A pointing technique for rapidly collecting localization responses in auditory research

ROBERT H. GILKEY
AL/CFBA, Wright-Patterson Air Force Base, Ohio
and Wright State University, Dayton, Ohio

MICHAEL D. GOOD
Wright State University, Dayton, Ohio

MARK A. ERICSON
AL/CFBA, Wright-Patterson Air Force Base, Ohio

JOHN BRINKMAN
Systems Research Laboratories, Inc., Dayton, Ohio

and

JOHN M. STEWART
Wright State University, Dayton, Ohio

A technique is described for rapidly collecting responses in auditory-localization experiments. Subjects indicate the perceived direction of the sound by pointing at a 20-cm-diam spherical model. In Experiment 1, the subjects judged the direction of a broadband signal, which could originate from any of 239 directions ranging through 360° of azimuth and between −45° and +90° of elevation. Using this technique, the subjects responded 2–8 times more rapidly than previous subjects who used either a verbal-reporting technique or a head-pointing technique. Localization judgments were as accurate as they had been with verbal reports, but were not as accurate as judgments collected using the head-pointing technique. In Experiment 2, the signal was turned off and the experimenter read the spherical coordinates of the signal location to the subjects. The subjects pointed to these coordinates more accurately than they had judged the direction of the sounds in Experiment 1, suggesting that the response technique had not been the limiting factor in that experiment. Circumstances relevant to the choice of response techniques for auditory-localization experiments are discussed.

There has been a recent resurgence of interest in the phenomena of human sound localization (e.g., Butler, 1987; Doll, Gerth, Engelman, & Folds, 1986; Makous & Middlebrooks, 1990; Middlebrooks & Green, 1991; Musicant & Butler, 1984; Oldfield & Parker, 1984a, 1984b; Wightman & Kistler, 1989a, 1989b). The need for further

research was sparked by the realization that traditional models of sound localization, based solely on interaural differences in stimulation, are inadequate to explain the observed phenomenology, such as vertical localization and the perceived externalization of sound sources. In addition, the development of three-dimensional virtual auditory displays (e.g., McKinley & Ericson, 1989; Wenzel, Wightman, & Foster, 1988) has created the need for a more detailed specification of the sound-localization abilities of human observers in various applied settings.

The effective rate of data collection for the field as a whole limits the speed and efficiency with which our understanding of sound localization can develop. In most experiments, the subject must report the perceived direction of the sound source on every trial. This type of response is complicated in comparison with the binary responses that are collected with most psychophysical procedures. With currently available techniques, localization judgments are collected at very slow rates. Wightman and Kistler (Kistler & Wightman, 1992; Wightman & Kistler, 1989b; Wightman & Kistler, 1992) required their subjects to report verbally the spherical coordinates

that corresponded to the perceived direction of the sound.[1] For example, a subject might call out the coordinates "minus 45, 30," indicating that the sound came from 45° to the left of the subject and from 30° above the horizontal plane. On each trial, the experimenter recorded the judgment by typing these coordinates into a computer. Using this technique, Wightman and Kistler (1989b) were able to collect 2–3 judgments per min. Makous and Middlebrooks (1990) used a head-pointing technique. In their "open-loop" condition, the subjects held their heads still during the sound presentation, and then turned their heads to point their noses in the direction from which they believed the sound had been presented. The position of the subject's head was monitored using a Polhemus Isotrak head-tracker. With this technique, Makous and Middlebrooks were able to collect approximately 3–4 localization judgments per min. The speed of trial presentation in these experiments was also limited by other factors, such as the stimulus duration and the need to move speakers between trials. Nevertheless, it is clear that these response techniques can be quite time consuming.

Because performance typically needs to be measured for each of a few hundred sound-source locations, more judgments are required from each subject under each experimental condition in a sound-localization experiment than would be the case in most psychophysical studies. Thus, considerable time and effort is required to complete even simple experiments. With breaks, somewhere between 150 and 300 responses can be expected in a 2-h session using techniques such as those employed by Wightman and Kistler (1989b) and Makous and Middlebrooks (1990). If, for example, the experimental design required 10 responses to be collected under a single condition from each of 225 speaker locations, data for four subjects could be collected in about two weeks, running 8 h per day (2 h per subject), 5 days per week. The duration of an experiment with a modest 3 × 3 factorial design would extend well beyond a single academic semester.

Our own research questions have often suggested experiments with factorial designs, and it therefore became critical to find procedures that would yield data at higher rates. It was not possible to run multiple subjects using our apparatus, nor was it financially feasible to reproduce the apparatus. Because we were uncertain a priori about the types of localization errors we would encounter, significantly reducing the number of speaker locations did not seem advisable. We therefore sought a procedure that would increase the number of responses that a single subject could produce per min. The God's eye localization pointing (GELP) technique reported here has proven to be quite effective. Using this technique, subjects can produce accurate responses at much higher rates than they could with previous procedures.

VALIDATION EXPERIMENTS

The GELP technique is quite simple in strategy: the subject indicates the perceived direction of the sound by

pointing at a 20-cm-diam spherical model of auditory space. Thus, rather than using an egocentric frame of reference to point from or to determine spherical coordinates, the subject responds using a "God's-eye" frame of reference. There was some concern that this change in frame of reference might be difficult for the subjects. However, the subjects report that the translation is quite easy and natural. Their data indicate that they can map spatial locations onto this model rapidly and accurately.

Two experiments were performed to validate the GELP technique. Experiment 1 was designed to determine its effectiveness as a response technique in a simple sound-localization experiment. On each trial, a broadband click train was presented from a randomly chosen location around the subject. The subject used the GELP technique to indicate the perceived direction of the sound. The results of this experiment were compared with those of similar experiments in the literature that had used different response techniques.

At least two possible sources of variability could influence the localization judgments from Experiment 1: (1) variability in the subject's perception of the stimulus location; and (2) variability in the motor response. The second experiment was designed to estimate the variability inherent to the motor response. On each trial of Experiment 2, the experimenter read a pair of spherical coordinates to the subject, whose task was to point to the location on the spherical model that corresponded to these coordinates. The results of Experiment 2 were compared with those of Experiment 1.

EXPERIMENT 1

Method
Subjects. Two males and one female served as subjects. Their ages ranged from 22 to 25 years. They all had clinically normal hearing and had participated in previous free-field listening experiments. They received several thousand practice trials using the GELP technique before data collection began. Subject M.G. is one of the authors.

