ELSEVIER

NeuroImage

www.elsevier.com/locate/ynimg
NeuroImage 28 (2005) 132 – 139

Listening to a walking human activates the temporal biological motion area

Aurélie Bidet-Caulet,[1] Julien Voisin,[1] Olivier Bertrand, and Pierre Fonlupt*

INSERM U280, IFNL, UCB-Lyon1, Centre Hospitalier Le Vinatier, 69500 Bron, France

Received 14 December 2004; revised 24 May 2005; accepted 1 June 2005
Available online 18 July 2005

A vivid perception of a moving human can be evoked when viewing a few point-lights on the joints of an invisible walker. This special visual ability for biological motion perception has been found to involve the posterior superior temporal sulcus (STSp). However, in everyday life, human motion can also be recognized using acoustic cues. In the present study, we investigated the neural substrate of human motion perception when listening to footsteps, by means of a sparse sampling functional MRI design. We first showed an auditory attentional network that shares frontal and parietal areas previously found in visual attention paradigms. Second, an activation was observed in the auditory cortex (Heschl's gyrus and planum temporale), likely to be related to low-level sound processing. Most strikingly, another activation was evidenced in a STSp region overlapping the temporal biological motion area previously reported using visual input. We thus propose that a part of the STSp region might be a supramodal area involved in human motion recognition, irrespective of the sensory modality input.

Keywords: fMRI; Biological motion; Auditory; Footsteps; Superior temporal sulcus

Introduction

Perceiving conspecific actions forms a key feature to survival for social species. This implies a fine ability towards movement recognition. Indeed, humans can recognize 'biological motion' from the minimal information displayed by a set of point-lights attached to the joints of an invisible walking actor (Johansson, 1973). The neural support of this visual ability has been repeatedly studied using positron emission tomography (Bonda et al., 1996), functional MRI (Grèzes et al., 2001; Grossman and Blake, 2002; Grossman et al., 2000; Howard et al., 1996; Pelphrey et al., 2003; Ptito et al., 2003; Servos et al., 2002; Vaina et al., 2001), electroencephalography (Hirai et al., 2003), magnetoencephalography (Pavlova et al., 2004), and neurological observations (Battelli et al., 2003; Pavlova et al., 2005). All these works point to a specific activation along the posterior extent of the superior temporal sulcus (STSp) and its ascending limb in the inferior parietal cortex, a region that could be the human homologue of the monkey superior temporal polysensory area (Puce and Perrett, 2003). This region is more active during biological motion perception than during any other kind of movement tested, and seems better related to human motion per se than to movement perception in general (Pavlova et al., 2004; Beauchamp et al., 2002, 2003; Santi et al., 2003; Grossman and Blake, 2001; Johansson, 1973). Its location, anterior to and superior to the human MT/V5 complex (Grèzes et al., 2001; Grossman et al., 2000), at the temporo-parieto-occipital junction, supports the proposal of an integration area involved in extracting structure from visual motion (Beintema and Lappe, 2002).

In addition, the STSp region involved in human motion perception as seen using visual inputs has also been reported to play a key role in social cognition (Puce and Perrett, 2003), polysensory interaction (Beauchamp et al., 2004; Wright et al., 2003), and even environmental sound recognition (Lewis et al., 2004). Critically, all the previous biological motion studies looked at human motion perception through the visual modality only. In everyday life, however, human motion can also be recognized using acoustic cues. In the present study, we questioned whether STSp would be involved in biological motion perception from auditory input as well. This has been addressed by considering an ecological situation, i.e., footstep listening.

An fMRI sparse sampling design was used so as to avoid scanner noise interference with the auditory stimuli. Subjects were instructed to listen to the footsteps of two walking humans (one on each side) during 6 s before MRI scanning started. After a random delay (1–4 s), one of the walkers crossed the auditory scene from one side to the other. Subject's task was then to report the crossing direction using two response buttons.

* Corresponding author.
E-mail address: fonlupt@lyon.inserm.fr (P. Fonlupt).
[1] These authors contributed equally to this work.
Available online on ScienceDirect (www.sciencedirect.com).

1053-8119/$ - see front matter © 2005 Elsevier Inc. All rights reserved.
doi:10.1016/j.neuroimage.2005.06.018

Materials and methods

Participants

10 subjects with normal hearing were included (3 females, aged 20–29). They were all right-handed and without any history of neurological or psychiatric disorder. Informed consent was obtained for all subjects and the study was approved by National Ethical Regulation (no RBM 03-18).

Stimulus

Auditory stimulus lasted 6 s and consisted of footsteps of two persons, one walking on the left side of the subject, and the other one walking on the right side. After a random delay, one of the two walkers started moving across the auditory scene from one side to the opposite side, while the other one kept on walking straight in his initial side. The beginning of the crossing (left-to-right or right-to-left, equiprobably) randomly occurred 1, 2, 3, or 4 s after stimulus onset and was completed 2 s later.

Each type of human footsteps was recorded individually, free from any background noise. These recordings were then digitized in mono mode (16 bits). Amplitude normalization, lateralization, linear crossing, and mixing of footsteps were obtained with SoundForge 6.0 (Sonic Foundry, Madison, Wisconsin, USA). Inter-aural intensity differences were used to give a realistic impression of spatially moving sounds. We used 24 auditory stimuli differing by the type of shoes and ground/floor (sand, mud, cobblestones, carpet, tiled or wooden floors). The walkers' footsteps were kept coherent within each trial in regards to the general scene (either indoor or outdoor) and to the sound intensity. Across trials, footstep speed was randomized but did not differ by more than 20%.

