

Premotor Cortex Is Sensitive to Auditory–Visual Congruence for Biological Motion

Sophie M. Wuerger¹, Laura Parkes², Penelope A. Lewis²,
Alex Crocker-Buque³, Roland Rutschmann⁴,
and Georg F. Meyer¹

Abstract

■ The auditory and visual perception systems have developed special processing strategies for ecologically valid motion stimuli, utilizing some of the statistical properties of the real world. A well-known example is the perception of biological motion, for example, the perception of a human walker. The aim of the current study was to identify the cortical network involved in the integration of auditory and visual biological motion signals. We first determined the cortical regions of auditory and visual coactivation (Experiment 1); a conjunction analysis based on unimodal brain activations identified four regions: middle temporal area, inferior parietal lobule, ventral premotor cortex, and cerebellum. The brain activations arising from bimodal motion stimuli (Experi-

ment 2) were then analyzed within these regions of coactivation. Auditory footsteps were presented concurrently with either an intact visual point-light walker (biological motion) or a scrambled point-light walker; auditory and visual motion in depth (walking direction) could either be congruent or incongruent. Our main finding is that motion incongruency (across modalities) increases the activity in the ventral premotor cortex, but only if the visual point-light walker is intact. Our results extend our current knowledge by providing new evidence consistent with the idea that the premotor area assimilates information across the auditory and visual modalities by comparing the incoming sensory input with an internal representation. ■

INTRODUCTION

When an object moves in the real world, its movement is usually associated with a sensory signal in both the auditory and visual modalities (Baumann & Greenlee, 2007). These signals are merged to yield a unified percept of the object in motion. The auditory and visual perception systems have developed special processing strategies for ecologically valid motion stimuli, utilizing some of the statistical properties of the real world (for a recent review, see Blake & Shiffrar, 2007). A prime example is the perception of biological movement, that is, the perception of human body motion, such as walking or running.

The cortical mechanisms underlying the processing of visual biological motion signals (such as point-light walkers) have received much attention, and a network encompassing occipital, parietal, and temporal areas has been implicated in the processing of visual biological motion, including the posterior superior temporal gyrus and STS (Pelphrey, Morris, Michelich, Allison, & McCarthy, 2005; Thompson, Clarke, Stewart, & Puce, 2005; Pelphrey et al., 2003; Grossman & Blake, 2002; Servos, Osu, Santi, & Kawato, 2002; Grossman & Blake, 2001; Grossman et al., 2000; Bonda, Petrides, Ostry, & Evans, 1996; Howard

et al., 1996), the lingual gyrus (Vaina, Solomon, Chowdhury, Sinha, & Belliveau, 2001), motion-sensitive areas middle temporal (MT) and MT+ (Grezes, 2001; Vaina et al., 2001), parietal areas (Grezes, 2001; Vaina et al., 2001; Bonda et al., 1996), and other areas including the amygdala (Bonda et al., 1996).

The involvement of the pSTS/superior temporal gyrus in biological motion processing is the most robust finding and consistent with macaque physiology (for a review see Puce & Perrett, 2003). Many areas that are selective for visual biological motion are also responsive to auditory biological motion signals. The pSTS is activated by auditory footsteps (Bidet-Caulet, Voisin, Bertrand, & Fonlupt, 2005), hence suggesting that pSTS may be a supramodal integration area for human biological motion.

More recent experiments suggest that, in addition to the STS, premotor areas play an important role in the processing of visual biological motion (Schubotz & von Cramon, 2004) and studies using a clinical (Saygin, 2007) or non-clinical population (Saygin, 2007; Saygin, Wilson, Hagler, Bates, & Sereno, 2004) confirm that the premotor cortex is necessary for intact biological motion perception. Neuroimaging studies on humans have demonstrated that premotor cortex is activated during action observation (e.g., Bonini et al., 2010; Buch, Mars, Boorman, & Rushworth, 2010; Jastorff, Begliomini, Fabbri-Destro, Rizzolatti, & Orban, 2010; Pilgramm et al., 2010; Calvo-Merino, Glaser,

¹University of Liverpool, ²University of Manchester, ³University of Edinburgh, ⁴University of Regensburg

Grezes, Passingham, & Haggard, 2005) and that auditory and visual motion signals converge in the premotor cortex (Bremmer et al., 2001). Taken together, these studies suggest that the human premotor cortex is a good candidate for the perceptual integration of auditory and visual actions, such as human body motions.

Behavioral evidence suggests that different integration mechanisms are at work for highly familiar auditory and visual signals (Arrighi, Marini, & Burr, 2009; Saygin, Driver, & de Sa, 2008; Arrighi, Alais, & Burr, 2006). RT studies with biological motion stimuli (point-light walkers) showed that the integration of biological motion stimuli is constrained by the direction of the auditory and visual motion signals and shorter RTs are reported for congruent biological motion (Brooks et al., 2007); the integration of random motion sequences is not affected by the inconsistency of the auditory–visual motion direction (Brooks et al., 2007; Meyer, Wuerger, Roehrbein, & Zetzsche, 2005; Wuerger, Hofbauer, & Meyer, 2003; Meyer & Wuerger, 2001). In the present imaging study, we looked for neural correlates of these differential auditory–visual integration mechanisms for biological and nonbiological motion signals that have been demonstrated behaviorally. As visual biological motion stimuli, we used point-light walkers (Johansson, 1973), because they give a compelling percept of a person walking and yet are highly controllable; a “scrambled” walker was obtained by randomizing the starting position of each limb, hence keeping the local motion signals intact but destroying the percept; the auditory stimulus consisted synchronized footsteps. We focussed on the question whether the incongruent auditory and visual motion direction has a differential effect on the brain activity arising from the integration of biological (point-light walker and synchronized footsteps) and nonbiological motion signals (“scrambled” walker and synchronized footsteps). Our hypothesis was that inconsistent motion across the auditory and visual modalities (auditory: looming motion; visual: receding motion) should have a greater effect when both modalities signal biological motion.

METHODS

Experimental Design

First, we identified candidate regions (ROIs) of auditory–visual coactivation (Experiment 1: Localizer); we then tested within these ROIs whether such differential neural activities were found for biological compared with scrambled motion sequences (Experiment 2). In Experiment 1 (localizer), subjects were presented with visual (point-light walkers), auditory (footsteps), or bimodal motion sequences, and their task was to detect motion in depth (looming or receding motion). fMRI scans were performed to reveal cortical activations common to the auditory and visual modalities (Harrison, Wuerger, & Meyer, 2010; Bremmer et al., 2001). The main purpose of the localizer experiment was to identify areas of

auditory–visual coactivation by performing a conjunction analysis (Friston, Penny, & Glaser, 2005) of the unimodal (auditory only, visual only) brain activations. In Experiment 2, we tested our main hypothesis by asking whether auditory–visual motion congruency (same vs. different directions of motion in the two modalities) yields a differential effect on neural responses to biological motion in comparison with meaningless motion sequences. fMRI was performed while subjects were presented with incongruent and congruent bimodal motion sequences. The statistical analysis of the effect of motion congruency on biological versus nonbiological motion is then performed within the ROIs defined by Experiment 1 (Meyer, Greenlee, & Wuerger, 2011; Szyck, Tausche, & Münte, 2008). Behavioral performances for both experiments were obtained at least 1 day before the scanning sessions under closely matched experimental conditions.

Subjects

Eighteen (15 naive and 3 authors) healthy volunteers (eight women) with normal or corrected-to-normal vision participated in the experiments (mean age = 24 years, $SD = 5$ years). All subjects gave written consent and were screened for MRI contraindications. The study was approved by the Sefton Liverpool Research Ethics Committee.

Apparatus

Auditory stimuli were played back using a real-time signal processor (RM1, Tucker-Davis Technologies, Alachua, FL) and presented via MRI-compatible MR Confon Optime 1 headphones (MR Confon, Magdeburg, Germany). Visual stimuli were generated using a visual stimulus generator (ViSaGe; Cambridge Research Systems Ltd., Kent, United Kingdom), which was controlled by a standard PC (Dell Precision 390). Stimuli were back projected with an LCD projector (Panasonic PT-L785U) onto a translucent circular screen, placed inside the scanner bore at 70 cm from the observer. The projector ran at a refresh rate of 60 Hz and a resolution of 800×600 pixels. The TDT system and the ViSaGe system were interfaced via triggers to ensure that the auditory and visual stimuli were synchronized. For stimulus presentation (auditory and visual) MatLab 7 (Mathworks) was used. Responses were acquired using an MRI-compatible response box.

Behavioral data were obtained at least 1 day before the scanning session using a similar experimental setup (ViSaGe interfaced with a TDT system). Subjects were seated in a sound-proof booth (IAC 404-A), at a distance of 100 cm from a CRT monitor (Mitsubishi DiamondPro 2070SB), running at a refresh rate of 60 Hz. Auditory stimuli were presented via conventional headphones (Sennheiser HD25SP). RTs were acquired using an infrared response box (Cambridge Research Systems Ltd., Kent, United Kingdom).

