Available online at www.sciencedirect.com

SCIENCE DIRECT®

Hearing Research 199 (2005) 124–134

HEARING RESEARCH

www.elsevier.com/locate/heares

Human sound localization at near-threshold levels

Andrew T. Sabin, Ewan A. Macpherson, John C. Middlebrooks *

Central Systems Laboratory, Kresge Hearing Research Institute, 1301 E. Ann Street, Ann Arbor, MI 48109-0506, USA

Received 21 May 2004; accepted 1 August 2004
Available online 6 October 2004

Abstract

Physiological studies of spatial hearing show that the spatial receptive fields of cortical neurons typically are narrow at near-threshold levels, broadening at moderate levels. The apparent loss of neuronal spatial selectivity at increasing sound levels conflicts with the accurate performance of human subjects localizing at moderate sound levels. In the present study, human sound localization was evaluated across a wide range of sensation levels, extending down to the detection threshold. Listeners reported whether they heard each target sound and, if the target was audible, turned their heads to face the apparent source direction. Head orientation was tracked electromagnetically. At near-threshold levels, the lateral (left/right) components of responses were highly variable and slightly biased towards the midline, and front vertical components consistently exhibited a strong bias towards the horizontal plane. Stimulus levels were specified relative to the detection threshold for a front-positioned source, so low-level rear targets often were inaudible. As the sound level increased, first lateral and then vertical localization neared asymptotic levels. The improvement of localization over a range of increasing levels, in which neural spatial receptive fields presumably are broadening, indicates that sound localization does not depend on narrow spatial receptive fields of cortical neurons.

Keywords: Sound localization; Psychoacoustics; Spatial coding; Auditory cortex

1. Introduction

Human listeners rely on multiple acoustical cues to determine the location of a sound source (see Middlebrooks and Green, 1991 for a review). The differential path from a sound source to the two ears results in interaural differences in arrival time (interaural time difference; ITD), and sound level (interaural level difference; ILD). The auditory system mainly uses these two cues to determine the lateral (left/right) location of a sound. Location in the polar (up/down and front/back) dimension is determined by the location-specific spectral filtering caused by a sound's interaction with the head and outer ears.

Whereas the acoustical cues used in sound localization are well known, the representation of sound location in the central nervous system is still debated. One hypothesis that has motivated several physiological studies is that locations in auditory space are represented by spatially tuned neurons, each neuron representing a particular location (e.g. Middlebrooks and Pettigrew, 1981; Imig et al., 1990; Rajan et al., 1990; Brugge et al., 1996; Reale et al., 2002). In the cat, neurons in the auditory cortex show fairly restricted spatial tuning at low sound levels. For example, one such unit (Fig. 1) exhibited spatial tuning to −60° azimuth at a sound level 20 dB above the neuron's threshold. However, as sound level increases, that tuning broadens.

Abbreviations: ITD, Interaural time difference; ILD, Interaural level difference; DTF, Directional transfer function; HRTF, Head-related transfer function; MAF, Minimum audible field; SPL, Sound pressure level; SL, Sensation level; MAA, Minimum audible angle

* Corresponding authors. Tel.: +1 734 764 5167; fax: +1 734 764 0014.

E-mail address: jmidd@umich.edu (J.C. Middlebrooks).

0378-5955/$ - see front matter © 2004 Elsevier B.V. All rights reserved.
doi:10.1016/j.heares.2004.08.001

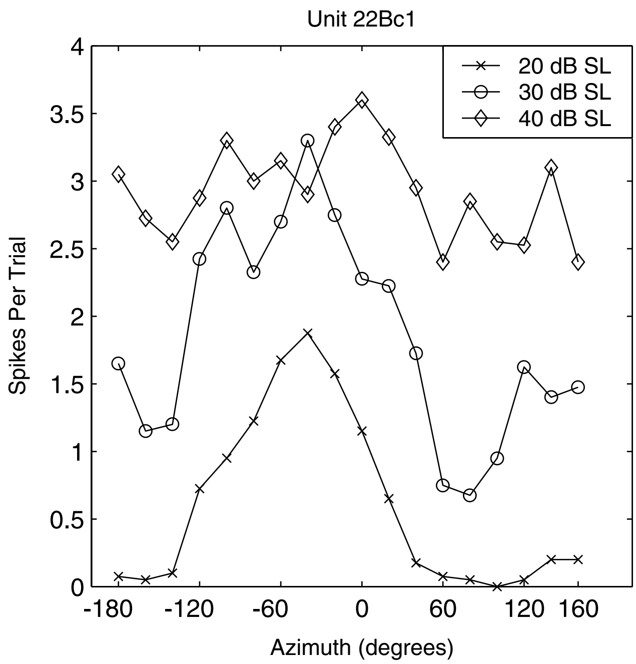

Fig. 1. Spatial tuning of a single neuron in the cat primary auditory cortex (area A1). The symbol types indicate sound level relative to the neuron's threshold. As sound level increases, the neuron's spatial tuning broadens (unpublished data from Stecker et al., 2003).

Any model relying on neurons with restricted spatial receptive fields would predict that sound localization will be most accurate at near-threshold sound levels and should *degrade* with increasing level. Contrary to this (counter-intuitive) prediction, sound localization accuracy generally seems to *improve* with increasing sound level.

There have been relatively few psychophysical studies focusing on the level dependence of sound localization. Altshuler and Comalli (1975) and Comalli and Altshuler (1976) tested the effect of stimulus level on lateral localization in the horizontal plane. The left-right position of a narrow-band noise source was manipulated, and blindfolded subjects were instructed to classify the sound's location as being centered on, or to the left or right of the vertical midline. The range of locations incorrectly classified as "centered" increased as sound level decreased. This observation suggests that lateral localization accuracy declines with decreasing sound level. There also have been several studies that suggest that localization in the vertical (polar) dimension worsens with decreasing sound level. Both Davis and Stephens (1974) and Hebrank and Wright (1975) studied polar localization by playing broadband noise bursts through one of nine loudspeakers located on the front portion of the vertical midline. The listener's task was to identify the loudspeaker from which the sound had originated. Both studies found that listeners reported an incorrect loudspeaker more often at lower sound levels. Inoue (2001) conducted a similar identification task in which broadband noise was played though a loudspeaker located on either the horizontal plane or the upper half of the vertical median plane. Subjects reported incorrect loudspeakers more often in both dimensions at lower sound levels. Vliegen and van Opstal (2004) used a head-pointing task to study the influence of sound level and duration on localization. At lower levels, vertical localization judgments decreased in accuracy with decreasing sound level down to the lowest tested level (26 dB SPL).