Apparatus. The Auditory Localization Facility of the Armstrong Laboratory at Wright-Patterson Air Force Base is housed in a large anechoic chamber with 1.1-m fiberglass wedges. The internal dimensions of the chamber are approximately 6.7 × 6.7 × 6.7 m from wedge tip to wedge tip. The facility includes a 4.3-m-diam geodesic sphere with 272 Bose 4.5-in. Helical-Voice-Coil, full-range drivers (Model 118038) mounted at the vertices of the sphere. An additional five drivers are mounted on the surface of the sphere, one directly in front of the subject (0° azimuth, 0° elevation), one directly to the subject's left (−90° azimuth, 0° elevation), one directly to the subject's right (90° azimuth, 0° elevation), one directly above the subject (0° azimuth, 90° elevation), and one directly behind the subject (180° azimuth, 0° elevation). The metal frame of the sphere is covered with acoustic foam to reduce sound reflections. A simple visual display, consisting of 15 light-emitting diodes (LEDs), mounted on the surface of the sphere directly in front of the subject, is used to provide information to the subject about the trial timing.

Implementation. Figure 1 shows a subject in our apparatus. The subject is positioned with her head in the center of the geodesic sphere of 277 speakers, and straddles an 81 × 30-cm padded bench attached to an 87 × 60-cm platform. A bite bar is mounted on the end of a PVC tube that protrudes vertically through the bench. The height of the platform and the length of the PVC tube are both

Figure 1. A subject using the GELP technique.

adjustable, so that subject comfort is maintained, while assuring that the subject's head is positioned in the center of the sphere. The spherical model is attached to the top of a second PVC tube 22 cm in front of the first tube. The model is plastic and approximately 20 cm in diam. The center of the model is 31 cm above the bench and approximately (depending on the height of the subject) 50 cm below the subject's ears. The subject's task on each trial is to position the tip of a Polhemus stylus (Model 4A0310-0X)[2] at a point on the surface of the spherical model, such that a vector from the center of the sphere to the stylus tip points in the same direction as a vector from the center of the subject's head to the perceived location of the sound source. The stylus is connected to a Polhemus Fastrak (Model 4A0316-01). The transmitter (Model 3A0369-03) for the Fastrak is mounted on a wood block directly below the model. Houselights in the chamber are kept on during the experimental sessions, and the subjects are not blindfolded. However, because the subjects hold the bite bar during the experiment, their view of the model is somewhat limited. Therefore, three great circles are etched on the surface of the spherical model, corresponding to the horizontal, frontal, and median planes. The subject can feel these lines and use them as a guide. Thus, we expect that the subjects can, and possibly do, complete this task without vision. When the tip of the stylus is positioned at the desired location, the subject depresses a foot-switch to record his/her judgment.

The Fastrak is connected to the serial-communications port of an 80486-based personal computer. When the foot-switch is depressed, the software on the computer sends a query to the Fastrak, which returns coordinates indicating the current position and orientation of the stylus. The software ignores the orientation coordinates and translates the x-, y-, z-coordinates of the stylus tip into spherical coordinates relative to the center of the spherical model.[3] If the radius is greater than 10.4 cm, the tip of the stylus is deemed to be too far from the surface of the sphere (as if the subject had accidentally depressed the foot-switch at the wrong moment), and the judgment is discarded. Otherwise, the azimuth and elevation are stored as the subject's judgment. As currently implemented, this procedure does not allow the subject to reenter a discarded response or to correct an incorrectly entered response. Although neither of these problems seemed to occur frequently, the former problem could be corrected by merely waiting for a response within the 10.4-cm radius, and the latter problem could be solved by using a second foot-switch.

Procedure. On each trial, the signal location was randomly chosen from a pool of 239 possible speaker locations. The pool of speakers included azimuths that completely surrounded the subject, but elevations were limited to values between $-45°$ and $+90°$. The angular separation of any speaker from its nearest neighboring speaker ranged from 8° to 15°.

The target signal was a train of 25-μsec pulses repeated at a 100-Hz repetition rate. The signal was windowed with 25-msec $\cos^2$ ramps to have a duration of 268 msec at the half-power points, and was played through a 16-bit digital-to-analog converter (Tucker-Davis Technologies Model DD1). The analog waveform was then passed through a 20-kHz lowpass antialiasing filter (Tucker-Davis Technologies Model FT5-11) and bandpass filtered from .53 to 11.0 kHz (Krohn-Hite Model 3700).

Each trial began with a 100-msec warning light (a single red LED on the display), followed, after a 200-msec pause, by the observation interval, during which the signal was presented. To mark the observation interval, a row of four LEDs was illuminated for 300 msec, beginning 66 msec before the signal onset. After the observation interval, the subjects were allowed to take as long as they needed to indicate the perceived direction of the signal, using the GELP technique. A row of three yellow LEDs was illuminated for 150 msec to inform the subjects that their judgments had been recorded. No feedback was provided concerning the accuracy of this judgment. After a 200-msec pause, the next trial began.

During each block, 50-65 trials were presented. Data were collected in sets of five blocks, with a 1-2-min break between blocks within a set and a 10-min break between sets. Within each experimental session, 3-4 sets of blocks were presented. Each session lasted approximately 1.5-2 h.

Data analysis. During the course of the experiment, 8 localization judgments were recorded in response to signals from each speaker location. To allow comparison of our data with those of Wightman and Kistler (1989b), the judgment centroid, the average angle of error, and κ^{-1} were computed for each location. The judgment centroid is computed as the average of all of the judgment vectors for a specific target location, where the judgment vector is the vector from the center of the spherical model to the position indicated by the GELP technique. The average angle of error is computed for each target location and is the mean of the unsigned angles between the judgment vectors and the target vector (the vector from the center of the spherical model to the point on the surface corresponding to the location of the speaker). κ^{-1} is a measure of dispersion of the judgments and compares the magnitude of the vector that results from summing all of the judgment vectors for a particular target location with the maximum possible resultant. An unbiased estimate of κ^{-1} for the unit sphere is provided by

$$\kappa^{-1} = \frac{N(N-R)}{(N-1)^2},$$

where N is the number of judgments and R is the length of the resultant vector (Fisher, Lewis, & Embleton, 1987, pp. 129-130). In keeping with the procedures of Wightman and Kistler, all front-back reversals are resolved before these statistics are computed (i.e., if the target vector is in the frontal hemisphere and the judgment vector is in the rear hemisphere, the response azimuth is reflected through the frontal plane to the corresponding azimuth in the frontal hemisphere; the opposite operation is performed when the target vector is in the rear hemisphere and the judgment vector is in the frontal hemisphere).

