CALL FOR PAPERS | Active Sensing

Dependence of auditory spatial updating on vestibular, proprioceptive, and efference copy signals

Daria Genzel,1,2 Uwe Firzlaff,2,3 Lutz Wiegrebe,1,2 and Paul R. MacNeilage2,4

1Department Biology II, Ludwig-Maximilian University of Munich, Planegg-Martinsried, Germany; 2Bernstein Center for Computational Neuroscience Munich, Planegg-Martinsried, Germany; 3Chair of Zoology, Technische Universität München, Freising-Weihenstephan, Germany; and 4Deutsches Schwindel- und Gleichgewichtszentrum, University Hospital of Munich, Munich, Germany

Submitted 19 January 2016; accepted in final form 9 May 2016

Genzel D, Firzlaff U, Wiegrebe L, MacNeilage PR. Dependence of auditory spatial updating on vestibular, proprioceptive, and efference copy signals. J Neurophysiol 116: 765–775, 2016. First published May 11, 2016; doi:10.1152/jn.00052.2016.—Humans localize sounds by comparing inputs across the two ears, resulting in a head-centered representation of sound-source position. When the head moves, information about head movement must be combined with the head-centered estimate to correctly update the world-centered sound-source position. Spatial updating has been extensively studied in the visual system, but less is known about how head movement signals interact with binaural information during auditory spatial updating. In the current experiments, listeners compared the world-centered azimuthal position of two sound sources presented before and after a head rotation that depended on condition. In the active condition, subjects rotated their head by ~35° to the left or right, following a pretrained trajectory. In the passive condition, subjects were rotated along the same trajectory in a rotating chair. In the cancellation condition, subjects rotated their head as in the active condition, but the chair was counter-rotated on the basis of head-tracking data such that the head effectively remained fixed in space while the body rotated beneath it. Subjects updated most accurately in the passive condition but erred in the active and cancellation conditions. Performance is interpreted as reflecting the accuracy of perceived head rotation across conditions, which is modeled as a linear combination of proprioceptive/efference copy signals and vestibular signals. Resulting weights suggest that auditory updating is dominated by vestibular signals but with significant contributions from proprioception/efference copy. Overall, results shed light on the interplay of sensory and motor signals that determine the accuracy of auditory spatial updating.

AUDITORY LOCALIZATION depends on binaural cues such as interaural time differences and interaural level differences, which provide information about the sound location in a head-centered reference frame. Numerous studies have examined auditory localization under conditions in which the head is held stationary (e.g., Lewald et al. 1999; Makous and Middlebrooks 1990). However, when the head moves, knowledge of head movement is required to map head-centered representations into the world-centered reference frame, thereby facilitating spatial constancy (Wallach 1940). This process is referred to as spatial updating, and it has most often been investigated using visual targets (for reviews see Klier and Angelaki 2008; Medendorp 2011).

Fewer studies have investigated purely auditory spatial updating. Goossens and van Opstal (1999) reported accurate saccades to auditory targets despite intervening eye and head movements. Further studies using saccadic responses have demonstrated that auditory updating also is accurate (Van Grootel et al. 2011) when sounds are presented during active eye-head gaze shifts (Vliegen et al. 2004), but not for briefly presented sounds during passive movement (Van Barneveld et al. 2011). Yost et al. (2015) did not use saccadic responses but also reported that updating performance was impaired for passive rotations with eyes closed compared with eyes open.

A central question is the source of the head movement signals that drive the updating process. Vestibular signals provide information about the angular velocity and linear acceleration of the head in space, which could be used to derive an estimate of head movement. However, inaccurate updating following passive movements (Van Barneveld et al. 2011; Yost et al. 2015) suggests that vestibular signals alone are insufficient. Other signals, including neck proprioceptive signals from stretch receptors in muscles and efference copies of motor commands, provide information about how the head is positioned on the body. Assuming that the body is stationary, vestibular, proprioceptive, and efference copy signals will provide redundant information.

Several previous studies have investigated the independent contributions of these various signals for estimating head motion (e.g., Crowell et al. 1998; Cullen and Roy 2004). Inaccurate updating during normal active head turns, consistent with overestimation of head turn angle.

NEW & NOTEWORTHY

Comparison of updating performance across active, passive, and cancellation head movement conditions reveals that vestibular signals dominate auditory spatial updating. However, neck proprioceptive and/or efference copy signals also play a role and appear to be responsible for...
studies focused specifically on visual spatial updating (e.g., Nakamura and Bronstein 1995; Mergner et al. 1998, 2001), subjects were presented with a brief visual target, after which the subject’s head is rotated. Finally, the subject makes a saccade back to the original target location in space, which required them to account for the experienced change in head orientation. Precision and accuracy of these saccadic responses depended on the nature of the intervening head rotation. Spatial updating was most accurate when head rotation was active such that vestibular, proprioceptive, and efference copy signals were all available. Gain was significantly reduced during passive movement when only proprioceptive or vestibular signals indicated rotation. Also, the precision of responses was greatest when all modalities signaled the rotation. Thus the evidence suggests that diverse signals are combined to estimate head rotation in space for the purposes of maintaining spatial constancy.

