Exp Brain Res (2003) 148:458–470
DOI 10.1007/s00221-002-1320-5

RESEARCH ARTICLE

P. M. Hofman · A. J. Van Opstal

Binaural weighting of pinna cues in human sound localization

Received: 22 April 2002 / Accepted: 22 October 2002 / Published online: 6 December 2002

Abstract Human sound localization relies on binaural difference cues for sound-source azimuth and pinna-related spectral shape cues for sound elevation. Although the interaural timing and level difference cues are weighted to produce a percept of sound azimuth, much less is known about binaural mechanisms underlying elevation perception. This problem is particularly interesting for the frontal hemifield, where binaural inputs are of comparable strength. In this paper, localization experiments are described in which hearing for each ear was either normal, or spectrally disrupted by a mold fitted to the external ear. Head-fixed saccadic eye movements were used as a rapid and accurate indicator of perceived sound direction in azimuth and elevation. In the control condition (both ears free) azimuth and elevation components of saccadic responses were well described by a linear regression line for the entire measured range. For unilateral mold conditions, the azimuth response components did not differ from controls. The influence of the mold on elevation responses was largest on the ipsilateral side, and declined systematically with azimuth towards the side of the free ear. Near the midsagittal plane the elevation responses were clearly affected by the mold, suggesting a systematic binaural interaction in the neural computation of perceived elevation that straddles the midline. A quantitative comparison of responses from the unilateral mold, the bilateral mold and control condition provided evidence that the fusion process can be described by binaural weighted averaging. Two different conceptual schemes are discussed that could underlie the observed responses.

Keywords Sound localization · Human · Eye movements · Pinna · Spectral cues · Binaural weighting

P.M. Hofman · A.J. Van Opstal (✉)
Department of Biophysics, University of Nijmegen,
Geert Grooteplein 21, 6525 EZ Nijmegen, The Netherlands
e-mail: johnvo@mbfys.kun.nl
Tel.: +31-24-3614251
Fax: +31-24-3541435

Introduction

Because the inner ear represents sounds tonotopically rather than spatially, sound localization relies on the neural processing of acoustic cues. Interaural differences in time and sound level provide robust information regarding sound-source *azimuth*. The pinnae provide spectral shape cues, enabling extraction of sound *elevation*, and frontal versus rear locations (Oldfield and Parker 1984; Wightman and Kistler 1989; Middlebrooks and Green 1991; Middlebrooks 1992; Blauert 1997). Other factors may also contribute to spatial hearing. These include vision (e.g., the 'ventriloquist illusion', e.g., Stein and Meredith 1993), head movements (Perrott et al. 1987; Goossens and Van Opstal 1999), and expectation about upcoming target locations.

To construct a consistent percept of sound location, the various sources of information must be combined by appropriate fusion and selection mechanisms. This poses interesting problems, especially when cues are ambiguous. The latter happens, e.g., when interaural phase and level differences point to more than one azimuth, or when the binaural pinna cues refer to different elevations.

Early dichotic lateralization studies have reported that when phase and level disparities point to opposing lateral positions, subjects may perceive an average location ('time-intensity trading'; e.g., Blauert 1997, for review). This averaging phenomenon depends in a complex way on stimulus parameters like level and bandwidth (Harris 1960). When differences are large, subjects report separated time and intensity images, often leading to a split percept (Whitworth and Jeffress 1961; Hafter and Jeffress 1968). However, more recently it has been reported that under more realistic acoustic conditions time-intensity trading is not observed. For example, when dichotic stimulation accurately accounts for the pinna-related spectral shape cues, the acoustic image is perceived outside rather than inside the head (Wightman and Kistler 1989). Under these circumstances, low-frequency interaural time differences appear to overrule contradicting level differences (Wightman and Kistler 1992).

Several studies have measured the contribution of the spectral pinna cues to binaural 2D sound localization (Hebrank and Wright 1974; Flannery and Butler 1981; Oldfield and Parker 1984; Wightman and Kistler 1989; Makous and Middlebrooks 1990; Middlebrooks 1992; Butler and Humanski 1992; Butler and Musicant 1993; Hofman et al. 1998). Binaurally applied molds to the pinna cavities (Oldfield and Parker 1984; Hofman et al. 1998), narrow-band sounds (Middlebrooks 1992), as well as roving high-frequency bands in the stimulus spectrum (Wightman and Kistler 1997) all affect localization of sounds in elevation, including a marked increase in front-back confusions. Monaural patients, as well as normal-hearing subjects with a plugged ear canal, can still localize sounds in elevation, especially on the hearing side, albeit slightly worse than normal-hearing binaural subjects (Hebrank and Wright 1974; Oldfield and Parker 1986; Slattery III and Middlebrooks 1994). This suggests that binaural processes may also influence the localization of sound-source elevation.

A comparison of localization performance of lateral targets (beyond 45 deg from the medial plane) when each ear was either free, or equipped with a mold, indicated a strong dominance of the near ear to perceived elevation (Humanski and Butler 1988). Recently, Morimoto (2001) reinvestigated this issue in more detail by measuring localization of targets in the upper hemifield at different lateral angles interspaced at 30-deg intervals. It was shown that localization errors in elevation, induced by a unilateral mold, were also observed on the side contralateral to the mold. In line with the findings of Humanski and Butler (1988), this binaural interaction disappeared for contralateral azimuths beyond 60 deg.

The present study extends these findings in several ways. First, the experiments concentrate on the central 35 deg of the frontal hemifield. Particularly in this target range, sound localization has a high resolution (Mills 1958; Makous and Middlebrooks 1990; Frens and Van Opstal 1995) and the acoustic inputs to both ears have comparable power. Thus, strongest (changes in) binaural interactions of pinna-related cues may be expected in this range. Although earlier studies have shown that binaural interactions exist, the actual contribution of either ear to sound elevation was not quantified. We therefore attempted to assess in more detail the contribution of either ear to perceived target elevation for targets near the midsagittal plane. Manipulation of the available spectral cues was obtained by applying both unilateral and bilateral molds.

Second, to measure the fastest available spatial percept of subjects, saccadic eye movements toward sounds presented at random locations within the 2D oculomotor range (about 35 deg in all directions) were recorded. Subjects were instructed to respond as fast and as accurately as possible. Typical response latencies are well below 300 ms, in contrast with the response method of Morimoto (2001), needing 4 s for each response.

Our results show that the binaural pinna cues are averaged to construct a unified elevation percept, where perceived stimulus azimuth acts as a weighting factor. Possible neural mechanisms are discussed.

Materials and methods

Subjects

This study involved four male subjects (S1–S4; 24–42 years of age). S2 and S3 were the authors of this paper. Subjects S1 and S4 were students from the laboratory. Subject S4 was naive regarding the purpose of these experiments. All four subjects had prior experience with sound localization studies in the laboratory and had no hearing problems of any kind. As to the procedures of the experiments, their informed consent was obtained. Audiograms, obtained by a standard staircase procedure for each ear with pure tones of 0.5–11.3 kHz at half-octave intervals, were found to be within the normal range. The paradigms and experiments were approved by the local ethics committee of the University of Nijmegen.

Apparatus

Experiments were conducted in a completely dark and sound-attenuated room with dimensions of 3×3×3 m. The walls, floor and ceiling were covered with acoustic foam, which eliminated reflections above 500 Hz. The room had an A-weighted ambient background noise level of 30 dB. The orientation of the subject's left or right eye was measured with the scleral search coil technique (Collewijn et al. 1975). Details of the equipment needed for this method are described in earlier papers from this laboratory (Frens and Van Opstal 1995; Hofman and Van Opstal 1998). An acoustically transparent frontal hemisphere (consisting of a thin wire framework, covered with black silk cloth) with 85 red light-emitting diodes (LEDs) was used for calibration of the eye-coil measurements and for providing a fixation light at the start of each localization trial.

Sound stimuli were delivered through a broad-range lightweight speaker (Philips AD-44725) mounted on a two-link robot (see also Fig. 1 in Hofman and Van Opstal 1998). By incorporating a random movement of at least 20 deg between trials, it was ensured that sounds from the robot engines provided no cues about the next stimulus position (Frens and Van Opstal 1995). The experiment was controlled by two PCs (80486). The first PC ('master') was equipped with hardware for data acquisition, stimulus timing and control of the LEDs. The second PC ('slave') controlled the robot and generated the auditory stimuli. A stimulus was stored in memory prior to a trial and, upon receiving a trigger from the master PC, was passed through a DA converter (Data Translation DT2821) at a sampling rate of 50 kHz. The output of the board was bandpass filtered (Krohn-Hite 3343; passband 0.2–20 kHz), and subsequently equalized in order to flatten the spectrum within 5 dB (Behringer Ultra-Curve). Finally, the signal was amplified (Luxman A-331) and passed to the speaker.

