Sound Localization Under Perturbed Binaural Hearing

Marc M. Van Wanrooij1,2 and A. John Van Opstal1
1Department of Biophysics, Radboud University Nijmegen, Nijmegen; and 2Department of Neuroscience, Erasmus University Rotterdam Medical Center, Rotterdam, The Netherlands

Submitted 10 March 2006; accepted in final form 20 October 2006

Van Wanrooij MM, Van Opstal AJ. Sound localization under perturbed binaural hearing. J Neurophysiol 97: 715–726, 2007. First published October 25, 2006; doi:10.1152/jn.00260.2006. This paper reports on the acute effects of a monaural plug on directional sound localization in the horizontal (azimuth) and vertical (elevation) planes of human listeners. Sound localization behavior was tested with rapid head-orienting responses toward brief high-pass filtered (>3 kHz; HP) and broadband (0.5–20 kHz; BB) noises, with sound levels between 30 and 60 dB, A-weighted (dBA). To deny listeners any consistent azimuth-related head-shadow cues, stimuli were randomly interleaved. A plug immediately degraded azimuth performance, as evidenced by a sound level–dependent shift (“bias”) of responses contralateral to the plug, and a level-dependent change in the slope of the stimulus–response relation (“gain”). Although the azimuth bias and gain were highly correlated, they could not be predicted from the plug’s acoustic attenuation. Interestingly, listeners performed best for low-intensity stimuli at their normal-hearing side. These data demonstrate that listeners rely on monaural spectral cues for sound-source azimuth localization as soon as the binaural difference cues break down. Also the elevation response components were affected by the plug: elevation gain depended on both stimulus azimuth and sound level and, as for azimuth, localization was best for low-intensity stimuli at the hearing side. Our results show that the neural computation of elevation incorporates a binaural weighting process that relies on the perceived, rather than the actual, sound-source azimuth. It is our conjecture that sound localization ensues from a weighting of all acoustic cues for both azimuth and elevation, in which the weights may be partially determined, and rapidly updated, by the reliability of the particular cue.

INTRODUCTION

Sound localization relies on the neural processing of acoustic cues that result from the interaction of sound waves with the torso, head, and ears. Directional hearing in the horizontal plane (azimuth) depends on binaural differences in sound arrival time and ongoing phase for relatively low (<1.5 kHz) frequencies (so-called interaural time differences [ITDs]). At higher frequencies (>3 kHz) the head-shadow effect causes differences in sound level (interaural level differences [ILDs]). Localization in the vertical plane (elevation) and front–back discrimination require an analysis of spectral shape cues that arise from direction-dependent reflections within the pinna (described by so-called head-related transfer functions [HRTFs]). The latter mechanism essentially constitutes a monaural localization cue for sound frequencies exceeding about 3–4 kHz. However, several studies suggested that the computation of sound-source elevation also involves binaural interactions, as a monaural perturbation of the spectral cues (e.g., by inserting a mold in one pinna), and has a systematic detrimental effect on elevation performance contralateral to the manipulated ear (Hofman and Van Opstal 1998; Humanski and Butler 1988; Morimoto 2001; Van Wanrooij and Van Opstal 2005).

A large body of experimental evidence supports the notion that the processing of the azimuth and elevation components of a sound’s location is embedded in independent neural pathways. In mammals, the ITDs emerge in the medial superior olive (MSO), whereas the ILDs are extracted in another nucleus of the superior olivary complex, the lateral superior olive (LSO; for a recent review see Yin 2002). Evidence also suggests that the first neural correlates of spectral shape analysis may be found in the dorsal cochlear nucleus, which receives monaural input from the ipsilateral ear (Young and Davis 2002).

Psychophysical evidence supports the hypothesis of independent processing of the acoustic cues. Experimental manipulations can considerably degrade elevation localization performance, whereas azimuth localization is far more robust: e.g., by inserting molds, either binaurally (Hofman et al. 1998; Oldfield and Parker 1984) or monaurally (Hofman and Van Opstal 2003; Morimoto 2001; Van Wanrooij and Van Opstal 2005), by introducing background noise (Good and Gilkey 1996; Zwis et al. 2001) or by extensively varying sound levels and sound duration (Hartmann and Rakerg 1993; Hofman and Van Opstal 1998; MacPherson and Middlebrooks 2000; Vliegen and Van Opstal 2004).

Yet, the idea of independent pathways for the processing of a sound’s azimuth and elevation coordinates may be too simple. Clearly, at more central neural stages, such as the midbrain inferior colliculus (IC) and beyond, the outputs of the different cue-processing pathways converge (Chase and Young 2006) and are therefore likely to interact.

Recent psychophysical evidence suggests that the computations underlying the extraction of azimuth and elevation may indeed interact. In a study with unilateral deaf listeners we demonstrated that only listeners who used the spectral-shape cues from their intact ear to localize azimuth could also localize elevation (Van Wanrooij and Van Opstal 2004). In particular, failure of the other listeners to localize elevation was remarkable because these results indicated that under chronic monaural conditions the azimuth and elevation components are not processed independently.