In the present study we apply a methodological approach similar to those described above, but instead of an oculomotor (saccadic) response, we employ a two-interval forced-choice (2-IFC) task that avoids response-based errors due to remapping of perceived location into a pointing (e.g., saccadic) response. We use the results to infer the relative weights given to vestibular and proprioceptive and/or efference copy signals for auditory spatial updating. Auditory updating is well suited for isolating the contributions of proprioceptive and vestibular signals because auditory spatial information is sensed directly in head coordinates, unlike visual information, which depends on oculomotor processing. Comparison of our results with previous studies allows us to examine if and how spatial updating processes generalize across visual and auditory modalities and across response type (e.g., saccadic vs. perceptual).

MATERIAL AND METHODS

Subjects

Ten subjects (6 men and 4 women, age range 21–25 yr), participated in the experiment. No subject reported auditory, vestibular, or sensory-motor impairments. All signed an informed consent protocol.

Experimental Setup

The experiments were performed in a sound attenuated chamber 2 × 2 × 2.2 m (IAC Industrial Acoustics, Niederkrüchten, Germany) covered with 20-cm foam wedges. An item Profile array (item Industrietechnik, Solingen, Germany) with 9 loudspeakers (Canton Plus XS.2; Canton Electronic, Weilrod, Germany) was mounted horizontally along one wall at a height of 1.7 m. The speakers were termed 1 through 9 with angles ranging from –40° (left) to 40° (right) in 10° steps (Fig. 1). A rotating chair (ekida, Buggingen, Germany) was positioned 97 cm from the center speaker. One additional loudspeaker was mounted to the ceiling directly above the chair. Also mounted on the ceiling was the tracking system (IS-900 System; Intersense, Billerica, MA), which tracked both chair- and head-fixed markers at a sampling rate of 200 Hz. The experimenter sat outside the chamber and monitored the procedure via computer control and video from a camera mounted inside the chamber.

Audio signals were computer generated (MATLAB 2007b; The MathWorks, Natick, MA) and sent out by digital-analog converters (MOTU 24IO audio interface; MOTU, Cambridge, MA) that were connected to the computer via a PCI card (MOTU PCI-424; MOTU). The sound presentation was controlled using the SoundMexPro toolkit (HörTech, Oldenburg, Germany). Signals were passed to amplifiers (NAD C9120; NAD Electronics International, Ontario, Canada) and then to the speakers. The tracking system sent its acquired data to the computer over a serial port. An additional PCI card (PCI-2517; Measurement Computing, Norton, MA) controlled the rotating chair. The script controlling the experimental procedure was written in MATLAB 2007b with an incorporated MEX file that linked to the tracking system and the PCI card responsible for the chair rotation.

Procedure

In a two-interval forced choice (2-IFC) paradigm, human listeners evaluated whether a probe signal was to the left or right of a preceding reference signal in world coordinates. Subjects were blindfolded to prevent the use of visual cues to self-motion. Before a trial could begin, the listeners had to ensure that their head was oriented toward the center speaker. Head orientation was tracked and a pip tone confirmed when an acceptable head orientation had been achieved (±1° tolerance). Following this tone, the subject could initiate a trial by pressing a joystick button (Logitech, Lausanne, Switzerland). After the start of a trial, the first (reference) noise burst was presented over the central speaker (Fig. 1A). Depending on the rotation condition (see below), subjects then performed an active head rotation, were passively rotated by the chair, or did not rotate (Fig. 1B). Finally, the second (probe) burst was presented over one of the nine speakers (Fig. 1C). The subjects then indicated whether the probe burst was perceived to the left or to the right of the reference burst by pressing the appropriate joystick button.

![Diagram](https://example.com/diagram.png)

Fig. 1. Schematic of the experimental setup and procedure. Initially, listeners always face the central speaker (A) over which the reference signal is presented (A). The listeners have to perform 1 of 4 different rotation conditions (D–G) between the 2 signal presentations (B). The probe signal is presented over 1 of the 9 speakers (C), after which the listeners decide whether it is left or right of the reference signal.
The two noise bursts consisted of 0.1 s of pink noise with a sampling rate of 48 kHz. The reference burst was always played over loudspeaker 5 (center speaker, 0° azimuth), and the probe burst was randomly presented over 1 of the 9 speakers. The interval between the two bursts was 2.2 s. This duration was required because average duration of head movement was 1.8 s. In addition to speaker-individual equalization (using 512-point FIR equalization filters), the bursts were manipulated individually from presentation to presentation in their frequency content by a roving spectral filter (10th-order FIR filter with 17 magnitude points roved with a standard deviation of ±5 dB and equally spaced on a log-frequency axis between 80 Hz and 20 kHz) to prevent listeners from recognizing residual timbre cues from individual speakers.

\[ \text{Equation} \]

**Rotation Conditions**

The rotation conditions manipulated the availability of vestibular and proprioceptive/efference copy signals (Fig. 1, D–G). We treat proprioception and efference copy as one signal because our methods do not allow us to distinguish their separate effects. The conditions were as follows.