Hearing conditions

Localization performance was tested under four different open-loop hearing conditions. Each ear could be either *free* (F), i.e., normal hearing, or equipped with a *mold* (M). For further reference, a label was assigned to each hearing condition. For example, condition MF refers to a mold in the left ear, and a free right ear. In the control condition (C) both ears were free. In the mold condition, the acoustic input to the ear was spectrally modified by means of a precisely fitting rubber mold. The mold was created by making a rubber cast of the ear canal and the outer ear (applying Otoform Otoplastik-K/c). The molds covered the entire scapha and filled up a large part of the concha, while leaving the ear canal unobstructed.

To ensure that a mold would stay in place during the experiment, wax was applied to make it stick better to the pinna. Each subject had his own personal set of two molds, which was used during the entire series of localization sessions.

Sound stimuli

Three different types of stimuli were used: broadband noise (BB), low-pass noise (LP), and high-pass noise (HP). The pass bands for these stimuli were 0.2–20 kHz, 0.2–1.5 kHz, and 3.0–20 kHz, respectively. The power spectra fell off steeply at the start and stop frequencies. Spectral power dropped by more than 120 dB/octave between 1.5 kHz and 2 kHz for the low-pass stimulus, and was about equally rapid for the high-pass spectrum from 3.0 kHz down to 2.5 kHz. The stimuli were created offline (Matlab 5.0, the Mathworks). Smooth on- and offsets were created by modulating the first and last 5 ms of the waveform by a squared sine function. The duration of the stimuli was 500 ms. The three types of noise stimuli were presented at two different intensities. The A-weighted sound level was either 50 or 60 dB, measured at the subject's head (amplifier BK2610 and microphone BK4144). The resulting set of six stimuli (three spectra, two intensities) were presented randomly interleaved in all sessions.

In this way, there was a large amount of prior uncertainty for the subjects, with respect to both the upcoming stimulus location and its acoustic properties. In the data analysis of response accuracies, the results from the HP and BB, as well as the low- and high-intensity stimuli, were pooled, as no systematic differences in the responses were found. The results for the LP stimuli were analyzed separately (intensities pooled), as the elevation responses to these stimuli differed markedly from the HP and BB results.

Stimulus positions

Stimulus positions are defined in the double-pole azimuth-elevation coordinate system with the origin at the center of the head. Azimuth, α, is defined as the (horizontal) angle relative to the vertical median plane. Similarly, elevation, ε, is defined as the (vertical) angle relative to the horizontal plane that contains the interaural line. A positive (negative) azimuth corresponds to the right (left) side of the subject. A positive (negative) elevation corresponds to the upward (downward) direction.

Stimulus positions were confined to 25 boxes centered at azimuths α=0, ±13, and ±26 deg, and elevations ε=0, ±13, and ±26. The dimension of each box was 8×8 deg. Consequently, the target range extended from –30 deg to +30 deg, for both azimuth and elevation. Sets of 25 target positions for each stimulus type were generated by selecting a position at random within each box. The purpose of this procedure was to ensure a high degree of uncertainty, while maintaining a homogeneous distribution over the target range (Hofman and Van Opstal 1998).

Paradigms

A typical session consisted of 72 calibration trials, 150 test trials, and 50 control trials. Molds were inserted before the start of the experimental session. In the first run, data were collected for offline calibration of eye position (head fixed), which served as a pointer to the perceived sound location (see also Frens and Van Opstal 1995). Subjects were instructed to generate an accurate saccade from the central fixation LED at 0 deg eccentricity to the peripheral visual target, and to maintain fixation as long as the target was visible. The LEDs at 2 deg eccentricity were not used. After calibration, the eye position was known with an absolute accuracy of 3%, or better, over the full oculomotor range (thus within 1 deg).

Sound stimuli were presented in the subsequent test run(s) (with the mold(s) in situ) and control run (both ears free). A trial always started with the central visual fixation stimulus. Then, after a random period of 0.4–0.8 s, the LED was switched off and the sound stimulus was presented during the following 500 ms at a randomly selected position within the target range. The eye position was measured for 1.5 s from the onset of the fixation spot. The subject's task was to redirect the eyes *as fast and as accurately as possible* toward the apparent sound location without moving the head. A firm head rest enabled the subject to stabilize his head position throughout the session.

Finally, a control run was performed with both ears free. This run consisted of 50 trials containing the broadband (BB) noise targets at two sound levels (50 and 60 dBA) presented at each of the 25 target boxes. Stimulus types, intensities, and target positions were randomized throughout the test run and the control run. Note that experiments were carried out under *open-loop* conditions, as subjects never received feedback about their performance. Subjects typically participated in about six experimental sessions in total, each carried out on a different day. In this way, a total of at least 300 trials per hearing condition (>50 per stimulus type) was obtained.

Audiograms (see above) were also obtained with the molds in situ. It was verified that these audiograms were indistinguishable from the free-ear condition over the entire frequency range tested (0.5–11.3 kHz).

Data analysis

Eye position was calibrated on the basis of responses to the 72 LEDs. From this run, sets of raw eye position signals (horizontal and vertical position channels) and the corresponding LED positions (in azimuth and elevation) were obtained (Hofman and Van Opstal 1998). These data were used to train a feedforward neural network (one hidden layer) that mapped the raw data signals to calibrated eye position signals.

A custom-made PC program identified saccades on the basis of preset velocity criteria for saccade on- and offset. The program enabled interactive correction of the detection markings. The endpoint of the first or second saccade after stimulus onset was defined as the response position (see also "Results"). If saccade latency regarding stimulus onset was less than 80 ms or exceeded 1,000 ms, the response was discarded from further analysis. If the first two saccades in a trial both met these requirements, then the endpoint of the second saccade, rather than the first saccade, was selected as the response position.

Computation of optimal model parameters

In order to describe the relation between target direction (α_T, ε_T), and response direction (α_R, ε_R), two different functions were used in this study, a *linear model*, and a so-called *non-linear twist model*. Both models are described by a generic expression of the following form:

$$\varepsilon_R(\alpha_T, \varepsilon_T) = a'(\alpha_T) + b'(\alpha_T) \cdot \varepsilon_T \quad (1)$$

for response elevation. A similar expression described the azimuth responses. Parameters *a'* and *b'* are constants in the linear model, and polynomials in α_T of order one or two in the non-linear twist model (see also "Results", Figs. 4, 6, Eqs. 2, 3, 4).

Optimal model parameters were determined by the minimal distance between predictions and actual data, according to the least squares criterion. Confidence intervals of the fit parameters were estimated by bootstrapping, in which the parameters were determined at least 150 times from different random realizations of the data set (Press et al. 1992). Correlation coefficients were computed according to Spearman's rank correlation (Press et al. 1992).

Results

For each subject, at least 1,200 saccadic eye movement responses were collected for the four hearing conditions (C, FM, MF and MM). For all conditions tested, the response accuracies of the 50- and 60-dB stimulus intensities were indistinguishable. The responses toward these stimuli were therefore pooled. The same holds for the HP and BB stimulus spectra. In this way, the results from at least 200 responses for each hearing condition could be pooled for each subject. The LP stimuli were analyzed separately (see below).

Since the localization responses of our subjects in the control condition (both ears free) were quantified by applying a new descriptive model to the data (see Table 1 for a summary of the results, and Fig. 4A for an example), the rationale for applying this model to our data is more conveniently introduced by first describing the results of the unilateral mold experiments.

Unilateral mold condition

In the current experiments, the question was investigated to what extent perturbation of the spectral cues on one side would influence localization responses over the entire 2D oculomotor range. To that end, subjects were tested for six stimulus types during a series of experimental sessions, while wearing a unilateral mold that fitted into the concha of either ear.

To gain a first qualitative impression of the subjects' response behavior, Fig. 1 shows a number of individual saccadic eye movement responses of subject S2 toward one small target area around coordinates (α_T, ε_T)=(26, 26) deg for different hearing conditions and stimulus spectra. The filled response curves correspond to the control responses (C) for either stimulus type, whereas the open symbols depict the unilateral mold response traces (FM and MF). Note that the control responses toward the HP and BB noise stimuli are quite accurate, but that the responses to the LP noise stimuli are only accurate in azimuth. The responses toward the LP noise stimuli with either mold appear to be indistinguishable from the control responses.