Stimuli

The auditory stimuli were natural recordings of footsteps (male walker) on gravel and lasted 1.8 sec (four footsteps; diotic presentation, $F_s = 44,100$ Hz, 64 dB(A)). The visual stimuli were either “point-light walkers” (PLW; biological motion) or “scrambled point-light walkers” (SCR), subtending a visual angle of 3.8° (width) \times 10° (height). The mean luminance of the display was fixed at 50 cd/m^2 ; the contrast of the PLWs was 100% (black on gray). The PLW was defined by 13 points (indicating the main joints and the head) representing the motion of the particular position of the body over four steps. PLWs were always presented in their front/back view. The view we presented was consistent with a front and back view because of the inherent orthographic ambiguity of PLWs (Vanrie & Verfaillie, 2006); it is also known that a concurrent auditory looming/receding sound can bias the observer’s interpretation (Schouten, Troje, Vroomen, & Verfaillie, 2011). Each point had a size of 3×3 pixels ($0.09 \times 0.09^\circ$), and one stimulus trial lasted for 1.8 sec. The “scrambled” walkers were generated by using the same local limb movements as present in the PLW, but the starting positions of the limb movements were randomized within a kernel defined by the extent of the original figures, for example, the knee movement could start near the elbow and vice versa. New scrambled motion was generated on each trial to avoid that observers learned the constellation of the scrambled walkers. The advantage of this control stimulus is that it contains the same local motion signals (hence the same spatio-temporal profile) as the point-light walker but is not recognized as a walker (Grossman & Blake, 2002). Auditory and visual motion stimuli could either be looming, receding, or neither looming nor receding. In the latter case, the point-light walker is walking “on a treadmill” (“no motion”). Receding visual motion was generated by contracting the visual stimuli by a factor of 0.25; receding auditory motion was generated by linearly decreasing the amplitude of the footsteps by the same factor. Looming motion was generated by linearly increasing the amplitude/size. We added dynamic visual noise to the visual stimuli in an attempt to roughly equate the saliency in both modalities, because the scanner noise was always present in the auditory modality. New dynamic visual noise was generated on each trial. To match the behavioral study (this was a separate experiment conducted before the brain scans) as closely as possible with the scanning conditions, we recorded the scanner noise using an optical microphone (MR Confon, Sennheiser, Germany) and then replayed the scanner noise in the sound-proof booth using loud speakers throughout the experiment. The auditory stimulus (footsteps) was presented via headphones. The onset of the (audio) footstep coincided with the (visual) foot touching the ground; this synchronization was performed manually.

Task and Procedure

We performed two experiments. In Experiment 1, we presented unimodal motion stimuli: auditory footsteps (A), visual biological motion (VBIO), visual scrambled motion (VSCR), and congruent bimodal stimuli (CONG_BIO = A + VBIO, CONG_SCR = A + VSCR). All five experimental stimuli conveyed the same motion direction (receding), and each experimental condition was presented 12 times. We included a control condition of no interest, which consisted “no motion” (walking on a treadmill) stimuli, presented either bimodally or unimodally. Each of the five control stimuli was presented four times and the task of the participant was to press a button when no motion was present. In addition, we included 20 null events (fixation target only) at random times. The stimuli (experimental, control, null) were presented in a randomized order; each stimulus was presented for 1.8 sec, and the average times between stimuli was 3 sec with a randomized jitter between -0.5 and $+0.5$ sec. Altogether, Experiment 1 consisted 100 trials and lasted just under 7 min (200 scans).

In Experiment 2 (main experiment), we tested whether auditory–visual congruency produces differential brain responses to VBIO compared to VSCR. In the four experimental conditions, auditory and visual motion could either move in the same direction (both receding: CONG_BIO, CONG_SRC) or in different directions (auditory looming and visual receding: INCONG_BIO, INCONG_SCR). Within a single scan, each of the experimental stimuli was presented 16 times. As in the localizer, we included two control conditions of no interest, consisting bimodal “no motion” stimuli (A + VBIO or A + VSCR), and each of the two control stimuli was presented 12 times. Twenty-two null events were included, and all stimuli were presented in a randomized order. Altogether, Experiment 2 consisted 110 trials and lasted slightly longer than 7 min (219 scans).

Each subject was in the scanner for less than 1 hr. First, the participant performed a short practice experiment (less than 5 min); then two scan sessions of Experiment 1 were run (each about 7 min) followed by a structural scan (12 min) and two sessions of Experiment 2 (each about 7 min). For half of the participants, the order of Experiments 1 and 2 was reversed. In the scanner, the observers’ task was to press a button (with the right index finger) only when there was “no motion” present (control condition). This ensures that the brain activity in response to the motion conditions is not confounded with the button presses.

For RT measurements, apparatus, stimuli, and procedure were the same as in the scanning session; the only difference was that observers were asked to press one button when the stimulus contained any motion and another button when no motion was present to match the motor activity between the conditions. Participants were instructed to respond as fast and accurately as possible. Collecting behavioral RT data before the scanning ensured that subjects were familiar with the stimuli and

the task, and no additional learning occurred during scanning. To ascertain that the auditory and visual motion stimuli elicited reliable and comparable motion percepts, performance for discriminating between looming (receding) motion and “no motion” was measured before the main experiments with the same set of observers. Performance for discriminating between auditory motion and “no motion”: 93% correct (for looming motion), 86% (for receding motion), and 71% (for “no motion” stimuli); visual biological motion versus “no motion”: 96% (looming), 91% (receding), and 96% (no motion); visual scrambled motion versus “no motion”: 72% (looming), 89% (receding), and 88% (no motion). In the main neuroimaging experiment (Experiment 2), we used auditory receding and visual receding motion to yield the congruent bimodal motion condition and auditory looming and visual receding motion to yield the incongruent biomodal motion condition. We are, therefore, confident that the stimuli used in the scanner elicited reliable and comparable auditory and visual motion percepts. This was confirmed in the localizer analysis (Figure 1; Supplementary Table S1), which showed activation patterns typical for the perception of auditory (Bidet-Caulet et al., 2005) or visual motion (e.g., Bremmer et al., 2001).

Data Acquisition

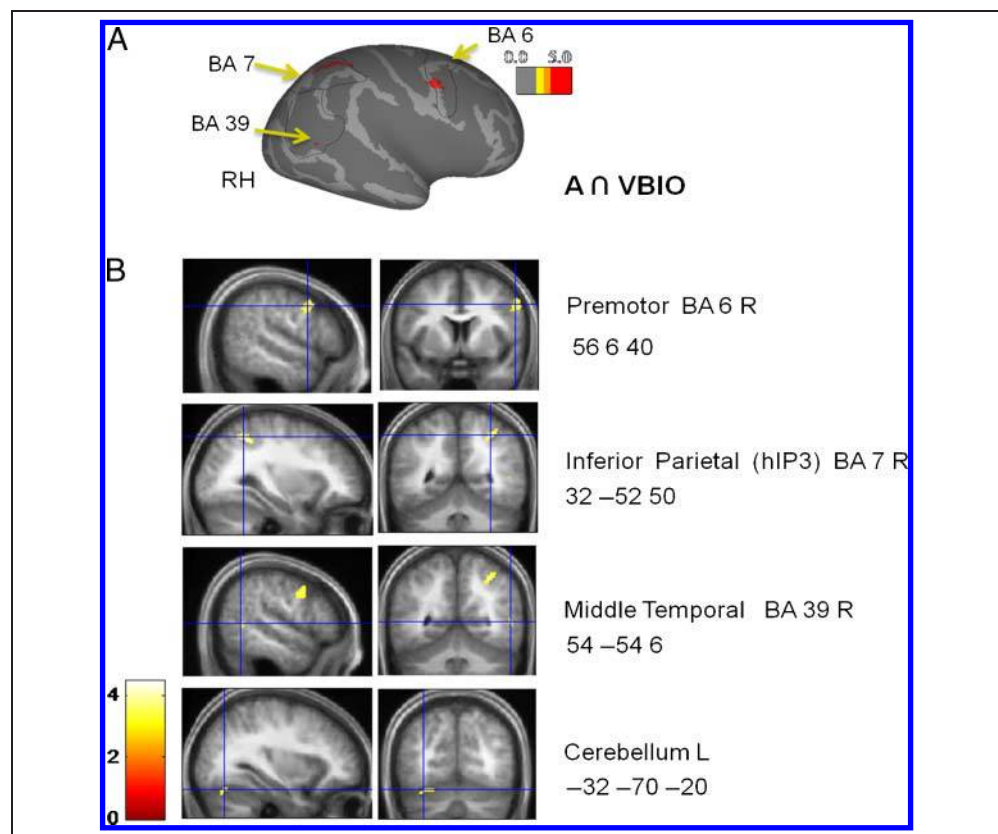
Imaging was performed using a 3-T MR whole body scanner (Siemens Trio, Erlangen, Germany) located at MARIARC,

University of Liverpool. In the functional scans, BOLD responses were measured using a T2*-weighted EPI sequence (echo time = 30 msec, volume repetition time = 2.0 sec, in-plane resolution = 3×3 mm, number of slices = 33, interleaved and ascending, slice thickness = 3 mm, gap between slices = 0.3 mm, flip angle = 80°). 3-D structural images of the whole brain were acquired using a T1-weighted MDEFT Sequence of 1-mm isotropic resolution.

