

NEUROSYSTEMS

Absence of compensation for vestibular-evoked passive head rotations in human sound localization

Denise C. P. B. M. Van Barneveld, Floor Binkhorst and A. John Van Opstal

Radboud University Nijmegen, Donders Institute for Brain, Cognition and Behaviour, Department of Biophysics, Geert Grooteplein 21, 6525 EZ Nijmegen, The Netherlands

Keywords: eye movements, human, models, multisensory integration, spatial updating, vestibular system

Abstract

A world-fixed sound presented to a moving head produces changing sound-localization cues, from which the audiomotor system could infer sound movement relative to the head. When appropriately combined with self-motion signals, sound localization remains spatially accurate. Indeed, free-field orienting responses fully incorporate intervening eye-head movements under open-loop localization conditions. Here we investigate the default strategy of the audiomotor system when localizing sounds in the absence of efferent and proprioceptive head-movement signals. Head- and body-restrained listeners made saccades in total darkness toward brief (3, 10 or 100 ms) broadband noise bursts, while being rotated sinusoidally ($f = 1/9$ Hz, $V_{\text{peak}} = 112$ deg/s) around the vertical body axis. As the loudspeakers were attached to the chair, the 100 ms sounds might be perceived as rotating along with the chair, and localized in head-centred coordinates. During 3 and 10 ms stimuli, however, the amount of chair rotation remained well below the minimum audible movement angle. These brief sounds would therefore be perceived as stationary in space and, as in open-loop gaze orienting, expected to be localized in world-centred coordinates. Analysis of the saccades shows, however, that all stimuli were accurately localized on the basis of imposed acoustic cues, but remained in head-centred coordinates. These results suggest that, in the absence of motor planning, the audio motor system keeps sounds in head-centred coordinates when unsure about sound motion relative to the head. To that end, it ignores vestibular canal signals of passive-induced head rotation, but incorporates intervening eye displacements from vestibular nystagmus during the saccade-reaction time.

Introduction

Acoustic cues, generated by the interaction of sound waves with the head and pinnae, specify sound locations with respect to the head. Interaural timing differences (ITDs) and interaural level differences (ILDs) refer to horizontal plane locations (azimuth), whereas spectral pinna cues define vertical directions (elevation) (Wightman & Kistler, 1989; Middlebrooks, 1992; Blauert, 1997). Thus, programming a goal-directed eye movement toward a sound should incorporate the initial eye-in-head orientation ('spatial updating') (Poppel, 1973; Jay & Sparks, 1984, 1987; Van Grootel & van Opstal, 2010). Indeed, changes in eye and/or head posture influence spatial hearing (Lewald, 1997; Goossens & Van Opstal, 1999; Kopinska & Harris, 2003; Koenigs *et al.*, 2007; Van Barneveld & Van Opstal, 2010), and open-loop gaze-orienting studies have shown that the auditory system accurately incorporates intervening eye-head movements made after, or during, sound presentation (Goossens & Van Opstal, 1999; Vliegen *et al.*, 2004; Van Grootel & van Opstal, 2010).

How does the auditory system dissociate target motion from self-motion? Under natural conditions, neck-muscle proprioception, corollary discharges of planned movements and vestibular signals could

all contribute to these transformations (e.g. Angelaki & Cullen, 2008; Armstrong *et al.*, 2008; Crapse & Sommer, 2008; for reviews). Such signals could, in principle, determine to what extent changes in acoustic cues correlate with (estimated) motion of the head. Because target and head movements would typically be uncorrelated, the default assumption could be that sounds originate from a world-centred reference frame. Indeed, free-field eye localization of tones (Goossens & Van Opstal, 1999) and eye-head orienting studies to brief noise bursts have supported this idea (Goossens & Van Opstal, 1999; Vliegen *et al.*, 2004).

The present study concerns saccade responses towards brief noise bursts (3, 10 and 100 ms) under passive whole-body rotation to investigate the vestibular contribution to audiospatial updating when corollary discharges and proprioceptive signals of head movements are absent. We analysed saccades (ΔE) by applying models that differ in the amount of compensation for intervening passive head displacement (ΔH_S) and eye-in-head position (E_H) at saccade onset (Fig. 1A).

$$\Delta E = T_H + b \cdot \Delta H_S + c \cdot E_H \quad (1)$$

with T_H the head-centred location at sound onset. Localization in world coordinates requires full compensation for intervening movements ($b = c = -1$, Model I). When targets remain in head-centred

Correspondence: A. John Van Opstal, as above.
E-mail: j.vanopstal@donders.ru.nl

Received 27 June 2011, revised 20 July 2011, accepted 21 July 2011

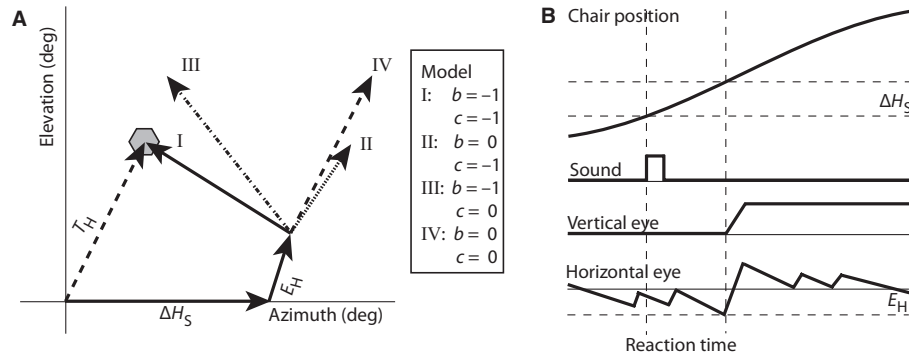


FIG. 1. (A) Four models for a saccadic eye movement toward an auditory target during passive whole-body rotation. At time = 0 a brief auditory target (T_H) is presented while the head is at $[\alpha, \varepsilon] = [0, 0]$. During the reaction time the head (and body) is passively rotated (ΔH_S) to the right. At the response onset the eyes are not centred in the orbit (E_H) due to the ongoing vestibular ocular reflex and the absence of visual landmarks. Model I predicts a response in world-centred coordinates and fully incorporates intervening head- and eye-displacement signals. Model II predicts a head-centred response, as it incorporates only the change in eye position. Model III only accounts for the head displacement, whereas Model IV keeps the target in the initial head-centred reference frame. Note that Model II is the ideal model, as the auditory targets were rotating together with the listeners. The ideal values of b and c of Eqn 1 are presented in Table 1. (B) Temporal order of chair position, auditory target and response.

coordinates only the eye-in-head position is accounted for ($b = 0$, $c = -1$, Model II), whereas Model IV lacks any spatial updating ($b = c = 0$).

Although sounds rotated along with the listener, we show that the amount of head displacement during the briefest sounds ($< 1^\circ$) remained well below the minimal audible movement angle (MAMA) (about 5°); the auditory system could thus not detect any changes in acoustic localization cues. These sounds are therefore perceived as stationary in space and, as in Vliegen *et al.* (2004), expected to be localized in world-centred coordinates. Longer-duration (100 ms) stimuli could induce measurable changes in ITDs/ILDs that would anticorrelate with self-generated head movements, and the auditory system might thus correctly infer that these sounds moved along with the head. Interestingly, our results demonstrate that listeners localized all sounds in head-centred coordinates (Model II), suggesting that head movements are not incorporated in sound localization when only vestibular cues are present, and/or the system is unsure about stimulus motion.

Materials and methods

Listeners

Nine listeners, all with normal hearing, participated in the experiments. Three of the listeners (the authors) were familiar with the purpose of the experiment. All listeners had normal or corrected-to-normal vision, except for JO, who is amblyopic in his right, recorded eye. Experiments were conducted after obtaining full understanding and written consent from the listener. The experimental procedures were approved by the Local Ethics Committee of the Radboud University Nijmegen, and adhered to The Code of Ethics of the World Medical Association (Declaration of Helsinki), as printed in the British Medical Journal of July 18, 1964.

Apparatus

Vestibular setup

Experiments were conducted in a completely dark room ($4.05 \times 5.15 \times 3.30$ m). The listener sat in a computer-controlled vestibular stimulator, with the head firmly stabilized in an upright position with a padded adjustable helmet. Chair position was

measured using a digital position encoder with an angular resolution of 0.04° (Van Beuzekom & Van Gisbergen, 2002).

Auditory stimuli emanated from two loudspeakers (SC5.9; Visaton, GmbH, Haan, Germany) mounted inside the vestibular chair at 10° above ear level, both left and right at an azimuth angle of 37° . The distance from the ears to the plane of the loudspeakers was 40 cm. Sounds were played at 70 dBA Sound Pressure Level (A-weighted; calibrated with a Brüel & Kjær BK2610 sound amplifier and BK4144 microphone at the position of the subject's head) and were well discernible from the low-frequency sinusoidal background noise produced by the motor (at the zenith) of the vestibular chair.