Saccades toward the HP and BB stimuli with the mold applied to the left ear (MF; i.e., contralateral to the stimulus) are also very similar to the controls. The saccadic responses to these stimuli with an ipsilateral mold (FM),

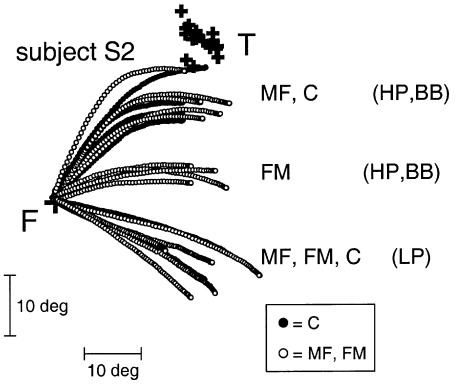

Fig. 1 Example traces of a number of individual saccadic eye movements toward targets located in a small 8×8-deg box centered around (α_T, ε_T)=(26, 26) deg, during control (*C*), and unilateral left (*MF*) or right (*FM*) mold conditions, for high-pass-, and low-pass noise (*solid circles* control responses, *open circles* unilateral mold responses). Note that the control and MF responses (mold contralateral to the stimulus) toward HP and BB stimuli are quite accurate, but that the responses for the ipsilateral mold condition (*FM*) are inaccurate in their elevation component. All conditions yield wrong elevation responses for LP stimuli. Azimuth is accurate for all stimulus and hearing conditions. Data are from subject S2

Table 1 Parameters from fits of the control responses to the twist model (Eq. 4) as applied to both azimuth and elevation. Stimulus is high-pass or broadband noise (pooled). Stimulus intensity is 60 dB SPL. *Column 1* subject, columns 2–5 twist model parameters bias *a*, bias drift Δa, gain *b* and gain drift Δb. Both the parameter and its estimated error (in *parentheses*) are indicated. *Column 6* $\Delta b/b$ ratio, indicating the relative variation of the gain within the target range; an *asterisk* marks when the ratio exceeds two times the estimated error, indicating that the ratio is significantly non-zero. *Column 7* correlation of the fitted responses and the actual data. *Column 8* number of responses used for the fit. Note that, for both azimuth and elevation, bias *a* is small (<5 deg), gains *b* are high (>0.72), and both bias drift and gain drift are relatively small. Also, note that correlations between actual responses and fitted responses are high (>0.95). Confidence intervals were estimated by means of bootstrapping

Subject	*a*	Δa	*b*	Δb	$\Delta b/b$	*r*	*N*
Azimuth response component							
Control							
S1	0.5 (0.2)	–0.3 (0.3)	1.06 (0.01)	–0.08 (0.02)	–0.08 (0.02)*	0.99	336
S2	3.0 (0.3)	–0.6 (0.6)	1.08 (0.02)	0.02 (0.03)	0.02 (0.03)	0.96	380
S3	–5.5 (0.2)	0.8 (0.4)	1.18 (0.02)	–0.12 (0.02)	–0.10 (0.02)*	0.98	398
S4	0.9 (0.3)	–0.6 (0.4)	0.72 (0.02)	0.01 (0.03)	0.01 (0.04)	0.96	273
Elevation response component							
Control							
S1	4.8 (0.3)	1.6 (0.5)	0.82 (0.02)	–0.13 (0.03)	–0.16 (0.04)*	0.95	336
S2	–4.8 (0.3)	–0.5 (0.4)	0.75 (0.02)	–0.06 (0.03)	–0.08 (0.04)*	0.95	380
S3	2.4 (0.3)	2.8 (0.4)	0.82 (0.02)	–0.02 (0.02)	–0.02 (0.03)	0.96	398
S4	–1.6 (0.3)	–0.7 (0.4)	0.77 (0.02)	0.03 (0.03)	0.04 (0.03)	0.96	273

however, appear to be inaccurate (i.e., undershooting) in their elevation components. All conditions yield comparable accuracy in their azimuth components.

To demonstrate that the responses shown in Fig. 1 are exemplary for the unilateral mold condition, Fig. 2 shows the saccadic eye-displacement components of subject S4 to all target positions applied (right-mold (FM) hearing condition; intensities pooled). The left-hand column shows the azimuth components of the saccades as function of the azimuth target components, whereas the right-hand column shows the results for elevation. To quantify the behavior, each response component was first fitted by a linear regression line. For elevation:

$$\varepsilon_R = a + b \cdot \varepsilon_T \qquad (2)$$

where a is the response bias (in degrees) and the slope b (dimensionless) represents the gain of the responses. An analogous description was used for the azimuth components (see "Materials and methods").

The top panels show the responses to the LP noise stimuli. Note that the azimuth components are quite accurate, as the fitted regression line has a gain close to unity, a bias close to zero, and a high correlation coefficient (r=0.98; there was no significant difference with the control condition; data not shown). The elevation response components, however, are poorly related to target elevation (gain b=0.07, and r=0.10). This is to be expected, as the LP noise stimuli do not carry elevation-specific spectral shape information (see also Fig. 1). The considerable downward bias in the response elevation components, $a \approx -13$ deg, was typical for all subjects.

The center and bottom panels (Fig. 2C–F) show the responses of this subject for the pooled HP and BB spectra (also pooled for the two intensities), separated for the free and mold ear. In this case, the responses were described with linear regression lines (Eq. 2), but separately for the left and right hemifields. The center panels (Fig. 2C, D) show the saccade components toward targets presented on the side of the free (i.e., left) ear ($\alpha_T < -4$ deg). Note that both the azimuth (Fig. 2C) and the elevation (Fig. 2D) responses are very consistent (a=4 deg, b=1.0 for azimuth, and a=1 deg, b=0.68 for elevation) and close to the control condition.

The bottom panels (Fig. 2E, F) show the responses for stimuli presented on the side of the mold ($\alpha_T > +4$ deg). A comparison with the leftward responses shows that the azimuth components for these rightward saccades reach a similar accuracy. Although the elevation components of the rightward saccades still correlate well with the actual target elevation (r=0.77), the regression analysis of Eq. 2 shows that both the bias (a=5 deg) and the gain (b=0.47) differ from the saccades contralateral to the mold.

A similar result for the HP and BB stimuli was obtained for localization with a mold in the left ear for this subject: in this condition, the gain of the elevation responses was lower for the leftward saccades (bias: a=+3 deg; gain: b=0.36), whereas now the rightward saccades were significantly closer to the control responses (bias: a=–3 deg; gain: b=0.59).

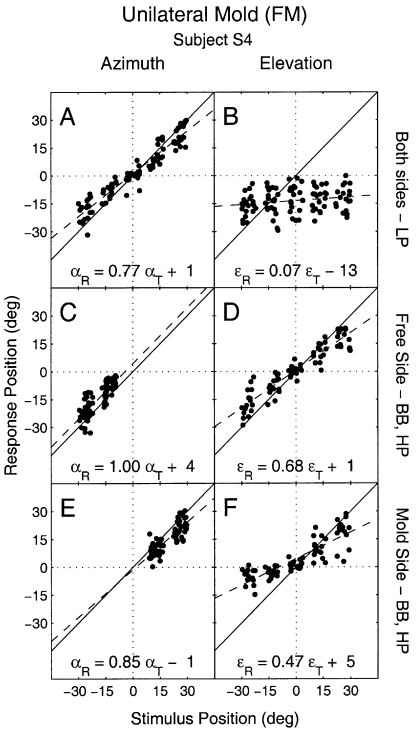

Fig. 2A–F Typical example of the linear regression analysis (Eq. 2, given in panels) applied to the data from subject S4 for azimuth (*left-hand column*) and elevation (*right-hand column*) response components, respectively, in the unilateral mold condition (*FM*). High- and low-intensity evoked responses are pooled. *Top row* Saccades to LP stimuli. Note that azimuth responses are accurate, but the subject is unable to extract target elevation. *Center row* Regression results, pooled for BB and HP stimuli, for the free (*left*) side only. Both azimuth and elevation responses are accurate. *Bottom row* Same as center row, for responses to the side of the mold. Azimuth is as accurate as on the left side, but elevation gain is significantly reduced

A summary of the linear regression analysis (Eq. 2) for the elevation response components of left and right mold conditions for all four subjects is provided in Fig. 2. In this figure, the gains of the linear regression analysis, performed separately for the mold side and the free side, are plotted against each other for each individual recording session, and for all subjects. The data have been separately analyzed for the three different stimulus spectra: LP (open squares), HP (solid triangles), and BB (open circles). Each data point is obtained from a different session, and is based on a regression of at least 50 data points.

