

Research Report

An event-related fMRI study of auditory motion perception: No evidence for a specialized cortical system

Kevin R. Smith, Kourosh Saberi, Gregory Hickok*

Department of Cognitive Sciences, Center for Cognitive Neuroscience, 3151 Social Science Plaza, University of California Irvine, CA 92697-5100, USA

ARTICLE INFO

Article history: Accepted 1 March 2007 Available online 7 March 2007

Keywords: fMRI Imaging Auditory motion Planum temporale fMRI adaptation

ABSTRACT

The existence of a specialized human cortical area for the processing of auditory motion is still a matter of debate. Initial functional imaging studies identified the planum temporale as being motion selective. Recent data contrasting spatially varying stationary stimuli with moving stimuli found no difference in the amount of activation between the two types of stimuli in the planum temporale. The present study re-examines this issue using an event-related paradigm. Ten subjects were scanned while listening to pairs of stimuli that were either both moving or both stationary. Consistent with the aforementioned study, we found no difference in the altionary conditions.

© 2007 Elsevier B.V. All rights reserved.

1. Introduction

The existence of a separate auditory motion processing area in humans is still under debate. Initial evidence for a specialized auditory motion processing area came from animal models where single cell studies have identified brainstem and cortical neurons that show preferential activation to a specific direction of motion and do not respond to stationary sounds (Spitzer and Semple, 1991; Moiseff and Haresign, 1992). Thus far, psychophysical evidence has been equivocal. Evidence from spatial acuity thresholds suggests a common mechanism for the processing of stationary and motion stimuli (Grantham, 1984; Grantham, 1989). These studies propose that auditory motion is inferred from an analysis of position changes of discretely sampled loci in space. Such a snapshot model holds that computational mechanisms supporting sound localization are used to compute sound-source movement. Other psychophysical studies point to the perception of velocity, acceleration,

Corresponding author. Fax: +1 949 824 2307.
E-mail address: gshickok@uci.edu (G. Hickok).

0006-8993/\$ – see front matter © 2007 Elsevier B.V. All rights reserved. doi:10.1016/j.brainres.2007.03.003

and Doppler effect, as evidence for a specialized motion processing system (Perrott, 1989; Perrott et al., 1993).

In the past ten years, several human brain imaging studies have weighed in on the debate. The initial imaging studies contrasted various sorts of moving and stationary auditory stimuli to identify human cortical areas that are more responsive to moving sounds than stationary sounds. These studies initially identified three areas as motion selective: the planum temporale (bilaterally) (Baumgart et al., 1999; Krumbholz et al., 2005; Lewis et al., 2000; Pavani et al., 2002; Warren et al., 2002), the premotor cortex (Bremmer et al., 2001; Griffiths and Green, 1999; Griffiths et al., 2000), and the right parietal cortex (Bremmer et al., 2001; Griffiths and Green, 1999; Griffiths et al., 2000; Griffiths et al., 1998). The most consistent activations have been found in the planum temporale (bilaterally). However, these studies did not fully rule out a snap-shot interpretation of motion processing. In particular, all of these experiments contrasted moving stimuli with

stationary stimuli that do not vary in spatial location within a stimulation block. Thus, the increase in neural activity during the motion conditions relative to the stationary conditions may simply result from a snap-shot mechanism having to calculate spatial location more frequently in the motion conditions (multiple calculations required) than in the stationary conditions (one calculation required).

In a previous study (Smith et al., 2004), we sought to examine the possibility that auditory motion activations found in previous studies may simply be a consequence of changes in spatial location, as per the snapshot model. We conducted an fMRI experiment, where we contrasted blocks of moving Gaussian noise bursts (by way of a dynamic interaural level difference) with blocks of stationary (Gaussian) noise bursts sampled from eight different locations. This study found that there was no difference in the activations of the planum temporale bilaterally during motion perception when compared to a spatially varying stationary condition.

The present study sought to replicate this result with two important changes that should provide us with additional sensitivity to detect a possible motion-specific processing region. First we used an event-related, instead of a block design. Secondly, we employed an adaptation manipulation, which relies on the phenomenon of neural adaptation or habituation. In an adaptation design, two stimuli are presented with close temporal proximity. The stimuli are either identical (same) or they vary on some stimulus feature (different). It has been shown that the BOLD response to different pairs is greater than to same pairs, presumably because a greater degree of neural adaptation takes place for same pairs. This paradigm has been used successfully in several domains (Henson, 2003).