We measured two-dimensional eye movements with the scleral search-coil technique (Collewyn *et al.*, 1975), using oscillating magnetic fields generated by two sets of orthogonal coils (0.77×0.77 m) inside the vestibular stimulator. The horizontal and vertical eye-position signals were amplified, demodulated by tuned lock-in amplifiers (Princeton Applied Research, NJ, USA, model PAR 128A), and subsequently sampled at 500 Hz per channel for storage on the computer's hard disk.

Sound stimuli

Coordinate system

We express the coordinates of auditory target locations as well as the eye-in-head position in a double-polar azimuth-elevation coordinate system, in which the origin coincides with the centre of the head (Knudsen & Konishi, 1979; Hofman & Van Opstal, 1998). In this system the azimuth angle, α , is defined as the angle within a horizontal plane with the vertical midsagittal plane. The elevation angle, ε , is defined as the direction within a vertical plane with the horizontal plane through the listener's ears. The straight-ahead position is defined by $[\alpha, \varepsilon] = [0, 0]^\circ$.

Stationary free-field sounds

Six listeners localized stationary free-field sounds, which were digitally generated with MATLAB software and consisted of broadband (0.3–12 kHz) Gaussian white noise of 3, 10 or 100 ms. The 3 and 10 ms sounds had onset and offset ramps of 0.5 ms; the 100 ms sounds had 1 ms ramps. As all auditory stimuli were produced by two loudspeakers at $[\alpha, \varepsilon] = [\pm 37, 10]$ with respect to the listener's head,

we simulated free-field sound locations by filtering the stereo broadband noise bursts with the listener's own ITD and frequency-dependent ILD information (listeners DB, JO and FB), or with the ITD and ILDs of listener JO (listeners AJ, PB and JT). In this way, simulated sounds originated at $\alpha = [-20, -17.5, \dots, 17.5, 20]^\circ$, at elevation $\varepsilon = 10^\circ$. Details on ITD and ILD measurements are provided in the Supporting Information.

Moving free-field sounds

We measured the minimum audible movement angle (MAMA) for five of our listeners in the setup, by simulating moving sounds with the two fixed loudspeakers through variation of the amplitude and timing of each sample of the sound according to a linear velocity profile. Sounds thus moved symmetrically through the straight-ahead location ($\alpha = 0^\circ$, where the spatial resolution of the auditory system is best) at $[\pm 25, \pm 50, \pm 100, \pm 150, \pm 200, \pm 250, \pm 400]$ deg/s, for a duration of 3, 10, 50 and 100 ms.

Experimental paradigms

Dynamic localization of free-field head-fixed sounds

Six listeners participated in two different dynamic localization experiments that were performed on different days. The first experiment contained only 100 ms sounds. The second experiment consisted of 3 and 10 ms sounds randomly intermingled. Each experiment consisted of a calibration run, two stationary runs and two dynamic runs.

Calibration. Each experimental session started with a calibration run in which 37 Light Emitting Diodes were presented (direction with respect to horizontal, $\Phi = 0\text{--}360^\circ$ in 30° steps; eccentricity with respect to straight ahead, $R = 0, 10, 20$ and 30° for listeners JO and PB, $R = 0, 13.2, 25$ and 35° for listeners DB, FB, AJ and JT), which the listeners had to fixate. These data were used for offline calibration of the eye-coil signals to veridical eye-in-head orientations.

Stationary condition. We assessed the listener's baseline sound-localization behaviour in two stationary runs, performed at the beginning and end of an experimental session, by presenting simulated free-field sounds in a semi-random order with an interstimulus interval of 3.5 s (total: 68 stimuli per run, four repetitions for each location). The listener had to redirect gaze as fast and as accurately as possible to the perceived location of the sound source, keep gaze there for a moment and then return to straight ahead. During localization trials we did not present an initial fixation light at straight ahead.

Dynamic condition. In the dynamic condition, listeners underwent sinusoidal rotation around the earth-vertical axis at a frequency of $1/9$ Hz, with a peak amplitude of 160° (peak chair velocity 112 deg/s; except for JO in the 100 ms condition: peak amplitude 90° , peak chair velocity 63 deg/s). To avoid discontinuities in chair velocity and acceleration at motion onset, the angular chair velocity increased linearly over the first two sinusoidal periods, during which no sound stimuli were presented. After these two periods, 68 (four repetitions of 17 sounds) stimuli were presented during 27 sinusoidal periods, at an interstimulus interval of 3.5 s. Because the chair had a cycle time of 9 s, stimuli were presented at nine different chair positions, and at 18 different chair velocities and chair accelerations. The task of the listener was to make a saccadic eye movement to the perceived location of the sound as fast and as accurately as possible, briefly fixate this position and then return to the estimated straight-

ahead location. In general, listeners were not able to return to straight ahead due to the ongoing vestibulo-ocular reflex and corresponding ocular nystagmus, and to a potential bias in their estimate. Note, however, that this variation in eye-in-head position is important to disentangle the different potential models (Fig. 1). We gave no additional instructions regarding the reference frame (head-centred, world-centred or otherwise) of the responses. Note that we did not present a fixation light at straight ahead to counteract the vestibulo-ocular reflex. We obtained the same results in pilot experiments with the use of an initial fixation light (data not shown).

Measuring the minimum audible movement angle

Five listeners (two of whom also participated in the dynamic localization experiments) participated in the MAMA experiment, which was designed to test whether listeners could reliably detect the movement direction of 3, 10, 50 and 100 ms sounds, and to estimate the MAMA from the 100 ms data for a stationary listener. To that end, listeners performed four runs of sound-movement discrimination in a two-alternative forced-choice task, in which they were asked to indicate whether the sound moved leftward or rightward. Stimuli were presented at an interstimulus interval of 2 s. Each run consisted of 168 stimuli of a single sound duration (12 repetitions of 14 moving sounds), resulting in four psychometric curves representing the probability of a rightward response as a function of sound velocity, for each sound duration.

In an additional experiment, four listeners (two of whom also participated in the dynamic localization experiments) indicated the stimulus movement direction of 50 ms sounds during the same sinusoidal whole-body rotation as in the dynamic localization experiments. Stimuli were presented at peak chair velocity (i.e. at chair amplitude 0°) in both leftward and rightward chair rotation directions. For each of these two chair positions we presented 12 repetitions of 14 moving sound stimuli, resulting in two psychometric curves.

Psychometric curve fitting

The psychometric data of the MAMA experiment were modelled by a cumulative Gaussian by using the method of maximum likelihood (Wichmann & Hill, 2001). The psychometric curve, $\psi(x)$ (x is stimulus velocity), is thus given by

$$\psi(x) = \lambda + (1 - 2\lambda) \cdot \text{erf}(x; \mu, \sigma) \quad (2)$$

in which $\text{erf}(x; \mu, \sigma)$ is the error function (mean μ , SD σ). The lapse parameter, λ , represents stimulus-independent errors that may be due to mistakes, a bias or random guessing of the listener. This parameter was restricted to be maximally 10%. The 95% confidence intervals of thresholds were determined by bootstrapping ($N = 1000$). We defined the just-notable difference (JND) as half of the distance between the 0.25 and 0.75 fraction rightward judgements. We then estimated the MAMA from the 100 ms moving stimuli by multiplying the stimulus duration with the JND.

Head displacement during stimulus presentation

As explained in the Introduction, stimuli during which the amount of head rotation remains well below the auditory movement detection threshold (the MAMA, see above) are perceived as stationary in space. As the peak chair velocity was 112 deg/s, the maximum chair displacements (and hence potential sound with respect to head displacements) during the 3, 10 and 100 ms were 0.33, 1.1 and 11° ,

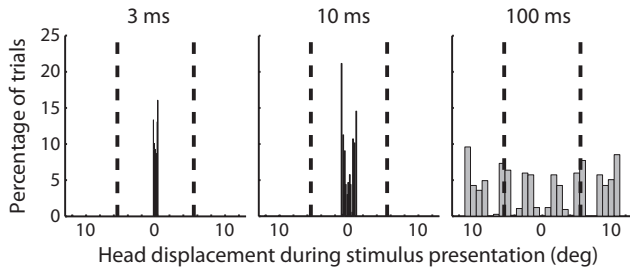


FIG. 2. Histograms of head displacement during stimulus presentation together with the average minimum audible movement angle (dashed lines) of $5.5 \pm 1.4^\circ$ as measured in our setup (see Results).

respectively. Figure 2 summarizes the actual passive head rotations during all applied stimulus presentations, together with the MAMA measured in our setup. Although the MAMA depends on target velocity (Perrott & Saberi, 1990; Chandler & Grantham, 1992), it is always higher than the minimum audible angle of stationary sounds ($5.5 \pm 1.4^\circ$ in our setup, see also Mills, 1958).