Data Analysis

Preprocessing and statistical data analysis were performed using SPM5 (Wellcome Department of Imaging Neuroscience, London, United Kingdom, www.fil.ion.ucl.ac.uk/spm/) running under Matlab 7 (Mathworks, Natick, MA). Functional images of each participant were corrected for residual head motion and realigned to the first image. Subsequently, all functional images were coregistered and normalized to the MNI-152 template and resampled to $2 \times 2 \times 2$ mm³ spatial resolution. Spatial smoothing was applied to the functional images using an isotropic Gaussian kernel with a FWHM of 8 mm. A general linear model was constructed for each participant to analyze the hemodynamic responses captured by the functional images. In all functional scans, an event-related design was used; regressors were generated by convolving unit impulses with the canonical hemodynamic function and also with the temporal derivative of this function (e.g., Henson, Price,

Figure 1. Experiment 1. The conjunction analysis for auditory footsteps and biological visual motion ($A \cap VBIO$) revealed four regions of neural activity common to the auditory and visual modalities ($p_{FWE} < .05$; cf. Table 1). (A) SPM t maps are depicted on an inflated PALS-B12 standard brain (Caret 5.6; Van Essen et al., 2001). (B) The SPM t maps are projected onto the average of the normalized brains of all 18 participants. The color represents the t values for each cortical location as indicated by the key on the left.



Rugg, & Friston, 2002). A random effect analysis was used for the statistical fMRI data analysis.

Experiment 1 was used to localize modality-unspecific motion-sensitive areas. The design matrix consisted 10 regressors, the five experimental stimulus conditions (A, VBIO, VSCR, A + VBIO, A + VSCR, all depicting receding motion) and the five control conditions (A, VBIO, VSCR, A + VBIO, A + VSCR, all depicting a stationary “treadmill” walker). A second-level global null analysis (as defined by Friston et al., 2005) was used to reveal areas that respond significantly (whole-brain family-wise error < 0.05) to motion in the auditory or in the visual modality. We confirmed that a conjunction null (as defined by Friston et al., 2005) revealed the same areas of coactivation (at a different family-wise error); hence, in our particular case, this was not a critical issue. These brain areas identified in Experiment 1 by the global null analysis are then used as ROIs in Experiment 2. These ROIs were extracted using the MarsBaR 0.38 toolbox for SPM (Brett, Anton, Valabregue, & Poline, 2002).

In Experiment 2, we tested our main hypothesis, namely whether there is an interaction between auditory–visual congruency (CONG vs. INCONG) and motion type (BIO vs. SCR). The design matrix consisted six regressors: the four experimental conditions (CONG_BIO, CONG_SCR, INCONG_BIO, INCONG_SCR) and the two control conditions. Individual contrast estimates, within the ROIs defined by Experiment 1, were extracted for each observer and for each ROI individually. They were then analyzed with a two-way ANOVA (factor 1, Motion type: BIO or SCR; factor 2, Motion Congruency: congruent or incongruent). Stereotaxic Montreal Neurological Institute (MNI) coordinates are used throughout this report. For the parietal lobe activations, the centers of gravity of suprathreshold regions were localized using the Anatomy toolbox for SPM (Eickhoff et al., 2005). For cortical areas where no probability maps were available in the Anatomy toolbox, we used the WFU_PickAtlas toolbox for SPM (Maldjian, Laurienti, Kraft, & Burdette, 2003).

To compute the correlations between the behavioral data (RTs) and the brain activations, we use the mean RTs for each individual observer for each of the four experimental conditions (CONG_BIO, CONG_SCR, INCONG_BIO, INCONG_SCR) and the individual contrast values associated with the four experimental conditions in each of the four ROIs. These contrast values are proportional to signal change and were extracted with MarsBaR (Brett et al., 2002); for the correlation analysis, the mean contrast value averaged across all voxels within the ROI was used. To test for interactions between Motion Type (BIO/SCR) and Motion Congruency (CONG/INCONG) both in the behavioral RTs and the fMRI contrasts, we performed a within-subject two-way ANOVA (MatLab statistics toolbox).

The main hypothesis was tested as described in the previous paragraphs. For visualization purposes (figures are supplied as Supporting Material), a whole-brain analysis was conducted. Using a flexible factorial design, sev-

eral contrasts (CONG_BIO vs. null; CONG_SRC vs. null; INCONG_BIO vs. null; INCONG_SCR vs. null) were calculated. The resulting SPM *t* maps were superimposed with the selected threshold (family-wise error < 0.05) onto the population average landmark and surface-based (PALS-B12) standard brain (Van Essen, 2005) using Caret 5.6 (Van Essen et al., 2001).

RESULTS

Localizer Experiment: Areas of Auditory–Visual Coactivation

In the localizer experiment, we observe very similar activation patterns for biological and scrambled visual motion. The main purpose of the localizer experiment is to define ROIs in which the main hypothesis can be tested. The conjunction (“global null”) analyses (Friston et al., 2005) were performed on the unimodal brain activations, (A > Rest) \cap (V > Rest), for both biological and scrambled visual motion, following Meyer et al. (2011). The conjunction “A \cap VBIO” revealed four areas of significant coactivations common to the auditory and visual modalities: the right ventral premotor area (vPM; BA 6, bordering on BA 44), the right inferior parietal lobule (IPL; BA 7) on the border to the superior parietal lobule (SPL), the right MT area (BA 39, bordering on BA 22 and BA 37), and the left cerebellum. Figure 1A shows the SPM *t* maps of this conjunction analysis (group results) superimposed on an inflated standard brain; Figure 1B shows the sagittal and coronal views. The coactivity in the premotor cortex, the IPL, and area MT is lateralized in the right hemisphere; common activity in the cerebellum is only present in the left hemisphere. The corresponding figure for the conjunction “A \cap VSCR” is shown in the supporting material (Supplementary Figure S1); the same regions of coactivations are revealed. Table 1 depicts the label of the ROI, the type of conjunction (A \cap VBIO or A \cap VSCR), the cortical location (MNI), and the number of significant voxels. Both *t* and *z* values are given; all neural activations are significant at $p < .05$ (family-wise error). Because both localizers reveal very similar ROIs, we will report the results of our main experiment for the BIO localizer only; the corresponding (and identical) results for the SCR localizer can be found in the Supplementary Material.

Bimodal Activations

Differential Effects of Auditory–Visual Motion Incongruency on Biological and Scrambled Visual Motion

The purpose of the main experiment (Experiment 2) was to test whether the type of visual motion (biological or scrambled) interacts with motion incongruency (auditory and visual motion signal the same direction = congruent motion; auditory and visual motion signal different motion directions = incongruent motion). We measured activations

Table 1. Conjunction Analysis Revealing Activations Common to Auditory and Visual Modalities (Experiment 1)

Location		Localizer	Position (MNI)	Voxels	<i>t</i>	<i>z</i>	<i>p</i> _{FWE}
<i>Frontal Lobe</i>							
BA 6 R	Premotor (vPM)	A ∩ VBIO	56 6 40	152	3.74	5.53	.002
BA 6/44 R	Premotor (vPM)	A ∩ VSCR	48 4 32	521	4.89	7.0	<.001
			48 0 42		4.20	6.12	<.001
<i>Parietal Lobe</i>							
BA 7 R	IPL (hIP3: 40%; SPL (7PC): 30%; SPL (7A): 20%)	A ∩ VBIO	32 -52 52	207	4.46	6.46	<.001
			36 -44 54		3.74	5.53	.002
BA 7 R	IPL (hIP3: 30%; SPL (7PC): 30%; hIP1: 10%)	A ∩ VSCR	32 -50 50	282	4.66	6.70	<.001
			40 -40 52		3.68	5.45	.002
<i>Temporal Lobe</i>							
BA 39 R	MT	A ∩ VBIO	54 -54 6	10	3.33	5.00	.020
BA 39 R	MT	A ∩ VSCR	54 -54 6	7	3.27	4.92	.029
L	Cerebellum	A ∩ VBIO	-32 -70 -20	47	3.54	5.28	.006
L	Cerebellum	A ∩ VSCR	-30 -74 -20	12	3.21	4.85	.040

The conjunction analysis revealed four areas of auditory–visual coactivation (family-wise error < 0.05). “A ∩ VBIO” refers to the conjunction between the brain activations in response to auditory footsteps (A) and the brain activations in response to the visual point-light walker (VBIO); “A ∩ VSCR” refers to the conjunction analysis based on auditory footsteps and the scrambled point-light walker (VSCR). The conjunction analysis was performed using SPM5. For anatomical labeling of premotor cortex, the border between dorsal and vPM cortex was assumed at a *z* level of 50 in Talairach coordinates (Rizzolatti & Craighero, 2004); we converted the Talairach coordinates into MNI coordinates for our analysis.

for the four bimodal conditions: congruent biological motion (CON BIO), incongruent biological motion (INCON BIO), congruent scrambled motion (CON SCR), and incongruent scrambled motion (INCON SCR), and tested within each ROI (determined in Experiment 1 using our localizer) whether there is an interaction between motion type (BIO vs. SCR) and auditory–visual motion incongruency (congruent vs. incongruent), that is, whether the differential activation $(INCON - CON)_{SCR} - (INCON - CON)_{BIO}$ differs from zero. Our main finding is that significant interactions are found only in the right vPM.