We also computed horizontal and vertical errors in the manner employed by Makous and Middlebrooks (1990). The vertical error is simply the difference between the judgment and target elevations when represented as spherical coordinates. The horizontal error is the angle between the target vector and the vector from the center of the sphere to a point on the surface of the sphere whose longitude is equal to the judgment longitude, but whose latitude corresponds to the target latitude. In keeping with the procedures of Makous and Middlebrooks, judgments that represented front-back reversals were eliminated from the data set before these measures were computed.

Results

A major motivation for this project was the need to develop a technique that would allow subjects to produce localization judgments more rapidly. Wightman and Kistler (1989b) reported that 2-3 judgments could be collected per minute with the verbal-reporting technique. Makous and Middlebrooks (1990) were able to collect 3-4 judgments per minute using the head-pointing technique. Our subjects were able to produce 16-20 judgments per minute using the GELP technique, a substantial increase in speed.[4]

Figures 2 and 3 compare our data with those of Wightman and Kistler (1989b). On each trial, Wightman and Kistler presented eight 250-msec bursts of Gaussian noise separated by 300-msec periods of silence. The noise was bandpass filtered from .2-14.0 kHz. The energy spectrum of each noise sample was scrambled (i.e., the energy in each critical band was separately randomized over a 20-dB range). Wightman and Kistler presented stimuli from 72 speaker locations selected from a set of 144 possible locations, which ranged through 360° of azimuth and from -36° to +54° of elevation. Their subjects wore blindfolds during the experiment.

Figure 2 plots the azimuth component of the judgment centroid as a function of the actual azimuth of the target. The top three panels show data for our three subjects, and the bottom two panels show data for two of the subjects of Wightman and Kistler (1989b). The overall localization performance of Subject S.D.O. (bottom left panel) was as good as or better than that of most of Wightman and Kistler's subjects, whereas the overall localization performance of Subject S.D.E. (bottom right panel) was worse than that of most of their subjects. Because Wightman and Kistler only presented elevations between -36° and +54°, the data shown in the top three panels of this figure have also been limited to this range. Note that ideal performance is represented by the positive slope diagonal. Table 1 shows the slope and intercept of the lines that result from the regression of azimuth coordinates of the judgment centroids on the actual azimuths. The proportion of variance accounted for by the relation is also shown. The value of r^2 is greater than .98 for each subject. As can be seen, the performance of our subjects is good, and is similar to the performance obtained by Wightman and Kistler with the verbal-reporting technique.

Figure 3 plots the elevation component of the subjects' judgment centroids as a function of the actual elevation of the target. As in Figure 2, the top three panels show the results for our three subjects, and the bottom two panels show the results for the same two subjects from Wightman and Kistler (1989b). Values of slope, intercept, and r^2 are shown in Table 1. When compared with the azimuth judgments, the subjects' elevation judgments show more spread around the best-fitting line (values of r^2 for our subjects range from .75 to .83). Moreover, the slopes are all less than 1.0 deg/deg (slopes range from 0.67 to 0.77 deg/deg), and the intercepts are all greater than 0.0° (intercepts range from 2.1° to 11.3°). This pattern of re-
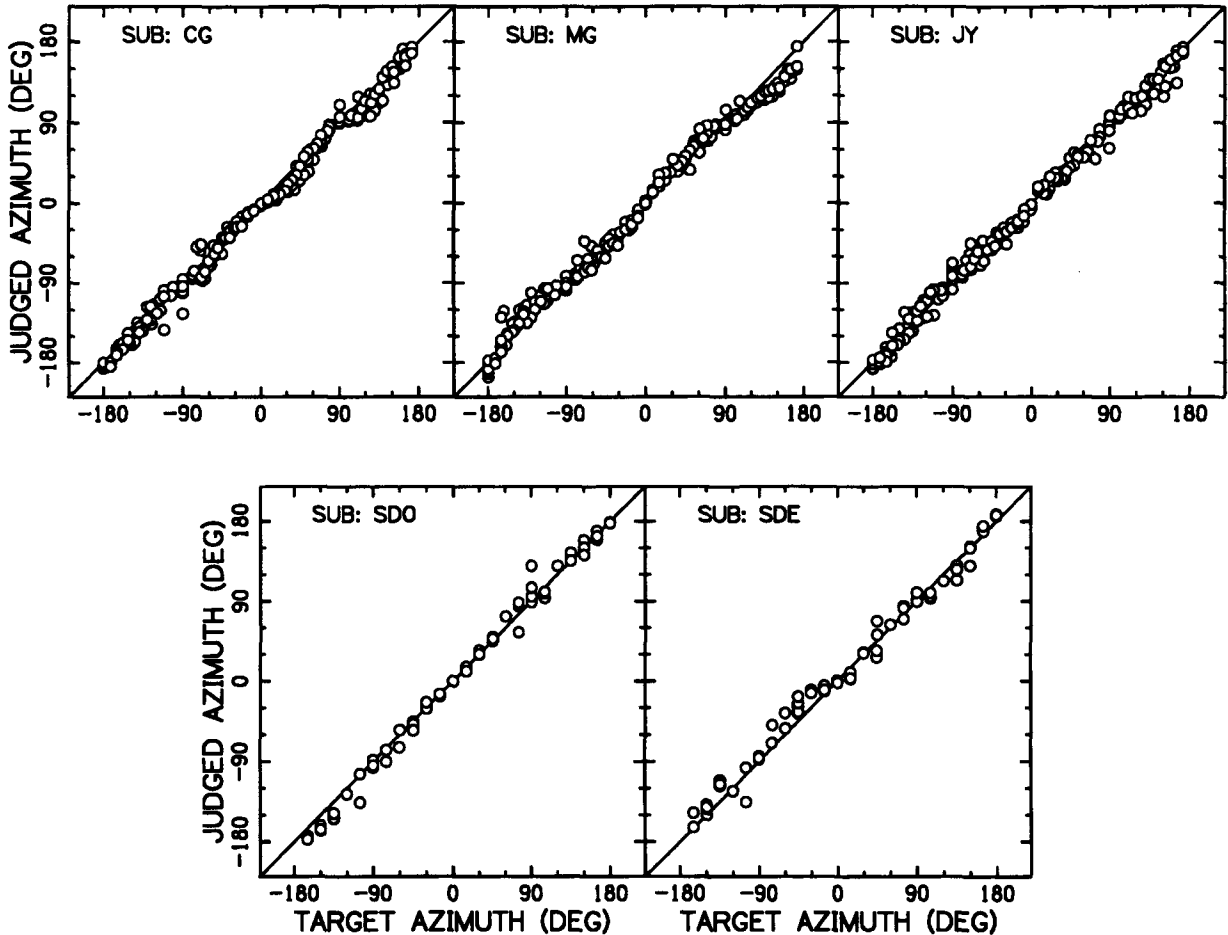
Figure 2. Judgment centroids for the sound-localization condition of Experiment 1. The azimuth coordinate of the judgment centroid for each target location is plotted as a function of the azimuth of the target. The top three panels show data for each of the three subjects in the present experiment. The bottom two panels show data for two of the subjects of Wightman and Kistler (1989b; the left panel shows data from a subject whose performance was better than, and the right panel shows data from a subject whose performance was worse than, that of most of their subjects). The centroids in the top panels are based on 8 judgments at each speaker location; those in the bottom panels are based on either 6 or 12 judgments at each speaker location. Front-back reversals have been resolved.