To study the mechanisms underlying the integration of the different acoustic cues, this paper reports on the acute effects of a monaural plug on localization performance of normal-
hearing listeners in the two-dimensional (2D) frontal hemifield. Earlier studies assessed the effect of a monaural plug on human sound localization performance (Flannery and Butler 1981; Musicant and Butler 1984b; Oldfield and Parker 1986; Slattery and Middlebrooks 1994). They all reported as a major effect a horizontal shift ("bias") of localization to the side of the unpluged ear. However, it has remained unclear whether the spectral cues contribute to the localization of azimuth. Although Musicant and Butler (1984a,b, 1985) reported that localization of far-lateral targets relied on monaural spectral cues, none of the plugged listeners in the study of Slattery and Middlebrooks (1994) was able to localize in azimuth. Moreover, Wightman and Kistler (1997) showed that a complete removal of the binaural cues in a dichotic setup abolished sound localization performance altogether. That study therefore suggested that the spectral cues are not sufficient for sound localization in the horizontal plane. In this study we took a different approach, by studying the effect of a plug on localization responses to a variety of acoustic stimuli that varied both in bandwidth and over a considerable range in sound level. The insertion of a plug perturbs the binaural difference cues in normal listeners in a frequency- and level-dependent way and is thus expected to affect the ITDs, ILDs, and HRTFs in different ways. We measured localization across the 2D frontal hemifield immediately after inserting the plug and quantified the changes in localization responses as a function of the acoustic parameters. Our analysis shows that the shift in azimuth responses depends on sound-source location, the sound spectrum, and on sound level. Moreover, performance in elevation is influenced by both sound level and the perceived azimuth location, rather than by the quality of the spectral cues defined by actual stimulus azimuth in the plugged condition. Our data therefore support the hypothesis that the processing of both sound-source azimuth and elevation involve weighted contributions from binaural difference cues as well as from spectral-shape cues. The relative weights of the acoustic cues are adjusted under acoustic perturbations that render a given cue unreliable.

METH ODS

Participants

Five listeners (ages 25–47 yr) participated in the experiments (including both authors, listeners MW and JO). All listeners were experienced with the type of sound localization studies carried out in the laboratory and all had normal hearing (within 20 dB of audiometric zero) as determined by an audiogram obtained with a standard staircase procedure (10 tone pips, 0.5-octave separation, between 500 Hz and 11.3 kHz). None of these listeners had any auditory or uncorrected visual disorder, except for listener JO who is amblyopic in his right eye.

Apparatus

During the experiments, the listener was seated comfortably in a chair in the center of a completely dark, sound-attenuated room (H × W × L = 2.45 × 2.45 × 3.5 m³). The walls, ceiling, floor, and every large object present were covered with black acoustic foam that eliminated echoes for sound frequencies >500 Hz. The room had an ambient background noise level of 25 dB, A-weighted (dBA).

The seated listener faced an array of 58 small broad-range loudspeakers (MSP-30; Monacor International, Bremen, Germany) containing light-emitting diodes (LEDs) in their center. These speakers were mounted on a thin wooden frame that formed a hemispheric surface 100 cm in front of the listener, at polar coordinates R = [0, 15, 30, 45, 60, 75] deg and Φ = [0, 30, …, 300, 330] deg, R is the eccentricity relative to the straight-ahead viewing direction (defined in polar coordinates as [R, Φ] = [0, 0] deg) and Φ is the angular coordinate, where Φ = 0 deg is rightward from the center location and Φ = 90 deg is upward. The lower three speakers (at Φ = 75 deg, and Φ = [240, 270, 300] deg) were left out to allow room for the listener’s legs (see also Van Wanrooij and Van Opstal 2005; their Fig. 2 shows an illustration of the speaker setup).

Head movements were recorded with the magnetic search-coil induction technique (Robinson 1963). To that end, the listener wore a lightweight (150 g) “helmet” consisting of two perpendicular 4-cm-wide strips that could be adjusted to fit around the listener’s head without interfering with the ears. On top of this helmet, a small coil was attached. From the left side of the helmet a 40-cm-long, thin aluminum rod protruded forward with a dim (0.15 cd/m²) red LED attached to its end. This LED could be positioned in front of the listener’s eyes by bending the rod. Two orthogonal pairs of 2.45 × 2.45-m² coils and one pair of 2.45 × 3.5-m² coils were attached to the room’s edges to generate the left–right (60 kHz), up–down (80 kHz), and front–back (40 kHz) magnetic fields, respectively. This arrangement allows for a precise recording of head orientations in all directions, including the rear hemifield. The head-coil signal was amplified and demodulated (Remmel Labs, Katy, TX), after which it was low-pass filtered at 150 Hz (model 3343; Krohn-Hite, Brockton, MA) before being stored on hard disk at a sampling rate of 500 Hz/channel for off-line analysis.

Auditory stimuli

Acoustic stimuli were digitally generated using Tucker-Davis System II hardware (Tucker-Davis Technologies, Alachua, FL), with a TDT DA1 16-bit D/A converter (50-kHz sampling rate). A TDT PA4 programmable attenuator controlled sound level, after which the stimuli were passed to the TDT HB6 buffer, and finally to one of the speakers in the experimental room.

All acoustic stimuli consisted of Gaussian noise and had 0.5-ms sine-squared on- and offset ramps. The auditory stimuli were either broadband (BB, flat characteristic between 1 and 20 kHz) or high-pass (HP, high-pass filtered at 3 kHz) stimuli with a duration of 150 ms. Sound levels ranged from 30 to 60 dBA (see following text). Absolute free-field sound levels were measured at the position of the listener’s head with a calibrated sound amplifier and microphone (BK2610/ BK4134; Bruel & Kjaer, Norcross, GA). Plugs

Listeners were equipped with a precisely fitting plug in their left ear canal to perturb their binaural cues. The plugs were manufactured by filling the ear canal with rubber casting material (Otiform Otoplastik-KÌ; Dreve, Unna, Germany).

Measurement of audiograms

To determine the attenuation provided by the custom-made plugs, audiograms (10 tone pips, 0.5-octave separation, between 500 Hz and 11.3 kHz) were taken of the listeners’ ears, with and without the plug (Fig. 1A). Although some plugs attenuated more than others, the attenuation was always considerable (>20 dB). For high frequencies the mean attenuation provided by the plugs (>3 kHz: 25–50 dB) was equal to or higher than that for low frequencies (<3 kHz: about 25 dB, Fig. 1B).

Paradigms

CALIBRATION EXPERIMENT. Head-position data for the calibration procedure were obtained by instructing the listener to make an
accurate head movement while redirecting the dim red LED in front of the eyes from the central fixation LED to each of the 57 peripheral LEDs, illuminated as soon as the fixation point extinguished. Each experimental session started with a calibration run.