Figure 2 shows the ROIs revealed by the localizer experiment (cf. Figure 1) superimposed onto an MNI normalized flat map template (Van Essen et al., 2001). BOLD contrasts within each ROI were extracted for each individual observer and the mean contrast differences between incongruent and congruent bimodal motion signals (“INCON – CON”) for biological (green) and scrambled (purple) motion are shown in the bar graphs for all four ROIs (for the numerical values of the contrast differences consult Table 2). In the right vPM, incongruent auditory–visual motion leads to a larger BOLD contrast increase when both modalities convey a biological motion signal in comparison with scrambled visual motion; the interaction is significant only in the vPM (within-subject two-way ANOVA: $F(1, 17) = 5.74$; $p = .028$). No significant interactions were found in IPL ($F(1, 17) = 0.54$; $p = .47$), in

MT ($F(1, 17) = 0.23$; $p = .63$), or in the cerebellum ($F(1, 17) < 0.0001$; $p = .97$). The significant interaction in vPM results from different BOLD contrasts for congruent and incongruent biological motion (BIO: top left of Supplementary Figure S3a, in the Supplementary Material); for the scrambled condition, congruent and incongruent motion yield the same BOLD contrasts (SCR: Supplementary Figure S3a). No significant contrast differences between congruent and incongruent motion were found in MT and the cerebellum; in IPL, there was a trend for incongruent biological motion to yield a higher BOLD contrast than congruent biological motion ($p = .066$; Supplementary Figure S3a).

We obtain almost identical results when we use a localizer defined by A ∩ VSCR, because the ROIs are almost completely overlapping (see Experiment 1): Only the interaction in vPM is significant (see Supplementary Material: Supplementary Figures S1, S2, and S3b; cf. with Figures 1 and 2 and Supplementary S3a). This differential effect of motion incongruency on biological motion can also be seen in the whole-brain group analysis: Incongruent motion is associated with an increased vPM (BA 6) activity for biological motion only and only in the right hemisphere (Supplementary Material: compare Supplementary Figure S4a: RH with S4b: LH).

In summary, our ROI analysis revealed a significant interaction in vPM (precentral; BA 6) in the right hemisphere

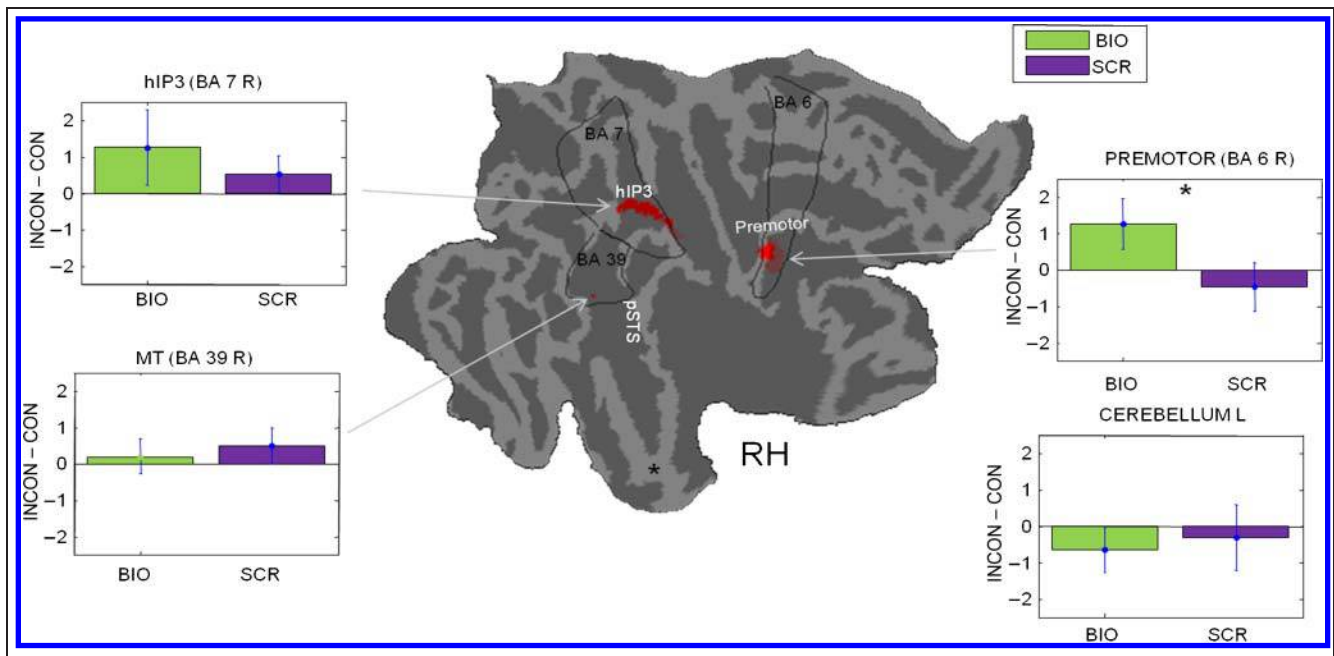


Figure 2. The location of the ROIs defined by the conjunction analysis ($A \cap VBIO$) are superimposed onto MNI normalized flat map template (Van Essen et al., 2001) and are shown in red. The fourth region is located in the cerebellum and is not shown here. The black lines represent the borders of the Brodmann's areas from the PALS-B12 atlas. The bar graphs show the contrast difference (INCONGRUENT – CONGRUENT) for biological (green) and scrambled (purple) motion. Only in the premotor cortex (vPM), incongruent auditory–visual motion leads to significant increase in the BOLD contrast when both modalities convey a biological motion signal as opposed to the visual scrambled condition. No significant interactions were found in IPL, MT, or cerebellum.

only: Incongruent motion in the auditory and visual modalities leads to an increase in the activation in these areas only if the auditory and visual modalities depict biological motion signals.

RTs and Their Neural Correlates

Figure 3 shows the differences in RTs (INCON – CON) for biological and scrambled visual motion. For biological

Table 2. Differential Activations for Biological and Scrambled Motion in ROIs

Location	Localizer	INCON BIO – CON BIO			INCON SCR – CON SCR		
		Contrast	t	p	Contrast	t	p
<i>Frontal</i>							
BA 6 R/premotor	$A \cap VBIO$	1.25	1.75	.041	–0.48	–0.84	.799
BA 6/44 R/premotor	$A \cap VSCR$	1.30	1.92	.028	–0.16	–0.30	.618
<i>Parietal</i>							
BA 7 R/IPL	$A \cap VBIO$	1.24	1.51	.066	0.52	0.79	.216
BA 7 R/IPL	$A \cap VSCR$	1.14	1.45	.075	0.47	0.74	.229
<i>Temporal</i>							
BA 39 R/MT	$A \cap VBIO$	0.20	0.30	.380	0.47	0.92	.178
	$A \cap VSCR$	0.23	0.35	.362	0.59	1.14	.128
Cerebellum L	$A \cap VBIO$	–0.67	–0.78	.781	–0.31	–0.46	.677
Cerebellum R	$A \cap VSCR$	–1.06	–1.14	.871	–0.04	–0.06	.997

No significant activation differences were found for scrambled motion, that is, the difference “INCON SCR – CON SCR” does not reach significance in any of the four ROIs. Only when the both modalities signal biological motion, significant differential activations are found in the Premotor cortex (BA 6) and to a lesser extent in IPL (BA 7).

motion, observers are slowed down (by 74 msec) when the auditory and visual modalities signal different directions of motion; when the visual point-light walker was scrambled, there is no significant RT difference between incongruent and congruent motion sequences (RT difference = -32 msec). There is a weak interaction between type of motion (BIO/SCR) and motion incongruency ($F(1, 17) = 3.73; p = .07$). In summary, observers are slowed down by incongruent information from the auditory and visual modalities if and only if both the auditory and the visual motion sequences depict biological motion, which is consistent with Brooks et al. (2007) and replicates our previously reported behavioral results (Wuerger, Crocker-Buque, & Meyer, in press).

Comparison of the differential brain activations (Figure 2) with the differential RTs (Figure 3) reveals that the BOLD contrast in vPM (BA 6) shows the same pattern as the RT, that is, an increase in RTs because of incongruent motion information from the auditory and visual modalities is associated with an increased activation in the premotor cortex. To quantify the strength of association between RTs and BOLD contrasts, we calculate the correlation between the individual brain activations within the ROIs and the individual RTs ($n = 18$) for all four experimental conditions (CON BIO; CON SCR; INCON BIO; INCON SCR). We predict an association between RTs and brain activity for all four conditions, but only in vPM. An ANCOVA (MatLab Statistics Toolbox) revealed that, when separate lines are fitted for each of the four conditions, the slopes of these lines do not differ significantly from each other (vPM: $F(1, 3) = 0.31; p = .82$; IPL: $F(1, 3) = 0.69; p = .56$; MT: $F(1, 3) = 0.05; p = .98$; cerebellum: $F(1, 3) = 0.65; p = .58$). When fitted in isolation for each condition separately (see Supplementary Material, Supplementary Figure S5a, b), the correlation between fMRI contrast and RT does not reach statistical significance. We therefore fitted a single line to all data, but separately for each ROI. Only

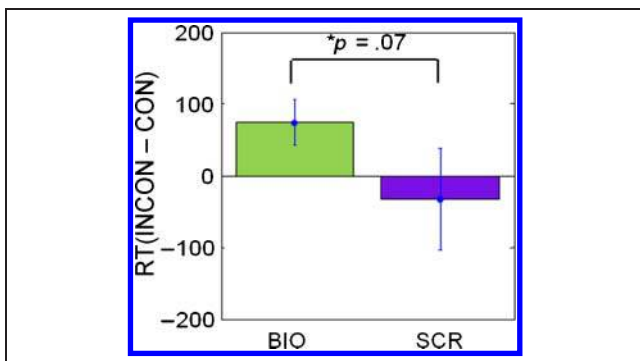


Figure 3. Behavioral data. RT differences (incongruent auditory–visual – congruent auditory–visual motion) are plotted for biological and scrambled motion signals. Incongruency of auditory and visual motion signals has an effect only when the audio-visual sequences depict biological motion; for scrambled motion no significant difference is observed between the incongruent and congruent condition. Error bars indicate standard errors of the mean.