- **No-rotation condition.** Listeners remained positioned with their head directed toward the central speaker. This condition provided a measure of baseline performance.
- **Active condition.** Listeners actively rotated their head ~35° to either the left or right following a previously trained rotation profile (see below). Vestibular and proprioceptive/efference copy signals were all available and signaled the same degree of head rotation.
- **Passive condition.** Listeners did not rotate their head, but the chair was rotated ~32° to the left or right following a rotation profile that was comparable to the listeners’ trained head rotation in the Active condition. Note that proprioceptive and efference copy signals were still present in this condition. Indeed, muscle contractions were necessary to overcome the inertia of the head and keep the head stationary on the moving body; we assume these signals indicated zero net head-on-body motion.
- **Cancellation condition.** Listeners rotated their head to either the left or right as in the Active condition, but head motion was tracked and used to counter-rotate the chair so as to cancel the active movement and stabilize the head in space. This resulted in a body-under-head rotation with the listeners’ head orientation remaining directed toward the central speaker. Subjects were not informed about how these compensatory chair rotations were performed. There was a latency of about 60 ms between the onset of head movement and the onset of the counter-rotation of the chair. Therefore, head movement was not canceled completely; vestibular signals were still present, but taken over the entire movement, they indicated zero net motion (Fig. 2, A and C, magenta traces).

In total, there were seven combinations of condition and rotation direction: No Rotation, Active-left, Active-right, Passive-left, Passive-right, Cancellation-left, and Cancellation-right. For each of these, a total of 270 trials were collected, 30 repetitions for each of the 9 speaker positions, so the total number of trials per subject was 1,890. These trials were divided into blocks of 126 trials, each of which could be completed in ~20 min. Total testing time for the 15 blocks (plus initial training, see below) was about 7 h per subject. Trials for all conditions were presented interleaved within a given experimental block in a predefined random sequence. Before each trial, subjects were informed about what kind of movement was required with an acoustic command from a speaker directly above the rotating chair. The acoustic commands were “no move” (for the No Rotation condition), “chair left/right” (for the Passive condition), or “head left/right” (for the Active and Cancellation conditions). Thus subjects did not know whether the chair would try to cancel the active head rotation after a “head left/right” command.

**Rotation Training and Head Rotation Analysis**

To ensure comparable results, all subjects underwent rigorous training concerning the active head rotations that they had to perform in the Active and Cancellation conditions. Small plastic balls (4-cm diameter) were positioned to contact the jaw/cheek when the head had been rotated to ±30°, and these remained present during the experiment. Note that these balls may have provided an additional tactile cue to head-on-body position, which would have supplemented proprioceptive and efference copy signals.

During training, the experimenter informed the subject if the movement did not match the template trajectory. The main experiment could begin only after subjects had learned to reliably reproduce head rotations with the required displacement and velocity; all subjects were able to learn the desired head rotation trajectories. During the main data acquisition, trials were excluded if they did not meet a nested set of criteria that quantified deviations of the executed rotation in that trial from the desired rotation. These trials were repeated again at a later time until at least 30 trials per condition were obtained. For all subjects and conditions, 3.43% of trials were rejected, with the

Fig. 2. Average head rotation trajectories for all subjects and conditions. Plotted are the mean head rotations for each subject color coded according to the experimental condition (Act, Active condition, green; Pass, Passive condition, blue; Canc, Cancellation condition, magenta). A and B show the head position in degrees along the time axis in seconds; C and D show the velocity in degrees per s. 

\[ \text{Equation} \]
maximum rejections for an individual subject being 9.53%. Average head rotation trajectories for all subjects and conditions are shown in Fig. 2.

Psychophysical Analysis

For each subject, condition, and rotation direction, data were expressed as a percentage of rightward responses for each speaker position and cumulative Gaussian psychometric functions were fit to the 270 trials using the Palamedes toolbox, PAL_PFML_fit (Palamedes toolbox MATLAB 2011a; Prins and Kingdom 2009) (Fig. 3). We refer to the mean parameter of the Gaussian fit as the point of subjective equality (PSE). The PSE indicates the probe speaker position perceived equal to the reference speaker position (i.e., probe position that elicits 50% rightward responses). Nonzero values of the PSE reflect a bias in auditory spatial updating. We refer to the standard deviation parameter of the fit as sigma (σ). It is inversely related to the slope of the function and provides a measure of the variability in auditory spatial updating.

Control Experiment

To test for the influence of the vestibular and proprioceptive inputs on spatial updating of the world-centered position independent of auditory stimuli, we designed a control experiment. Blindfolded subjects were passively rotated in total darkness via the rotating chair with the same rotation profile employed for the main experiment. The chair rotation track was scaled by factors of 0.67, 1, and 1.33, with the same rotation profile employed for the main experiment. Before chair rotation, the subject’s head orientation was monitored via the tracking system to ensure the head was oriented straight ahead (±1° tolerance), i.e., that it was facing the central loudspeaker. A pip tone confirmed an acceptable head position had been achieved. After the passive rotation, the subjects were asked to rotate only their head back to the starting position and confirm their end position via the joystick by pressing a button; we refer to this as “back rotation.” For each subject, a total of 60 trails were collected for each rotation direction, 30 trails with gain of 1, and 15 trials each with gain of 0.67 and 1.33. Trials were presented in a predefined random sequence.

