Spatial Localization of Auditory Stimuli in Human Auditory Cortex is Based on Both Head-Independent and Head-Centered Coordinate Systems

Eitan Schechtman,¹ Talia Shrem,² and Leon Y. Deouell^{1,2}

¹Edmond and Lily Safra Center for Brain Sciences, and ²Department of Psychology, The Hebrew University of Jerusalem, Jerusalem 91905, Israel

In humans, whose ears are fixed on the head, auditory stimuli are initially registered in space relative to the head. Eventually, locations of sound sources need to be encoded also relative to the body, or in absolute allocentric space, to allow orientation toward the sounds sources and consequent action. We can therefore distinguish between two spatial representation systems: a basic head-centered coordinate system and a more complex head-independent system. In an ERP experiment, we attempted to reveal which of these two coordinate systems is represented in the human auditory cortex. We dissociated the two systems using the mismatch negativity (MMN), a well studied EEG effect evoked by acoustic deviations. Contrary to previous findings suggesting that only primary head-related information is present at this early stage of processing, we observed significant MMN effects for both head-independent and head-centered deviant stimuli. Our findings thus reveal that both primary head-related and secondary body- or world-related reference frames are represented at this stage of auditory processing.

Introduction

Sounds reach the ears from any direction, only mildly obstructed by intervening objects. Hence, the auditory system is optimal for monitoring the environment and orienting other senses toward salient events (Arnott and Alain, 2011). As auditory spatial information is extracted from the comparison between left- and right-ear signals (for azimuth) and by the spectral filtering properties of the fixed pinnae (for elevation) (Blauert, 1997), initial encoding of space in humans must be anchored to the head [head-centered (HC)]. Yet behavioral studies have shown that the human nervous system eventually compensates for head motion, producing a veridical head-independent (HI) stable representation of auditory sources, which allows proper direction of motor responses, including gaze shifts (Goossens and van Opstal, 1999; Vliegen et al., 2004; Van Grootel et al., 2011). Is this representation a late, higher-order/cognitive effect or does it emerge early in auditory processing?

Two contrasting answers were recently reported in this journal. Van Grootel et al. (2011) modeled the HC and HI frameworks using patterns of gaze shifts. Elegantly using the fact that pure tones' perceived elevation is unrelated to the true elevation, their results reaffirmed the emergence of an HI representation, and moreover, localized it to tonotopic, likely

brainstem, regions. These results stand in striking contrast to those of a recent EEG study that failed to find evidence for HI (allocentric) representation at the level of auditory cortex (Altmann et al., 2009).

Altmann et al.'s (2009) subjects listened, through headphones, to a sequence of sounds virtually located straight ahead, and then rotated their heads by 30°. This was followed by another sound, which was either located in the exact same position in absolute space as the previous sounds, yet was now at an angle to the new head orientation (egocentric deviant), or was relocated by 30° along with the head, so that it was in the same position relative to the new head orientation, but now deviated in its absolute location (allocentric deviant). Only the egocentric HC deviant elicited the mismatch negativity (MMN) event-related potential, which indexes neural detection of deviations in the acoustic environment. Thus, the conclusion was that only head-centered representations were present at the auditory cortex probed by MMN.

We surmised that the absence of MMN to allocentric, head-independent deviance was due to specific features of Altmann et al.'s (2009) study, mainly the fact that the sounds were produced through headphones rather than in real space. Using nonindividualized virtual auditory space produced over headphones, sounds may appear lateralized, but within the head rather than in external space (Plenge, 1974; Møller et al., 1996; Hammershøi and Møller, 2002). This would naturally mitigate the effect of an HI reference frame. We thus used Altmann et al.'s (2009) essential design, using sounds elicited by loudspeakers in free field around the subjects. As described below, several other changes to the design were also introduced to control for alternative interpretations. We found that both HC and HI spatial changes elicit the MMN, resolving the discrepancy.

Received March 16, 2012; revised June 18, 2012; accepted July 23, 2012.

Author contributions: E.S. and L.Y.D. designed research; E.S. and T.S. performed research; E.S. and L.Y.D. analyzed data; E.S. and L.Y.D. wrote the paper.

This work was supported by Grant 823/2008 from the Israel Science Foundation to L.Y.D.

The authors declare no financial conflicts of interest.

Correspondence should be addressed to Prof. Leon Y. Deouell, Department of Psychology, The Hebrew University of Jerusalem, Jerusalem 91905, Israel. E-mail: leon.deouell@huji.ac.il.

DOI:10.1523/JNEUROSCI.1315-12.2012

Copyright © 2012 the authors 0270-6474/12/3213501-09\$15.00/0

Figure 1. Experimental design. *a*, Example of the timeline for a single trial. Note that although the figure describes a trial with three standard stimuli (Stim), half of the trials included four standard stimuli. *b*, Illustration of the setup and three stimuli. Standard stimuli: the eyes and head are oriented toward an LED straight ahead (gray dot, labeled CL). The sound is emitted by the speaker 30° to the right of the midline (RS). HC-deviant: the head and eyes point toward a LED 30° to the left of the midline (LL), the sound is emitted by RS. HI-deviant: the head and eyes point toward an LED 30° to the left of the midline, the sound is emitted by the CS. The insets show the top view of a single subject's head during the standard (bottom) and deviant (top) stimuli, as captured by the overhead video camera during the experiment.

