Emotion modulates activity in the 'what' but not 'where' auditory processing pathway

James H. Kryklywyd,e, Ewan A. Macpherson c,f, Steven G. Greening b,e, Derek G.V. Mitchell a,b,d,e,*

1 Department of Psychiatry, University of Western Ontario, London, Ontario N6A 5A5, Canada
2 Department of Anatomy & Cell Biology, University of Western Ontario, London, Ontario N6A 5A5, Canada
3 National Centre for Audiology, University of Western Ontario, London, Ontario N6A 5A5, Canada
4 Graduate Program in Neuroscience, University of Western Ontario, London, Ontario N6A 5A5, Canada
5 Brain and Mind Institute, University of Western Ontario, London, Ontario N6A 5A5, Canada
6 School of Communication Sciences and Disorders, University of Western Ontario, London, Ontario N6A 5A5, Canada

Abstract

Auditory cortices can be separated into dissociable processing pathways similar to those observed in the visual domain. Emotional stimuli elicit enhanced neural activation within sensory cortices when compared to neutral stimuli. This effect is particularly notable in the ventral visual stream. Little is known, however, about how emotion interacts with dorsal processing streams, and essentially nothing is known about the impact of emotion on auditory stimulus localization. In the current study, we used fMRI in concert with individualized auditory virtual environments to investigate the effect of emotion during an auditory stimulus localization task. Surprisingly, participants were significantly slower to localize emotional relative to neutral sounds. A separate localizer scan was performed to isolate neural regions sensitive to stimulus location independent of emotion. When applied to the main experimental task, a significant main effect of location, but not emotion, was found in this ROI. A whole-brain analysis of the data revealed that posterior-medial regions of auditory cortex were modulated by sound location; however, additional anterior-lateral areas of auditory cortex demonstrated enhanced neural activity to emotional compared to neutral stimuli. The latter region resembled areas described in dual pathway models of auditory processing as the 'what' processing stream, prompting a follow-up task to generate an identity-sensitive ROI (the 'what' pathway) independent of location and emotion. Within this region, significant main effects of location and emotion were identified, as well as a significant interaction. These results suggest that emotion modulates activity in the 'what,' but not the 'where,' auditory processing pathway.

© 2013 Elsevier Inc. All rights reserved.

Introduction

The ability to interact effectively in an environment requires the accurate recognition and localization of surrounding objects and the capacity to prioritize these objects for behavior. One characteristic known to modulate this is the emotional nature of the stimuli (Adolphs, 2008; Lang and Davis, 2006; Pessoa and Ungerleider, 2004; Vuilleumier, 2005). Considerable evidence suggests that emotional visual stimuli gain rapid and often preferential access to the brain's processing resources. At the behavioral level, emotional visual stimuli are detected faster than neutral stimuli (Graves et al., 1981), are more likely to enter into awareness (Amting et al., 2010; Mitchell and Greening, 2012) and can cause significantly greater influence on task-relevant behaviors (Mitchell et al., 2008; Vuilleumier and Driver, 2007). These effects are thought to be conferred by enhanced sensory processing; thus, in the visual domain, emotional stimuli elicit greater activity than similar neutral stimuli within areas of the visual cortex (Morris et al., 1998; Vuilleumier and Driver, 2007). Just like in the visual domain, studies of auditory processing have demonstrated that the analysis of emotional auditory stimuli occurs rapidly (Gydeke et al., 2004; Sauter and Eimer, 2009) and is associated with enhanced activity in sensory (i.e., auditory) cortices (Pecteau et al., 2007; Viinikainen et al., 2012). Despite some emerging work concerning the influence of emotion on the representation of auditory objects, essentially nothing is known about how emotion influences auditory stimulus localization. There is accumulating evidence that auditory processing occurs within two separate cortical streams (Ahveninen et al., 2006; Alain et al., 2001; Barrett and Hall, 2006; Clarke et al., 2002; Lomber and Malhotra, 2008; Mathiak et al., 2007; Rauschecker, 2012; Rauschecker and Tian, 2000) that may share some similarities with the well-established dorsal and ventral processing streams of the visual system (Haxby et al., 1991; Milner and Goodale, 1993). Spatial cues used for localization are processed primarily in posterior-medial regions of the auditory cortex (Arnott et al., 2004; Bushara et al., 1999; Lomber...
et al., 2007) including the posterior superior temporal gyrus (STG) and the transverse temporal gyrus. In contrast, sound identity cues, including pitch and language features, are processed in anterior-lateral regions of auditory cortex along the anterior STG (Altmann et al., 2008; Warren and Griffiths, 2003). However, despite continuous advances toward understanding the neural mechanisms underlying both enhanced representation of emotion within sensory cortices and our representations of auditory space, the impact of emotion during auditory localization remains unknown. Specifically, it remains unclear whether evidence of enhanced activity observed in prior studies to emotional relative to neutral, non-spatialized auditory stimuli (Fecteau et al., 2007; Viinikainen et al., 2012) would also translate into enhanced auditory stimulus localization and enhanced activity in areas of auditory cortex sensitive to object location.

