Eye movements during auditory attention predict individual differences in dorsal attention network activity

Rodrigo M. Braga1, 2, 3*, Richard Z. Fu1, Barry M. Seemungal1, Richard J. Wise1, Robert Leech1

1Computational, Cognitive and Clinical Neuroimaging Laboratories, Imperial College London, United Kingdom, 2Center for Brain Sciences, Harvard University, USA, 3Aathinoula A. Martinos Center for Biomedical Imaging, USA

Submitted to Journal:
Frontiers in Human Neuroscience

ISSN:
1662-5161

Article type:
Original Research Article

Received on:
20 Dec 2015

Accepted on:
01 Apr 2016

Provisional PDF published on:
01 Apr 2016

Frontiers website link:
www.frontiersin.org

Citation:
Braga RM, Fu RZ, Seemungal BM, Wise RJ and Leech R(2016) Eye movements during auditory attention predict individual differences in dorsal attention network activity. Front. Hum. Neurosci. 10:164. doi:10.3389/fnhum.2016.00164

Copyright statement:
© 2016 Braga, Fu, Seemungal, Wise and Leech. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution and reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.
Eye movements during auditory attention predict individual differences in dorsal attention network activity

Rodrigo M. Braga*1,2,3, Richard Z. Fu1, Barry M. Seemungal1, Richard J. S. Wise1 & Robert Leech1

1Computational, Cognitive and Clinical Neuroimaging Laboratory, Division of Brain Sciences, Imperial College London, Hammersmith Hospital Campus, Du Cane Road, London, UK.
2Center for Brain Science, Harvard University, Cambridge, Massachusetts, USA.
3Aathinoula A. Martinos Center for Biomedical Imaging, Charlestown, Massachusetts, USA.

* Correspondence: Rodrigo M. Braga, Computational, Cognitive and Clinical Neuroimaging Laboratory C3NL Burlington Danes Building, Hammersmith Hospital Campus, Du Cane Road, Imperial College London, W12 0NN
rodrigo.braga@imperial.ac.uk

Keywords: Auditory attention, Listening, Eye Movements, Saccades, Dorsal Attention Network, Spatial Attention.

Abstract

The neural mechanisms supporting auditory attention are not fully understood. A dorsal frontoparietal network of brain regions is thought to mediate the spatial orienting of attention across all sensory modalities. Key parts of this network, the frontal eye fields (FEF) and the superior parietal lobes (SPL), contain retinotopic maps and elicit saccades when stimulated. This suggests that their recruitment during auditory attention might reflect crossmodal oculomotor processes; however this has not been confirmed experimentally. Here we investigate whether task-evoked eye movements during an auditory task can predict the magnitude of activity within the dorsal frontoparietal network.

A spatial and non-spatial listening task was used with on-line eye-tracking and functional magnetic resonance imaging. No visual stimuli or cues were used. The auditory task elicited systematic eye movements, with saccade rate and gaze position predicting attentional engagement and the cued sound location, respectively. Activity associated with these separate aspects of evoked eye-movements dissociated between the SPL and FEF. However these observed eye movements could not account for all the activation in the frontoparietal network. Our results suggest that the recruitment of the SPL and FEF during attentive listening reflects, at least partly, overt crossmodal oculomotor processes during non-visual attention. Further work is needed to establish whether the network’s remaining contribution to auditory attention is through covert crossmodal processes, or is directly involved in the manipulation of auditory information.
1 Introduction

The visual and auditory sensory systems can be thought of as serving a single role – to gather information about our surroundings so that we may adapt our behavior accordingly. When a loud sound alerts us to a potentially dangerous situation, our eyes instinctively orient towards the source of that sound to gain further knowledge of its identity. This example highlights that the two sensory systems are intimately linked, with attention-capture in one modality often leading to recruitment of the other. Despite this, we are able to control which features of a given sensory modality we wish to pay attention to. This suggests some degree of modal separation in the ‘top-down’ or ‘endogenous’ attentional modulation of sensory information. This nuanced relationship poses a conundrum for establishing the neural correlates of auditory and visual attention, if indeed they are subserved by separate systems.

In vision, a great body of research has established that a dorsal frontoparietal network comprised of the superior parietal lobe (SPL) and frontal eye fields (FEF) activated during top-down attention (Kincade et al. 2005; Vossel et al. 2006, Corbetta et al. 2008). This network is sometimes called the “dorsal attention network” (DAN; Corbetta et al. 2008), and is active for example during the maintenance of attention to a visual field location in anticipation of a visual stimulus (Kastner et al. 1999). There is complementary evidence that the core nodes of this network, the SPL and FEF, do have a primarily visuospatial role. The regions are known to be involved in eye movement control (Büttner-Ennever and Horn 1997) and visuospatial processing (Behrmann 2004). Retinotopic maps have been located in both the FEF and SPL using direct stimulation and functional neuroimaging (Moore et al. 2003; Ruff et al. 2008; Saygin and Sereno 2008). The FEF and neighboring supplementary eye fields (SEF) were shown to be activated during the planning of subsequent saccades using electrophysiology and functional MRI (Isoda and Tanji 2003; Hu and Walker 2011). Further, transcranial magnetic stimulation (TMS) of the FEF delays voluntary saccades (Muggleton et al. 2011). These different lines of evidence converge on the DAN having a role in oculomotor control and visual orienting.

In hearing, the networks subserving top-down attention are not as well understood. A frontotemporal network consisting of the middle and inferior frontal gyri (IFG) and regions near the posterior superior temporal sulcus (pSTS), has been proposed to mediate the orienting of attention to non-spatial features of sounds, such as frequency and identity (Braga et al. 2013, Maeder et al. 2001, Salmi et al. 2007, Seydell-Greenwald et al. 2013). In support of this, activity in the SPL and FEF is notably absent from tasks that require auditory attention such as speech and music perception (Hickok et al. 2003; Warren 2008). However, during sound localization dorsal frontoparietal activity is often observed (e.g. Alho et al. 2015) even in the absence of visual stimuli (Maeder et al. 2001; Shomstein and Yantis 2006; Hill and Miller 2010; Alain et al. 2001; Lewis et al. 2000; Petit et al. 2007). The FEF even shows preparatory activity for spatial listening in the absence of auditory or visual stimuli (Lee et al. 2013). This has led to the theory that the dorsal frontoparietal network is “amodal” and directly mediates attentional orienting to all sensory modalities (Posner and Petersen 1990; Driver and Spence 1998; Macaluso 2010).

It is difficult to reconcile the DAN’s role in eye movement control with its recruitment during auditory orienting without recourse to a possible crossmodal cause (Driver and Spence, 1998). Auditory attention is likely to involve both direct modality-specific as well as indirect cross-modal processes, and the role of the DAN in this regard is not clear. On the one hand, activation of the DAN could represent direct manipulation of auditory processes, such as the top-down tuning of auditory
spatial receptive fields in the auditory cortex (Fritz et al. 2010). Alternatively, the activation of the
DAN during listening may be representative of indirect processes such as visual spatial orienting or
task-induced eye movements. Such processes may be facilitatory for auditory attention, even if not
directly involved in modulating auditory information.