Materials and Methods

Subjects. The participants in the main experiment were 17 students from the Hebrew University of Jerusalem (10 male; 16 right-handed) aged 19–26 years (mean age, 23.1 years; SD, 2.18 years), with reportedly normal hearing and no history of neurological disorders. The students were paid or given course credits for participation in the study. The experiment was approved by the ethical committee of the faculty of social science at the Hebrew University of Jerusalem, and informed consent was obtained after the experimental procedures were explained. One participant's data were excluded from analysis due to excessive artifacts.

Stimuli and apparatus. Subjects sat in a comfortable reclining armchair in a dimly lit, sound-attenuated and echo-reduced chamber (C-26; Eckel). They were surrounded by three loudspeakers (model 821615 midrange 122M; Peerless) mounted on a semicircular bar with a fixed radius of 90 cm. One was placed at an angle of 30° to the right of the subjects' midsagittal plane [right speaker (RS)], another 30° to the left of the midsagittal plane [left speaker (LS)], and a third central speaker (CS) straight ahead of the subjects on their midsagittal plane. LS was not active in the experiment and was present for symmetry and for mounting the LED. Attached to the top of the speakers were three identical green LEDs (RL, LL, and CL, respectively; Fig. 1*b*). All the auditory stimuli used in this experiment were acoustically identical (in frequency, intensity, and duration) and differed only in the position in space from which they were produced. Stimuli were composed of frozen band-passed noise (250–4000 Hz) with duration of 50 ms.

Procedure. Throughout the experimental blocks, the subjects' only instructions were to look at the currently lit LED (only one was lit at each time) and to reorient their head to face the lit LED at all times, ignoring any auditory stimuli. They were asked to avoid reorienting their eyes with no head movement (i.e., making saccades) and to refrain from body motion and massive blinking. Being in a semi-reclined position, the armchair used in the experiment did not allow trunk reorientation in any simple way. A camera located above the subjects' head was used to monitor the head and body position during the experiment (Fig. 1).

The experiment included five identical blocks, each lasting ~10 min and 40 s, and including 80 trials. Short subject-paced breaks were provided between blocks. Each trial began with the activation of the CL and with subjects orienting their head to face it. This was followed by a sequence of three or four sounds from the RS ("standards"), with a fixed

stimulus onset asynchrony (SOA) of 1700 ms (Fig. 1*a*). In each sequence, 350 ms after the last standard (LastStd), the CL turned off and the LL turned on, and subjects turned their heads to face it. After 1350 ms, the deviant stimulus was produced from one of the speakers. In the head-independent deviant (HI-Dev) condition (Fig. 1*b*, right), this stimulus was produced from the CS. Thus, the sound's angle relative to the head was the same as in the standard sounds, but its position in space was different. In the head-centered deviant (HC-Dev) condition, this stimulus was produced from RS (Fig. 1*b*, left). That is, the sound's angle relative to the head was now 60° instead of 30°, but its absolute position in space was the same as for the standards. Three hundred milliseconds after deviant onset, LL turned off and CL turned on again, indicating the onset of a new trial. Each run consisted of 200 HC-Dev trials and 200 HI-Dev trials. Half of the deviants in each condition were preceded by three standards and half by four standards. All trial types were randomly mixed within blocks.

Note that this design follows the principles of Altmann et al.'s (2009), but it is not identical. In that study, in HI deviant trials [called "allocentric" by Altmann et al. (2009)], as in the standard trials, the visual and auditory stimuli were at the same position in space (straight ahead), whereas in an HC ("egocentric") trial, they were at different positions (the LED was now on the left while the sound remained central). Thus, the egocentric HC MMN could have reflected a cross-modal change, rather than a merely spatial change (Fort et al., 2002). Therefore, in the present study, the visual and auditory stimuli were always disparate to mitigate cross-modal deviation.

EEG acquisition and postprocessing. The EEG was recorded using an Active 2 system (Biosemi) from 64 preamplified electrodes mounted on an elastic cap according to the extended 10–20 system, with the addition of two mastoid electrodes and a nose electrode. Eye movements were recorded using two electrodes at the outer canthi of the right and left eyes and two above and below the center of the right eye. The EEG was continuously sampled at 256 Hz and stored for off-line analysis. All data were digitally filtered (zero-phase 24 dB/octave Butterworth filter) with a bandpass of 1–12 Hz. The data were referenced offline to the nose channel. Ocular artifacts (blinks and saccades) were removed using the independent component analysis method (Jung et al., 2000) implemented using Vision Analyzer 1.05 (Brain Products). Segments contaminated by other artifacts were discarded (rejection criteria: >100 mV absolute dif-

ference between samples within segments of 200 ms; absolute amplitude >100 mV). Only trials in which subjects moved their head on time after the LED cues to the correct orientation were analyzed. Head movement was detected in the EEG data based on head and eye movement artifacts [the EEG threshold was set as (right horizontal EOG) – (left horizontal EOG) >45 mV]. These criteria were verified using samples of each subject's overhead video recording (Fig. 1).