2. Results

Selection of voxel time courses within our left and right planum temporale ROIs resulted in a data set that consisted of 4032 data points, 2352 in the right hemisphere, and 1680 in the left. Outliers were excluded by calculating the mean and standard deviation across subjects and voxels at each sample point in each condition. Data point values that were 2 standard deviations above or below the mean were considered outliers and excluded from further analysis. These calculations were performed separately for each hemisphere. Excluded voxels amounted to 3.4% of the entire sample, and were equally distributed across hemispheres and conditions. The median value was then calculated for each sample time point, in each condition, in each ROI (left and right hemispheres), and in each subject. The median was chosen to further reduce the effects of outliers on signal amplitude estimations. (Calculating mean values yielded qualitatively identical results but with reduced power.) Thus, a cluster of voxels in a given subject was reduced to a single time course for each condition (motion-same, motion-different, stationary-same, stationary-different) and ROI (left vs. right planum). Fig. 1a shows the average time course for motion vs. stationary trials in the left (n=7) vs. right (n=10) hemisphere ROIs. Signal amplitude was then estimated for each condition, in each ROI, in each subject by calculating the difference between the local maximum and minimum values.

The resulting data set was then submitted to a 2 (motion vs. stationary)×2 (same vs. different) ×2 (left vs. right hemisphere) analysis of variance. This analysis involved only those seven subjects who had activations in both the left and right hemispheres. The main effect of primary interest, that between motion and stationary stimuli, did not reach significance (F(1,6)=0.858, p=0.39), indicating no difference overall in the amplitude to moving vs. stationary sound sources (Fig. 1b). The main effect of hemisphere, however, approached significance (F(1,6) = 4.081, p = 0.09), and indeed was highly reliable in a paired t-test (p < 0.003) (Fig. 1c), suggesting a right hemisphere bias in the response to spatial auditory stimuli generally. The hemisphere factor did not interact with the motion vs. stationary factor (F(1,6)=2.18), p=0.19) (Fig. 1d); however, the interaction between hemisphere and the same vs. different factor approached significance (F(1,6)=5.159, p=0.064) (Fig. 1e); post hoc contrasts showed that this effect was driven primarily by a significantly greater response to different than same stimulus pairs in the left hemisphere (p=0.01, Bonferroni corrected), same vs. different stimuli did not differ in the right hemisphere (p=0.14, uncorrected). Thus, the left but not the right hemisphere appeared to show an adaptation effect for both motion and stationary stimuli. No other effects approached significance.

A separate 2 (motion vs. stationary)×2 (same vs. different) ANOVA was carried out on the right hemisphere data only. This analysis included all 10 participants. The main effect of stimulus type (motion vs. stationary) approached significance (F(1,9)=3.862, p=0.08) (Fig. 1f). This finding hints at a possible weak preference for motion over stationary sound sources in the right hemisphere ROI. However, the magnitude of this effect is very small (12% difference), and (if real) is more parsimoniously explained in terms of the increase in spatial locations sampled by a moving vs. a stationary trial, rather than in terms of a dedicated auditory motion-computation system. No other effects approached significance.

3. Discussion

Using an event-related adaptation design, we assessed whether regions of the planum temporale show preferential responses to moving auditory sound sources when compared to spatially varying but non-moving sound sources. Consistent with much previous work (Baumgart et al., 1999; Bremmer et al., 2001; Griffiths et al., 1998; Griffiths and Green, 1999; Griffiths et al., 2000; Krumbholz et al., 2005; Lewis et al., 2000; Pavani et al., 2002; Smith et al., 2004; Warren et al., 2002), we found activation foci within the planum temporale of each participant that was responsive to moving sound sources when compared with listening to background scanner noise (rest). Such foci were found in the right hemisphere in all 10 subjects, and in the left hemisphere of 7 of the 10 subjects. However, these same regions were also highly responsive to stationary sound sources. The signal amplitude for moving vs. stationary stimuli failed to reach statistical significance in the overall analysis (left and right ROIs combined), and also failed to reach significance in the within hemisphere analyses. Right hemisphere activation