There is behavioral evidence that auditory attention elicits systematic eye movements (Paulsen and
Ewertsen 1966; Rolfs et al. 2005; Valsecchi and Turatto 2009; Kerzel et al. 2010; Yuval-Greenberg
and Deouell 2011; Zou et al. 2012). For example, an auditory stimulus leads reliably to visual
saccades towards the source of the sound (Zahn et al. 1978; Zambarbieri et al. 1982; Van Grootel and
Van Opstal 2009). Rotating a sound about a subject’s head can induce nystagmus (Paulsen and
Ewertsen 1966). Further, the presentation of an auditory stimulus can reduce the rate of saccades
(Kerzel et al. 2010; Yuval-Greenberg and Deouell 2011; Zou et al. 2012; Rolfs et al. 2005). There is
also evidence that gaze position can affect auditory localization accuracy (Maddox et al. 2013).
However, neuroimaging studies rarely consider the influence of eye movements on auditory
attention, meaning that parts of the networks implicated in auditory attention may in fact be
mediating these crossmodal effects.

Given that the DAN, and particularly the FEF, is known to be involved in the generation of saccades,
it is possible that its recruitment during listening tasks reflects task-induced indirect oculomotor
processes. To test this hypothesis, eye movements need to be recorded in an auditory attention task in
the absence of visual stimuli and without any requirement for saccades or fixation. In contrast, the
usual approach has been to employ fixation conditions to investigate how gaze position affects
auditory processing (e.g. sound localization performance; Maddox et al. 2013). If DAN activity
during listening is associable with indirect processes such as increased eye movement control while
other parts of the auditory network are not, this would provide evidence that DAN activation is the
result of an indirect attentional route. This evidence would be particularly strong if the magnitude of
DAN activity represents the magnitude of indirect crossmodal influences. In the present manuscript
we recorded eye movements during a purely auditory attention task using functional magnetic
resonance imaging (fMRI) and in-scanner eye-tracking (Figure 1). Our hypothesis was that attentive
listening would be associated with systematic effects on eye movements, and that these effects would
be associated with the magnitude of activity within visuospatial regions of the DAN.

2 Materials and Methods

2.1 Subjects

Twenty healthy right-handed volunteers (9 female, mean age 26.2, range 21 to 36). All participants
reported no hearing problems and had normal or corrected vision (via contact lenses or MRI
compatible glasses). The study was conducted in accordance with the guidelines of Imperial College
Research Ethics Committee, and written consent was obtained from all volunteers before their
participation. Participants were screened for contraindications to MRI, and were excluded on the
basis of hearing difficulties and previous psychiatric or neurological disorders. Four additional
participants were scanned but had to be excluded due to technical issues with the eye tracking
equipment (3) and excessive motion (1).

2.2 Auditory Task

Subjects listened to 12 different looped melodies. Each melody lasted 2s, and was repeated four
times in each trial so that each trial lasted 8s. The melodies were generated by manually selecting
sequences of diatonic notes within one octave using prepackaged synthesizer sounds from the Logic Pro X software (version 10.2.2). Each melody contained between 7 and 12 staccato notes (mean 10 notes). Six of the melodies had a low tonal center (F above C1), and six had a high tonal center (C3), with no overlap in pitch between high and low pitch sequences. In each trial, two looped melodies, one high and one low, were presented simultaneously, either dichotically (one sound in each ear) or diotically (both sounds in both ears equally; Figure 1). The competing melodies overlapped in terms of note onset and duration, but not pitch. Stimuli were presented using Sensimetrics S14 sound-attenuating in-ear MR-compatible headphones. Subjects were trained outside the scanner to listen out for an oddball target in the form of a “pitch change”, which was in fact a transposition of the whole 2s melody to a tonal center 7 semitones above the original key. This key change made the two melodies incongruous and the target detectable. The task was split into 2 identical blocks. Each block contained 70 trials, of which 23 contained a pitch change in the cued melody (the target), 7 contained a distractor pitch change in the non-cued melody (catch trials), and 10 were silent rest trials. Targets were not presented in 30/70 trials. Targets and distractors were presented in either the second (7/70), third (9/70) or fourth (14/70) repeat of the 2s diatonic melody, in a pseudo-randomized order to avoid long repeats and ensure an even distribution of Rest trials. Distractors and targets were never presented in the same trial. The order of stimulus presentation was changed halfway through the experiment (6 subjects received one order and 14 received the other) to control for order-effects. We performed a confirmatory analysis with balanced groups (n=6 subjects receiving each presentation order) which confirmed that the eye movement behavioral results reported were not due to order-effects. Before each trial, subjects were presented with a diotic auditory spoken word (“Right”, “Left”, “High” or “Low”) which cued them to listen to the melody presented in their right or left ear, or that was higher or lower in pitch, respectively. The cue period lasted 2s, and the spoken cue onset was at the start of those 2s (not centered within the 2s). The task period lasted 8s, and the response period lasted 3s (including an auditory ‘Please respond’ cue). Each trial was followed by a period of silence lasting between 1-3s. Ten silent ‘Rest’ trials were also interspersed between listening trials. These were preceded by an auditory spoken cue (“Rest”), and no auditory stimuli were presented for the same duration as a normal trial and response period. No “Respond” cue was presented after ‘Rest’ trials. Subjects were instructed to keep their eyes open throughout the listening experiment. A featureless black screen was displayed during the whole experiment and no instructions to fixate were given. Subjects were naïve to the purpose of the experiment, and were told that the eye tracker would be used for a separate visual task that took place between the two blocks of the auditory task.

2.3 Saccade Distractor task

In between each block of the auditory task, subjects performed a visual distractor task that had 2 components: 1) visual fixation to a central cross (white on black background), and 2) forced saccades to a white cross that appeared unpredictably on each corner of the black screen (see Eye Tracking). These two tasks lasted 32 seconds each and were repeated 4 times. Four rest periods of 32 seconds duration were interspersed between tasks, wherein a blank screen was presented. The distractor block served as an explanation to subjects for the presence of the eye tracking equipment, for calibration of the eye tracker, and to functionally localize the DAN.

2.4 Eye Tracking

Vertical and horizontal gaze displacements were recorded at 500 Hz using a MR-compatible head-mounted infrared camera (Jazz-NOVO, Ober Consulting, Eye movement range - vertical: +/- 20º, horizontal +/-35º; sampling frequency - 500Hz). The voluntary saccades element of the visual distractor task served as a four-point calibration and was performed in between the two runs of the
Eye movements during auditory attention

auditory task to be close in time to both runs. In this task, white crosses were presented in each corner of a black 7.5” IFIS-SA LCD screen. The screen was at a viewing distance of 13 cm. The crosses subtended a horizontal angle of 60° and vertical angle of 40° from one another. Gaze displacements to the left, right, upper and lower visual spaces were quantified relative to a center point, which was defined as the average vertical and horizontal gaze position across the run.

Eye movements were analyzed using the Jazz-Manager software (saccade detection, blink removal) and using in-house software based on MATLAB (normalization and gaze displacement measurement). For each participant, gaze displacement along both axes was detrended to remove low frequency drifts, and normalized by dividing by the standard deviation within each 17min run. The saccade detection algorithm involved first a differentiation of the raw eye position signal (in degrees of angle) to velocity (°/s). A saccade was then determined if it satisfied all of the following criteria:

(i) an initial velocity of 35 °/s or greater; (ii) a minimum peak velocity of 100 °/s; (iii) a maximum duration of 300 ms; (iv) a minimum duration of 20 ms; (v) an inter-saccadic interval of 50 ms or greater (since very short inter-saccadic intervals of <50 ms would indicate either artifact or pathological eye movements such as ocular flutter or opsoclonus).