Table 3. Correlations between RTs and Brain Activations

Location	Localizer	Pearson Correlation	
		Correlation Coefficient	<i>p</i>
<i>Frontal</i>			
BA 6 R/vPM	A ∩ VBIO	.29	.013
BA 6/44 R/vPM	A ∩ VSCR	.32	.006
<i>Parietal</i>			
BA 7 R/IPL	A ∩ VBIO	.17	.151
BA 7 R/IPL	A ∩ VSCR	.15	.196
<i>Temporal</i>			
BA 39 R/MT	A ∩ VBIO	-.14	.236
	A ∩ VSCR	-.16	.186
Cerebellum L	A ∩ VBIO	.16	.170
Cerebellum L	A ∩ VSCR	.11	.360

The correlation coefficients between contrast level (which is proportional to the data in **bold** font) in the four ROIs and the mean RTs are shown. Only the activation in the premotor area (BA 6) is significantly correlated with RTs ($r \sim 0.3; p < .05$; two-tailed test). Importantly, note that RT data were acquired outside the scanner before the experiment.

premotor activity is significantly correlated with RTs ($r \sim 0.3; p < .05$; Table 3).

DISCUSSION

Our aim was to identify the cortical network that differentiates between biologically plausible and implausible auditory–visual inputs. We first determined the cortical regions of auditory–visual coactivation by performing a conjunction analysis based on unimodal brain activations (Experiment 1: Localizer). The regions identified by this conjunction analysis were MT, IPL, and vPM. The brain activations arising from bimodal (auditory–visual) motion stimuli (Experiment 2) were then analyzed within these regions of coactivation. Our main finding is that the incongruency in the auditory and visual motion direction of the walker only affects the activity in the right vPM and only if the visual walker is intact. We therefore conclude that the right vPM not only plays a role in recognizing motion sequences in the visual and auditory modality in isolation but is also selective to the familiarity of the combined auditory–visual input.

Areas of Auditory and Visual Coactivation in the Right Hemisphere

Our conjunction analysis (Experiment 1) revealed four regions of auditory–visual coactivation: area MT (BA 39

bordering on BA 22 and BA 37), vPM (BA 6), and IPL (BA 7; at the border to SPL) in the right hemisphere and the cerebellum in the left hemisphere (see Table 1; also Supplementary Table S2 in supporting material). The strong right lateralization of brain activity in response to auditory footsteps is consistent with the findings that auditory motion in depth (looming/receding) is encoded in the right hemisphere (Seifritz et al., 2002; Baumgart, Gaschler-Markefski, Woldorff, Heinze, & Scheich, 1999), in particular, in the right premotor cortex (Schubotz & von Cramon, 2002). Brain activation for the (visual) point-light walker was also right-lateralized, in accordance with experiments by Pelphrey et al. (2005). Lateralization of auditory–visual coactivation in the right ventral intraparietal cortex and premotor cortex has also been found for random visual and auditory motion stimuli (Bremmer et al., 2001); the right IPL has been identified as a region of higher-level visual motion processing (Claeys, Lindsey, De Schutter, & Orban, 2003). In our experiments, the intact as well as scrambled point-light walkers were embedded in dynamic visual noise (to ensure comparable difficulty level to the auditory footsteps) which might also contribute to the lateralization in the right hemisphere as previously reported (Decety et al., 1997).

Auditory–Visual Coactivation in the Parieto-premotor Network

All three cortical ROIs identified as areas of auditory and visual coactivation (Experiment 1; Table 1; Figure 1) are known to be part of the controversial “mirror neuron system” (Dinstein, Gardner, Jazayeri, & Heeger, 2008; Dinstein, Thomas, Behrmann, & Heeger, 2008; Rizzolatti & Craighero, 2004). vPM (Iacobini et al., 1999; Decety et al., 1997; Rizzolatti, Fadiga, Gallese, & Fogassi, 1996) and IPL neurons (Buccino et al., 2001) are activated by the passive observation of actions. This parieto-premotor network (IPL, vPM) is thought to receive input from the MTG/pSTS; pSTS neurons are selective for biological motion, such as body, hand, and lip movements (Barracough, Xiao, Baker, Oram, & Perrett, 2005; Puce & Perrett, 2003) and are engaged in the perception of animacy (Schultz, Friston, O’Doherty, Wolpert, & Frith, 2005). The particular MT region identified by our conjunction analysis (BA 39/BA 22/BA 19) is close to areas engaged in the processing of body motions (Puce & Perrett, 2003) and is sometimes labeled as pSTS because of functional similarities with pSTS (Materna, Dicke, & Thier, 2008); in this study, we refer to it as MT region. Although all three areas, MT, IPL, and vPM, play a significant role in passive observation, imitation, and motion imagery (Hamzei et al., 2002), their connectivity is still a matter of debate (Bien, Roebroek, Goebel, & Sack, 2009). A simple common framework for action observation and imitation (Stanley & Miall, 2007) starts with a visual representation of action in the pSTS, an area that is active during observation but not execution (Barracough et al., 2005). Visual information is then passed on to the IPL,

which codes for the predicted outcome of the action and, subsequently, the intended action is translated into a motor program in vPM; an efferent copy of the planned action then returns to pSTS where it is compared with the original visual representation. In addition, direct bidirectional connections exist between the MT/pSTS and both the vPM and IPL (for a review, see Pineda, 2008). Our localizer experiment suggests that MT, IPL, and vPM are areas that receive both auditory and visual input. The fourth ROI defined by our localizer as an area of auditory–visual coactivation is the cerebellum. The cerebellum may play a role in converting the visual representation into motor codes, the “inverse model” (Stanley & Miall, 2007; Miall, 2003), by receiving information from the parietal lobe and forwarding it to the premotor cortex. The observed auditory–visual coactivation suggests that the involvement of the cerebellum in the inverse model may not be restricted to visual representations.

Increased Activity for Incongruent Auditory–Visual Biological Motion Signals in vPM

In our main experiment (Experiment 2), we compared the brain activation resulting from congruent (same motion direction in the auditory and visual modalities) with the activation resulting from incongruent motion (different motion direction in the auditory and visual modalities) within the areas of auditory–visual coactivation (derived in Experiment 1). Incongruent auditory–visual motion resulted in an increased brain activity only when both modalities signal biological motion; for scrambled visual motion, congruent and incongruent auditory–visual motion is associated with the same brain activations (Figure 2). A significant interaction is found only in one of the four ROIs, namely in the vPM (BA 6). The vPM not only plays a role in visual action observation and action imagery (Schubotz & von Cramon, 2001) but also responds to auditory actions (Kaplan & Iacoboni, 2007; Bidet-Caulet et al., 2005; Schubotz & von Cramon, 2002). A common vPM region is activated by visual motion imagery (Grafton, Arbib, Fadiga, & Rizzolatti, 1996), the observation of biologically meaningful actions (Bien et al., 2009), and the observation of meaningless (nonbiological) sequences (Schubotz & von Cramon, 2004), consistent with our findings that both biological and scrambled motion leads to vPM activation (Supplementary Figure S1 and Supplementary Table S2, first row). Schubotz and von Cramon (2002, 2004) concluded that the vPM is able to generate short-term action templates and that the vocabulary of motor acts stored in vPM is flexible and not innate. In our experiment, we find an increased premotor activity for incongruent biological motion in comparison with congruent biological motion (Figure 2; Supplementary Figure S2a,b); this increased premotor activity is associated with longer RTs (Figure 3; Table 3). Increased right premotor activity and associated increased RTs have also been reported for incongruent visuomotor conditions (Blakemore & Frith, 2005; Grezes,

Armony, Rowe, & Passingham, 2003) and for directionally incompatible or antiphase limb movements (Wenderoth, Debaere, Sanaert, Van Hecke, & Swinnen, 2004; de Jong, Leenders, & Paans, 2002). Increased right premotor activity (Jeannerod, 2001) is, therefore, likely to reflect conflicting or incompatible signals within or across sensory modalities as well as incompatible motor patterns. A very recent fMRI study using an entirely different set of biological motion stimuli (auditory and visual drumming actions) showed similar locations and patterns of activity changes as a function of expertise (Petrini et al., 2011): In the right IPL and the right premotor cortex, incongruent auditory–visual drumming actions leads to an increase in neural activity, but only in expert drummers as opposed to novices.

One possible explanation for the increased premotor activity for incongruent (i.e., a auditory–visual discrepancy in motion direction) biological motion is, in accordance with Schubotz and von Cramon (2004), the generation of novel motor templates based on the (inconsistent) sensory inputs across the auditory and visual modalities. Because, in this experimental condition, the auditory system signals a looming walker and the visual system signals a receding walker, no stored amodal action template provides a match to the bimodal sensory inputs hence necessitating the need for the generation of novel motor patterns. Congruent biological motion, on the other hand, yields auditory and visual motion signals that are likely to be matched to a single existing amodal template in the observer’s motor repertoire, yielding less premotor activity and shorter RTs (cf. Figures 2 and 3). This account is consistent with equal vPM activation for both congruent and incongruent scrambled motion (Supplementary Material S4a,b) because this hypothesis predicts that bimodal scrambled motion does not result in conflicting motion information in vPM. An alternative explanation is that the incongruent auditory–visual walker triggers two motor templates: one for a receding walker (based on the visual input) and one for a looming walker (based on the auditory input). Either explanation predicts increased activity (in the bimodal motion conditions) in vPM for incongruent biological motion only.