In Fig. 3A it can be seen that for all subjects and all recording sessions, the fitted gains for the LP stimuli were close to zero. For the BB and HP stimuli the gains all differed significantly from zero. Note, however, that the gains obtained from responses to targets presented at the mold side were systematically lower than for the targets presented on the free side. This trend was similar for both the BB and the HP stimuli.

This is further illustrated in Fig. 3B, in which the cumulative probabilities of the gain differences between the free side and the mold side are plotted. The gain

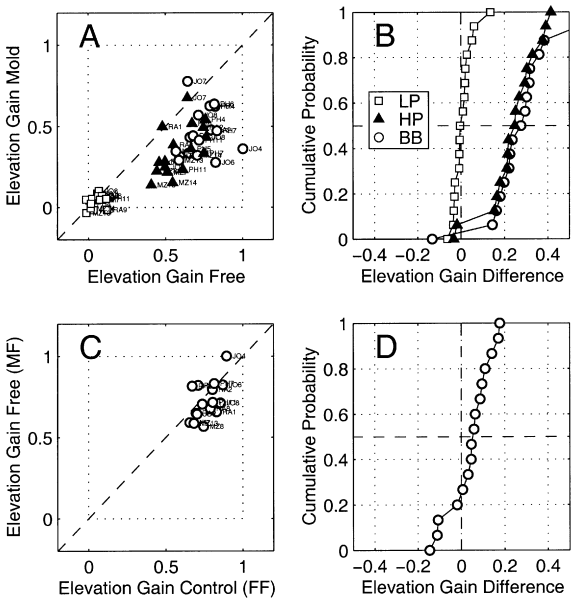

Fig. 3 **A** Elevation gains obtained for saccades into the free vs. the mold side in the unilateral mold localization condition (as in Fig. 2). Left- and right-side mold conditions are shown for all subjects and recording sessions. *Symbols* refer to different stimulus spectra (see *inset*). For both HP and BB stimuli, the elevation gains on the free side are higher than for the mold side, in all but three cases. Gains for the LP stimuli scatter around zero for both conditions. **B** Cumulative probabilities of elevation gain differences in the unilateral mold condition. Curves are systematically shifted rightward for the HP and BB stimuli. Both curves superimpose. On average, the gain on the free side is about 0.25 higher than on the mold side, as $P(diff \geq 0.25)=50\%$. No significant difference was obtained for the LP stimuli, as the cumulative probability curve is symmetric around zero. **C** Elevation gains to sounds in the free side are compared with control responses (both ears free) to the same side. Note slightly lower gains for the unilateral mold conditions. **D** The cumulative distribution of gain differences (control minus mold condition) is slightly shifted rightward

difference is taken as the vertical distance from each data point in Fig. 3A to the identity line. Since the gain differences scatter symmetrically around zero for the LP stimuli, localization performance is the same for either the mold or the free side. The cumulative distribution is therefore symmetrical around zero, and $P(diff \leq 0)=50\%$. However, the cumulative distributions are shifted rightward for both the HP and the BB stimuli by a similar amount. Thus, for the great majority of experimental sessions, the gain difference is positive [for these stimulus spectra, $P(diff \geq 0.25)=50\%$, Kolmogorov-Smirnov test: $P<0.001$].

The panels in Fig. 3C, D compare the elevation gains for sounds presented on the side of the *free* ear for both the FM and MF conditions with the control responses to that same side. Note that the majority of values lie slightly below the diagonal (panel C) and that the cumulative distribution of gain differences is shifted rightward by approximately 0.05. Thus, in the unilateral mold condition, responses to the free side appear to be slightly, but consistently, affected by the mold too (KS test, $P<0.01$).

Thus, the impression gained from these data is that in the unilateral mold condition the elevation components of responses for HP and BB sounds toward the side of the mold are affected in both their gain and their bias parameters. Responses toward the side of the free ear are only slightly affected. Typically, the response gain is lower on the side of the mold, and the bias often differed substantially (not shown). Azimuth responses remain unaltered for either side. Thus, these data suggest that response elevation not only depends on target elevation (ε_T; as in Eq. 2), but also on target azimuth (α_T).

The data shown in Figs. 2 and 3, however, do not allow for a more accurate assessment of this effect, as the saccades were selected for either ipsilateral or contralateral targets with respect to the median plane ($|\alpha_T| \geq 4$ deg) and gains were determined for the pooled responses on each side. Therefore, to investigate in a quantitative way how response elevation depends on target azimuth over the *entire* oculomotor range, the linear regression model was extended by allowing the bias and gain of Eq. 2 to be linear functions of target azimuth, as follows:

$$\begin{aligned} \varepsilon_R &= a'(\alpha_T) + b'(\alpha_T) \cdot \varepsilon_T \\ &= (a + c \cdot \alpha_T) + (b + d \cdot \alpha_T) \cdot \varepsilon_T \\ &= a + c \cdot \varepsilon_T + c \cdot \alpha_T + d \cdot \alpha_T \cdot \varepsilon_T \end{aligned} \quad (3)$$

where the four parameters *a, b, c,* and *d* are multiple linear regression coefficients (see "Materials and methods"). Parameter *a* is the bias (in deg), *b* is the gain of the response component. Parameter *c* introduces an azimuth-dependent bias to the elevation responses, whereas *d* (in deg^{-1}) ensures that the gain of the elevation responses depends on azimuth too.

To allow for a direct comparison of the azimuth-dependent changes (*c* and *d*) with the static bias (*a*) and gain (*b*) values, Eq. 3 was normalized:

$$\varepsilon_R = (a + \Delta a \cdot \hat{\alpha}_T) + (b + \Delta b \cdot \hat{\alpha}_T) \cdot \varepsilon_T \quad \text{where}$$
$$\hat{\alpha}_T \equiv \alpha_T / \alpha_{max} \quad (4)$$

and $\alpha_{max}=30$ deg. Note that the normalized target azimuth, $\hat{\alpha}_T$, lies between –1 and +1. In what follows, the terms 'bias drift' and 'gain drift' refer to parameters Δa and Δb, respectively. An analogous description was applied to the azimuth data. This model, which is the simplest extension of Eq. 2, yet is capable of describing the azimuth-dependent effects, will be referred to as the 'twist model'.

Figure 4 shows the results of applying Eq. 4 to both the azimuth and elevation response components of subject S1, pooled for HP/BB spectra and low/high stimulus intensities, for the two unilateral mold conditions (Fig. 4B, C), as well as for the control responses (Fig. 4A). Although the control responses of this subject were accurate, a small but systematic asymmetry in the data distribution is also apparent. The bias drift was slightly positive (Δa=+1.6 deg, which is visible in the $\varepsilon_R=0$ line running upward), and the gain drift had a significant negative value (Δb=–0.13). The latter property is visible as a small compression of the elevation range for the

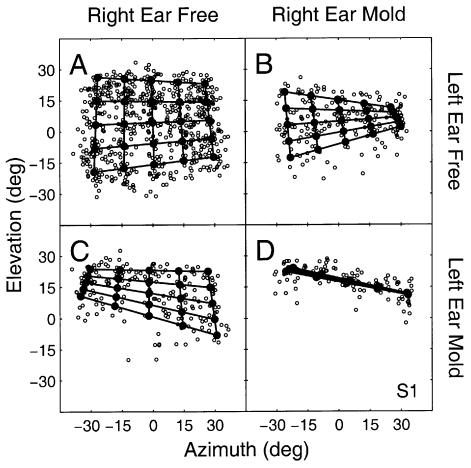

Fig. 4A–D Results of applying the twist model fit (grids) to the localization data of subject S1 for four hearing conditions. The grids are constructed by connecting the predicted responses (α_R, ε_R) (*black dots*; Eq. 4) to target locations at (α_T, ε_T)=0, ±15, ±30 deg, after which neighboring targets are interconnected. *Open circles* indicate the individual responses. **A** The control condition (both ears free); **B** a unilateral mold in the right ear; **C** a unilateral mold in the left ear; **D** molds are applied bilaterally. See text for details

rightward responses. The drift in gain over the entire azimuth domain is about $|\Delta b|/b$=16% of the static value. A similar observation can be made for the azimuth components: the isoazimuth lines of the fit are slightly compressed for the upward targets, which emerges as a small, but significant, gain drift (Δb=–0.08). The bias drift, however, is not significantly different from zero, since the α_R=0 line remains vertical. This particular response behavior appeared to be idiosyncratic, as can be deduced from the fit results presented in Table 1, and may possibly reflect a property of oculomotor behavior, rather than a subject's sound localization ability. For two of the subjects (S3 and S4), the control data are equally well described by Eq. 2 (Δa and Δb are essentially zero for both components in these subjects). For the other two subjects, the small corrections provided by Eq. 4 yielded a slightly better description of the control data.