AUDITORY LOCALIZATION. The listener started a trial by fixating the central LED with the head-fixed LED pointer. After a pseudorandom period of 1.5 to 2.0 s, this fixation LED disappeared and an auditory stimulus was presented 400 ms later. The listener was asked to redirect the head by pointing the dim red LED, which was on continuously throughout the experiment, as accurately and as fast as possible to the perceived location of the sound stimulus. Because the response reaction times typically exceeded 200 ms, all responses were made under open-loop conditions.

Hearing conditions

Sound localization experiments were run under two different hearing conditions. In the free condition, both ears had normal hearing. In the plug condition, the left ear was sealed with a plug. Listeners wore their plug only once, during the experiment, which started immediately after insertion of the plug. Care was taken not to provide plugged listeners with any acoustic input other than the experimental stimuli.

In one experimental session (either with plug or without), two different stimuli of different bandwidths were tested, in two subsequent runs. During one run listeners had to localize BB stimuli of various intensities (40, 50, and 60 dBA). One such BB run consisted of three stimulus intensities × 57 locations = 171 targets.

In the other run HP stimuli were presented, which were typically more attenuated by the plug than the lower frequencies (see Fig. 1B). In the HP run a larger range of stimulus intensities ([30, 35, . . . , 55, 60] dBA) than that in the BB run was used. In this way, one complete experimental HP run consisted of seven stimulus intensities × 57 locations = 399 targets that were randomized across trials. After 200 trials a short break was introduced in which the lights in the experimental room were turned on.

Data analysis

DATA CALIBRATION. The calibration experiment provided a set of 58 LED/speaker locations and raw head position signals. These locations were all transformed into the double-pole azimuth-elevation coordinate system (Knudsen and Konishi 1979). In this system, azimuth $\alpha$ is defined as the angle between the sound source (or response direction), the center of the head, and the midsagittal plane. Elevation $\varepsilon$ is defined as the angle between the sound source, the center of the head, and the horizontal plane. The origin of the $(\alpha, \varepsilon)$ coordinate system corresponds to the straight-ahead speaker location.

Azimuth and elevation can be calculated from the polar coordinates $(R, \Phi)$ by

$$\alpha = \arcsin (\sin R \cos \Phi) \quad \text{and} \quad \varepsilon = \arcsin (\sin R \sin \Phi) \tag{1}$$

These 58 fixation points and raw head-position signals were used to train two three-layer neural networks that served to calibrate the head-motion data, using a back-propagation algorithm based on the gradient descent method of Levenberg–Marquardt (MathWorks, Natick, MA).

The networks corrected for small inhomogeneities in the magnetic fields and could adequately cope with minor cross talk between the channels, resulting from small deviations from orthogonality of the magnetic field coils. The trained networks were subsequently used to map the raw data to calibrated 2D head-positions, yielding azimuth and elevation response components with an absolute accuracy within 4% over the entire response range.

HEAD MOVEMENT DETECTION. Saccadic head movements were detected from the calibrated head-motion signals by setting thresholds to the vectorial head velocity for on- and offset, respectively, using a custom-made program (onset velocity = 20 deg/s; offset velocity = 15 deg/s). Detection markings from the program were visually checked by the experimenter and could be adjusted manually, when deemed necessary. Head movements with reaction times <80 ms, or >1,000 ms, were discarded because responses with extremely short latencies may be regarded as anticipatory and responses with excessive latencies usually arise from the inattentiveness of listeners.

STATISTICS. Responses of each listener were quantified by determining the optimal linear fit for the following stimulus–response relations

$$\alpha_R = a + b \cdot \alpha_T \quad \text{and} \quad \varepsilon_R = c + d \cdot \varepsilon_T \tag{2}$$

for the azimuth and the elevation components, respectively, by minimizing the least-squares error (Press et al. 1992). In Eq. 2, $\alpha_R$ and $\varepsilon_R$ are the azimuth and elevation response components; $\alpha_T$ and $\varepsilon_T$ are the azimuth and elevation coordinates of the target. Fit parameters $a$ and $c$ are the response biases (offsets, in degrees), whereas $b$ and $d$ are the overall response gains (slopes, dimensionless) of the azimuth and elevation response components, respectively. An ideal listener yields a gain of one and an offset of zero deg. Also, Pearson’s linear correlation coefficient and residual errors around the regression line were calculated. Because in the plugged hearing condition the regression results heavily depended on the applied sound level, regressions were performed separately for each sound level.

Localization performance was also quantified by determining the mean absolute error (MAE) of the responses.
for the azimuth and the elevation components, respectively, with \( N \) the total number of trials.

To account for the strong azimuth dependency of the azimuth and elevation response components in the plugged hearing conditions (see results), data analysis was also performed within restricted regions of azimuth space (local regression/local MAE). To that end, responses were collected within 25-deg-wide azimuth bins (pooled across elevation), each shifted in 5-deg steps (thus with 20-deg overlap between adjacent bins), from which we determined a smooth estimate of the local azimuth and elevation regression parameters and the MAE (e.g., Fig. 4; Zwiers et al. 2003).

We also compared the gain and bias between the plug and the control session by computing the relative gain

\[
 b_{rel} = \frac{b_{plug}}{b_{control}} \quad \text{and} \quad d_{rel} = \frac{d_{plug}}{d_{control}}
\]

and the change in bias

\[
 \Delta b = b_{plug} - b_{control} \quad \text{and} \quad \Delta d = d_{plug} - d_{control}
\]

for azimuth (\( b_{rel}, \Delta b \)) and elevation (\( d_{rel}, \Delta d \)) components, respectively.

The bootstrap method was applied to determine confidence limits for the optimal fit parameters in the regression analyses. To that end, 100 data sets were generated by randomly selecting (with replacement) data points from the original data set. Bootstrapping thus yielded a set of 100 different fit parameters. The SDs in these parameters were taken as an estimate for the confidence levels of the parameter values obtained in the original data set (Press et al. 1992).