The potential of auditory virtual environments (AVEs) as a method to examine neural pathways associated with auditory stimulus localization has been described in previous studies (Bohil et al., 2011; Fujiki et al., 2002; Langendijk and Bronkhorst, 2000; Wightman and Kistler, 1989a,b). Previous neuroimaging studies investigating auditory localization have created AVEs using generic head-related transfer functions (HRTFs) generated from measurements of mannequins or a prototypical head shape (Alveninen et al., 2006; Bushara et al., 1999; Krumhholz et al., 2009). These, however, fail to accommodate individual differences in head size and pinnae structure that alters sound as it enters the ear canals, resulting in imperfect perception of spatialized sounds (Middlebrooks et al., 2000; Wenzel et al., 1993). Such variables have been shown to influence reactions to and ratings of emotional auditory stimuli (Vastfjall, 2003). Despite its potential importance, we are not aware of any neuroimaging studies utilizing unique AVEs created from individualized HRTFs.

In the present study, we investigated whether the emotion-related enhancements observed in the visual domain at the behavioral (Anting et al., 2010; Graves et al., 1981) and neural levels (Morris et al., 1998; Vuilleumier and Driver, 2007) would also be found during auditory stimulus localization. We hypothesized that positive and negative auditory cues would receive prioritized processing relative to neutral stimuli. We predicted that this prioritization would be reflected by increased accuracy, decreased reaction time, and increased neural activity within the posterior-medial ‘where’ pathways of auditory processing during the localization of emotional compared to neutral sounds. However, as will be described below, the data did not fit this prediction, and instead we found slower response times for emotional auditory stimuli compared to neutral ones. Additionally, consistent with previous studies involving non-spatialized emotional auditory cues (Fecteau et al., 2007), we predicted that anterior-lateral areas of auditory cortex (i.e., the putative ‘what’ processing pathway) would also show enhanced activity for emotional compared to neutral sounds. Furthermore, in light of lesion data suggesting that the what/where pathways are doubly dissociable (Lomber and Malhotra, 2008), we predicted that anterior-lateral regions of auditory cortex would not be modulated by sound location.

To test these predictions, we created AVEs by generating sounds based on each individual’s unique HRTFs. While undergoing fMRI, participants located or identified a series of auditory stimuli presented in these virtual environments. The current study consisted of three related tasks. Task 1 was designed as a functional localizer, aimed at independently identifying ROIs specifically related to sound localization while controlling for object identity. Task 2 was conducted in the same scanning session as Task 1. In this task, participants were required to identify the source locations of positive, negative and neutral sounds presented within a virtual auditory environment. This task served two purposes. First, the ‘where’ ROI derived from the Task 1 localizer was applied to the data in Task 2 and interrogated to determine potential effects of emotion on location-sensitive areas of auditory cortex. Second, Task 2 allowed us to perform an exploratory whole-brain analysis examining the effects of, and interactions between, emotion and location during auditory stimulus localization. Contrary to expectations, the results showed that emotion did not modulate regions of auditory cortex sensitive to location. However, a distinct region of anterior lateral temporal cortex identified in this exploratory study was modulated by emotion. This area strongly resembled regions associated with sound-identity processing in previous studies (i.e., the putative ‘what’ pathway; Barrett and Hall, 2006; Warren and Griffiths, 2003). To help determine the degree to which this area could be characterized as part of the ‘what’ auditory pathway, a follow-up localizer was conducted in a subset of participants in a subsequent session. This functional ‘what’ pathway localizer identified ROIs that were modulated by sound identity while location and emotion were held constant. The resulting ROI was extracted and applied to the data generated from Task 2, allowing us to independently test the effects of emotion on the resulting ‘what’ pathway.

Methods

Subjects

Eighteen healthy human subjects, (9 male, 9 female) with a mean age of 23.56 (range 19–35, SD 4.51), completed Tasks 1 and 2. All subjects granted informed consent and were in good mental health, as assessed by a Structured Clinical Interview for DSM-IV (Diagnostic and Statistical Manual of Mental Disorders, 4th Edition). All subjects had normal hearing, normal or corrected-to-normal vision and were fluent English speakers. Ten of these subjects (5 male, 5 female), with a mean age of 24.3 (range 19–35, SD 5.42), returned to complete Task 3. All participants were reimbursed for their time at the end of the study. All experiments were approved by the Health Science Research Ethics Board at the University of Western Ontario.

Stimuli and apparatus

Stimuli

Twelve sounds were chosen from the International Affective Digitized Sound (IADS) stimuli set that were of a neutral, negative or positive affective nature as defined by standardized ratings (Bradley and Lang, 1999). Each stimulus category contained two single-source non-verbal human vocalizations, one multi-source non-verbal human vocalization, and one non-human sound. All sounds were presented with a variable duration of 2000–3000 ms (balanced across stimuli; variable durations were used to facilitate deconvolution of the BOLD signal). Importantly, all stimuli were matched for their onset amplitude and root mean-square amplitude, which ensures that the power and energy were consistent. Positive and negative stimuli were balanced for arousal ratings (mean positive = 6.58, mean negative = 6.92) and valence levels (positive = 7.56, negative = 2.36, absolute neutral = 5). In addition, to create a novel unrecognizable noise for Task 1, the 12 task sounds of Task 2 were merged into a single audio file, segmented into <3 ms fragments, and subsequently scrambled, reconstituted and cropped to a duration of 15,000 ms. This sound maintains the average long-term power spectrum of the stimulus set of Task 2, while remaining unidentifiable.