The eye blink detection algorithm uses the fact that during an eye blink, the eye moves primarily in the vertical plane, first up and then down (“Bell’s phenomenon”) and hence this algorithm used only the vertical eye signal. It also follows that blinks are bi-phasic with two velocity peaks per blink. The eye blink detection algorithm used the following criteria: (i) a first peak minimum velocity of 200 °/s; (ii) a second peak minimum velocity of 100 °/s; (iii) a maximum inter-velocity-peak duration of 150 ms; (iv) a total blink duration of between 100 ms (minimum) and 500 ms (maximum); (v) an amplitude of 15 ° or greater; (vi) since the vertical eye position at the end of a blink is typically close to the pre-blink position, a ratio of the final to initial vertical eye position for a blink should be close to 1. This ratio was set to between 0.6 (minimum) and 1.4 (maximum).

2.5 MRI acquisition

MRI data were obtained using a Phillips Intera 3T MRI system with an 8-element phased array head coil and sensitivity encoding. High-resolution (1 mm × 1 mm × 1 mm) T1-weighted whole-brain structural images were obtained for each participant to allow accurate spatial registration of the functional images. Functional MRI data were acquired using an echoplanar imaging (EPI) sequence. Continuous data acquisition was used to collect whole-brain images in 44 axial slices with a slice thickness of 3.5mm, and a repetition time (TR) of 3s (TE=45ms, FOV=220×143×190mm, 44 slices, 128 x 128 voxels per slice, voxel size: 2.18 x 2.18 x 3.25 mm). The first three dummy scans were discarded. A total of 674 whole brain functional images were acquired for each subject, split into two runs of 337 images. Paradigms were programmed using MATLAB and stimuli presented through an IFIS-SA system (In Vivo Corporation). Eye tracking data and stimulus presentation were synchronized to the initial scanner pulse. Responses were recorded through a fiber optic response box (Nordicneurolab, Norway), interfaced with the stimulus presentation PC running MATLAB.

2.6 fMRI image analysis
Standard preprocessing was carried out using FSL (FMRIB's Software Library (Smith et al. 2004), www.fmrib.ox.ac.uk/fsl). Image pre-processing involved realignment of EPI images to reduce the effects of motion between scans, spatial smoothing using a 8mm full-width half-maximum Gaussian kernel, pre-whitening using FILM and temporal high-pass filtering using the default cut-off frequency of 100 Hz. FMRIB's Linear Image Registration Tool (FLIRT) was used to register EPI functional datasets into a standard MNI space using the participant's individual high-resolution anatomical images. We also ran three separate confirmatory analyses using nonlinear registration and spatial smoothing at 6 and 10mm to test that our choice of registration and smoothing parameters were not affecting the result. These analyses produced qualitatively similar activation patterns to those reported.

The listening task was divided into four conditions depending on the cue and dichotic/diotic method of stimulus presentation. Seven variables were entered into a general linear model with the onsets and durations of the Cue, Response, Listen-Left, Listen-Right, Listen-High/Low (diotic) and Listen-High/Low (dichotic) trial periods, and one additional Error variable that modeled the trials in which subjects responded incorrectly (across all conditions). The model included the full duration of each condition. ‘Rest’ trials were not modeled in the general linear model and served as the implicit baseline along with the silent periods following each trial. To rule out that the observed behavioral and neuroimaging results might be a result of participants closing their eyes during the trials, we ran a confirmatory analysis where we removed any trials during which no saccades were detected. In this separate confirmatory analysis, no-saccade trials were included in the Error variable and excluded from the task condition variables. This analysis produced similar neuroimaging and eye tracking results. Individual variability in the number of trials without saccades also did not correlate with any of the behavioral rankings (gaze position bias, saccade inhibition or task inhibition). A synthetic double-gamma hemodynamic response function was convolved with each explanatory variable and its first temporal derivative was included to account for variability in the hemodynamic delay function. Six motion parameters were included in the general linear model as confound regressors. To conclusively rule out motion as a potential confound, we ran a confirmatory analysis using 24 motion regressors which produced qualitatively similar results.

2.7 Group fMRI analysis

Mixed effects analysis of session and group effects was carried out using FLAME (FMRIB's Local Analysis of Mixed Effects(2003)). Final statistical images were thresholded using Gaussian Random Field based cluster inference with an initial cluster-forming threshold of Z>2.3 and a cluster significance threshold of p<0.05. This resulted in statistical maps of clusters significantly activated by the task. Group-mean images were produced by giving each subject equal weighting. Subjects were also ranked by three behavioral variables; 1) mean task performance, 2) mean difference in saccade rate between all task and ‘Rest’ trials, and 3) mean gaze position difference between ‘Listen-Left’ and ‘Listen-Right’ conditions. These rankings were zero-meaned and used as weightings for each subject to probe individual differences in each behavioral measure. These weightings were entered into the same higher-level model. We also carried out a confirmatory group-level analysis using each subject’s (demeaned) t-values for gaze position bias and saccade inhibition instead of their relative ranking. This analysis revealed qualitatively similar results to those reported.

2.8 Functional connectivity analysis

To test the network membership of our eye-movement-derived regions of interest, we used resting state data from 20 participants data from the Human Connectome Project (Smith et al. 2013; Van...
Eye movements during auditory attention

Essen et al. 2013). This dataset was used because of its high quality, and because resting state data was not collected from the participants who took the auditory task. The data consisted of the first 20 subjects from the third phase public release from the Human Connectome Project. Each resting state acquisition was composed of 1200 whole-brain volumes, with a TR=0.72, collected on a 3T MRI scanner. We used the spatially and temporally preprocessed version of the data. Each run was corrected for spatial distortions from gradient nonlinearity and from motion by registration to a reference image as well as corrected for B0 distortion before being registered to a high-resolution structural image and into MNI standard space. A liberal 2000s cut-off for a high-pass temporal filter was applied to the data. Twenty-four motion parameters were then aggressively temporally filtered out of the data, along with other non-neural structured noise identified with FIX automatic independent component denoising approach (Salimi-Khorshidi et al. 2014), as described by Smith et al (Smith et al. 2013). In addition, the data was downsampled into 4x4x4mm space, to reduce the computational overhead. Functional connectivity was calculated using the dual regression (Beckmann et al. 2009) pipeline from FSL version 5. Two regions of interest (the activation pattern corresponding to individual variability in either eye gaze displacement or saccade inhibition) were entered separately into the dual regression as the design matrix for a general linear model with the HCP rest data as the dependent variable. This resulted in a timecourse which was then regressed again with the rest data, resulting in a whole-brain spatial map of regression coefficients estimating functional connectivity with the initial pattern of activation. The functional connectivity map for each subject was then entered into a higher-level general linear model. Results were thresholded using a family-wise error correction for multiple comparisons.