Data analysis

Calibration of eye data

The relation between raw eye-position signals and the corresponding Light Emitting Diodes positions was obtained by training two neural networks for the horizontal and vertical eye-position components, respectively. The trained networks were subsequently used to calibrate all eye-position signals (for details, see Goossens & Van Opstal, 1997).

Saccade detection

A custom-made program detected saccades and vestibular quick phases from the calibrated eye-movement signals offline by setting separate eye-velocity thresholds for saccade onset (70 deg/s) and offset (60 deg/s). To differentiate between quick phases of vestibular nystagmus and goal-directed saccades, we required the dynamic goal-directed saccades to have a vertical component (mean \pm SD: $15.5 \pm 10.9^\circ$), as the sounds emanated from loudspeakers at 10° elevation, and the vestibular quick phases had a negligible vertical component. We visually checked the saccade detection markings and made manual changes when deemed necessary. We discarded saccades with latencies shorter than 80 ms and longer than 800 ms. Responses with extremely short latencies were regarded as anticipatory, and very long reaction times as inattentiveness of the listener. Eye positions exceeding 30° (listeners JO and PB) or 35° (listeners DB, FB, JT and AJ) were excluded because of the calibration range (see Calibration section in Experimental paradigms). Typically, <10% of the data was excluded from this analysis. Listeners sometimes responded by making several correction saccades. We report on the first goal-directed saccade in each trial only.

Statistics

For the stationary localization condition, the eye-in-head endpoint of the saccadic response in the azimuth direction, E_{stat} , was quantified by determining the optimal linear fit through the data

$$E_{stat} = g_{stat}\alpha_T + b_{stat} \quad (3)$$

where α_T is the simulated target azimuth, b_{stat} is the bias (offset, in degrees) and g_{stat} the corresponding gain (slope, dimensionless).

Parameters were found by minimizing the mean-squared error (Press *et al.*, 1992). From the linear fit we also determined the correlation coefficient between data and model prediction.

Ideally, for stationary localization, the gain is 1.0 and the bias is 0° . However, parameters g_{stat} and b_{stat} could deviate from the ideal values in an idiosyncratic way. To enable data pooling across listeners and conditions, we normalized the simulated target locations such that response gain and bias for all three stationary conditions were 1.0° and 0.0° , respectively

$$T_H = g_{stat}\alpha_T + b_{stat} \quad \text{and hence} \quad E_{stat} = T_H \quad (4)$$

The normalized target locations, T_H , were then used to perform regression on the saccade endpoints of the dynamic localization responses during vestibular stimulation in world coordinates

$$E_{dyn} = g_{dyn}T_H + b_{dyn} \quad (5)$$

Models

To determine to what extent the auditory system incorporated the intervening vestibular-induced eye and head movements during the reaction time period (see Fig. 1), we performed a multiple linear regression analysis on the horizontal saccadic eye-displacement responses (ΔE) by a linear combination of the normalized initial target location in head-centred coordinates (T_H , Eqn 4), the passive head displacement in space between sound onset and response onset (ΔH_S), and the onset position of the eyes in the head at the start of the response (E_H)

$$\Delta E = aT_H + b\Delta H + cE_H + d \quad (6)$$

in which a , b and c are dimensionless response gains, and d is the response bias (in degrees). In this study we considered four potential spatial updating models to explain auditory-evoked saccade responses (Fig. 1). In Model I, full compensation of eye- and head-displacement signals corresponds to a world-centred target representation. In Model II, only the vestibular-induced change in eye position is accounted for and the target remains in an updated head-centred reference frame. Model III only incorporates the passive change in head orientation, while the audiomotor system is unaware of the intervening vestibular nystagmus. Finally, in Model IV, none of the intervening movement signals are accounted for, and the target remains in its initial head-centred reference frame.

Histograms

The bin width of the histograms (see Fig. 5) was determined by $BW = \sqrt{\text{range}/N}$, where range is the difference between the largest and smallest values (excluding the two most extreme points) and N is the number of included points.

Results

Stationary sound localization (baseline)

To assess the baseline sound-localization performance of the listeners towards the simulated free-field noise bursts in the vestibular setup, they responded to brief sounds without vestibular stimulation during the first and last run of each experimental session. The first two columns of Fig. 3 show typical sound-localization trials of listener DB to 3 ms (top row), 10 ms (centre row) and 100 ms (bottom row) sounds. The left-most panels show

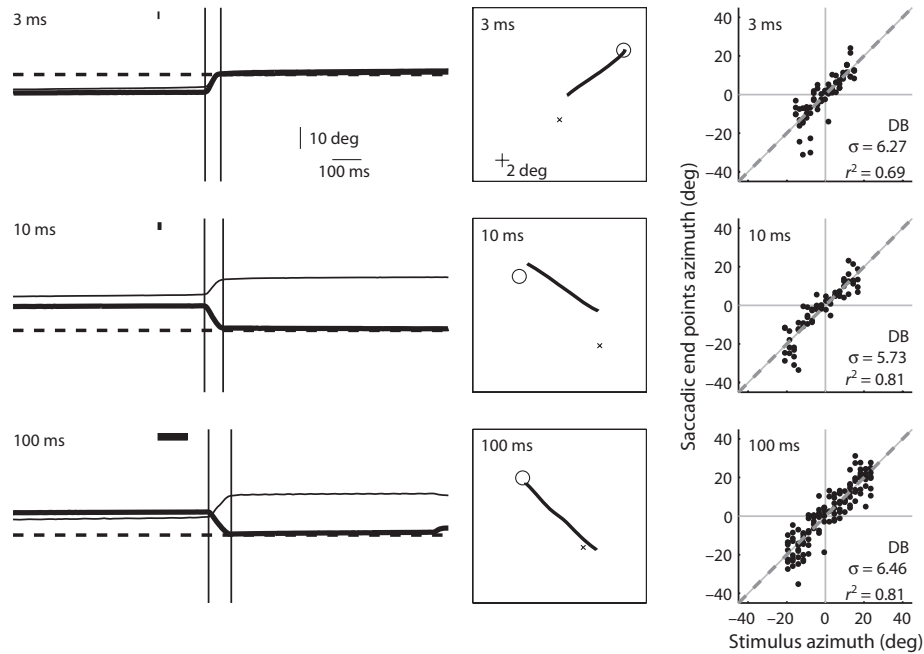


FIG. 3. Stationary sound localization behaviour of listener DB for the three sound durations (rows). First two columns: example eye movements in the stationary condition to 3, 10 and 100 ms noise bursts presented in darkness. In the temporal plots (first column), azimuth (thick trace) and elevation (thin trace) components of eye position are plotted relative to target onset. The thin vertical lines show the localization saccade onsets and offsets. The thick horizontal bar shows target presentation time. The horizontal dashed line corresponds to the normalized stimulus azimuth (T_H ; Eqn 4). The spatial plots (second column) present the saccades of the temporal data of the first column. The small \times denotes straight ahead [α, ε] = [0, 0]. Circles correspond to the target location. In the third column, the endpoints of all saccades of listener DB are plotted as a function of normalized stimulus azimuth in world coordinates. The dashed line shows a linear regression on these endpoints. Coefficients of determination (r^2) and response variability (σ) are given in the lower-right corners.

the calibrated azimuth (bold) and elevation (thin) eye-movement traces relative to target onset, together with the sound's timing (black horizontal bar), and normalized (Eqn 4) azimuth target locations (dashed horizontal line). These examples show that the azimuth responses were reliably directed towards the normalized target locations, even for the briefest stimulus duration. The second column shows the spatial trajectories of the saccades (samples taken between the solid vertical lines in the left-most column) together with the normalized sound locations (circles). Note that the azimuth components of the responses were accurate. In this study, we do not discuss the elevation response components, as listeners were rotated around the earth-vertical axis only, and stimulus elevation was always fixed at $+10^\circ$.

As described in Materials and methods (Eqn 4), we normalized the stimulus locations such that the linear regression line had a slope of 1.0 and no bias. Figure 3 (right-hand column) shows all endpoints of localization responses of listener DB plotted against the normalized stimulus azimuth in the stationary condition. The response endpoints correlated well with the target location, indicating that the responses were indeed goal-directed, even for the very brief click-like sounds of 3 ms (top) and 10 ms (centre) ($r^2 > 0.7$). The left-hand side of Table 1 shows that all listeners had good baseline sound localization performance (uncorrected: $0.65 < \text{gains} < 1.7$, biases $< 13^\circ$, $r^2 > 0.65$ and variances $< 12^\circ$), showing that their responses were indeed solely guided by simulated binaural acoustic information.

