig. 5 were inconclusive. Subjects 2, 4, and 8 maintained the bias towards the near wall (i.e., they reversed the handedness bias), as expected if the acoustical explanation is correct. Subjects 9, 10, and 11 reversed their bias towards the near wall (maintained handedness bias), as expected if the psychological explanation is correct. The anomalous listener, subject 7, with a rightward bias maintained that handedness bias in the mirror reversed condition. The question of bias would seem to merit further study. However, the bias effects were small in our geometry. Sensitive experiments and perhaps a special geometry may be required to gain further information on this effect. Two points are worth mentioning. First, bias effects cannot be observed in minimum-audible-angle experiments, as introduced by Mills (1958); one must use an identification method such as ours. Second, there are strong visual effects on auditory localization (Warren, 1979). We have found it possible to introduce large bias effects by directing the gaze of listeners (Hartmann, 1983).

4. Learning effect

In many psychoacoustical tasks the subjects' performance improves with exposure to the task. We were interested to know whether this is true of the localization task. We calculated the rms error for the first eight judgments ever made by each of the subjects, statistic $\overline{D8}$. Seven of these were in the reflecting room, six were in the absorbing room. The average across subjects ($8\times 13 = 104$ judgments) was less than 0.1° larger than the average rms error for all absorbing and reflecting room judgments ($80\times 39 = 3120$ judgments.) The difference is insignificant, and we conclude that localization is a task that subjects can do right away. It is possible that extensive training in the task would reveal long-term improvement, but, unlike other tasks, signal detection for example, there is no short-term improvement.

II. EXPERIMENT 2: LOCALIZATION OF SOUNDS WITHOUT ATTACK TRANSIENTS

In contrast to the pulsed tones of experiment 1, the stimuli of experiment 2 had no onset transients. Whereas experiment 1 employed only one kind of stimulus with different room conditions, experiment 2 explored several different stimuli and, unless otherwise noted, the room was always the absorbing room described above. The experimental geometry was identical to that in experiment 1, shown in Fig. 2.

In experiment 2 the stimulus was turned on slowly, requiring 6 to 10 s to reach maximum intensity. The stimulus remained on until the subject gave his response. To make the experiment more efficient we allowed subjects to make their responses at any time during the stimulus rise. After the response was given the stimulus was turned off. Eight subjects, numbers 1, 2, 3, 4, 6, 7, 8, and 12 from experiment 1 participated in experiment 2, though only five of them listened to all the conditions. The subjects were seated, above the X and facing the far corner, in such a way that they were able to maintain the head in one position or to move it as they desired. In all other respects the experiment protocol was identical to that for experiment 1.

A. Localization of sine tones

It is a common experience that it is hard to localize a continuous low frequency sine tone in a room. Our first experiment provided a quantitative study of that difficulty. As for experiment 1 the sine tone frequency was 500 Hz. The maximum reading on a sound level meter moved about in the listeners' square meter was 70 dB. Eight subjects attempted the frustrating task of localizing the sine tone. Data are shown in Fig. 7. The rms error $\langle \overline{D} \rangle$ averaged across subjects, was 12.6° with a standard deviation of 1.9°. The number itself, however, is meaningless because $\overline{D}$ for random guessing is 12.9°. It seems likely that if the angular separation of the sources were increased the $\langle \overline{D} \rangle$ would increase as well.

But although their accuracy appears to be no better than chance, subjects tended to be consistent in their judgments. The standard deviation averaged across subjects was $\langle \overline{s} \rangle = 6.5°$ with a standard error of 0.9°. This can be com-

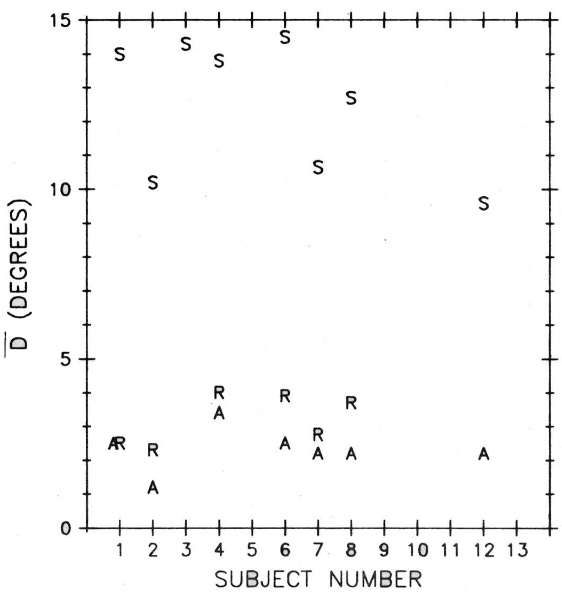

FIG. 7. The rms error, $\overline{D}$ for eight subjects for localizing sounds with slow attacks. *S* 500-Hz sine tones in the absorbing condition, *A* broadband noise in the absorbing condition, *R* broadband noise in the reflecting condition.

