# The nature and distribution of errors in sound localization by human listeners

Simon Carlile [a,c,*], Philip Leong [b], Stephanie Hyams [a]

[a] Auditory Neuroscience Laboratory, Department of Physiology, University of Sydney, Sydney, NSW 2006, Australia
[b] SEDAL, Department of Electrical Engineering, University of Sydney, Sydney, NSW 2006, Australia
[c] Institute for Biomedical Research, University of Sydney, Sydney, NSW 2006, Australia

## Abstract

Measurement of localization performance will reflect errors that relate to the sensory processing of the cues to sound location and the errors associated with the method by which the subject indicates the perceived location. This study has measured the ability of human subjects to localize a short noise burst presented in the free field with the subject indicating the perceived location by pointing their nose towards the source. Subjects were first trained using a closed loop training paradigm which involved instantaneous feedback as to the accuracy of head pointing which resulted in the reduction of residual localization errors and a rapid acquisition of the task by the subjects. Once trained, 19 subjects localized between 4 and 6 blocks of 76 target locations. The data were pooled and the distribution of errors associated with each target location was examined using spherical methods. Errors in the localization estimates for about one third of the locations were rotationally symmetrical about their mean but the remaining locations were best described by an elliptical distribution (Kent distributed). For about one half of the latter locations the orientations of the directions of the greatest variance of the distributions were not aligned with the azimuth and elevation coordinates used for describing the spatial location of the targets. The accuracy (systematic errors) and the distribution of the errors (variance) in localization for our population of subjects were also examined for each test location. The size of the data set and the methods of analysis provide very reliable measures of important baseline parameters of human auditory localization.
© 1997 Elsevier Science B.V.

*Key words:* Sound localization; Spectral cue; Pinna; Binaural processing

## 1. Introduction

Sound source localization by binaurally hearing individuals has been shown to be dependent on a number of stimulus related factors (e.g. Middlebrooks, 1992; Butler, 1986; Butler and Humanski, 1992; Butler and Belendiuk, 1977) including the acoustic conditions under which we measure performance (e.g. Hartmann, 1983; see Carlile, 1996a for recent review). Measurement of localization performance involves an estimate of the errors associated with this process and will reflect a combination of errors dependent on the method of

measurement and errors associated with sensory component of the task. In this study we were interested in (i) examining methods which minimize as much as possible the errors associated with measuring localization performance and (ii) examining the nature of the distributions of the errors associated with localization performance. An overall objective of this research program is to combine this latter information with what is known about the spatial variation in the cues to sound location. In this way we hope to make inferences about the process(es) involved in sound localization, the utility of different cues to a sounds location and the resolution of their encoding by the auditory system.

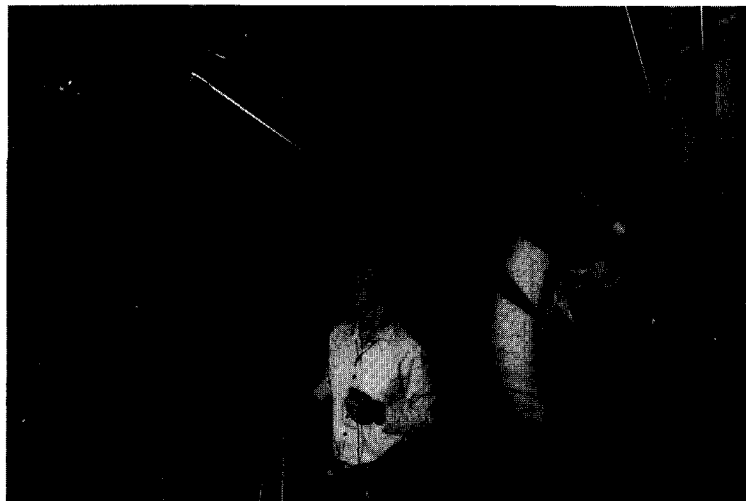There have been a number of recent developments in automated techniques for presenting stimuli as well as

Fig. 1. Subject demonstrating the sound localization task in the anechoic chamber. The chamber is 4 m a side with a clear working area of 2.5 m³. The double hoop robot arm supporting the sound source can be seen hanging from the ceiling of the chamber. The receiver for the Polhemus head tracker is mounted on a modified support from a welding helmet. The transmitter can be seen placed behind the subject on a wooden post. The subject indicates the completion of localization or training task by pressing the subject response button seen in this subject's left hand.

determining how the subject indicates the perceived location (e.g. Makous and Middlebrooks, 1990; Gilkey et al., 1995; see Carlile, 1996b for review). Recent data suggest that aspects of these methodologies may also affect the resulting measured localization performance. For instance, in many previous studies of sound localization, the most practical way of varying the sound location has been to switch the stimulus between a number of fixed speakers. Under these conditions the most straightforward means of indicating the perceived locations was for the subject to identify the speaker perceived to be generating the sound (e.g. Butler et al., 1990; Hammershoi and Sandvad, 1994). However, when subjects are aware of the potential target locations, then restricting the number or location of targets will constrain the subjects responses to these categories (Perrett and Noble, 1995). This has the undesirable effect of forcing subjects to indicate a possible target location that does not actually correspond to the perceived location and will also result in quantization of the estimates of the accuracy of spatial localization. Restricting the targets to a limited region of space may also provide cues to resolve perceptual ambiguities such as those involved in front-back confusions. In addition, the particular arrangement of potential targets has also been found to bias judgements for some spatial locations (Butler and Humanski, 1992; Perrett and Noble, 1995; for recent discussion and review of these issues see Carlile, 1996c).

An alternative procedure is to use an unseen and movable sound source together with a method of pointing to indicate the perceived location (Makous and Middlebrooks, 1990; Oldfield and Parker, 1984a,b, 1986; Gilkey et al., 1995). As both of these methods

are spatially continuous rather than quantized they provide more sensitive measures of localization accuracy and avoid the methodological pitfalls of 'category' localization discussed above. Recent advances in automating the stimulus positioning systems and the methods used in tracking subject pointing has also increased the popularity of these methods. The availability of relatively inexpensive electromagnetic 3D tracking devices that can be mounted on the top of a subject's head has made possible the use of 'head pointing' as a way of indicating the perceived location. In these experiments the subject is instructed to point his or her nose towards the target location (e.g. Makous and Middlebrooks, 1990). One previous group of studies used 'gun' pointing but the optical methods used in assessing pointing location were not automated (Oldfield and Parker, 1984a,b, 1986).

Turning to face towards the source of a sound is a highly ecological behavior. The functional consequence is to bring the source of the sound into the visual field. A possible source of error associated with such a technique is that the eyes are also free to move in the head. For sound locations that require pointing up to the mechanical extremes of movement of the head (e.g. extremes of elevation) or for locations that are close to the resting location of the head, there is a strong tendency for subjects to also use movements of the eyes to 'capture' the location of the auditory target. As the sensor is mounted on the top of the subject's head and monitors head position and not eye position, this would result in systematic errors in estimating the perceived location of the target. In a previous study using head pointing, training was provided for subjects in the form of localization of a visible sound source (Makous
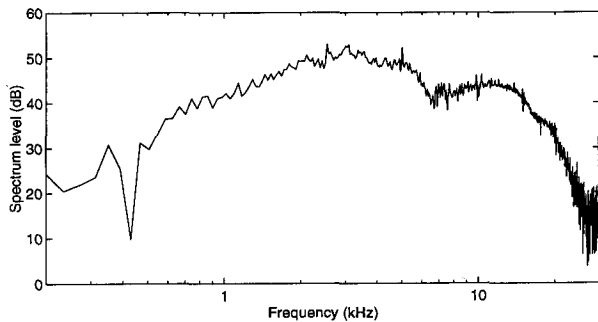
Fig. 2. The transfer function of the VIFA tweeter used as the sound source. Note that for frequencies below 3 kHz, the speaker rolls off at about 6–7 dB per octave. For frequencies below 700 Hz there is a rapid fall off in transmission. The higher frequency roll off is also quite steep but transmission remains high for frequencies up to about 15 kHz.

and Middlebrooks, 1990). This approach would provide subjects some opportunity to correct for eye pointing, however, in the absence of feedback to the subject regarding the absolute accuracy of their pointing, the systematic errors described above may still persist.

In this first part of this paper we describe a method of training subjects using closed loop training coupled with immediate sensory and cognitive feedback as to the absolute accuracy of head pointing. We show that this results in more accurate pointing compared to other closed loop training paradigms. Using subjects trained in this way we have examined free field localization accuracy for 76 locations around the subject. We have also examined the distribution of localization errors using a number of spherical statistical methods. We report here that the errors are not arranged symmetrically about the mean and, for some spatial locations, the axes of distribution do not follow the azimuth and elevation coordinate axes usually used in describing the locations of auditory targets in space. The distribution of errors is discussed in terms of recent models of localization processing which integrate binaural information with monaural spectral cues to the location of a sound source.

## 2. Methods

### 2.1. Testing environment

All localization testing was carried out in a sound attenuated chamber, anechoic down to 150 Hz (better than 99% absorption measured at greater than 0.3 m from the wall). The triple walled chamber provides an insertion loss of better than 40 dB for frequencies greater than 250 Hz. The clear working area within the chamber was 2.5×2.5×2.5 m. Subjects were placed on a platform at the center of the chamber such that their head was centered in the middle of this clear

working area (see Fig. 1). A two-way intercom system provided communication between the subject inside the chamber and the experimenter outside.

The location of the sound stimulus was varied using a computer controlled positioning system based on a suspended double hoop design (Fig. 1). A loudspeaker mounted on the inner hoop could be placed at almost any location on an imaginary sphere one metre radius from the center of the subject's head. Because of mechanical restrictions, the stimulus could not be placed directly below the subject. A small light emitting diode was also fitted at the center of the speaker to provide visual information as to the location of the sound source. The stimulus positioning system was driven by high resolution, high torque stepper motors to control the azimuth and elevation locations of the stimulus. Location could be varied at rates of up to 90°/s with azimuth acceleration/deceleration of $36°/s^2$ and elevation accelerations/deceleration of $33°/s^2$. This allowed rapid stimulus placement over a wide range of locations about the subject. The dimensions of all of the mechanical hardware within the chamber were kept to a minimum to ensure that the anechoic environment was not disturbed over the frequency range of interest.