The EEG data were parsed into 500 ms segments starting 100 ms before stimulus onset. The voltage was measured relative to the mean of the 100 ms prestimulus period (baseline correction). To measure the response to the standard stimuli, we considered the EEG signal of the second to last (STL) standard stimulus, i.e., the third stimulus in four standard stimuli trials and the second stimulus in three standard stimuli trials. MMN amplitudes were measured by comparing the ERPs elicited by the STL standard stimulus in each trial to that elicited by each type of deviant stimulus (HI-Dev or HC-Dev). We report here the results using the STL stimuli and not the last standard (LastStd) stimuli to avoid the possible contamination by head motion (which always followed the LastStd) on the ERP waveform. However, we conducted all statistical analyses on both STL and LastStd data and found similarly significant results. Note that the STL (as well as the LastStd) was identical across conditions, and that the subject did not know at the time of its onset which deviant will ensue.

ERP statistical analysis. The MMN was calculated for each subject as the average amplitude of the difference wave between the responses to STL and the response to the deviant in each condition at electrode Fz, 160–220 ms after the stimulus onset. Fz was chosen a priori based on established MMN literature, and the time window was chosen to span the peak negativity between 100 and 200 ms, as observed in across-subject grand averaging across the HI-Dev and HC-Dev difference waves. To establish the presence of the MMN, MMNs for each condition were compared with zero using one-tailed *t* tests.

The above analysis, which is the standard in most neuroimaging studies, uses only one statistic (Student's *t*) per condition, considering only the variability between subjects' mean MMNs. It does not consider the variability between trials within a subject, which could be substantial, nor the number of artifact-free trials in every subject. Since establishing the presence of MMN in each condition is critical in this study, we further introduce a complementary statistical approach, which takes into account intrasubject, and not merely intersubject, variability. To do so, we used a method akin to a meta-analytic approach, where each subject is considered as an experiment. We first calculated a *p* value for each subject and for each condition using one-tailed paired *t* tests, which compared, within subject, each trial's STL amplitude with that trial's deviant amplitude (HI-Dev or HC-Dev), at the time of the peak of the grand average MMN. Trials in which either the STL or the deviant waveform were contaminated by an artifact were not included in this analysis. The *p* values obtained for each subject and each condition were combined into one *p* value using Fisher's combined test, frequently used for meta-analyses (Hedges and Olkin, 1985):

$$\chi^2 = -2 \sum_{k=1}^n \log(P_k)$$

where *n* is the number of subjects and *p_k* is the *p* value calculated for the *k*th subject. Log is the natural logarithm. The results belong to a χ^2 distribution with 2**n* degrees of freedom. Since each subject's *p* value is affected by the size of the effect, by the within-subject variance, and by the degrees of freedom (determined by the number of trials entering the analysis), this method intrinsically gives appropriate weights to the results obtained from different subjects. We will refer to this method of performing within-subject *t* tests and combining their *p* values as the Subject Variance Inclusion Test (SVIT).

This method was also used to calculate the significance of the difference between the effects of the two deviants, but this time, since deviant trials cannot be paired, a single nonpaired two-tailed test was used within each subject, and subsequently combined using the SVIT approach.

Scalp distribution calculations. To check whether the MMN effects for the HI-Dev and HC-Dev conditions resulted from different intracranial

source configurations, we compared the scalp distributions of the difference waves at the peak of both MMNs. For this analysis, we clustered the electrodes into 3 × 3 regions [left anterior region (electrodes F7, F3, and AF3), medial anterior region (electrodes FC2, FC1, and Fz), right anterior region (electrodes F8, F4, and AF4), left central region (electrodes CP5, C3, C5, and FC5), medial central region (electrodes CP1, CP2, and Cz), right central region (electrodes CP6, C6, C4, and FC6), left posterior region (electrodes O1, P7, and P3), medial posterior region (electrodes Oz, PO4, PO3, and Pz), and right posterior region (electrodes O2, P8, and P4)] (cf. Bell et al., 2010) and tested the interaction between regions and conditions after scaling the amplitudes (McCarthy and Wood, 1985).

Control experiments. A cardinal aspect of our experiment is that sounds were produced from real loudspeakers located in space, rather than virtually over headphones. However, this leaves possible alternative explanations for any MMN effect found in the HI-Dev condition.

In this setting, the deviant stimulus was produced by one loudspeaker—CS—whereas the standard stimulus was produced by a different loudspeaker—RS (Fig. 1*b*). Therefore, one alternative explanation is that if the speakers had different frequency responses, the MMN could be a response to deviance in frequency or intensity of the sounds, rather than to deviance in spatial location of the sound source. We examined this possible alternative explanation in three different ways.

First, we compared the power spectrum of the speakers using an omnidirectional microphone (model 46AE; G.R.A.S.) located at identical distances from the two examined speakers, and at the same location where the middle of the subject's head would be. Specifically, we examined whether there are systematic differences between the two speakers.

Second, we conducted a behavioral discrimination experiment to check whether subjects could detect any differences between sounds emitted by the different speakers if the two speakers were in the same location in space (i.e., eliminating the spatial difference). To do so, a single subject was equipped with two miniature electret microphones (KE4–211-2; Sennheiser) embedded in ear plugs. The microphones were placed in the external auditory canal, pointing outwards, with their front end aligned with the external auditory meatus. The same 250–4000 Hz bandpassed noise stimulus used in the main experiment was played 250 times from one speaker, placed straight ahead from the subject, and recorded stereophonically at 44.1 Khz/16 bit through the microphones (DMP3 preamplifier; M-Audio; Audigy 2 ZS sound card; Creative). Without moving the subject, the same was repeated with the other speaker positioned in the exact same straight-ahead place. Recorded this way, the stereophonic sounds genuinely reflected the sound pressure at the external auditory meati of the subjects, with all individual binaural and pinna-related spatial cues embedded (Møller et al., 1996; Hammershøj and Møller, 2002; cf. Deouell et al., 2007). Five subjects (*n* = 5, four male, with reported normal hearing) listened to 80 pairs of sounds through headphones (HD 25; Sennheiser), which were calibrated in closed loop with the microphones using a graphical equalizer to genuinely reproduce the recorded sounds. The subjects determined whether the two are similar (“belong to the same group”) or different (“belong to different groups”). Half of the pairs were similar (i.e., the members of the pair were each chosen randomly from the pool of recordings of the same speaker) and half different (chosen randomly, each from the pool of a different speaker). Subjects' responses were statistically compared with chance selection using a one-way *t* test.