3 Results

3.1 Behavioral results:

3.1.1 Natural gaze position is biased towards attended sound location

Despite the absence of visual cues and stimuli, following spatial (‘Left’ or ‘Right’) auditory cues, subjects tended to shift their gaze left or right consistent with the auditory spatial task (left/right; Figure 2A). A significant difference in mean gaze position was observed between trials when subjects were cued to listen to their left vs. right ear (t-test of “Left” vs “Right” cue trials, t19 = -4.54, p < 0.001, Figure 3A). This gaze position bias was not observed for dichotic trials preceded by a spectral (i.e. “High” or “Low”) cue (Figure 2B). No effects on vertical gaze position were observed for Right-Left or High-Low discriminations (Left-Right: t19 = -0.42 p = 0.68; High-Low: t19 = 0.25 p = 0.81).

3.1.2 Saccade rate decreases during auditory attention

We used a within-subject design to compare saccade rate between task trials and the ten silent ‘Rest’ trials which were interspersed across each block. The auditory attention task was associated with a lower saccade rate compared to rest (t19 = 1.973, p < 0.05, Figure 3B). This was observed across all trials, regardless of cue type.
3.1.3 Individual variability in eye movements and performance

Subjects performed the task with high accuracy (average 92.1% correct responses, standard deviation 8.3%). No significant differences in performance were observed between spatial (Left-Right) and spectral (High-Low) discrimination conditions (Left-Right: 93.9 %, High-Low: diotic 91.4 %, dichotic 91.0%, pairwise t-tests: all p > 0.05, n.s.). Subjects ignored the distractor pitch change successfully in the majority of catch trials (percentage of catch trials with false alarms: 8.2 %). In terms of eye movements, high variability in effect of task was observed across participants for both saccade rate and gaze displacement following spatial cues (Figure 3C). We assessed whether the individual differences in eye movement control predicted individual differences in performance on the auditory task. The task performance scores did not correlate with the inhibition of saccade rate (Figure 3D, R2 < 0.01, n.s.) or gaze position bias (R2 = 0.01, n.s.) across subjects. Gaze position bias and saccade rate inhibition across participants were also not correlated with each other (R2 all < 0.01, n.s.).

3.2 Neuroimaging results:

3.2.1 Activity in frontotemporal and DAN areas relates to auditory attention and individual variation in task performance

Across all task conditions, activation in widespread regions was observed during attentive listening (Figure 4). This network included the SPL, FEF and middle frontal gyrus (MFG) in both hemispheres, as well as both superior temporal gyri and sulci, and regions of the cerebellum. We ranked subjects by their performance scores (% percentage correct) in the pitch-change detection task (Figure 3C), and assessed which brain regions had activity relating to this rank order during the task. Increased activation of the same widespread DAN and frontotemporal network was associated with improved performance on the task. No differences in the mean (unranked) signal were observed between specific task conditions, such as “Left”, “Right”, “High” or “Low” (diotic and dichotic) trials.

3.2.2 Gaze position bias mediated by the SPL

To determine which brain regions mediated the behavioral gaze bias evoked by the spatial task (which were observed during the ‘Left’ and ‘Right’ cue conditions; Figure 3A) whilst controlling for some of the listening requirements and auditory input, we contrasted the spatial and spectral listening conditions (‘Listen-Left’ + ‘Listen-Right’ > ‘Listen-High or Low’ dichotic). Although no differences were found in the mean signal for this contrast, this may have been a consequence of the individual variability present in gaze position bias (Figure 3C). Therefore, we ranked subjects by their difference in mean gaze displacement between ‘Listen-Left’ and ‘Listen-Right’ trials (average gaze position during all “Listen-Right” trials minus average gaze position during all “Listen-Left” trials). Subjects that showed the greatest mean gaze position difference between left and right trials were given the highest rank, and displayed higher activation of the posterior parietal and suprolateral occipital lobes bilaterally during the spatial compared to spectral listening trials (Figure 4). No differences were found for the direct contrast of ‘Listen-Left’ > ‘Listen-Right’.
3.2.3 Saccade inhibition effect mediated by the FEF

Substantial individual differences in the task-induced inhibition of saccades were also observed (Figure 3B). Therefore, to assess which brain regions mediated this effect, we ranked the participants by the amount of reduction in their saccade rate during the task (saccade rate during listening trials minus saccade rate during rest periods; Figure 3C). Subjects with the largest reduction in saccade rate between rest and task trials were given the highest rank. No regions of activation were observed during the task (compared to the implicit baseline) for this rank analysis. However, during the cue period immediately preceding the task subjects that showed the greatest reduction in saccade rate displayed higher activation of the FEF and MFG bilaterally, as well as some activation in the left superior temporal gyrus (Figure 4).

3.2.4 Auditory task-evoked eye movement regions overlap with and dissociate anterior and posterior components of the DAN

The visual distractor task was used to functionally localize the DAN in our dataset (Figure 5). The activation patterns obtained from the rank analyses of gaze position bias and saccade inhibition were located primarily within regions of the DAN, as evoked by the independent visual task acquired with the same participants. To provide a more detailed description of these two patterns of activation, their functional connectivity with the rest of the brain was assessed using resting state data acquired as part of the Human Connectome Project (Van Essen et al. 2013). The regions activated by variability in gaze position were functionally connected to the DAN, including SPL, FEF and supplementary eye fields near the midline. There was also extensive connectivity with the dorsal and ventral visual streams bilaterally, extending from the SPL via the occipital lobes to the fusiform gyri (Figure 5). Regions activated by variability in saccade inhibition were also functionally connected to the whole DAN, including FEF and SPL, and also the MFG, and visual streams to a lesser extent. There was also extensive connectivity with posterior superior and posterior middle temporal cortices, and anterior regions of the lateral prefrontal cortices. Both functional connectivity maps overlapped considerably with the DAN as evoked by the visual task (Figure 5).

4 Discussion

The study demonstrates that attentive listening is associated with changes in eye movements that are independent of visual stimuli or visual demands. In addition, this study shows that these crossmodal effects are associated with increased activity in core regions of the dorsal frontoparietal network, the FEF and SPL. Rather than being directly involved in the attentional selection of auditory information, the present results suggest that the role of the DAN during auditory attention is at least partly an indirect one; mediating task-evoked eye movements. It remains to be determined whether the remaining role of the DAN in auditory attention also represents crossmodal process (such as covert multimodal attentional orienting without eye movements, the formation of supramodal spatial maps, or the inhibition of non-auditory inputs) or truly represents direct attentional modulation of auditory information.

Although the neural correlates of eye movements have been investigated previously (e.g. Nobre, et al., 2000), to our knowledge, this is the first study that has simultaneously shown that eye movements are evoked by a purely auditory task and that these eye movements are associated with the DAN. This has important implications for the supposed role of the DAN as an amodal
Eye movements during auditory attention

Attentional system in that it suggests that the DAN mediates crossmodal oculomotor processes, which may or may not be intrinsic to attention to all modalities. Further work is therefore necessary to establish the exact contribution of this brain system to non-visual attention. It is likely that these frontoparietal regions, which are activated under many conditions and have been associated with numerous cognitive processes (Cabeza and Nyberg, 2000), may play multiple roles even within the context of visual and auditory processing.