RESULTS

Acute effects of a unilateral plug

The attenuation of sounds with a plug in the left ear (Fig. 1), although leaving the right ear unperturbed, quite profoundly changed a listener’s sound localization ability. In Fig. 2 the acute effect of the plug is exemplified by comparing the localization performance of a typical listener (RK) to HP stimuli during normal hearing (Fig. 2, A and B; pooled across intensities), with the acute plug condition (Fig. 2, C and D).

With normal binaural hearing, this listener was quite precise, accurately localizing both sound azimuth (Fig. 2A) and elevation (Fig. 2B), regardless of the stimulus intensity. Despite the large range in stimulus levels, including some as low as 30 dBA, regression lines were nearly optimal (as indicated by an overall regression gain near 1.0 and a bias near 0.0 deg), with modest scatter of the data around the regression line (residual errors: ±10.8 and ±11.4 deg, for azimuth and elevation, respectively).

The plugged condition, however, introduced a large shift in the listener’s azimuth localization responses toward the unplugged ear (Fig. 2C; bias \( a = 45.2 \) deg). In addition, the response gain (\( b = 0.18 \)) and correlation between target and response location degraded substantially. Despite this clear detriment in sound localization (residual error: ±16.3 deg), however, the stimulus–response correlation was still positive and significant (\( r = 0.3, P < 0.05 \)). The listener’s elevation responses, on the other hand, were affected much less because there was only a slight decrease in both the gain (\( d = 0.76 \)) and the correlation (\( r = 0.85, P < 0.01 \); residual error: ±14.5 deg). As such, these data are in line with previous reports on sound localization to fixed-intensity stimuli under acutely plugged listening conditions (Oldfield and Parker 1986; Slattery and Middlebrooks 1994). The observed effects also underline the strong dominance of the binaural cues for normal-hearing sound localization in the horizontal plane.

Because the plug appeared to have a dominant effect on the azimuth response components, we will first focus on the results of azimuth localization. The effects on elevation localization will be dealt with later in this section.

Influence of sound intensity and source azimuth on azimuth localization

Although pooling the data across intensities clearly demonstrates the deterioration of azimuth localization, it ignores a potential systematic effect of sound level. As an illustration, Fig. 3 shows the localization responses of listener RK for three different intensities (30, 45, and 60 dBA, HP stimuli). For normal binaural hearing, no noticeable effect of intensity appeared (Fig. 3, A–C). In the acute plug condition, however, changes in sound level dramatically influenced the listener’s responses (Fig. 3, D–F). For all intensities, sound azimuth localization was clearly perturbed, but the overall localization bias toward the unaffected ear increased strongly with intensity. Higher intensities induced a larger shift in the acute plug condition (e.g., at 60 dBA: \( a = 56 \) deg; Fig. 3D) than did lower intensities (e.g., at 30 dBA: \( a = 36 \) deg; Fig. 3F). Note that the azimuth response gain was highest for the low-intensity stimulus (\( b = 0.14 \) at 60 dBA, but \( b = 0.32 \) for the 30 dBA sounds). This is remarkable because at low intensities the listener was effectively monaural, whereas for the higher intensities the plugged ear still received acoustic input, albeit strongly attenuated (see Fig. 1). Thus although the strongest perturbations of the ILDs were obtained for low-intensity...
sounds, localization performance was better than that for the higher sound levels.

Figure 4A shows the systematic influence of sound level on the azimuth localization bias of HP stimuli. Note that all listeners demonstrated a similar effect: the bias is at a minimum for the lowest sound level, and increased monotonically with intensity. Although the actual bias values varied considerably from listener to listener, the variation of the bias with sound level was quite similar for each of the listeners: over a 30 dBA intensity range, the response bias shifted by about 20 deg.

Also the spatial gain (quantified by the azimuth regression slope) changed in a systematic way with sound level. Figure 4B shows for each individual listener the gain as a function of sound level, after subtracting each listener’s mean azimuth gain (which is shown in the inset in Fig. 4B). The data show that azimuth gain tended to be highest for the lowest stimulus levels, with a dip at intermediate levels, for all listeners, except

FIG. 3. Influence of sound level. Sound azimuth localization responses from listener RK toward HP stimuli of 60 dBA (A, D), 45 dBA (B, E), and 30 dBA (C, F), when both ears were free (A, B, C), and when his left ear was plugged (D, E, F). Thick solid lines denote linear regression lines. Note decreasing bias and increasing gain for low intensities in the plugged condition, which is not seen in the control data.

FIG. 4. Linear regression analysis on plugged azimuth localization. Regression parameters bias (A) and gain (B), minus the overall mean gain (see inset) to remove individual differences, for the HP stimuli in the plugged condition are plotted against stimulus level for all listeners. Note the increase in performance as evidenced by a lower bias (and a tendency for a higher gain) for the lowest sound levels. C: gain and bias (for all stimulus levels and listeners) are plotted against each other for broadband (BB, dark markers) and high-pass (HP, open markers) stimuli and plugged (circles) and control (diamonds) conditions. Note high correlation between the bias and gain for both the BB and HP stimuli in the plugged condition. D: mean plugged azimuth bias (± SD, pooled across intensities) vs. subjective plug attenuation (mean attenuation pooled across all frequencies) for BB stimuli (dark circles) and for HP stimuli (open circles; see Fig. 1A) for all listeners. Note absence of correlation.
listener MW (whose mean azimuth gain was not significantly different from 0; see inset).

Interestingly, the azimuth gain and bias in the plugged condition were highly negatively correlated, as shown in Fig. 4C (BB and HP stimuli, and listeners pooled; for HP stimuli: \( r^2 = 0.58, P \ll 0.01 \); for BB stimuli: \( r_{BB} = 0.31, P < 0.05 \)). This indicates that a smaller shift in the azimuth responses covaried with a larger spatial gain.