Fig. 1 – Time course and mean amplitude activations in the Planum Temporale (ROI). (a) Time course for motion vs. stationary sound sources in left vs. right ROIs. (b) Signal amplitude for motion vs. stationary sound sources collapsed across left and right ROIs. (c) Signal amplitude for left vs. right hemisphere ROIs. (d) Signal amplitude for motion vs. stationary sound sources in left vs. right hemisphere ROIs. (e) Signal amplitude for same vs. different stimulus pairs in left vs. right hemisphere ROIs. (f) Signal amplitude for motion vs. stationary stimuli in the right hemisphere ROI.

to spatial auditory stimuli (both moving and stationary) was significantly greater than left hemisphere activation in post hoc contrasts among subjects who had both left and right planum activations. This observation, coupled with the fact that only 70% of participants showed left hemisphere activation, compared to 100% with right hemisphere activations, suggests a right hemisphere bias in spatial auditory processing (Bushara et al., 1999; Clarke et al., 2000; Griffiths et al., 1998; Griffiths et al., 2000; Weeks et al., 1999). Finally, although an overall adaptation effect was observed in the left hemisphere ROI (different trials > same trials), this adaptation did not interact with stimulus type (motion vs. stationary). It is not clear why the adaptation effect would be lateralized to the left hemisphere. What is clear, however, is that the adaptation manipulation did not lead to any motion-specific effects.

The present findings confirm and extend our previous report (Smith et al., 2004) showing no difference in response between moving and spatially varying but non-moving sound sources. This lack of a difference held up even in the present study, which used a potentially more sensitive event-related design. Furthermore, the possibility that motion-selective processing might be revealed by employing an adaptation paradigm was also assessed, but failed to detect motionspecific processing. It is possible, however, that the feature we selected in our adaptation design (direction of motion) was not a relevant computation parameter, and that adaptation designs using other features may prove more successful.

Our analysis failed to yield a significant motion vs. stationary difference in either the left or right planum temporale. However, the right planum temporale showed a very small difference in amplitude favoring motion stimuli that reached a *p*-value of 0.08. One might be tempted to argue that this finding supports the claim for a dedicated auditory motion processing region. This would be a rather weak argument, though, for several reasons. First, the effect is far from robust. No difference was found in the right planum in our previous block design study, and only a marginal effect was found in the present study which used a ROI approach. One would expect a dedicated motion processing system to yield robust and replicable differences. Second, even if one accepts the difference as real, the magnitude of the effect is small. By comparison, the magnitude of motion effects in visual motion areas compared to stationary stimuli is typically on the order of a 3 to 1 ratio (Tootell et al., 1995). The present effect is only a 1.13 to 1 ratio for moving vs. stationary sound sources. While there is no reason to think that the auditory system should behave analogously to the visual system, the small effect size suggests that other, more parsimonious, explanations may be preferred. For example, one possibility is that the motion > stationary difference (again, if real) reflects the fact that the moving stimuli sampled more spatial

Table 1 – Experimental design					
	Same trials		Different trials		
Motion Stationary	$L \rightarrow R, L \rightarrow R$ L, L	$R \rightarrow L, R \rightarrow L$ R, R	$L \rightarrow R, R \rightarrow L$ L, R	$R \rightarrow L, L \rightarrow R$ R, L	
Depiction of the various conditions implemented in our 2×2 design.					

locations than our stationary stimuli, which only sampled two locations. This explanation predicts that if we sampled more locations in the stationary conditions, the motion vs. stationary difference would evaporate. Our previous study (Smith et al., 2004) did, in fact, sample more spatial locations in the stationary condition, and no differences were observed.

One recently published study (Poirier et al., 2005) claims to have observed greater activation for moving than nonmoving but spatially varying auditory sound sources. An impressively large portion of the cerebral cortex was activated in the motion minus stationary contrast in that study, but critically for our purpose, the planum temporale was among the activated regions. The authors suggest that portions of the planum are indeed specialized for auditory motion processing. There are several problems with that study however. First, activation of planum was observed in a fixed effect group analysis with only six subjects. It is therefore not clear whether this effect generalizes across subjects. Second, the task used in that study introduced confounds between response complexity and the motion vs. stationary manipulation. Two types of sounds were used, pure tones and complex tones, and these were either moving or non-moving but at various spatial locations. Responses were made, as follows, by pressing one of two switches, one in the left hand and one in the right: (a) stationary pure tone stimuli in any position, right switch; (b) stationary complex stimuli in any position, left switch; (c) pure tone moving to right, right switch; (d) complex sound moving to the left, left switch; (e) pure tone moving to the left, both switches; (f) complex sound moving to the right, both switches. Effectively what this entails is that for stationary stimuli, subjects simply classify the sounds in terms of the tone quality irrespective of spatial location (pure vs. complex=right vs.