Sound-motion detection ability

As argued in the Introduction, adequate spatial updating requires the system to dissociate self-motion from target motion, for which it needs accurate information about movement of the sound source

with respect to the moving head. We expected that the 100 ms stimuli could potentially provide sufficient acoustic information to determine that the sound source moves along with the head, whereas the briefest sounds would be perceived as stationary in space. To test the auditory system's sensitivity for sound-source motion for a stationary and a rotating head, we performed two-alternative forced-choice psychophysics on sounds of 3, 10, 50 and 100 ms duration, moving at a range of leftward and rightward constant velocities (see Materials and methods). As an example, Fig. 4A shows the psychometric curves (Eqn 2) of listener DB for the four stimulus durations. The subject could reliably detect the direction of motion for 100 ms sounds, as the probability of rightward responses for rightward moving sounds is one for speeds exceeding about 75–100 deg/s and is 0 for sounds moving leftward at 75–100 deg/s and higher. The measured threshold is close to 0 deg/s. Figure 4C shows that the JND for the 100 ms sounds was well below 100 deg/s for all five listeners (mean \pm SD: 55 ± 14 deg/s), indicating that the movement direction of these sounds could in principle be detected when moving through the straight-ahead direction. Because in the dynamic localization experiment the sounds were head-fixed, the auditory system might thus be confident about the detected absence of motion relative to the head, and should thus ignore the vestibular signals indicating head-through-space motion. These JNDs correspond to a MAMA of $5.5 \pm 1.4^\circ$ (Fig. 2).

In contrast, the psychometric curves for the 3 and 10 ms sounds were completely flat (Fig. 4A), as the listener could not hear any motion in these short click-like sounds. This can also be appreciated from the JND determined for all listeners (mean \pm SD: 3 ms: $5701 \pm 5964^\circ$, 10 ms: $2454 \pm 1825^\circ$; Fig. 4C), which were much larger (yielding unnaturally high values) than for the 100 ms sounds.

TABLE 1. Localization parameters for all subjects: gain, bias and variance of uncorrected stationary localization (Eqn 3) and dynamic localization (Eqn 5)

Duration (ms)	Listener	Stationary				Dynamic			
		Uncorrected gain (g_{stat})	Bias (b_{stat})	r^2	Variance	Gain (g_{dyn})	Bias (b_{dyn})	r^2	Variance
3	DB	0.76	-0.25	0.69	6.95	0.61	3.79	0.07	21.3
	FB	1.65	1.34	0.87	11.03	0.87	1.14	0.42	20.6
	JO	0.68	0.97	0.80	5.81	0.82	3.65	0.09	22.7
	PB	1.03	-8.07	0.72	7.71	0.73	5.28	0.19	18.0
	JT	1.02	6.15	0.70	8.41	0.80	-1.73	0.05	38.8
	AJ	1.03	9.88	0.72	7.90	0.46	5.91	0.02	36.8
10	DB	0.94	-2.12	0.81	5.77	0.41	1.05	0.05	21.6
	FB	1.69	-1.39	0.86	11.94	0.84	-0.19	0.37	21.4
	JO	0.66	0.25	0.67	7.15	0.75	-1.48	0.05	26.4
	PB	1.42	-9.96	0.76	9.78	0.51	-0.53	0.17	21.1
	JT	1.47	3.93	0.81	10.40	0.72	-10.24	0.11	37.0
	AJ	1.26	4.28	0.83	7.87	0.48	2.02	0.04	34.5
100	DB	1.08	2.02	0.81	6.53	0.67	3.78	0.28	15.3
	FB	1.37	-0.80	0.84	8.64	0.91	-1.11	0.42	18.1
	JO	1.38	-6.22	0.91	6.72	0.67	7.00	0.15	26.8
	PB	0.97	-12.69	0.87	4.67	0.92	-0.31	0.46	11.8
	JT	1.46	3.23	0.85	9.31	0.81	-0.56	0.20	28.3
	AJ	1.44	2.05	0.79	10.27	1.06	2.36	0.25	30.8

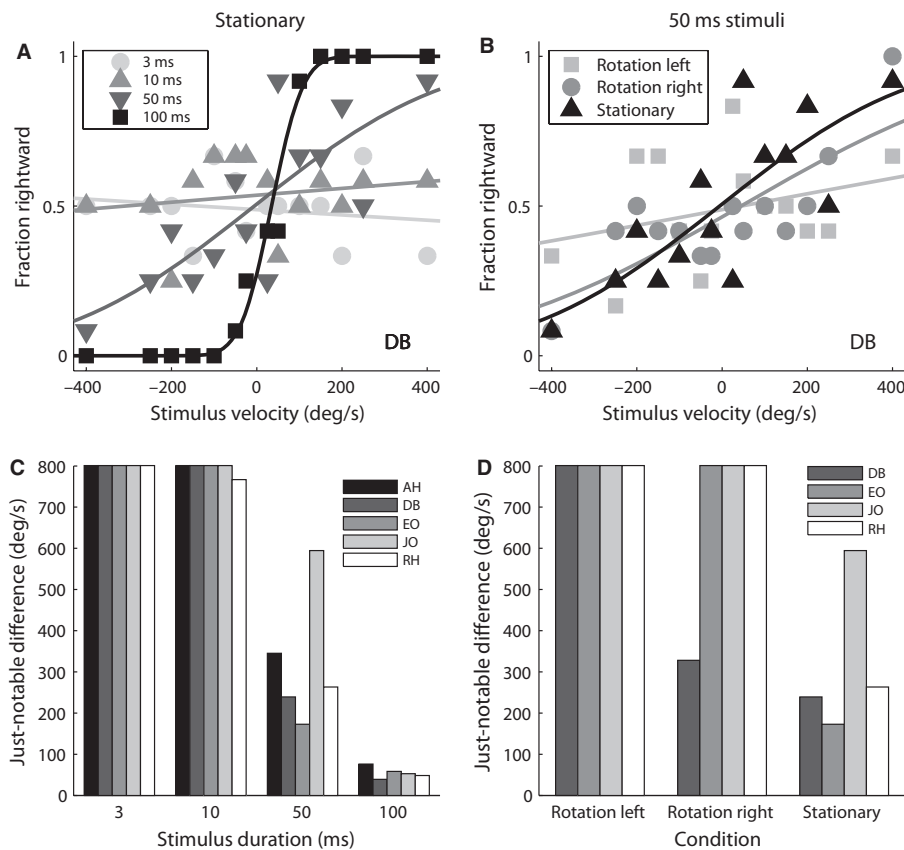


FIG. 4. Detection of direction of sound motion. (A) Fraction of rightward judgements as a function of stimulus velocity (negative: leftward motion) for the four different sound durations for listener DB. (C) Individual JNDs for the four stimulus durations. Note that the bars for 3 and 10 ms sounds extend well above 800 deg/s. (B) Fraction of rightward judgements as a function of stimulus velocity for 50 ms sounds when listener DB was stationary (black triangles) or during rotation (grey squares and circles). (D) Individual JNDs for 50 ms sounds when listeners were stationary or rotating. Note that the bars for rotation left and rotation right extend well above 800 deg/s (except for listener DB rotation right).

The JND for the 50 ms sounds appeared to fall around 200–300 deg/s, indicating that a 50 ms stimulus moving at about 100 deg/s would not be reliably perceived as moving.

To check whether the JND would somehow strongly improve when the listener was subjected to whole-body rotation (at a maximum velocity of 112 deg/s), four listeners performed the same motion-

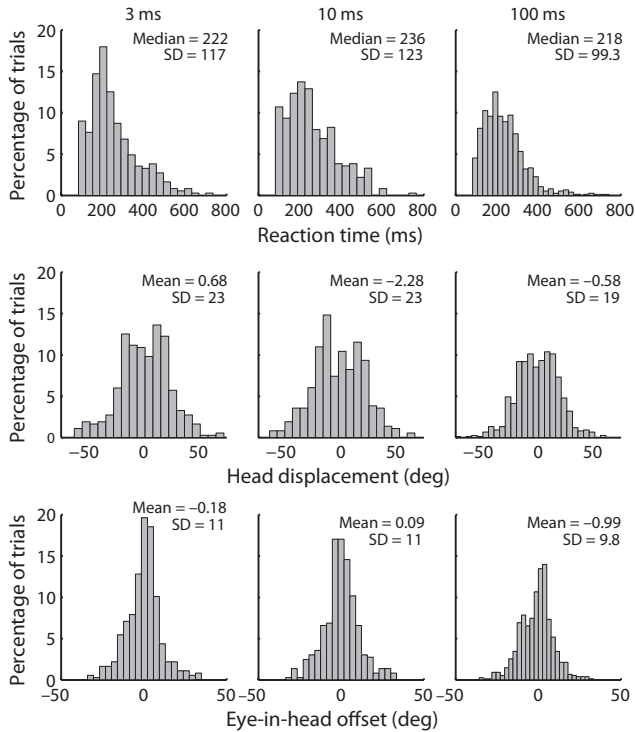


FIG. 5. Distributions of reaction times (latency), head displacement during the reaction times and eye-in-head position at the time of the saccade response for the different sound durations, pooled over all listeners.

discrimination task on the 50 ms sounds in the dynamic rotation condition. We specifically used the 50 ms sounds to detect any improvement in performance, as this stimulus duration lay closest to

threshold performance (Figs 4A and C). Although the example curves in Fig. 4B suggest that, during rotation, the psychometric curves of listener DB may actually have had a lower slope value (i.e. a higher JND, decreased sensitivity) than for the stationary condition, when pooled across listeners, the differences were not significant (t -test: $P > 0.12$). We conclude that, during whole-body rotation, the auditory system is not more sensitive to moving sounds than during the head-stationary condition.