Task

Subjects were asked to indicate, as soon as possible, the motion direction of the crossing walker by clicking the right or left button of a response box (right hand), and then to wait passively until the end of the stimulus. The task was designed so as to direct subject's attention towards the footsteps, and to ensure that the two walkers were clearly identified and their trajectory followed. The difficulty of the task was set at a level allowing all subjects to perform it easily. Between the beginning of the sound and the button press, subject's attention was oriented towards the footsteps; we called this period the "attention period".

Procedure

Two 36-trial runs were acquired for each subject, i.e., 24 auditory footstep stimuli and 12 resting states. Among the 24

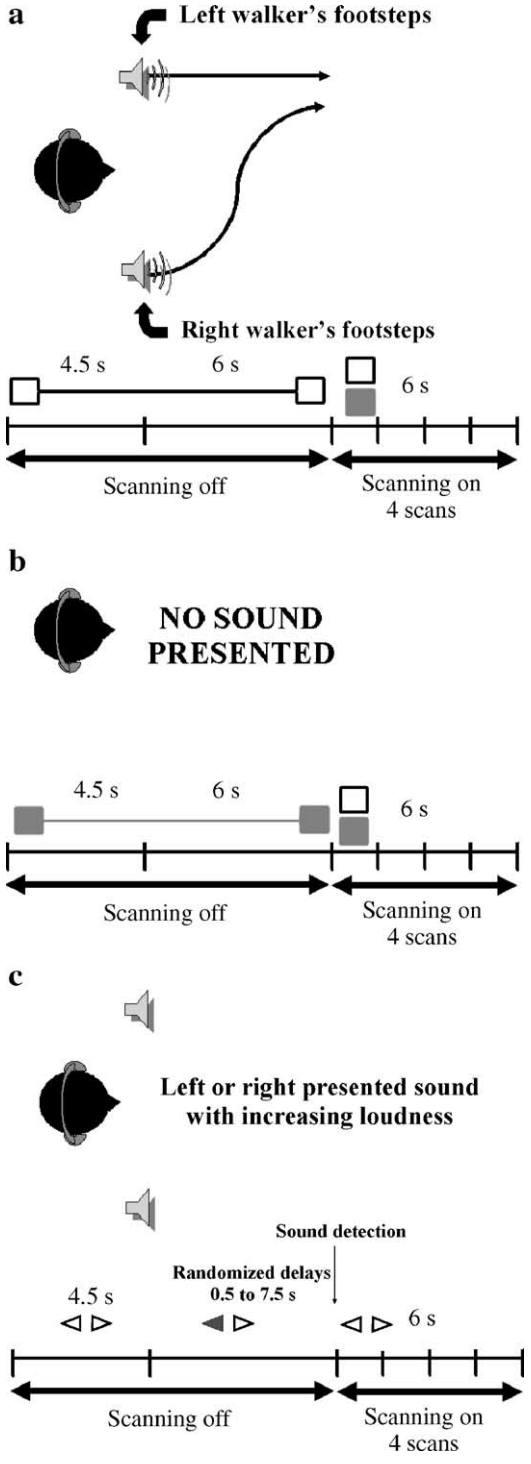

Fig. 1. Trial design. (a) Footstep perception. A green square (white on the figure) displayed on the screen instructs the subject to actively listen to the auditory scene. The sounds of two walkers' footsteps are delivered (one on each side). After a random delay (1–4 s), one of the walker crosses the auditory scene during 2 s towards the opposite side. As soon as the subject perceives the crossing, he has to press the response button. The total duration of an auditory stimulus is 6 s whatever delay duration before footstep crossing and subject's reaction time. When the sound ends, four MRI scans are acquired during 6 s and the color of the square is changed according to the next trial condition. (b) Resting state. A grey square instructs the subject to rest until the next color switches. Four MRI scans are acquired with the same timing as for the footstep detection condition. (c) Noise detection. A green (white on the figure) double sided arrow instructs the subject to detect the emergence of a sound from the silence. The noise is delivered to the right or left ear with increasing intensity. As soon as the subject detects the sound, he presses the response button and the scanning begins.

footstep trials, the crossing started 1/2/3/4 s after sound onset, in 6/6/3/9 trials, respectively. In each trial, a green square indicating the forthcoming auditory stimulus was presented during 4.5 s to the subjects while a grey square indicated that the next trial would be a resting state trial (no acoustic stimulus, no task during 6 s). MRI acquisition started at the end of the auditory stimulus (Fig. 1a) or resting period (Fig. 1b), and four scans were successively acquired. The trial order was randomized.

Subjects were equipped with earplugs covered by piezoelectric circumaural headphones. Before fMRI acquisition, an empirical method has been used to adjust stimulus intensity. A series of footsteps and band-passed noises were delivered to the subject to adapt loudness, specifically for each subject, allowing a comfortable hearing level and a clear sound perception with this setting.

fMRI acquisition

A sparse sampling acquisition protocol was used following the procedure described by Seifritz et al. (2002). Images were acquired using a 3T whole-body scanner MEDSPEC 30/80 AVANCE (Brücker, Ettlingen, Germany) equipped with a circular polarized head coil. For each participant, we first acquired high resolution T1-weighted anatomical images (inversion–recovery sequence, 1 × 0.75 × 1.22 mm voxel size, sagittal orientation), covering the whole brain. For functional imaging, an EPI sequence was used with TE = 30 ms, TR = 1580 ms, acquisition bandwidth 123 kHz, 64 × 64 matrix, 192 × 192 mm field of view. The resulting in-plane resolution was 3 × 3 mm. Twenty-six adjacent axial slices (3 mm thickness, 0.5 mm interslice gap) were acquired parallel to the bicommissural plane.