From the data obtained from the unilateral mold conditions a different picture emerges. As can be clearly seen in Fig. 4B (FM condition) and Fig. 4C (MF condition), the fitted response surfaces are now strongly asymmetric in the elevation domain, whereas they are close to the control situation for the azimuth components (data from subject S1; see also Table 2). Note that the bias

Table 2 Parameters from fits of the unilateral-mold and bilateral-mold data to the twist model for both azimuth and elevation. HP and BB noise stimuli have been pooled. Stimulus intensity is 60 dB SPL. Note that the relative variation in elevation gain, $\Delta b/b$, is significant (*) and large for all subjects in both unilateral-mold conditions. Moreover, note that the sign of this ratio is consistent within each mold condition. *Dashes* indicate cases of small gain values for which this ratio could not be computed ('division by zero'). See Table 1 for further details

Subject	a	Δa	b	Δb	$\Delta b/b$	r	N
Azimuth response component							
Unilateral mold left							
S1	–1.9 (0.3)	0.1 (0.5)	1.18 (0.02)	–0.10 (0.03)	–0.08 (0.02)*	0.99	197
S2	1.6 (0.2)	2.2 (0.4)	1.12 (0.02)	0.09 (0.02)	0.08 (0.02)*	0.99	277
S3	–8.6 (0.3)	2.5 (0.5)	1.20 (0.02)	–0.07 (0.03)	–0.06 (0.03)*	0.99	199
S4	2.5 (0.3)	1.1 (0.4)	0.73 (0.02)	–0.02 (0.03)	–0.02 (0.03)	0.97	200
Unilateral mold right							
S1	1.5 (0.3)	–2.2 (0.5)	1.00 (0.02)	–0.02 (0.03)	–0.02 (0.03)	0.98	197
S2	3.6 (0.3)	–3.3 (0.5)	1.14 (0.02)	0.03 (0.03)	0.03 (0.03)	0.99	200
S3	–2.3 (0.3)	1.2 (0.5)	1.16 (0.02)	–0.05 (0.03)	–0.05 (0.03)	0.98	200
S4	0.2 (0.3)	–2.8 (0.5)	0.79 (0.02)	0.01 (0.04)	0.02 (0.05)	0.97	195
Bilateral mold							
S1	4.6 (0.4)	0.9 (0.6)	1.09 (0.02)	–0.05 (0.03)	–0.04 (0.03)	0.98	100
S2	4.2 (0.4)	–0.2 (0.5)	1.13 (0.02)	0.03 (0.03)	0.03 (0.03)	0.99	99
S3	–7.1 (0.4)	1.0 (0.7)	0.87 (0.03)	–0.05 (0.05)	–0.06 (0.05)	0.97	100
S4	2.6 (0.3)	–1.0 (0.5)	0.72 (0.02)	–0.01 (0.03)	–0.01 (0.04)	0.97	200
Elevation response component							
Unilateral mold left							
S1	12.2 (0.5)	–5.8 (0.7)	0.42 (0.03)	0.20 (0.04)	0.47 (0.09)*	0.81	197
S2	–7.1 (0.4)	–0.6 (0.6)	0.69 (0.03)	0.14 (0.04)	0.20 (0.06)*	0.91	277
S3	7.2 (0.6)	–4.7 (1.0)	0.52 (0.04)	0.26 (0.07)	0.50 (0.13)*	0.80	199
S4	–0.2 (0.4)	–4.6 (0.7)	0.48 (0.03)	0.20 (0.04)	0.41 (0.08)*	0.90	200
Unilateral mold right							
S1	5.1 (0.5)	2.1 (0.8)	0.39 (0.03)	–0.25 (0.05)	–0.64 (0.13)*	0.80	197
S2	–6.6 (0.4)	–2.1 (0.6)	0.58 (0.03)	–0.28 (0.04)	–0.49 (0.07)*	0.92	200
S3	1.5 (0.4)	–0.1 (0.6)	0.66 (0.02)	–0.08 (0.04)	–0.12 (0.06)*	0.94	200
S4	2.6 (0.4)	2.9 (0.6)	0.57 (0.03)	–0.14 (0.04)	–0.25 (0.07)*	0.91	195
Bilateral mold							
S1	17.5 (0.5)	–6.8 (1.0)	0.03 (0.04)	–0.01 (0.06)	–	0.65	100
S2	–3.4 (0.5)	–5.8 (0.8)	0.07 (0.04)	0.00 (0.05)	0.06 (0.64)	0.61	99
S3	7.9 (0.5)	–4.4 (0.9)	–0.03 (0.03)	–0.09 (0.06)	–	0.51	100
S4	25.0 (0.3)	1.6 (0.4)	0.10 (0.02)	–0.04 (0.03)	–0.39 (0.22)	0.48	200

drift for the elevation components (Δa) contributes more to the description of the data in Fig. 4C than in Fig. 4B: Δa=–5.8±0.7 deg vs. Δa=2.1±0.7 deg, respectively. This is also visible in the response patterns for the far-left targets in Fig. 4C versus the far-right targets in Fig. 4B: the average response elevation is higher in the former as compared to the latter. Also the gain drift (Δb) is significantly different from zero in both unilateral mold conditions: Δb=0.20±0.04 (left mold; 47% of the static gain) and Δb=–0.25±0.05 (right; 64% of the static gain), respectively.

The description provided by Eq. 4 is far from perfect, as deviations are still apparent. Yet, the simple first-order extension of Eq. 2 already leads to a considerable improvement of the data fits. For the example of Fig. 4C, the correlation between fit and data increased from r=0.71 for Eq. 2 to r=0.81 for Eq. 4, whereas the mean absolute residue decreased from 5.9 to 4.8 deg. This improvement of the twist model over the simple linear model was verified for all stimulus types, recording sessions and subjects. The fits for the azimuth components indicate that the linear model Eq. 2 performs equally well, since both Δa and Δb remain close to zero.

The results from the non-linear twist model confirm and further extend the findings presented in Figs. 2 and 3: in the unilateral mold condition, localization performance of targets presented on the ipsilateral side of the free ear is close to the control result, whereas elevation performance depends in a systematic way on azimuth for targets presented at other locations.

All four subjects yielded similar results, although quantitative differences between subjects, and between the effects of the molds on the left and right side, may also be noted (see Table 2).

As expected, the low-pass stimuli yielded no azimuth-dependent elevation responses. For all conditions, the gains and biases for elevation were indistinguishable from the control values (not shown).

Bilateral mold condition

The results of the unilateral mold experiments suggest that the percept of sound-source elevation is not only based on the extraction of spectral shape information by the ipsilateral ear. In that case, the elevation responses toward the side of the mold should have yielded no correlation with target elevation. As these correlations were highly significant, the results instead indicate that the percept of target elevation is weighted by the spectral information from *both* ears. For this conclusion to be valid, however, it has to be demonstrated that the molds were indeed effective in removing the elevation-related spectral cues. To that end, experiments were conducted in which subjects wore a mold in both ears.

The responses toward the HP and BB stimuli differed dramatically from the unilateral mold conditions, in that the elevation responses were now completely compressed into a narrow response range (see also Hofman et al. 1998), which was typically at an upward elevation in the current study (see Table 2). Interestingly, the elevation response bias depended systematically on perceived stimulus azimuth. An example of the particular response pattern for this condition is provided in Fig. 4D for subject S1. Note the systematic decrease in mean elevation responses as target azimuth varies from left to right. This is expressed by a substantial bias drift, Δa=6.8±1.0 deg. Since the responses are all compressed into a thin region, the fitted surface almost resembles a line in this plot. This is quantified by the negligible gain b=–0.01±0.06. As shown in Table 2, the response gains for all four subjects were very low, and insignificant in three out of four.

In summary, the data presented so far suggest a relative weighting of spectral cues from either ear. An interesting question is whether this weighting process changes gradually as a function of target azimuth.