Head and eye movements during the saccade reaction time

The dynamic experiments were designed to ensure considerable and variable passive head movements during the saccadic reaction time of the listeners, who were instructed to make a rapid saccadic eye movement toward the perceived sound location. Figure 5 quantifies the relevant variables from the experimental data, pooled for all listeners. The top row of Fig. 5 shows the reaction-time distributions for the three stimulus durations (mean \pm SD: 3 ms: 222 ± 117 ms, 10 ms: 236 ± 123 ms, and 100 ms: 235 ± 100 ms). The centre row of Fig. 6 shows distributions of head displacements during the measured reaction-time periods. Note that there was a considerable amount of head rotation; the mean was around 0° but the distributions had an SD of about 20° , corresponding to a range from about -70° to $+70^\circ$. The histograms in the bottom row show that the eyes were also not stationary during the reaction time. Although listeners were instructed to redirect their eyes toward the perceived straight ahead in preparation for the next localization response, the eyes deviated significantly from this location due to the ever-present vestibulo-ocular reflex and ocular nystagmus, and to a potential bias in their own estimate of straight ahead. The eye-position variance at response-saccade onset was about 10° , so that initial eye positions almost covered the entire oculomotor range of $\pm 35^\circ$. Note that these variations in eye-in-head position and head displacements were

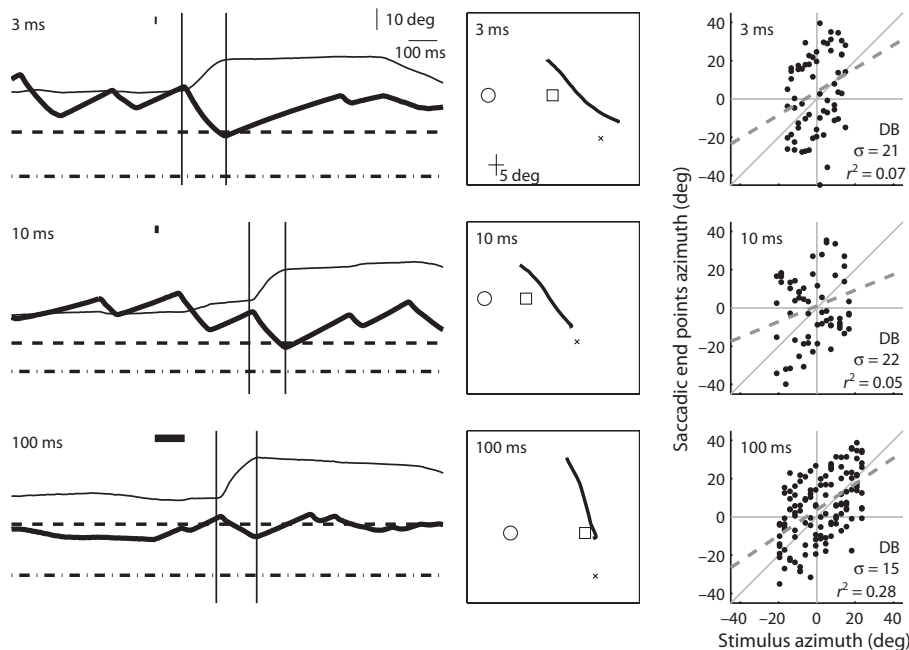


FIG. 6. Dynamic sound localization behaviour of listener DB for the three sound durations (rows). Same conventions as Fig. 3. Target locations are presented in head coordinates (T_H ; Eqn 4; horizontal dashed lines in the first column and squares in the second column) and in world-coordinates ($T_W = T_H - \Delta H_S$; horizontal dot-dashed lines in the first column and circles in the second column).

required to disentangle the different models (outlined in Fig. 1) on dynamic sound-localization performance. Thus, the dynamic stimulation condition indeed challenged the auditory system to a substantial amount of spatial remapping to represent the target in either world coordinates, or to keep it in head-centred coordinates.

Dynamic sound localization

The left-hand side of Fig. 6 shows three trials of subject DB during vestibular stimulation (same format as Fig. 3). Note the clear horizontal nystagmus pattern in the eye-movement traces (bold), which were not accompanied by vertical eye movements (thin). The sound-localization responses, in contrast, had a clear elevation component. Note that the responses in the three examples appear to be directed to the chair-fixed stimulus locations. In the following section we will further quantify these qualitative observations of the response patterns with respect to the different spatial updating models of Fig. 1.

Figure 6 (right-hand column) shows the azimuth localization responses of listener DB in the dynamic conditions for the three stimulus durations (rows); we expressed the final eye position after the localization saccade in world coordinates (Eqn 5). Unlike the stationary condition (Fig. 3, right), the azimuth responses correlated poorly with the normalized world-centred target location for the dynamic vestibular condition. Linear regression yielded a large variability ($\sigma > 28^\circ$) and a very low correlation ($r^2 < 0.2$). The right-hand side of Table 1 shows that this was the case for all listeners

($r^2 < 0.46$, variances $> 12^\circ$). The horizontal scatter in the responses was significantly larger than in the static condition (Kolmogorov-Smirnov test: $P < 10^{-7}$ for all listeners).

Testing the spatial updating models

The linear regression analysis on the response data in Fig. 6 suggests that listener DB did not incorporate the passive head rotation in her localization responses. This holds for all listeners (Table 1). To test which of the four models described in Materials and methods (Fig. 1, Eqn 6) best accounted for the behaviour of all listeners, we applied the optimal regression coefficients of the models to the head- and eye-movement data summarized in Fig. 5 (centre and bottom rows) to predict the saccadic response. Figure 7 presents the predicted horizontal saccadic eye displacement of each model against the measured saccade component pooled for all listeners. The results make clear that, for all sound durations, the craniocentric Model II ($\Delta E = T_H - E_H$, see Materials and methods) yields by far the highest correlation (r^2) and lowest variability (σ) between predicted and measured saccadic responses (Kolmogorov-Smirnov test on Model I, III and IV errors compared with Model II errors: $P < 10^{-3}$ for all stimulus durations). This suggests that, whereas the audio motor system accounted for the vestibular-induced changes in eye-in-head position (Fig. 5, bottom row), the substantial passive head rotations during the listener's reaction time (Fig. 5, centre row) were not incorporated in the responses.