Activity in vPM is also increased in the unimodal (vision only) condition when the visual point-light walker is not intact [scrambled point-light walker (SCR) vs. intact point-light walker (BIO); Supplementary Table S1, top row; see also Thompson et al., 2005]. Although neurons in vPM are likely to respond to the components of the scrambled point-light walker such as legs, arms, and so forth, the overall configuration is unlikely to match an existing action template hence generating more activity in right vPM. Because new scrambled motion was generated on each trial, observers could not learn specific constellations (see Methods). The involvement of the vPM in human body processing has been shown using TMS: The body inversion effect is absent when TMS is applied in this area, hence suggesting that the vPM is involved in configural processing of human body shapes (Urgesi, Calvo-Merino, Haggard, & Aglioti, 2007). In line with our findings, increased right-lateralized vPM

activity has been reported during the observation of meaningless hand sequences (Decety & Grezes, 2006; Grezes, Costes, & Decety, 1999; Decety et al., 1997); parietal areas (BA 7) may have a role in selecting and monitoring motion sequences with on-line reference to a working memory in the right premotor cortex (Sadato, Campbell, Ibanez, Deiber, & Hallett, 1996). The increased activation of the right vPM in response to scrambled point-light walkers is consistent with the role of the right parieto-premotor network in the processing of novel and complex visual stimuli (Schubotz & von Cramon, 2002). Such an increase in stimulus complexity and novelty can be brought about by conflicting information within or across modalities. This is consistent with the idea that the right premotor network is not only involved in recognizing meaningful actions within a single modality but assimilates the information across the auditory and visual modalities by comparing it with a motor template, possibly residing in the premotor area (Schwarzbach, Sandrini, & Cattaneo, 2009; Sadato et al., 1996).

Specialized Neural Machinery for Biological Motion?

Numerous studies have shown an increased activity for visual biological motion in pSTS (for a review, see Puce & Perrett, 2003) and also identified pSTS as an area for the integration of auditory and visual biological motion signals. Our conjunction analysis (Figure 1) did not identify pSTS as an area of auditory–visual coactivation, but area MT (BA 39, bordering on BA 22 and BA 37), IPL (BA 7), and vPM (BA 6). Within these areas of auditory–visual coactivation, activity for the intact point-light walker was less (vPM, IPL) or equal (MT) to the activity in response to the scrambled walker (Supplementary Figure S3a,b and Supplementary Table S1). Equal activation in MT in response to intact and scrambled point-light walkers has been reported previously (Jastorff & Orban, 2009) and is at odds with the proposed role of MT for biological motion (e.g., Grossman, Battelli, & Pascual-Leone, 2005; Grossman et al., 2000). Furthermore, Jastorff and Orban (2009) proposed that the lack of differential activation for biological vs scrambled motion in pSTS could be associated with task complexity. This is consistent with the findings by Meyer et al. (2011), who documented a role of the pSTS in the processing of biological motion stimuli closely matched to the ones used in this experiment but crucially employing a one-back task.

Another significant methodological difference between our study and previous studies using PLW was that we used looming and receding PLWs (instead of a PLW walking on a “treadmill”), hence signaling motion in depth, which is not a stimulus feature STS is very sensitive to (Perrett, Harries, Benson, Chitty, & Mistlin, 1990). The task of our observers was to judge whether there was any motion in depth present as opposed to categorizing or identifying the biological motion (Meyer, Crocker-Buque, & Wuerger, 2007); our task therefore also favors the involvement of the vPM (Ochiai, Mushiake, & Tanji,

2005; Schubotz & von Cramon, 2002; Kakei, Hoffman, & Strick, 2001). Finally, to equate the auditory and visual PLWs in difficulty, we added dynamic noise to the visual PLWs which might also bias the activation toward area MT and the right parieto-premotor network (Pelphrey et al., 2005; Bremmer et al., 2001).

The increased activity in the right vPM for scrambled compared with intact point-light walkers is in line with more recent imaging studies showing increased right-lateralized activity for incoherent vs coherent action sequences in the right vPM (Bien et al., 2009). A right-lateralized decrease in neural activity when novel stimuli become more familiar via training or prolonged observation (Vogt et al., 2007; Downar, Crawley, Mikulis, & Davis, 2002) is consistent with the idea that learned meaningless movements generate less cortical activity than unlearned meaningless sequences because the neural population that represents the familiar stimuli have become more selective during learning. Biological motion stimuli are special configurations of highly familiar local limb movements; whereas numerous neurons are likely to respond to individual limb movements (such as contained in a scrambled PLW), a small population of neurons is likely to respond to the particular configuration of limb movements depicted in an intact PLW.

Our current findings are consistent with the idea that the right vPM is involved in the processing of body movements by comparing sensorimotor representations of familiar body movements with incoming sensory input. It extends our current knowledge by suggesting that vPM is also involved in the integration of sensory inputs across the auditory and visual modalities and compares information across modalities with an amodal template, possibly residing in the premotor area (Schwarzbach et al., 2009; Sadato et al., 1996).

Previous studies identified both vPM areas, BA 6, a homolog to monkey F4, and BA 44, which is assumed to be a homolog to monkey F5, as areas activated by hand or arm movements (for a review, see Rizzolatti, Fogassi, & Gallese, 2002). In particular, there is evidence that the vPM contains also motor-related presentations of space, in relation to one's own body. Makin, Holmes, and Zohary (2007) showed that vPM plays a role in representing perihand space; this study is also consistent with the premotor cortex as a site of sensory convergence, because strong PMv activation required concurrent visual and tactile stimulation. Our own data show that vPM (border of BA 6 and BA 44) is activated by a walker which is approaching or receding in relation to the participant, whether the motion is defined by auditory or visual stimulation is irrelevant (see Supplementary Table S1 in the supporting material). Hence, an alternative interpretation of our data is that vPM is encoding information about the closeness of objects/individuals in relation to one's body, instead of containing general motor templates as outlined above. In either case, vPM is a site that contains both visual and auditory representations of moving stimuli and is involved in the consolidation of these representations.

Acknowledgments

S. M. W.'s stay at the University of Regensburg (in Professor Greenlee's laboratory) was supported by a Wellcome Trust Sabbatical Grant (GR/082831). The ViSaGe system was cosponsored by Cambridge Research Systems Ltd., Kent, United Kingdom, and the Wellcome Trust (GR/080205). Scanning costs were covered by the Faculty of Medicine at the University of Liverpool. We thank Ingo Keck for helpful comments on the manuscript.

Reprint requests should be sent to Sophie M. Wuerger, Department of Experimental Psychology, University of Liverpool, Eleanor Rathbone Building, Bedford Street South, Liverpool, L69 7ZA, United Kingdom, or via e-mail: s.m.wuerger@liverpool.ac.uk, web: www.liv.ac.uk/Psychology/staff/swuerger.html.

REFERENCES

- Arrighi, R., Alais, D., & Burr, D. (2006). Perceptual synchrony of audiovisual streams for natural and artificial motion sequences. *Journal of Vision*, *6*, 260–268.
- Arrighi, R., Marini, F., & Burr, D. (2009). Meaningful auditory information enhances perception of visual biological motion. *Journal of Vision*, *9*, 1–7.
- Barracough, N. E., Xiao, D., Baker, C. I., Oram, M. W., & Perrett, D. I. (2005). Integration of visual and auditory information by superior temporal sulcus neurons responsive to the sight of actions. *Journal of Cognitive Neuroscience*, *17*, 377–391.
- Baumann, O., & Greenlee, M. W. (2007). Neural correlates of coherent audiovisual motion perception. *Cerebral Cortex*, *17*, 1433–1443.
- Baumgart, F., Gaschler-Markefski, B., Woldorff, M. G., Heinze, H.-J., & Scheich, H. (1999). A movement-sensitive area in auditory cortex. *Nature*, *400*, 724–726.
- Bidet-Caulet, A., Voisin, J., Bertrand, O., & Fonlupt, P. (2005). Listening to a walking human activates the temporal biological motion area. *Neuroimage*, *28*, 132.
- Bien, N., Roebroeck, A., Goebel, R., & Sack, A. T. (2009). The brain's intention to imitate: The neurobiology of intentional versus automatic imitation. *Cerebral Cortex*, *19*, 2338–2351.
- Blake, R., & Shiffrar, M. (2007). Perception of human motion. *Annual Review of Psychology*, *58*, 47–73.
- Blakemore, S. J., & Frith, C. (2005). The role of motor contagion in the prediction of action. *Neuropsychologia*, *43*, 260–267.
- Bonda, E., Petrides, M., Ostry, D., & Evans, A. (1996). Specific involvement of human parietal systems and the amygdala in the perception of biological motion. *Journal of Neuroscience*, *16*, 3737.
- Bonini, L., Rozzi, S., Serventi, F. U., Simone, L., Ferrari, P. F., & Fogassi, L. (2010). Ventral premotor and inferior parietal cortices make distinct contribution to action organization and intention understanding. *Cerebral Cortex*, *20*(10), 1372.
- Bremmer, F., Schlack, A., Shah, N. J., Zafiris, O., Kubischik, M., Hoffmann, K., et al. (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.
- Brett, M., Anton, J.-L., Valabregue, R., & Poline, J. B. (2002). *Region of interest analysis using an SPM toolbox*. Paper presented at the 8th International Conference on Functional Mapping of the Human Brain, Sendai, Japan.
- Brooks, A., van der Zwan, R., Billard, A., Petreska, B., Clarke, S., & Blanke, O. (2007). Auditory motion affects visual biological motion processing. *Neuropsychologia*, *45*, 523–530.
- Buccino, G., Binkofski, F., Fink, G. R., Fadiga, L., Fogassi, L., Gallese, V., et al. (2001). Action observation activates