left switches); whereas for moving stimuli subjects must classify the tone quality and indicate direction of motion (pure+right motion=right switch, pure+left motion=both switches, etc.). It seems likely that this compound classification requirement for the motion stimuli vs. the binary classification for the stationary stimuli, is what led to the extensive motion>stationary activations.

In conclusion, despite many claims to the contrary, there is very little evidence to support the claim that a specialized auditory motion processing system exists in human cortex. The present study replicated and extended our previous experiment showing that areas implicated in auditory motion processing (portions of the planum temporale in particular) are equally responsive to non-moving, but spatially varying sound sources. This finding does not preclude the possibility that such a specialized mechanism exists, but it does show clearly that existing neuroimaging evidence has not yet provided convincing support for the hypothesis because of a failure to control for non-motion spatial processing.

4. Experimental procedures

4.1. Participants

Ten (6 male and 4 female) subjects participated in this study. Subjects gave informed consent under a protocol approved by the Institutional Review Board at the University of California, Irvine.

4.2. Materials

Stimuli were pairs of 1000 ms bursts of Gaussian noise presented through electrostatic headphones (STAX SR-001) at a sampling rate of 44.1 kHz. On each trial, a 1000 ms burst of noise was followed by a 300 ms silent interstimulus interval (ISI) followed by another 1000 ms burst of noise. Each noiseburst had a 10 ms linear rise-decay time. Motion was simulated by dynamically changing the stimulus interaural time difference (ITD). Stimuli were generated in the frequency domain using a pair of complementary Discrete Fourier



Fig. 2 - Activation maps depicting the central coordinate of activation for each subject on a group template.

Table 2 – Talairach coordinates				
Subject	Right hemi coordinates	Left hemi coordinates		
1	51, -24, 4	-53, -24, 13		
2	58, -18, 7	-55, -20, 11		
3	58, -21, 10	-60, -25, 11		
4	52, -40, 13			
5	49, -37, 16	-51, -20, 11		
6	49, -34, 8	-55, -32, 10		
7	57, -21, 10			
8	61, -34, 9	-54, -40, 13		
9	50, -27, 10			
10	45, –20, 8	-50, -21, 6		
Talairach coordinates for the center of activation from each ROI.				

Transforms (DFT) for which the component spacing for one series was different than that of the other (Saberi, 2004). For motion stimuli, the ITD linearly changed from $-600 \ \mu s$ to $600 \ \mu s$ during the 1 s presentation of each stimulus, where the negative and positive signs denote ITDs favoring the left and right ears, respectively. The stimulus thus simulated a constant velocity of 1200 $\ \mu s/s$, which is equivalent to an angular velocity of approximately 150°/s associated with a real-source moving in the freefield. This velocity was selected because it produces a strong percept of motion and low motion-detection thresholds (Saberi et al., 2002; Saberi et al., 2003). The stationary (no-motion) stimuli consisted of the same 1 s bursts of Gaussian noise, but with a constant ITD of either 600 or $-600 \ \mu s$ on each presentation.

4.3. Design and procedure

We employed a 2×2 factorial design. One factor was a manipulation of stimulus type, motion vs. stationary sound sources. The other factor was an adaptation manipulation, that is, whether the two stimuli in a single trial were the same (possibly producing an adaptation effect) or different. For motion stimuli, the direction of motion was either the same for both stimuli in a trial (e.g., left to right motion followed by left to right motion), or different (e.g., left to right followed by right to left). For stationary stimuli, the spatial position of a sound source was either identical (e.g., right sound source followed by right sound source) or in opposite hemifields (e.g., right followed by left). Left to right direction of motion, and left and right spatial positions were equally represented across conditions and fully counterbalanced (see Table 1). Subjects were presented with 24 trials in each condition which were presented in a fixed random order within each run. The order of runs was counterbalanced across subjects. Subjects performed a same-different discrimination task on the stimulus pairs.