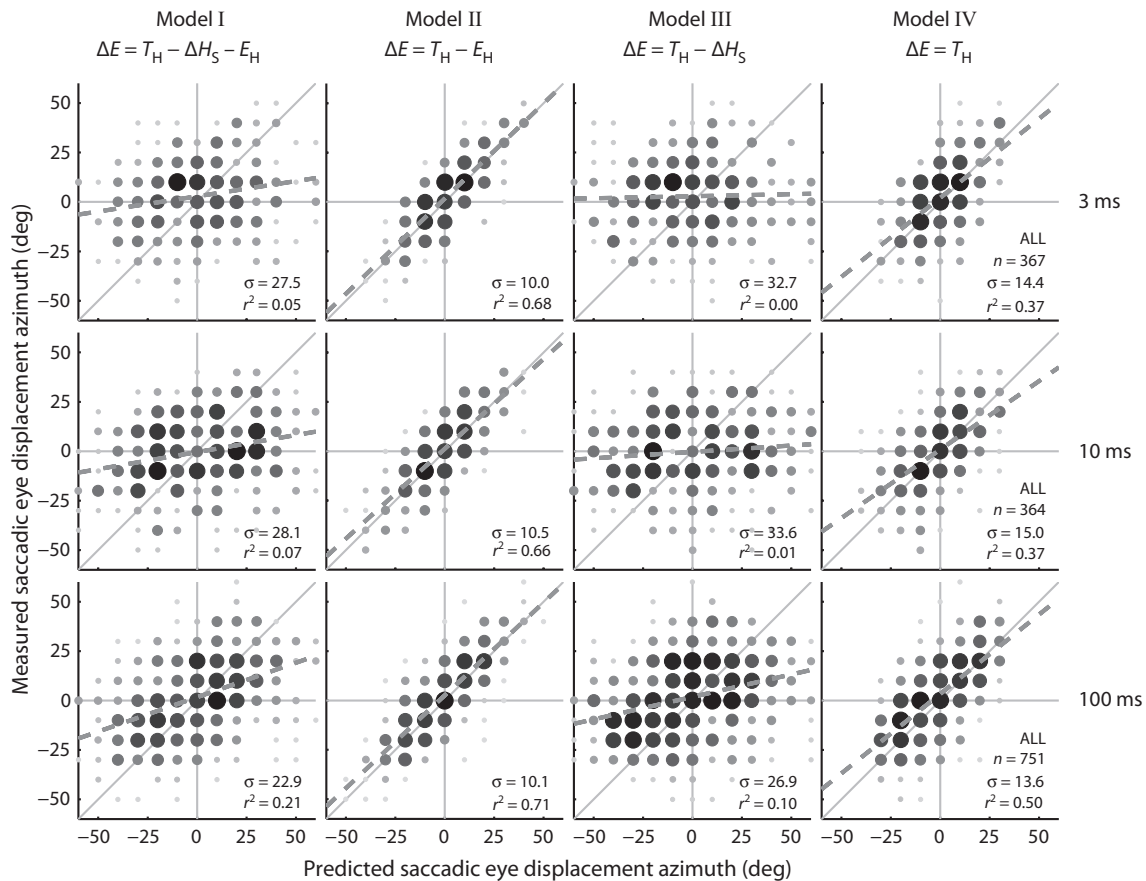


FIG. 7. Predicted saccadic azimuth for the four models described in Materials and methods plotted against the measured saccade azimuth pooled for all listeners. If the model predicted the listener's responses perfectly, responses would fall on the unity line and r^2 would be one. Coefficients of determination (r^2) and response variability (σ) are given in the lower-right corners. Model II has the highest correlation and smallest variability between predicted and response saccade azimuth.

TABLE 2. Multiple linear regression parameters of individual listeners

Duration (ms)	Listener	Multiple linear regression analysis				r^2
		T_H	ΔH	E_H	Bias	
		$a \pm \text{SD}$	$b \pm \text{SD}$	$c \pm \text{SD}$	$d \pm \text{SD}$	
3	AJ	0.90 ± 0.10	0.20 ± 0.06	-0.52 ± 0.17	3.94 ± 1.69	0.76
	DB	0.96 ± 0.07	0.16 ± 0.08	-0.65 ± 0.18	1.67 ± 0.97	0.75
	FB	0.98 ± 0.04	0.18 ± 0.06	-0.92 ± 0.15	-0.77 ± 1.06	0.87
	JO	0.89 ± 0.09	0.39 ± 0.10	-0.63 ± 0.17	8.06 ± 1.72	0.72
	JT	1.58 ± 0.12	0.24 ± 0.06	-0.63 ± 0.14	-6.60 ± 1.84	0.73
	PB	0.66 ± 0.08	0.24 ± 0.05	-0.91 ± 0.20	2.74 ± 0.67	0.75
10	AJ	0.96 ± 0.10	0.06 ± 0.05	-0.92 ± 0.21	1.61 ± 1.53	0.62
	DB	0.81 ± 0.06	0.32 ± 0.06	-0.56 ± 0.18	2.28 ± 0.90	0.75
	FB	1.11 ± 0.07	0.25 ± 0.08	-0.92 ± 0.20	2.85 ± 1.30	0.84
	JO	0.66 ± 0.09	0.30 ± 0.09	-0.75 ± 0.16	3.60 ± 2.43	0.63
	JT	1.03 ± 0.09	0.18 ± 0.06	-0.63 ± 0.11	-4.00 ± 1.68	0.74
	PB	1.03 ± 0.09	0.20 ± 0.04	-0.94 ± 0.20	0.77 ± 0.65	0.79
100	AJ	1.05 ± 0.06	0.04 ± 0.06	-0.73 ± 0.16	4.60 ± 1.45	0.65
	DB	0.87 ± 0.05	0.01 ± 0.04	-0.67 ± 0.12	2.23 ± 0.57	0.78
	FB	1.12 ± 0.04	0.13 ± 0.04	-0.95 ± 0.10	-0.44 ± 0.64	0.88
	JO	0.93 ± 0.04	0.42 ± 0.07	-0.74 ± 0.16	3.42 ± 0.79	0.83
	JT	1.06 ± 0.05	0.07 ± 0.05	-1.00 ± 0.10	0.53 ± 1.10	0.79
	PB	0.71 ± 0.05	0.26 ± 0.04	-0.89 ± 0.16	8.57 ± 0.79	0.70

The horizontal component of the saccade was described as a function of the horizontal target position relative to the head (target, T_H), head displacement in the reaction time (ΔH) and eye-in-head offset at the response (E_H , see Eqn 5). The table lists the regression coefficients (slopes a , b and c , bias d in degrees).

Multiple linear regression analysis

To quantify the actual contributions of initial sound location, eye-in-head position and passive head displacement in sound-evoked saccades during vestibular stimulation, we performed a multiple linear regression analysis on the listener's saccadic eye displacements (Eqn 6). Table 2 provides the resulting coefficients for the individual listeners, and Fig. 8 shows the averaged results across listeners for each of the three sound durations. The data show that there is no effect of stimulus duration on the responses, as the gain of the craniocentric target location (T_H) (ANOVA: $F_{2,15} = 0.12$, $P = 0.89$), gain of the passive head displacement during the reaction time (ΔH) (ANOVA: $F_{2,15} = 0.83$, $P = 0.45$) and gain of the eye-in-head position at saccade onset (E_H) (ANOVA: $F_{2,15} = 0.92$, $P = 0.42$) did not depend on stimulus duration. These results show that for all sounds, and for all listeners, the T_H gain did not differ from the ideal value of +1 (t -test: $P = 0.44$), whereas the E_H gain was close to -1. The ΔH gain,

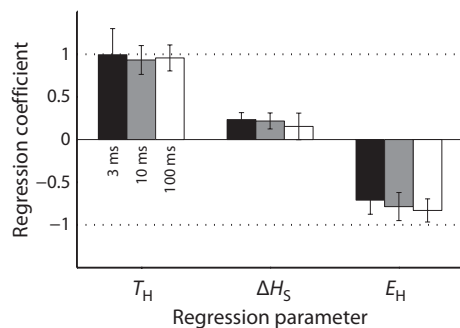


FIG. 8. Multiple linear regression results (Eqn 6) on the listener's responses to 3, 10 and 100 ms noise bursts during passive whole-body rotation. Bars represent the regression coefficients of Eqn 6 averaged across listeners. Error bars represent SD. The horizontal lines at +1 and -1 correspond to ideal regression values (see Fig. 1).

however, remained close to 0 (even slightly positive, $P < 0.05$). In other words, the initial craniocentric target location (derived from simulated acoustic cues), as well as the change in eye-in-head position at the time of the response, were both incorporated in planning the saccadic response, although the eye-in-head orientation appeared to be slightly underestimated (mean ± SD: $c = -0.77 \pm 0.16$). Importantly, the passive head displacements were not accounted for. These results therefore contrast markedly with the findings of the dynamic localization study by Vliegen *et al.* (2004), in which listeners made spatially accurate saccades to brief sound bursts under head-unrestrained conditions.

Variability of stationary vs. dynamic conditions

Although listeners reported finding the localization task difficult, especially for the briefest noise bursts, their responses appeared to be quite consistent with only modest variability. To quantify this aspect of the data, Fig. 9 compares the saccadic endpoints of all listeners pooled for the stationary and dynamic conditions against the normalized target in head-centred coordinates. Note that the response variability for the dynamic condition (mean ± SD over subjects and stimulus durations: $9.59 \pm 0.38^\circ$) is higher than for the stationary condition ($7.02 \pm 0.09^\circ$) (t -test: $P < 10^{-4}$).

Discussion

We investigated the spatial updating of auditory-evoked saccades towards short (3 and 10 ms) and long (100 ms) noise bursts during passive whole-body rotation. Although the auditory system could not detect the presence or absence of stimulus motion with respect to the head for the shortest stimuli (Fig. 4), all sounds were localized in head-centred coordinates (Figs 7 and 8). In doing so, it ignored vestibular information about passive-induced head rotations, but accounted for changes in eccentric eye-in-head position imposed by