Su and colleagues tested the effect of level on threshold for detecting a change in location both in monkeys (Su et al., 2000) and in humans (Su and Recanzone, 2001). Their procedure yielded a form of minimum audible angle (MAA), which is the just noticeable difference in sound source location. In these experiments, MAA was calculated for each subject at a variety of levels independently in the lateral and vertical dimensions. This method allowed for a direct comparison of lateral- and vertical-angle performance in the same listener. MAA broadened in both lateral and vertical dimensions as sound level decreased, although that broadening occurred at a higher level in the vertical dimension. This observation suggests that vertical localization is more vulnerable to low sound levels than is lateral localization. The greater degradation of vertical localization at low levels can also be seen in the data presented by Inoue (2001). Su and Recanzone (2001) proposed that the spectral cues necessary for vertical localization were inaudible for sound levels at which the binaural cues necessary for lateral localization were audible, thus creating the lateral/vertical discrepancy.

For certain stimuli, increases in sound level have a negative effect on localization performance. Several studies have shown that vertical sound localization accuracy decreases at high sound levels for very brief stimuli such as 3-ms noise bursts or clicks (Hartmann and Rakerd, 1993; Hofmann and van Opstal, 1998; Macpherson and Middlebrooks, 2000; Vliegen and van Opstal, 2004). That "negative effect of level" has been attributed variously to insufficient temporal integration (Vliegen and van Opstal, 2004) or cochlear distortion (Macpherson and Middlebrooks, 2000). Since the present study used longer duration (250 ms) stimuli, no negative effect of sound level was expected.

All of the previously mentioned studies suggest that sound localization is degraded at low levels. Aside from Vliegen and van Opstal (2004), however, they cannot be regarded as tests of localization per se but rather of discrimination. MAA studies merely tested the detectability of any change in the perceived sound, which may or may not reflect a change in the perceived location. The loudspeaker identification tasks limited the number of response choices, and it has been shown that doing so can bias the listener's response (Perrett and Noble, 1995).

Most of the studies that have looked at the effects of sound level on localization tested a relatively small range of sound levels, and only one (Inoue, 2001) tested near detection threshold. By limiting the stimuli to a relatively small range of sound levels, the investigators may have analyzed data that did not fully reveal the level-localization relationship. In addition, the previously mentioned studies tested lateral and vertical dimensions independently with stimuli that originated on either the horizontal plane or the vertical midline. In the present study we attempted to overcome these limitations by using a free response localization task with target locations distributed throughout the full range of auditory space, and by employing a wide range of sound levels extending down to the detection threshold.

2. Materials and methods

2.1. Participants

Six (one male and five female) University of Michigan students aged 19–23 years were paid for their participation. All listeners had hearing thresholds within 10 dB of audiometric normal for pure tones at octaves from 0.25 to 8 kHz. None of the subjects had previous experience in psychoacoustic experiments. Prior to data collection, listeners were required to participate in three 90-min training sessions in which they learned the task and demonstrated localization proficiency for all regions of auditory space at supra-threshold intensities. All procedures used in this study were approved by the University of Michigan's Institutional Review Board for Human Research (IRBMED).

2.2. Stimuli

All sounds presented in this experiment were 250-ms, broadband (0.5–16 kHz), flat-spectrum, random-phase noise bursts with 20-ms raised cosine onset and offset ramps. The stimuli were synthesized at a sampling rate of 50 kHz, output through a digital-to-analog converter (Tucker-Davis Technologies [TDT] DD1), attenuated (TDT PA4), and amplified (Adcom GFA-535II). That signal was then routed to one of two outputs through a power multiplexer (TDT PM1) and transduced by one of two loudspeakers (Infinity 32.3 CF). All stimuli were filtered by the inverse of the loudspeaker's transfer function, which removed the spectral characteristics of the loudspeaker from the acoustic stimulus (see Macpherson and Middlebrooks (2000) for details). The loudspeakers resided on either side of a 1.2-m radius hoop that was covered in sound-absorbing foam. That hoop was attached to computer-controlled stepping motors which could manipulate both horizontal and vertical position. The motors were covered with sound-absorbing foam to eliminate reflections. Motor noise was audible while the hoop was moving, however it gave the listener no cue as to the hoop's orientation because the sound always came from the same overhead location and because the trials were ordered so that the hoop moved for approximately the same duration between all trials.

2.3. Detection thresholds

Detection thresholds were measured at eight evenly spaced locations around the horizontal plane. First, the listener indicated when a sound, which began at 0 dB SPL and increased by 5 dB with each presentation, became audible. Following this, the listener began a three-down, one-up two-interval forced choice procedure (Levitt, 1971) starting at the indicated level and continuing until six reversals had occurred. For the first two reversals, the step size was 5 dB and for the subsequent four, the step size was 3 dB. The levels at which the last four reversals occurred were averaged. This process was repeated five times at each location. The highest and lowest thresholds were discarded, and the mean of the remaining three was computed. The resulting intensity was called the detection threshold. All sensation levels are expressed relative to the detection threshold for a source positioned straight in front of the listener.

2.4. Localization task

Each block of trials in the localization task contained either low- (0, 5, 10, 15, or 20 dB SL) or high- (20, 30, 40, 50, or 60 dB SL) level stimuli. The repetition at 20 dB SL allowed for a comparison of performance with the same stimulus between the two conditions. Stimuli were presented from 200 locations distributed roughly evenly on the surface of an imaginary sphere centered on the listener's head (radius = 1.2 m). Spatial location was quantified using the horizontal polar coordinate system (Fig. 2). The locations were approximately 10–14° apart. Each level-location combination was tested once for a total of 2000 trials. The low-level stimuli were divided randomly into 13 blocks of 70 trials and one block of 90 trials. The same was done to the high-level stimuli for a total of 28 blocks. The listener generally alternated between low- and high-intensity trial blocks in one session.

All testing was conducted in a darkened anechoic chamber ($2.6 \times 3.7 \times 3.2$ m). The listener stood on a platform that was adjusted in height to place his/her head in the center of the hoop. Head orientation was measured by an electromagnetic head tracker (Polhemus FASTTRAK) that was mounted on top of a cap worn by the listener.