### 2.2. Stimulus generation

Software for stimulus generation and data collection exploited the Tucker Davis Technology (TDT) system II hardware and software platform. Broadband white noise stimuli were generated using D/A conversion at 80 kHz and delivered to a power amplifier (Quad 306) via a programmable attenuator (TDT: PA4). The noise stimuli were regenerated for each stimulus presentation rather than using a single 'frozen' noise stimulus. The loudspeaker (VIFA-D26TG-35) mounted on the hoop positioning system had a frequency response of 1 kHz to 16 kHz (± 5 dB: see Fig. 2). Broadband stimuli were presented at 70 dB SPL measured at the center of the stimulus positioning system (microphone B&K 4165; pressure level meter B&K 2203).
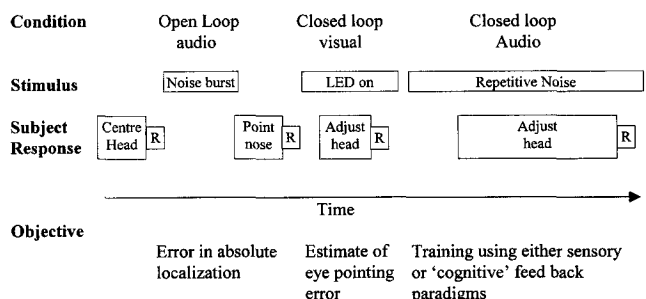


Fig. 3. Time line showing the sequence of stimuli and subject responses in one training cycle. 'R' indicates when the subject depresses the response button to indicate the completion of the preceding portion of the training task.
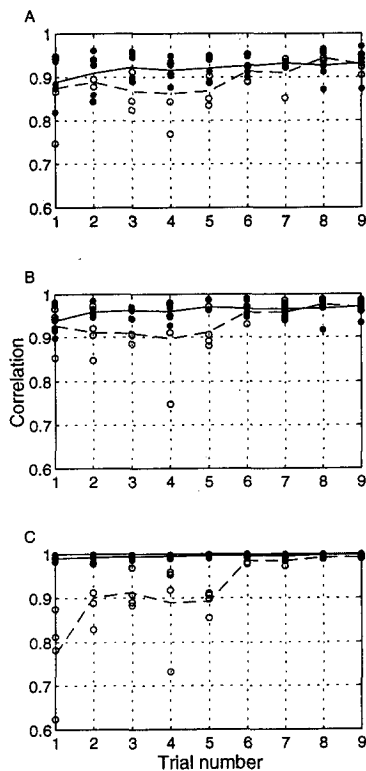
Fig. 4. The effects of training on the localization performance for (A) open-loop localization of a single burst of broadband noise, (B) orientation of the head towards a light source and (C) orientation towards a repetitive broadband noise burst. Localization performance has been summarized using the spherical correlation coefficient of the indicated location and the actual location of the target. In all plots the open circles indicate those subjects who have undergone training with the 'sensory' closed loop paradigm (CLS) up to the fifth training block after which they were transferred to the 'cognitive' closed loop training paradigm (CLC: see text for details). The broken line indicates the mean of this group. The filled circles indicate those subjects who have undergone a 'cognitive' closed loop training paradigm for the whole nine training blocks.

## 2.3. Measurement of perceived sound location

Head pointing was used as the method by which subjects indicated the perceived location of the sound source. Two advantages of this approach are (i) head pointing per se is a highly ecological behavioral response to a sound and (ii) tracking the position of the head in space is relatively straightforward. A Polhemus Isotrack tracking device was used to measure head position. The receiver was attached to the top of the subject's head using the plastic frame modified from a welding mask. The transmitter was located behind the subject on a wooden support (see Fig. 1) and the position of the head was continuously tracked (TDT Isotrak interface, HTI 1). Feedback to the subject as to her head orientation with respect to the stimulus coordinate system could also be provided via a series of colored light emitting diodes placed in front of the subject. When the subject was properly positioned at the cali-

brated start position (azimuth 0°, elevation 0°, see below) a central green light on the LED array was illuminated. Deviations of the head by more than 1.5° in azimuth, elevation or roll from the calibrated zero position was indicated by illumination of one or more of six other red LEDs arranged around the central green LED and indicated up-down, right-left or clockwise-anticlockwise roll errors respectively. The subject indicated the completion of a particular task (e.g. positioning the head) by pressing a hand held response button.

At the beginning of each trial the subject was positioned in the center of the support platform so that her head was at the center of the hoop-stimulus system. To assist the subject, head position feedback was provided via the light-emitting diode array. To provide an initial alignment for the head and head tracking system at the beginning of each measurement session the subject stood in the center of the support platform and pointed her face towards the visible target at the start position (located at 0° azimuth, 0° elevation). The subject was thus aligned with the hoop-stimulus coordinate system using a perceptual task. The head tracker was then placed on the subjects head and aligned with the hoop-stimulus system with the aid of the light-emitting diode array. Once in the start position the subject initiated the trial by depressing the start button. The ability of subjects to reliably place themselves at the start location between blocks of trials was determined for four subjects by fitting the head tracker as described above. The subject was then requested to place her head at the start location a number of times but without the aid of the LED array. The output of the head tracker was noted for each repeat. These data showed that subjects can reliably place their head in the start position with an error of less than 2–3°.

## 2.4. Task training

A problem with the head pointing method is that subjects also tend to move their eyes to visually 'capture' the location of a target. As a consequence we have established a training protocol to eliminate, as much as possible, this eye pointing error. During training each subject accomplished a series of tasks for a large number of stimulus locations (Fig. 3). After centering the head, a 150 ms burst of noise was presented and the subject was required to point her face towards the perceived location of the noise burst. All stimuli were presented in complete darkness. This was referred to as the open-loop estimate of location. Once the subject had orientated her head appropriately and depressed the response button, a small light-emitting diode on the speaker was activated. The subject was allowed to readjust her head to correct for any perceived error in head pointing. This was referred to as the closed-loop visual condition and allowed the subject to adjust for gross
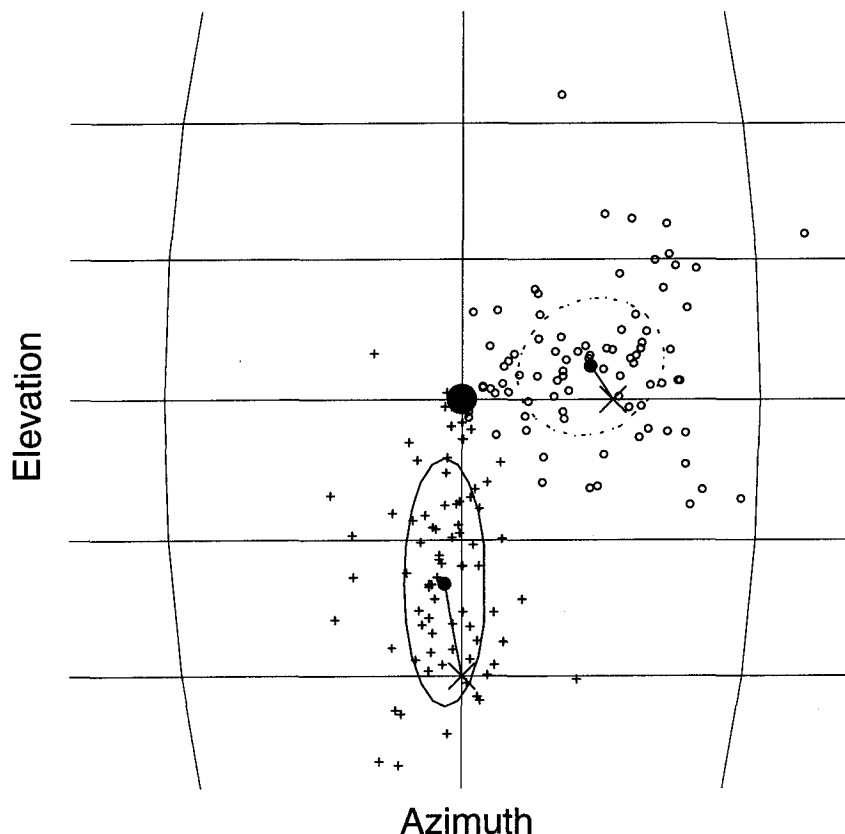
Fig. 5. Two examples of the distribution of errors in the pooled location estimates for two spatial locations: Az 10°/El 0° (small open circles) and Az 0°/El −20° (small crosses). The large filled circle at the center of the figure indicates the direction directly ahead of the subject (Az 0°/El 0°). The large 'X' indicates the actual location of the target. The centroids of both clusters of data are indicated by the filled circles and the standard deviation about the centroid is indicated by the ellipses. The broken lined ellipse indicates a Fisher (symmetrical) distribution of the data while the solid line indicates a Kent (elliptical) distribution of the data.

errors in head pointing and provided an estimate of the eye pointing error for that location. This was followed by a close-loop audio component where the sound stimulus was presented repetitively allowing the subject to adjust for any further head pointing error. Once satisfied that she was pointing towards the target, the subject depressed the response button for the third time and a new trial was initiated.

Each training block was composed of 36 locations drawn from a total of 324 different positions. Ten subjects were trained on a total of nine blocks and a further nine subjects were trained on four or five blocks (see below). The target locations were chosen to be roughly equally distributed on the surface of the sphere and were intended to provide training for a large number of positions between −50° and 40° elevation. All psychophysical procedures used in this study were approved by the Human Ethics Committee at the University of Sydney, Australia.

2.5. 'Sensory' and 'cognitive' closed loop training

Two different closed-loop audio-visual conditions

have been tested in this study. The first was the closed-loop sensory (CLS) condition where the repetition rate of the closed loop stimulus was kept constant at 1 Hz (Fig. 3). Four subjects were initially trained using this condition and their performance was compared with the results from six other subjects who underwent a closed-loop cognitive (CLC) condition. In the CLC condition the repetition rate of the stimulus was dependent upon the absolute pointing accuracy of the subject (minimum about 0.5 Hz, maximum 5.5 Hz). In this case, the position of the head was monitored continuously and the repetition rate of the stimulus varied according to the absolute accuracy of head pointing. The presentation rate was at its highest when the subject was pointing directly at the sound source. Both groups of subjects trained on nine blocks of 36 trials. On the sixth training block the subjects assigned to the CLS group were switched to the CLC method. Those in the CLC group did not switch, but remained with the same training method during the whole course of the training.