4.4. fMRI acquisition

Sixteen axial slices were collected using a 1.5 T Marconi/Picker Eclipse whole body MRI scanner using an EPI pulse sequence (FOV=240 mm, matrix=128×128, TE=40 ms, TR=2 s, slice thickness=5 mm). For each subject, a high resolution anatomical image was acquired with a 3D SPGR pulse sequence.

To correct for subject motion artifacts, the image volumes of each subject were aligned to the sixth volume in the series using a 3D rigid body, six parameter model in the AIR 3.0 program (Woods et al., 1998). The volumes were then coregistered to the high resolution anatomical image. After alignment, each volume was spatially smoothed (Gaussian spatial filter 4 mm FWHM) and the time course of the blood oxygen level dependent (BOLD) signal was temporally filtered (bandpass 0.1667 to 0.6667 Hz).

Previous studies of auditory motion perception have identified the planum temporale as a critical site. Thus, our analysis strategy was region of interest (ROI) based, and proceeded as follows. First, for each subject, an impulse response function (IRF) corresponding to each condition for all of the motion trials was generated in AFNI using a deconvolution analysis [B.D. Ward, Deconvolution Analysis of fMRI time series data, http://afni.nimh.nih.gov/afni]. These IRFs were used as regressors and used to calculate an F statistic at each voxel. Activation maps from the F statistics were thresholded at p < 0.001 (uncorrected) and contiguous clusters of four voxels or more were used as ROIs (Forman et al., 1995). Next, two ROIs, one in the left hemisphere and one in the right, were defined in the planum temporale of each subject (defined on the subject's own sulcal and gyral anatomy) by selecting the top five activated voxels in each hemisphere in the motion-minus-rest contrast. For this contrast, all motion conditions were combined to avoid selection bias based on adaptation conditions. Motion stimuli were used to select ROIs because such a procedure should, if anything, bias our findings against our hypothesis that planum activations are not motion specific. That is, if a motion-selective region exists, using motion stimuli to define our ROIs will more likely identify this region than using stationary stimuli. ROIs that reached threshold were identified in the right planum temporale in all 10 subjects. In the left planum temporale, seven of the 10 subjects had activations that reached threshold. Fig. 2 shows the locations for the center of activation of each subject in each hemisphere (for Talairach coordinates see Table 2). Voxel time course data for these ROIs in all conditions were the focus of all subsequent analyses.

Acknowledgment

This work was supported by NSF grant BCS0477984.

REFERENCES

- Baumgart, F., Gaschler-Markefski, B., Woldorff, M.G., Heinze, H.J., Scheich, H., 1999. A movement sensitive area in auditory cortex. Nature 4000, 724–726.
- Bremmer, F., Schlack, A., Shah, N., Zafiris, O., Kubischik, M., Hoffmann, K.P., Zilles, K., Fink, G.R., 2001. Polymodal motion processing in posterior parietal and premotor cortex: a human fMRI study strongly implies equivalencies between humans and monkeys. Neuron 29, 287–296.
- Bushara, K.O., Weeks, R.A., Ishii, K., Catalan, M.J., Tian, B., Rauschecker, J.P., Hallett, M., 1999. Modality-specific frontal and parietal areas for auditory and visual spatial localization in humans. Nat. Neurosci. 2, 759–766.