Each trial began with the illumination of a light-emitting diode (LED) positioned at eye level approxi-

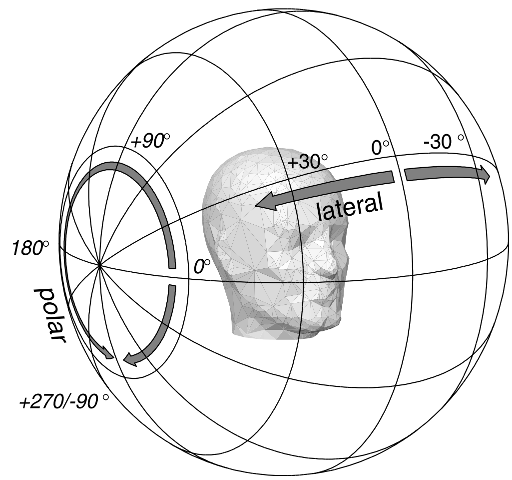

Fig. 2. The horizontal polar coordinate system. Zero degrees lateral is on the vertical midline. Positive and negative lateral angles were to the right and left of the vertical midline respectively. Zero degrees polar is the horizontal plane, 90° is directly overhead, 180° is directly behind, 270° or −90° is directly beneath.

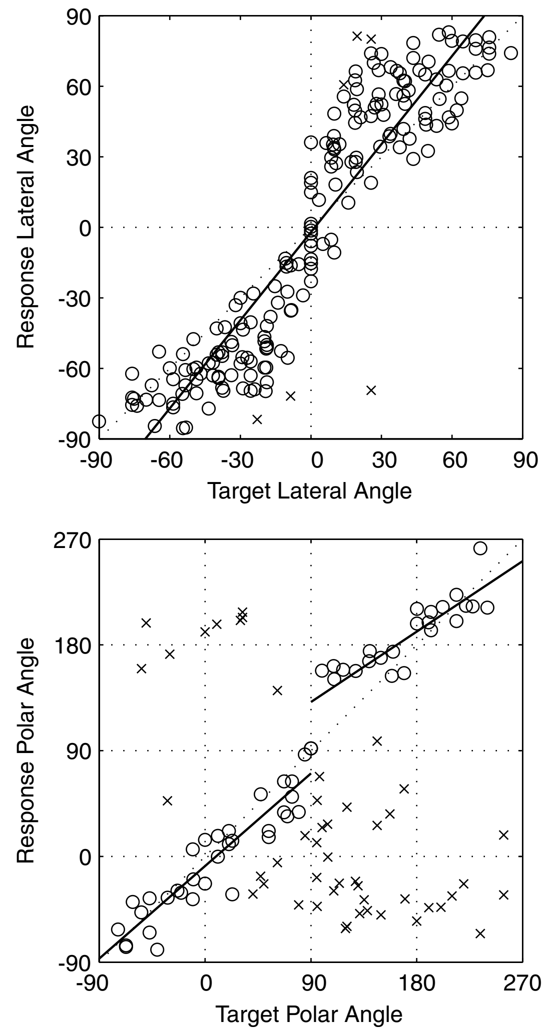

Fig. 3. Data analysis for the localization task. Lateral and polar response components were separated and plotted with respect to target angle. Independent regression lines were calculated for lateral, front polar, and rear polar data. The *slope* of the regression line was termed the subject's *gain*. Responses that were within 45° of the regression line (○) were termed *quasi-veridical.* Responses more than 45° from the regression line are plotted using the *x* symbol. The *standard deviation* of the distance from the quasi-veridical responses to the regression line was termed the subjects *variability*. These statistics were calculated for each of the tested intensities.

mately 60 cm in front of the listener (i.e. inside the radius of the loudspeaker-moving system). The listener oriented his/her head toward the LED and pressed one button on a two-button response box. Immediately after the button was pressed, the light was extinguished, and following a random delay of 500–1500 ms, a target stimulus was presented at a particular location and intensity. If the sound was inaudible, the listener pressed the left button. If the sound was audible, the listener oriented his/her head towards the sound and pressed the right button. The computer recorded both the actual location of the target stimulus and the reported location as was indicated by the orientation of the listener's head. Following the button press, the hoop rotated to put the loudspeaker in the next target location. When the loudspeaker reached the next target location, the LED was again illuminated and the next trial began. This procedure was coordinated by custom software written in MATLAB (The Mathworks, Inc.) running on an Intel-based computer. A block of trials generally took 15 min. After every block, the listener took a short (2–5 min) break. Most subjects were able to complete four blocks in 90 min.

Lateral and polar angle response components were analyzed separately. Scatterplots with target location on the abscissa and reported location on the ordinate were created for each intensity level. Sample data for one listener at 10 dB SL are plotted in Fig. 3. The particularly poor performance seen in this plot serves well for demonstration of our analyses. For the lateral component, if more than five trials were audible, a regression line was fit to the scatterplot. The *slope* of the regression line was called the listener's *lateral angle gain.* A gain of 1 would indicate bias-free performance, and a gain of 0 would indicate a strong bias towards the vertical midline. Responses that were more than 45° away from the regression line (× in Fig. 3) were not included in further calculations. The remaining responses (○ in Fig. 3) were termed *quasi-veridical.* The *standard deviation* of the distance from the quasi-veridical responses to the regression line indicated the listener's *variability.*

Analysis of the polar angle component differed in a few respects. In the horizontal polar coordinate system, the range of vertical locations is compressed as lateral angle proceeds further away from the vertical midline (Fig. 2). For this reason polar analyses were restricted to stimuli with target locations from −35° to 35° lateral. Independent regression lines were calculated for front

(−90° to +90°) and rear (+90° to +270°) target sounds. Trials in which either front targets elicited rear responses or rear targets elicited front responses were omitted from the regression line calculation but were included in all subsequent analyses. A regression line was calculated if there were more than five non-confused audible trials. When there were fewer than five such trials, a line with slope zero and y-intercept 0° for the front or 180° for the rear was used in determining quasi-veridicality. The slopes of the regression lines were termed the listener's *polar angle gains.* A positive polar angle gain of less than 1 would indicate a bias towards the horizontal plane. Again, the responses that were within 45° of the regression line were termed quasi-veridical (○ in Fig. 3), and the listener's *polar variability* was calculated by computing the standard deviation of the distance from the quasi-veridical responses to the regression line.