4.1 Eye movement control during attentive listening

In this study, participants reduced their rate of saccadic eye movements when they were required to listen attentively (Figure 3B). As no visual cues or stimuli were presented at any stage of the auditory task, and there were no behavioral requirements for eye movements, this natural inhibition of saccades was driven by the requirement to listen attentively. The inhibition of eye movements when attentional resources are required in the auditory modality might serve to reduce the amount of novel incoming visual information which could interfere with the auditory task. Another explanation could be that the resources normally used to plan and execute saccades are diverted to auditory feature selection or spatial orientation during attentive listening. Either way, the auditory-evoked saccade rate inhibition suggests that auditory attention competes with visual control processes during natural listening conditions.

In addition, participants tended to look towards the direction of the cued sound after receiving spatial auditory cues (i.e. to listen to the melodies in their left or right ear; Figure 2). Previous research (Zahn et al. 1978; Zambarbieri et al. 1982) has shown that saccades are often made towards a presented sound, an instance of crossmodal effects during ‘bottom-up’ attention capture. In our study, subjects were presented with sounds in both ears, making it unlikely that ‘bottom-up’ auditory attention capture determined the gaze position bias. This suggests that top-down or endogenous auditory attention, driven by the spoken instructions, influenced gaze position during the present task. Previous behavioral studies have shown that gaze position has a substantial effect on the accuracy of auditory spatial localization accuracy (Razavi et al. 2007; Pavan et al. 2008; Van Grootel and Van Opstal 2009; Pages and Groh 2013, Maddox et al. 2013). Subjects with no functioning visual system were found to be severely impaired on a spatial but not a non-spatial auditory attention task (Gori et al. 2014). These findings suggest that auditory spatial maps are likely to be calibrated or fine-tuned by gaze position, meaning that the interaction between oculomotor and auditory processes may be an intrinsic facilitatory mechanism for auditory localization.

4.2 Individual variability across separate dimensions of eye movement control reveals different listening strategies

Importantly, although some subjects showed consistent crossmodal effects across trials, not all subjects displayed systematic eye movements during attentive listening. This heterogeneity across participants is consistent with previous behavioral research (Yuval-Greenberg and Deouell 2011). In the present study, individual variability in saccade rate inhibition was not correlated with individual variability in gaze position bias (Figure 3C), meaning that subjects that displayed one eye movement effect did not necessarily display the other. Therefore, our data suggest that saccade rate and gaze position bias were two separate manifestations of auditory attention affecting eye movements, possibly reflecting underlying listening strategies.

It is possible that the auditory-induced gaze-position bias we observed was facilitatory, helping some subjects to focus on the attended sounds and ignore the competing sounds. However, no relationship
between gaze position and task performance was observed in this study. We were therefore unable to
distinguish whether gaze position aided performance or was an epiphenomenon. Here, we observed
a left-right gaze bias only in trials when subjects received spatial (“Left” or “Right”) cues, not for
non-spatial cues (“High” or “Low”), even though these could also be discriminated spatially when
presented dichotically. This suggests that the gaze position bias reflected a top-down strategy that
some subjects were employing during the spatial task. Although pitch can be discriminated without
any spatial information, there is prior evidence for an associations between ‘High’ and ‘Low’ pitch
and upward and downward visual space (Chiou and Rich, 2012). In this study we did not find any
vertical gaze position effects when subjects listened out for different pitch qualities. This could be
due to many reasons (outside the remit of this study), such as the nature of the sounds that were used
as stimuli, the difference in pitch between high and low competing sounds, as well as the level of
musical training of the subjects.

4.3 SPL activity correlates with gaze position bias during spatial listening

Subjects that displayed the largest difference between leftward and rightward gaze position during
spatially cued trials showed the greatest activation in the posterior SPL and superirolateral occipital
cortex. This suggests that the posterior SPL mediates the spatial orientation of visual fixation during
spatial listening. This has implications for studies attempting to isolate the cortical networks
supporting auditory attention (Hallett et al. 1999; Shomstein and Yantis 2006; Salmi et al. 2007;
Kong et al. 2012; Salmi et al. 2009) as our data suggest that the amount of activation in the SPL
corresponds with how much subjects biased their eye movements during spatial listening. It is
noteworthy however that the SPL regions associated with spatial gaze bias in the present study did
not overlap with the listening task activations (Figure 4), but did fall within the DAN as defined
using a visuospatial orienting task (Figure 5). Previous visual studies have shown that gaze position
is encoded in the posterior parietal lobe (Williams and Smith 2010), and crossmodal audiovisual
salience maps were located in the SPL (Nardo et al. 2013). It is possible that the SPL mediates the
supra-modal formation of spatial maps, partly through the cuing of gaze position (Nardo et al. 2013).
The recruitment of the SPL during auditory spatial attention may therefore reflect specifically this
cross-modal tuning of spatial maps, indicating that the SPL’s role in auditory attention may also be
via the indirect route. The SPL is a candidate for mediating this cross-modal tuning, as it is activated
during visual and auditory spatial searching (Corbetta et al. 2008; Hill and Miller 2010). In addition,
both the SPL and FEF are more strongly activated with auditory spatial tasks than auditory tasks
involving pitch discrimination (Maeder et al. 2001; Hill and Miller 2010).

4.4 FEF activity correlates with saccade rate inhibition during attentive listening

When we probed the brain systems that might mediate crossmodal saccade inhibition effects, subjects
that displayed the largest saccade rate difference between rest and task trials showed higher FEF
activity during the auditory cue period (Figure 4). These activations overlapped with both the
auditory task activations (Figure 4) and the DAN as defined by a visual task (Figure 5). No individual
differences in brain activity were observed during the Task period. There are different reasons why
this may have been the case. For one, it is not clear whether the inhibition of saccades, or in other
words the control of fixation duration should be expected to cause increased activity in higher-order
cognitive networks (Henderson and Choi, 2015). The FEF have been proposed to mediate the
planning of subsequent saccades (Isoda and Tanji 2003; Hu and Walker 2011). As such it is possible
that the observed FEF activity in the cue period represents the planning of or preparation for the
inhibition of the rate of saccades during the upcoming trial. It is also possible that the analysis
techniques deployed were not sensitive to what may be subtle neural correlates of inhibiting
Eye movements during auditory attention

saccades. Nonetheless, the finding of elevated FEF activity in those subjects which were prone to
showing overt oculomotor differences again suggests that the FEF may be, in part, mediating
crossmodal factors during listening, even though this is unlikely to be its only role (Bharadwaj, et al.,
2014). For example, the FEF shows elevated activity when attention is maintained to spatial
locations, even if those locations fall outside the visual field (Tark and Curtis 2009). Previous work
also suggests that the FEF mediates internally guided saccades, while the SPL is involved in both
internally and visually guided saccades (Bender et al., 2013).

4.5 The DAN mediates auditory task-evoked eye movement control

The activation patterns obtained for saccade inhibition and gaze position were located predominantly
within the DAN, and each displayed widespread functional connectivity with the remaining DAN
regions (Figure 5). This suggests that the regions associated with eye control during auditory
attention form core parts of the DAN, as shown by their intrinsic connectivity. The gaze position seed
in the SPL produced a posterior-loaded DAN which had stronger connectivity with visual regions.
The saccade inhibition seed in the FEF produced a more front-loaded DAN with increased prefrontal
cortex connectivity, but also interestingly with stronger connectivity to auditory regions in the
temporal lobes. One interpretation is that the FEF communicates intrinsically with both auditory and
visual regions, which makes it a stronger candidate for an amodal center than the DAN as a whole.
However, our results suggest that the FEF influences auditory attention, at least in part, through
indirect mechanisms such as the control of eye movements (possibly through suppressing eye
movements or altering eye movement planning) rather than through direct top-down control, e.g. the
modulation of auditory receptive fields (Fritz et al. 2010). It is possible, although speculative, that the
observed auditory effects on eye movement are primarily mediated by the functional connectivity
between auditory regions and the FEF, which then exerts an effect on the SPL via its strong
functional connectivity (i.e. the DAN). The SPL may then mediate crossmodal spatial orienting, in
part through the cuing of gaze position.