sults is similar to that observed for the subjects of Wightman and Kistler, except that their Subject S.D.O. shows a slope greater than 1.0. Note also that most of Wightman and Kistler's other subjects showed slopes closer to 1.0 than did Subject S.D.E. or any of our subjects.

Table 1 also shows median values (across speaker locations) of the average angle of error and κ^{-1}. The average angle of error is influenced both by the variability of the judgment vectors and by any systematic bias; κ^{-1}, on the other hand, is determined by the dispersion of the judgment vectors, and is less influenced by systematic bias in the judgments (in both cases, smaller values indicate more accurate performance). Median values of average angle of error and κ^{-1} observed for our subjects are comparable to, or smaller than, those observed for the subjects of Wightman and Kistler (1989b).

Makous and Middlebrooks (1990) used a head-pointing procedure, in which their subjects were to point their noses in the perceived direction of the sound source. On each

trial, two 150-msec bursts of reproducible noise were presented. An inverse filter was used to assure a flat amplitude spectrum for the noise at the output of each speaker, within the 1.8- to 16.0-kHz passband of the noise; a single pseudorandom phase spectrum was used for all presentations. The first noise was presented from directly in front of the subjects, at 0° elevation, and the subjects used this to "center" their heads. The second stimulus was the target; the subjects' task was to point their noses in the perceived direction of this sound. Makous and Middlebrooks presented sounds from 249 speaker locations that ranged in azimuth from −170° to + 170° and in elevation from −45° to +55°. Their experimental sessions were conducted in the dark.

Makous and Middlebrooks (1990) reported substantially smaller localization errors than Wightman and Kistler (1989b). Table 2 compares the localization errors reported by Makous and Middlebrooks with the localization errors observed in Experiment 1. For this analysis,
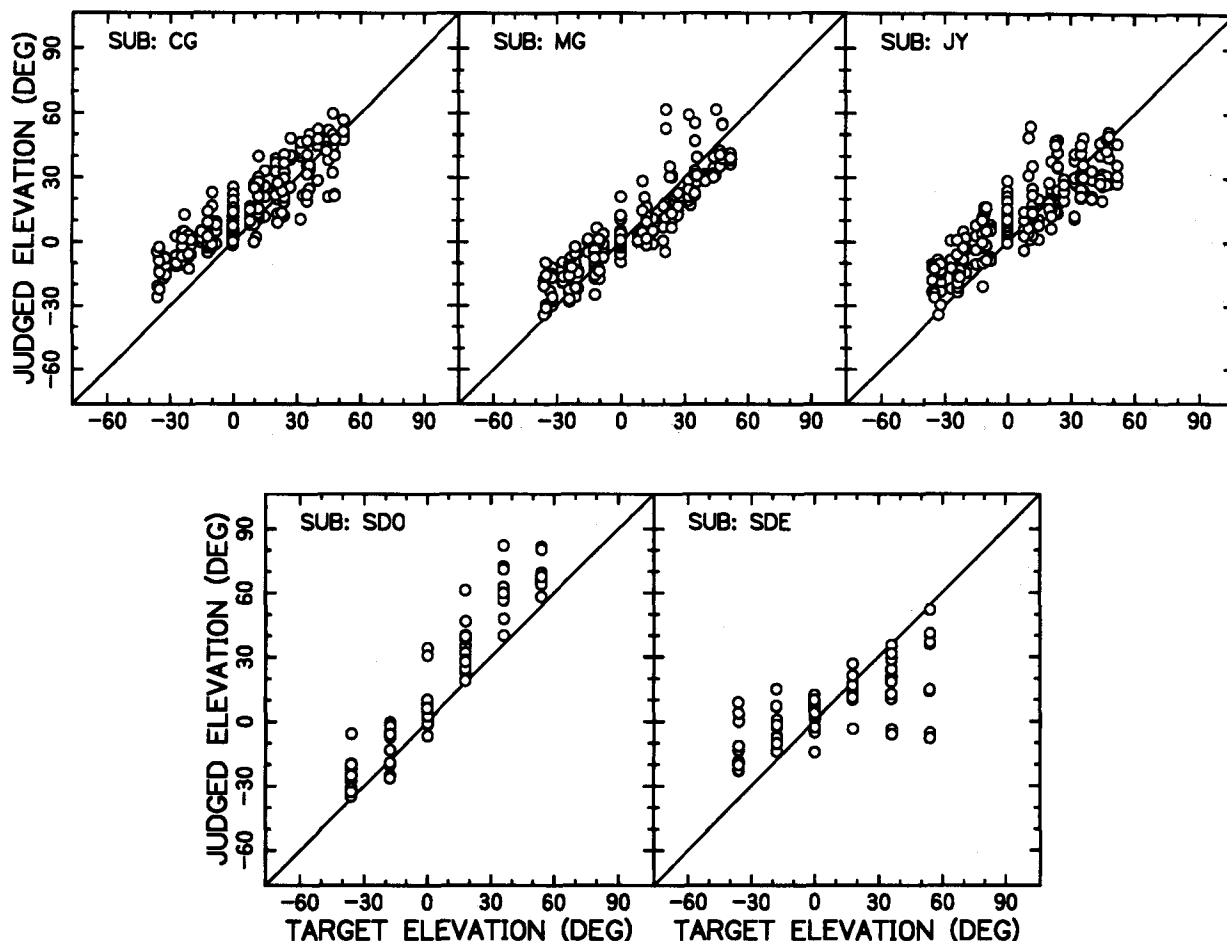
Figure 3. Judgment centroids for the sound-localization condition of Experiment 1. The elevation coordinate of the judgment centroid for each target location is plotted as a function of the target elevation. Note that, relative to Figure 2, the range of values on the axes has been reduced substantially. Other details are as in Figure 2.