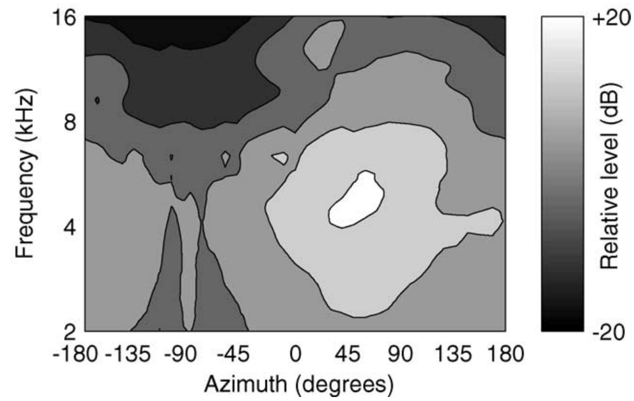

Fig. 5. One subject's right ear HRTF at locations around the horizontal meridian. The HRTF characterizes the spectral filtering caused by the head and external ear. The location with the greatest gain was at ∼45° azimuth and the frequency range with the greatest gain at that location was centered at ∼4 kHz.

3. Results

3.1. Audibility

Detection threshold means and standard errors across all listeners for eight evenly spaced locations around the horizontal plane are plotted in Fig. 4. Detection thresholds were roughly left-right symmetric. The highest threshold was for a sound located directly behind the listener's head (±180° azimuth). Listeners were most sensitive (i.e. had the lowest detection thresholds) for targets at ±45° azimuth. This location corresponds to the region with the most gain in the head-related transfer function (HRTF) (the acoustic axis of the pinna), plotted in Fig. 5. The frequency band with the most gain at this location was centered slightly above 4 kHz.

In the localization task, sound levels were specified as sensation level relative to the threshold at 0° lateral and polar. Since detection thresholds vary with location, and low sensation levels were tested, some of the targets were inaudible. The percentage of audible front and rear trials with respect to sound intensity was pooled across all subjects and is plotted in Fig. 6. As expected, the number of audible trials increased with sound level, and at equal sound levels, more front-presented stimuli were audible than rear. It was not until 15 dB SL that nearly all stimuli were audible.

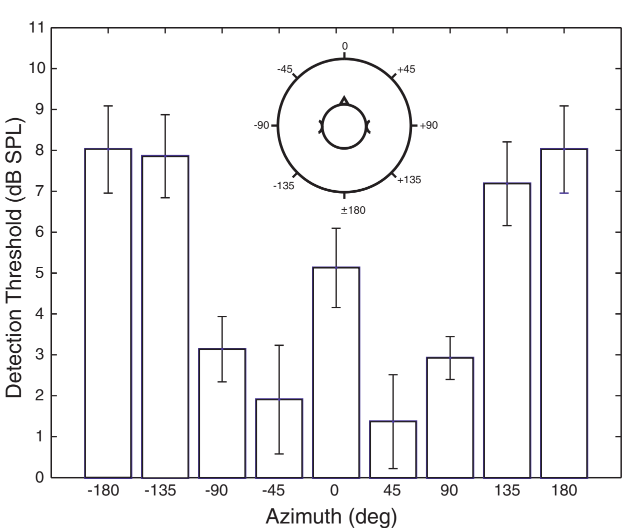

Fig. 4. Pooled detection thresholds for eight locations around the horizontal meridian. Detection thresholds were roughly left-right symmetrical about the vertical midline (azimuth = 0°). −180° and 180° azimuth both indicate the rear midline.

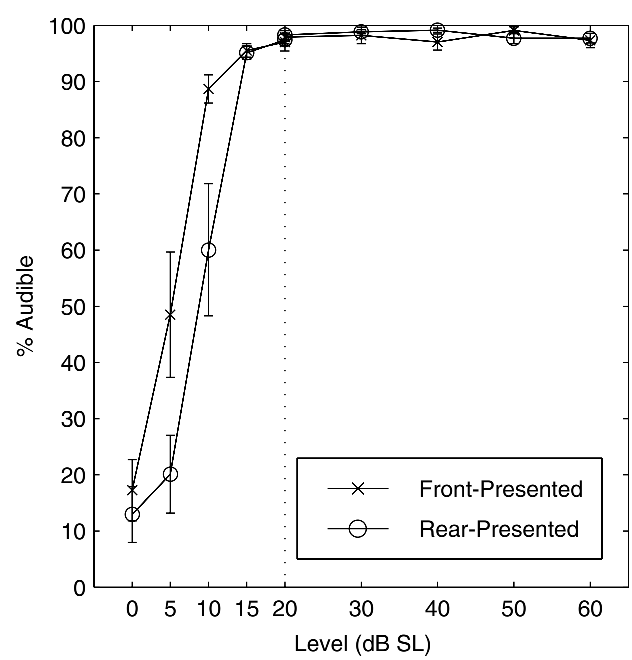

Fig. 6. Percentage of audible trials as a function of sensation level. At equal sound levels, more front-presented stimuli were audible than rear-presented stimuli.

3.2. Localization

Audible trials were sorted by sound level and localization judgments were plotted. The raw data at each sound level from one representative subject are plotted in Fig. 7. Inspection of these graphs reveals that localization performance became both more accurate and more consistent as sound intensity increased. At near-threshold levels, there appeared to be a very strong bias in the polar component towards the front horizontal plane (0° polar). At slightly higher levels, (e.g. 10 dB SL) the lateral angle response component was quite accurate, while the polar angle component remained inaccurate.

Each listener's gain (regression line slope) for both the lateral and polar response components was computed at each sound level. The mean and standard error of the gain across all subjects were computed at each sound level and plotted in Fig. 8. At the lowest sound levels, responses in the lateral dimension were slightly biased towards the midline (gain ~0.5). That bias decreased (gain approached 1) with increasing sound level and was extinguished (gain ~1) by approximately 10 dB SL. At levels greater than 10 dB SL lateral gain was slightly greater than 1, suggesting that listeners overshot the lateral location.

In the polar dimension, the front-presented stimuli exhibited a strong bias toward the front horizontal plane (gain ~0) at threshold. That bias decreased (gain neared 1) with increasing sound level and reached a plateau by ~30 dB SL. The highest value that the front polar gain reached was approximately 0.75. The small polar gain suggests that the listeners under-estimated the polar deviation from the horizontal plane even at the highest sound levels.

The rear polar gain could not be calculated at 0 dB SL because there were not enough audible, non-confused trials for any listener. At 5 dB SL, only one listener had five audible non-confused trials (signified by the open circle in Fig. 8). In contrast with the front and lateral data, the rear polar gain remained relatively constant regardless of sound level. Rear polar gain stayed below 1 for all sound levels and therefore rear responses exhibited the same under-estimation as the front polar data.

The percentage of quasi-veridical trials (audible and within 45° of the regression line) was computed for each subject at each sound level for both lateral and polar response components. The mean and standard error of those values were computed across subjects and are plotted with respect to sound level in Fig. 9. The lateral angle function showed a clear increase with respect to sound level for intensities from 0 to 15 dB SL. For stimulus intensities from 15 to 60 dB SL, nearly all responses were quasi-veridical.