To verify whether the observed differences in the azimuth regression parameters across listeners could be attributed to intersubject differences in the perceptual attenuation of the plug we compared the regression results with each listener’s audiogram (cf. Fig. 1A). The result is shown in Fig. 4D, in which for each listener the azimuth bias (averaged across intensity) is plotted against the subjective attenuation of the plug (averaged across the HP and LP frequency bands; see Fig. 1A). Interestingly, we obtained no correlation for either the HP stimuli (containing only potential ILDs) or the BB sounds (both ITDs and ILDs may be present). A similar result was obtained for the azimuth gains (not shown). Thus the plug’s attenuation, which corresponds to a fixed, but frequency-dependent perturbation of the ILDs, does not predict the value of the shift in perceived sound-source azimuth.

We next studied whether listeners with a monaural plug could have relied on spectral-shape cues to localize azimuth. Note that the plug caused an immediate effect on the quality of the spectral cues across the azimuth domain: on the side ipsilateral to the plug the spectral cues were strongly attenuated (and often abolished altogether), whereas on the side of the normal-hearing ear the spectral cues remained unperturbed.

Thus if listeners would rely only on spectral cues to extract the sound-source azimuth, response accuracy should depend on the sound’s azimuth location. To exemplify our analysis, Fig. 5 quantifies the dependency of azimuth performance on the acoustic parameters for all HP stimuli for listener MW, for both the control (top) and plugged (bottom) hearing conditions. In each \( 7 \times 10 \) matrix, an entry corresponds to a particular combination of sound level (abscissa) and target azimuth (ordinate), whereas the average response azimuth (Fig. 5, A, B, and D) and the mean absolute error (MAE, Eq. 3; Fig. 5, C and E) values are color encoded. Figure 5A illustrates the appearance of the localization matrix for an ideal listener, whose responses do not depend on sound level and whose response azimuth corresponds exactly to target azimuth (\( \alpha_R = \alpha_T \)). Note that in the control hearing condition (Fig. 5B) the responses of listener MW corresponded quite well to those of the ideal listener. As a result, the azimuth MAE was small (Fig. 5C, about 5–10 deg) and did not vary systematically across the azimuth-intensity parameter space.

In the plugged condition, however, the listener’s responses showed a marked shift toward the unplugged ear (the dark-red color indicates far-rightward responses). Note, however, that the shift is slightly less (light red) for the low-intensity stimuli on the side of the unplugged ear (Fig. 5D). This systematic difference becomes more evident when the MAE in azimuth is plotted as a function of stimulus intensity and target azimuth (Fig. 5E). The MAE is minimal for low-intensity stimuli on the unplugged hearing side (bottom right, blue voxels), whereas it increased systematically for targets toward the plugged side and for the higher sound levels (left and top voxels).

**Influence of intensity and azimuth on elevation localization**

Despite some quantitative intersubject variability, all listeners demonstrated a similar effect for the HP stimuli (Fig. 6): the MAE reached a minimum for low-intensity sounds presented on the side of the normal-hearing ear (blue, bottom right voxels) and increased with both increasing intensity and increasing distance from the unplugged ear (red, top left voxels). The results for the BB stimuli were indistinguishable from the HP data (not shown). We also performed a local regression analysis on these data (see METHODS), which revealed a similar trend: the local azimuth bias was lower and the local azimuth gain was highest for the low-intensity stimuli contralateral to the plug (data not shown).

In summary, in the acute monaural plug condition listeners localized low-intensity sounds better than high-intensity sounds on the unplugged hearing side (evidenced by a lower MAE for azimuth), despite a total absence of the binaural difference cues for the low-intensity stimuli. On the plugged side, azimuth localization was always poor regardless of the stimulus intensity. These data therefore strongly suggest that plugged listeners relied on the monaural spectral cues from their unplugged ear, especially for the low-intensity stimuli for which the ILDs and ITDs were either poor or nonexistent. Furthermore, these spectral cues appeared to be binaurally weighted because their contribution depended in a systematic and gradual way on azimuth.

**FIG. 5. Influence of sound intensity and target azimuth on free and plugged azimuth localization.** Color encodes averaged, interpolated response location (deg) for each of the stimulus azimuth-intensity bins for (A) an ideal listener (\( \alpha_R = \alpha_T \)), (B) listener MW in the unplugged condition, and (D) listener MW in the plugged condition. Red bins refer to responses made into the right hemifield for the unplugged (A, B) and plugged (D) conditions, and blue bins are responses into the left hemifield. C and E: color encodes azimuth mean absolute error (MAE, deg) for the unplugged (C) and plugged condition (E), where red refers to a large error and blue to a small error. Note that the low errors in the plugged condition on the free, unplugged side at low intensities are comparable to the control MAs.
by a larger amount of scatter and a lower gain, even for the pooled responses. A detrimental effect on the pooled data was to be expected, given that no spectral cues survived at the plugged side for nearly all sounds. In what follows, we will quantify in greater detail the effect of sound level and azimuth on the subject’s elevation performance.

As a first step in our analysis, Fig. 7 plots the linear regression results on elevation for listener RK for three sound levels for control hearing (top row) and for the plugged hearing condition (bottom). The data are pooled for azimuth locations across the frontal hemifield. The insertion of a plug immediately reduced the overall gain for all three sound levels. The elevation response bias appeared to change slightly, but systematically, with sound level when compared with the control experiment (change in bias $\Delta c$, Eq. 5, equaled $-8.2$, $-4.0$, and $+1.5$ for 60, 45, and 30 dBA stimuli, respectively).