The DAN has been shown to be active under a variety of task conditions, including auditory attention
(Corbetta, et al., 2008). Further evidence for an auditory role comes from findings that the FEF can
show frequency-tagged responses to sounds (Bharadwaj, et al., 2014). The present results do not rule
out that the DAN plays an important part in attentional orienting to all modalities. Or indeed that the
DAN, and in particular the FEF, may have multiple roles. Rather, the present findings suggest that
one of the roles that the DAN does perform is to orient the visual system to comply with auditory
task demands. This might explain why the DAN is activated during orienting to both auditory and
visual stimuli, but is not present during the maintenance of attention to auditory stimuli (Salmi et al,
2007; 2009). This visual orienting process may be intrinsic to auditory orienting, particularly
considering how interlinked the auditory and visual systems are (Driver and Spence, 1998), and that
there is competition for resources between the systems (Saults and Cowan, 2007). Given that
crossmodal processes are likely to play a large part in successful attentional orienting, it is likely that
the DAN is crucial for auditory attention. However, in order to establish the DAN’s exact role in
listening we propose that the distinction between direct and indirect mechanisms merits further study.
For example, this could help make sense of why stroke lesions resulting in visual neglect often do not
lead to auditory neglect for spectral features, but do impair auditory localization (Pavani, et al.,
2002). In this case, damage to the DAN has a dramatic effect on supramodal processes such as spatial
orienting, but would not compromise non-spatial auditory processes as its contribution to listening is
an indirect, and therefore not necessary one. Thus the location of lesions within the SPL (and not
FEF) would largely determine the selective deficits.
Eye movements during auditory attention

In conclusion, this study shows that auditory attention induces overt eye movements, and that these eye movement effects are mediated by activity in core components of the DAN, the SPL and FEF. Our data suggest that the activation of DAN regions during auditory attention is at least partly attributed to oculomotor control. This is evidence for the DAN being indirectly involved in auditory attention. However, we do not rule out that the DAN plays a larger role in auditory attention, though it remains to be determined whether this remaining role is also indirect (e.g., through covert visual system modulation without eye movements) or direct (e.g., through the modulation of auditory receptive fields). As auditory attention is associated with the inhibition of non-auditory sensory inputs (Langner et al. 2011), it is likely that DAN activation is essential to auditory attention, even if its role is predominantly to modulate and limit interference from visual input. Nonetheless, our data adds to the growing body of evidence (Braga et al. 2013; Seydell-Greenwald et al. 2013, Maeder, et al., 2001; Salmi, et al., 2007, Michalka et al. 2015) that the role of the DAN in auditory attention is not as clear as in vision. This evidence suggests that there may be parallel but interacting networks for attention to visual and auditory modalities (Salmi et al. 2007; Braga et al. 2013, Michalka et al. 2015), whose mechanisms should be further studied. We propose that the dorsal frontoparietal network mediates intrinsic but crossmodal aspects of auditory attention by virtue of its predominantly visual role.

5 Acknowledgments

This research was funded by the Medical Research Council of the UK and the Wellcome Trust. We wish to thank Laura Buck and Yulia Nigmatullina for assistance with the eye tracking analysis, and all our participants for taking part. A subset of the data analyzed was provided by the Human Connectome Project, WU-Minn Consortium (Principal Investigators: David Van Essen and Kamil Ugurbil; 1U54MH091657) funded by the 16 NIH Institutes and Centers that support the NIH Blueprint for Neuroscience Research; and by the McDonnell Center for Systems Neuroscience at Washington University.

6 References

Alain C, Arnott SR, Hevenor S, Graham S, Grady CL (2001) "What" and "where" in the human auditory system. Proc Natl Acad Sci U S A 98:12301-12306.

Alho K, Salmi J, Koistinen S, Rinne T. (2015) Top-down controlled and bottom-up triggered orienting of auditory attention to pitch activate overlapping brain networks. Brain Research 1626:136–145.

Beckmann CF, Jenkinson M, Smith SM. (2003) General multilevel linear modeling for group analysis in FMRI. Neuroimage. 20:1052–1063.

Beckmann CF, Mackay CE, Filippini N, Smith SM. (2009) Group comparison of resting-state FMRI data using multi-subject ICA and dual regression. Hum Brain Map Conf 181.

Bender J, Tark KJ, Reuter B, Kathmann N, Curtis E (2013) Differential roles of the frontal and parietal cortices in the control of saccades. Brain Cogn 83(1):1-9.

Bharadwaj, H.M., Lee, A.K., Shinn-Cunningham, B.G. (2014) Measuring auditory selective attention using frequency tagging. Frontiers in integrative neuroscience, 8:6.

Behrmann M. (2004) Parietal cortex and attention. Curr Opin Neurobiol. 14:212–217.
Eye movements during auditory attention

Braga RM, Wilson LR, Sharp DJ, Wise RJS, Leech R. (2013) Separable networks for top-down attention to auditory non-spatial and visuospatial modalities. Neuroimage. 74:77–86.

Büttner-Ennever JA, Horn AK. (1997) Anatomical substrates of oculomotor control. Curr Opin Neurobiol. 7:872–879.

Cabeza, R., Nyberg, L. (2000) Imaging cognition II: An empirical review of 275 PET and fMRI studies. Journal of cognitive neuroscience, 12:1-47.

Chiou, R., Rich, A.N. (2012) Cross-modality correspondence between pitch and spatial location modulates attentional orienting. Perception, 41:339-53.

Corbetta M, Patel G, Shulman GL. (2008) The reorienting system of the human brain: from environment to theory of mind. Neuro. 58:306–324.

Driver, J., Spence, C. (1998) Cross-modal links in spatial attention. Philosophical transactions of the Royal Society of London. Series B, Biological sciences, 353:1319-31.

Fritz JB, David SV, Radtke-Schuller S, Yin P, Shamma SA. (2010) Adaptive, behaviorally gated, persistent encoding of task-relevant auditory information in ferret frontal cortex. Nat Neurosci. 13:1011–1019.

Gori M, Sandini G, Martinoli C, Burr DC. (2014) Impairment of auditory spatial localization in congenitally blind human subjects. Brain. 137:288–293.

Hallett M, Bushara KO, Weeks RA, Ishii K, Catalan M-J, Tian B, Rauschecker JP. (1999) Modality-specific frontal and parietal areas for auditory and visual spatial localization in humans. Nat Neurosci. 2:759–766.

Henderson, J.M., Choi, W. (2015) Neural Correlates of Fixation Duration during Real-world Scene Viewing: Evidence from Fixation-related (FIRE) fMRI. Journal of cognitive neuroscience, 27:1137-45.