Data analysis

The data were pre-processed using SPM2 software (http://www.fil.ion.ucl.ac.uk/spm/). Each subject's image was realigned to a scan halfway through the time series to correct for motion, and then normalized using the template of the Montreal Neurological Institute (MNI). The scans were then smoothed with a Gaussian kernel (6 × 6 × 7 mm half-width) and finally resampled in the space defined by the MNI anatomical template (2 × 2 × 2 mm voxel size). The functional data were analyzed using general linear model (Friston et al., 1995; Wicker and Fonlupt, 2003). Analysis was performed on the average of the 4 scans to avoid any assumption about the shape of the hemodynamic response. We used a mixed effect model allowing for inference about the population by taking into account inter-subject variance (called "random effect analysis" in SPM software).

The contrast of interest was the difference between the footstep condition and the resting state condition, we called this contrast: "footstep−rest". For the group analysis, the averaged contrasts across the 10 subjects were then compared to 0 using a Student's *t*

Table 1
Characteristics of the clusters activated during footstep perception versus resting state

Side region	(Broadman's area)	(Cluster name)	MNI coordinates			*t* max	Cluster size	Number of subjects
			x	*y*	*z*			
Frontal								
R inferior frontal gyrus/precentral gyrus	(BA 44/BA 6)	(*r*DLPF)	55	19	29	7.04	243	8
R inferior frontal gyrus	(BA 47)	(*r*IFG)	43	19	−2	5.69	137	5
L precentral gyrus	(BA 6)	(*l*DLPF)	−47	−5	19	7.84	258	6
L inferior frontal gyrus	(BA 47)	(*l*IFG)	−44	16	−5	5.63	146	5
Superior frontal gyrus, supplementary motor area	(BA 6)	(*l*DLPF)	4	7	57	8.01	507	7
Parietal								
R supramarginal gyrus/inferior parietal lobule	(BA 40)	(*r*SG-IPL)	49	−50	37	6.88	352	8
L supramarginal gyrus/inferior parietal lobule	(BA 40)	(*l*SG-IPL)	−35	−65	40	5.86	291	7
Temporal								
R superior temporal gyrus	(BA 21/BA 22)	(*r*STG-HG)	49	−2	−13	10.59	175	6
R posterior superior temporal gyrus	(BA 21/BA 22)	(*r*STG-MTG)	58	−47	15	6.23	155	9
L superior temporal gyrus	(BA 22)	(*l*STG-HG)	−41	−23	−9	7.06	184	6
L posterior middle temporal gyrus	(BA 21/22)	(*l*STG-MTG)	−59	−41	5	6.35	116	8
L posterior middle temporal gyrus	(BA 21/BA 37)	(*l*STG-MTG)	−59	−62	1	7.24	219	7
Sub-cortical								
R thalamus	(ventral anterior nucleus)		16	−2	12	7.20	86	5
L thalamus	(ventral anterior nucleus)		−14	−5	12	8.25	104	3
Thalamus	(Pulvinar)		1	−29	−5	8.15	398	7

Results are derived from a mixed effect analysis (10 subjects). *x*, *y*, and *z* correspond to the stereotaxic coordinates (mm) in the MNI reference brain provided with SPM2. The statistical map is thresholded at $t = 3.4$ which corresponds to $P < 0.01$. All *t* max (the maximum value of *t* among the voxels of the cluster) reported in the table correspond to $P < 0.001$. Cluster size represents the number of contiguous voxels showing significant activation (2 × 2 × 2 mm voxels). For each cluster (found in the group analysis) and subject, the "footstep − rest" contrast is computed from the mean activity of the cluster. The last column (number of subjects) gives the number of subjects showing a significant effect ($P < 0.05$). Cluster names are given according to the cluster extent depicted in Fig. 2. DLPF: dorsolateral prefrontal cortex; IFG: inferior frontal gyrus; SMA: supplementary motor area; SG-IPL: supramarginal gyrus–inferior parietal lobule; STG-MTG: superior temporal gyrus–middle temporal gyrus; STG-HG: superior temporal gyrus–Heschl's gyrus. *r* or *l* stands for right or left hemisphere.

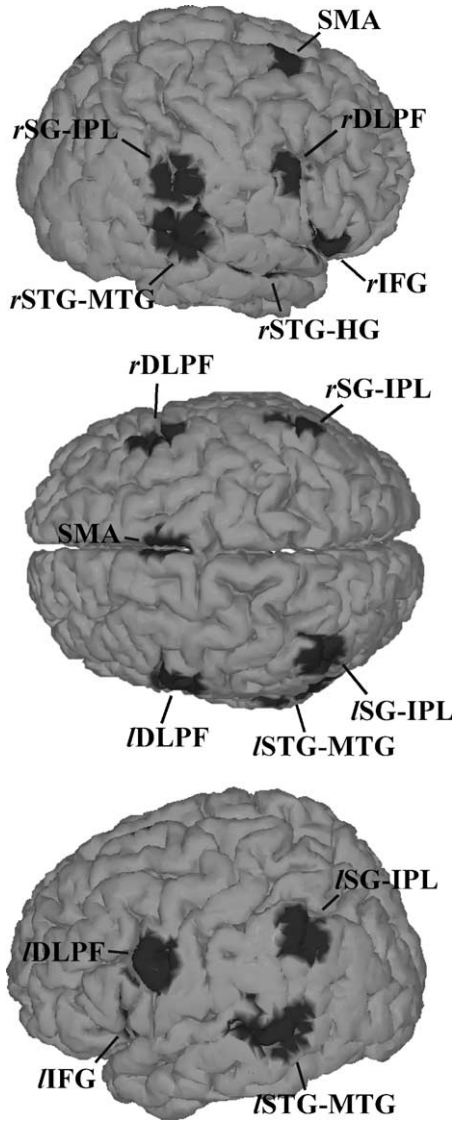

Fig. 2. Topography of the brain activations found when contrasting footstep perception to resting state, displayed on a pseudo-3D brain render. See Table 1 for the detailed characteristics of these areas. DLPF: dorsolateral prefrontal cortex; IFG: inferior frontal gyrus; SMA: supplementary motor area; SG-IPL: supramarginal gyrus–inferior parietal lobule; STG-MTG: superior temporal gyrus–middle temporal gyrus; STG-HG: superior temporal gyrus–Heschl's gyrus. *r* or *l* stands for right or left hemisphere.