In order to localize neural regions that were sensitive to stimulus identity, a novel set of nine neutral sounds were chosen from the IADS (mean valence 5.28, SD 0.98) for use in Task 3. These nine sounds were human, animal, or machine in origin (3 of each). An additional three segments of scrambled noise (identical to that used in Task 1) comprised a fourth sound class. All sounds in this set were 5000 ms in duration, and were matched for onset amplitude and root-square mean amplitude.

Auditory virtual environment

Throughout the experiment, all sounds were presented within an auditory virtual environment through Sensimetric MRI-Compatible
Insert Earphones. Volume was controlled by a Dayton Audio Class T Digital Amplifier. Initial volume was set to 88–90 dBs and adjusted slightly to the comfort level of each individual participant. To induce the perceptual experience of spatialized sounds using insert-style headphones, HRTFs were measured individually for each subject.

To obtain the HRTF measurements, miniature electret microphones (Knowles FG3629) were mounted facing outsides in foam earplugs inserted flush with the ear canal entrances. The participant stood on an adjustable platform which positioned his or her head at the height of a 1.5-m radius array of 16 loudspeakers (Tannoy i5 AW) spaced in 22.5-degree intervals around the listener. The array was located in a large hemi-anechoic chamber, and the floor area within the array was covered with acoustical foam to attenuate reflections. The impulse response from each loudspeaker to each ear microphone was measured using the averaged response to 32 periods of a 2047-point maximum-length sequence signal (Rife and Vanderkooy, 1989) presented at a sampling rate of 48,828 Hz via a Tucker Davis Technologies RX6 real time processor and QSC CX168 power amplifiers. During the measurement procedure, head motion was minimized by monitoring head position with an electromagnetic tracker (Polhemus FASTRAK) while participants were asked to aim the light from a head-mounted LED accurately at a frontal target position. To correct for individual loudspeaker characteristics, each HRTF measurement was equalized in the frequency domain by dividing by the appropriate loudspeaker transfer function measured with a reference microphone (Brueel & Kjaer 4189) placed at the center of the array in the absence of the listener’s head. The impulse responses were windowed in post-processing to remove any residual reflections. The resulting HRTF measurements were resampled to 44.1 kHz and combined with the equalization filters for the headphones supplied by the manufacturer to create a new set of auditory filters limited to the 10-kHz bandwidth of the headphones. These individualized filters were then applied to each sound by time-domain convolution to create the experience of a virtual 3-dimensional auditory space; sounds presented with headphones were perceived to originate from controlled physical locations external to the participants.

Sounds for Tasks 1 and 2 were spatialized to four locations along the horizontal plane (−90°, −22.5°, 22.5° and 90°) along the sagittal midline; negative = left). Sounds for Task 3, along with an additional 500-ms artificially generated white noise were spatialized to a location directly along the sagittal midline, in front of the listener. The entire set of stimuli (across locations, sound sources, and listeners) was normalized such that each stimulus had the same root-mean-square level computed on the concatenated left- and right-ear signals.

Procedure

Task 1: isolating location-sensitive regions of auditory cortex (the ‘where’ pathway)

Prior to entering the scanner, participants were acclimatized to the auditory virtual environment by completing modified versions of task 1 and task 2. For a detailed review of these training tasks, see Supplementary materials. In Task 1, 18 participants (9 female, 9 male) performed a pure auditory localization task to help identify neural regions involved in representing auditory space. This task was designed as a functional localizer allowing us to generate an ROI corresponding to the ‘where’ pathway of auditory processing. Participants were instructed to close their eyes and listen to a series of 15,000-ms scrambled sounds. Each sound was presented in one of the four virtual locations (8 times per location for a total of 32 events broken into 2 runs). Participants were instructed to actively fixate on the sound location for the duration of its presentation, and to indicate with a button press the location of the perceived source during the sound presentation. Between each sound presentation, there was a 15,000-ms period of silence. Order of presented sound location was randomized for each run.

Task 2: locating emotional sounds within auditory space

In Task 2, the same 18 participants (9 female, 9 male) completed an emotional auditory localization task. Participants were asked to close their eyes for the duration of each scan to reduce possible confounds of visual feedback. Each trial began with a white noise burst (500 ms) spatialized to a location along the sagittal midline, in front of the participant. This acted as an ‘auditory fixation cross’ to reorient the subject’s attention to a standard central point and was immediately followed by a spatialized target sound (2000–3000 ms) and a period of silence (2000–3000 ms). The target was randomly presented in one of the four possible locations. The participant was required to locate the target sound with a button press indicating the perceived location as quickly and accurately as possible. Over the course of a single run, participants heard each of the 12 sounds in every location a single time for a total of 48 trials (16 trials per emotion condition). Additionally, there were 16 silent ‘jitter’ trials incorporated into each run, for a total of 64 trials per run. All trials were selected in a random order within each run. The task run was repeated six times for a total of 384 trials.