Third, we tested whether any as yet undetected differences between the loudspeakers could elicit an MMN effect when the two loudspeakers were in the same location in space (“Loudspeaker MMN Control”). We again used individualized in-ear recorded stimuli (see above) in a typical MMN design. Three types of stimuli were recorded: (1) standards: 250–4000 Hz bandpassed noise recorded from one of the two speakers placed straight ahead from the subject, (2) spectral deviants: 1000–4000 Hz bandpassed noise recorded from the same speaker, and (3) loudspeaker deviants: 250–4000 Hz bandpassed noise recorded from the other speaker positioned at the same location in space as was the standard stimulus. After performing the in-ear recordings as described above, subjects (*n* = 9, six male, 23 ± 2.18 years old, with reported normal hearing) were seated comfortably in an acoustic room and watched a silent movie, while sounds were played by the microphone-calibrated

headphones. The subjects were instructed to disregard the sounds and focus on the movie. The EEG acquisition process and the postprocessing procedure were similar to those of the main experiment. Each block consisted of 2400 stimuli (SOA = 400 ± 50 ms). In each block, 80% of the stimuli were standards and 20% were deviants—either loudspeaker deviants (in the first block) or spectral deviants (in the second block). Stimuli were mixed in random order with the restriction that at least three standards preceded a deviant. Since our prediction was that no MMN would be elicited by the mere exchange of loudspeakers, the spectral deviation condition was used to confirm that the tested subjects indeed showed MMN responses to another auditory deviation. One subject did not show a clear MMN response to the spectral deviance and so was excluded from further analysis.

A further putative concern was that since the standards and the HI deviant in the main experiment were presented at different angles to the subject's body, the interactions of the sound with the body, as well as objects in the acoustic chamber (e.g., chair), could induce subtle spectral or phase differences that may elicit an MMN. To examine the possibility that such differences, rather than perceived spatial change, could be responsible for the MMN elicited in the HI condition, we conducted another control MMN experiment ("Body-interaction MMN Control"). The stimuli used in this control were three types of sounds recorded from each subject's ears (in the manner described above) in the same room and setup as in the main experiment: (1) simulated standard sounds (sSTD): the subject faced the central speaker and the sounds (250–4000 Hz noise) were presented from a speaker 30° to the right, just like in the standard condition of the main experiment; (2) simulated HI (sHI) deviants: the subject faced 30° to the left and the same sounds were presented by the central speaker, just like in the HI condition of the main experiment; and (3) spectral deviant sounds: the subject faced the central speaker and narrower band noise (1000–4000 Hz) was presented from the speaker 30° to the right. These individualized recorded stimuli were then played via calibrated headphones in the body-interaction MMN control experiment. While EEG was recorded, the subjects watched a silent movie presented on the straight-ahead screen and were instructed to ignore the sounds and refrain from head motion. In the first block, sSTD served as frequent (probability = 0.8) standard and sHI served as deviants (probability = 0.2). The differences between the recorded sSTD and sHI presumably encompassed all the acoustic differences between the standard and HI-Dev of the main experiment, but since the head was static in the control experiment, no HI spatial deviance was present. MMNs elicited under these conditions would imply that the subtle physical differences indeed account for at least some of the effect found in the original experiment's HI-Dev condition. In the second block, sSTD was again used as the frequent standard and the spectral deviant sound was used as a deviant. The purpose of this block was to confirm that the tested subjects indeed showed MMN responses when clear spectral changes were introduced (see the description of the loudspeaker MMN control experiment, above), and to allow the characterization of this group's spectral MMN in time and space for comparison with the sHI condition. Twelve subjects participated (25.6 ± 8.43 years old, five male, with reported normal hearing; an additional subject was omitted due to recording problems). The procedure, deviant probabilities, analyses, and rejection criteria were similar to those of the previously described loudspeaker MMN control experiment.

Results

The grand average waveforms of STL standards, HI-Dev and HC-Dev, and the difference waves are illustrated in Figure 2, *a* and *c*. After rejection of contaminated trials, an average of 154 trials was analyzed per subject per condition (range: 87–193). For both head-centered and head-independent conditions, the grand averages show a mismatch negativity starting at ~ 100 ms and peaking close to 170 ms. The topographies of the effects are typical of the MMN, showing a frontal negativity with slight rightward tendency and a reversal of polarity at the mastoids (Fig. 2).

Both head-related and head-independent deviations elicit an MMN

Traditional analysis

Both egocentric and allocentric conditions elicited a significant MMN (one-tailed *t* test against zero: HI-Dev, $t_{(15)} = 5.00$, $p < 0.0001$; HC-Dev, $t_{(15)} = 2.50$, $p < 0.02$).