A  Left                                                    Right

B  Front                                                   Back

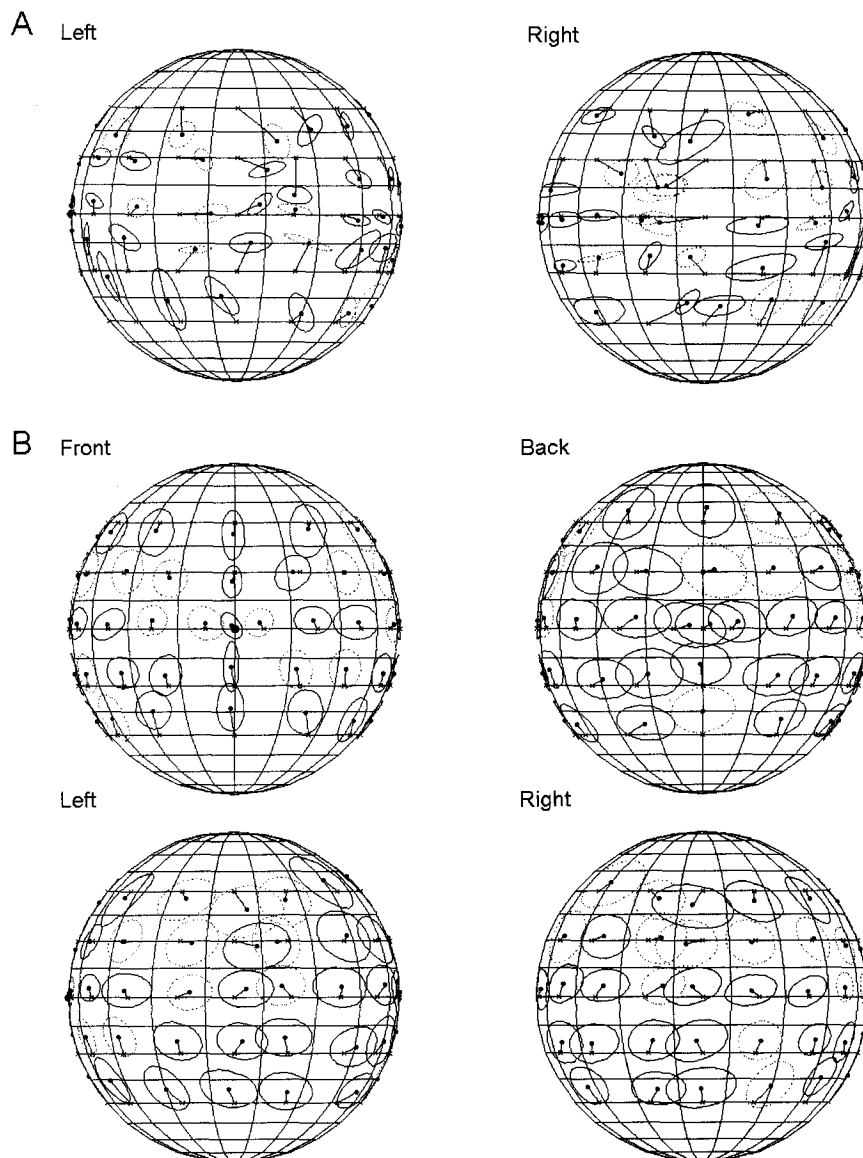   Left                                                    Right

Fig. 6. The centroids and the standard deviations of the localization estimates of 76 target locations are plotted for (A) one subject on spherical plots for the right and left hemispheres and (B) pooled from 19 subjects and plotted for the front, back, left and right hemispheres. All other details as for Fig. 5.

## 2.6. Open loop testing conditions

In the second part of this study the localization accuracy of 19 subjects was examined using the head pointing technique. Ten of the subjects had participated in the first part of the study described above. An additional nine subjects were given four or five blocks of training in head pointing using the CLC condition so that their performance as measured using the spherical correlation coefficient had reached an asymptote (see below). The stimulus in all localization tests was restricted to a single 150 ms broad-band noise presentation with a rise and fall time of 5 ms. The 150 ms duration was chosen to ensure that the subjects were not able to move their heads during stimulus presenta-

tion and thereby re-sample the sound field. Subjects positioned themselves at the start position with the aid of the LED array and initiated a trial by pressing the response button. Following the stimulus presentation the subject was required to point her face towards the location of the target and depress the response button. All localization testing and training was carried out in complete darkness so the subject had no visual cues as to the location of the target during this open-loop location estimate. As the stimulus positioning system made some noise during stimulus repositioning (typically < 40 dB SPL A-weighted), a two step repositioning was employed where the first position of each pair was drawn randomly from the list of test locations. The experiments were carried out in complete darkness so

that the subject was unaware of whether the steps were additive or subtractive (see also Makous and Middlebrooks, 1990).

## 2.7. Localization errors

In general there are two types of localization errors. The first type is referred to as a local error where the perceived location is within about 20° of the actual target location. The second type of error is referred to as a front-back confusion error, or a cone of confusion error, and represent a few percent of the total localization estimates in any block of trials. A front-back localization error is where the angle of estimate with respect to the median plane is correct but the hemisphere of the target is confused. For instance, a location 10° to the left of the anterior mid-line could be confused with a location 10° left of the posterior mid-line. Because of the large qualitative differences in these errors, in this and other studies, the front-back confusion errors have been extracted from the responses and are dealt with separately. This has the added benefit of ensuring that the distribution of errors are unimodal rather than bimodal which simplifies the statistical analysis. For analytical purposes a front-back confusion was defined as any error estimate that crossed the interaural axis.

## 2.8. Data analysis

### 2.8.1. Spherical coordinate system

A single pole spherical coordinate system was used to describe points on a unit sphere centered about the subject's head (see Carlile, 1996a). A point directly in front of the subject is described as being 0° azimuth and 0° elevation. The azimuth coordinate increases in a clockwise direction from 0° azimuth and elevation increases upwards from 0° elevation.

### 2.8.2. Examination of the distribution of errors using spherical statistics

Localization errors have been analyzed using spherical statistical methods and the distributions modelled as either a Fisher (symmetrical) distribution or a Kent (elliptical) distribution (Kent, 1982; Fisher et al., 1987). The Kent distribution is a generalization of the Fisher distribution which can deal with asymmetrically distributed data. Using this method of modelling, we are attempting to make less restrictive assumptions about the nature of the statistical distribution. These methods and their implementation are described in detail elsewhere (Leong and Carlile, 1997) but the main points are summarized below.

The Kent distribution is described by the parameters $\mathbf{G}$, $\kappa$, $\beta$, where $\mathbf{G}$ is a $3 \times 3$ matrix containing the three $3 \times 1$ column vectors ($\xi_1$, $\xi_2$, $\xi_3$). $\xi_1$ is the mean direction of the distribution, $\xi_2$ is the direction in which the

data density is the highest (major axis), and $\xi_3$ is the direction of least data density (minor axis). The parameters were calculated using estimation by moments (Fisher et al., 1987). It is most convenient to think of $\mathbf{G}$ as being the rotation matrix which best aligns the sample mean direction to the 'north pole' of the sphere. The $\kappa$ parameter describes the degree of concentration of the data about the pole of the distribution, and $\beta$ is the ovalness parameter which is small for circular data and increases as the data becomes more ovoid. Using the shape parameters $\kappa$ and $\beta$, unimodal distributions ($\kappa/\beta \geq 2$) can be distinguished from bimodal distributions ($\kappa/\beta < 2$). In order to obtain the ellipses used in Fig. 6, the data were rotated using the $\mathbf{G}$ matrix to align them with the pole

$$\begin{pmatrix} x_j' \\ y_j' \\ z_j' \end{pmatrix} = \mathbf{G} \begin{pmatrix} x_i \\ y_i \\ z_i \end{pmatrix}^{\mathrm{T}}$$

Following rotation, the principal components (direction of greatest variance) of the data distribution align with the azimuth and elevation axes, centered about the pole. The standard deviation along the polar directions is then calculated and the points of an ellipse about (0,0,1) with major and minor axes one standard deviation in size are computed. These points are then rotated back to the original position again using the $\mathbf{G}$ matrix to generate the plotting coordinates to produce an ellipse which has its major and minor axes in the principal directions of data variance, and with the size being proportional to the standard deviation of the data. The method of rotating the data before computing the statistic avoids the problem that azimuth values are distorted at elevations away from the equator.

A general metric of localization performance is the spherical correlation coefficient (SCC) of the actual and perceived locations (Fisher et al., 1987; e.g. Wightman and Kistler, 1989). If $\mathbf{X}_i$ is a $n \times 3$ matrix of the direction cosines of $n$ perceived locations, and $\mathbf{X}_i^*$ are the corresponding actual target locations, then we can calculate the SCC $\rho$ as follows

$$S_{\mathrm{XX}^*} = \det\left(\sum_{i=1}^{n} \mathbf{X}_i \mathbf{X}_i^{*'}\right)$$

$$S_{\mathrm{XX}} = \det\left(\sum_{i=1}^{n} \mathbf{X}_i \mathbf{X}_i'\right)$$

$$S_{\mathrm{X}^*\mathrm{X}^*} = \det\left(\sum_{i=1}^{n} \mathbf{X}_i^* \mathbf{X}_i^{*'}\right)$$

$$\rho = \frac{S_{\mathrm{XX}^*}}{\sqrt{(S_{\mathrm{XX}} S_{\mathrm{X}^*\mathrm{X}^*})}}$$

A



Azimuth

B



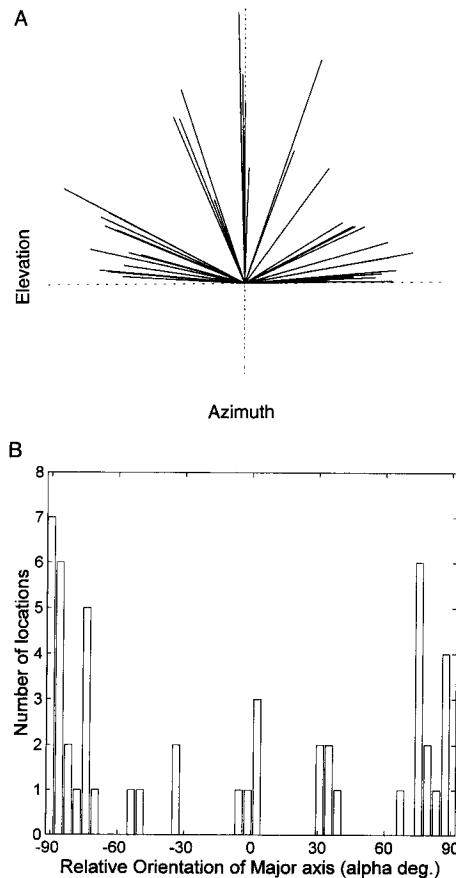Relative Orientation of Major axis (alpha deg.)