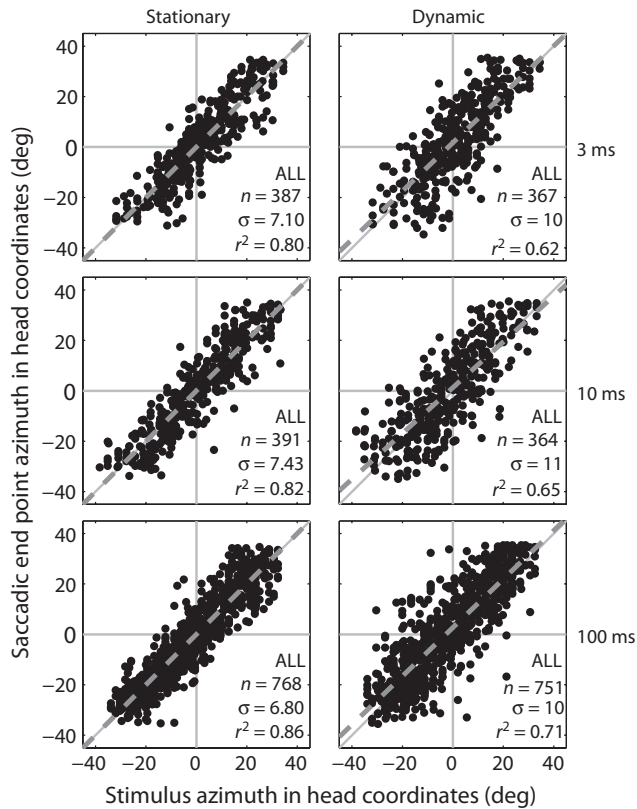


Fig. 9. Comparison of stationary and dynamic localization conditions regarding response variability. Data are pooled across listeners and expressed in normalized head-centred coordinates. Despite the subjects' reported difficulty in the task, response variability in the dynamic condition ($\sigma \approx 10^\circ$) is only slightly larger than in the stationary condition ($\sigma \approx 7^\circ$).

vestibular nystagmus. These results contrast markedly with previous studies using brief (3–50 ms) sound bursts, presented either prior to, or during, actively generated intervening eye-head gaze shifts. Under such conditions, sound-localization performance was spatially accurate (Goossens & Van Opstal, 1999; Vliegen *et al.*, 2004). Below we will discuss possible explanations for this discrepancy.

Comparison with previous studies

The major differences between our study and previous studies are as follows: (i) we denied the audiomotor system use of corollary discharge information of planned head movements, (ii) there was no proprioceptive information regarding changes of head-on-trunk orientation, (iii) the head was rotated passively and (iv) the sound source moved along with the head.

The visual-auditory double-step paradigm applied to head-unrestrained eye-head gaze shifts demonstrated that the audio motor system updates its sound-localization responses in world (or body)-centred coordinates (Goossens & Van Opstal, 1999; Vliegen *et al.*, 2004). Under such conditions, all sources of information about eye and head movements (neck proprioception, corollary discharges of planned movements and vestibular signals) are available and consistent, and the listener is engaged in the active planning of a sequence of goal-directed gaze shifts. Vliegen *et al.* (2004) showed that accuracy (i.e. mean localization error) and precision (i.e. response variability) for static (sound presented before the voluntary head movement) and dynamic (sound presented during the head saccade) acoustic condi-

tions were almost indistinguishable from sound-localization responses without intervening movements. In the current experiments, response variability was slightly higher in dynamic localization trials than in static trials (Fig. 9). However, given that during the dynamic localization task the listener had to combine the acoustic localization cues with dynamic changes in eye position, while at the same time filtering out acoustic background noise from the vestibular apparatus, as well as the vestibular head-movement signal, the difference in response variance ($\sim 37\%$ increase of the SD) was modest.

Interestingly, Vliegen *et al.* (2004) also showed that the auditory system continues to process acoustic-cue information under dynamic localization conditions, as gaze shifts were systematically better for longer-duration (50 ms) noise bursts, than for very brief (3 ms) sounds. The auditory system thus appeared to accumulate the processing of acoustic input as time progresses (Frens & Van Opstal, 1995; Hofman & Van Opstal, 1998; Vliegen *et al.*, 2004). During head movements the acoustic localization cues change in a systematic and predictable way. Thus, for accurate localization, the head movement should be incorporated to compensate for these dynamic acoustic changes. This is not a trivial problem because of the different representations of the involved signals (e.g. tonotopic sensory codes vs. rate coding of neck muscle contractions).

Under head-fixed stationary conditions, self-generated dynamic changes in eye position are almost fully incorporated in sound localization (with a gain of about -0.95) (Van Grootel & Van Opstal, 2009; Van Grootel & van Opstal, 2010). Because ocular responses also account for the involuntary drift of the eyes in darkness, spatial updating of sounds is based on an accurate representation of instantaneous eye position. The current experiments further extend these results to the eye-position changes of ocular nystagmus imposed by passive vestibular stimulation, with an eye-position gain of about -0.8 (Figs 5 and 7).

Two other studies have investigated spatial oculomotor behaviour during whole-body vestibular stimulation around the yaw axis. Van Beuzekom & Van Gisbergen (2002) measured saccades to briefly-flashed (4 ms) visual targets that moved along with the head, and reported that subjects correctly localized stimuli in head-centred coordinates. We recently extended these experiments to visual flashes of different durations, presented at either head-fixed or world-fixed locations (Van Barneveld *et al.*, 2011a). The results showed that spatial localization of brief visual flashes was only accurate if the visual system could reliably infer the direction of stimulus motion across the retina. Thus, long-duration (100 ms) flashes produced substantial visual streak patterns that could be updated in head-centred coordinates for head-fixed targets, and in world coordinates for world-stationary targets. In contrast, short flashes (0.5 and 4 ms) provided no detectable retinal motion cues and were kept in their initial sensory (i.e. retinal) coordinates. As a result, these brief stimuli were mislocalized.

Why no head-movement compensation?

We consider several possibilities to explain why, in our experiments, the audiomotor system did not compensate for the passive-induced head rotations, and instead kept sounds in an updated head-centred reference frame.

First, one might argue that listeners knew that the sounds were moving along with them, and could therefore have voluntarily adopted a craniocentric strategy. We believe that this explanation is unlikely, because listeners were responding without any audiovisual feedback and were not instructed to localize the target in any particular reference frame. If they followed a voluntary strategy, part of the trials (or listeners) would have violated this strategy (either erroneously or by adopting a different default). This was not observed. Moreover, if

knowledge of the chair-fixed speaker locations played a role, part of the responses should have been directed toward the actual speaker locations (at $\pm 37^\circ$), which never occurred. Instead, responses were reliably guided by the virtual, simulated sound locations at short reaction times, meanwhile incorporating the highly variable initial eye-position signal (Fig. 5). From this we infer the use of bottom-up acoustic and motor information sources, rather than top-down cognitive signals. Furthermore, although listeners made reliable and fast responses towards the sounds, they judged the task to be very difficult and were unable to reflect on their performance during the open-loop experiments. We therefore do not consider it feasible that listeners voluntarily disconnected the remapping of a (vestibularly-induced) dynamic head-movement signal, while at the same time remapping unpredictable eye positions.

Second, in the dynamic localization experiments, listeners were rotated at variable speeds during sound presentation, but received no dynamically changing acoustic cues, as the loudspeakers were attached to the chair. Could the auditory system have detected the absence of changing acoustic cues? For 100 ms sounds during high-velocity rotation, the answer might be yes, as our psychophysical results showed that fast-moving 100 ms sounds were correctly perceived as moving (Figs 4A and C). For the 3 and 10 ms sounds, however, the amount of head rotation remained well below the MAMA (Fig. 2) (Mills, 1958; Saberi & Perrott, 1990; Chandler & Grantham, 1992), and we also verified that whole-body rotation did not improve JNDs (Figs 4B and D). As a result, the auditory system appears to regard these short stimuli as stationary in space. Note also that, as the head movements followed sinusoidal acceleration profiles, the head displacements also did not reach the MAMA in the majority of trials for 100 ms sounds (Fig. 2). We therefore consider it improbable that the auditory system could have reliably discerned whether or not sounds were moving along with the chair on the basis of acoustic input. However, as demonstrated by the near-perfect regression coefficients for the head-centred target location [a close to one (Eqn 6); Fig. 8], subjects accurately extracted the head-centred coordinates from the simulated acoustic cues during static and dynamic localization trials. Under natural hearing conditions, the most likely situation for a world-stationary sound is a change in acoustic cues that anticorrelates with ongoing head movements, and it is highly unlikely that sounds are fixed to the head. We therefore reasoned that, in the absence of evidence for dynamic changes in the acoustic cues, the auditory system's default assumption ('prior') would be that such sounds are regarded as stationary in space. This was indeed observed for head-unrestrained orienting responses (Vliegen *et al.*, 2004), but our results indicate a different response mode for passive-induced head movements.

A third explanation might be that the auditory system uses signals for a coordinate transformation on the basis of their reliability. When motion cues are in conflict (in this study: vestibular information reports head motion in the absence of efference copies, neck-proprioceptive cues and acoustic changes), the vestibular signal may contribute with only a small weight to the required coordinate transformations for world-centred stability. Although the eye movements were consistent with the interpretation of a vestibularly-induced head movement, the slow-phase vestibulo-ocular reflex showed that this compensation was far from perfect, as its gain was well above -1.0 (see also Van Barneveld *et al.*, 2011a). Moreover, ocular nystagmus may also occur in the absence of vestibular stimulation, and would therefore be unreliable as a single source for spatial updating. Thus, if head-movement evidence is unreliable, it may be ignored altogether and the sound would be kept in its initial, head-centred reference frame. Similar ideas, based on Bayesian statistical

inference, have been proposed for perceptual tasks, multisensory integration and sensorimotor performance (Battaglia *et al.*, 2003; Niemeier *et al.*, 2003; Alais & Burr, 2004; Kording *et al.*, 2007; De Vrijer *et al.*, 2009; Van Barneveld *et al.*, 2011b).