pared with a value of 8.7° for random guessing. No subject approached the chance value, indicating that each subject developed some strategy for identifying binaural cues with positions in space. But different listeners used different strategies. This is clear from examining the mean response for a given source, statistic $R(k)$. There are occasional similarities among listeners, e.g., six of eight listeners placed source number one to the right of the midline, but in general there are no common features to the identification data. The lack of correspondance may be caused by differences in acoustic cues, because head positions for the different subjects were somewhat different, or it may be caused by different mappings of the acoustic cues onto the set of eight response positions. Perhaps both differences were involved.

As a control we studied the localization of continuous sine tones with frequencies of 200 and of 5000 Hz, for five of the eight subjects. The results are shown, together with a repeat of the 500-Hz data, in Fig. 8. For all the subjects the error $\overline{D}$ was smaller for 200 than for 500 Hz, but, with the exception of subject 6, the difference was insignificant. At 5000 Hz, however, the performance was much improved. Apparently it was possible for all subjects to localize this high-frequency continuous tone to some extent, despite the fact that wall reflections, as indicated by the depth of standing-wave minima, were still strong. Further, the upper three data sets in Fig. 8, for the continuous sine tones, show considerable parallelism. This suggests that some subjects are better than others in localizing continuous tones, and that even at 500 Hz some subjects were not merely guessing. Subject 2, whose rms errors are consistently below those of other subjects, frequently moved his head during the presentation of a tone and took a long time to give his response. Possibly the multiple looks aided his performance. In sum, subjects

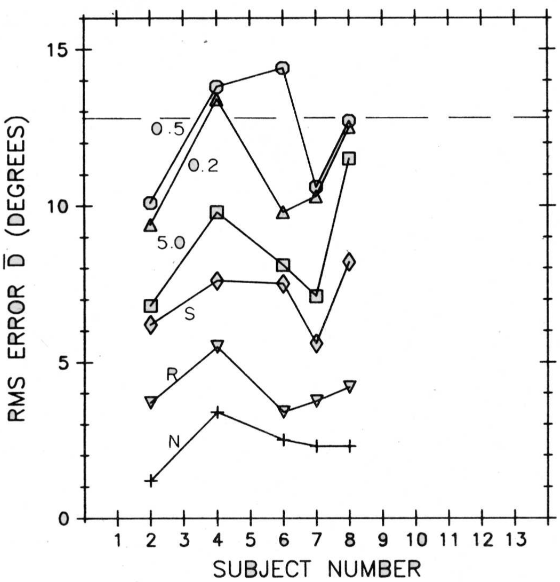

FIG. 8. The rms error $\overline{D}$ for five subjects for localizing sounds without attack transients in the absorbing room condition. Stimuli are "0.5" 500-Hz sine, "0.2" 200-Hz sine, "5.0" 5000-Hz sine, "S" spectrally sparse complex 200-Hz tone, "R" 10% rectangular 200-Hz tone, "N" broadband noise. Points are connected to aid the eye. The dashed horizontal line indicates random performance.

can gain localization information from sine tones in rooms, but they never gain much.

B. Localization of broadband noise

We studied the localization of broadband noise presented without attack transients. The experiment was identical to the sine localization experiments in the absorbing room except for the signal source. The broadband noise level was 70 dB at the listeners' square. The noise bandwidth was limited only by the loudspeaker response. (See Appendix A for details.)

1. Absorbing room

Seven subjects participated in the experiment. Their data are shown in Fig. 7. The rms error, averaged across subjects, $\langle \overline{D} \rangle$ was 2.3° (0.6), a full degree smaller than $\langle \overline{D} \rangle$ for pulsed sines in the absorbing room. Evidently the lack of an attack transient does not prevent successful localization of broadband noise.

2. Reflecting room

We repeated the noise localization experiment in the reflecting room. Six of the seven subjects from the absorbing room experiment participated. The rms error, averaged over subjects was $\langle \overline{D} \rangle = 3.22°$ (0.75), nearly a degree larger than $\langle \overline{D} \rangle$ for the absorbing room. All subjects performed less well in the reflecting room, except for subject 1 who performed equally well. The difference is significant. The statistic $\langle \overline{D}$ reflecting$-\overline{D}$ absorbing$\rangle$ has a value of 0.9°, and a standard deviation of only 0.5°.