horizontal and vertical errors were computed separately, following the procedures of Makous and Middlebrooks. The values shown are the unsigned errors between the judged position and the actual position. The errors are measured on each trial and averaged across trials and subjects. The Makous and Middlebrooks results are based on the data of six subjects, with five or six presentations at each speaker location. Our results are based on the data of three subjects, with eight presentations at each speaker location. These measures are influenced both by the variability of the subjects' responses and by any systematic bias (e.g., a tendency to respond too high). Although the unsigned horizontal errors are similar in magnitude across the two studies, the horizontal errors measured in the present study are generally somewhat larger (24 of 36 untied comparisons indicate larger horizontal errors in the present study), particularly for locations in the frontal hemisphere that are not on the median plane. The pattern for vertical errors is similar to that in the study of Makous and Middlebrooks, but in most cases, the vertical errors are substantially larger in the present study (35 of 38 comparisons show larger errors).

Discussion

The results of Experiment 1 and the comparison with the studies of Wightman and Kistler (1989b) and Makous and Middlebrooks (1990) indicate that the GELP technique is much faster than either the verbal-reporting technique or the head-pointing technique. The accuracy of our subjects' judgments collected with the GELP technique is similar to that of the judgments collected with the verbal-reporting technique by Wightman and Kistler.[5] However, both of these sets of data appear to be less accurate than the data of Makous and Middlebrook, collected with the head-pointing technique.

A strict comparison of the data from these three studies is probably inappropriate. A number of procedural differences that are not directly related to the response technique are likely to have affected the comparability of the results in these studies. There were differences in the auditory stimuli, the room acoustics, the number and range of possible source locations, the presence of visual stimulation, and the use of head restraints. Any of these factors could have influenced the perceived location of the stimulus. In terms of the present paper, the relevant ques-

Table 1
Slope, Intercept, and r^2 for the Best-Fitting Lines,
Together With Average Angle of Error and κ^{-1}

	Slope		Intercept		r^2			
Subjects	Azm	Elv	Azm	Elv	Azm	Elv	Error	κ^{-1}
Auditory Localization								
Experiment 1								
CG	.99	.67	−4.2	11.3	.992	.799	20.3	.032
MG	.95	.77	−0.9	2.1	.989	.825	15.5	.027
JY	.97	.67	−2.3	7.2	.993	.748	18.8	.045
Wightman and Kistler (1989b)								
SDO	1.05	1.15	−2.8	12.3	.994	.906	19.6	.034
SDE	.96	.39	4.5	4.6	.988	.499	22.3	.060
Verbal Coordinates								
Experiment 2								
CG	1.00	.83	−3.0	0.0	.997	.965	7.5	.006
MG	.97	.98	−1.8	1.1	.993	.959	9.2	.008
JY	.96	.73	−0.8	−0.4	.942	.938	10.7	.012

Note—Slope (degrees/degree), intercept (degrees), and r^2 for best-fitting lines obtained by the regression of the judgment centroids on the target locations are shown separately for the azimuth (Azm) and elevation (Elv) coordinates of the centroids. Median values are shown for the average angle of error (degrees) and κ^{-1}.

tion is whether the GELP technique can be used to report accurately the subjects' perception of the stimulus location, and not whether that perception accurately reflects the true stimulus location. For example, Butler (1987) and Middlebrooks (1992), using narrowband stimuli, have shown that the perceived location of a signal can be strongly influenced by the source spectrum. Recent results from our laboratory and from F. L. Wightman (personal communication, October 1993) indicate that *systematic* errors in the perceived elevation of broadband stimuli are quite common. Thus, it is possible that the recorded accuracy of the judgments of our subjects is limited by the accuracy of their perceptions, not by the accuracy of their motor responses. Experiment 2 was designed to evaluate this possibility further, by dramatically reducing the variability in the subjects' perception of stimulus location.

EXPERIMENT 2

Method

The subjects, apparatus, and procedure were the same as in Experiment 1, except that the signal channel was unplugged, so that instead of presenting the click-train signal, on each trial, the experimenter read the coordinates of the speaker location through the laboratory intercom. The subject's task was to point to the location on the surface of the spherical model corresponding to these coordinates. Four trials were presented for each speaker location.

Results and Discussion

Figure 4 shows the results when the azimuth and elevation of the judgment centroids are plotted as a function of the actual azimuth and elevation coordinates read by the experimenter. As can be seen, the subjects were able to point quite accurately at the location corresponding to the spherical coordinates read by the experimenter. Values of slope, intercept, and r^2 for the best-fitting straight line through the data are shown in Table 1. For

Subjects CG and MG, the values of r^2 observed for the azimuth data are *slightly* larger than those observed in Experiment 1. For Subject JY, the opposite is true. A comparison of the correlations, using the technique described by Steiger (1980), indicates that these small differences are significant (for Subject CG, $Z_1 = 6.6$, $p \ll .001$; for Subject MG, $Z_1 = 3.2$, $p \ll .01$; and for Subject JY, $Z_1 = -15.0$, $p \ll .001$). The difference between Experiments 1 and 2 is much more noticeable for the elevation component of the judgment centroids. For all three subjects, the slope of the best-fitting line is closer to 1.0 deg/deg (but is still less than 1.0), the intercept is closer to 0.0°, and the values of r^2 are significantly larger (for Subject CG, $Z_1 = 12.2$, $p \ll .001$; for Subject MG, $Z_1 = 10.2$, $p \ll .001$; and for Subject JY, $Z_1 = 9.8$, $p \ll .001$).