As in the main experiment, trials were rejected according to defined criteria when subjects rotated their head in a manner that was not consistent with the instructions. First, we checked for a sufficiently good match between chair- and head-tracking data during the passive first interval, to ensure that subjects did not rotate the head relative to the body while they were rotated by the chair. Second, we checked that the subject did not rotate their head between the end of the first passive interval and the beginning of the second active interval. If the difference between the end point of passive rotation and the beginning of the active back rotation was greater than 2°, the trial was rejected. Finally, we checked that the entire active back rotation covered at least 5°. This guaranteed that an active back rotation had taken place. These strict criteria led to a large percentage of rejected trials, 40% on average across subjects, with the maximum rejections for an individual subject being 54%.

Modeling

The purpose of the four experimental conditions is to tease apart the relative influence of vestibular and proprioceptive signals on auditory spatial updating. The modeling presented below explains the framework used to derive these relative influences on the basis of the PSEs measured in each condition.

At the PSE the actual auditory eccentricity (A) is equal to the head rotation (H) minus the PSE (Fig. 4A; A = H − PSE). Despite any discrepancies in the actual values, these quantities are perceived to be equal (Å = H, thus the term PSE). We relate perceived to actual stimulus values by multiplicative gain factors (Å = gA,H; H = gH,H). From the PSE in each condition, it is therefore possible to calculate the ratio of these perceptual gains (Eq. 1):

\[
g_a = \frac{H - PSE}{H}.
\]

This gain ratio will vary across conditions, but if we assume that \( g_a \) is constant, comparison of gain ratios across conditions reveals how perceived head rotation depends on the available head rotation signals. More generally, the gain ratio constitutes an updating index that expresses the accuracy of updating performance (PSE) relative to the magnitude of head movement (H).

To quantify the relative influence of vestibular and proprioceptive/effference copy signals, we model the perceived head rotation as a weighted linear combination of these two unbiased signals with weights that sum to 1:

\[
g_v H = w_H V + w_p P
\]

(2)

\[
w_H + w_p = 1
\]

(3)

In the Active condition, vestibular (V) and proprioceptive/effference copy (P) signals are equal to the total head rotation (V = P = H). However, based on results from the Passive and Cancellation conditions, it is possible to solve for the vestibular (and proprioceptive) weights because proprioceptive and vestibular signals, respectively, are set to zero:

\[
w_{H_P} = \frac{g_{v\text{Pass}}}{g_{H\text{Act}}}
\]

(4)

Fig. 3. Example data from a single subject. Depicted is the proportion of rightward responses as a function of the azimuth angle in degree of the probe speaker position for the 4 conditions (−40° is the leftmost position, +40° is the rightmost position). The PSE and σ for the Passive condition are indicated. NoRo, No-Rotation condition; Gaussfit, Gaussian fit.

Fig. 4. Illustration of the gain ratios. A shows the auditory eccentricity (Å; magenta) after a head rotation (H; green) to the right. B illustrates the inequality when an active head rotation (H act; green) compensates for a previous passive head rotation (H pass; blue).

\[
\frac{g_v}{g_a} = \frac{H - PSE}{H}.
\]

(1)
In this way, performance in the Passive and Cancellation conditions (numerator) is essentially normalized relative to performance in the Active condition (denominator) when all signals are available.

RESULTS

The current experiments quantified observers’ ability to correctly update the allocentric position of a sound source after a head rotation of \(\sim 30^\circ\). Auditory stimuli are sensed in head coordinates, but the task required judgments to be made in world coordinates. This depends on the nervous system’s ability to monitor head rotation in space and thereby maintain an accurate mapping between coordinates systems. Comparison across conditions reveals the relative roles of vestibular and proprioceptive/efference copy signals for monitoring head rotation in space.

Updating Accuracy and Weights

Example psychometric functions are shown in Fig. 3. For each subject and condition we extracted the PSE (mean) and \(\sigma\) (SD) values of the cumulative Gaussian psychometric fits, which quantify the accuracy and precision of updating performance, respectively. Updating accuracy for all subjects and conditions is summarized in the top row of Fig. 5. The gain ratios plotted constitute an updating index calculated according to Eq. 1 using the head movements and PSE values reported in Table 1 and 2, respectively.

Gain ratios were significantly greater than 1 following normal active head movements (Active condition; \(t\)-test diff from 1, \(P = 0.0016\)), indicating that auditory targets had to be shifted in the direction opposite the head turn to be perceived to be at the same position as the original central reference target. Values were significantly less than 1 in the Cancellation condition (\(t\)-test diff from 1, \(P = 0\)), consistent with a reduced gain on perceived head rotation, when vestibular stimulation was reduced. Values in the Passive condition were not significantly different from 1 (\(t\)-test diff from 1, \(P = 0.3811\)), consistent with accurate updating performance. Average values in the Passive condition were reduced compared with the Active condition, but the difference was not significant (paired \(t\)-test, \(P = 0.6534\)) due to high across-subject variability in the Passive condition.