Task 3: isolating identity-sensitive regions of auditory cortex (the ‘what’ pathway)

Task 3 was conducted during a follow-up scanning session in a subset of the original sample to identify auditory areas associated with auditory object identification. This additional functional localizer scan was initiated in response to the results of Task 1 and 2, outlined below, in order to better explore the potential doubly dissociable effects of emotion on the auditory ‘what’ and ‘where’ processing pathways. Specifically, it was designed to independently derive functional ROIs corresponding to the ‘what’ pathway of auditory processing. This type of cross-validation, using separate sources for region identification and signal estimation, avoids problems of circularity in interpreting neuro-imaging results (Vul and Pashler, 2012), allowing us to interrogate the ROIs over the time course of Task 2 without fear of statistical bias (Esterman et al., 2010; Vul and Kanwisher, 2010). In the scanner, 10 participants (5 female, 5 male) performed a block design auditory identification task similar to that used in Task 1. During this task, each 15,000-ms sound presentation consisted of a triad of stimuli belonging to a single object class presented directly in front of the listener along the mid-sagittal plane within the AVE. There were four possible object classes: human, animal, machine and scrambled sounds. Over the course of two runs, each object class was presented 8 times, for a total of 32 events. Participants were instructed to actively attend to the identity of the sound objects for the duration of their presentation, and to indicate the identity of the presented sounds (human, animal, machine or scrambled) via button press. Between the presentations of each triad, there was a 15,000-ms period of silence. Presentation order was randomized.

Behavioral data analysis

To investigate the possible effect of emotional content on sound localization, we recorded reaction times and accuracy level for the duration of Task 2. A 4 (location) × 3 (emotion) ANOVA was conducted for each of the behavioral measures. Follow-up pair-wise t-tests were performed to delineate the nature of any significant effects.

Imaging

MRI data acquisition

Subjects were scanned during all task performances using a 3 T Siemens Scanner with a 32-channel head coil. fMRI images were taken with a T2*-gradient echo-planar imaging sequence (repetition time [TR] = 2500 ms, echo time [TE] = 36 ms; field of view [FOV] = 18.8 cm, 78 × 78 matrix). Tasks 1 and 2 took place over a single scan session. Task 3 took place in a separate session. For all
Table 1

Results of Experiment 1: accuracy and time to localization for all regressor conditions.

<table>
<thead>
<tr>
<th>Emotion</th>
<th>Location</th>
<th>Time to localization (ms; correct trials only)</th>
<th>Localization accuracy (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Positive</td>
<td>−90°</td>
<td>775.1 (242.2)</td>
<td>83.8 (19.1)</td>
</tr>
<tr>
<td></td>
<td>−22.5°</td>
<td>756.2 (199.2)</td>
<td>85.9 (16.6)</td>
</tr>
<tr>
<td></td>
<td>22.5°</td>
<td>804.5 (199.4)</td>
<td>88.2 (14.1)</td>
</tr>
<tr>
<td></td>
<td>90°</td>
<td>808.5 (251.9)</td>
<td>82.2 (20.0)</td>
</tr>
<tr>
<td>Negative</td>
<td>−90°</td>
<td>736.6 (210.0)</td>
<td>87.0 (15.3)</td>
</tr>
<tr>
<td></td>
<td>−22.5°</td>
<td>776.2 (218.2)</td>
<td>84.3 (15.4)</td>
</tr>
<tr>
<td></td>
<td>22.5°</td>
<td>779.9 (210.1)</td>
<td>83.1 (18.1)</td>
</tr>
<tr>
<td></td>
<td>90°</td>
<td>733.6 (209.1)</td>
<td>84.7 (17.8)</td>
</tr>
<tr>
<td>Neutral</td>
<td>−90°</td>
<td>686.5 (153.3)</td>
<td>89.4 (12.0)</td>
</tr>
<tr>
<td></td>
<td>−22.5°</td>
<td>736.5 (194.7)</td>
<td>85.6 (20.9)</td>
</tr>
<tr>
<td></td>
<td>22.5°</td>
<td>751.7 (218.4)</td>
<td>88.2 (17.5)</td>
</tr>
<tr>
<td></td>
<td>90°</td>
<td>727.5 (224.6)</td>
<td>85.9 (18.5)</td>
</tr>
</tbody>
</table>

Standard deviations are in parentheses.

a Significant effects of emotion were identified in the reaction time data. These effects of emotion were driven by a significantly slower localization of positive and negative stimuli when compared to neutral stimuli (p < 0.001 and p < 0.05 respectively).