Both front and rear polar response components also showed a clear increase in the percentage of quasi-veridical trials with increasing sound level. The percentage of quasi-veridical trials for front and rear responses reached plateaus of ~85% and ~75% respectively. The

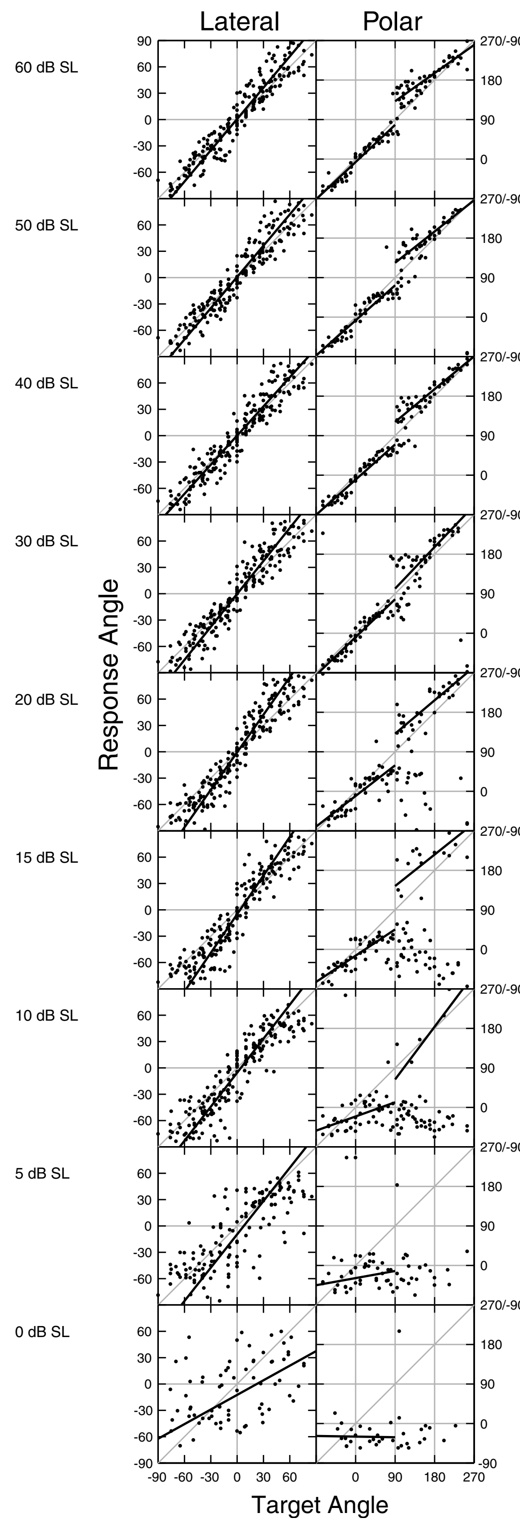

Fig. 7. Raw data from one representative subject (S153) separated by intensity level. Accuracy in both the lateral and polar response components increased with sound level. There were levels (e.g. 10 dB SL) at which lateral localization was accurate while polar was inaccurate.

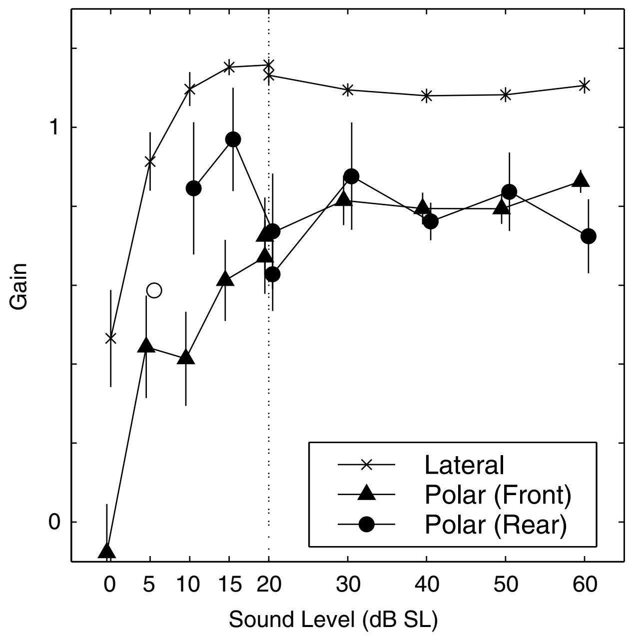

Fig. 8. Gain (regression line slope) as a function of sensation level averaged across all subjects. Lateral and front polar gain increased with level and became asymptotic around gain ~1. Since bias-free performance would be indicated by a gain of 1, the listener's bias decreased with increasing sound level. The bias for rear polar responses remained relatively constant regardless of level. The open circle represents data from only one listener. Data points are slightly staggered horizontally for ease of viewing.

corresponding percentage for the lateral response component, ~97%, was considerably higher.

Variability (standard deviation of the distance from the quasi-veridical responses to the regression line) was calculated at each sound level for each subject in both the lateral and polar dimensions. The mean and standard error of these variability values were then computed across all subjects at each sound level and are plotted in Fig. 10. Lateral angle variability was ~25° at threshold and decreased with increasing sound level, reaching an asymptote of 12.5° at approximately 30 dB SL.

The polar variability for front-presented stimuli was quite small at near-threshold levels. That variability increased with sound level until 10 dB SL. At levels higher than 10 dB SL, variability decreased with increasing sound level, reaching an asymptote of ~12° at 40 dB SL. The low variability at low levels is somewhat misleading as it does not indicate accurate localization. Variability indicates the consistency of the responses. At low levels, listeners consistently pointed to an incorrect location, near 0° polar angle, as can be seen by the near-zero gain and small variability in the polar data for front-hemisphere targets.