- premotor and parietal areas in a somatotopic manner: An fMRI study. *European Journal of Neuroscience*, *13*, 400–404.
- Buch, E. R., Mars, R. B., Boorman, E. D., & Rushworth, M. F. S. (2010). A network centered on ventral premotor cortex exerts both facilitatory and inhibitory control over primary motor cortex during action reprogramming. *The Journal of Neuroscience*, *30*, 1395.
- Calvo-Merino, B., Glaser, D. E., Grezes, J., Passingham, R. E., & Haggard, P. (2005). Action observation and acquired motor skills: An fMRI study with expert dancers. *Cerebral Cortex*, *15*, 1243–1249.
- Claycs, K. G., Lindsey, D. T., De Schutter, E., & Orban, G. A. (2003). A higher order motion region in human inferior parietal lobule: Evidence from fMRI. *Neuron*, *40*, 631–642.
- de Jong, B. M., Leenders, K. L., & Paans, A. M. J. (2002). Right parieto-premotor activation related to limb-independent antiphase movement. *Cerebral Cortex*, *12*, 1213–1217.
- Decety, J., & Grezes, J. (2006). The power of simulation: Imagining one's own and other's behavior. *Brain Research*, *1079*, 4–14.
- Decety, J., Grezes, J., Costes, N., Perani, D., Jeannerod, M., Procyk, E., et al. (1997). Brain activity during observation of actions. Influence of action content and subject's strategy. *Brain*, *120*, 1763–1777.
- Dinstein, I., Gardner, J. L., Jazayeri, M., & Heeger, D. J. (2008). Executed and observed movements have different distributed representations in human aIPS. *Journal of Neuroscience*, *28*, 11231–11239.
- Dinstein, I., Thomas, C., Behrmann, M., & Heeger, D. J. (2008). A mirror up to nature. *Current Biology*, *18*, R13–R18.
- Downar, J., Crawley, A. P., Mikulis, D. J., & Davis, K. D. (2002). A cortical network sensitive to stimulus salience in a neutral behavioral context across multiple sensory modalities. *Journal of Neurophysiology*, *87*, 615–620.
- Eickhoff, S. B., Stephan, K. E., Mohlberg, H., Grefkes, C., Fink, G. R., Amunts, K., et al. (2005). A new SPM toolbox for combining probabilistic cytoarchitectonic maps and functional imaging data. *Neuroimage*, *25*, 1325–1335.
- Friston, K. J., Penny, W. D., & Glaser, D. E. (2005). Conjunction revisited. *Neuroimage*, *25*, 661–667.
- Grafton, S. T., Arbib, M. A., Fadiga, L., & Rizzolatti, G. (1996). Localization of grasp representations in humans by positron emission tomography. *Experimental Brain Research*, *112*, 103–111.
- Grezes, J. (2001). Does perception of biological motion rely on specific brain regions? *Neuroimage*, *13*, 775–785.
- Grezes, J., Armony, J. L., Rowe, J., & Passingham, R. E. (2003). Activations related to “mirror” and “canonical” neurones in the human brain: An fMRI study. *Neuroimage*, *18*, 928–937.
- Grezes, J., Costes, N., & Decety, J. (1999). The effects of learning and intention on the neural network involved in the perception of meaningless actions. *Brain*, *122*, 1875–1887.
- Grossman, E. D., Battelli, L., & Pascual-Leone, A. (2005). Repetitive TMS over STSp disrupts perception of biological motion. *Vision Research*, *45*, 2847.
- Grossman, E. D., & Blake, R. (2001). Brain activity evoked by inverted and imagined biological motion. *Vision Research*, *41*, 1475.
- Grossman, E. D., & Blake, R. (2002). Brain areas active during visual perception of biological motion. *Neuron*, *35*, 1167.
- Grossman, E. D., Donnelly, M., Price, R., Pickens, D., Morgna, V., Neighbour, G., et al. (2000). Brain areas involved in the perception of biological motion. *Journal of Cognitive Neuroscience*, *12*, 711–720.
- Hamzei, F., Dettmers, C., Rijntjes, M., Glauche, V., Kiebel, S., Weber, B., et al. (2002). Visuomotor control within a distributed parieto-frontal network. *Experimental Brain Research*, *146*, 273–281.
- Harrison, N. R., Wuerger, S. M., & Meyer, G. F. (2010). Reaction time facilitation for horizontally moving auditory–visual stimuli. *Journal of Vision*, *10*, 1–21.
- Henson, R. N. A., Price, C., Rugg, M. D., Turner, R., & Friston, K. (2002). Detecting latency differences in event-related BOLD responses: Application to words versus nonwords, and initial versus repeated face presentations. *Neuroimage*, *15*, 83–97.
- Howard, R. J., Brammer, M., Wright, I., Woodruff, P. W., Bullmore, E. T., & Zeki, S. (1996). A direct demonstration of functional specialization within motion-related visual and auditory cortex of the human brain. *Current Biology*, *6*, 1015.
- Iacoboni, M., Woods, R. P., Brass, M., Bekkering, H., Mazziotta, J. C., & Rizzolatti, G. (1999). Cortical mechanisms of human imitation. *Science*, *286*, 2526–2528.
- Jastorff, J., Begliomini, C., Fabbri-Destro, M., Rizzolatti, G., & Orban, G. A. (2010). Coding observed motor acts: Different organizational principles in the parietal and premotor cortex of humans. *Journal of Neurophysiology*, *104*, 128.
- Jastorff, J., & Orban, G. A. (2009). Human functional magnetic resonance imaging reveals separation and integration of shape and motion cues in biological motion processing. *The Journal of Neuroscience*, *29*, 7315–7329.
- Jeannerod, M. (2001). Neural simulation of action: A unifying mechanism for motor cognition. *Neuroimage*, *14*, S103–S109.
- Johansson, G. (1973). Visual perception of biological motion and a model for its analysis. *Perception and Psychophysics*, *14*, 201–211.
- Takei, S., Hoffman, D. S., & Strick, P. L. (2001). Direction of action is represented in the ventral premotor cortex. *Nature Neuroscience*, *4*, 1020–1025.
- Kaplan, J., & Iacoboni, M. (2007). Multimodal action representation in human left ventral premotor cortex. *Cognitive Processing*, *8*, 103–113.
- Makin, T. R., Holmes, N. P., & Zohary, E. (2007). Is that near my hand? Multisensory representation of peripersonal space in human intraparietal sulcus. *The Journal of Neuroscience*, *27*, 731–740.
- Maldjian, J. A., Laurienti, P. J., Kraft, R. A., & Burdette, J. H. (2003). An automated method for neuroanatomic and cytoarchitectonic atlas-based interrogation of fMRI data sets. *Neuroimage*, *19*, 1233–1239.
- Materna, S., Dicke, P. W., & Thier, P. (2008). Dissociable roles of the superior temporal sulcus and the intraparietal sulcus in joint attention: A functional magnetic resonance imaging study. *Journal of Cognitive Neuroscience*, *20*, 108–119.
- Meyer, G., Crocker-Buque, A., & Wuerger, S. (2007). Auditory-visual integration of biological motion. *Perception Supplement*, *36*, 171.
- Meyer, G., & Wuerger, S. (2001). Cross-modal integration of auditory and visual motion signals. *NeuroReport*, *12*, 2557–2600.
- Meyer, G., Wuerger, S., Roehrbein, F., & Zetzsche, C. (2005). Low-level integration of auditory and visual motion signals requires spatial co-localisation. *Experimental Brain Research*, *166*, 538–547.
- Meyer, G. F., Greenlee, M., & Wuerger, S. (2011). Interactions between auditory and visual semantic stimulus classes: Evidence for common processing networks for speech and body actions. *Journal of Cognitive Neuroscience*, *23*, 2271–2288.
- Miall, R. C. (2003). Connecting mirror neurons and forward models. *NeuroReport*, *14*, 2135–2137.
- Ochiai, T., Mushiake, H., & Tanji, J. (2005). Involvement of the ventral premotor cortex in controlling image motion of the hand during performance of a target-capturing task. *Cerebral Cortex*, *15*, 929–937.