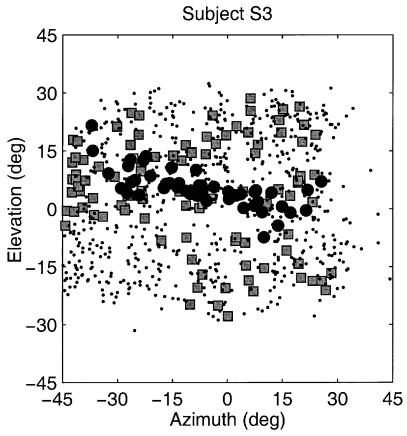

Fig. 5 Saccadic eye movements of subject S3, plotted as response azimuth against response elevation, are compared for three different hearing conditions: MF, MM, and C (*filled gray squares* MF condition, *small dots* control responses, *large filled dots* MM responses). Note that for the control responses there is no systematic dependence on stimulus azimuth. The binaural mold responses, however, gradually change from about +15 deg on the far left, to near zero deg for far-right target locations. The distribution of MF responses closely corresponds to the MM responses for far-left targets, but approaches the control condition for the far-right stimuli

Binaural weighting of spectral shape cues

The unilateral mold experiments show that elevation gains for targets ipsilateral to the mold are systematically lower than for the free side (Fig. 3A, B). In addition, elevation responses for targets presented at the mold side are much more accurate than for the bilateral mold condition (Fig. 4D).

To enable a direct comparison between the different hearing conditions, Fig. 5 shows all elevation localization responses of subject S3 in the MF condition (gray squares), MM condition (large filled dots), and control condition (small dots) as a function of response azimuth.

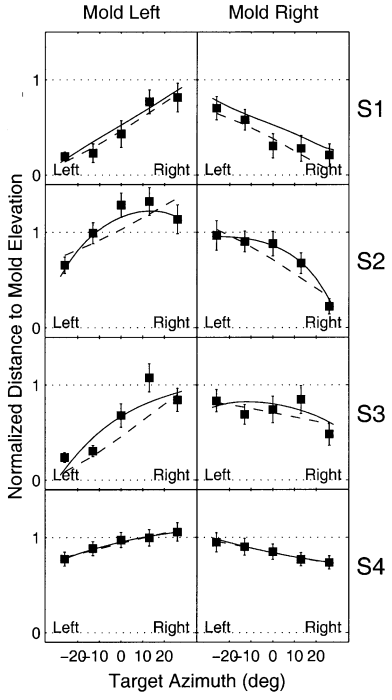

Fig. 6 Normalized weighting index for the unilateral mold conditions, described by index. Data are from all four subjects (*rows*). Left and right-mold hearing conditions correspond to the *left- and right-hand columns*, respectively. *Filled squares* depict the measured indices (with 1 SD). Note the gradual increase (decrease) of the index with target azimuth, for the left- (right-) mold hearing condition, which was observed in all four subjects. The *dashed-dotted lines near the bottom of each panel* show the lower bounds for the indices, corresponding to the scatter in the mold-induced illusory elevation (see also text). The *dashed lines* are the computed weighting indices as predicted by the twist model (Eq. 4) fitted to the response data. The *solid lines* represent the index as predicted by an extension of the model (*a'* and *b'* as second-order polynomials; see "Materials and methods")

Note that the control responses do not appear to depend in a systematic way on target azimuth as the data points are evenly distributed across the entire elevation range for all azimuth locations. Note the much smaller variance, and the systematic decrease in response bias from +15 (left) to 0 deg (right), for the MM condition. The MF data lie between these two extremes: at far-left azimuths, both the low variance and the positive average bias of the MF responses approach those of the MM condition. For far-right targets, however, the distribution of data points is more similar to the control data. This is particularly apparent for the increased variance of the data, but also for the comparable average elevation biases. For intermediate azimuth positions, the MF data appear to change gradually from one type of behavior (MM) to the other (C).

These data suggest that elevation responses to either side result from a *binaural weighting* process, rather than from an ipsilateral *monaural* analysis. In the latter case no localization ability for targets ipsilateral to the mold is expected, and no change in accuracy regarding control for sounds presented on the side of the free ear.

To further quantify this putative binaural interaction, we introduce a normalized azimuth-dependent weighting index, $\delta(\alpha)$, which relates the observed elevation responses for each individual mold to the perceived azimuth position. The index is defined such that it approaches zero when the response distribution for the unilateral mold condition is indistinguishable from the binaural mold condition, and such that it equals one when responses are distributed as in the control condition (see "Appendix" for quantitative details).

Figure 6 shows the computed weighting index for all four subjects, together with the prediction resulting from the twist model (Eq. 4, dashed line). In addition, a higher-order data fit is also included (solid curves; see "Materials and methods"). Note that for all measured ears, the weighting index varies systematically with target azimuth. This indicates that *gradual weighting* of inputs from the left and right ear underlies the measured elevation percepts. As stimuli move into the free side, the index approaches one, whereas on the side of the mold it is lowest, with the median-plane data lying between these extremes.

Discussion

Summary

This study investigated binaural interactions of spectral shape cues subserving sound localization of targets within the oculomotor range. To that end, the spectral shape cues were altered by inserting a mold in either ear, or in both ears, while measuring the saccadic eye movement responses toward targets of varying spectral content and intensity.

The bilateral mold experiments demonstrate that the molds effectively perturbed the original spectral pinna cues, as elevation localization was completely abolished on both sides (Fig. 4D). In addition, unilateral molds degraded the localization of sound elevation for targets presented contralateral to the mold (Fig. 3C, D). Conversely, spectral cues from the free ear improved elevation localization on the side carrying the mold, as compared to the bilateral mold condition (Figs. 1, 2, 3). Disrupting the spectral shape cues had no measurable effect on azimuth localization in either the unilateral (Fig. 2) or the bilateral mold condition (Table 2). This further underlines the dominance of (undisturbed) interaural phase and intensity differences for the percept of stimulus azimuth.

A quantitative comparison of the accuracy of elevation responses in the unilateral mold condition with those of the bilateral mold and control conditions (Fig. 6) indicated a gradual change of the weighting index with target azimuth.

We conclude from these findings that spectral shape cues from the left and right ear are weighted to construct an elevation percept, and that perceived stimulus azimuth acts as a weighting factor.

Comparison with earlier studies

The saccade accuracy results of the binaural mold (MM) experiment confirm earlier studies (Oldfield and Parker 1984; Hofman et al. 1998; Morimoto 2001). The subject's percept of sound elevation is abolished, whereas sound-source azimuth is still localized as accurately as in the control condition. These results are also in line with the more subtle experimental manipulations of Wightman and Kistler (1989, 1992, 1997), who applied random roving of high-frequency sub-bands in their dichotic stimuli.

In our experiments, subjects typically responded to a narrow range of elevations (Fig. 4D), the value of which depended in an idiosyncratic, but systematic, way on stimulus azimuth. In our quantitative description of Eq. 4, this dependence emerges as a significant drift bias, Δa. Note that this interesting feature is not captured by the linear regression of Eq. 2.

A possible explanation for this dependence on perceived stimulus azimuth might be that the molds provided each ear with different spectral cues (none of which is related to the original pinna cues), which are still interpreted by the auditory system as two different sound elevations. As can be seen in Fig. 4, for example, the left-ear mold leads to an elevation response of approximately +24 deg, for far-left ($\alpha \approx -26$ deg) targets, and to about +11 deg for far-right ($\alpha \approx +26$ deg) targets. It may also be noted that the data show a systematic decrease from the far-left elevation estimate, to the far-right elevation estimate, which could be interpreted as a result of the weighting process. Thus, for targets presented near the midline, the overall estimate of elevation in this subject is about +17 deg (see also Table 2). According to the twist function, the drift bias for this subject in the binaural-mold condition was $\Delta a = -6.8$ (Table 2), which captures this azimuth-dependent change of the elevation percept. The values obtained for the other subjects were different, which is likely caused by differences in the exact spectral cues produced by the molds, in combination with the way each subject's auditory system has learned to interpret these cues on the basis of the original pinna spectra (e.g., Middlebrooks 1992).

Our monaural mold experiments (MF/FM) extend earlier binaural localization studies to the frontal oculomotor range. As argued by Morimoto (2001), this is not a trivial extension, as in this target domain the acoustic inputs to either ear are comparable and therefore likely to interact.