Table 1 also presents median values of the average angle of error and κ^{-1} under this condition. As can be seen, the median values of average angle of error for the three subjects are smaller by a factor of 1.6 or more than those observed in Experiment 1. A sign test indicates that these differences are significant [for Subject CG, $z(P = .92, N = 200) = -11.8$, $p \ll .001$; for Subject MG, $z(P = .83, N = 200) = -9.3$, $p \ll .001$; and for Subject JY, $z(P = .86, N = 200) = -10.1$, $p \ll .001$]. Similarly, the median values of κ^{-1} are smaller by a factor of 3.3 or more under the verbal-coordinate condition. Again, a sign test shows this difference to be significant [for Subject CG, $z(P = .88, N = 200) = -10.5$, $p \ll .001$; for Subject MG, $z(P = .85, N = 200) = -9.7$, $p \ll .001$; and for Subject JY, $z(P = .83, N = 200) = -9.3$, $p \ll .001$].

Horizontal and vertical errors were computed from these data and are shown in Table 2. As can be seen, in general, the horizontal errors in Experiment 2 are smaller than those in Experiment 1 (28 of 38 comparisons show smaller errors in Experiment 2), and they are

Table 2
Comparison Between Unsigned Horizontal and Vertical Errors (in Degrees) in the "Open-Loop" Condition
of Makous and Middlebrooks (1991; M&M) and Those in Present Experiments 1 (Ex1) and 2 (Ex 2)

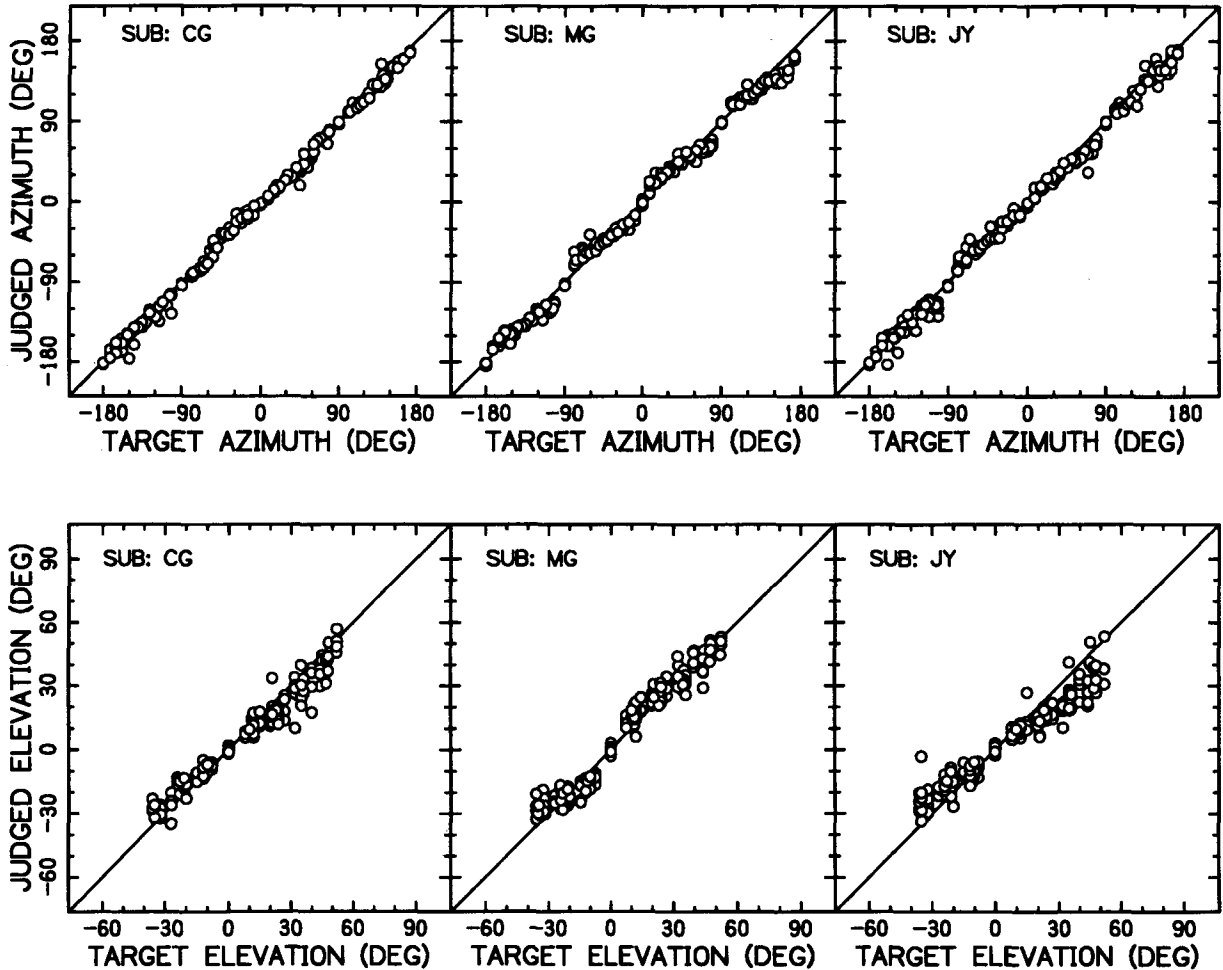|	Azimuth																										
	0°			20°			40°			60°			80°			100°			120°			140°			160°		
Elevation	M&M	Ex1	Ex2	M&M	Ex1	Ex2	M&M	Ex1	Ex2	M&M	Ex1	Ex2	M&M	Ex1	Ex2	M&M	Ex1	Ex2	M&M	Ex1	Ex2	M&M	Ex1	Ex2	M&M	Ex1	Ex2
								Horizontal Error (deg)																		
45°	3.5	2.5	3.2	6.0	8.0	4.3	8.7	12.3	4.8	8.1	11.8	9.5							8.1	9.8	6.3	15.9	19.1	5.6	16.3	13.5	4.1
25°	2.4	1.8	3.0	5.0	6.7	6.8	7.4	13.8	7.1	9.7	11.5	7.0							10.3	15.9	6.6	12.0	16.2	7.9	14.5	13.8	9.4
5°	2.2	2.2	2.4	3.8	7.6	5.4	6.1	8.2	6.3	9.4	11.5	7.0	5.2	5.6	5.0	6.2	7.3	13.0	10.4	14.7	6.2	13.0	14.1	8.7	15.6	13.0	11.9
−5°	1.5	2.4	2.1	3.7	7.5	4.5	5.5	8.5	4.9	7.1	8.6	12.3	5.5	7.2	5.2	7.4	7.3	13.0	9.6	8.9	9.1	10.7	9.2	5.2	12.7	12.7	10.9
−25°	2.5	1.8	1.3	5.2	9.2	4.8	5.2	8.2	5.5													9.4	9.0	5.5	12.8	9.1	10.4
−45°	3.6	1.7	1.9																								
								Vertical Error (deg)																		
45°	10.2	18.4	5.0	8.4	16.2	7.0	7.8	13.8	9.3	6.8	14.6	8.2							5.9	13.1	7.2	11.3	14.4	7.4	11.4	11.3	11.3
25°	6.0	19.0	7.4	5.4	10.0	6.3	7.0	14.9	6.6	4.2	7.0	1.6							7.3	8.2	1.5	7.3	12.0	1.8	19.1	18.0	4.9
5°	4.1	6.2	2.0	4.6	7.3	2.6	3.9	6.6	3.3	4.0	7.0	1.6	3.2	8.4	2.6	6.5	11.4	0.8	7.8	6.2	1.0	8.0	13.3	1.8	9.2	16.6	2.6
−5°	3.5	8.7	2.2	4.1	7.8	2.6	3.4	7.0	2.4	5.4	12.4	5.5	4.6	10.7	2.3	7.2	11.4	0.8	7.4	17.4	5.3	7.9	16.4	6.9	6.9	16.4	2.8
−25°	6.5	11.5	3.8	8.0	9.8	4.6	5.9	9.9	5.9													10.5	25.6	10.2	7.9	18.8	7.4
−45°	7.5	14.7	10.5																								