Fig. 7. A: The orientations of the principal components of the Kent distributed localization estimates are shown. To facilitate comparison, each vector has been rotated to Az 0°/El 0° and the length is relative to the standard deviation indicated by that component. B: The orientation of the rays in 'a' have been calculated with respect to the positive vertical axis ($\alpha$°) and plotted as a distribution histogram. Note that approximately half of the rays are in excess of 10° from the vertical or the horizontal axes.

The value of $\rho$ is between $-1$ and 1, being 1 if one set can be transformed to the other by a rotation, and $-1$ for a reflection.

## 3. Results

### 3.1. 'Sensory' verses 'cognitive' closed loop training

In this study the two methods of training were evaluated by calculating the spherical correlation coefficient of the actual and perceived locations of the targets following extraction of the front-back confusions in the data. Spherical plots of the responses were generated at the end of each block and provided additional feedback to subjects as to their performance during training. Feedback was not routinely provided during testing.

The results from open loop localization (Fig. 4A)

demonstrate that in the first trial block all subjects achieved similar correlation coefficients regardless of group allocation. However, as the block number increased, so too did the separation in the correlation coefficients for each training paradigm. The mean correlation coefficient for CLC subjects (Fig. 4A: filled circles; mean solid line) after the first trial block increased to values of between 0.9 and 0.95. There was some scatter, particularly in the earlier trial blocks as some subjects mastered the task more quickly than others. However, even on trial block 2, one subject showed correlation coefficient values greater than 0.95. On the other hand, those in the CLS group (open circles, mean: dashed line) demonstrate comparatively lower correlation coefficient scores over trial blocks 1-5. The mean correlation coefficient score remained below 0.9, and individual values generally ranged from 0.8 (two values were below 0.8) to 0.9. At trial block 6 the CLS subjects changed to the CLC training paradigm. It is evident that for trial blocks 6-9, there was a rapid increase in the mean correlation coefficient for this group to values greater than 0.9. The significance of the differences between each training paradigm was calculated using non-parametric statistics (Mann-Whitney $U$, $P < 0.05$) for training blocks 1-5. Significance between CLC and CLS was demonstrated at trial blocks 3 and 5. There were no significant differences demonstrated after trial block 6, that is, where those in the CLS group changed to the CLC training method.

As noted in Section 2, the stimulus locations used for training varied from block to block. Therefore, some of the differences in the spherical correlations calculated for the open-loop localization between blocks might be attributable to the variations in the locations tested. As the positions for each training block were selected pseudorandomly from the 324 equally spaced spatial locations this variation is likely to be equally dispersed amongst each of the nine training blocks.

Fig. 4B shows the spherical correlation coefficient values obtained when subjects were required to point at the LED located at the stimulus speaker (visual feedback). Again there was a clear separation in the mean value between the two groups over trial blocks 1-5. Those in the CLC group demonstrated values generally greater than 0.95, compared to those in the CLS group where mean values were around 0.9. These differences were statistically significant for trial blocks 2, 3 and 5. Once those in the CLS group changed to the CLC paradigm at trial block 6, there was a marked increase in the level of performance. The former CLS group demonstrated performance levels that were statistically indistinguishable from those who had received CLC training for the entire set of nine training blocks.

Fig. 4C shows correlation coefficient values for localizing and pointing utilizing the audio feedback. The results demonstrate a fairly distinct separation in the

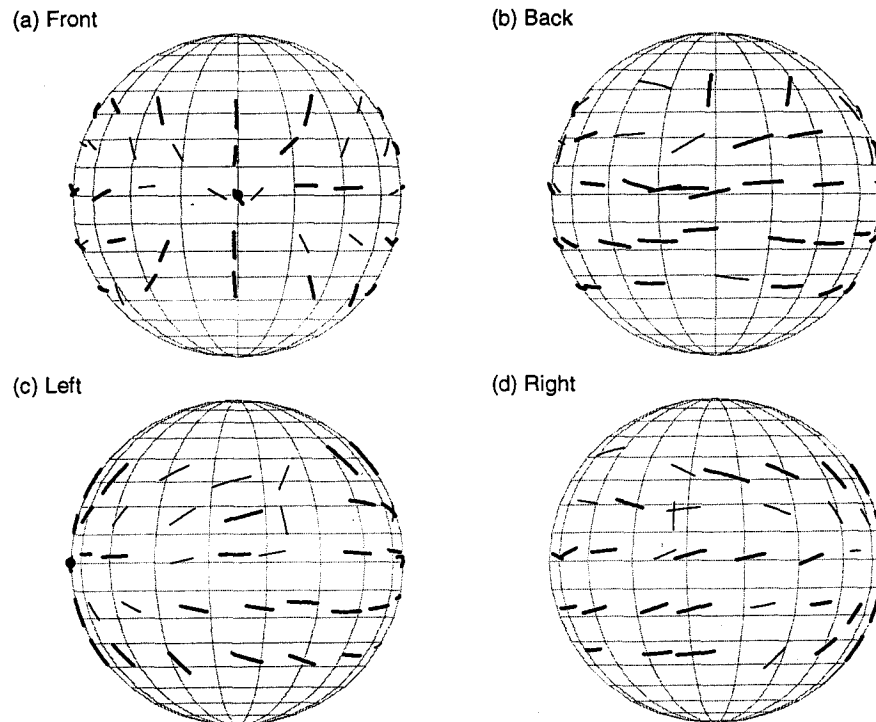(a) Front

(b) Back

(c) Left

(d) Right

Fig. 8. The orientations of the major axes of the Kent distributions have been plotted as solid lines on spherical plots for the front, back, left and right hemispheres. The longest axis of the Fisher distributed data has also been plotted (broken lines) although the differences between the major and minor axes did not reach significance ($\alpha < 0.05$). All other details as for Figs. 5 and 6.

correlation coefficient values between each training method up to trial number 5. Those in the CLC group demonstrate values of greater than 0.97 throughout the entire trial set. Once again, there was a greater scatter of individual values for those in the CLS group as compared to the CLC group. Beyond trial block 6, both groups demonstrate correlation coefficient values much greater than 0.96. The differences between the two groups were statistically significant up to trial blocks 6. From trial blocks 6 to 9, both groups were statistically indistinguishable.

## 3.2. Open loop testing

Once subjects had been appropriately trained, open loop localization data were then collected for 76 spatial locations in a single block of localization trials. Each subject typically performed between four and six blocks of the test locations, the order of which were randomized from block to block. Testing for each block lasted about 20 min and typically each subject carried out two to three blocks of testing per visit to the laboratory.

### 3.2.1. Distribution of localization errors

The 6909 localization responses were pooled from all 19 subjects. The nature of the distribution of errors about each stimulus location was examined to determine if the data were symmetrically distributed about the centroids for each location (Fisher distribution) or

if the distributions were best described by two principal components (Kent distribution; for computational details see Fisher et al., 1987; Leong and Carlile, 1997). In the former case the centroid is a faithful representation of the mean azimuth and mean elevation of the pooled responses and the localization error can be usefully expressed in terms of the degrees azimuth and elevation differences between the centroid and the actual location of the auditory target. However, in the case of the Kent distributed data (Fig. 5: small '+' and solid lined ellipse), the axis of the first and second principal components of the distribution of the data will define the coordinate system of the errors. In this case, the validity of calculating the average azimuth and elevation error will depend entirely on how the axes of the distribution are related to the coordinate sphere used to describe the locations of the targets and the responses. In the case of the Kent distributed data illustrated in Fig. 5 the coordinate system is aligned with the spherical coordinate system so that the major component corresponds to the elevation error in localization and the minor component corresponds to the azimuth error in localization. However, this was not the case for about a third of the locations tested in this study (see below).

### 3.2.2. Orientation of the major axes of the distribution of localization errors

The localization errors for a typical subject are plotted for the right and left hemispheres (Fig. 6A) and the
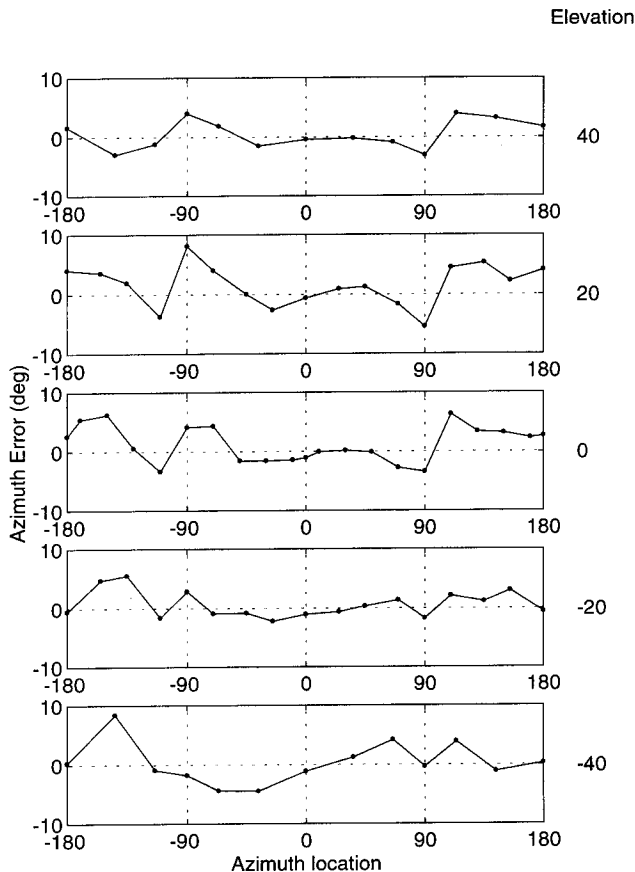
Fig. 9. The differences between the azimuth location of the centroid of the pooled localization estimates and the actual target location is plotted as a function of the azimuth location of the target for each elevation used in this study (40°, 20°, 0°, −20° and −40°). The degrees azimuth have been corrected so that they are all equivalent to degrees azimuth at 0° elevation (see text). Positive errors for azimuth locations between −180° and 0° indicate a shift in the perceived location away form the midline and for locations 0° to 180°, positive errors indicates shifts towards the midline.