- Pelphrey, K. A., Mitchell, T. V., McKeown, M. J., Goldstein, J., Allison, T., & McCarthy, G. (2003). Brain activity evoked by the perception of human walking: Controlling for meaningful coherent motion. *Journal of Neuroscience*, *23*, 6819–6825.
- Pelphrey, K. A., Morris, J. P., Michelich, C. R., Allison, T., & McCarthy, G. (2005). Functional anatomy of biological motion perception in posterior temporal cortex: An fMRI study of eye, mouth and hand movements. *Cerebral Cortex*, *15*, 1866.
- Perrett, D. I., Harries, M. H., Benson, P. J., Chitty, A. J., & Mistlin, A. J. (1990). Retrieval of structure from rigid and biological motion: An analysis of the visual responses of neurones in the macaque temporal cortex. In A. Blake & T. Troscianko (Eds.), *AI and the eye* (pp. 181–200). London: John Wiley & Sons Ltd.
- Petrini, K., Pollick, F. E., Dahl, S., McAleer, P., McKay, L., Rocchesso, D., et al. (2011). Action expertise reduces brain activity for audiovisual matching actions: An fMRI study with expert drummers. *Neuroimage*, *56*, 1480–1492.
- Pilgramm, S., Lorey, B., Stark, R., Munzert, J., Vaitl, D., & Zentgraf, K. (2010). Differential activation of the lateral premotor cortex during action observation. *BMC Neuroscience*, *11*, 89.
- Pineda, J. (2008). Sensorimotor cortex as a critical component of an “extended” mirror neuron system: Does it solve the development, correspondence, and control problems in mirroring? *Behavioral and Brain Functions*, *4*, 47.
- Puce, A., & Perrett, D. (2003). Electrophysiology and brain imaging of biological motion. *Philosophical Transactions of the Royal Society of London, Series B, Biological Sciences*, *358*, 435–445.
- Rizzolatti, G., & Craighero, L. (2004). The mirror-neuron system. *Annual Review of Neuroscience*, *27*, 169–192.
- Rizzolatti, G., Fadiga, L., Gallese, V., & Fogassi, L. (1996). Premotor cortex and the recognition of motor actions. *Cognitive Brain Research*, *3*, 131–141.
- Rizzolatti, G., Fogassi, L., & Gallese, V. (2002). Motor and cognitive functions of the ventral premotor cortex. *Current Opinion in Neurobiology*, *12*, 149–154.
- Sadato, N., Campbell, G., Ibanez, V., Deiber, M., & Hallett, M. (1996). Complexity affects regional cerebral blood flow change during sequential finger movements. *Journal of Neuroscience*, *16*, 2691–2700.
- Saygin, A. P. (2007). Superior temporal and premotor brain areas necessary for biological motion perception. *Brain*, *130*, 2452–2461.
- Saygin, A. P., Driver, J., & de Sa, V. R. (2008). In the footsteps of biological motion and multisensory perception: Judgments of audiovisual temporal relations are enhanced for upright walkers. *Psychological Science*, *19*, 469–475.
- Saygin, A. P., Wilson, S. M., Hagler, D. J., Jr., Bates, E., & Sereno, M. I. (2004). Point-light biological motion perception activates human premotor cortex. *Journal of Neuroscience*, *24*, 6181–6188.
- Schouten, B., Troje, N. F., Vroomen, J., & Verfaillie, K. (2011). The effect of looming and receding sounds on the perceived in-depth orientation of depth-ambiguous biological motion figures. *PLoS One*, *6*, e14725.
- Schubotz, R. I., & von Cramon, D. Y. (2001). Functional organization of the lateral premotor cortex: fMRI reveals different regions activated by anticipation of object properties, location and speed. *Cognitive Brain Research*, *11*, 97–112.
- Schubotz, R. I., & von Cramon, D. Y. (2002). Predicting perceptual events activates corresponding motor schemes in lateral premotor cortex: An fMRI study. *Neuroimage*, *15*, 787–796.
- Schubotz, R. I., & von Cramon, D. Y. (2004). Sequences of abstract nonbiological stimuli share ventral premotor cortex with action observation and imagery. *Journal of Neuroscience*, *24*, 5467–5474.
- Schultz, J., Friston, K. J., O’Doherty, J., Wolpert, D. M., & Frith, C. D. (2005). Activation in posterior superior temporal sulcus parallels parameter inducing the percept of animacy. *Neuron*, *45*, 625–635.
- Schwarzbach, J. V., Sandrini, M., & Cattaneo, L. (2009). *Neural populations in the parietal and premotor cortices of humans perform abstract coding of motor acts: A TMS-adaptation study*. Paper presented at the ECVF Abstract Supplement Regensburg.
- Seifritz, E., Neuhoff, J. G., Bilecen, D., Scheffler, K., Mustovic, H., Schächinger, H., et al. (2002). Neural processing of auditory looming in the human brain. *Current Biology*, *12*, 2147.
- Servos, P., Osu, R., Santi, A., & Kawato, M. (2002). The neural substrates of biological motion perception: An fMRI study. *Cerebral Cortex*, *12*, 772.
- Stanley, J., & Miall, R. C. (2007). Functional activation in parieto-premotor and visual areas dependent on congruency between hand movement and visual stimuli during motor-visual priming. *Neuroimage*, *34*, 290–299.
- Szycik, G. R., Tausche, P., & Münte, T. F. (2008). A novel approach to study audiovisual integration in speech perception: Localizer fMRI and sparse sampling. *Brain Research*, *1220*, 142–149.
- Thompson, J. C., Clarke, M., Stewart, T., & Puce, A. (2005). Configural processing of biological motion in human superior temporal sulcus. *Journal of Neuroscience*, *25*, 9059–9066.
- Urgesi, C., Calvo-Merino, B., Haggard, P., & Aglioti, S. M. (2007). Transcranial magnetic stimulation reveals two cortical pathways for visual body processing. *Journal of Neuroscience*, *27*, 8023–8030.
- Vaina, L. M., Solomon, J., Chowdhury, S., Sinha, P., & Belliveau, J. W. (2001). Functional neuroanatomy of biological motion perception in humans. *Proceedings of the National Academy of Sciences, U.S.A.*, *98*, 11656–11661.
- Van Essen, D. C. (2005). A population-average, landmark- and surface-based (PALS) atlas of human cerebral cortex. *Neuroimage*, *28*, 635–662.
- Van Essen, D. C., Drury, H. A., Dickson, J., Harwell, J., Hanlon, D., & Anderson, C. H. (2001). An integrated software suite for surface-based analyses of cerebral cortex. *Journal of the American Medical Informatics Association*, *8*, 443–459.
- Vanrie, J., & Verfaillie, K. (2006). Perceiving depth in point-light actions. *Attention, Perception, & Psychophysics*, *68*, 601–612.
- Vogt, S., Buccino, G., Wohlschläger, A. M., Canessa, N., Shah, N. J., Zilles, K., et al. (2007). Prefrontal involvement in imitation learning of hand actions: Effects of practice and expertise. *Neuroimage*, *37*, 1371–1383.
- Wenderoth, N., Debaere, F., Sunaert, S., Van Hecke, P., & Swinnen, S. P. (2004). Parieto-premotor areas mediate directional interference during bimanual movements. *Cerebral Cortex*, *14*, 1153–1163.
- Wuerger, S. M., Crocker-Buque, A., & Meyer, G. F. (in press). Evidence for auditory–visual processing specific to biological motion. *Seeing and Perceiving*.
- Wuerger, S. M., Hofbauer, M., & Meyer, G. F. (2003). The integration of auditory and visual motion signals at threshold. *Perception & Psychophysics*, *65*, 1188–1196.

This article has been cited by:

1. Tristan A. Chaplin, Marcello G. P. Rosa, Leo L. Lui. 2018. Auditory and Visual Motion Processing and Integration in the Primate Cerebral Cortex. *Frontiers in Neural Circuits* **12**. . [[Crossref](#)]
2. Wagner Souza Silva, Gayatri Aravind, Samir Sangani, Anouk Lamontagne. 2018. Healthy young adults implement distinctive avoidance strategies while walking and circumventing virtual human vs. non-human obstacles in a virtual environment. *Gait & Posture* **61**, 294-300. [[Crossref](#)]
3. Michael Andric, Ben Davis, Uri Hasson. 2017. Visual cortex signals a mismatch between regularity of auditory and visual streams. *NeuroImage* **157**, 648-659. [[Crossref](#)]
4. Jeroen J. Stekelenburg, Mirjam Keetels. 2016. The effect of synesthetic associations between the visual and auditory modalities on the Colavita effect. *Experimental Brain Research* **234**:5, 1209-1219. [[Crossref](#)]
5. Andy Bevilacqua, Fred Paas, Genomary Krigbaum. 2016. Effects of Motion in the Far Peripheral Visual Field on Cognitive Test Performance and Cognitive Load. *Perceptual and Motor Skills* **122**:2, 452-469. [[Crossref](#)]
6. Mark Laing, Adrian Rees, Quoc C. Vuong. 2015. Amplitude-modulated stimuli reveal auditory-visual interactions in brain activity and brain connectivity. *Frontiers in Psychology* **6**. . [[Crossref](#)]
7. Matthias Bischoff, Karen Zentgraf, Sebastian Pilgramm, Rudolf Stark, Britta Krüger, Jörn Munzert. 2014. Anticipating action effects recruits audiovisual movement representations in the ventral premotor cortex. *Brain and Cognition* **92**, 39-47. [[Crossref](#)]
8. Yi-Huang Su. 2014. Peak velocity as a cue in audiovisual synchrony perception of rhythmic stimuli. *Cognition* **131**:3, 330-344. [[Crossref](#)]
9. Georg F Meyer, Neil R Harrison, Sophie M Wuerger. 2013. The time course of auditory-visual processing of speech and body actions: Evidence for the simultaneous activation of an extended neural network for semantic processing. *Neuropsychologia* **51**:9, 1716-1725. [[Crossref](#)]
10. Timothy D. Sweeny, Steve Haroz, David Whitney. 2012. Reference repulsion in the categorical perception of biological motion. *Vision Research* **64**, 26-34. [[Crossref](#)]