test (corrected for multiple comparisons using the Scheffe method). All the statistical maps were thresholded at $P < 0.01$ ($t = 3.4$). Only clusters of at least 20 contiguous significant voxels were taken into consideration. For the individual subject analysis, the "footstep − rest" contrast was studied independently for each subject.

The anatomical location of the clusters was determined using "MAsks for Region of INterest Analysis" (http://www.bion.de/Marina.htm) and "Talairach Daemon" (http://ric.uthscsa.edu/projects/talairachdaemon.html) software.

A parametric regression analysis was then performed on each cluster identified from the contrast: "footstep − rest". The method allows the determination of brain areas where the activity was increasing when the task duration or load was increasing. In the present study, delays between stimulus onset and footstep crossing were randomized between 1 and 4 s. When subject had detected the footstep crossing (button press), he had no longer to focus his attention on footsteps. Thus, the period between stimulus onset and subject response corresponded to the "attention period" when subject's attention was oriented towards footsteps. We aimed at identifying brain areas where the activity was increasing when the "attention period" was increasing. In practice, the duration of the "attention period" was used as a continuous regressor in the general linear model. A detailed description of such a procedure has been previously reported (Mazoyer et al., 2002).

Functional localizer

To identify the temporal regions activated by auditory attention directed to sounds which do not represent any structured object, each subject undertook a functional localizer run. It was based on an experimental design similar to that of the "footstep" run (Fig. 1c), except that subject's task was to detect the emergence of a band-passed filtered noise from silence (24 semitone-wide, starting frequency varying from 175 to 831 Hz across trials). The MRI acquisition of four scans started after the emerging sound was detected by the subject. Auditory attention-dependent responses were obtained by contrasting those scans (28 trials) to those obtained in resting state trials (20 trials), we called this contrast: "noise − rest".

Table 2
Individual subject analysis of the clusters activated in the STS during footstep perception versus resting state

Subject	*x*	*y*	*z*	*t* max	Cluster size	% of signal change
Left						
#1	−44	−44	12	5.24	51	0.26
#2	−44	−42	−2	14.55	587	0.99
#3	−60	−32	−2	6.29	176	0.76
#4	−60	−42	12	11.72	547	0.71
#5	−62	−38	−2	6.37	249	0.62
#6	−56	−36	−2	8.05	524	0.77
#7	−56	−36	8	13.22	380	0.69
#8	−56	−30	8	18.56	792	0.64
#9	−54	−32	0	19.24	728	0.96
#10	−50	−48	12	4.39	31	0.09
Right						
#1	60	−42	12	7.12	332	0.28
#2	60	−48	0	8.16	553	0.45
#3	52	−54	18	5.02	166	0.62
#4	54	−50	18	8.95	400	0.52
#5	46	−42	−6	5.81	162	0.25
#6	64	−36	4	13.01	884	0.50
#7	54	−44	12	12.54	341	0.59
#8	60	−36	12	12.90	654	0.42
#9	58	−42	−6	13.04	516	0.38
#10	52	−54	14	6.01	103	0.09

For each of the 10 subjects, the effect of footstep perception (contrast: "footstep − rest") was analyzed. This shows the characteristics of the clusters located in the STS (STG-MTG clusters in Fig. 2). *x*, *y*, and *z* correspond to the stereotaxic coordinates (mm) in the MNI reference brain provided with SPM2. All *t* max reported in the table correspond to a $P < 0.001$. Cluster size represents the number of contiguous voxels showing significant activation ($2 \times 2 \times 2$ mm voxels). % of signal change represents the increase ("footstep − rest") expressed as percent of global activity. This percentage was calculated for the mean activity of the cluster.

Results

All the subjects performed the task with very few errors (less than 2 errors on 48 trials, for each subject). The reaction times for the 4 delays (1, 2, 3, 4 s) were (mean ± SEM) 1525.5 ± 81, 1350.4 ± 94, 1436.5 ± 95, and 1204.8 ± 62 ms, respectively.

The perception of the human footsteps, compared to the resting state, was associated with a significant increase of the BOLD signal in the areas described in Table 1 and illustrated in Fig. 2. These regions included bilateral frontal and parietal areas. They also included areas of the temporal lobe extending bilaterally from the antero-medial part (Heschl's gyrus) to the posterior end of the superior temporal sulcus (superior and middle temporal gyrus). To show the reproducibility of these effects, the "footstep − rest" contrast was computed in each individual subject, for each cluster identified in the group study. The number of subjects showing a significant effect, measured on the mean activity of each of those clusters, is indicated in Table 1. In addition, each subject was analyzed individually. The "footstep − rest" contrast allowed characterizing the individual STS clusters (Table 2). Their location on the individual MRIs is provided in Fig. 3. One can note that the locations of the maximum t value (Table 2) were in the vicinity of those found in the clusters identified from the group analysis (Table 1).