localization data pooled for all subjects are plotted in more detail in Fig. 6B. When the distributions of the errors in the pooled data were examined for each of the stimulus test locations 66% (50/76) of the distributions were Kent distributed (Fig. 6B: solid lined ellipses). The Fisher or symmetrically distributed errors did not seem to be associated with any particular group of spatial locations although their incidence was higher for the upper hemisphere. More importantly, however, the principal components of the distribution of errors for the Kent distributed data were not always aligned with the coordinate sphere. This can be seen in both the individual data (e.g. Fig. 6A) as well as the pooled data. For each location where the errors were Kent distributed, we have plotted a vector representing the orientation of the major axis of the distribution with length relative to one standard deviation of the error along that orientation (Fig. 7A). To facilitate comparison each vector starts at the origin. The angle ($\alpha$) is

defined as the angle between the major axis and the lines of longitude with vertical up as 0°. The distribution of $\alpha$ for all of the Kent distributed data is shown as a histogram in Fig. 7B. If the major and minor axes of the ellipses describing the Kent distributed data had been aligned with the spherical plotting coordinates then the lines in Fig. 7A would have fallen on (or close to) the axes of the figure and the histogram would have demonstrated sharp peaks corresponding to −90°, 0° or 90°. Furthermore, the relative sizes of the peaks would correspond to the percentage of the distributions aligned to either of the cardinal axes (azimuth or elevation). In contrast, the data show that for about half of the locations $\alpha > 10°$ from the cardinal axes. For these data, measurement of localization accuracy based on the calculation of the mean and standard deviations of the azimuth and elevation coordinates of the subjects responses are likely to result in significant errors of estimation of the centroid of the distribution.

To examine if there was any particular spatial pattern in the orientation of the major axes of the ellipses for the Kent distributed data, these were plotted on spherical plots (Fig. 8: thick lines). The major axis of
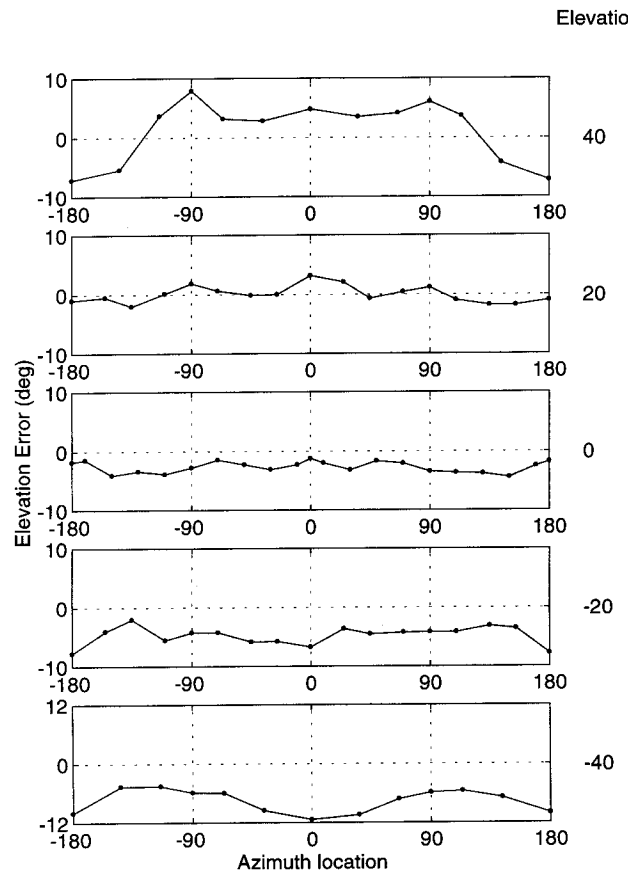


Fig. 10. The differences between the centroid elevation and the target location is plotted as a function of the azimuth location of the target for elevations 40°, 20°, 0°, −20° and −40°. A negative error indicates an estimate above the target.

the Fisher distributed data have also been plotted (thin lines) although the differences between the major and minor axes did not reach significance level ($P < 0.05$). The orientations of the major axes are most markedly deviant from the cardinal axes of the sphere for test locations in the frontal hemisphere. For locations around the interaural axes and in the posterior hemisphere the major axes are principally orientated parallel to the lines of co-latitude. With increasingly more frontal locations, the major axes are increasingly orientated towards the location directly ahead (0°,0°) forming a radial pattern with (0°,0°) as the origin.

### 3.2.3. Accuracy of auditory localization

Systematic errors in localization that result from the measurement procedures or sensory bias in our subjects are indicated by the differences between the target location (Fig. 6, '+') and the centroid of the localization data associated with each target location (filled circle at the center of each ellipse). The azimuth error (Fig. 9) has been plotted for each test elevation (40°, 20°, 0°, −20°, −40°) as a function of the azimuth location of
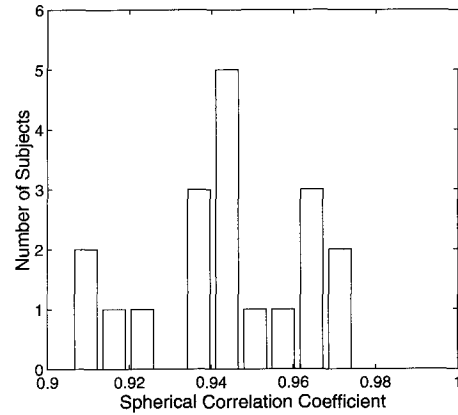


Fig. 12. The spherical correlation of the centroid of the perceived location estimates and the actual target locations have been calculated for each individual subject. The frequency distribution of these correlation values are plotted.

the test stimulus. As a result of using the single pole coordinate system to describe location, the distance on the sphere represented by degrees azimuth varies as function of the elevation: e.g. 10° of azimuth at elevation 0° represents a greater distance along the sphere than 10° azimuth at say 60° elevation. To facilitate comparison the azimuth errors were corrected prior to plotting so that they equivalent to azimuth at elevation 0°. Positive errors for locations −180° to 0° indicate shifts in the perceived location away from the anterior midline while for locations 0° to 180° positive errors indicate shifts towards the anterior midline. These data are fairly mirror symmetrical about the anterior midline (azimuth 0°) with the exception that for the posterior hemisphere there is the suggestion of a slight overall anti-clockwise shift in the perceived locations for the middle three elevations. For locations in the anterior hemisphere there are very slight shifts in the azimuth of the perceived locations towards the anterior midline, although this is predominantly seen at the lowest elevations. There is a sharp shift in systematic errors about the interaural axis at all but the lowest elevations; locations on the interaural axis show a strong tendency to be perceived further behind the subject, particularly for the lower elevations.

The systematic errors in the elevation of the centroids of the perceived location were also determined for each elevation and plotted as a function of azimuth (Fig. 10). In this plot a negative error indicates a location estimate above the target location. In general, the centroids for the upper hemisphere locations were lower than the target locations and for the lower hemisphere were higher than the target locations. That is, there was a general trend to bias locations towards the audio-visual horizon. For locations on the audio-visual horizon (azimuth 0°) the centroids were above the target locations by a few degrees (1.3–4.4°).
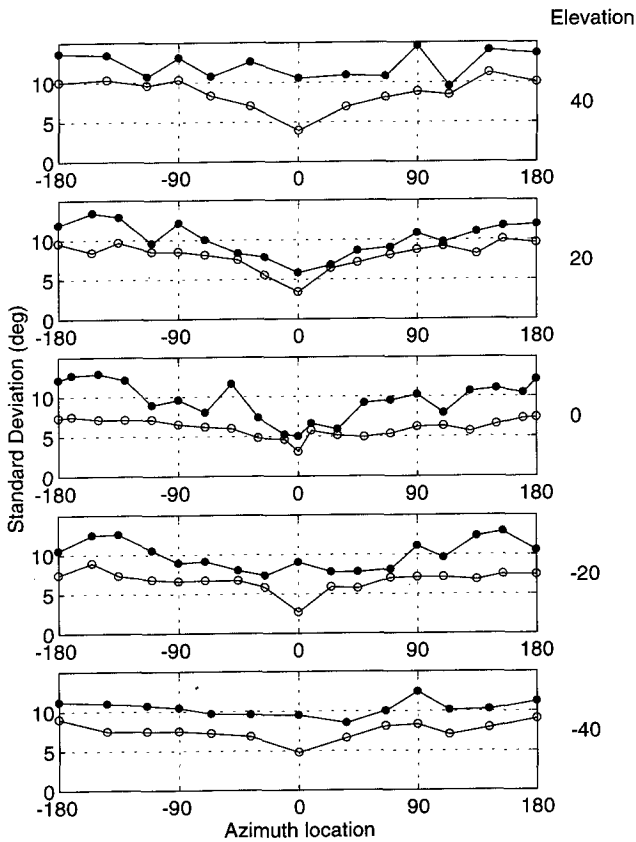


Fig. 11. The standard deviation of the localization errors associated with the first and second principal components of the distribution of the localization data are plotted as a function of target azimuth for each elevation used in this study (40°, 20°, 0°, −20° and −40°). The closed circles indicate the major axis in each distribution.
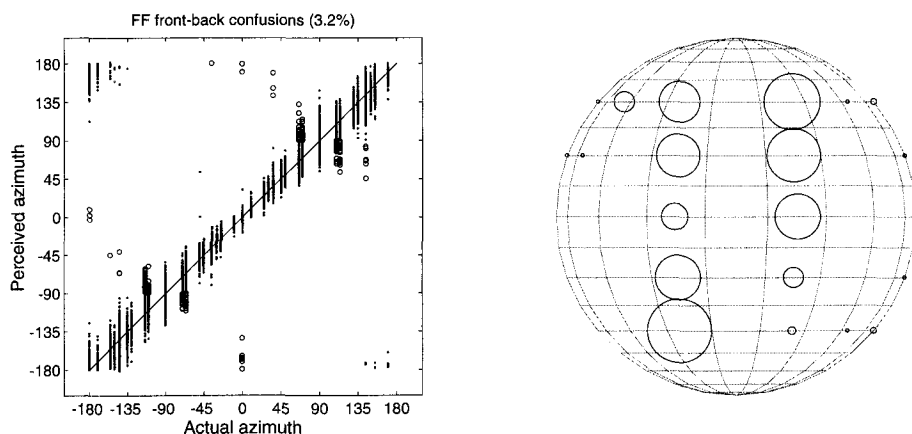
FF front-back confusions (3.2%)

Fig. 13. Left: The front-back confusion (open circles) and the remaining localization estimates (small crosses) have been collapsed across elevation and plotted as a function of actual target location. The solid diagonal line indicates a perfect response. Right: The target locations that resulted in a front-back localization estimate are indicated by the open circles, the radius of the circle indicates the relative number of front-back confusions for each location. The data for both left and right hemispheres have been collapsed on to one side. The filled circle indicates the position directly ahead of the subject (Az 0°/El 0°).