3. Discussion

We think it reasonable to imagine continuous noise as a series of small impulses, random amplitude fluctuations which serve as transients that enable one to localize a sound by interaural time differences. Of course, the broadband noise does not sound impulsive, it sounds smooth. But the smoothness may result from temporal integration of loudness by the auditory system. By contrast the binaural system can identify interaural time differences of tens of microseconds, at least two orders of magnitude smaller than minimum auditory integration times. The binaural system then may be able to use temporal cues present in noise of which we are not otherwise aware.

We suppose that the small noise impulses in the direct sound from the loudspeakers provide the localization cues. The noise reflected from the walls of the room then appears as a masker. Therefore we expect poorer localization performance in the reverberant environment compared with the absorbing environment. Comparison of the results in Subsecs. 1 and 2 above shows that this is indeed the case.

Further evidence supporting this view came from observing the subjects' strategy during the experiment. Listeners found it easier to localize the noise source before the sound level had grown to its maximum. By making their judgments while the level was rising listeners improved the ratio of direct signal to reflected noise.

4. Signal-to-noise ratio

One can quantify the concept of localization as a masking experiment, by calculating a "signal-to-noise ratio." This can be regarded as simply the ratio of direct to reverberant noise. Let the effective absorbing area of the room surfaces be A, and suppose that the direct sound, at distance $r = 12$ m, is spread over an area G. Then the signal-to-noise ratio in dB is $10 \log (A/4G)$. The effective absorbing area can be calculated from the measured reverberation time T,

$$A = 0.161\ V/T,$$

where V is the volume of the room, 4278 m^3. The area for the direct sound is

$$G = 4\pi r^2/Q,$$

where Q is the directivity factor of the loudspeaker (Beranek, 1954).

The signal-to-noise ratio depends strongly upon frequency. Generally the A increases and the G decreases with increasing frequency so that the ratio increases faster than either of the two areas alone. To calculate a realistic signal-to-noise ratio requires that one know the frequency range that is appropriate for the hypothesized mechanism, localization by timing of micostructure. Recent experiments by Henning (1974) and by McFadden and Passanen (1976, 1978) suggest that useable timing information is present in both low and high frequency components. Arbitrarily then we calculate the signal-to-noise ratio at 1000 Hz. For the absorbing room the reverberation time is 1.3 s so that the effective absorbing area is 530 m^2. The directivity of the speaker, measured in an anechoic room is about 4.0. Therefore, the direct signal is spread over an area of 450 m^2 in the forward direction. From the two areas we find that the signal-to-noise ratio is about -5.7 dB. For the reflecting room the reverberation time at 1000 Hz is larger by about a factor of 4 so that the signal-to-noise ratio is smaller by 6 dB, i.e., it is -11.7 dB. From our data and this discussion we would conclude that an decrease of 6 dB in the ratio of direct signal to reverberant noise leads to a degradation in localization of at least one degree.

The application of the signal-to-noise ratio concept to the localization task needs to be questioned. First, it is a steady-state value, whereas subjects tended to make localization judgments before the reverberant noise had reached its steady-state value. Second, it is not *a priori* clear that a simple intensity ratio adequately describes the way that reflected noise impulses interfere with localization, though the critical-band effects observed by Canevet *et al.* (1980) suggest that the intensity ratio is indeed important.

The concept of signal-to-noise ratio needs to be put to experimental test. As mentioned above, the signal-to-noise ratio depends upon frequency through the directivity of the source and the absorption by the walls. If the signal-to-noise ratio is a valid concept for localization then this frequency dependence should show up in experiments on the localization of filtered noise bands.

C. Localization of a broadband complex signal

1. Spectrally dense complex tones

We studied the localization of a complex tone with fundamental frequency of 200 Hz with no attack transient. The waveform was rectangular with a duty factor of 10%. The spectrum of the tone, measured with a microphone in the subjects' square, averaged over the eight speakers had every 10th harmonic missing, as expected. The 15th harmonic was 15 dB down from the fundamental, the 25th and 35th were 20 dB down. The tone was slowly turned on to deprive subjects of onset information. Five subjects, numbers 2, 4, 6, 7, and 8 participated. The average rms error $\langle \overline{D} \rangle$ was 4.2° (0.8). Everyone of the subjects performed less well than with the broadband noise. To measure the difference we compared rms errors, $\overline{D}_R$ (for the rectangular pulse) and $\overline{D}_N$ (for the noise). The average $\langle \overline{D}_R - \overline{D}_N \rangle$ was 2.1° with a standard error of 0.4°. With the exception of one subject, performance with the rectangular tone was poorer than with the 500-Hz pulsed sine from experiment 1.