No variability was calculated for the polar response component for rear-hemisphere targets at 0 and 5 dB SL due to a lack or shortage of non-confused audible trials. The standard deviation at 10 dB SL was quite large (~32°). That variability decreased with increasing

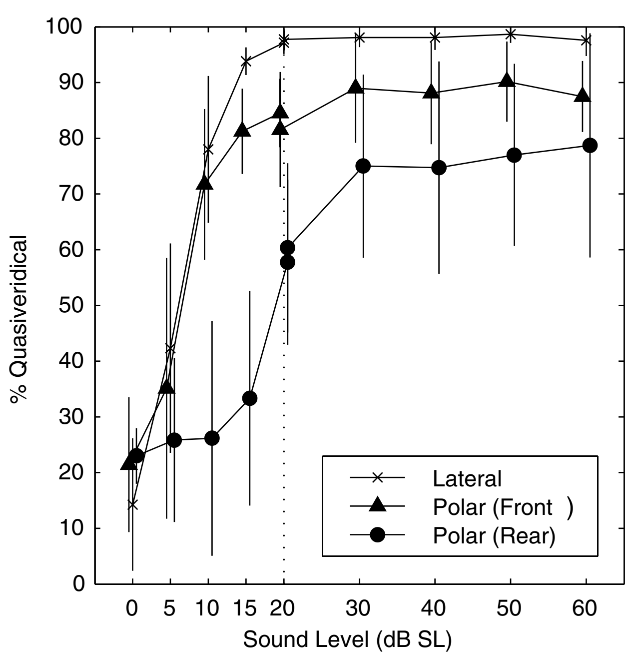

Fig. 9. Percentage of quasi-veridical responses (audible and within 45° of the regression line) as a function of sensation level averaged across all subjects. The open circle represents data from one listener. Data points are slightly staggered horizontally for ease of viewing.

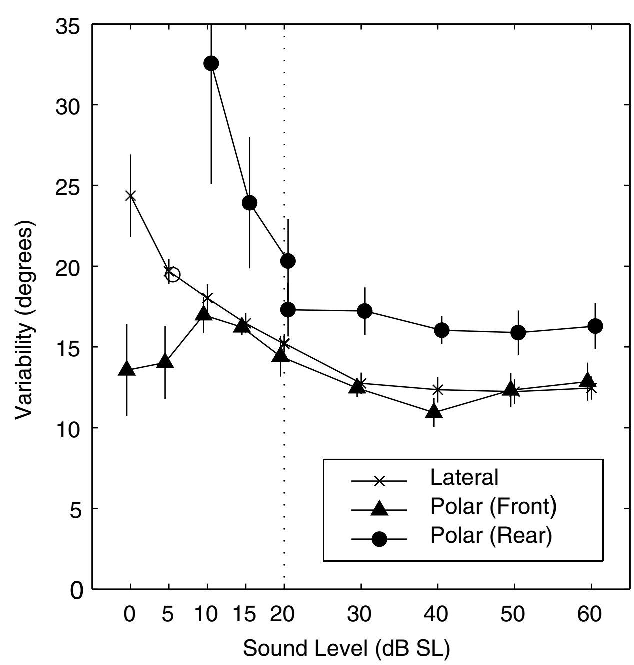

Fig. 10. Variability of quasi-veridical responses as a function of sensation level averaged across all subjects. Response variability decreased with increasing level for lateral and polar response components suggesting that listeners' responses became more consistent as level increased. The open circle represents data from only one listener. Data points are slightly staggered horizontally for ease of viewing.

sound level and plateaued at 20 dB SL. The lowest polar variability for rear presented responses was approximately 17°. There was little difference in performance between the 20-dB-SL stimuli presented in the high and low conditions.

4. Discussion

4.1. Overview

The results confirm that sound localization is inaccurate at levels near the detection threshold but rapidly improves with level, approaching an asymptote at ~30 dB SL. This observation accords with all psychophysical studies that have examined the dependence of location discrimination on sound level with stimuli of similar duration. The present study extends the previous results to a free response localization task, and shows that localization in the lateral and polar dimensions are differentially sensitive to low sound levels. The level dependence of lateral and polar localization is most likely related to the audibility of particular sound localization cues at low stimulus levels; this has been suggested by Su and Recanzone (2001) to account for their results in a study of location discrimination. With regard to the physiological motivation of this study, the results are inconsistent with models that posit representations of sound source location by single neurons with spatial receptive fields like those that have been observed in the auditory cortex.

4.2. Cue audibility at low sound levels

The salience of all directional sound localization cues varies with frequency. Nineteenth-century physicist Lord Rayleigh found that interaural differences in sound level (ILDs) were used to determine lateral angle location of pure tones at high frequencies (Strut, 1907). In addition Rayleigh demonstrated that listeners are sensitive to interaural phase differences of low-frequency sounds, which suggested that ITDs might be used to localize those sounds. The frequency dependence of lateral localization cues was further studied by Mills (1958) who showed that listeners had a deficiency in pure tone lateral localization for frequencies between 2 and 4 kHz. Presumably, this frequency region is above the range in which ITDs are detectable and below the range of large, azimuth-dependant ILDs. Macpherson and Middlebrooks (2002) tested the relative importance of ITD and ILD in broadband, high-pass, and low-pass conditions. When stimuli contained energy at low frequencies (0.5–2 kHz), listeners localization judgments were heavily weighted by ITD cues and showed little contribution of ILD cues. The dominance of ITD cues in stimuli containing low frequencies has also been shown by Wightman and Kistler (1992). For high-pass filtered (4–16 kHz) noise, localization judgments were heavily weighted by ILD and showed little contribution of ITD. Polar angle localization has also been shown to be frequency dependant. Morimoto and Aokata (1984) observed that polar localization accuracy for broadband noise decreases dramatically when the sound is low-pass filtered at 4.8 kHz, but is barely affected if the sound is high-pass filtered at the same frequency thus suggesting that the useful spectral cues lie above 4.8 kHz.

Since different frequency regions carry different combinations of localization cues, the effects of sound level on localization performance might be explained by the manner in which various portions of the spectrum rise above the threshold of audibility as sound level increases. This is illustrated schematically in Fig. 11, in which excitation patterns for pinna-filtered, flat-spectrum noise at four levels (thin lines) are compared to the minimum audible field (MAF, thick line). An excitation pattern represents the output of a bank of auditory filters as a function of filter center frequency, and thus takes into account the increase in auditory filter bandwidth (and consequent loss of spectral resolution and summation of energy) at higher center frequencies. The excitation patterns were calculated using the method described by Glasberg and Moore (1990) assuming a fixed stimulus level of 20 dB SPL, and then were offset vertically to represent different stimulus levels. Pinna filtering is represented here by the application of a directional transfer function (DTF), which captures the