Hickok G, Buchsbaum B, Humphries C, Muftuler T. (2003) Auditory-motor interaction revealed by fMRI: speech, music, and working memory in area Spt. J Cogn Neurosci. 15:673–682.

Hill KT, Miller LM. (2010) Auditory attentional control and selection during cocktail party listening. 20:583–590.

Hu Y, Walker R. (2011) The neural basis of parallel saccade programming: an fMRI study. J Cogn Neurosci. 23:3669–3680.

Isoda M, Tanji J. (2003) Contrasting neuronal activity in the supplementary and frontal eye fields during temporal organization of multiple saccades. J Neurophysiol. 90:3054–3065.

Kastner S, Pinsk MA, De Weerd P, Desimone R, Ungerleider LG. (1999) Increased activity in human visual cortex during directed attention in the absence of visual stimulation. Neuron. 22:751–761.

Kerzel D, Born S, Souto D. (2010) Inhibition of steady-state smooth pursuit and catch-up saccades by abrupt visual and auditory onsets. J Neurophysiol. 104:2573–2585.

This is a provisional file, not the final typeset article
Eye movements during auditory attention

Kincade JM, Abrams RA, Astafiev SV, Shulman GL, Corbetta M. (2005) An event-related functional magnetic resonance imaging study of voluntary and stimulus-driven orienting of attention. J Neurosci. 25:4593–4604.

Kong L, Michalka SW, Rosen ML, Sheremata SL, Swisher JD, Shinn-Cunningham BG, Somers DC. (2012) Auditory Spatial Attention Representations in the Human Cerebral Cortex. Cereb Cortex. 24:773-784.

Langner R, Kellermann T, Boers F, Sturm W, Willmes K, Eickhoff SB. (2011) Staying responsive to the world: Modality-specific and -nonspecific contributions to speeded auditory, tactile, and visual stimulus detection. 33:398–418.

Lee, A. K. C., Rajaram, S., Xia, J., Bharadwaj, H., Larson, E., Hämäläinen, M. S., & Shinn-Cunningham, B. G. (2013) Auditory Selective Attention Reveals Preparatory Activity in Different Cortical Regions for Selection Based on Source Location and Source Pitch. Frontiers in Neuroscience. 6:Article 190.

Leech R, Braga RM, Sharp DJ. (2012) Echoes of the brain within the posterior cingulate cortex. Journal of Neuroscience. 32:215–222.

Lewis J.W., Beauchamp M., DeYoe E.A. (2000) A comparison of visual and auditory motion processing in human cerebral cortex. Cereb Cortex 10:873-888.

Macaluso E. (2010) Orienting of spatial attention and the interplay between the senses. Cortex. 46:282–297.

Maddox, R. K., Pospisil, D. A., Stecker, G. C., & Lee, A. K. C. (2014) Directing Eye Gaze Enhances Auditory Spatial Cue Discrimination. Current biology. 24:748-752.

Maeder PP, Meuli RA, Adriani M, Bellmann A, Fornari E, Thiran JP, Pittet A, Clarke S. (2001) Distinct pathways involved in sound recognition and localization: a human fMRI study. Neuroimage. 14:802–816.

Michalka SW, Kong L, Rosen ML, Shinn-Cunningham BG, Somers DC (2015) Short-term memory for space and time flexibly recruit complementary sensory-biased frontal lobe attention networks. Neuron 87(4):882-892.

Moore T, Armstrong KM, Fallah M. (2003) Visuomotor origins of covert spatial attention. Neuron. 40:671–683.

Muggleton NG, Kalla R, Juan C-H, Walsh V. (2011) Dissociating the contributions of human frontal eye fields and posterior parietal cortex to visual search. J Neurophysiol. 105:2891–2896.

Nardo D, Santangelo V, Macaluso E. (2013) Spatial orienting in complex audiovisual environments. Hum Brain Mapp. 35: 1597-1614.

Nobre, A.C., Gitelman, D.R., Dias, E.C., Mesulam, M.M. (2000) Covert visual spatial orienting and saccades: overlapping neural systems. NeuroImage, 11:210-6.

Pages DS, Groh JM. (2013) Looking at the ventriloquist: visual outcome of eye movements
Eye movements during auditory attention

615 calibrates sound localization. PLoS ONE. 8:e72562.

616 Paulsen J, Ewertsen HW. (1966) Audio-visual reflex: determination of the audio-visual reflex in
directional hearing by employment of electroneystagmography. Acta Otolaryngol. Suppl224:211.

618 Pavani F, Husain M, Driver J. (2008) Eye-movements intervening between two successive sounds
disrupt comparisons of auditory location. Exp Brain Res. 189:435–449.

620 Pavani, F., Ladavas, E., Driver, J. (2002) Selective deficit of auditory localisation in patients with
visuospatial neglect. Neuropsychologia, 40:291-301.

623 Petit L., Simon G., Joliot M., Andersson F., Bertin T., Zago L., Mellet E., Tzourio-Mazoyer N.
(2007) Right hemisphere dominance for auditory attention and its modulation by eye position:
An event-related fMRI study. Restorative Neurology and Neuroscience 25:211-225.

626 Posner MI, Petersen SE. (1990) The attention system of the human brain. Annu Rev Neurosci.
13:25–42.

628 Razavi B, O'Neill WE, Paige GD. (2007) Auditory spatial perception dynamically realigns with
changing eye position. Journal of Neuroscience. 27:10249–10258.

630 Rolfs M, Engbert R, Kliegl R. (2005) Crossmodal coupling of oculomotor control and spatial
attention in vision and audition. Exp Brain Res. 166:427–439.

632 Ruff CC, Bestmann S, Blankenburg F, Bjoertomt O, Josephs O, Weiskopf N, Deichmann R, Driver J.
(2008) Distinct causal influences of parietal versus frontal areas on human visual cortex:
evidence from concurrent TMS-fMRI. Cereb Cortex. 18:817–827.

635 Salimi-Khorshidi G, Douaud G, Beckmann CF, Glasser MF, Griffanti L, Smith SM. (2014)
Automatic denoising of functional MRI data: Combining independent component analysis and
hierarchical fusion of classifiers. Neuroimage. 90:449–468.

641 Salmi J, Rinne T, Koistinen S, Salonen O, Alho K. (2009) Brain networks of bottom-up triggered and
top-down controlled shifting of auditory attention. Brain Research 1286: 155-164.

643 Saults, J.S., Cowan, N. (2007) A Central Capacity Limit to the Simultaneous Storage of Visual and
Auditory Arrays in Working Memory. Journal of experimental psychology. General, 136:663-
84.

646 Saygin AP, Sereno MI. (2008) Retinotopy and attention in human occipital, temporal, parietal, and
frontal cortex. Cereb Cortex. 18:2158–2168.

648 Seydell-Greenwald A, Greenberg AS, Rauschecker JP. (2013) Are you listening? Brain activation
associated with sustained nonspatial auditory attention in the presence and absence of
stimulation. Hum Brain Mapp. 35: 2233-2252.