- Clarke, S., Bellmann, A., Meuli, R.A., Assal, G., Steck, A.J., 2000. Auditory agnosia and auditory spatial deficits following left hemispheric lesions: evidence for distinct processing pathways. Neuropsychologia 38, 797–807.
- Forman, S.D., Cohen, J.D., Fitzgerald, M., Eddy, W.F., Mintun, M.A., Noll, D.C., 1995. Improved assessment of significant activation in functional magnetic resonance imaging (fMRI): use of a cluster-size threshold. Magn. Reson. Med. 33, 636–647.
- Grantham, D.W., 1984. Discrimination of dynamic interaural intensity differences. J. Acoust. Soc. Am. 76, 71–76.
- Grantham, D.W., 1989. Auditory motion perception via successive 'snapshot' analysis. Sound localization by human observers. Proc. Natl. Acad. Sci. 35.
- Griffiths, T., Green, G., 1999. Cortical activation during perception of a rotating wide-field acoustic stimulus. NeuroImage 10, 84–90.
- Griffiths, T.D., Rees, G., Rees, A., Green, G.G., Witton, C., Rowe, D., Buchel, C., Turner, R., Frackowiak, R.J., 1998. Right parietal cortex is involved in the perception of sound movement in humans. Nat. Neurosci. 1, 74–78.
- Griffiths, T., Green, G., Rees, A., Rees, G., 2000. Human brain areas involved in the analysis of auditory movement. Hum. Brain Mapp. 9, 72–80.
- Henson, R.N.A., 2003. Neuroimaging studies of priming. Prog. Neurobiol. 70, 53–81.
- Krumbholz, K., Schonwiesner, M., Rubsamen, R., Zilles, K., Fink, G.R., Voon Cramon, D.Y., 2005. Hierarchical processing of sound location and motion in the human brainstem and planum temporale. Eur. J. Neurosci. 21 (1), 230–238.
- Lewis, J., Beauchamp, M., DeYoe, E., 2000. A comparison of visual and auditory motion processing in human cerebral cortex. Cereb. Cortex 10, 873–888.
- Moiseff, A., Haresign, T., 1992. Response of auditory units in the barn owl's inferior colliculus to continuously varying interaural phase differences. J. Neurophysiol. 67, 1428–1436.
- Pavani, F., Macaluso, E., Warren, J.D., Driver, J., Griffiths, T.D., 2002. A common cortical substrate activated by horizontal and vertical sounds movement in the human brain. Curr. Biol. 12, 1584–1590.

- Perrott, D.R., 1989. Are there motion detectors in the auditory system? Sound localization by human observers. Proc. Natl. Acad. Sci. 36.
- Perrott, D.R., Costantino, B., Ball, J., 1993. Discrimination of moving events which accelerate or decelerate over the listening interval. J. Acoust Soc. Am. 93, 1053–1057.
- Poirier, C., Collignon, O., DeVolder, A.G., Renier, L., Vanlierde, A., Tranduy, D., Scheiber, C., 2005. Specific activation of the V5 brain area by auditory motion processing: an fMRI study. Cogn. Brain Res. 25, 650–658.
- Saberi, K., 2004. Fast Fourier-based DSP algorithm for auditory motion experiments. Behav. Res. Methods Instrum. Comput. 36, 585–589.
- Saberi, K., Tirtabudi, P., Petrosyan, A., Perrott, D.R., Strybel, T.Z., 2002. Concurrent motion detection based on dynamic changes in interaural delay. Hear. Res. 174, 149–157.
- Saberi, K., Tirtabudi, P., Petrosyan, A., Perrott, D.R., Strybel, T.Z., 2003. Detection of dynamic changes in interaural delay. Acta Acustica 89, 333–338.
- Spitzer, M.W., Semple, M.N., 1991. Interaural phase coding in auditory midbrain: influence of dynamic stimulus features. Science 254, 721–723.
- Smith, K.R., Okada, K., Saberi, K., Hickok, G., 2004. Human cortical auditory motion areas are not motion selective. NeuroReport 15, 1523–1526.
- Tootell, R.B., Reppas, J.B., Kwong, K.K., Malach, R., Born, R.T., Brady, T.J., Rosen, B.R., Belliveau, J.W., 1995. Functional analysis of human MT and related visual cortical areas using magnetic resonance imaging. J. Neurosci. 15, 3215–3230.
- Warren, J.D., Zielinski, B.A., Green, G.G., Rauschecker, J.P., Griffiths, T.D., 2002. Perception of sound-source motion by the human brain. Neuron 34, 139–148.
- Weeks, R.A., Aziz-Sultan, A., Bushara, K.O., Tian, B., Wessinger, C.M., Dang, N., Rauschecker, J.P., Hallett, M., 1999. A PET study of human auditory spatial processing. Neurosci. Lett. 262, 155–158.
- Woods, R.P., Grafton, S.T., Watson, J.D., Sicotte, N.L., Mazziotta, J.C., 1998. Automated image registration: II. Intersubject validation of linear and nonlinear models. J. Comput. Assist. Tomogr. 22, 153–165.