Because the rectangular tone is not impulsive like the broadband noise or the pulsed sine the poorer performance which we observed was expected. We did not expect however that performance with the rectangular tone would be as good as it turned out to be. The rms error of 4.2° is much smaller than the rms error for sine tones without attack transients.

One wonders what information is present in the 200-Hz rectangular tone and is not present in the sine tones that causes the former to be so much more localizable. The rectangular tone has many spectral components which may give the subject multiple cues, affording a statistical advantage. Alternatively one might note that critical bands above 1.4 kHz contain more than one harmonic of the rectangular tone. The output of a critical band filter therefore exhibits beats at 200 Hz, a form of envelope oscillation, which may allow the precedence effect to operate, even for the steady tone, as we believe it does for steady noise.

2. Spectrally sparse complex tone

To try to decide which of the above two explanations is correct we performed the localization experiment with a spectrally sparse complex tone. This tone had a fundamental frequency of 200 Hz and 11 harmonics, up to 5800 Hz. The harmonics were chosen such that none of them fell in the same critical band. There was a separation of at least 1.5 bark between any two harmonics. (See Appendix A for details.) As a result temporal cues associated with beats within a critical band should be absent in the case of the sparse complex tone. If these beats mediate localization of a complex tone then we would expect performance in localizing the sparse complex tone to fall to the level of performance for the 5000-Hz sine tone.

The same five listeners participated in the sparse-complex-tone experiment. Their results are shown in Fig. 8. Clearly localization accuracy decreased considerably, for all listeners, compared to the complex tone, with a dense spectrum. The average rms error was $\langle \overline{D} \rangle = 7.0°$ (1.0) to be compared with 4.2° (0.8) for the dense spectrum. We are inclined to conclude that the mixing of components within a single critical band plays a significant role in localization. However, for no subject did the performance with the sparse complex tone ever fall to the level of performance for the 5000-Hz sine tone. Possibly some intra-critical-band mixing was still present despite the 1.5-bark separation. Our conclusion then must be a tentative one. However, we can conclude in general that localization accuracy increases monotonically with increasing spectral density.

III. CONCLUSIONS

We have performed experiments on the localization of sounds in rooms. By using a variable acoustics room we were able to alter the room absorption and the room geometry, and we employed source signals of different types. Our major conclusions are as follows:

(1) The localization of brief impulsive tones is unchanged if the reverberation time is reduced from 5 to 1 s by adding absorption. These two reverberation times span the range of useful reverberation times in concert halls.

(2) The precedence effect, as it operates in rooms, does not absolutely exclude the effects of early reflections upon localization. The effects of early reflections can be modified by changes in the room geometry which reorder the sequence of reflections.

(3) Early reflections which come from the same direction as the direct sound reinforce the sense of localization of the source.

(4) The favored early lateral reflections in concert halls tend to delocalize sources.

(5) Localization judgments exhibit biases, which we were unable to explain in any single satisfactory way.

(6) There is no short-term learning effect in localization.

(7) It is almost impossible to localize a steady low frequency sine tone in a room, but not totally impossible. Localization is less impossible for high frequency sine tones. This is possibly the result of monaural localization processes.

(8) Localization performance for steady sounds improves monotonically with increasing spectral density of the source.

(9) A steady broadband noise in a dry environment is the most easily localized of all our sources. We explain the localization of a steady noise source by supposing that the binaural system regards noise to be a series of small impulses.

(10) The localization of steady noise can be significantly degraded by increasing reverberation. We introduced the notion of direct signal to reverberant noise ratio to deal with the latter effect.

(11) Our experimental procedure employed a source-identification method. This procedure is able to measure bias effects whereas the minimum-audible-angle method can not. The relationship between the source-identification method and the minimum-audible-angle method is discussed in Appendix C.

ACKNOWLEDGMENTS

I am grateful to René Caussé, Didier Arditi, and Daniel Raguin for help with the experimental equipment, and to Xu Yaying who performed the loudspeaker measurements in an anechoic room. The experiments reported here would have been impossible without the extraordinary assistance of Alexis Barsacq and the ESPRO team. Stephen McAdams, Yves Potard, Christine Hartmann, and Mitra Hartmann served as experimenters while the author was the subject. Dr. David Wessel, Dr. David M. Green, Dr. A. H. Benade, and Dr. M. R. Schroeder provided helpful discussions.