651 Shomstein S, Yantis S. (2006) Parietal cortex mediates voluntary control of spatial and nonspatial
This is a provisional file, not the final typeset article
Eye movements during auditory attention

Smith SM, Beckmann CF, Andersson J, Auerbach EJ, Bijsterbosch J, Douaud G, Duff E, Feinberg DA, Griffanti L, Harms MP, Kelly M, Laumann T, Miller KL, Moeller S, Petersen S, Power J, Salimi-Khorshidi G, Snyder AZ, Vu AT, Woolrich MW, Xu J, Yacoub E, Ugurbil K, Van Essen DC, Glasser MF, WU-Minn HCP Consortium. (2013) Resting-state fMRI in the Human Connectome Project. Neuroimage. 80:144–168.

Smith SM, Jenkinson M, Woolrich MW, Beckmann CF, Behrens TEJ, Johansen-Berg H, Bannister PR, De Luca M, Drobnjak I, Flitney DE, Niazy RK, Saunders J, Vickers J, Zhang Y, De Stefano N, Brady JM, Matthews PM. (2004) Advances in functional and structural MR image analysis and implementation as FSL. Neuroimage. 23 Suppl 1:S208–S219.

Tark K-J, Curtis CE (2009) Persistent neural activity in the human frontal cortex when maintaining space that is off the map. Nat Neuro 12: 1463-1468.

Valsecchi M, Turatto M. (2009) Microsaccadic responses in a bimodal oddball task. Psychol Res. 73:23–33.

Van Essen DC, Smith SM, Barch DM, Behrens TEJ, Yacoub E, Ugurbil K, WU-Minn HCP Consortium. (2013) The WU-Minn Human Connectome Project: an overview. Neuroimage. 80:62–79.

Van Grootel TJ, Van Opstal AJ. (2009) Human sound-localization behaviour after multiple changes in eye position. Eur J Neurosci. 29:2233–2246.

Vossel S, Thiel CM, Fink GR. (2006) Cue validity modulates the neural correlates of covert endogenous orienting of attention in parietal and frontal cortex. Neuroimage. 32:1257–1264.

Warren J. (2008) How does the brain process music? Clin Med. 8:32–36.

Williams AL, Smith AT. (2010) Representation of eye position in the human parietal cortex. J Neurophysiol. 104:2169–2177.

Yuval-Greenberg S, Deouell LY. (2011) Scalp-recorded induced γ-band responses to auditory stimulation and its correlations with saccadic muscle-activity. Brain Topogr. 24:30–39.

Zahn JR, Abel LA, Dell'Osso LF. (1978) Audio-ocular response characteristics. Sens Processes. 2:32–37.

Zamarbieri D, Schmid R, Magenes G, Prablanc C. (1982) Saccadic responses evoked by presentation of visual and auditory targets. Exp Brain Res. 47:417–427.

Zou H, Müller HJ, Shi Z. (2012) Non-spatial sounds regulate eye movements and enhance visual search. J Vis. 12.

7 Figure legends
Eye movements during auditory attention

Figure 1: Schematic of task design. Two competing melodies were presented simultaneously using either dichotic (one melody played to each ear) or diotic (both melodies played to both ears equally) listening. Subjects were cued to listen to the melody on the left or right ear (“Left” or “Right”), or to listen to the melody that was higher or lower in pitch (“High” or “Low”). After the melodies were completed, subjects were cued to Respond (RESP) with button presses to indicate whether they heard a pitch change in the cued melody.

Figure 2: Auditory spatial attention influences natural gaze position. A) Plot of gaze position for a single subject over all spatially cued (“Listen-Left” or “Listen-Right”) trials. B) Gaze position for all subjects grouped by cue and target sound location. Spatial cues influenced gaze position during the subsequent listening task. No gaze biases were observed for spectral discriminations (“Listen-High” or “Listen-Low”) regardless of whether stimuli were presented dichotically (“R” and “L”) or diotically (“C”). Positive values on y-axis represent rightward, and negative values represent leftward gaze displacements. Displacement axes are in arbitrary units.

Figure 3: Individual variability in auditory task-evoked eye movement control and performance. Across all 20 subjects, A) attentive listening following spatial cues resulted in a gaze position bias towards cued sound, and B) attentive listening reduced saccade rate compared to interleaved rest trials. C) Graphs showing t-value (tval) for each subject for the comparison of saccade rate (green; Task vs Rest) and mean gaze position difference (blue; Listen Left vs Listen Right trials). High inter-subject variability was found in saccade inhibition (all trials) and gaze position bias (spatially cued trials). D) Individual differences in saccade reduction and gaze position bias did not correlate to individual differences in task performance or with each other.

Figure 4: Neuroimaging correlates of auditory attention task and variability in eye movement. Orange: Across all trials, attentive listening was associated with activation of a widespread auditory network (superior temporal gyri and sulci bilaterally; STG) as well as the dorsal attention network (superior parietal lobes; SPL, frontal eye fields; FEF) and middle frontal gyri (MFG) bilaterally. However, only the superior parietal lobes and frontal eye fields were associated with individual differences in eye movement control. Blue: Subjects that displayed the largest left-right gaze position bias following spatial auditory cues had greater activation of the SPL during the attentive listening periods. Green: Subjects that displayed the largest reduction in saccade rate during listening trials compared to rest had greater activation of the FEF and left posterior temporal lobe during the cue period that preceded each trial. Colorbars show cluster-corrected z-scores.

Figure 5: Auditory task-evoked eye movement regions overlap with dorsal attention network (DAN). Orange and Red: A visual cued saccade task was used to functionally define the DAN in our dataset. The regions (from figure 4) associated with increased gaze position bias (blue) and saccade inhibition (green) were used as seeds to assess intrinsic functional connectivity (FC) in the resting state. Both FC maps overlapped considerably with the DAN. The saccade inhibition FC map also revealed stronger connectivity with anterior prefrontal and auditory regions in the temporal lobe. The gaze position bias FC map also revealed strong connectivity with the dorsal and ventral visual streams.

Table 1: MNI coordinates for local maxima of task activation map and individual differences in eye movement maps shown in Figure 4. R, right; L, Left; STG, superior temporal gyrus; planum t, planum temporale; FEF, frontal eye fields; MFG, middle frontal gyrus; p, posterior; Sup, superior; Lat, lateral; Occ, Occipital.
Eye movements during auditory attention

| Auditory Task Peaks | x   | y   | z  |
|---------------------|-----|-----|----|
| R STG               | 70  | -28 | 12 |
| R planum t          | 46  | -32 | 12 |
| L planum t          | -60 | -30 | 8  |
| L STG               | -62 | -34 | 10 |

| Saccade Effect Peaks | x   | y   | z  |
|----------------------|-----|-----|----|
| R FEF                | 44  | 2   | 54 |
| L FEF                | -34 | 0   | 44 |
| R-MFG                | 56  | 22  | 28 |
| L MFG                | -46 | 4   | 52 |
| L STG                | -56 | -34 | 26 |
| L p Insula           | -50 | -22 | 12 |
| L pMTG               | -54 | -56 | 4  |

| Gaze Position Effect | x   | y   | z  |
|----------------------|-----|-----|----|
| R Sup Occ            | 52  | -78 | 16 |
| R SPL                | 18  | -78 | 50 |
| R Lat Occ            | 34  | -86 | 24 |
| L Lat Occ            | -28 | -88 | 24 |
| L Sup Occ            | 12  | -86 | 40 |
**Dichotic**

Cue: "Right" "Left"  "High" "Low"

---

**Diotic**

Cue: "High" "Low"