Given the previously outlined assumptions underlying the weighted linear combination model (Eqs. 2 and 3), it is possible to calculate the relative weights given to vestibular and proprioceptive/efference copy signals for monitoring head rotation in space.

Table 1. Average head rotation for each subject and condition

<table>
<thead>
<tr>
<th>Subject</th>
<th>Active</th>
<th>Passive</th>
<th>Cancellation</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>35.4</td>
<td>32.0</td>
<td>37.5</td>
</tr>
<tr>
<td>2</td>
<td>35.3</td>
<td>32.9</td>
<td>35.9</td>
</tr>
<tr>
<td>3</td>
<td>29.3</td>
<td>32.5</td>
<td>29.7</td>
</tr>
<tr>
<td>4</td>
<td>33.6</td>
<td>33.0</td>
<td>34.6</td>
</tr>
<tr>
<td>5</td>
<td>38.7</td>
<td>33.1</td>
<td>41.9</td>
</tr>
<tr>
<td>6</td>
<td>37.7</td>
<td>33.5</td>
<td>37.5</td>
</tr>
<tr>
<td>7</td>
<td>31.8</td>
<td>32.4</td>
<td>33.2</td>
</tr>
<tr>
<td>8</td>
<td>32.6</td>
<td>32.9</td>
<td>37.3</td>
</tr>
<tr>
<td>9</td>
<td>39.0</td>
<td>32.6</td>
<td>39.4</td>
</tr>
<tr>
<td>10</td>
<td>38.4</td>
<td>32.7</td>
<td>38.7</td>
</tr>
<tr>
<td>Mean</td>
<td>35.3</td>
<td>32.8</td>
<td>37.4</td>
</tr>
</tbody>
</table>

Values are average head rotation (H) in degrees for each subject and condition. For the Passive condition, \(H_{\text{Pass}}\) is equal to the passive chair rotation. For the Active condition, \(H_{\text{Act}}\) is equal to the average magnitude of the voluntary head rotation for a given subject. For the Cancellation condition, \(H_{\text{Canc}}\) is equal to the average magnitude of the voluntary head-on-body rotation.
propiocceptive/efference copy signals in estimating head rotation in the Passive and Cancellation conditions (Eqs. 4 and 5). These weights are plotted and compared in Fig. 6. Weights less than 1 in Passive condition suggest a possible role for proprioception or efference copy signals even during passive rotation. Similarly, vestibular weights less than 1 in the Cancellation condition suggest that all cues play a role, with vestibular cues being weighted most heavily. Weights were significantly smaller than 1 in the Passive condition (P = 0.0103), but not in the Passive condition due to high across-subject variability (P = 0.5457).

Vestibular weights calculated from the Passive and Cancellation conditions are compared in Fig. 6C. Vestibular cues were given significantly more weight in the Passive compared with the Cancellation condition (paired t-test, P = 0.0339), as might be expected for passive compared with active head rotation. Nevertheless, the significant across-listener correlation between weights calculated in the two conditions (P = 0.0321) indicates that, despite the difference in weights, the same or similar processes are at work in the two conditions.

### Updating Precision

Psychometric functions provide a measure of the variability of updating performance across the conditions through the standard deviation parameter (σ) of the cumulative Gaussian fit. Updating variability for all subjects and conditions is shown in the bottom row of Fig. 5. Variability was least in the No-Rotation condition. This reflects underlying auditory discrimination ability because there was no intervening head rotation, and therefore no need for coordinate transformation and no additional noise due to head rotation signals. Increased variability was observed in the other conditions. Variability was greatest in the Passive condition, significantly greater than in either the Active (t-test, P = 0.0032) or Cancellation (t-test, P = 0.0006) conditions, suggesting greater variability on vestibular compared with proprioceptive and/or efference copy head rotation signals.

### Control Experiment

Given the differences observed between the Active and Passive conditions, we ran a control experiment to explore the hypothesis that these differences are due to differences in head rotation estimation only, and not to differences in the readout of head-centered, binaural spatial auditory cues. In this control condition, subjects experienced passive rotations and were asked to actively rotate their head back to the initial position. Average head rotation responses as a function of passive rotation angle are plotted in Fig. 7A. Head rotation responses were significantly less than required to reach the initial position (paired t-test, passive vs. return, P < 0.001), suggesting that the perceptual gain on active head rotation is greater than that on passive head rotation.

The finding of the control experiment is consistent with results of the main experiment, which found a greater gain on active compared with passive head rotation (g_{return} > g_{pass}), as expected for passive compared with active head rotation. Nevertheless, the significant across-listener correlation between weights calculated in the two conditions (P = 0.0321) indicates that, despite the difference in weights, the same or similar processes are at work in the two conditions.
DISCUSSION

We have examined auditory spatial updating in the horizontal plane in response to yaw head rotations and observed the following: 1) the gain on auditory updating during active head rotations is greater than unity (Active condition), meaning that auditory targets must be displaced in the direction opposite the head rotation to be perceived world-fixed; 2) gain is reduced during passive head rotations, when only vestibular signals indicate rotation (Passive condition); and 3) gain is nonzero if the trunk is rotated under the stationary head to stabilize the head in space such that only proprioceptive/efference copy signals indicate rotation (Cancellation condition). These last two findings are interpreted to reveal that both vestibular and proprioceptive/efference copy signals contribute to auditory spatial updating even if one or the other signal is set to zero. In the following discussion, we first relate these findings to prior behavioral results concerning both visual and auditory space perception and then speculate about the neurophysiological underpinnings of these behaviors.