APPENDIX A: MEASUREMENTS OF STIMULI

To determine the temporal structure of the 500-Hz impulsive sine tone we measured the signal onset with an AKG 45 E microphone, 20 cm from the woofer cone; we stored it with a Gould OS 4020 digital dual-trace oscilloscope, triggered by the input. The onset appeared to be the sum of one and a half cycles of high-frequency resonant response, damped within 1 ms, and a growing driven response at 500 Hz. After 2 ms the driven response had reached its final amplitude and phase.

Further measurements were made of the power spectra of the stimuli. For these we used a Neuman U 87 microphone (omnidirectional configuration) placed near the position of the listener at a height of 1.8 m. Power spectra were obtained with a Spectral Dynamics SD 345 real-time spectrum analyzer.

1. Pulse

The spectrum of the pulsed sine wave, with frequency $f_0 = 500$ Hz and duration $T = 50$ ms was measured, in the absorbing condition, by averaging the power of six impulses from each of the eight speakers. The reader should note that averaging power, as we did, avoids the interference effects that occur if one averages microphone signals. The averaging time was 50 s.

The spectrum is shown in Fig. A1. In this figure the oscillations, of the form

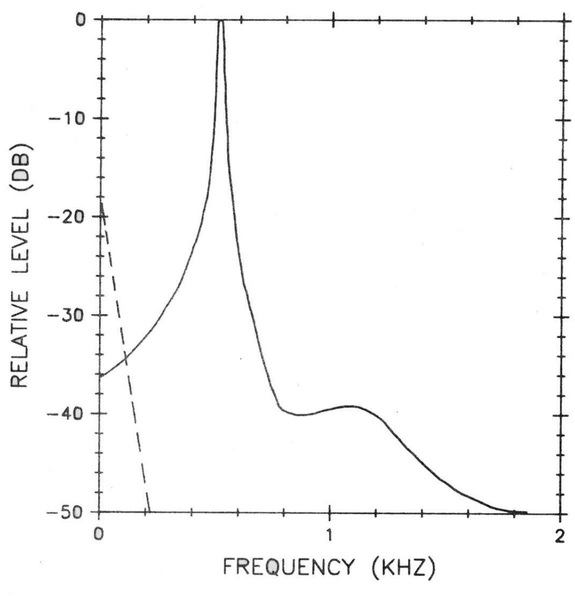

FIG. A1. Power spectrum of the 50-ms, 500-Hz sine pulse averaged over six pulses and eight speakers. The dashed line shows the spectrum of the background noise in the ESPRO, present during the experiments.

$\sin[\pi(f-f_0)T]/[(f-f_0)T]$

have been suppressed. These oscillations had a period of 40 Hz and a magnitude of about 6 dB, limited by the resolution of our analyzer. They contributed a fine-grained hash to the average curve shown in Fig. A1. The broad peak near 1100 Hz in Fig. A1 was caused by the room. It corresponds to the first nonzero peak of a comb filter with a delay time of about 0.9 ms, representing reflection from the floor. No other room effects are apparent; they were eliminated by the average over speakers. Figure A1 also shows the background noise spectrum.

2. Broadband signals

The spectrum of the broadband noise showed irregular oscillations of about 6 dB with a dominant period of 1100 Hz, corresponding to floor reflection. The average spectrum was flat to within 3 dB from 50 to 6000 Hz and dropped at a rate of -16 dB per octave above 6000 Hz. The average spectrum of the 200-Hz, 10% pulse train included all harmonics except for the 10th, 20th, etc. The harmonic amplitudes were as expected, given the spectrum of the original pulse train and given the filtering by the loudspeakers and the room as observed for the broadband noise.

3. Sparse complex tone

The spectrally sparse complex tone was generated digitally. It had a fundamental frequency of 200 Hz and ten other harmonics, all separated by at least 1.5 critical bands. As generated the harmonics were all of equal amplitude and all in sine phase. The measured harmonics are shown in Table AI together with critical band numbers [determined from Zwicker (1961)] and spectral levels in the room, averaged over the eight loudspeakers. There were two small distortion components observable, one at 10.2 kHz the other at 15.8 kHz. Both had a level of -40 dB.

TABLE AI. Components of the sparse complex tone.

Harmonic number	Frequency (Hz)	Critical band (Bark)	Relative level (dB)
1	200	2.0	0
2	400	4.1	3.6
3	600	5.7	2.9
4	800	7.2	2.1
6	1200	9.6	5.1
8	1600	11.5	1.8
10	2000	13.0	-0.8
13	2600	14.8	-3.2
17	3400	16.4	-0.4
22	4400	18.0	-3.3
29	5800	19.5	3.5

APPENDIX B: SOURCE IDENTIFICATION STATISTICS

This appendix defines the statistical quantities used in this paper.