### 3.2.4. The distributions of errors about the population centroid

For our population of subjects, we have calculated the standard deviations of the localization estimates about the centroid for each test location (see above). For one third of the test positions the principal components of the distributions of localization estimates departed from the cardinal axes. As a consequence we chose not to summarize the error estimates for each location in terms of the azimuth and elevation components of the data. Rather, the spherical angular extents of the minor and major axes of the distributions of the data have been plotted as a function of azimuth (Fig. 11).

The smallest standard deviations in the pooled localization data are associated with the anterior median plane, on or just above the audio-visual horizon. At these locations the standard deviations are between 3° and 6°. Both the major and minor axes of the standard deviations of the distributions increase as azimuth locations move progressively towards the interaural axis. For azimuth locations progressively further behind the interaural axis, there are only relatively small increases in the length of the minor axis. However, the major axis of the distributions show a progressive increase in size over the same range of azimuth locations although they generally remain below 12° for elevations of −20°, 0° and 20°. The major axes errors are relatively large (10°) at elevations of −40° and 40° regardless of azimuth, although there are some azimuth dependent changes in the magnitude of the minor axes.

### 3.2.5. Spherical correlation as a metric of localization accuracy

A further method by which the accuracy of localiza-

tion can be expressed is using the spherical correlation between the actual and perceived locations of the test stimuli (see Wightman and Kistler, 1989). This measure collapses the data across the whole sphere of space and so acts as a global metric of localization accuracy. In this study the spherical correlation for the pooled data is 0.98 indicating a very high correspondence between actual and the average perceived location. It is instructive to calculate this value for each individual subject in this study. This allows us to look for systematic biases within individuals that result in lower individual spherical correlation coefficients that cancel out when the data is pooled across a number of individuals. The distribution of spherical correlation coefficients for our population of subjects is plotted as a histogram in Fig. 12. While this method may underestimate the systematic biases in our population this distribution indicates a range of localization skills in our population of 19 subjects.

### 3.2.6. Front-back confusions

The front-back confusion were extracted from the data prior to the calculation of the distribution of localization errors. A number of previous studies have documented the front-back confusion rate in localization data collected under similar conditions (Makous and Middlebrooks, 1990; Wightman and Kistler, 1989). A front-back confusion was defined as any localization estimate which represented an error reflected about the interaural axis. Locations on the interaural axis were not counted. Special care was taken to trap those front-back confusions where the actual targets were on or close to the midline but the perceived locations crossed the midline to the opposite hemisphere.

As it has been previously argued that front-back confusion errors are a special case of cone-of-confusion

errors it was decided to also examine the pooled data set for up-down errors. If there was a similar incidence of up-down errors this would be consistent with the idea of a generic cone-of-confusion error. Alternatively, a lack of similarity might suggest that front-back confusion and up-down confusion errors are dependent on different processes, or that there was some fundamental differences in the cues available for resolving cone of confusion errors over these different segments of space. Up-down confusion errors were defined as any localization error that crossed the audio-visual horizon. Similar to the front-back confusion errors, locations on the audio-visual horizon were not counted. The overall error rates for these two types of error differed markedly. There were 3.2% front-back confusion errors but, to our surprise, **no** up-down confusions using this stimulus.

The front-back confusion errors have been plotted in two ways (see also Makous and Middlebrooks, 1990). By collapsing across elevation, a general indication of the incidence can be obtained (Fig. 13). The greatest
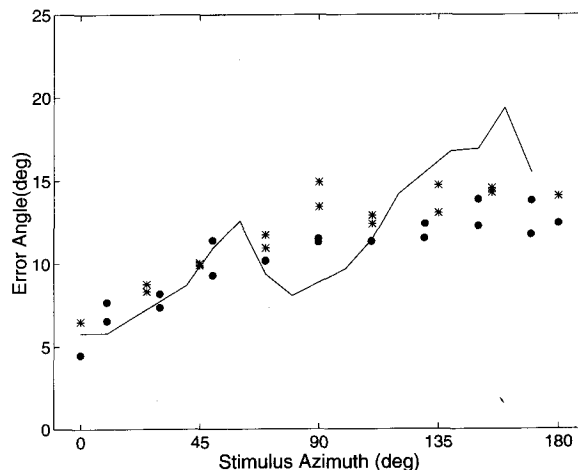


Fig. 15. The average spherical errors have been calculated for locations on the audio-visual horizon for the right and left hemispheres (filled circles) and for locations 20° above the horizon (*). For comparison, the solid line indicates the mean spherical errors obtained at +5° and +15° above the audio-visual horizon by Makous and Middlebrooks (1990), their Figure 7b.

majority of these errors can be seen to have occurred for locations within 30° of the vertical plane containing the interaural axis. Relatively few of these confusions have occurred for locations on the median plane. These data have also been plotted in terms of the original locations of the target stimuli (Fig. 13). The majority of front-back confusions occurred for locations at extremes of elevation.

## 4. Discussion

### 4.1. Head pointing to indicate perceived location

A fundamental problem in many areas of psychophysical research, including auditory localization, is the selection of the response measure employed to indicate the perceptual phenomena under study. Any response measure will have associated with it a variance which will contribute to the overall measure of performance. The challenge is to develop methods that reduce that component of the response variance and which provide some estimate of that variance.

In this study we have used head pointing as the method by which the subject indicates the perceived location of the stimulus. Any measure employing a pointing response is, to some extent, likely to be confounded by (a) errors in the motor component of the measured behavior and (b) variations in the time between the stimulus and the completion of the response for different spatial locations. Locations requiring large movements from the start position could show greater motor and memory related errors. The smaller localization errors evident in the frontal field are consistent
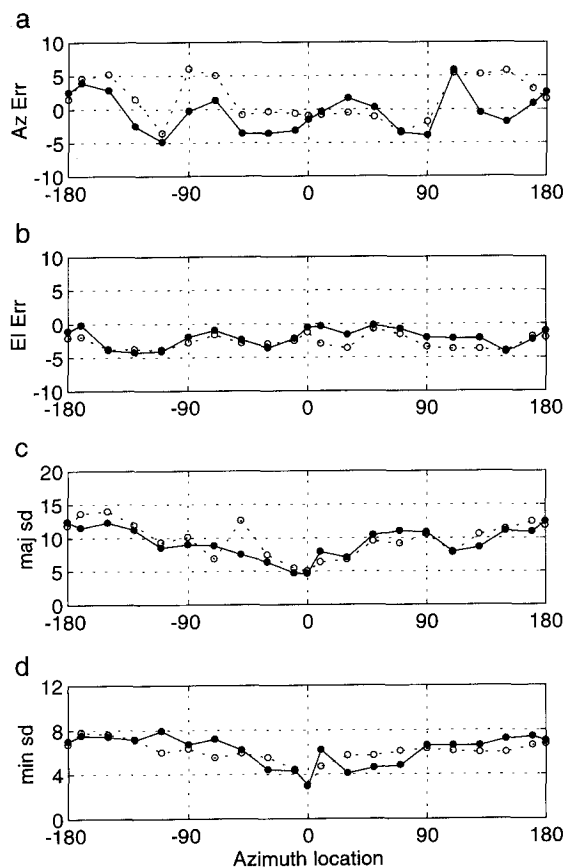


Fig. 14. The systematic errors in azimuth and elevation and the standard deviations about centroids are plotted for the subjects who underwent head pointing training on a total of 324 training locations (filled circles and solid lines) compared with the group who were trained on 180 or less positions (open circles dotted lines). Systematic errors in (a) azimuth and (b) elevation and (c) major and (d) minor axes of the distributions about the centroid.

with this argument (Figs. 6, 9 and 10). However, these data are qualitatively similar to other studies which have not required a gross motor response. For instance, Wightman and Kistler (1989) required their subjects to call out numerical estimates of the perceived location. In a study by Hammershoi and Sandvad (1994) the subjects indicated perceived location of a target using a stylus and a computer input tablet. Likewise, Good and Gilkey (1996), Gilkey et al. (1995) and Gilkey and Anderson (1995) required subjects to indicate the location of a stimulus using a stylus and a sphere representing the surrounding sound field. Despite some difficulties in making quantitative comparisons between studies, some general observations can be made. Firstly, in each of these previous studies the magnitude of the errors for each location was larger than those found in this study or the previous studies employing head pointing (Makous and Middlebrooks, 1990). Secondly, these studies all report the same qualitative change in the magnitude of the localization errors, i.e. small in the frontal field increasing to larger errors behind. This suggests that, for the large part, the spatially dependent variations in the localization performance found in the studies using head pointing probably reflect sensory rather than motor or memory related procedures.