Consistent changes in bias and gain are better observed for the relative elevation gain ($d_{rel}$, Eq. 4) and the change in elevation bias ($\Delta c$, Eq. 5) for all sound levels and all participants (Fig. 8). To account for a potential systematic azimuth dependency, we first performed the regression analysis for two separate regions: for stimuli on the side of the plug ($\alpha_T < -20$ deg, Fig. 8, A and C), versus stimuli on the normal-hearing side ($\alpha_T > 20$ deg, Fig. 8, B and D). The mean relative elevation gain (averaged across sound levels) on the plugged side varied somewhat from listener to listener, although it was clearly worse than that in the control condition for all listeners ($d_{rel} < 1$, Fig. 8A, inset). All listeners, except for JO (open circles), exhibited a similar effect on the elevation gain as a function of sound level (Fig. 8A): higher sound levels elicited a higher response gain. Note that listener JO was not able to localize elevation at all on the plugged side (see Fig. 8A, inset, white bar).

Also on the normal-hearing side, the mean relative elevation gain decreased for all listeners (Fig. 8B, inset), although not as much as on the plugged side. Surprisingly, however, the effect of sound level on elevation gain was now reversed (Fig. 8B): on the hearing side, low stimulus intensities elicited higher response gains. Apparently, both sound level and azimuth determined gain in the elevation direction.

The overall change in elevation bias (Eq. 5, averaged across sound levels) was characterized by a substantial decrease on the plugged side (Fig. 8C, inset), whereas for all listeners, except MW, it was barely affected on the normal-hearing side (Fig. 8D, inset). Furthermore, the change in bias was influenced by sound level in a similar way on both the plugged and the normal-hearing side: lower stimulus intensities elicited a higher (i.e., more upward) change in elevation bias than louder sounds.

To quantify the influence of sound-source azimuth on the localization of elevation with a monaural plug in more detail, we adopted a local regression analysis similar to that for the azimuth responses. Figure 9, A–D shows, in a format similar to...
that in Figs. 5 and 6, the local elevation gain (left) and bias (right) as a function of sound azimuth and level for listener MW. The top row shows the data for the control hearing condition. Note that local elevation gain is uniformly high (dark red) across the frontal hemifield (Fig. 9A), whereas the local bias is around +15 deg, with a tendency to slightly increase for higher HP sound levels (Fig. 9B). For the plugged hearing condition, the response bias decreased to about 5 deg (Fig. 9D). The local elevation gain (Fig. 9C) depended systematically on both the azimuth and intensity parameters: the maximal local elevation gain was obtained for sounds on the normal-hearing side, at the lowest intensities, whereas it decreased for higher sound levels and for stimuli on the plugged side. The bottom row shows the same data, now expressed as the local relative elevation gain (Eq. 4, Fig. 9E) and the local change in bias (Eq. 5; Fig. 9F). Whereas the relative elevation gain showed a trend similar to that of the absolute elevation gain, the overall change in bias was negative and did not systematically depend on the stimulus parameters for this listener.

Figure 10 shows the results of this analysis for the other four listeners. Despite the variability in absolute values, listeners yielded quite consistent results; elevation localization had its highest spatial gain at locations far into the normal-hearing hemifield and for the lowest intensity stimuli (bottom right voxels). Higher stimulus intensities as well as locations on the plugged side yielded lower-elevation gains. The plug also influenced the elevation bias (Fig. 10, E–H), in a pattern that was qualitatively similar across listeners (with the exception of MW; Fig. 9F). Listeners responded more upward for the low-intensity stimuli on their normal-hearing side, while systematically pointing to more downward locations for high-intensity stimuli and for stimuli on the plugged side. Results for the BB stimuli were very similar (data not shown).

**Weighting of cues**

Our results on the localization of sound-source azimuth (Figs. 2–6) give rise to the idea that the computation of azimuth is not solely determined by binaural difference cues, but also by the spectral-shape cues of the normal-hearing ear. In turn, the elevation data (Figs. 7–10) suggest that apart from spectral-shape cues, azimuth is also an important factor that contributes to the elevation percept. Here, one should distinguish the actual source azimuth from the perceived azimuth location because under plugged hearing these locations are very different (e.g., Fig. 2C). Previous studies with molds have shown that the computation of elevation involves binaural interactions (see INTRODUCTION), but it remains unclear whether this concerned the actual azimuth or the perceived azimuth location because a mold does not perturb the azimuth percept.

Note that the spectral cues on the hearing side were unperturbed under monaural plugged hearing. In other words, if the actual source azimuth would be the determining factor, elevation would be perceived at its veridical location for all azimuths on the hearing side. As Figs. 8–10 show, this was not the case, and we therefore wondered whether in fact the perceived azimuth location might have determined elevation. To assess this point, we selected all responses for HP targets presented on the far-lateral hearing side (source azimuths between 40 and 60 deg) because for these locations the spectral cues would be optimal for all hearing conditions. Separately for each listener, and each stimulus level, we then took the azimuth MAE of these responses as a measure for the perceived azimuth location at these nearly fixed azimuth positions.

![Fig. 8. Influence of sound level on elevation localization. Relative elevation gain (A, B; see Eq. 4) and change in bias (deg, C, D; see Eq. 5) for the HP stimuli in the plugged condition vs. stimulus level for all listeners, for the plugged (A, C) and the normal-hearing side (B, D). Insets: mean regression parameter (averaged over intensities) has been subtracted to remove intersubject variability. Note the increase in gain for increases in sound level on the plugged side, whereas on the unplugged side the reverse is true. Elevation bias tends to be higher for the low-intensity stimuli on both sides.](image-url)
if the MAE is low, the perceived azimuth approaches the veridical azimuth, whereas for large MAEs the perceived azimuth is very different. If the actual azimuth would determine the perceived elevation, elevation gain should not depend on the MAE. Instead, Fig. 11A shows that for all listeners the elevation gain varied with the MAE: the larger the MAE, the lower the elevation gain, which suggests that perceived azimuth is the relevant parameter.

Note, that as the change in stimulus intensity (and thus, perhaps, the integrity of the spectral cues) is implicit in Fig. 11A, one might suspect that the low elevation gains could have resulted from a poor signal-to-noise ratio (SNR) of the spectral shape cues at low intensities. However, Fig. 11B shows that in fact the opposite is true: the lower stimulus intensities produced the highest elevation gains for all listeners (see also Figs. 9 and 10). Therefore the spatial resolution in elevation depended on the MAE and thus was determined by the perceived azimuth location, rather than by the actual azimuth location.