1. Given quantities

Speaker angular separation	$A = 4°$,
Number of speakers	$L = 8$,
Number of trials per speaker	$M = 10$,
Number of trials per run	$N = ML = 80$,
Stimulus speaker for trial i	$S_i (= 1,2,\ldots,L)$,
Response for trial i	$R_i (= 1,2,\ldots,L)$,
Error for trial i	$e_i = R_i - S_i$.

2. Indicator

The indicator function $\Delta_i(k)$ has the value 1 if the stimulus speaker on trial i is speaker number k. Otherwise it is zero. The indicator is the Kronecker delta.

In a balanced experiment such as ours the normalization sums are

$$\sum_{i=1}^{N} \Delta_i(k) = M \quad (\text{for all } k),$$

$$\sum_{k=1}^{L} \Delta_i(k) = 1 \quad (\text{for all } i).$$

3. Derived quantities for a given speaker k

Error $$E(k) = A \sum_{i=1}^{N} \Delta_i(k) e_i / M,$$

rms error $$D(k) = A\left(\sum_{i=1}^{N} \Delta_i(k) e_i^2 / M\right)^{1/2},$$

Mean response $$\overline{R}(k) = \sum_{i=1}^{N} \Delta_i(k) R_i / M,$$

Standard deviation $$s(k) = A\left(\sum_{i=1}^{N} \Delta_i(k)[R_i - \overline{R}(k)]^2/M\right)^{1/2}.$$

4. Overall quantities for a given subject

Run error $$\overline{E} = \sum_{k=1}^{L} E(k)/L,$$

Run rms error $$\overline{D} = \left(\sum_{k=1}^{L} D^2(k)/L\right)^{1/2},$$

Run standard deviation $$\overline{s} = \left(\sum_{k=1}^{L} s^2(k)/L\right)^{1/2}.$$

Quantities averaged over subjects are indicated by $\langle \ldots \rangle$.

APPENDIX C: RELATING IDENTIFICATION ERRORS TO THE MINIMUM AUDIBLE ANGLE

Mills (1958) introduced a measure of localization which he called the minimum audible angle, MAA. Other authors, e.g., Perrot (1969) have used this measure as well. In an MAA experiment the subject hears a tone from a reference source at azimuth θ. He then hears a tone from a second source which is either to the left or to the right of the reference, by an angle $\pm\delta\theta$. The subject must declare whether the second source is to the right or the left of the reference. The task is thus a two-alternative forced choice. The MAA, $\Delta\theta$, is defined as that value of $\delta\theta$ where the subject's responses are 75% correct. (More precisely, the MAA is half the angle between 75% correct azimuth points to the right and to the left of the reference.)

In the present paper we study localization by a source identification experiment. Performance is measured by the rms error $\overline{D}$ and the standard deviation $\overline{s}$, expressed in degrees. The purpose of this appendix is to relate the MAA, $\Delta\theta$, to $\overline{D}$ and $\overline{s}$ using a simple perceptual model.

In fact, the identification errors $\overline{D}$ and $\overline{s}$ cannot be immediately related to the MAA because the identification errors include the effects of bias whereas the MAA does not. To find the identification errors one needs to introduce the bias separately. The bias enters the calculation early, and including it tends to make calculations rather special. We therefore assume, for the purposes of this appendix, that there is no bias. In that case rms error $\overline{D}$ and standard deviation $\overline{s}$ are the same. We formally solve here for $\overline{D}$. The following paragraph sketches our procedure.

We suppose that there is a single psychological dimension ψ which is the internal representation of azimuth θ. We suppose further that for given θ, ψ is normally distributed with standard deviation Σ. Because the relation between θ and ψ is invertable Σ can be expressed as a standard deviation in azimuth which we call σ. To begin our procedure we determine the width of the distribution σ from the minimum audible angle. Knowing the distribution we are able to calculate the probability density for ψ, given that the source is at azimuth θ. (Note that the assumption of no bias corresponds to the assumption that the most probable internal azimuth coincides with the source. If bias is to be included then it must be introduced here in the probability density.) The probability that the subject chooses source number k', given source k, in an identification experiment, $P_k(k')$, can then be found if one assumes that the subject chooses the source which is closest to the value of ψ. The set of values of $P_k(k')$ for a given k permit one to calculate the identification errors D and s for given source k, and ultimately to calculate $\overline{D}$ and $\overline{s}$, using the definitions in Appendix B.