We then performed a parametric regression analysis on the mean activity of the clusters revealed in the group analysis, by the "footstep − rest" contrast. The "attention period" (i.e., the time from stimulus onset to subject's response during which subject's

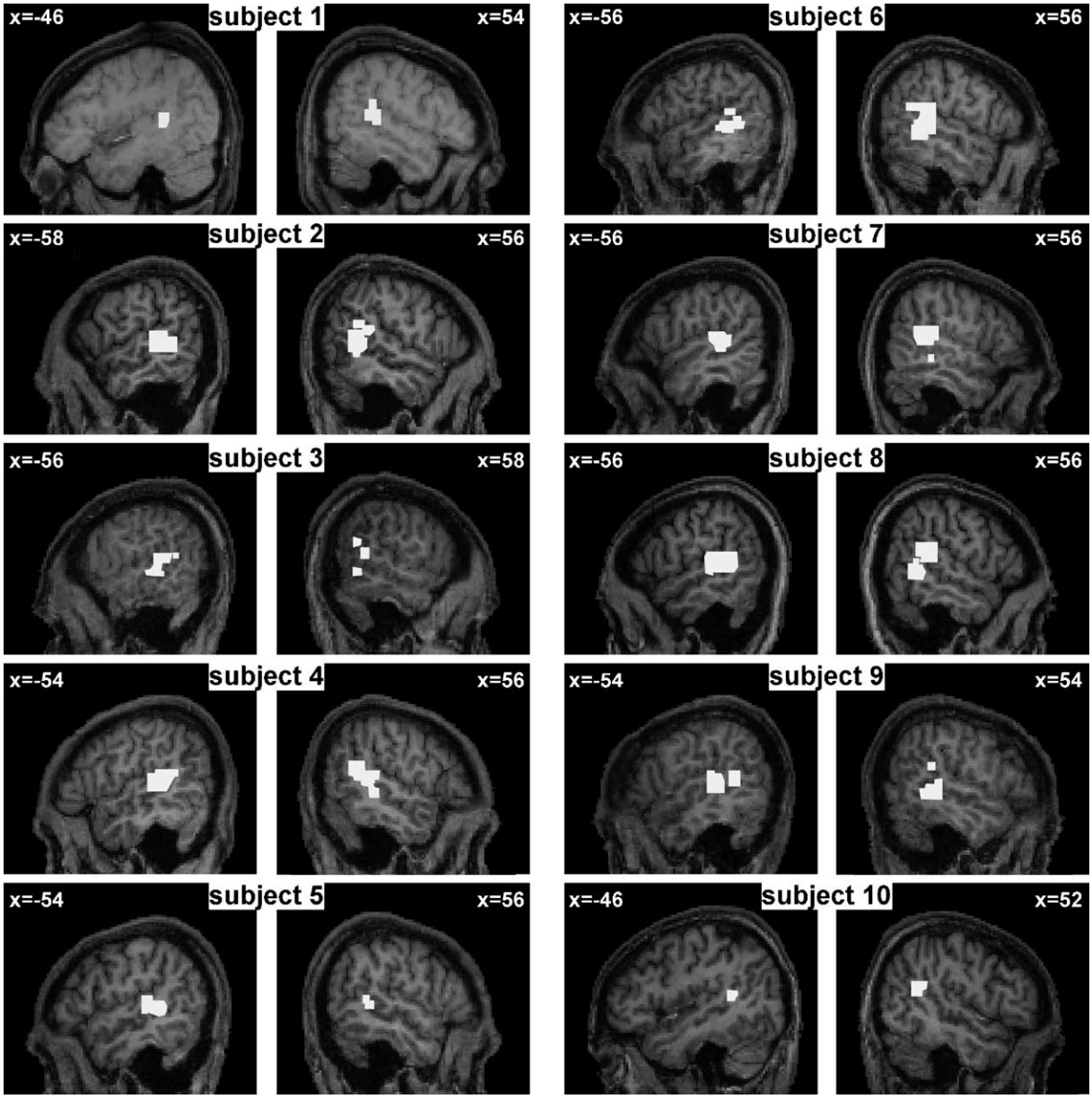

Fig. 3. Topographies of the temporal activations during footstep perception ("footstep − rest") in the 10 individual subjects. Activities are presented in white on left and right sagittal views for each individual MRIs. The statistical maps are thresholded at $t = 3.4$ ($P < 0.01$). See Table 2 for the detailed characteristics of these clusters.

Table 3
Characteristics of the temporal brain clusters activated during footstep perception or noise detection

	MNI coordinates			*t* max	
	x	*y*	*z*		
Footstep perception > noise detection					
R superior temporal sulcus	52	−50	9	7.00	139
L superior temporal sulcus	−50	−41	9	7.73	110
During both footstep perception and noise detection					
L planum temporale	−50	−28	7		88
R planum temporale	63	−31	11		37

Footstep perception > noise detection: Results are derived from a mixed effect analysis (10 subjects). *x*, *y*, and z correspond to the stereotaxic coordinates (mm) in the MNI reference brain provided with SPM2. The contrast used to detect the clusters is "(footstep − rest) − (noise − rest)". The map is thresholded at $P < 0.01$. *t* max is the maximum *t* found among the voxels constituting the cluster. *Both footstep perception and noise detection:* The clusters are defined by the voxels for which the values of the "footstep − rest" and "noise − rest" contrasts (intersection of the maps "footstep − rest" and "noise − rest") both correspond to $P < 0.05$.

attention was actually directed towards footsteps) was taken as the continuous factor. The activity of three regions only significantly correlated with the attention period ($P < 0.05$): the right parietal cortex (MNI coordinates: +49, −50, +37 mm; *t* value = 2.34), the right dorsolateral prefrontal cortex (+55, +19, +29 mm; *t* value = 1.84), and the supplementary motor area (SMA) (+4, +7, +57 mm; *t* value = 3.81).