Figure 4. Judgment centroids for the verbal-coordinate condition of Experiment 2. The top three panels show the azimuth coordinate of the judgment centroid as a function of the azimuth coordinate read by the experimenter, and the bottom three panels show the elevation coordinate of the judgment centroid as a function of the elevation coordinate read by the experimenter. Note that, relative to those in the top panels, the values on the axes in the bottom panels have been reduced substantially. Each centroid is based on four judgments. Front–back confusions have been resolved.

also smaller than the horizontal errors observed by Makous and Middlebrooks (1990; 26 of 38 comparisons show smaller horizontal errors for Experiment 2). The vertical errors are substantially smaller in Experiment 2 than in Experiment 1 (37 of 37 untied comparisons show smaller vertical errors in Experiment 2); these errors are also smaller than the errors observed by Makous and Middlebrooks (30 of 37 untied comparisons show smaller vertical errors for Experiment 2). These results indicate that the subjects were able to produce responses that are more accurate than the localization judgments recorded in Experiment 1. Thus, it seems likely that performance in Experiment 1 was limited by the accuracy of the subjects' perceptions and not by their ability to produce accurate motor responses using the GELP technique. Note, however, that the subjects found the task in Experiment 2 to be quite difficult and responded more slowly than they did in Experiment 1; it may be that the more slug-

gish manner of responding in Experiment 2 contributed to the increase in judgment accuracy.

GENERAL DISCUSSION

The results of Experiment 1 suggest that subjects can produce localization judgments at very high rates using the GELP technique. The obtained judgments were similar in accuracy to those obtained with the verbal-reporting technique of Wightman and Kistler (1989b), but were, in general, less accurate than those obtained with the head-pointing technique of Makous and Middlebrooks (1990). The results of Experiment 2 showed that subjects were able to produce judgments to verbal coordinates that were more accurate than the judgments they had produced in the auditory-localization condition of Experiment 1, and suggest that it was not the GELP technique that had limited their performance in Experiment 1.

Which response technique is best for a particular experiment is likely to depend on a number of experimental and nonexperimental factors. The verbal-reporting technique is easy and economical to implement, and no special devices or hardware need to be placed near the subjects that might cast an acoustic shadow or lead to unwanted reflections. However, response reporting is slow and requires manual entry on the part of the experimenter. Furthermore, this technique would seem to be impractical when head movements are to be constrained with a bite bar. The head-pointing technique seems to be a very accurate way to make localization judgments. Response recording is automatic and slightly faster than it is with the verbal-reporting technique. It is, however, somewhat costly to implement, may require considerable subject training (D. J. Kistler, personal communication, April 1994), and would appear to be awkward if a bite bar is required; in addition, this technique does require a small receiver and a transmitter to be positioned on and near the head. The major advantage of the GELP technique lies in the automatic and rapid recording of responses; further, it is comfortably used with a bite bar and produces accurate responses. However, it is more costly to implement than the verbal-reporting technique, and although it does not require additional pieces of apparatus to be placed on or near the head, the spherical model itself could cast an acoustic shadow for sound sources at low elevations in front of the subject.

An important limitation of the GELP technique (and also of the head-pointing technique) is that no mechanism for reporting distance is specified. We plan on extending the GELP technique to include distance judgments, either by replacing the solid model with a wireframe model or by incorporating a second response, in which the subject points to a linear scale indicating the perceived distance.

SUMMARY

A pointing technique has been described that allows subjects to report the perceived direction of sound sources in auditory localization experiments. A major advantage of this procedure is that data can be collected at substantially faster rates than they could be with other techniques. Our results indicate that, compared with other techniques, the number of responses collected per min can be increased by a factor of between 2 and 8. This increase in speed is of great practical consequence. Setting up the apparatus for a localization experiment requires a substantial investment in space, money, time, and effort. This investment is so large that it is almost never practical to run more than one subject at a time. Further, localization experiments require a large number of responses to be collected under each experimental condition. Finally, as our understanding of sound localization increases, so will the complexity of our research questions. The GELP technique will not only allow routine questions to be answered more rapidly, but also allow researchers to

address questions that would be too complex to answer using other techniques.

REFERENCES

BUTLER, R. A. (1987). An analysis of the monaural displacement of sound in space. *Perception & Psychophysics*, **41**, 1-7.