SVIT analysis

The SVIT method took into account the differences between conditions within subjects. Using Fisher's combined test, we found that both the HI-Dev and the HC-Dev significantly differed from the STL standard (HI-Dev, $\chi^2_{32} = 109.41$, $p < 0.00001$; HC-Dev, $\chi^2_{32} = 68.63$, $p < 0.0005$).

Comparison of effect sizes between the two conditions

The MMN elicited by the HI-Dev was somewhat larger on average than the MMN elicited by the HC-Dev. The difference proved significant in a traditional two-tailed paired-samples *t* test ($t_{(15)} = 2.86$, $p < 0.02$) but not when using the SVIT statistical method ($\chi^2_{32} = 34.94$, $p = 0.33$). Further inspection of this result shows that none of the subjects individually showed a significant difference between the effects of the two conditions, possibly because of the very large variation of effect sizes within each subject.

Comparison of the scalp distribution between the two conditions

Using ANOVA of region X condition, we expected to find a significant interaction effect if there were differences in intracranial source configurations between the conditions. We found a main effect for region ($p < 0.0001$), which is a trivial result of the nonuniform scalp distribution (Fig. 2*b,d*), a main effect for condition ($p < 0.02$), and no interaction effect ($p = 0.99$). This implies that scalp distributions were not affected by the type of deviance.

The main finding of our experiment is that the head-independent deviants elicit a significant MMN. However, since the sounds were played in real space (rather than by headphones), we note that the standard and the deviant sounds in the HI-Dev condition might have differed in their physical properties, either as a result of the use of two separate loudspeakers (RS and CS) or as a result of sound–body interactions. These two alternative explanations were examined in separate control experiments (see Materials and Methods, above).

Is head-independent MMN due to physical difference between loudspeakers?

Could the MMN elicited in the HI condition be explained by subtle differences between the sounds emitted by the two loudspeakers, due to their physical characteristics (i.e., independent of the spatial locations)? We addressed this question in three ways (see Materials and Methods, above). First, we found that the interspeaker power differences across the spectrum were minute, below the measurement accuracy of 0.2 dB SPL (the spectra will be provided upon request). Second, we found that subjects could not tell the two speakers from one another: the discrimination between similar (same speaker) and dissimilar (different speakers) pairs was not different from chance ($47.85 \pm 2.9\%$ correct classification, $t_{(4)} = -1.67$, $p > 0.9$). Finally, we found that rare stimuli from one loudspeaker do not elicit a significant MMN effect when played in the context of standards from the other loudspeaker when the two loudspeakers do not differ in location (loudspeaker MMN control; $t_{(7)} = -0.68$, $p = 0.74$ using the common statistical method; $p = 0.55$ using the SVIT method; Fig. 3). In contrast, spectral deviants elicited a significant MMN response in the same subjects ($p < 0.005$ using the common

Figure 2. Grand average ($N = 16$) of HC-deviant and HI-deviant MMNs. **a, c**, Response waveform for the second-to-last standard, the deviant, and their difference wave, which illustrates the MMN effect, as recorded in electrode Fz. Time 0 represents stimulus onset. Error bars on difference wave represent SEMs across subjects. **b, d**, Scalp distribution of the difference wave at the peak of the MMN. The white asterisk represents electrode Fz.

statistical method; $p < 0.00001$ using the SVIT method). Taken together, it seems unlikely that the HI-Dev MMN was due to differences in loudspeaker characteristics.

Is head-independent MMN due to sound interaction with the body?

The body-interaction MMN control addressed the possibility that the HI MMN in the main experiment was elicited by spectral or phase differences due to interactions of sounds emitted from different locations in space with the body (or environment). Rare stimuli recorded under conditions identical to those of the HI-Dev condition were played in the context of frequent stimuli recorded under conditions identical to those of the standards of the main experiment, and did not elicit a significant MMN effect ($t_{(11)} = -1.14$, $p = 0.14$; Fig. 4). While the deviants did elicit a very small and short-lived negativity relative to the standards (Fig. 4), which passed the threshold of the SVIT analysis ($p < 0.01$), inspection of the scalp distribution of this effect, on the group average as well as in each subject separately, clearly showed that this difference does not fit an MMN topography in any way (Fig. 4). In comparison, the spectral deviant condition showed a pronounced MMN ($t_{(11)} = -7.47$, $p < 0.00001$ using the tradi-

tional method; $p < 0.00001$ using the SVIT), with typical scalp distributions in all subjects, indicating that these subjects do generate a typical MMN scalp distribution under the appropriate conditions. Thus, it seems that any physical differences that may exist between the HI-Dev and the standard due to interactions with the environment in the main experiment are not large enough to elicit the robust HI MMN seen in the main experiment.

Discussion

Head-centered and head-independent representations in the auditory cortex

To maintain perceptual constancy and allow proper motor planning, moving organisms need to take into account changes in the spatial relations between their sensory organs, their bodies, and the environment. In audition, this entails a transformation from a primary head-anchored spatial representation to secondary body-centered or world-centered representations. Behaviorally, real or illusory unilateral lengthening of neck muscles, indicating contralateral head rotation, affects the subjective localization of both visual (Biguer et al., 1988) and auditory stimuli (Lewald and Ehrenstein, 1998; Lewald et al., 1999; Van Grootel et al., 2011).