To disclose, within the temporal lobe, the activities due to low-level sound processing from those related to human footstep processing per se, we compared the areas identified in the footstep protocol to those found in the auditory detection protocol in which sounds did not represent any known object. Computing the "(footstep − rest) − (noise − rest)" contrast allowed to detect a bilateral cluster in the STS, exhibiting a significant greater increase of activity during footstep perception than during noise detection (Table 3, upper panel). To determine the origin of this difference, we computed separately the "footstep − rest" and "noise − rest" contrasts, measured on the mean activity of each of the two clusters. A significant activation was found for "footstep − rest" (right: $t = 5.24$, $P < 0.001$; left: $t = 5.72$, $P < 0.001$) whereas no significant increase was found for "noise − rest" (right: $t = -0.93$; left: $t = 1.17$). This bilateral cluster included voxels from both the superior and the medial temporal gyri, in the depth of the STS. More precisely, it was located in the posterior part of the STS exhibiting a strong vertical inflexion (in the vicinity of the parieto-temporo-occipital junction). Besides, we have searched for voxels activated during both footstep perception and noise detection (Table 3, lower panel). The intersection of the maps "footstep − rest" and "noise − rest", thresholded at $t = 2.3$ ($P < 0.05$), showed a bilateral cluster in the superior temporal gyrus. By computing the "footstep − rest" and "noise − rest" contrasts, measured on the mean activity of each of these intersection clusters, we confirmed the significant increase during both footstep perception (right: *t* =

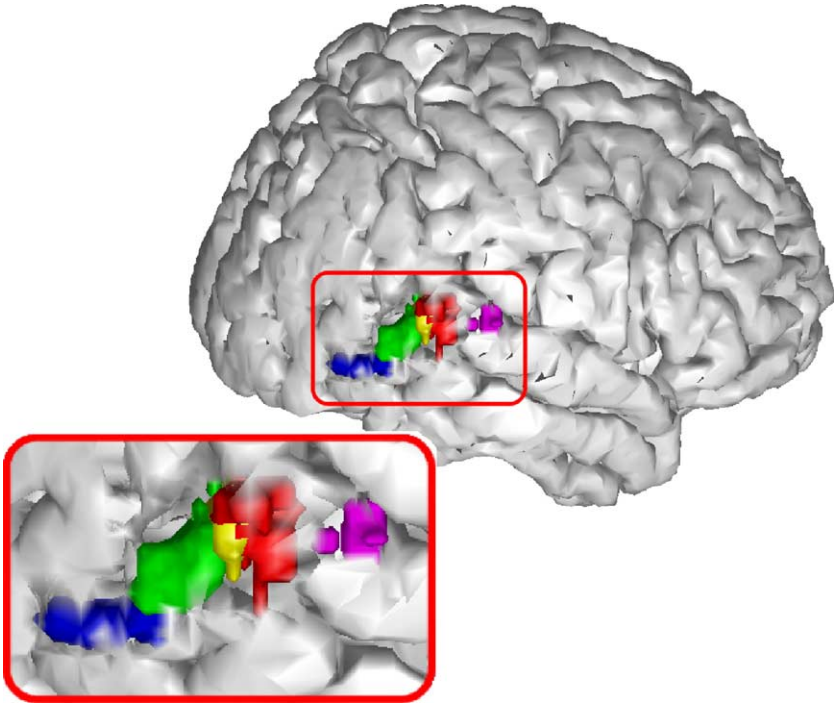

Fig. 4. Topography of occipito-temporal activations during footstep perception or (visual) biological motion perception. Activations related to biological motion perception via visual input are derived from the data obtained in a previous experiment (Grèzes et al., 2001). Blue: V5 area, showing activation during both rigid motion and biological motion perception (see Grèzes et al., 2001 for details). Magenta: planum temporale, activated by both meaningless sounds and footsteps. Green: posterior STS, showing activation during biological motion but not during rigid motion perception (Grèzes et al., 2001). Red: posterior STS, showing activation during footstep perception but not during meaningless sound detection. Yellow: posterior STS showing activation during biological motion but not during rigid motion perception (Grèzes et al., 2001) and activation during footstep perception but not during meaningless sound detection.

3.95, $P < 0.01$; left: $t = 3.19$, $P < 0.02$) and noise detection (right: $t = 4.99$, $P < 0.001$; left: $t = 4.59$, $P < 0.005$). These clusters were located in the superior part of the STG (in a location above and anterior to the STS) including the planum temporale.

Discussion

Listening to footsteps produced by walking humans was shown to activate a large network of brain areas. Several sub-systems could be dissociated according to their putative functional role.

A first sub-system, frequently associated with general attention processes, included thalamic nuclei, parietal areas, dorsolateral prefrontal cortex and supplementary motor area (SMA). SMA activation likely reflects the preparation of the motor response required by the task (Dum and Strick, 2002; Rushworth et al., 2004), and thalamic activation may relate to general arousal (Shipp, 2004; Pinault, 2004). The parietal and dorsolateral prefrontal areas found here are generally thought to modulate the activity of the visual pathway during attention related tasks (Chaminade and Fonlupt, 2003; Kanwisher and Wojciulik, 2000; Mazoyer et al., 2002; Rees and Lavie, 2001; Rees et al., 1997). Dorsolateral prefrontal area has also been found to be associated with auditory attention (Bushara et al., 1999; Lewis et al., 2000). Moreover, the areas best regressing with the "attention period" (which can be related to the amount of attention needed to perform the task), parietal and dorsolateral prefrontal cortices, are similar to those found when visually monitoring biological motion in a paradigm based on Johansson point-light walkers (Chaminade and Fonlupt, 2003; Mazoyer et al., 2002). This similarity is consistent with a role of these areas in a supramodal control of attentional processes.