Perceptual Measurement of Auditory Spatial Updating

Spatial updating research is essentially concerned with understanding the transformations between eye, head, body, and world coordinates that enable spatial constancy (Klier and Angelaki 2008; Medendorp 2011). To achieve a better understanding, it is helpful to exclude effects due to one or the other reference frame. Two distinctive features of the current study are 1) the use of purely auditory stimuli and 2) the use of a 2-IFC task rather than pointing responses (e.g., saccades) to investigate spatial updating. Auditory stimuli are sensed directly in head coordinates, so the transformation from retinal to head coordinates needed for visual targets is unnecessary. Additionally, our 2-IFC task does not require for the spatial representation of the auditory target to be remapped into a pointing response. In this way our design excludes the eye-centered reference frame in the stimulus domain and also excludes remapping errors in the response domain.

Relative Contribution of Vestibular, and Proprioceptive/Efference Copy Signals

When prior studies have directly examined separate contributions of vestibular and proprioceptive/efference copy signals, results have been similar to ours (Mergner et al. 1998, 2001; Nakamura and Bronstein 1995), even though targets were visual and responses were saccadic. Compensation for head rotation is greatest when both signals are available (e.g., Active condition) and reduced when either one or the other signal is available in isolation (e.g., Passive and Cancellation conditions). However, our interpretation differs from that of previous studies. For example, Mergner et al. (1991, 2001) propose that head rotation is underestimated in vestibular and proprioception/efference copy conditions because head rotation in space is referred to an erroneous estimate of trunk rotation. In contrast, we propose a model that assumes that the trunk is stationary in space, roughly analogous to stationarity assumptions proposed elsewhere (Wexler et al. 2001). With this assumption, the proprioceptive/efference copy signals can be taken as measures of head rotation in space.

We model estimation of head rotation as a simple weighted sum of proprioceptive/efference copy and vestibular signals (Eqs. 2 and 3), and we calculate the weights that best explain our results. Notice that this approach amounts to assuming forced fusion [proposed previously for visual-vestibular (de Winkel et al. 2015; Prsa et al. 2012) and vestibular-proprioceptive (Frissen et al. 2011) cue integration] in that gross conflicts across signals are ignored (i.e., when one or the other signal is set to 0). An additional assumption of the model is that vestibular and proprioceptive estimates (V and P in Eq. 3) are unbiased; in other words, these signals provide veridical information about the head movement. Consequently, the model cannot distinguish between effects due to weighting and those due to biased signals. The model must be further tested, but even if the assumptions are not valid, the weights provide a convenient method to quantify and compare effect size across conditions.

Weights were correlated across conditions (Fig. 6C), suggesting similar underlying processes, but vestibular weights were significantly higher in the Passive condition. This could reflect proprioceptive/efference copy downweighting to account for the passive nature of the rotation. Alternatively, weights could be fixed across conditions, and the observed mismatch could be an experimental artifact of incomplete nulling of the head rotation in the Cancellation condition (see Fig. 2, A and C). Even though head orientation returns to zero by the end of each trial, this is accomplished via an initial period of undercompensation, during which the head moved...
one way, followed by a catch-up period in which the head moved in the opposite direction. This nonmonotonic, out and back oscillation may lead to an erroneous nonzero vestibular displacement estimate that could be manifest as an increase in estimated proprioceptive weight in the model. In future it should be possible to design a control algorithm to cancel head motion more effectively.

It is additionally instructive to compare variability across conditions (Fig. 5, bottom row). Variability was least in the Active and Cancellation conditions, suggesting that proprioceptive/efference copy signals are more precise when the head is moved on the body than when it is stationary (Passive condition). This result also suggests that vestibular signals are typically more variable than proprioceptive/efference copy signals. Nevertheless, we observed a vestibular weight of ~0.85 for this presumably unreliable signal, a finding that is at odds with reliability-based weighting models of cue integration. One rationale for assigning a high vestibular weight despite the variability of this signal is that the vestibular cue provides a measure of head motion in space; use of efference copy and proprioception depends on the assumption that the body is fixed in space. High vestibular weight may reflect uncertainty associated with this assumption, and an associated downweighting of these cues. Further research is needed to examine relative cue reliability and weighting when conflicts across modalities are more moderate.

Further research is also needed to investigate the relative contributions of proprioceptive and efference copy signals. These signals covary during normal active head movements, such as those in our experiments. Our study also included a possible tactile cue to head-on-body position provided by the plastic balls used for head-movement training (see MATERIALS AND METHODS). We are not able to tease apart separate effects of these various head-on-body cues. Some prior studies have attempted to dissociate proprioceptive and efference copy cues, either by rotating the head passively on the body, leading to proprioceptive but not efference copy signals, or by blocking active head movements, leading to efference copy but no proprioception (Crowell et al. 1998; Cullen et al. 2011; Nakamura and Bronstein 1995). These studies conclude that both signals play a significant role. Such manipulations would also be possible in future auditory updating experiments.