We now follow in detail the procedure sketched above. From the psychometric function generated by a minimum audible angle experiment we learn d', the ratio of azimuth separations between the sources to the width of the distribution, σ. When the separation is at the minimum audible angle the percent correct is 75 and

$$d' = 0.95 = \Delta\theta/\sigma .$$

This simple formula is all that one needs from the MAA experiment.

In general, the MAA depends upon the azimuth of the source. Therefore, σ determined here, should depend upon source number k. This dependence creates no real problem for the mathematical steps which follow. However, Mills noted that for azimuths less than 20° or 30° the MAA is essentially constant. Because all the sources in our experiments fall into this range we can suppress the dependence of σ on source azimuth.

It is convenient to work in angular units given by the separation between the sources, assumed to be equally separated by A°. We therefore define the reduced width

$$W = \sigma/A = \Delta\theta/(0.95A) \approx \Delta\theta/A ,$$

and the reduced internal azimuth

$$X = \psi/A .$$

The normalized probability density for X, given source number k, is thus

$$f_k(X) = (2\pi W^2)^{-1/2} \exp[-(X-k)^2/2W^2] .$$

Because only integral numbered source positions are allowed responses we need to quantize the internal representation χ. The probability of a quantized internal response k' given source k, $P_k(k')$, is a function of the integer error $k'-k$, defined as δk. The probability of an error δk is then

$$P(\delta k) = (2\pi W^2)^{-1/2} \int_{\delta k - 1/2}^{\delta k + 1/2} \exp(-X^2/2W^2) dX .$$

The limits on the integral come from the assumption that the quantized representation is simply the number of the source whose azimuth is closest to the internal representation of azimuth.

The quantized perceptual error δk as described above, may correspond to perceptions which lie outside the azimuthal limits of the sources. The subject is not allowed to report source numbers that don't exist, and therefore the response error may be different from the quantized perceptual error. We deal with this problem by limiting the response error as follows. With a total of L sources, numbered 1 through L, and given that source number k is the correct choice, the largest possible response error is $L-k$, and the most negative possible response error is $1-k$. Values $L-k$ and $1-k$ are the limiting values of the range of allowed responses. Therefore, the response error, $K_k(\delta k)$, is given by δk when δk is within the range of allowed responses, and it is given by the appropriate limiting value when δk is outside the range.

The error D_k is then given by

$$D_k^2 = A^2 \sum_{\delta k = -\infty}^{\infty} P(\delta k) K_k^2(\delta k) .$$

In a balanced experiment, in which each source is presented an equal number of times, the overall rms error is

$$\overline{D} = A\left(\frac{1}{L}\sum_{k=1}^{L} D_k^2\right)^{1/2} .$$

The above relationship between the MAA and the errors of the identification experiment is reasonable so long as the normalized width of the distribution W is not too large. In the derivation of the relation we assumed that when the internal representation of a source azimuth lies outside the boundaries of the observed sources the subject responds with the appropriate extreme. But if the width of the distribution is large compared to the number of sources $\langle W > L \rangle$ much of

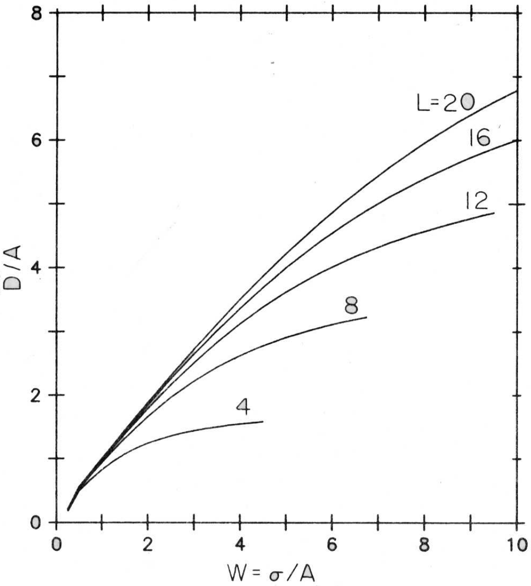

FIG. C1. The rms error $\overline{D}$, as a function of the width of the distribution σ, which is approximately equal to the minimum audible angle, according to the model of Appendix C. Both D and σ are normalized by the angular separation between sources A. Parameter L is the number of sources in the experiment. Curves for $L = 4$, 8, and 12 stop at the random-guessing limit.