### 4.2. Training effects of immediate cognitive feedback

The principal advantages of using head pointing as a means of indicating perceived location are that tracking the head position is straightforward and turning to face a sound source is seen as an ecological response to a sound. The principal disadvantage is that subjects also tend to move their eyes in addition to their heads to visually 'capture' the perceived location. This is most evident when the subject is required to move to the mechanical extremes of the head's movement (e.g. extremes of elevation). We have attempted to remove the 'eye pointing' errors as much as possible by training the subjects and providing a variety of feedback and encouragement. We have also examined the effects of providing both sensory and immediate cognitive feedback as to the accuracy of head pointing during training. The provision of accuracy feedback results in faster acquisition of the task and a greater overall accuracy in head pointing to the target (Fig. 4). Using the closed loop cognitive paradigm (CLC) subjects' open loop performance had generally plateaued by the third training block. Compared to the closed loop sensory condition, accuracy was also significantly greater as measured by the spherical correlation coefficient of the actual and perceived locations. There was a sharp jump in performance of subjects who had plateaued on the CLS condition when they were switched to the CLC condition.

As part of the training procedure, subjects were also

required to point their head to a light positioned at the center of the speaker. This provided an estimate of how much of the residual error in the open loop condition might be accounted for by eye pointing (Fig. 4B). In the case of the subjects receiving only sensory feedback during training, the spherical correlation for the light pointing was significantly lower than those in the CLC condition. This indicates that, despite the sensory feedback provided during training, these subjects had failed to eliminate eye pointing from their responses despite being strongly encouraged to do so. As a result it is likely that a significant fraction of the errors evident in the open loop test during subject training can be accounted for by eye pointing errors. All subjects employed in the second part of this study were trained using the closed loop cognitive feedback paradigm until the spherical correlation of the localization performance under the open loop condition plateaued. In nearly every case, this had been achieved within three to five training sessions.

Makous and Middlebrooks (1990) have also used head pointing as a means of indicating perceived location. Prior to testing, subjects were given 10–20 training sessions (83 trials each) which included visual feedback using a LED at the center of the target speaker. This training condition is most closely related to the visual feedback condition used in the present study but without subsequent closed loop auditory feedback. Closed loop performance was also measured by Makous and Middlebrooks for three subjects. The variations in the responses for the one subject illustrated are relatively small (2–3° in azimuth and 2–4° in elevation). However, the systematic errors in localization are consistent with a residual eye pointing errors, particularly for locations at the extremes of elevations in the posterior hemisphere (Makous and Middlebrooks, 1990; their Figure 3f).

The choice of stimulus was guided by the fact that broadband, spectrally stable stimuli are most easily localized (Butler, 1986; Makous and Middlebrooks, 1990; Middlebrooks, 1992). This is not surprising as such a stimulus is identical in the frequency domain to an impulse for which there is considerable evolutionary pressure to localize accurately (for discussion see Erulkar, 1972; Carlile and Pettigrew, 1987; Carlile, 1996a). In the present study, an important assumption of our approach was that, as the target stimulus was easily localizable, then the principal effects of the subject training was to increase the accuracy by which subjects indicated the perceived location rather than to increase the localization accuracy per se, i.e. that the training principally reduced the error associated with head and eye pointing rather than decreased the error associated with the sensory component of the task. Alternatively, the feedback as to actual target location provided during the course of the head pointing train-

ing may have allowed the subjects to correct for miss-perception of the location and have an impact on the subsequently measured localization accuracy. We can estimate some of the possible training effects on the sensory component of the task by comparing the localization performance of those subjects who had received the full nine blocks of training with those who had only received four or five blocks of training. The azimuth and elevation systematic errors and the standard deviations were calculated for each of these two groups and have been plotted for locations on the audio-visual horizon (Fig. 14). In most cases the centroids of these two groups are within 2–3° of each other and there are no clear systematic differences between them. There is some suggestion that for azimuth errors at this elevation the subjects with less training do better for the anterior hemisphere but poorer for the posterior hemisphere although this is not seen systematically for the other elevations (data not show). The failure to see a generalized improvement in performance suggests that the training in head pointing had not systematically altered the ability of the subjects to localize the target. However, it might be further argued that any learning effects may have saturated by the fourth or fifth training block. To examine this possibility we compared the average spherical error associated with those locations where the nine block group had received training but where the less trained group had not. That is, the nine block group had one experience of each of the 76 test locations in the course of exposure to the full 324 training locations, whereas, for the group that had received five or fewer blocks of training, there were 33 locations in the test set for which they had never previously experienced the target stimulus. Should the head pointing training have significantly improved subsequent localization performance then the average spherical error should be less for those who had received training on the test locations compared to those who had not. A statistical comparison indicates that these two populations were identical in terms of the average spherical error ($t$-test for paired comparisons: $t = 0.458$; df 32; $0.9 > \alpha > 0.5$). Therefore, although we cannot completely rule out the possibility that head pointing training had some kind of effect on the subsequent perception of the location of the broadband stimulus, we interpret the statistical equivalence above as providing good evidence that head pointing training had no significant effect on the sensory component of the task.

### 4.3. The accuracy of auditory localization of a brief broadband sound

There are two principal considerations with respect to describing the accuracy of localization performance. First the correspondence between the mean perceived location and the actual target location and second,

the distribution of the perceived locations about the mean perceived location. In this study the centroid of the cluster of responses associated with each actual target location was determined using methods appropriate for data distributed on a sphere (Fisher et al., 1987; Leong and Carlile, 1997). This calculation does not make any assumptions about rotational symmetry of the distributions of the data. In addition, to ensure that the data were unimodal and the centroids provided a straight forward interpretation the front-back confusions were removed from the pooled data prior to this calculation. By convention, spatial locations have been described by using coordinates of azimuth and elevation and for convenience this approach is also adopted here. However, this does not mean that we consider the underlying processes to be constrained to these axes. On the contrary, the orientations of the elliptical distributions of errors for a significant fraction of the locations tested suggest that the processes underlying localization are not constrained to the cardinal axes of our coordinate system (see below).

A number of systematic localization errors were evident in the pooled data in this study. Average miss-localizations were generally of the order of 3° degrees in azimuth and 4° in elevation. They were smallest for locations in the anterior hemisphere, particularly those associated with the audio-visual horizon. In general, for locations off the audio-visual horizon the mean estimates were shifted towards the audio-visual horizon, particularly for the highest elevation and the two lowest elevations (Figs. 6 and 10). Additionally, there were posterior shifts in the mean locations for targets around the interaural axis for both right and left hemispheres (Figs. 6 and 9). There was also the suggestion of an anti-clockwise shift in the perceived locations, particularly for locations in the posterior hemisphere (Fig. 6).

What is not clear is whether these systematic errors represent sensory distortions of auditory space (a decrease in the accuracy of spatial representations for these relative locations) or if they have their origins in the acoustic environment or the response measure. For instance, the systematic errors in elevation could simply represent the increasing recruitment of eye movements to 'capture' the auditory target at locations where head pointing approaches the comfortable mechanical limits of movements of the head. An analysis of the visual-only data in the training is consistent with this interpretation (data not shown): i.e. following training which resulted in a plateau of the spherical correlation coefficients for the open loop responses, some residual eye pointing errors were still evident. However, the pattern of systematic errors in the azimuth dimensions can not be as easily interpreted in this way. The small anti-clockwise rotation shift (Figs. 6 and 9) may have reflected some asymmetry in our measurement environment. In the face of this result we recalibrated the posi-

tion of the handrail on the subject support platform and found that it was rotated 3.7° clockwise with respect to the hoop coordinate system. This is a very similar sized shift as the anti-clockwise shift in the perceptual data and is in the appropriate direction. There are two small features on the front and back of the hand rail of the support that may have influenced some of the subjects in their front-back alignment.

### 4.4. The distribution of the localization errors

In this study we have shown that two thirds of the pooled localization data were more appropriately modelled using an elliptical distribution. As these data are not rotationally symmetric, then descriptive statistics assuming a normal or at least symmetric distribution need to be handled with care. The first two principal components of about 50% of the elliptical distribution were found to deviate from the cardinal axes of the coordinate system by more than 10°. The application of single variable descriptive statistics to such data will result in significant errors in the estimates of the parameters of the distributions and so spherical statistical approaches are indicated.

The preponderance of elliptical distributions in the pooled data in this study is consistent with previous studies where auditory localization accuracy was estimated using target locations varied in azimuth and elevation. Makous and Middlebrooks (1990) report larger standard deviations in elevation than in azimuth for locations about the anterior midline and the reverse for locations in the lateral hemispheres. Similar comparisons are difficult to make with two other similar studies because of the differences in the methods employed in illustrating the data (Wightman and Kistler, 1989; Oldfield and Parker, 1984a). However, those data also show the same qualitative changes in the spatial dependence of localization data: that is, greatest accuracy and smallest distributions for locations about the anterior midline close to the audio-visual horizon and an increase in the systematic errors in both elevation and azimuth for target locations in the posterior and upper quadrant. The largest errors were associated with locations close to the posterior midline and above and below the horizon (cf. Figs. 6, 9 and 10).

The use of spherical statistics in describing these data also has a further analytical advantage: namely that the calculation of the parameters of the distributions of the estimates is not tied to the spherical coordinate system used for describing the spatial location of the data. One of the assumptions in these kinds of analysis is that the distribution of errors in localization to some extent reflects errors in the processing accuracy of the underlying sensory processes. In addition, the errors must also reflect spatially dependent changes in the nature of the physical cues to a sounds locations. On the one hand, it

is unlikely that the resolution of the neural processes involved in encoding the cues to a sounds location vary in a spatially dependent manner. Rather, it is the nature of the physical cues to a location that vary. The resolving power of the neural processes may vary as a function of the magnitude of a particular parameter. (For instance, sensitivity to ILD varies with the absolute magnitude of the ILD (Hafter et al., 1977) and it is this that is dependent on the nature and location of the stimuli.