**Inaccuracy in Auditory Spatial Updating During Active Head Rotation**

Inaccurate updating in the Active condition (Fig. 5A) was unexpected. Nevertheless, this performance (gain ratio approximately equal to 1.1) is roughly consistent with prior reports of bias in auditory localization, which is roughly 10% of the head rotation magnitude (Lewald et al. 2000). Similar findings are reported by Kopinska and Harris (2003) for both visual and auditory localization during active head turns. Because the effect transfers across modalities, they attribute it to an overestimation of perceived head turn angle. In other words, the gain on perceived head motion, \( g_s \), which is the numerator of the gain ratio (Eq. 1), is large relative to the denominator. This agrees with reports of overestimation of active head-turn angle assessed with a variety of methods (Becker and Saglam 2001). Also consistent with this view, results of the present control experiment found that degree of overestimation during active relative to passive head rotation was similar even when measured with a nonauditory task (Fig. 7C). Because the overestimation is not observed during passive rotation, it likely results from proprioceptive and/or efference copy input.

An alternative explanation is that the gain ratio greater than unity in the Active condition results from a reduced gain on perceived auditory eccentricity, \( g_a \), the denominator of the gain ratio (Eq. 1). A similar reduced gain on perceived auditory motion was reported recently (Freeman et al. 2015) and is hypothesized to reflect a Bayesian prior for slow or stationary auditory targets. An analogous prior for straight ahead auditory targets would lead to underestimation of auditory eccentricity in the current experiments. However, such a prior on auditory eccentricity cannot explain overestimation of head-turn angle observed in the control experiment or differences in performance between active and passive conditions.

A possible explanation for the measured updating errors might be distortion and compression of auditory space or biases due to representational momentum during head turns (Cooper et al. 2008; Leung et al. 2008; Feinkohl et al. 2014). Leung et al. (2008) state that for compressive spatial errors of auditory or visual space during head turns to be revealed, high head velocities of at least 200°/s are needed. In our study, maximal head velocities were around 50°/s (see Fig. 2, bottom). Similar head velocities were used in a study by Jackson et al. (2005), where spatial and temporal errors during saccade head movements were not found. Cooper et al. (2008) measured distortions of auditory space during rapid head motions, where the acoustic target was presented during certain head-turning points. They found that distortions are larger for acoustic targets presented toward the end of a turn and maximal for targets located around the head-centered midline. Again, much higher head velocities were used (mean of 124°/s), and furthermore, the angular head displacements of ±60° were twice as large as the displacements we used in our study (+/−30°). As reported by Feinkohl et al. (2014), the auditory representational momentum is affected in a linear fashion by the velocity of moving targets but is mainly biased toward the eye position. Because the subjects in our study were blindfolded and did not receive a visual target, we can exclude a localization bias toward a visual target position. Furthermore, it is important to note that the velocities the vestibular system experienced during the Active and Passive conditions are very similar because subjects were trained to perform movements that matched the passive rotation trajectory. Therefore, if the measured updating errors were only due to compression of auditory space or effects on representational momentum, the errors in both conditions should be comparable, which is not the case (see example psychometric response function in Fig. 3).

Another consideration is that differences in performance across Passive and Active conditions could be influenced by differences in eye movements accompanying active and passive head movement. During Active conditions, subjects may have turned the eyes in the direction of the head turn, as during an eye-head gaze shift. If so, these eccentric eye-in-head positions could have impacted auditory localization of the second auditory target. However, we consider this to be an unlikely explanation. The present experiments were conducted with blindfolded subjects, so eye movements are not functional as they are during a typical eye-head gaze shift to an eccentric
visual target. Thus we would expect that by the end of the 2.2-s interstimulus interval, subjects would have adopted a neutral eye position in all conditions; however, we did not record eye movements, so we cannot say for sure. Even if subjects maintained eccentric eye positions following active head movement, it would not be sufficient to explain our results because the effect of eye position on auditory localization is small (3° for 45° eye turn) compared with differences in auditory localization observed across conditions in the present study (8° difference between Active and Passive conditions for 30° head turn); the small effect of eye position is further reduced in complete darkness (Lewald and Ehrenstein 1996).

Although larger effects of eye movements have been reported (Razavi et al. 2007), these effects are typically observed only after an eccentric eye position is maintained for an extended period of time. Thus the effect of eye position signals on the current results is likely negligible.