Our data also support and further extend Humanski and Butler's (1988) and Morimoto's (2001) demonstration of a near-ear dominance for far-lateral targets. Indeed, the weight index appears to be lowest for the far-left (in the case of the MF condition) and far-right (FM) targets, indicating that for targets in that range the percept is nearly fully dominated by the mold. By showing that this dominance diminishes gradually from ipsilateral to contralateral locations (Fig. 6), even over a range as restricted as 30 deg on either side of the midsagittal plane, our data show that both ears are gradually weighted to the percept of elevation in the frontal target domain. By means of the non-linear description of the data, this gradual change is manifest from the significant values found for the azimuth-dependent drifts in both gain and bias (Table 2).

By analyzing the data through a direct comparison across all hearing conditions it is shown that the weight index $\delta(\alpha)$ changes in a gradual way as a function of perceived azimuth (Figs. 5, 6). Thus, the monaurally perceived elevations of either ear seem to be fused into a single, weighted percept of target elevation. By using rapid saccadic eye movements as a pointer, our data show that this binaural weighting takes place already at the earliest percept underlying spatial hearing (within about 200–250 ms). In the oculomotor literature, similar rapid behavior towards visual stimuli has been described as *weighted averaging*.

Possible neural mechanisms

Weighted averaging has been studied extensively for saccades evoked by two simultaneously presented visual targets. In such experiments, saccades are a weighted averaging response that is determined by the relative size and intensity of the two stimuli, and by their eccentricity relative to the fovea (Findlay 1982; Ottes et al. 1984). This phenomenon has been explained by local neural interactions within topographically organized visuomotor maps, most notably in the midbrain superior colliculus (SC), a structure known to be crucial for the generation of saccadic gaze shifts to targets of multiple modalities (see Sparks and Mays 1990; Stein and Meredith 1993, for reviews).

Response averaging to combined visual-auditory stimuli occurs when the salience of the visual stimulus is reduced (e.g., Lueck et al. 1990; Frens et al. 1995), suggesting that visual input typically contributes with a large weight to the audiovisual response. Although little evidence exists for the neural basis underlying this phenomenon, it has also been explained by spatial interactions within the SC (Lueck et al. 1990).

Visuomotor studies have shown that averaging only occurs when the spatial separation of the targets is not too large. Otherwise, short-latency visuomotor responses become *bistable* (Ottes et al. 1984) in the sense that saccades are directed to either target. Such bistable behavior, however, was not observed in our results.

The observed elevation responses in the present study indicate that conflicting spectral cues induce the percept of a single, weighted, target position. At present it is not known whether this binaural fusion process might follow similar rules to visually evoked responses. If so, large apparent separation (spatial or spectral, see below) of acoustic inputs would induce the percept of two stimuli at different elevations, rather than one averaged stimulus location. Indeed, earlier dichotic lateralization studies have indicated that time-intensity trading dissolves as interaural intensity and timing localization cues are

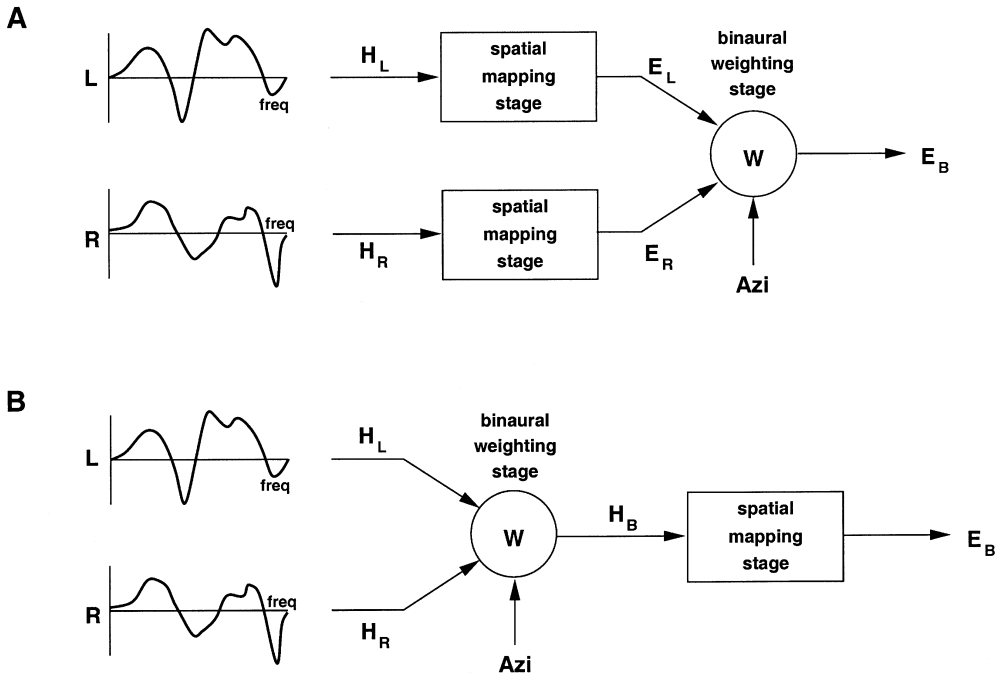

Fig. 7A, B Two conceptual schemes that may account for the response elevation data. Two different stages underlie the acoustic orienting response: a spectral-to-spatial mapping stage in which the pinna-related spectral cues are transformed into an estimate of sound elevation, and a binaural fusion stage, in which the information from both ears is combined to produce a binaural estimate, which is weighted by perceived target azimuth. The two schemes differ in the order in which these two stages are implemented. **A** The spatial weighting scheme. **B** The spectral weighting scheme. See text for further details

separated. Subjects then typically perceive two spatial images: a 'time image' and an 'intensity image' (Whitworth and Jeffress 1961; Hafter and Jeffress 1968).

An interesting question is how the observed binaural weighting of elevation (Fig. 6) might be implemented in the auditory system. Two conceptually different schemes are outlined in Fig. 7. To appreciate the problem, it is convenient to identify two different stages in the transformation from acoustic input to spatial localization: a stage in which spectral input (tonotopically represented) is transformed into an estimate of target elevation (e.g., by determining the maximal spectral correlation between stored (learned) pinna filters and the acoustic input at the eardrums, as proposed by Middlebrooks (1992), and a stage in which binaural inputs are fused, and weighted by perceived azimuth (as found in this study, and reported by Morimoto 2001).

As a first possibility, the binaural interactions could occur at a stage where acoustic information is already represented *spatially*, rather than *tonotopically*. Such a scheme is represented in Fig. 7A (the spatial weighting scheme), and has been described earlier by Middlebrooks (1992). Alternatively, the weighting could occur within the tonotopic representations of the auditory system (Fig. 7B; the spectral weighting scheme), prior to the spectral-spatial transformation. Note that although the details of the two stages remain unspecified, both the spectral-spatial mapping and the binaural fusion and weighting stage are non-linear processes. As a corollary, the order in which these two mechanisms are implemented matters for the overall transformation that relates the spectral inputs to binaurally perceived target elevation. Thus, the two schemes represent different models that could, at least in principle, be dissociated on the basis of an input-output analysis.

The current experiments, however, do not allow a distinction between the two schemes. So far, it is not known whether the information from both ears gives rise to two separate representations of left-ear and right-ear elevations (as in Fig. 7A) or to only one, unified or weighted representation (as in Fig. 7B).

Recently, lesions in the cat dorsal cochlear nucleus (DCN) have suggested a role of DCN in encoding the spectral localization cues (May 2000) that may support the scheme proposed in Fig. 7A. In the primate auditory system, the existence of a spatial map of head-centered auditory space has not been demonstrated so far, although the mammalian auditory cortex and adjacent areas have been implicated in sound localization (Middlebrooks et al. 1998; Rauschecker 1998; Xu et al. 1998; Recanzone et al. 2000). So far, however, an explicit (binaural, 2D) spatial map has not been demonstrated in these structures. The midbrain SC could be an alternative candidate for its known 2D spatial representation of auditory space which, however, appears to represent targets in oculocentric, rather than in craniocentric, coordinates (Jay and Sparks 1984). It would therefore be of interest to determine whether the weighting indices of Fig. 6 are better expressed in head-centered, eye-centered or space-centered coordinates.

Acknowledgements We thank Ton van Dreumel and Hans Kleijnen for their valuable technical support. This research was supported by the Netherlands Organization for Scientific Research (P.H.; 805–33.705-P), the Human Frontier Science Program (RG 0174/1998-B, P.H. and A.J.V.O.) and the University of Nijmegen (A.J.V.O.).