DOLL, T. J., GERTH, J. M., ENGELMAN, W. R., & FOLDS, D. J. (1986). *Development of simulated directional audio for cockpit applications* (AAMRL-TR-86-014). Wright-Patterson Air Force Base, OH: Armstrong Aerospace Medical Research Laboratory.

ERICSON, M. A., D'ANGELO, W., & MCKINLEY, R. (1994). Response measures used in free-field localization. *Journal of the Acoustical Society of America*, **95**, 2897.

FISHER, N. I., LEWIS, T., & EMBLETON, B. J. (1987). *Statistical analysis of spherical data*. Cambridge: Cambridge University Press.

KISTLER, D. J., & WIGHTMAN, F. L. (1992). A model of head-related transfer functions based on principal components analysis and minimum-phase reconstruction. *Journal of the Acoustical Society of America*, **91**, 1637-1647.

MAKOUS, J. C., & MIDDLEBROOKS, J. C. (1990). Two-dimensional sound localization by human listeners. *Journal of the Acoustical Society of America*, **87**, 2188-2200.

MCKINLEY, R. L., & ERICSON, M. A. (1989). A real-time digital auditory localization cue synthesizer. In *Sound localization by human observers symposium proceedings* (p. 45). Washington, DC: National Academy Press.

MIDDLEBROOKS, J. C. (1992). Narrow-band sound localization related to external ear acoustics. *Journal of the Acoustical Society of America*, **92**, 2607-2624.

MIDDLEBROOKS, J. C., & GREEN, D. M. (1991). Sound localization by human listeners. *Annual Review of Psychology*, **42**, 135-159.

MUSICANT, A. D., & BUTLER, R. A. (1984). The influence of pinnae-based spectral cues on sound localization. *Journal of the Acoustical Society of America*, **75**, 1195-1200.

OLDFIELD, S. R., & PARKER, S. P. A. (1984a). Acuity of sound localisation: A topography of auditory space. I. Normal hearing conditions. *Perception*, **13**, 581-600.

OLDFIELD, S. R., & PARKER, S. P. A. (1984b). Acuity of sound localisation: A topography of auditory space. II. Pinna cues absent. *Perception*, **13**, 601-617.

STEIGER, J. H. (1980). Tests for comparing elements of a correlation matrix. *Psychological Bulletin*, **87**, 245-251.

VALENCIA, G., CALHOUN, G. L., ERICSON, M. A., & AGNEW, J. R. (1990). *Localization performance with synthesized directional audio* (AAMRL-TR-90-025). Wright-Patterson Air Force Base, OH: Armstrong Aerospace Medical Research Laboratory.

WENZEL, B. M., WIGHTMAN, F. L., & FOSTER, S. H. (1988). A virtual display system for conveying three-dimensional acoustic information. *Proceedings of the Human Factors Society*, **32**, 86-90.

WIGHTMAN, F. L., & KISTLER, D. J. (1989a). Headphone simulation of free-field listening. I: Stimulus synthesis. *Journal of the Acoustical Society of America*, **85**, 858-867.

WIGHTMAN, F. L., & KISTLER, D. J. (1989b). Headphone simulation of free-field listening. II: Psychophysical validation. *Journal of the Acoustical Society of America*, **85**, 868-878.

WIGHTMAN, F. L., & KISTLER, D. J. (1992). The dominant role of low-frequency interaural time differences in sound localization. *Journal of the Acoustical Society of America*, **91**, 1648-1661.

NOTES

1. Wightman and Kistler (1989a, 1989b) used a coordinate system that is similar to standard spherical coordinates. This system is also employed in the present paper. In this system, *0° azimuth and 0° elevation* refers to a location that is directly in front of the subject within the horizontal plane (distance from the center of the coordinate system—i.e., from the center of the subject's head—is typically ignored). Positive azimuths (from 0° to 180°) refer to locations to the subject's right; negative azimuths (from 0° to −180°) refer to locations to the

subject's left; positive elevations (from 0° to 90°) refer to locations above the horizontal plane; and negative elevations (from 0° to −90°) refer to locations below the horizontal plane.

2. The Polhemus Fastrak generates and measures electromagnetic fields to monitor the location of the stylus. It is important that no metal objects are positioned close to the spherical model (the distance between the transmitter and the nearest metal object should be at least 2.5 times the maximum distance between the transmitter and the receiver/stylus). Therefore, the bench, bite bar, and spherical model are constructed primarily of wooden and plastic parts and fasteners.

3. The x-, y-, z-coordinates of the center of the spherical model are determined before the experimental session by measuring and averaging the x-, y-, z-coordinates of four points in the horizontal plane. This is accomplished by placing the stylus tip at (0° azimuth, 0° elevation), (90° azimuth, 0° elevation), (180° azimuth, 0° elevation), and (−90° azimuth, 0° elevation).

4. The values presented are rates per min within blocks. Differences in procedure make a strict comparison of effective data rates difficult. First, the trial structures in the three experiments were somewhat different. Makous and Middlebrooks (1990) used brief stimuli (150 msec), but essentially required the subjects to make two responses per trial— a centering response to start the trial, and the actual response. Wightman and Kistler (1989b) used a long period for stimulus presentation, 4,400 msec. We used a single brief stimulus presentation (268 msec).

Second, Wightman and Kistler and Makous and Middlebrooks both used long blocks, 10–25 min, while we used short blocks, 2.5–3 min, separated by 1–2 min breaks. Finally, the apparatus of both Wightman and Kistler and Makous and Middlebrooks required that speakers be moved between trials. Some portion of the total trial duration was devoted to this activity. D. J. Kistler (personal communication, April 1994) has recently estimated the response rate for the verbal-reporting technique at 6–10 responses per minute, using stimuli more comparable to those employed here. Ericson, D'Angelo, and McKinley (1994) directly compared the verbal-reporting and GELP techniques in a task where subjects were encouraged to make very accurate responses and concluded that the GELP technique was slightly more accurate and two times faster. All things considered, the GELP technique probably allows 2 to 8 times the number of responses to be collected in a single session.

5. Valencia, Calhoun, Ericson, and Agnew (1990) measured the ability of subjects to judge the azimuth of virtual sound sources. In this study, they compared circle-pointing (the two-dimensional analog of the GELP technique) with the verbal-reporting technique. They also found that the two techniques yield data that are comparable in accuracy.