More generally, direct neurophysiological evidence for efference copy processes in auditory spatial updating has not been presented, so far. Efference copies have been shown to play a role in auditory processing during Doppler-shift compensation in bats (Metzner 1989, 1993) and also during vocal learning in songbirds (e.g., Prather et al. 2008). That such a process is very likely to be an underlying mechanism for spatial updating of acoustic input during ear movements has been proposed for bats, where a map-like representation of pinna position has been described in the superior colliculus (Valentine et al. 2002). It can be assumed that this might be a general feature of all mammals with large moveable pinnae (Kanold and Young 2001; Kanold et al. 2011). An efference copy mechanism like that would resemble findings for the visual system, where spatial receptive fields in the frontal eye field are shifted toward the direction of saccades due to corollary discharges provided by superior colliculus via the medial dorsal nucleus of the thalamus (Sommer and Wurtz 2002). For the vestibular system and proprioception, corollary discharges have been demonstrated to play a role in, for example, the suppression of undesired reflexes (Roy and Cullen 2004; Seki et al. 2003; Voss et al. 2006). It is well conceivable that these neural signals might also be used for the purpose of spatial updating of sound source position during head and body movements.

Auditory vs. Multimodal Spatial Representations: Behavioral and Neurophysiological Evidence

Initial spatial processing of visual and auditory signals is obviously different. For example, although retinal signals convey spatial information, peripheral auditory signals must be combined centrally to extract spatial information. However, prior behavioral studies suggest that ultimately visual and auditory updating depend on common principles or common spatial representations. For example, Lackner and DiZio (2010) showed that localization of both auditory and visual targets are similarly affected by unusual gravitoinertial acceleration stimuli that lead to illusions of perceived body orientation. Similarly, Klatky et al. (2003) reported similar patterns of error in updating of both visual and auditory targets. Likewise, as discussed above, our present results are comparable to those obtained when targets are visual and responses are based on saccades (Mergner et al. 1998, 2001; Nakamura and Bronstein 1995). On these grounds, it appears that spatial updating processes generalize across both visual and auditory modalities and across both perceptual and motor responses.

This could be because visual and auditory spatial representations are separately updated but in similar ways. For auditory processing, early convergence occurs at the first synaptic levels in the mammalian brain stem: anatomical studies in guinea pigs (Burian and Gstoettner 1988; Burian et al. 1989, 1991; Gstoettner et al. 1991), gerbils (Kevetter and Perachio 1986, 1989), rats (Barker et al. 2012), and macaque monkeys (Newlands and Perachio 2003) have shown that the vestibular nerve and lateral vestibular nucleus project to the cochlear nucleus, the first station of the central auditory pathway. Proprioceptive/efference copy input is present already in the vestibular nuclei (Cullen and Roy 2004), as are visual motion signals (Waespe and Henn 1977), so vestibular pathways are likely to also convey proprioceptive and visual information. However, proprioceptive information also converges directly with auditory information early in the auditory pathway (reviewed in Shore and Zhou 2006). In particular, projections from the dorsal column convey head and trunk information and might be an important substrate underlying the findings of the current study. Similarly, visual sensory processing is likely influenced early in the stream of processing by vestibular, proprioceptive, and efference copy signals. For example, the lateral geniculate nucleus, the first central stage of visual processing, is known to receive projections from the vestibular nuclei (Magnin and Putkonen 1978).

Alternatively, a common spatial representation could be used to drive both auditory and visual responses. The superior olivary complex and the external cortex of the inferior colliculus are likely candidates as they receive not only auditory and visual inputs but also proprioceptive information (Shore et al. 2000; Shore and Zhou 2006). This is of special importance, because auditory midbrain structures such as the brachium of the inferior colliculus and the inferior colliculus are known to be important for the formation of multisensory maps of space in the midbrain of mammals and birds (Knudsen and Konishi 1978; Schnupp and King 1997).

Conclusion

Overall, results complement former studies showing both vestibular and proprioceptive/efference copy signals play a significant role in spatial updating (e.g., Mergner et al. 1998; Nakamura and Bronstein 1995). Inaccuracy following active head movements appears to be a consequence of proprioceptive/efference copy signals that result in either an underestimation of perceived auditory eccentricity (Lewald et al. 1999) or an overestimation of perceived head turn (Becker and Saglam 2001; Kopinska and Harris 2003). Results of our Control condition support the later conclusion. We observed accurate updating performance following passive rotations, in contrast with prior reports (Van Barneveld et al. 2011; Yost et al. 2015); these differences are likely due to the different methodologies employed. In future it will be important to investigate factors that determine weights given to vestibular vs. proprioceptive and efference copy signals, preferably using paradigms in which conflicts among signals are not so pronounced, and also to investigate the physiological underpinnings of these interactions. In general, spatial updating during
head movement is an excellent model system for examining the role of motor signals in reference frame transformations during sensory processing.

ACKNOWLEDGMENTS

We thank Humberto Alvarez and Lisa Rohrbacher for their assistance in technical development and data acquisition, respectively.

GRANTS

P. MacNeilage was supported by Bundesministerium fur Bildung und Forschung Grant 01EO1401. The research was funded by a transfer-project grant from the Bernstein Center for Computational Neuroscience Munich.

DISCLOSURES

No conflicts of interest, financial or otherwise, are declared by the authors.

AUTHOR CONTRIBUTIONS

D.G., U.F., L.W., and P.M. interpreted results of experiments; D.G., L.W., and P.M. analyzed data; D.G., L.W., and P.M. approved final version of manuscript.

REFERENCES