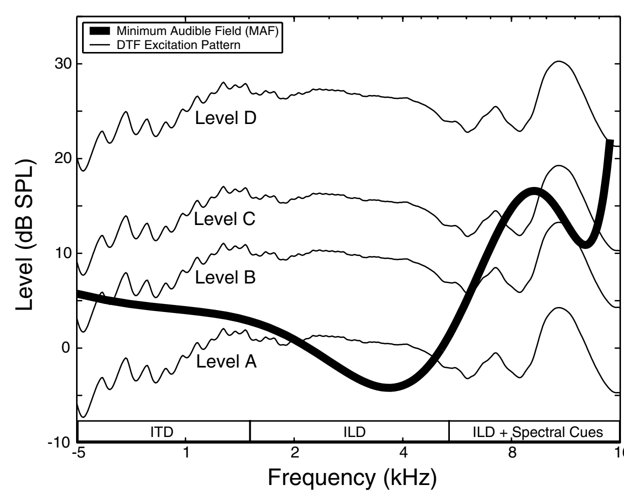

Fig. 11. A schematic offering an explanation for the effect of sound level on localization. The thick line is the MAF, which is a function of pure tone detection thresholds with respect to frequency. The thin lines are the excitation pattern for broadband flat-spectrum noise filtered by one listener's directional transfer function plotted at four levels (designated by letter). Any portion of the excitation pattern that lies above the MAF is considered audible. The audible portion of the frequency spectrum widens as sound level increases. Since localization cues are distributed throughout a wide range of frequencies, it is not until the full spectral range is audible that all cues are useable and optimal localization can be achieved.

directional sensitivity of the external ear for a particular source location but excludes spatially invariant contributions to the overall source-to-eardrum transfer function, such as the ear canal resonance. The MAF is the frequency-specific detection threshold for a front-positioned pure tone (data taken from Glasberg and Moore (1990)), and reflects the combined effects of the resonance of the pinna and ear canal and the frequency sensitivity of the cochlea (among other factors). The lowest point on the MAF curve (thick line), and therefore most sensitive frequency region, is at approximately 4 kHz.

At any particular sound level, one can regard the frequency range of available spatial cues as the range of the excitation pattern that lies above the MAF function. The audible portion of the frequency spectrum widens as sound level increases. Since necessary localization cues are distributed throughout a wide range of frequencies, it is not until the full spectral range is audible that all cues are useable and optimal localization can be achieved.

The only portion of the excitation pattern at level A (which represents a near-threshold stimulus) in Fig. 11 that lies above the MAF curve is a narrow band centered at 4 kHz. Lateral localization is impaired in this frequency range (Mills, 1958) as the only spatial cue that is detectable is ILD. Even ILDs are small in this frequency region compared to higher frequencies. Degraded lateral localization performance was seen in the data by the high variability and moderate bias present at the lowest levels (Figs. 8 and 10, respectively).

Essentially no polar location cues reside in the narrow band centered at 4 kHz. Indeed, polar responses were grossly inaccurate at near-threshold levels. Listeners consistently responded on the horizontal plane, regardless of where the sound was presented. This is seen by the low variability and near-zero polar angle gain (Figs. 8 and 10, respectively). Since no polar cues were audible, it is possible that listeners were oblivious to the polar location and gave a default response on the horizontal plane. It is also possible that the listeners actually perceived the sound to have originated on the horizontal plane. Previous investigations have shown that supra-threshold narrow-band noise bursts with center frequencies of 4 kHz are often reported to have originated on the horizontal meridian regardless of presented location (Middlebrooks, 1992; Butler, 1971; Blauert, 1969/1970).

The excitation pattern at level B in Fig. 11 represents a stimulus that is slightly above the detection threshold. A much greater portion of the excitation pattern lies above the MAF curve. The audible sound can now be roughly described as low-pass filtered at 5 kHz. The audible frequency range contains both of the cues used in lateral localization (ITD and ILD). Accurate lateral localization should be possible at this sound level. This increase in lateral localization performance is seen in the data by the dramatic decrease in bias (increase in gain) and variability as sound level increased. By 10 dB SL, lateral angle response components were unbiased and consistent. The frequency region containing the spectral cues necessary for polar localization is still sub-threshold at level B, which accords with the observation that polar angle responses retain their bias towards the horizontal plane at sound levels greater than the detection threshold.

The excitation pattern at level C in Fig. 11 represents a stimulus that is moderately above threshold. The majority of the spectrum at this level lies above the MAF curve and therefore all localization cues should be audible. There should be little increase in lateral angle localization accuracy, as all of the necessary cues were audible at a lower sound level. All lateral data become asymptotic around 20 dB SL (Figs. 8–10).

At level C, some bands in the frequency region containing the polar angle spectral cues become audible. However, it is not sufficient that some frequencies in this region are audible. Polar localization is achieved by comparing the energy level across various frequencies (i.e. by determining spectral shape). To achieve accurate polar localization, enough dynamic range must be audible for the listener to detect the spectral peaks and notches. Since only part of the spectral shape at level C is audible, polar localization should be sub-optimal. Polar localization should gradually improve with level as a greater portion of the spectral shape becomes detectable. The front polar angle gain curve (Fig. 8) increases much more gradually than the lateral.

Rear polar angle gain exhibits little effect of level (Fig. 8). The spectral cues used in determining vertical and front/rear position reside in roughly the same frequency region. If there is enough energy to recognize a rear presentation, there should also be enough energy to determine the elevation component. Since rear polar angle gain was only calculated for non-confused rear responses, one would expect little level dependence.

The excitation pattern at level D in Fig. 11 represents a stimulus that is well above threshold. At this point the entire spectrum is audible, so optimal polar localization can be achieved. There should be little change in localization performance at levels higher than D. All polar and lateral angle functions presented in this study at some point became asymptotic (Figs. 8–10). One might suspect that at very high sound levels, polar localization would be degraded due to decreased spectral resolution in the auditory system (Shailer et al., 1990; Rosen et al., 1998; Supin et al., 2003).

Vliegen and van Opstal (2004) examined elevation gain as a function of level and duration. For stimuli at the lowest tested sound level (26 dB SPL) and the shortest duration (3 ms) localization judgments were heavily biased towards the horizontal plane. Whereas the schematic presented here in Fig. 11 attempts to explain why localization should be impaired, the model proposed by Vliegen and van Opstal (2004) attempts to explain why localization judgments occurred where they did. Briefly, those authors proposed that the auditory system begins with a default location on the horizontal plane and that spatial estimates are then integrated over the duration of the stimulus moving the perceived location closer to the target. This model predicts that stimuli with poorly represented spectra or short durations should be localized on or near the horizontal plane as was observed in the present study for stimuli at near-threshold levels.