fMRI analysis

An analysis of the fMRI data was conducted using Analysis of Functional Neuroimages (AFNI) software (Cox, 1996) at both the individual and group levels. Motion correction was performed by registering all volumes to the relevant functional volume acquired temporally adjacent to the anatomical scan. The data set for each participant was spatially smoothed (using an isotropic 4 mm Gaussian kernel) to reduce the influence of individual differences. The time series data were normalized by dividing the signal intensity of each voxel at each time point by the mean signal intensity of that voxel for each run and multiplying the result by 100. Resultant regression coefficients represent the percent signal change from the mean activity. Regressor files modeling the presentation time course of relevant stimuli were created for each of the 12 conditions of Task 2 (4 locations × 3 emotions) during correct trials only, and for each of the four conditions of Task 1 (4 locations) and Task 3 (4 classes of objects). The relevant hemodynamic response function was fit to each regressor to perform linear regression modeling. To account for voxel-wise correlated drifting, a baseline plus linear drift and quadratic trend were modeled for each time series. This resulted in a β coefficient and t value for each voxel and regressor. To facilitate group analysis, each individual’s data were transformed into the standard space of Talairach and Tournoux. Following this, a 4 (location) × 3 (emotion) ANOVA was performed on the imaging data from Task 2, while two separate one-way ANOVAs were performed on the imaging data from Tasks 1 and 3 to examine the effects of location and sound-identity respectively (4 levels of each). ANOVAs at this level were conducted using the AFNI function GroupAna, yielding an F value for each main effect and interaction at each voxel. Percent signal change from the mean activity was extracted from significant clusters of activation for each relevant regressor using the AFNI function 3Dmaskave. To correct for multiple comparisons, a spatial clustering operation was performed using AlphaSim with 1000 Monte Carlo iterations on the whole brain EPI matrix.

Results

Behavioral results

To delineate the effects of emotion on auditory localization, participants were presented a series of positive, negative and neutral sounds within a virtual auditory environment while undergoing fMRI. Participants were instructed to localize these sounds as quickly and accurately as possible (Task 2).

A 4 (location) × 3 (emotion) ANOVA was conducted on the reaction time and error data (Table 1) to determine the impact of stimulus emotion on auditory localization at a behavioral level. This yielded a significant main effect of emotion (F(2, 34) = 12.617, p < 0.005) on the reaction times. We had originally predicted that emotion would enhance the localization of auditory stimuli. However, contrary to these predictions, follow-up t-tests revealed that this effect was driven by a
significantly slower localization of emotional sounds (both positive and negative) when compared to neutral sounds (p < 0.001 and p < 0.05 respectively). There was no significant effect of sound location (F(3, 51) = 0.616, p > 0.05) or location × emotion interaction (F(6, 102) = 1.891, p > 0.05) on reaction time. The same analysis applied to the localization error data yielded no significant effects for either sound location (F(3, 51) = 0.158, p > 0.05) or emotion (F(2, 34) = 1.783, p > 0.05), nor did it yield a significant location × emotion interaction (F(6, 102) = 1.320, p > 0.05).

**Imaging results**

**Interrogating location-sensitive regions of auditory cortex (the 'where' pathway)**

In order to determine the impact of sound location independent of sound identity, location-related activity was assessed during the localization of unrecognizable scrambled sounds (Task 1). A one-way ANOVA (four locations) conducted on the whole brain EPI data obtained during Task 1 identified a significant effect of sound location (p < 0.005; corrected at p < 0.05) within regions of the temporal cortex (Table 2). This activation included primary auditory cortex (BA 41/42) and extended medially along the transverse temporal gyrius (BA 13; Fig. 1A). This region showed greatest activity to stimuli positioned far in the contralateral hemisphere (p < 0.01; corrected at p < 0.05), while in the occipital lobe, it was not (p > 0.05). These results are illustrated in Fig. 3B.

Bilateral precuneus and occipital lobe displayed strikingly similar patterns of activation. Within both regions, the impact of location within emotion was significant during the presentation of negative sounds (p < 0.005) but not neutral sounds (p > 0.05). Within bilateral precuneus, the impact of location during the presentation of positive stimuli was significant (p < 0.05), while in the occipital lobe, it was not (p > 0.05). These results are illustrated in Fig. 3D/F. The impact of emotion within each location for both regions was also examined, yielding similar patterns of activation. Sounds coming from the contralateral hemisphere (i.e., 90° and 22.5° left of the sagittal midline) elicited greater activity in the right auditory cortex for emotional relative to neutral sounds (positive > neutral, p < 0.001; negative > neutral, p > 0.01; positive > negative, p < 0.05). These results are illustrated in Fig. 3B.

**Whole brain analysis: locating emotional sounds in auditory space**

An exploratory whole brain analysis by way of a 4 (location) × 3 (emotion) ANOVA was conducted on the data from Task 2 to identify neural regions that varied as a function of location and emotion. A significant main effect of location (p < 0.005; corrected at p < 0.05) was identified in regions of temporal and parietal cortices (Table 2). This activation extended from the primary auditory cortices (BA41/42) posteriorly and medially along the transverse temporal gyrius and into the inferior parietal lobule (Fig. 2A). Percent signal change from the mean activity for each significantly activated voxel was extracted from all regions displaying significant main effects or interactions using 3Dmaskev. Contributing to this effect, the activity in the posterior superior temporal gyrius (posterior areas of BA 13) varied as a function of location; the activity was greatest to the stimuli presented in the far contralateral hemisphere of the auditory virtual space, and decreased progressively with a distance from that location (Fig. 2B). This general pattern of activation emerged bilaterally, with each hemisphere showing an inverse pattern of activation to the other. This result is consistent with studies that describe auditory stimulus location encoding as involving differential activation of units across the two hemispheres, as opposed to local encoding within specific nuclei or regions (Grothe et al., 2010; Stecker et al., 2005). Similar patterns of activity were identified in the areas of precuneus, inferior parietal lobule, pre/post-central gyrius and medial occipital gyrius that exhibited main effects of sound location.