In Fig. 11C we tested whether the perceived azimuth angle was in turn determined by the spectral-shape cues. To that end, we plotted azimuth performance (i.e., the local MAE for each of the ten different azimuth regions) as a function of the relative elevation gain, but now for the low-intensity stimuli only (30 dBA). Note that for all listeners, the MAEs for these stimuli were consistently correlated with the elevation gain: the higher the relative elevation gain (i.e., the better the spectral cues were resolved by the auditory system), the lower the azimuth error. These data therefore strongly support the idea that for plugged hearing, the spectral cues do indeed contribute to azimuth localization.

**DISCUSSION**

We studied the effect of an acute perturbation of the binaural localization cues on 2D sound localization behavior. All listeners immediately mislocalized sounds predominantly in the azimuth direction (e.g., Fig. 2C). Localization of target elevation was also affected, albeit in a more subtle way (Fig. 2D). We found that listeners were best at localizing low-intensity sounds on the hearing side for both the azimuth and elevation directions, despite a better SNR for the spectral cues at higher intensities. Note that at the higher intensities binaural inputs will start to overcome the plug’s attenuation, although they are highly perturbed (Fig. 1). The plug thus creates ILDs that far exceed the normal physiological range provided by the head-shadow. Yet, these erroneous cues were incorporated in forming the azimuth percept, giving rise to larger localization errors than for low-intensity stimuli for which the binaural cues were absent (Figs. 3–6). The presence of perturbed binaural difference cues also affected elevation localization at higher intensities (Figs. 8–10).

We conclude from these findings that under plugged hearing conditions, listeners make use of spectral-shape cues from the normal-hearing ear to localize sound-source azimuth (Fig. 11C). We also conclude that target elevation is not only determined by the spectral-shape cues of the ear ipsilateral to the sound, but also results from a binaural weighting process. However, this weighting is strongly influenced by the per-
ceived azimuth location of the sound, rather than by the actual azimuth (Fig. 11, A and B).

Spectral cues for azimuth

Several arguments support our conclusion that plugged azimuth localization depends on spectral cues. First, for low-intensity sounds binaural cues were virtually absent (Fig. 1A), suggesting that listeners had no other possibility but to use the spectral cues. Furthermore, the extent of the plug’s attenuation (20–55 dB) suggests that listeners did not exclusively use the perturbed ILDs to generate their responses. The ILDs of a human head reach a maximum of about 20 dB for the extreme lateral positions (e.g., Blauert 1997; Van Wanrooij and Van Opstal 2004). Thus an ILD of 20 dB corresponds to a location beyond 60 deg from the midline. However, plugged listeners responded to far less extreme locations, even for the HP stimuli (around 30 deg; e.g., Figs. 2C and 4A), suggesting that they also relied on other cues. Indeed, the insertion of a mold in the normal-hearing ear, when a plug plus an additional muff attenuated sounds by ≥40 dB, completely abolished sound localization performance in azimuth and elevation (Hofman 2003), showing the importance of spectral cues for monaural hearing.

Our findings show that for plugged hearing the spectral cues dominated for the lowest sound levels (e.g., Figs. 5 and 6). Indeed, azimuth localization performance gradually decreased toward the plugged side, in concordance with a decrease in elevation performance (Fig. 11C). Thus when binaural difference cues were severely perturbed, or became ambiguous, the spectral localization cues became increasingly important.

We believe that the weighting of the different cues changes instantaneously because in our experiments the stimuli were randomized. Thus although the plug introduces a large and frequency-dependent conflict between the binaural localization cues at higher sound levels, they were still favored over monaural spectral-shape cues. Under normal binaural hearing the difference cues dominate entirely because azimuth localization is robust against large variations in sound level and SNRs (e.g., Fig. 3, A–C; see also Good and Gilkey 1996; Hofman and Van Opstal 1998; MacPherson and Middlebrooks 2000; Van Wanrooij and Van Opstal 2005; Vliegen and Van Opstal 2004; Zwiers et al. 2001) and does not depend on the integrity of the spectral cues (Hofman and Van Opstal 2003; Morimoto 2001; Oldfield and Parker 1984; Van Wanrooij and Van Opstal 2005; Wightman and Kistler 1997).

Azimuth cues for elevation

Our results show that a monaural plug had a systematic effect on the localization of sound-source elevation, not only on the side of the plug, but also on the normal-hearing side.

FIG. 11. Influence of spectral cue saliency and azimuth percept on plugged localization. A: relative elevation gain vs. change in azimuth MAE (both parameters corrected for individual differences by subtracting their means) for stimuli between 40 and 60 deg azimuth, for all listeners and sound levels. B: relative elevation gain vs. stimulus intensity (dBA). Note that the highest levels (indicative for high signal-to-noise ratios) elicit the lowest gains. C: change in azimuth MAE (deg) vs. relative elevation gain (means subtracted) for all listeners and different azimuth sectors (the lower the elevation gain, the higher the MAE, and vice versa). Sound level: 30 dBA.
Although the effects were smaller than the dramatic localization deficits for azimuth, they were systematic and consistent across listeners. Butler et al. (1990) showed that under plugged hearing, advance knowledge of the sound’s azimuth may enhance elevation performance. In line with this, findings obtained with a monaural mold (Hofman and Van Opstal 2003; Humanski and Butler 1988; Morimoto 2001; Van Wanrooij and Van Opstal 2005) indicated that the localization of elevation involves binaural interactions, the strength of which varies gradually with azimuth. The present study extends these results by showing that the binaural weighting depends on the perceived azimuth location, which in the case of plugged hearing may be quite erroneous (Fig. 11, A and B). This is a remarkable finding given that the spectral shape cues, the sound level, and the SNR of the peaks and notches in the sound spectra at the normal-hearing side were all unaffected by the plug. When localizing elevation, acutely plugged listeners showed a marked decrease in performance on the plugged side for the lower sound intensities (Fig. 8A). In contrast, listeners performed better for low-intensity sounds at their normal-hearing side (Fig. 8B). We believe that two different factors underlie these seemingly different behaviors. First, the spectral cues of the plugged ear have a low SNR, which affects low-intensity sounds more than high-level sounds (Fig. 8A). Second, the elevation percept comes about by fusing the spectral cues from each ear with a weight determined by the perceived azimuth (see above), which on the normal-hearing side is perturbed more at higher intensities than at low intensities (e.g., Figs. 5, 6, and 8B).