Figure 3. Grand average waveforms and difference waves at electrode Fz ($n = 8$) for the loudspeaker MMN control experiment conducted to rule out a speaker-elicited MMN effect. **a, c**, Average response waveform for standard stimuli (dotted line), deviant stimuli (dashed line), and the difference between the two (solid line), which illustrates the MMN effect. Time 0 represents stimulus onset. Error bars represent SEMs across subjects. Note that the lack of clear N1 component in the raw waveforms is typical of the fast stimulation rate in the control experiment. **a**, Spectral deviation elicited an MMN. **c**, Change in loudspeaker did not elicit an MMN. **b, d**, Average scalp distribution of the difference wave at the peak of the spectral MMN's difference wave (152 ms poststimulus onset). Note the different scales. The distribution in **b** resembles the distribution of an MMN.

This suggests that head position is taken into account in determining the position of objects relative to the world or to the body. Furthermore, recent psychophysical results suggest that this head-independent representation involves tonotopic stages of auditory processing (Van Groote et al., 2011). However, a recent electrophysiological study failed to obtain evidence for a head-independent representation at the level of the auditory cortex, suggesting that this transformation occurs further downstream (Altmann et al., 2009), possibly at the level of multimodal or motor planning.

Here, we modified the scheme developed by Altmann et al. (2009) to readdress this conundrum. Following Altmann et al. (2009), we used the MMN event-related potential to index detection of change in sound-source location (Schröger and Wolff, 1996) and examined the representation of sounds in either head-centered or head-independent frameworks using free-field stimuli. The presence of an MMN response to a deviation in some dimension provides unequivocal evidence that the dimension in question has been processed by the time the MMN has been

elicited (for review, see Näätänen et al., 2007). The critical question was whether change in the location of sounds in space elicits an MMN even if the sounds' location relative to the head remains unchanged. Since the MMN has been shown by several methods, including EEG, MEG, and hemodynamic (PET and fMRI) and animal electrophysiology, to originate primarily from supratemporal auditory cortex (for review, see Alho, 1995; Escera et al., 2000), this would provide an upper boundary for the stage where a secondary, head-independent representation is formed. Indeed, unlike the findings of Altmann et al. (2009), a change in the spatial location of a sound relative to the body and the environment elicited a robust head-independent spatial MMN, as did head-centered changes.

Previous studies suggested that the sources of the spatial MMN are localized in the planum temporale (Deouell et al., 2006, 2007). Other studies suggested that the sources of the MMN elicited by different acoustic dimensions are close but distinct (Giard et al., 1995). Our findings do not support topographically distinct sources for head-centered and head-independent representations; however,

Figure 4. Grand average waveforms and difference waves at electrode Fz ($n = 12$) for the body-interaction control experiment. **a, c**, Average response waveform for standard stimuli (dotted line), deviant stimuli (dashed line), and the difference between the two (solid line), which illustrates the MMN effect. Time 0 represents stimulus onset. Error bars represent SEMs across subjects. **a**, Spectral deviation elicited an MMN. **c**, Simulated head position elicited a small, nonsignificant negativity (see Results). **b, d**, Average scalp distribution of the difference wave at the peak of the spectral MMN's difference wave (133 ms poststimulus onset). Note the different scales. The distribution in **b**, but not **d**, resembles the distribution of an MMN.

this may be due to the low spatial resolution of extracranial EEG. The primary locus for integration of binaural head-centered information and head-position information also awaits more fine-grain studies, possibly using electrophysiology in animals. Early stations, such as the dorsal cochlear nucleus, receive vestibular input, which could inform of head motion (Oertel and Young, 2004). The cat superior colliculus (SC), which harbors auditory as well as visual spatial maps, receives a major input of neck muscle afferents. A recent study in the monkey revealed head-position gain effects on activity of SC neurons during a delayed saccade task (Abrahams and Rose, 1975; Nagy and Corneil, 2010), but direct evidence for interaction of head-position with spatial auditory representation is lacking. Parietal neurons also reflect head position (Brotchie et al., 1995) and may provide this information either to the SC or directly to the auditory cortex.

What is the frame of reference of the head-independent representation that was found in our study? Previous studies that manipulated head positions, as was done here, used the terms world-centered representations (Goossens and van Opstal, 1999; Kopinska and Harris, 2003; Vliegen et al., 2004; Van Grootel et

al., 2011) or allocentric representations (Altmann et al., 2009). However, showing a truly world-centered, allocentric framework, rather than a trunk-related one, requires the manipulation of trunk-to-world position as well. In fact, Van Barneveld et al. (2011) found no compensation for passive whole-body rotation on auditory localization (but see Lewald and Karnath, 2001, who found a small effect of rotation on localization using headphones). Thus, the point at which a truly world-centered, listener-independent representation is obtained is not fully determined. Interestingly, body-centered spatial representations that are independent from the primary (head-centered) representation are reminiscent of the body-centered visual receptive fields in the parietal lobe of monkeys, as described by Andersen and colleagues (1993).

Methodological considerations

As already noted, Altmann et al. (2009) did not find an MMN in their allocentric condition. Van Grootel et al. (2011), trying to reconcile this negative result with their own positive behavioral evidence for a head-free representation, conjectured that the

head-free representation is too weak to be detected by EEG. We however found clear indication for a head-free representation. Following our initial hypothesis, we suggest that this is due to our use of more ecological free-field stimuli in contrast to the canonical head-related transfer functions used by Altmann et al. (2009). While the use of virtual sound space has advantages in some situations, its use may be suboptimal, especially when the aim is to explore representations of externalized, non-head-centered space (Getzmann and Lewald, 2010).