A second sub-system, more specific to active auditory processing, encompassed temporal areas, including Heschl's gyrus and planum temporale, and an inferior frontal area located at the anterior end of the insula. The inferior frontal area has been associated to several high-level auditory processes, especially in allocating auditory attention (Alain et al., 2001; Binder et al., 2004; Hall et al., 2000; Lipschutz et al., 2002). Heschl's gyrus and planum temporale, also activated during detection of meaningless sounds, are likely to concern low-level stages of auditory processing (Griffiths and Warren, 2002; Warren and Griffiths, 2003).

A third cluster of activation was found along the posterior STS during footstep perception but not during the detection of a meaningless sound that does not represent any known object. Its location is clearly superior and anterior to area V5 (as defined by Dumoulin et al., 2000), i.e., in the posterior STS. Similar STSp activation has been found in all of the fMRI studies addressing visual biological motion perception (Beauchamp et al., 2002, 2003; Grèzes et al., 2001; Grossman and Blake, 2002; Grossman et al., 2000; Howard et al., 1996; Pelphrey et al., 2003; Ptito et al., 2003; Santi et al., 2003; Servos et al., 2002; Vaina et al., 2001). This region has also been described as a key element in a network of areas responsible for social cognition (for review, Puce and Perrett, 2003). In addition, it has been involved in polysensory interaction (Beauchamp et al., 2004; Puce and Perrett, 2003; Van Atteveldt et al., 2004; Wright et al., 2003), environmental sound recognition (Lewis et al., 2004), and voice processing (Belin et al., 2002; Kriegstein and Giraud, 2004).

Furthermore, we paralleled the present results to those previously obtained in an analogous visual experiment on biological motion perception (Grèzes et al., 2001). Although both these experiments (the present one and Grèzes's) involved different subjects, we consider that this comparison is valid because (i) the very same analysis procedure was applied to the data, and (ii) the activated loci exhibited low inter-subject variability. This comparison suggests (Fig. 4) that the recognition of human motion via auditory or visual input might share the same two-level pattern of organization. The first component is specific of the sensory modality and could represent a low-level analysis of the input signal: area V5 (blue) and posterior planum temporale (magenta) for visual and auditory input, respectively. The second component in the posterior STS would be involved in higher-level integration of meaningful stimulus, consistent with recent neuroimaging findings (Beauchamp et al., 2004; Wright et al., 2003). More precisely, red and yellow areas in Fig. 4 are activated by footstep sounds, and green and yellow areas by visual biological motion. One can speculate that the yellow area, activated by both auditory and visual inputs, could be part of the network underlying the human motion concept. Additional work is needed to further validate this model, for instance by using functional connectivity analysis methods.

In conclusion, our results showing posterior STS activation when listening to human motion are in line with a role of this region in the recognition of other people's action and social perception.

Acknowledgments

We thank JL. Anton, B. Nazarian, and M. Roth from the fMRI Centre, Marseille, France for their helpful assistance. Funding from the Fondation Recherche Médicale (FRM fellowship to J. Voisin) is gratefully acknowledged.

References

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

Battelli, L., Cavangh, P., Thornton, I., 2003. Perception of biological motion in parietal patients. Neuropsychologia 41, 1808–1816.

Beauchamp, M., Lee, K., Haxby, J., Martin, A., 2002. Parallel visual motion processing streams for manipulable objects and human movements. Neuron 34, 149–159.

Beauchamp, M., Lee, K., Haxby, J., Martin, A., 2003. fMRI responses to video and point-light displays of moving humans and manipulable objects. J. Cogn. Neurosci. 15, 991–1001.

Beauchamp, M., Argall, B., Bodurka, J., Duyn, J., Martin, A., 2004. Unraveling multisensory integration: patchy organization within human STS multisensory cortex. Nat. Neurosci. 7, 1190–1192.

Beintema, J., Lappe, M., 2002. Perception of biological motion without local image motion. Proc. Natl. Acad. Sci. U. S. A. 99, 5661–5663.

Belin, P., Zatorre, R.J., Ahad, P., 2002. Human temporal-lobe response to vocal sounds. Brain Res. Cogn. Brain Res. 13, 17–26.

Binder, J., Liebenthal, E., Possing, E., Medler, D., Ward, D., 2004. Neural correlates of sensory and decision processes in auditory object identification. Nat. Neurosci. 7, 295–301.

Bonda, E., Petrides, M., Ostry, D., Evans, A., 1996. Specific involvement of human parietal systems and the amygdala in the perception of biological motion. J. Neurosci. 16, 3737–3744.

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

Chaminade, T., Fonlupt, P., 2003. Changes of effective connectivity between the lateral and medial parts of the prefrontal cortex during a visual task. Eur. J. Neurosci. 18, 675–679.

Dum, R., Strick, P., 2002. Motor areas in the frontal lobe of the primate. Physiol. Behav. 77, 677–682.

Dumoulin, S., Bittar, R., Kabani, N., Baker, C., Le Goualher Jr., G., Bruce Pike, G., Evans, A., 2000. A new anatomical landmark for reliable identification of human area V5/MT: a quantitative analysis of sulcal patterning. Cereb. Cortex 10, 454–463.

Friston, K., Holmes, A., Worsley, K., Poline, J., Frith, C., Frackowiak, R., 1995. Statistical parametric maps in functional imaging: a general linear model approach. Hum. Brain Mapp. 2, 189–210.

Grèzes, J., Fonlupt, P., Bertenthal, B., Delon-Martin, C., Segebarth, C., Decety, J., 2001. Does perception of biological motion rely on specific brain regions? Neuroimage 13, 775–785.