Monaural listening

Our results extend recent findings obtained from monaurally deaf listeners (Van Wanrooij and Van Opstal 2004). That study showed that all monaural listeners relied heavily on the head-shadow effect (i.e., absolute sound level at the hearing ear) to localize sounds in the horizontal plane, whereas half of the listeners also incorporated the spectral-shape cues of their intact ear to estimate azimuth. Interestingly, only listeners who had learned to use spectral cues to localize azimuth could also localize elevation on the side of their intact ear. Monaural listeners who did not make use of the spectral cues for azimuth could not localize elevation either, which hinted at the possibility that the ability to localize elevation strongly depended on the performance in azimuth.

Plugging the ear of an otherwise normal-hearing binaural listener is very different from real monaural hearing of the unilaterally deaf for a number of reasons: First, although the plug strongly attenuates sounds, the acoustic input will not be entirely abolished as in the monaurally deaf. Thus for sounds at a sufficiently high intensity, all localization cues are still present, albeit heavily perturbed. Second, the acoustic effect of the plug typically depends on frequency (Fig. 1), yielding ambiguous localization cues when compared with normal hearing. For example, not only will the ILDs differ for different frequency bands (and thus point to different azimuth locations), but the ITDs and ILDs are also affected in a different way. Thus the outputs of the two binaural localization streams will often not agree on sound-source azimuth either. For sufficiently low sound levels, however, plugged hearing approaches real monaural hearing. The plug’s intensity- and frequency dependencies therefore pose an interesting and non-trivial challenge to the sound-localization system. Third, the monaurally deaf have had long-term exposure to their hearing condition, allowing ample time for adaptive processes to reshape their localization behavior. This contrasts with the immediate and complex effect of a plug on localization of the normal-hearing listener. Conversely, binaural listeners have had ample experience in using the binaural difference cues and the detailed complex spectral shape cues from either ear.

Taken together, it is reasonable to expect that the central organization of the sound localization systems of the monaurally deaf and of binaural listeners may be quite different (see also Bilecen et al. 2000; Ponton et al. 2001; Scheffler et al. 1998).

Integration of acoustic cues

We propose that azimuth and elevation are both computed on the basis of evidence from all available acoustic cues, but that their relative weights depend on the acoustic conditions. Figure 12 depicts this conceptual model, which extends the classical idea that the stimulus coordinates are determined by independent, noninteracting pathways. The localization of azimuth is based on a weighting of binaural difference cues as well as of spectral shape cues. When the binaural cues become unreliable or ambiguous (e.g., for weak sounds at far-lateral locations or after plugging one ear) their weights are reduced, and the contribution of the spectral cues from the contralateral ear increases. In turn, the computation of sound-source elevation involves a weighting of the spectral shape cues from both ears. The strength of the binaural weighting is modulated by the perceived azimuth location.

Neural mechanisms

At the initial stages in the auditory system, the localization cues are processed by independent brain stem pathways, both...
in birds (e.g., the barn owl; for review see Takahashi 1989) and in mammals. In mammals, the medial superior olive (MSO) constitutes the ITD pathway, whereas the ILD pathway is processed in the lateral superior olive (LSO; for reviews see Irvine 1986; Yin 2002). The elevation pathway has yet to be identified, but recent evidence suggests that the first stages of spectral-shape analysis may already occur at the level of the dorsal cochlear nucleus (Reiss and Young 2005; Young and Davis 2002).

Our finding that the percept of sound-source azimuth is determined by spectral cues, when the binaural cues become unreliable, could arise from mechanisms that rely entirely on the acoustic properties of the signal. Because the different brain stem pathways all converge in the midbrain, the inferior colliculus (IC) could play a role in such preattentive computations. Although an explicit map of auditory space has not been demonstrated in the mammalian IC, unlike in the barn owl (Knudsen and Konishi 1978), recent evidence suggests that auditory space may be represented in the IC by space-specific modulations within a large population of neurons. For example, the majority of monkey IC cells are sharply tuned to sound frequency, although their firing rates are also modulated by sound level, by the location of the sound in azimuth and elevation, and by nonacoustic signals such as eye position (Groh et al. 2001, 2003; Zwiers et al. 2004).

On the other hand, our result that the spectral-cue weighting that determines elevation is influenced by the listener’s perceived azimuth, rather than by the purely acoustic effects of stimulus azimuth on the spectral cues (Fig. 11A), might suggest that higher, perhaps cortical, mechanisms are involved. Recent recordings indicated that populations of cells in the primate auditory cortex may encode sound locations in a way similar to that of the IC (Recanzone 2000; Werner-Reiss et al. 2003). Whether the auditory cortex may be involved in the perceptual integration of acoustic cues has to be established by future recording studies.

Acknowledgments

We thank G. Van Lingen, H. Kleijnen, G. Windau, H. Versnel and T. Van Dreumel for technical assistance.

Grants

This work was supported by Radboud University Nijmegen grants to A. J. Van Opstal and M. M. Van Wanrooij and by Human Frontiers Science Program Grant RG 0174–1998/B to M. M. Van Wanrooij.

References


J Neurophysiol • VOL. 97 • JANUARY 2007 • www.jn.org