Notwithstanding, since in our study stimuli were produced by distinct loudspeakers, a potential confound could be that subtle spectral differences between the loudspeakers, rather than true spatial changes, were responsible for the head-independent MMN. In other words, if one speaker produced softer sounds, or if its spectrum were different, the MMN elicited could reflect a frequency or intensity deviance, both known to elicit an MMN. However, three control experiments, using physical, psychophysical, and electrophysiological measures, suggest that this is unlikely. The loudspeakers were found to be perfectly matched in their frequency responses; subjects could not tell, above chance, whether two sequential stimuli were played from two different loudspeakers (when the speakers shared the same space) or from the same one; and finally, no MMN was elicited when one loudspeaker served as standard and the other as deviant when both were colocalized. An additional control experiment ruled out the possibility that subtle spectral differences, related to sound interaction with the body and environment, might contribute significantly to the HI MMN. Thus, it seems safe to conclude that mainly differences stemming from spatial factors played a role in the elicitation of the HI MMN in our study.

In a previous study, we showed that the amplitude of the MMN (derived by comparing deviant to standard) is linearly related to the magnitude of spatial change (Deouell et al., 2006; for the possible relevance of the derivation in this regard, see Horváth et al., 2008). Here, the MMN elicited by the head-independent changes tended to be larger than that elicited by the head-centered changes. While this may suggest higher sensitivity to body- or world-centered changes than to head-centered changes, we note that the difference was significant by the traditional statistical approach, but not in the meta-analytic approach. Nevertheless, the larger MMN in the critical head-independent condition serves to rule out the following alternative interpretation of our results. Even though we monitored head turns, if subjects did not turn their heads precisely toward the left LED, the angle between the head and the sound and the head-independent deviant stimulus would not be the same as in the standard sequence. If this were the case, the MMN elicited in the HI-Dev condition could be due merely to a small change in a head-centered framework. However, considering the relationship between MMN and magnitude of deviation, the MMN in this condition should have been much smaller than the MMN in the head-centered condition. This conclusion is also consistent with the null results of the body-interaction MMN control experiment.

To further determine the reliability of the MMN, we introduced a second statistical approach, the SVIT, which is seldom used in ERP research. In this approach, each subject is treated as an experiment, with its own significance value, and the overall significance is obtained using Fisher's meta-analysis method. This method uses nonparametric statistics to combine a group of statistical tests with identical null (H_0) and alternative (H_1) hypotheses and produce an overall p value signifying the probability of H_0 being true (Hedges and Olkin, 1985). The test has several potential benefits. First, the individual p values take into account

the intersubject variance, and not just the difference between the means of two (or more) conditions. Thus, the level of background noise within each single subject is considered, while this information is lost with simple averaging. Second, the individual p values consider the degrees of freedom (based on the number of valid trials) within each subject. This allows for appropriate weighting of differences between conditions in individual subjects. Third, the calculation of intersubject p values allows for more complex manipulation of the data. In our study, for example, we paired each standard with its deviant.

One of the most intriguing challenges in cognitive neuroscience is to elucidate the stages of construction of a stable representation of the world from the dynamic, multifaceted information we get from our sensory systems. A previous study (Schröger, 1996) showed that the two main head-centered auditory spatial cues, namely interaural level and time differences, are independently represented in auditory cortex. Our findings confirm that head-on-trunk information is combined with head-centered information relatively early (<200 ms) in auditory processing, and therefore head-independent space is represented in unimodal auditory cortex, alongside a head-centered representation.

Notes

Supplemental material for this article, including individual scalp distributions in the body-interaction control experiment, is available at <http://hcnl.huji.ac.il/supp/SchechtmanetalSuppl.pdf>. This material has not been peer reviewed.

References

- Abrahams VC, Rose PK (1975) Projections of extraocular, neck muscle, and retinal afferents to superior colliculus in the cat: their connections to cells of origin of tectospinal tract. *J Neurophysiol* 38:10–18. [Medline](#)
- Alho K (1995) Cerebral generators of mismatch negativity (MMN) and its magnetic counterpart (MMNm) elicited by sound changes. *Ear Hear* 16:38–51. [CrossRef Medline](#)
- Altmann CF, Wilczek E, Kaiser J (2009) Processing of auditory location changes after horizontal head rotation. *J Neurosci* 29:13074–13078. [CrossRef Medline](#)
- Andersen RA, Snyder LH, Li CS, Stricanne B (1993) Coordinate transformations in the representation of spatial information. *Curr Opin Neurobiol* 3:171–176. [CrossRef Medline](#)
- Arnott SR, Alain C (2011) The auditory dorsal pathway: orienting vision. *Neurosci Biobehav Rev* 35:2162–2173. [CrossRef Medline](#)
- Bell R, Dentale S, Buchner A, Mayr S (2010) ERP correlates of the irrelevant sound effect. *Psychophysiology* 47:1182–1191. [CrossRef Medline](#)
- Biguer B, Donaldson IM, Hein A, Jeannerod M (1988) Neck muscle vibration modifies the representation of visual motion and direction in man. *Brain* 111:1405–1424. [CrossRef Medline](#)
- Blauert J (1997) Spatial hearing. In: *The psychophysics of human sound localization*, revised edition. Cambridge, MA: MIT.
- Brotchie PR, Andersen RA, Snyder LH, Goodman SJ (1995) Head position signals used by parietal neurons to encode locations of visual-stimuli. *Nature* 375:232–235. [CrossRef Medline](#)
- Deouell LY, Parnes A, Pickard N, Knight RT (2006) Spatial location is accurately tracked by human auditory sensory memory: evidence from the mismatch negativity. *Eur J Neurosci* 24:1488–1494. [CrossRef Medline](#)
- Deouell LY, Heller AS, Malach R, D'Esposito M, Knight RT (2007) Cerebral responses to change in spatial location of unattended sounds. *Neuron* 55:985–996. [CrossRef Medline](#)
- Escera C, Alho K, Schröger E, Winkler I (2000) Involuntary attention and distractibility as evaluated with event-related brain potentials. *Audiol Neurotol* 5:151–166. [CrossRef Medline](#)
- Fort A, Delpuech C, Pernier J, Giard MH (2002) Early auditory-visual interactions in human cortex during nonredundant target identification. *Brain Res Cogn Brain Res* 14:20–30. [CrossRef Medline](#)
- Getzmann S, Lewald J (2010) Effects of natural versus artificial spatial cues on electrophysiological correlates of auditory motion. *Hear Res* 259:44–54. [CrossRef Medline](#)
- Giard MH, Lavikainen J, Reinikainen K, Perrin F, Bertrand O, Pernier J,