A significant main effect of emotion was identified in the bilateral temporal cortex (p < 0.005; corrected at p < 0.05; Table 2). Activation extended from the primary auditory cortices (BA 41/42) anterior and inferior along the superior temporal gyrius (to BA 22; Fig. 2C). This region showed significantly greater activation bilaterally during the presentation of emotional sounds (positive and negative), when compared to neutral sounds (p < 0.001 and p < 0.005 respectively; Fig. 2D).

A significant interaction was identified within regions of the right auditory cortex (BA 41/42/22; Fig. 3A), bilateral precuneus (BA 31/7; Fig. 3C), and bilateral occipital lobe (BA 17/18/19; Fig. 3E; p < 0.005, corrected at p < 0.05; Table 2; Figs. 3A–F). To identify the nature of these effects, a series of one-way ANOVAs were performed, and where significant (p < 0.05), these were followed-up with paired t-tests. First, the impact of location within each emotion was examined. The right auditory cortex showed greater activity to positive stimuli presented in the contralateral relative to ipsilateral hemisphere (p < 0.01); however, no such effect was observed for neutral or negative stimuli (p > 0.05). The impact of emotion within each location was also examined. Sounds coming from the contralateral hemisphere (i.e., 90° and 22.5° left of the sagittal midline) elicited greater activity in the right auditory cortex for emotional relative to neutral sounds (positive > neutral, p < 0.001; negative > neutral, p > 0.01; positive > negative, p < 0.05). These results are illustrated in Fig. 3B.

Interrogating identity-sensitive regions of auditory cortex (the ‘what’ pathway)

The whole-brain analysis (Task 2) identified a separate anterior-lateral area of auditory cortex that was modulated by emotion. This region resembled areas implicated in the putative ‘what’ auditory processing stream (Altmann et al., 2008; Warren and Griffiths, 2003), raising the question of whether emotion has dissociable effects on the ‘what’ versus ‘where’ auditory processing streams. To test this hypothesis, a follow-up localizer scan (Task 3) was performed to extract a sound identity sensitive ‘what’ pathway ROI (independent of location and emotion) that could then be applied to the Task 2 data. A one-way ANOVA (with four sound categories) was conducted on the EPI data revealing object-identity sensitive areas within the temporal cortex (p < 0.005; corrected at p < 0.05; Table 2). This activation included regions of cortex anterior and lateral to primary auditory cortex along the superior temporal gyrius (BA 22; p < 0.005; corrected at p < 0.05, Fig. 4A). Further exploration of this effect revealed that a significantly greater activity was elicited in this area for sounds in the human, animal and machine categories compared to scrambled sounds (p < 0.001; Fig. 4B). In addition, this region showed a significantly greater activity to biologically generated sounds (human and animal) relative to machine-generated sounds (p < 0.005, p < 0.001 respectively).
identity-related ROI was then applied to the BOLD data collected in Task 2, and the percent signal change was extracted from each mask for all locations and emotional categories using 3Dmaskave, and subjected to a 4 (location) × 3 (emotion) ANOVA. Within the ‘what’ pathway ROI, activity was found to be significantly modulated by both sound location ($F_{(3, 51)} = 26.223, p < 0.001$) and emotion ($F_{(2, 34)} = 8.914, p < 0.005$; Fig. 4c). Activity in this region was significantly increased for sounds presented in the contralateral hemisphere when compared with sounds presented in the ipsilateral hemisphere ($p < 0.001$). Additionally, this region showed enhanced activation for positive and negative versus neutral stimuli ($p < 0.005$ and $p < 0.05$ respectively). Lastly, a significant location × emotion interaction was observed in this region ($F_{(6, 102)} = 2.450, p < 0.05$).

To identify the nature of the interaction effects in the ‘what’ pathway, a series of one-way ANOVAs were performed. First, the impact of location within each emotion was examined. Within this ROI, there was a greater activity for both positive and negative stimuli presented in the contralateral relative to ipsilateral hemisphere ($p < 0.001$ and $p < 0.05$ respectively); however, no such effect was observed for neutral stimuli ($p > 0.05$). The interaction was further examined by comparing the impact of emotion within each location; however, no significant effects were identified for this contrast ($p > 0.05$). This effect contrasts with that found in the ‘where’ pathway ROI, which featured a significant effect of location, but no significant effect of emotion or location × emotion interaction. Furthermore, the coordinates of the interaction identified during Task 2 are more closely related to the coordinates of the ‘what’ ROI compared to the ‘where’ ROIs, so this effect is not unexpected. Conjunction analyses confirmed that ROIs generated from Task 1 (‘where’ localizer) and Task 3 (‘what’ localizer) overlapped with the main effects of location and emotion respectively in Task 2 (see Supplementary materials). It should be noted that Task 3 involved a subset of the original sample used in Tasks 1 and 2, and therefore had less power to define the ROI. Nevertheless, the fact that this independently derived ROI shows the same functional properties exhibited in the emotion-sensitive regions uncovered in Task 2 (despite reduced power) lends further support to the conclusion that the auditory ‘what’ pathway is modulated by both emotion and location characteristics.