Griffiths, T., Warren, J., 2002. The planum temporale as a computational hub. Trends Neurosci. 25, 348–353.

Grossman, E., Blake, R., 2001. Brain activity evoked by inverted and imagined biological motion. Vision Res. 41, 1475–1482.

Grossman, E., Blake, R., 2002. Brain areas active during visual perception of biological motion. Neuron 35, 1167–1175.

Grossman, E., Donnelly, M., Price, R., Pickens, D., Morgan, V., Neighbor, G., Blake, R., 2000. Brain areas involved in perception of biological motion. J. Cogn. Neurosci. 12, 711–720.

Hall, D., Haggard, M., Akeroyd, M., Summerfield, Q., Palmer, A., Elliot, M., Bowtell, R., 2000. Modulation and task effects in auditory processing measured using fMRI. Hum. Brain Mapp. 10, 107–119.

Hirai, M., Fukushima, H., Hiraka, K., 2003. An event-related potentials study of biological motion perception in humans. Neurosci. Lett. 344, 41–44.

Howard, R., Brammer, M., Wright, I., Woodruff, P., Bullmore, E., Zeki, S., 1996. Curr. Biol. 6, 1019–1050.

Johansson, G., 1973. Visual perception of biological motion and a model for its analysis. Percept. Psychophys. 14, 201–211.

Kanwisher, N., Wojciulik, E., 2000. Visual attention: insights from brain imaging. Nat. Rev. Neurosci. 1, 91–100.

Kriegstein, K.V., Giraud, A.L., 2004. Distinct functional substrates along the right superior temporal sulcus for the processing of voices. Neuroimage 22, 948–955.

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

Lewis, J., Wightman, F., Brefczynski, J., Phinney, R., Binder, J., DeYoe, E., 2004. Human regions involved in recognizing environmental sounds. Cereb. Cortex 14, 1008–1021.

Lipschutz, B., Kolinsky, R., Damhaut, P., Wikler, D., Goldman, S., 2002. Attention-dependent changes of activation and connectivity in dichotic listening. Neuroimage 17, 643–656.

Mazoyer, P., Wicker, B., Fonlupt, P., 2002. A neural network elicited by parametric manipulation of the attention load. NeuroReport 13, 2331–2334.

Pavlova, M., Lutzenberger, W., Sokolov, A., Birbaumer, N., 2004. Dissociable cortical processing of recognizable and non-recognizable biological movement: analyzing gamma MEG activity. Cereb. Cortex 14, 181–188.

Pavlova, M., Sokolov, A., Staudt, M., Marconato, F., Birbaumer, N., Krðgeloh-Mann, I., 2005. Recruitment of periventricular parietal regions in processing cluttered point-light biological motion. Cereb. Cortex 15, 594–601.

Pelphrey, K., Mitchell, T., McKeown, M., Goldstein, J., Allison, T., McCarthy, G., 2003. Brain activity evoked by the perception of human walking: controlling for meaningful coherent motion. J. Neurosci. 23, 6819–6825.

Pinault, D., 2004. The thalamic reticular nucleus: structure, function and concept. Brain Res. Rev. 46, 1–31.

Ptito, M., Faubert, J., Gjedde, A., Kupers, R., 2003. Separate neural pathways for contour and biological-motion cues in motion-defined animal shape. Neuroimage 19, 246–252.

Puce, A., Perrett, D., 2003. Electrophysiology and brain imaging of biological motion. Philos. Trans. R. Soc. London, B 358, 435–445.

Rees, G., Lavie, N., 2001. What can functional imaging reveal about the role of attention in visual awareness? Neuropsychologia 39, 1343–1353.

Rees, G., Frackowiak, R., Frith, C., 1997. Two modulatory effects of attention that mediate object categorization in human cortex. Science 275, 835–838.

Rushworth, M., Walton, M., Kennerley, S., Bannerman, D., 2004. Action sets and decisions in the medial frontal cortex. Trends Cogn. Sci. 8, 410–417.

Santi, A., Servos, P., Vatikiotis-Bateson, E., Kuratate, T., Munhall, K., 2003. Perceiving biological motion: dissociating visible speech from walking. J. Cogn. Neurosci. 15, 800–809.

Seifritz, E., Esposito, F., Hennel, F., Mustovic, H., Neuhoff, J., Bilecen, D., Tedeschi, G., Scheffler, K., Di Salle, F., 2002. Spatiotemporal pattern of neural processing in the human auditory cortex. Science 297, 1706–1708.

Servos, P., Osu, R., Santi, A., Kawato, M., 2002. The neural substrates of biological motion perception: an fMRI study. Cereb. Cortex 12, 772–782.

Shipp, S., 2004. The brain circuitry of attention. Trends Cogn. Sci. 8, 223–230.

Vaina, L., Solomon, J., Chowdhury, S., Sinha, P., Belliveau, J., 2001. Functional neuroanatomy of biological motion perception in humans. Proc. Natl. Acad. Sci. U. S. A. 98, 11656–11661.

Van Atteveldt, N., Formisano, E., Goebel, R., Blomert, L., 2004. Integration of letters and speech sounds in the human brain. Neuron 43, 271–282.

Warren, J., Griffiths, T., 2003. Distinct mechanisms for processing spatial sequences and pitch sequences in the human auditory brain. J. Neurosci. 23, 5799–5804.

Wicker, B., Fonlupt, P., 2003. Generalized least-squares method applied to fMRI time series with empirically determined correlation matrix. Neuroimage 18, 588–594.

Wright, T., Pelphrey, K., Allison, T., McKeown, M., McCarthy, G., 2003. Polysensory interactions along lateral temporal regions evoked by audiovisual speech. Cereb. Cortex 13, 1034–1043.