4.3. Relation to physiology

Models explaining the representation of auditory space in the central nervous system that rely on sharp tuning of single neurons predict that localization accuracy will increase as that neuron's receptive field narrows. Many studies have shown that the receptive fields of single neurons in the auditory cortex narrow as sound level decreases. Therefore, such models predict that sound localization will improve with decreasing sound level. That prediction clearly conflicts with the present data (and with common sense). Alternative models state that the representation of auditory space may be distributed across many broadly tuned neurons (Middlebrooks et al., 1998; Stecker and Middlebrooks, 2003; Jenison, 2001). A distributed model predicts that sound localization will decline at lower sound levels because fewer neurons are activated. This prediction accords with the present psychophysical results.

Acknowledgements

The authors would like to thank: Dr. Brian C.J. Moore and one anonymous reviewer for their useful comments on an earlier version of this paper; Zeikye Onsan for providing invaluable technical and administrative assistance; Dr. Jinro Inoue for providing information on his previous study; and Dr. Chris Stecker and Dr. Ian Harrington for providing constructive conversation. This study was supported by NIDCD grants R01 DC00420 and P30 DC05188.

References

Altshuler, M.W., Comalli, P.E., 1975. Effect of stimulus intensity and frequency on median horizontal plane sound localization. J. Aud. Res. 15, 262–265.

Blauert, J., 1969/1970. Sound localization in the median plane. Acustica 22, 205–213.

Brugge, J.F., Reale, R.A, Hind, J.E., 1996. The structure of spatial receptive fields of neurons in primary auditory cortex of the cat. J. Neurosci. 16, 4420–4437.

Butler, R.A., 1971. The monaural localization of tonal stimuli. Percept. Psychophys. 9, 99–101.

Comalli, P.E., Altshuler, M.W., 1976. Effect of stimulus intensity, frequency, and unilateral hearing loss on sound localization. J. Aud. Res. 16, 275–279.

Davis, J.R., Stephens, S.D.G., 1974. The effect of intensity on the localization of different acoustical stimuli in the vertical plane. J. Sound Vib. 35, 223–229.

Glasberg, B.R., Moore, B.C.J., 1990. Derivation of auditory filter shapes from notched-noise data. Hear. Res. 47, 103–138.

Hartmann, W.M., Rakerd, B., 1993. Auditory spectral discrimination and the localization of clicks in the sagittal plane. J. Acoust. Soc. Am. 94, 2083–2092.

Hebrank, J., Wright, D., 1975. The effect of stimulus intensity upon the localization of sound sources on the median plane. J. Sound Vib. 38, 498–500.

Hofmann, P.M., van Opstal, J., 1998. Spectro-temporal factors in two-dimensional human sound localization. J. Acoust. Soc. Am. 103, 2634–2648.

Imig, T.J., Irons, W.A., Samson, F.R., 1990. Single-unit selectivity to azimuthal direction and sound pressure level of noise bursts in cat high-frequency primary auditory cortex. J. Neurophysiol. 63, 1448–1466.

Inoue, J., 2001. Effects of stimulus intensity on sound localization in the horizontal and upper-hemispheric median plane. J. UOEH 23, 127–138.

Jenison, R.L., 2001. Decoding first-spike latency: A likelihood approach. Neurocomputing 38, 239–248.

Levitt, H., 1971. Transformed up-down methods in psychocacoustics. J. Acoust. Soc. Am. 49, 467–477.

Macpherson, E.A., Middlebrooks, J.C., 2000. Localization of brief sounds: effects of level and background noise. J. Acoust. Soc. Am. 108, 1837–1849.

Macpherson, E.A., Middlebrooks, J.C., 2002. Listener weighting of cues for lateral angle: the duplex theory of sound localization revisited. J. Acoust. Soc. Am. 111, 2219–2236.

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

Middlebrooks, J.C., Green, D.M., 1991. Sound localization by human listeners. Annu. Rev. Psychol. 42, 135–159.

Middlebrooks, J.C., Pettigrew, J.D., 1981. Functional classes of neurons in primary auditory cortex of the cat distinguished by sensitivity to location. J. Neurosci. 1, 107–120.

Middlebrooks, J.C., Xu, L., Eddins, A.C., Green, D.M., 1998. Codes for sound-source location in nontonotopic auditory cortex. J. Neurophysiol. 80, 863–881.

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

Morimoto, M., Aokata, H., 1984. Localization cues of sound sources in the upper hemisphere. J. Acoust. Soc. Jpn. 5, 165–173.

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

Rajan, R., Aitkin, L.M., Irvine, D.R., 1990. Azimuthal sensitivity of neurons in primary cortex of cats. II. Organization along frequency band strips. J. Neurophysiol. 64, 888–902.

Reale, R.A., Jenison, R.L., Brugge, J.F., 2002. Directional sensitivity of neurons in the primary auditory (AI) cortex: effects of sound-source intensity level. J. Neurophysiol. 89, 1024–1038.

Rosen, S., Baker, R.J., Darling, A., 1998. Auditory filter nonlinearity at 2 kHz in normal listeners. J. Acoust. Soc. Am. 103, 2539–2550.

Shailer, M.J., Moore, B.C.J., Glasberg, B.R., Watson, N., Harris, S., 1990. Auditory filter shapes at 8 and 10 kHz. J. Acoust. Soc. Am. 88, 141–148.

Stecker, G.C., Middlebrooks, J.C., 2003. Distributed coding of sound locations in the auditory cortex. Biol. Cybernet. 89, 341–349.

Stecker, G.C., Mickey, B.J., Macpherson, E.A., Middlebrooks, J.C., 2003. Spatial sensitivity in field PAF of cat auditory cortex. J. Neurophysiol. 89, 2889–2903.

Strut, J.W., 1907. On our perception of sound direction. Philos. Mag. 13, 214–232.

Su, T.K., Recanzone, G.H., 2001. Differential effect of near-threshold stimulus intensities on sound localization performance in azimuth and elevation in normal subjects. J. Assoc. Res. Oto. 2, 246–256.

Su, T.K., Woods, T.M., Recanzone, G.H., 2000. Effect of intensity on sound localization performance in macaque monkeys. Soc. Neurosci. Abstr. 26, 955.

Supin, A.Y., Popov, V.V., Milekhina, O.N., Tarakanov, M.B., 2003. Rippled-spectrum resolution dependence on level. Hear. Res. 185, 1–12.

Vliegen, J., van Opstal, J.V., 2004. The influence of duration and level on human sound localization. J. Acoust. Soc. Am. 115, 1705–1713.

Wightman, F.L., Kistler, D.J., 1992. The dominant role of low-frequency interaural time differences in sound localization. J. Acoust. Soc. Am. 91, 1648–1661.