If we concentrate on an analysis of the locations on the audio-visual horizon, the axes of the distributions of errors were roughly parallel to the coordinate system. In this case we are able to directly compare these data with the data obtained in the study by Makous and Middlebrooks (1990). One distribution free measure of localization error is the average angular error of localization (Fig. 15). This is determined from the average of the spherical angles between the perceived locations and the actual target location. This measure is comparable with the data of Makous and Middlebrooks (1990) who calculated the mean unsigned magnitude of errors as a spherical angle subtended at the center of the subjects head. Fig. 15 shows that for the anterior locations on the audio-visual horizon there is good correspondence between the data in these two studies. However, for locations around the interaural axis the spherical angular error is slightly larger in the current study. This probably reflects the way in which Makous and Middlebrooks (1990) dealt with the front back confusion in the analysis of their data to allow comparison with an earlier study (Makous and Middlebrooks, 1990; their Figure 7b): namely that the front-back confusions for target locations at azimuth 80° and 100° were 'resolved' before the average localization errors were calculated. In the present study, localization errors on the horizontal plane including the interaural axis demonstrate a mean standard deviation for the major axis of around 11° (Fig. 11). One possibility acknowledged by Makous and Middlebrooks (1990; p. 2196) is that the localization error may be underestimated in this form of analysis as locations that are 'resolved' as front-back confusion may indeed represent local errors that have simply crossed the interaural axis. Some reduction in the calculated local errors following removal of the front-back confusion errors might also be a problem in the present study, although to a much smaller extent because of the larger spacing of the test locations around the interaural axis (±20°). For locations in the posterior hemisphere, the localization errors were smaller in this study than that observed by Makous and Middlebrooks (1990) for the same area of space (Fig. 15). This undoubtedly reflects differences in the way the subjects were able to indicate the location of the stimulus. In the present study, subjects stood in the center of the anechoic chamber and were encouraged to

turn their whole body towards the stimulus. In the former study, the subjects were seated and needed to twist around to face the target. The movement restriction and the complex rotational translations are likely to increase the error in the motor component of the response for these locations.

## 4.5. Implications for the processing of localization cues

If the analysis of these kinds of data is constrained to a particular coordinate system this implies that the underlying processes share that coordinate system. For instance, a double pole coordinate system has been usefully employed in the analysis of front-back confusion errors. A front-back reversal is thought to result from the failure to resolve the cone of confusion using the spectral information provided by the filtering of the sound by the outer ear (Carlile and Pralong, 1994; Carlile and King, 1994; Middlebrooks, 1992). In the case of the double-pole system a cone of confusion corresponds to the azimuth coordinates specified by the lateral poles of the coordinate system. In an elegant combination of psychophysical experimentation, bioacoustics and numerical modelling Middlebrooks (1992) has provided strong evidence for ILD and spectral processing acting in this complimentary manner. Those experiments employed narrowband stimuli and clearly demonstrate that, under these conditions, the spatial location can be specified on the basis of the cone of confusion determined by the ILD and that the location on the cone of confusion is determined by a spectral analysis of the stimuli. Such data provide important insights into the way in which localization cues are processed within frequency bands.

A close examination of the pattern of distribution of errors for target locations about the anterior midline (Fig. 6) is consistent with the idea that two different processes might underlie the processing of localization cues for this region of space. In terms of the model described above the ILDs might determine the distributions of errors in the horizontal dimension and that errors in spectral analysis determines the vertical dimension of the distribution. If this was the case then the errors associated with an analysis of ILD would seem to result in smaller horizontal errors than the elevation errors resulting from an analysis of the spectral patterns. Furthermore, the horizontal errors on the audio-visual horizon increase with locations towards the interaural axis which is also consistent with the finding that the sensitivity to ILDs is greatest for those around zero (Hafter et al., 1977). An analysis of the spectral features of the filter functions of the outer ears for locations on the frontal midline indicate that the rate of the spatially dependent changes in the perceptually salient components are not high for this area of space (Carlile and Pralong, 1994). That study dem-

onstrates that for anterior midline stimuli, there is an elevation dependent change in the peak of the excitation pattern produced by a broadband stimulus. The predicted excitation peak shifts from around 3.5 kHz at elevation $-30°$ to around 5 kHz at elevation $+30°$. There is also a relative decrease in the power in the frequency band 9–10 kHz over the same range of elevations. While this analysis does not demonstrate a necessary role for the spectral information in these frequency ranges, it does provide a framework for subsequent psychophysical investigation.

The distributions of localization errors for locations off the midline are not as easily interpreted using a model of orthogonal processing of interaural and spectral cues. If, say, the major axes of the distributions indicate errors in the processing of spectral cues and the minor axes of the errors indicate processing of errors in the binaural cues, then orientation of the major axes of the distributions should follow the cones of confusion. That is, they should be arranged in concentric rings centered roughly on the interaural axis (see Middlebrooks, 1992). However, in the lateral and posterior hemispheres the major axes are generally arranged parallel to the lines of co-latitude. In the lateral and frontal hemispheres the major axes are arranged radially to the point directly ahead at azimuth $0°$/elevation $0°$. The orientation of the principal components of the distributions may be dependent on the motor programs involved in head pointing to the frontal hemisphere or on the spatial distributions of the cues to sound location or some other combination of the two. These later effects present a relatively non-trivial analytical problem and are currently the subject of a modeling exercise in our group (see Chung et al., 1996).

An important and related question is how the relative sets of narrowband binaural and broadband spectral cues are employed in localizing a broadband stimulus. The set of cues available for the localization of broadband stimuli are much richer than those available for narrowband stimuli. While this should not be construed as a criticism of the approaches discussed above it does indicate the need for caution in extending the results using simplified stimuli to spectrally and/or temporally complex stimuli. Where a particular stimulus provides a rich set of cues, it is possible that the auditory system may weight processing strategies in terms of the information content in the sound. If this were the case then the characteristics of the errors in localization may vary in a subtle and cue dependent manner. For instance, two different kinds of processing strategies have been identified in the processing of binaural signals depending on their duration. Very short stimuli are analyzed analytically while longer duration stimuli are analyzed synthetically (Dye et al., 1994). Variations in processing strategies might also extend to other types of binaural cues as well as the duration of the signals. The

large data set and the robust analytical routines reported in this paper provide a baseline for an examination of the effects of various types of stimuli constructed to probe the processing of the cues to a sounds location.

## Acknowledgments

## References

Butler, R.A., 1986. The bandwidth effect on monaural and binaural localization. Hear. Res. 21, 67–73.

Butler, R.A., Belendiuk, K., 1977. Spectral cues utilized in the localization of sound in the median sagittal plane. J. Acoust. Soc. Am. 61, 1264–1269.

Butler, R.L., Humanski, R.A., 1992. Localization of sound in the vertical plane with and without high-frequency spectral cues. Percept. Psychophys. 51, 182–186.

Butler, R.A., Humanski, R.A., Musicant, A.D., 1990. Binaural and monaural localization of sound in two- dimensional space. Perception 19, 241–256.

Carlile, S., 1996a. Auditory space. In: Carlile, S. (Ed.) Virtual Auditory Space: Generation and Applications. Landes, Austin, TX, Ch. 1.

Carlile, S., 1996b. The physical and psychophysical basis of sound localization. In: Carlile, S. (Ed.) Virtual Auditory Space: Generation and Applications. Landes, Austin, TX, Ch. 2.

Carlile, S. (Ed.), 1996c. Virtual Auditory Space: Generation and Applications. Landis, Austin, TX.

Carlile, S., King, A.J., 1994. The integration of monaural and binaural cues in the generation of a neural representation of auditory space. Proc. Aust. Neurosci. Soc. 5, 31.

Carlile, S., Pralong, D., 1994. The location-dependent nature of perceptually salient features of the human head-related transfer function. J. Acoust. Soc. Am. 95, 3445–3459.

Chung, W., Carlile, S., Leong, P., 1996. A computational model for auditory localization. Australian Conference on Neural Networks, Canberra, pp. 150–154.

Dye, R.H., Yost, W.A., Stellmack, M.A., Sheft, S., 1994. Stimulus classification procedure for assessing the extent to which binaural processing is spectrally analytic or synthetic. J. Acoust. Soc. Am. 96, 2720–2730.

Erulkar, S.D., 1972. Comparative aspects of spatial localization of sound. Physiol. Rev. 52, 238–360.

Fisher, N.I., Lewis, T., Embleton, B.J.J., 1987. Statistical Analysis of Spherical Data. Cambridge University Press, Cambridge.

Gilkey, R.H., Anderson, T.R., 1995. The accuracy of absolute localization judgements for speech stimuli. J. Vestib. Res. 5, 487–497.

Gilkey, R.H., Good, M.D., Ericson, M.A., Brinkman, J., Stewart, J.M., 1995. A pointing technique for rapidly collecting localization responses in auditory research. Behav. Res. Methods Instruments Comput. 27, 1–11.

Good, M.D., Gilkey, R.H., 1996. Sound localization in noise: The effect of signal-to-noise ratio. J. Acoust. Soc. Am. 99, 1108–1117.

Hafter, E.R., Dye, R.H., Nuetzel, J.M., Aronow, H., 1977. Difference thresholds for interaural intensity. J. Acoust. Soc. Am. 61, 829–834.

Hammershoi, D., Sandvad, J., 1994. Binaural Auralization. Simulating Free Field Conditions by Headphones. Audio Engineering Society, Amsterdam, pp. 1–19.

Hartmann, W.M., 1983. Localization of sound in rooms. J. Acoust. Soc. Am. 74, 1380–1391.

Kent, J.T., 1982. The Fisher-Bingham distribution on the sphere. J. R. Statist. Soc. B 44, 71–80.

Leong, P.H.W., Carlile, S., 1997. Methods for spherical data analysis and visualisation. J. Neurosci. Methods (submitted).

Makous, J., Middlebrooks, J.C., 1990. Two-dimensional sound localization by human listeners. J. Acoust. Soc. Am. 87, 2188–2200.

Middlebrooks, J.C., 1992. Narrow-band sound localization related to external ear acoustics. J. Acoust. Soc. Am. 92, 2607–2624.

Oldfield, S.R., Parker, S.P.A., 1984a. Acuity of sound localization: a topography of auditory space I. Normal hearing conditions. Perception 13, 581–600.

Oldfield, S.R., Parker, S.P.A., 1984b. Acuity of sound localization: a topography of auditory space II: Pinna cues absent. Perception 13, 601–617.

Oldfield, S.R., Parker, S.P.A., 1986. Acuity of sound localization: a topography of auditory space. III Monaural hearing conditions. Perception 15, 67–81.

Perrett, S., Noble, W., 1995. Available response choices affect localization of sound. Percept. Psychophys. 57, 150–158.

Wightman, F.L., Kistler, D.J., 1989. Headphone simulation of free field listening. II: Psychophysical validation. J. Acoust. Soc. Am. 85, 868–878.

Carlile, S., Pettigrew, A.G., 1987. Distribution of frequency sensitivity in the superior colliculus of the guinea pig. Hear. Res. 31, 123–136.