the statistical weight lies outside the boundaries. The above assumption results in a large number of extreme responses and the resulting values of errors become larger than the errors which result from random guessing among the sources. This can be seen as follows: In the limit of very large W all responses are extreme responses and

$$\left(\frac{\overline{D}}{A}\right)^2 = \frac{1}{2L}\sum_{k=1}^{L}\left[(k-1)^2 + (L-k)^2\right] = \frac{(L-1)(2L-1)}{6}.$$

For random guessing,

$$\left(\frac{\overline{D}}{A}\right)^2 = \frac{1}{L^2}\sum_{k=1}^{L}\sum_{k'=1}^{L}(k'-k)^2 = \frac{L^2-1}{6}.$$

The first quantity is larger than the second so long as there are more than two sources.

The relation between $\overline{D}$ and the width of the distribution, W ($\approx$ the MAA), computed from the above, is shown in Fig. C1, for $L = 4$, 8, 12, 16, and 20 sources. The figure shows that when the width of the distribution is less than the angular separation between the sources ($W < 1$), as it is for most of our experiments, and when the number of sources is eight or greater, the values of $\overline{D}$ are essentially equal to the values of the width, i.e., about 5% larger than the values of the MAA. Although the calculation has not taken proper account of the bias observed experimentally, it does help us understand why our observed values of $\overline{D}$ and $\bar{s}$ are similar to values of the minimum audible angle as reported by others.

Ando, Y. (**1977**). "Subjective preference in relation to objective parameters of music sound fields with a single echo," J. Acoust. Soc. Am. **62**, 1436–1441.

Barron, M., and Marshall, A. H. (**1981**). "Spatial impression due to early lateral reflections in concert halls; the derivation of a physical measure," J. Sound Vib. **77**, 211–232.

Benade, A. H. (**1976**). *Fundamentals of Musical Acoustics* (Oxford U. P., New York).

Beranek, L. L. (**1954**). *Acoustics* (McGraw–Hill, New York), pp. 109 ff.

Canevet, G., Germain, R., and Scharf, B. (**1980**). "Localisation d'une information sonore en presence de bruit masquant," Acustica **46**, 95–99.

Durlach, N. I., and Colburn, H. S. (**1978**). "Binaural phenomina," in *Handbook of Perception*, 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).

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

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

Henning, G. B. (**1974**). "Detectability of interaural delay in high frequency complex waveforms," J. Acoust. Soc. Am. **55**, 84–90.

Jonquet, A., and Pignon, J. P. (**1977**). "Localisation tridimensionelle," in Proc. 9th ICA (Madrid), p. 45.

McFadden, D., and Pasanen, E. G. (**1976**). "Lateralization at high frequency based on interaural time differences," J. Acoust. Soc. Am. **59**, 634–639.

McFadden, D., and Pasanen, E. G. (**1978**). "Binaural detection at high frequency with time-delayed waveforms," J. Acoust. Soc. Am. **63**, 1120–1131.

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

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

Peutz, V. M. A. (**1981**). "Nouvelle examen des theories de reverberation," Revue d'Acoustique **57**, 99–109.

Peutz, V. M. A., and Bernfeld, B. (**1980**). "Variable acoustics of the IRCAM concert hall in Paris," Proc. 10th ICA (Sidney), Abstract E1.3.

Plomp, R. (**1977**). "Acoustical aspects of cocktail parties," Acustica **38**, 186–191.

Sandel, T. T., Teas, D. C., Feddersen, W. E., and Jefress, L. A. (**1955**). "Localization of sound from single and paired sources," J. Acoust. Soc. Am. **27**, 842–852.

Schroeder, M. R. (**1978**). "Binaural dissimilarity and optimum ceilings for concert halls; More lateral sound diffusion," J. Acoust. Soc. Am. **65**, 958–963.

Schroeder, M. R. (**1980**). "Toward better acoustics for concert halls," Phys. Today **33** (10), 24–30.

Schroeder, M. R., Gottlub, D., and Siebrasse, K. F. (**1974**). "Comparative study of European concert halls: Correlation of subjective preferences with geometrical and acoustic parameters," J. Acoust. Soc. Am. **56**, 1195–1201.

Stevens, S. S., and Newman, E. B. (**1936**). "The localization of actual sources of sound," Am. J. Psychol. **48**, 297–306.

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

Warren, D. H. (**1979**). "Spatial localization under conflict conditions: Is there a single explanation?," Perception **8**, 323–337.

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

Zwicker, E. (**1961**). "Subdivision of the audible frequency range into critical bands (Frequenzgruppen)," J. Acoust. Soc. Am. 33, 248.