- Näätänen R (1995) Separate representation of stimulus frequency, intensity, and duration in auditory sensory memory: an event-related potential and dipole-model analysis. *J Cogn Neurosci* 7:133–143. [CrossRef](#)
- Goossens HH, van Opstal AJ (1999) Influence of head position on the spatial representation of acoustic targets. *J Neurophysiol* 81:2720–2736. [Medline](#)
- Hammershøi D, Møller H (2002) Methods for binaural recording and reproduction. *Acta Acustica Acustica* 88:303–311.
- Hedges LV, Olkin I (1985) *Statistical methods for meta-analysis*. Orlando: Academic.
- Horváth J, Czigler I, Jacobsen T, Maess B, Schröger E, Winkler I (2008) MMN or no MMN: no magnitude of deviance effect on the MMN amplitude. *Psychophysiology* 45:60–69. [CrossRef](#) [Medline](#)
- Jung TP, Makeig S, Westerfield M, Townsend J, Courchesne E, Sejnowski TJ (2000) Removal of eye activity artifacts from visual event-related potentials in normal and clinical subjects. *Clin Neurophysiol* 111:1745–1758. [CrossRef](#) [Medline](#)
- Kopinska A, Harris LR (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* 57:23–37. [Medline](#)
- Lewald J, Ehrenstein WH (1998) Influence of head-to-trunk position on sound lateralization. *Exp Brain Res* 121:230–238. [CrossRef](#) [Medline](#)
- Lewald J, Karnath HO (2001) Sound lateralization during passive whole-body rotation. *Eur J Neurosci* 13:2268–2272. [CrossRef](#) [Medline](#)
- Lewald J, Karnath HO, Ehrenstein WH (1999) Neck-proprioceptive influence on auditory lateralization. *Exp Brain Res* 125:389–396. [CrossRef](#) [Medline](#)
- McCarthy G, Wood CC (1985) Scalp distributions of event-related potentials— an ambiguity associated with analysis of variance models. *Electroencephalogr Clin Neurophysiol* 62:203–208. [CrossRef](#) [Medline](#)
- Møller H, Sørensen MF, Jensen CB, Hammershøi D (1996) Binaural technique: do we need individual recordings? *J Aud Engineer Soc* 44:451–469.
- Näätänen R, Paavilainen P, Rinne T, Alho K (2007) The mismatch negativity (MMN) in basic research of central auditory processing: a review. *Clin Neurophysiol* 118:2544–2590. [CrossRef](#) [Medline](#)
- Nagy B, Corneil BD (2010) Representation of horizontal head-on-body position in the primate superior colliculus. *J Neurophysiol* 103:858–874. [CrossRef](#) [Medline](#)
- Oertel D, Young ED (2004) What's a cerebellar circuit doing in the auditory system? *Trends Neurosci* 27:104–110. [CrossRef](#) [Medline](#)
- Plenge G (1974) On the differences between localization and lateralization. *J Acoust Soc Am* 56:944–951. [CrossRef](#) [Medline](#)
- Schröger E (1996) Interaural time and level differences: integrated or separated processing? *Hear Res* 96:191–198. [CrossRef](#) [Medline](#)
- Schröger E, Wolff C (1996) Mismatch response of the human brain to changes in sound location. *Neuroreport* 7:3005–3008. [CrossRef](#) [Medline](#)
- Van Barneveld DC, Binkhorst F, Van Opstal AJ (2011) Absence of compensation for vestibular-evoked passive head rotations in human sound localization. *Eur J Neurosci* 34:1149–1160. [CrossRef](#) [Medline](#)
- Van Grootel TJ, Van Wanrooij MM, Van Opstal AJ (2011) Influence of static eye and head position on tone-evoked gaze shifts. *J Neurosci* 31:17496–17504. [CrossRef](#) [Medline](#)
- Vliegen J, Van Grootel TJ, Van Opstal AJ (2004) Dynamic sound localization during rapid eye-head gaze shifts. *J Neurosci* 24:9291–9302. [CrossRef](#) [Medline](#)