**Discussion**

Considerable evidence from visual studies suggests that emotional stimuli gain rapid and often preferential access to the brain’s processing resources. Although less work has been conducted in the auditory domain, enhanced effects of emotion on auditory cortex have also been observed across multiple investigative techniques, including fMRI (Wiethoff et al., 2008) and NIRS (Plichta et al., 2011). However, essentially nothing is known about how emotion influences auditory stimulus localization. In the current study, we used fMRI in concert with individualized virtual auditory environments to investigate the effect of emotion on sound localization. Surprisingly, participants

![Neural regions significantly modulated by sound location (Task 1).](image URL)
were significantly slower to localize positive and negative sounds relative to neutral ones. Moreover, activity in an independently identified location-sensitive region of auditory cortex was not modulated by emotion. Subsequent whole-brain analyses were conducted on a task involving the localization of emotional and neutral stimuli. This analysis revealed that enhanced activity to positive and negative stimuli was observed in anterior-lateral areas of auditory cortex irrespective of location. In contrast, posterior-medial regions of auditory cortex, as well as the inferior parietal lobule and precuneus, were modulated by location, irrespective of emotion. Both the response of anterior-lateral regions of auditory cortex to sound location and the lack of response of posterior-medial region of auditory cortex to emotion ran contrary to original predictions. These unexpected results raised the possibility that emotional sounds augment activity in the anterior-lateral ‘what’ pathway, but not the posterior-medial ‘where’ pathway during auditory localization. To more clearly delineate the functional significance of the divisions identified, we conducted an additional functional localizer scan to independently identify regions of auditory cortex that responded to changes in sound identity (i.e., areas associated with encoding ‘what’ in dual pathway models of auditory processing; Alain et al., 2001; Altmann et al., 2008; Rauschecker and Tian, 2000; Warren and Griffiths, 2003). This functionally-derived ROI was then applied to the original study. Collectively, the results support the conclusion that whereas sound location modulates activity within the ‘what’ and ‘where’ functionally-derived pathways, emotion modulates neural activity in the ‘what’ but not the ‘where’ pathway.

**Regions associated with processing auditory object location**

A main effect of location was observed in bilateral temporal cortex (Tasks 1 and 2), precuneus, and inferior parietal cortex (Task 2 only). Of note, the location-related effects were characterized in posterior-medial auditory cortex by greatest activity to stimuli positioned furthest in the contralateral hemifield, decreasing activity to midline sounds, and significantly less activity to ipsilaterally positioned sounds. This general pattern of activation emerged bilaterally, with each hemisphere showing the inverse pattern of activation. ROIs generated from Task 1 confirmed that independently defined location-sensitive regions in STG were not modulated by emotion.

**Regions associated with processing auditory object identity**

In the current study, areas within STG were implicated in processing both the emotional and object identity features of auditory stimuli (Tasks 2 and 3). The identity-related effect identified in Task 3 was characterized in anterior-lateral auditory cortex by greater activity to biological relative to non-biological sounds, but did not distinguish between human and non-human sounds. This result appears to contradict previous work that implicates superior temporal sulcus (STS) specifically in human voice processing (Belin et al., 2000; Ethofer et al., 2012). One critical difference between these studies and the current study is the treatment of non-vocal human sounds (i.e., sounds that do not involve vocal fold vibration). Previous work found that contrasting vocal human sounds with non-vocal sounds yielded voice selective activity.
of the STS (Belin et al., 2000). Other studies yielding similar effects in STS contrasted human vocal sounds (human non-vocal sounds were excluded) with both animal and environmental sounds (Ethofer et al., 2012). The current study, however, classified all sound originating from a human source into the ‘human’ category (e.g., clapping and crying were both ‘human’). This regressor was designed to be an easily distinguishable auditory class, and not to represent neural changes associated with vocal processes. As such, it included one vocal and two non-vocal stimuli to eliminate any bias toward the identification of regions displaying voice-specific activation. The neural region observed in the present study may be located in an area slightly superior to that implicated in prior studies of human vocalizations. Interestingly, the use of linguistic utterances may also impact the location of emotion-sensitive regions. For example, Ethofer et al. (2012) demonstrated that an area of STG defined as selective for human voices is also sensitive to the emotional inflection of pseudo-linguistic stimuli. This region, while overlapping with the emotion-sensitive areas in the present study, extends into more posterior-lateral areas of temporal cortex. In future work, it will be interesting to determine whether verbal and non-verbal emotional sounds augment activity in dissociable areas of temporal cortex.

Regions associated with processing both auditory object identity and location

In addition to showing an effect of object identity, areas of right anterior-lateral STG also showed a significant effect of location, and a significant location X emotion interaction. Activity in this region was modulated by location for positive and negative stimuli, but not neutral stimuli. Furthermore, these regions showed increased activity to positive and negative relative to neutral stimuli, but only for those stimuli presented in the contralateral hemisphere. Thus, these areas show some evidence of location-related encoding for emotional, but not neutral stimuli. These findings raise the intriguing possibility that anterior-lateral regions of right auditory cortex may be involved in integrating information about object identity and location for emotionally salient stimuli. It also calls into question strict boundaries between spatial and object identity encoding in human STG. Further work will be required to delineate the underlying neuroanatomy and the stimulus parameters associated with these different processes in humans.