Implications for spatial perception and behaviour

Our recent visual-vestibular experiments demonstrated that the presence or absence of spatial updating of brief visual flashes relied on the integrity of the retinal motion signal (Van Barneveld *et al.*, 2011a), rather than on the presumed reliability of the head-motion signal. Thus, the vestibular-only signal can be used for adequate spatial updating, despite the absence of efference copies or neck proprioception. We here conjecture that this is also the case for spatial updating of auditory stimuli. We hypothesize that the default strategy of the visual and auditory systems under vestibular-only stimulation would be to keep targets in their initial reference frame (audition: head-centred; vision: eye-centred) until sufficient sensory evidence indicates stimulus motion relative to the head (or to the eye). The system employs accurate spatial updating only if the sensory evidence is reliable. These default strategies for auditory and visual remapping may seem surprising, or even suboptimal in the light of statistical inference models, as in daily life it is highly unlikely that sounds move along with the head or that visual stimuli are fixed on the retina.

Supporting Information

Additional supporting information may be found in the online version of this article:

Data S1. Simulation of free-field sound locations with two speakers. Please note: As a service to our authors and readers, this journal provides supporting information supplied by the authors. Such materials are peer-reviewed and may be re-organized for online delivery, but are not copy-edited or typeset by Wiley-Blackwell. Technical support issues arising from supporting information (other than missing files) should be addressed to the authors.

Acknowledgements

This research was supported by the Radboud University Nijmegen (A.J.V.O., F.B. and D.C.P.B.M.V.B.) and the Netherlands Organisation for Scientific Research, NWO, grant no. 805.05.003 ALW/VICI (A.J.V.O.). The authors thank Hans Kleijnen, Ger van Lingen and Stijn Martens for valuable technical assistance.

Abbreviations

ILD, interaural level difference; ITD, interaural timing difference; JND, just-notable difference; MAMA, minimum audible movement angle.

References

- Alais, D. & Burr, D. (2004) The ventriloquist effect results from near-optimal bimodal integration. *Curr. Biol.*, **14**, 257–262.
- Angelaki, D.E. & Cullen, K.E. (2008) Vestibular system: the many facets of a multimodal sense. *Annu. Rev. Neurosci.*, **31**, 125–150.
- Armstrong, B., McNair, P. & Taylor, D. (2008) Head and neck position sense. *Sports Med.*, **38**, 101–117.
- Battaglia, P.W., Jacobs, R.A. & Aslin, R.N. (2003) Bayesian integration of visual and auditory signals for spatial localization. *J. Opt. Soc. Am. A.*, **20**, 1391–1397.
- Blauert, J. (1997) *Spatial Hearing: The Psychophysics of Human Sound Localization*. MIT Press, Cambridge, MA.

- Chandler, D.W. & Grantham, D.W. (1992) Minimum audible movement angle in the horizontal plane as a function of stimulus frequency and bandwidth, source azimuth, and velocity. *J. Acoust. Soc. Am.*, **91**, 1624–1636.
- Collewijn, H., van der Mark, F. & Jansen, T.C. (1975) Precise recording of human eye movements. *Vision Res.*, **15**, 447–450, IN445-447-450, IN445.
- Crapse, T.B. & Sommer, M.A. (2008) Corollary discharge across the animal kingdom. *Nat. Rev. Neurosci.*, **9**, 587–600.
- De Vrijer, M., Medendorp, W.P. & Van Gisbergen, J.A.M. (2009) Accuracy-precision trade-off in visual orientation constancy. *J. Vis.*, **9**, 1–15.
- Frens, M.A. & Van Opstal, A.J. (1995) A quantitative study of auditory-evoked saccadic eye movements in two dimensions. *Exp. Brain Res.*, **107**, 103–117.
- Goossens, H.H.L.M. & Van Opstal, A.J. (1997) Human eye-head coordination in two dimensions under different sensorimotor conditions. *Exp. Brain Res.*, **114**, 542–560.
- Goossens, H.H.L.M. & Van Opstal, A.J. (1999) Influence of head position on the spatial representation of acoustic targets. *J. Neurophysiol.*, **81**, 2720–2736.
- Hofman, P.M. & Van Opstal, A.J. (1998) Spectro-temporal factors in two-dimensional human sound localization. *J. Acoust. Soc. Am.*, **103**, 2634–2648.
- Jay, M.F. & Sparks, D.L. (1984) Auditory receptive fields in primate superior colliculus shift with changes in eye position. *Nature*, **309**, 345–347.
- Jay, M.F. & Sparks, D.L. (1987) Sensorimotor integration in the primate superior colliculus. II. Coordinates of auditory signals. *J. Neurophysiol.*, **57**, 35–55.
- Knudsen, E.I. & Konishi, M. (1979) Mechanisms of sound localization in the barn owl (*Tyto alba*). *J. Comp. Physiol.*, **133**, 13–21.
- Koenigs, K., Knoll, J. & Bremmer, F. (2007) Localisation of auditory targets during optokinetic nystagmus. *Perception*, **36**, 1507–1512.
- Kopinska, A. & Harris, L.R. (2003) Spatial representation in body coordinates: evidence from errors in remembering positions of visual and auditory targets after active eye, head, and body movements. *Can. J. Exp. Psychol. Revue. Can. Psychol. Exp.*, **57**, 23–37.
- Kording, K.P., Beierholm, U., Ma, W.J., Quartz, S., Tenenbaum, J.B. & Shams, L. (2007) Causal inference in multisensory perception. *PLoS One*, **2**, e943.
- Lewald, J. (1997) Eye-position effects in directional hearing. *Behav. Brain Res.*, **87**, 35–48.
- Middlebrooks, J.C. (1992) Narrow-band sound localization related to external ear acoustics. *J. Acoust. Soc. Am.*, **92**, 2607–2624.
- Mills, A.W. (1958) On the Minimum Audible Angle. *J. Acoust. Soc. Am.*, **30**, 237–246.
- Niemeier, M., Crawford, J.D. & Tweed, D.B. (2003) Optimal transsaccadic integration explains distorted spatial perception. *Nature*, **422**, 76–80.
- Perrott, D.R. & Saberi, K. (1990) Minimum audible angle thresholds for sources varying in both elevation and azimuth. *J. Acoust. Soc. Am.*, **87**, 1728–1731.
- Poppel, E. (1973) Comment on “visual system’s view of acoustic space”. *Nature*, **243**, 231.
- Press, W.H., Flannery, B.P., Teukolsky, S.A. & Vetterling, W.T. 1992. *Numerical Recipes in C: The Art of Scientific Computing*. Cambridge University Press, Cambridge, MA.
- Saberi, K. & Perrott, D.R. (1990) Minimum audible movement angles as a function of sound source trajectory. *J. Acoust. Soc. Am.*, **88**, 2639–2644.
- Van Barneveld, D.C.P.B.M. & Van Opstal, A.J. (2010) Eye position determines audiovestibular integration during whole-body rotation. *Eur. J. Neurosci.*, **31**, 920–930.
- Van Barneveld, D.C.P.B.M., Kiemeneij, A.C.M. & Van Opstal, A.J. (2011a) Absence of spatial updating when the visuomotor system is unsure about stimulus motion. *J. Neurosci.*, **31**, 10558–10568.
- Van Barneveld, D.C.P.B.M., Van Grootel, T.J., Alberts, B. & Van Opstal, A.J. (2011b) The effect of head roll on perceived auditory zenith. *Exp. Brain Res.*, **213**, 235–243.
- Van Beuzekom, A.D. & Van Gisbergen, J.A. (2002) Interaction between visual and vestibular signals for the control of rapid eye movements. *J. Neurophysiol.*, **88**, 306–322.
- Van Grootel, T.J. & Van Opstal, A.J. (2009) Human sound-localization behaviour after multiple changes in eye position. *Eur. J. Neurosci.*, **29**, 2233–2246.
- Van Grootel, T.J. & van Opstal, A.J. (2010) Human sound localization accounts for ocular drift. *J. Neurophysiol.*, **103**, 1927–1936.
- Vliegen, J., Van Grootel, T.J. & Van Opstal, A.J. (2004) Dynamic sound localization during rapid eye-head gaze shifts. *J. Neurosci.*, **24**, 9291–9302.
- Wichmann, F.L. & Hill, N.J. (2001) The psychometric function: I. Fitting, sampling, and goodness of fit. *Percept. Psychophys.*, **63**, 1293–1313.
- Wightman, F.L. & Kistler, D.J. (1989) Headphone simulation of free-field listening. I: stimulus synthesis. *J. Acoust. Soc. Am.*, **85**, 858–867.