Appendix

In the unilateral-mold condition, weighting index δ represents the relative contribution of the normal pinna cues, with respect to the (spectral) mold cues, to the resulting elevation percept. For condition MF (mold in left ear), δ is computed as follows.

First, the mold-induced 'illusory' elevation is estimated from elevation responses in the bilateral mold condition (MM) to far-left targets, for which it is assumed that the perceived elevation depends largely on the spectral cues provided by the left mold (Humanski and Butler 1988). Therefore, left-mold elevation, ε_{L}, is computed as the average response elevation for targets in the far-left section, in this case at $\alpha_L \equiv -26$ deg (see Fig. 4C for an example). Then, for the MF responses, the target azimuth domain is divided into five 8-deg-wide sections, centered at $\alpha_{\mathrm{T}}=0$, ± 13, and ± 26 deg.

Next, it is estimated how responses compare to the illusory elevation in each condition. To that end, the average absolute distance between response elevation and the mean illusory left-mold elevation is computed for each target section. In general, for condition X (either MF, MM or C) and azimuth section, this yields for the average distance $\Delta_L^X(\alpha)$:

$$\Delta_L^X(\alpha) = \frac{1}{N}\sum_{i=1}^{N} \left|\varepsilon_i^X(\alpha) - \varepsilon_{\mathrm{L}}\right|$$

where $\varepsilon_i^X(\alpha)$ represents the elevation of the ith response in condition X within the section at α, and N is the number of targets presented in each sector. Note that ε_{L} is in fact the average of response elevations $\varepsilon_i^{MM}(\alpha_{\mathrm{L}})$, and that according to the definition of Eq. 6 $\Delta_L^{MM}(\alpha_{\mathrm{L}})$ equals the mean absolute error in ε_{L}. Finally, the unilateral mold distance, is normalized by the control distance $\Delta_L^C(\alpha)$, yielding the weight index $\delta_L^{MF}(\alpha)$:

$$\delta_L^{MF}(\alpha) = \frac{\Delta_L^{MF}(\alpha) - \Delta_L^{MM}(\alpha_{\mathrm{L}})}{\Delta_L^{C}(\alpha) - \Delta_L^{MM}(\alpha_L)}$$

Note that $\delta_L^{MF}(\alpha) = 0$ corresponds to the situation that responses at α are distributed similarly to those for the far-left bilateral mold condition. If, in contrast, $\delta_L^{MF}(\alpha) = 1$, the responses are indistinguishable from the control condition at α. (Computation of δ_R^{FM} for unilateral-mold condition FM is done in an analogous way.)

Finally, note that $\delta(\alpha)$ is not confined to the 0, 1 range: if the variance in the MF responses happens to exceed that of the controls, the index will be larger than one. Although rare, this was occasionally observed in our data (see Fig. 6). If, on the other hand, the MF variance was smaller than that of the MM responses, the index would be negative. This was never observed in our data.

References

Blauert J (1997) Spatial hearing. The psychophysics of human sound localization. Revised edn. MIT Press, Cambridge, MA

Butler RA, Humanski RA (1992) Localization of sound in the vertical plane with and without high-frequency spectral cues. Percept Psychophys 52:182–186

Butler RA, Musicant AD (1993) Binaural localization: influence of stimulus frequency and the linkage to covert peak areas. Hearing Res 67:220–229

Collewijn H, Van der Mark F, Janssen TC (1975) Precise recording of human eye movements. Vision Res 15:447–450

Findlay JM (1982) Global processing for saccadic eye movements. Vision Res 22:1033–1045

Flannery R, Butler RA (1981) Spectral cues provided by the pinna for monaural localization in the horizontal plane. Percept Psychophys 29:438–444

Frens MA, Van Opstal AJ (1995) A quantitative study of auditory-evoked saccadic eye movements in two dimensions. Exp Brain Res 107:103–117

Frens MA, Van Opstal AJ, Van der Willigen RF (1995) Spatial and temporal factors determine auditory-visual interactions in human saccadic eye movements. Percept Psychophys 57:802–816

Goossens HHLM, Van Opstal AJ (1999) Influence of head position on the spatial representation of acoustic targets. J Neurophysiol 81:2720–2736

Hafter ER, Jeffress LA (1968) Two-image lateralization of tones and clicks. J Acoust Soc Am 44:563–569

Harris GG (1960) Binaural interaction of impulsive stimuli and pure tones. J Acoust Soc Am 32:685–692

Hebrank JH, Wright D (1974) Are two ears necessary for localization of sound sources on the median plane? J Acoust Soc Am 56:935–938

Hofman PM, Van Opstal AJ (1998) Spectro-temporal factors in two-dimensional human sound localization. J Acoust Soc Am 103:2634–2648

Hofman PM, Van Riswick JG, Van Opstal AJ (1998) Relearning sound localization with new ears. Nat Neurosci 1:417–421

Humanski RA, Butler RA (1988) The contribution of the near and far ear toward localization of sound in the sagittal plane. J Acoust Soc Am 83:2300–2310

Jay MF, Sparks DL (1984) Auditory receptive fields in primate superior colliculus shift with changes in eye position. Nature 309:345–347

Lueck CJ, Crawford TJ, Savage CJ, Kennard C (1990) Auditory-visual interaction in the generation of saccades in man. Exp Brain Res 82:149–157

Makous JC, Middlebrooks JC (1990) Two-dimensional sound localisation by human listeners. J Acoust Soc Am 87:2188–2200

May BJ (2000) Role of the dorsal cochlear nucleus in the sound localization behavior of cats. Hearing Res 148:74–87

Middlebrooks JC (1992) Narrow-band sound localization related to external ear acoustics. J Acoust Soc Am 92:2607–2624

Middlebrooks JC, Green DM (1991) Sound localization by human listeners. Ann Rev Psychol 42:135–159

Middlebrooks JC, Xu L, Clock Eddins A, Green DM (1998) Codes for sound-source location in nontonotopic auditory cortex. J Neurophysiol 80:863–881

Mills AW (1958) On the minimum audible angle. J Acoust Soc Am 30:237–246

Morimoto M (2001) The contribution of two ears to the perception of vertical angle in sagittal planes J Acoust Soc Am 109:1596–1603

Oldfield SR, Parker SP (1984) Acuity of sound localization: a topography of auditory space. II. Pinna cues absent. Perception 13:601–617

Oldfield SR, Parker SP (1986) Acuity of sound localization: a topography of auditory space. III. Monaural hearing conditions. Perception 15:67–81

Ottes FP, Van Gisbergen JAM, Eggermont JJ (1984) Metrics of saccade responses to visual double stimuli: two different modes. Vision Res 24:1169–1179

Perrott DR, Ambarsoom H, Tucker J (1987) Changes in head position as a measure of auditory localization performance: auditory psychomotor coordination under monaural and binaural listening conditions. J Acoust Soc Am 82:1637–1645

Press WH, Flannery BP, Teukolsky SA, Vettering WT (1992) Numerical recipes in C, 2nd edn. Cambridge University Press, Cambridge, MA

Rauschecker JP (1998) Cortical processing of complex sounds. Curr Opin Neurobiol 8:516–521

Recanzone GH, Guard DC, Phan ML, Su TK (2000) Correlation between the activity of single auditory cortical neurons and sound-localization behavior in the macaque monkey. J Neurophysiol 83:2723–2739

Slattery WH 3rd, Middlebrooks JC (1994) Monaural sound localization: acute versus chronic unilateral impairment. Hearing Res 75:38–46

Sparks DL, Mays LE (1990) Signal transformations required for the generation of saccadic eye movements. Ann Rev Neurosci 13:309–336

Stein BE, Meredith MA (1993) The merging of the senses. MIT Press, Cambridge, MA

Wightman FL, Kistler DJ (1989) Headphone simulation of free-field listening. II: Psychophysical validation. J Acoust Soc Am 85:868–878

Wightman FL, Kistler DJ (1992) The dominant role of low-frequency interaural time differences in sound localization. J Acoust Soc Am 91:1648–1661

Wightman FL, Kistler DJ (1997) Monaural sound localization revisited. J Acoust Soc Am 101:1050–1063

Whitworth RH, Jeffress LA (1961) Time versus intensity in the localization of tones. J Acoust Soc Am 33:925–929

Xu L, Furukawa S, Middlebrooks JC (1998) Sensitivity to sound-source elevation in nontonotopic auditory cortex. J Neurophysiol 80:882–894