Representation of emotion in a dual pathway model of audition

Considerable evidence exists supporting the suggestion that the visual system is comprised of separable ventral and dorsal pathways responsible for processing objects for identity and action respectively (Milner and Goodale, 1998). It has also recently been proposed that auditory cortices can be separated into similar ‘what’ and ‘where’ pathways (Alain et al., 2001; Lomber and Malhotra, 2008; Rauschecker, 2012). In the present study, a functional dissociation between anterior-lateral and posterior-medial divisions of auditory cortices was observed that resembled the boundaries described in some dual pathway models of auditory processing (Bushara et al., 1999; Lomber 2012). Within both the precuneus and occipital cortex, the impact of location within each emotion was significant during the presentation of negative and positive (p < 0.005) but not neutral (p > 0.05) sounds. (*) p < 0.01; ** p < 0.005 *** p < 0.001.
and Malhotra, 2008; Warren and Griffiths, 2003). To further explore this possibility, results from Tasks 1 and 3 were used to independently generate functionally defined ROIs corresponding to these pathways. This analysis revealed that whereas sound location modulates neural activation within both the ‘what’ and ‘where’ auditory processing pathways, emotion modulates activity in the ‘what’ but not the ‘where’ processing pathway. Importantly, the location X emotion interaction observed in the anterior-lateral areas of auditory cortex showed that spatial processing in the ‘what’ pathway was present specifically for emotional stimuli.

Effects of emotion on localization behavior

It would seem advantageous for organisms to rapidly and accurately localize the source of emotional auditory stimuli in the environment. Contrary to expectations, participants in the present study were slower to locate positive and negative sounds relative to neutral ones; so why might emotional stimuli be associated with slower localization performance in the present study? One possibility is that important differences exist between the dual pathways of the visual and auditory systems. Our finding that participants were significantly slower to localize emotional relative to neutral sounds conflicts with evidence from the visual domain suggesting that dorsal stream guided behaviors are less susceptible to processing limitations (Fecteau et al., 2000; Pisella et al., 2000) and emotional interference (Ngan Ta et al., 2010). Another possibility is that, unlike in the visual system, both object perception and localization in the auditory domain may be sensitive to processing load. In the current study, emotion may have augmented the representation of object features at the expense of spatial-cues, thereby slowing localization performance. Consistent with this interpretation, we observed evidence that both spatial and emotional characteristics were integrated in anterior-lateral areas of auditory cortex, making this a potential site of competition between stimulus features. This integration of object and spatial features within the putative ‘what’ auditory area raises questions about the extent to which the ‘what’ and ‘where’ streams of the auditory system are functionally segregated.

It is important to note that the precise relationship between emotion and the putative ‘what/where’ pathways of the auditory system was only preliminarily addressed in the current study. First, the nature and anatomical mapping of the putative dual processing stream in the auditory system remains unclear, as is the degree to which its function resembles that of the visual system. In addition, the dissociable effects of emotional versus neutral stimuli within regions of auditory cortex observed were unexpected, and the interpretation is based in part on ROIs generated from follow-up scans involving a smaller sample. Although generated from activity in a sample smaller than the initial population, this ROI displayed similar functional properties to the regions identified in the original task, providing
cross-validation for our original interpretation. Nevertheless, in future work it would be beneficial to precisely map the boundaries of the ‘what/where’ auditory processing streams, and establish the impact of emotion on each.

Conclusions

Although the emotional content of a sound has been demonstrated to influence processing in auditory cortex, the neural and behavioral effects of emotion on stimulus localization was previously unknown. Our results indicate that during sound localization, the emotional content of a sound modulates activity in anterior-lateral regions of STG (areas corresponding to the putative ‘what’ pathway of auditory processing). In contrast, and contrary to predictions, emotion elicited no significant changes in neural activity to posterior-medial areas of STG (regions corresponding to the putative ‘where’ pathway). An unexpected interaction between emotion and location was also observed in anterior-lateral areas of STG suggesting that the boundaries between object identity and location decoding may be blurred; emotional sounds were associated with enhanced spatial processing in the ‘what’ pathway despite having no effect on the ‘where’ pathway. It is important to note that at the behavioral level, a significant delay in localization of emotional compared to neutral sounds was also observed. The idea was raised that this effect may have been driven by competition between spatial and emotional features for representation and control over behavior. Interestingly, anterior-lateral auditory cortex activity was modulated by both spatial location and emotion, suggesting this area as a potential site for such competitive interactions. Finally, this work demonstrates, for the first time, the feasibility of utilizing individualized virtual auditory environments as a stimulus presentation tool during fMRI. This technique, which involves presenting auditory stimuli through headphones while maintaining spatial integrity and perceptual realism, holds promise for future neuroimaging studies investigating the spatial component of sound.

Supplementary data to this article can be found at http://dx.doi.org/10.1016/j.neuroimage.2013.05.051.

Acknowledgments

The authors would like to thank Dr. Marc Joanisse and Dr. Stefan Everling for helpful discussion on experimental design and analysis, and David Grainger, of the National Centre for Audiology, for technical support. This work was funded by a Natural Science and Engineering Research Council Discovery Grant to Dr. Derek Mitchell.

Conflict of interest

The authors have no conflict of interest regarding the current experiment.

References


Vul, E., Kanwisher, N., 2010. Begging the Question: The Non-independence error in fMRI Data Analysis.


