Cogn Process (2007) 8:103–113
DOI 10.1007/s10339-007-0165-z

RESEARCH REPORT

Multimodal action representation in human left ventral premotor cortex

Jonas T. Kaplan · Marco Iacoboni

Received: 21 December 2006 / Revised: 2 March 2007 / Accepted: 6 March 2007 / Published online: 15 May 2007

Abstract We used functional magnetic resonance imaging (fMRI) to investigate the neural systems responding to the sight and to the sound of an action. Subjects saw a video of paper tearing in silence (V), heard the sound of paper tearing (A), and saw and heard the action simultaneously (A + V). Compared to a non-action control stimulus, we found that hearing action sounds (A) activated the anterior inferior frontal gyrus and middle frontal gyrus in addition to primary auditory cortex. The anterior inferior frontal gyrus, which is known to be activated by environmental sounds, also seems to be involved in recognizing actions by sound. Consistent with previous research, seeing an action video (V) compared with seeing a non-action video activated the premotor cortex, intraparietal cortex, and the pars opercularis of the inferior frontal gyrus. An A + V facilitation effect was found in the ventral premotor cortex on the border of areas 44, 6, 3a, and 3b for the action stimuli but not for the control stimuli. This region may be involved in integrating multimodal information about actions. These data provide evidence that the ventral premotor cortex may provide an action representation that abstracts across both agency (self and other) and sensory modality (hearing and seeing). This function may be an important precursor of language functions.

Introduction

There is growing evidence that the primate motor system plays an important role in action recognition. Mirror neurons, first discovered in area F5 of the monkey premotor cortex, are active during both execution and observation of a goal-directed action (Gallese et al. 1996; Rizzolatti et al. 1996). Imaging studies in humans have found activations throughout premotor cortex, Broca's area, and the intraparietal region during action observation (Buccino et al. 2001; Grezes et al. 2003; Grezes and Decety 2001; Iacoboni et al. 1999; Kohler et al. 2002; Koski et al. 2002; Molnar-Szakacs et al. 2005; see Rizzolatti and Craighero 2004 for review; Schubotz and von Cramon 2003). Since activity in this mirror neuron system seems to be agent-independent, it may provide an abstract representation that allows matching between self and other.

The mirror neuron system may be essential for predicting the goals and future actions of others. When we observe someone acting, the resulting activation of our own motor programs may also recruit circuitry for the motivations and goals that normally produce those actions. Blakemore and Frith (2005) for example, have argued that motor simulation facilitates our ability to predict the actions of others—when you are in someone's shoes, you know what they are going to do next. Once we frame an action in motor terms, we have at our disposal the sophisticated prediction mechanisms embedded in our motor system (Wolpert et al. 2003). In fact, we have shown that some mirror neurons in the inferior frontal gyrus are sensitive to contextual cues that suggest a specific intention (Iacoboni et al. 2005; Kaplan and Iacoboni 2006).

We are able to recognize actions not only by seeing them, but also by hearing them. For example, we can easily recognize the sound of footsteps coming toward us or a

J. T. Kaplan · M. Iacoboni
Ahmanson-Lovelace Brain Mapping Center,
Department of Psychiatry and Biobehavioral Sciences,
Neuropsychiatric Institute, David Geffen School of Medicine,
University of California, Los Angeles, CA, USA

J. T. Kaplan (✉)
UCLA Ahmanson-Lovelace Brain Mapping Center,
660 Charles Young Dr. South,
Los Angeles, CA 90095, USA
e-mail: jonask@ucla.edu

hand knocking on a door without seeing these actions. Recently, a subset of mirror neurons in monkey premotor cortex was shown to respond to auditory as well as visual stimuli (Keysers et al. 2003; Kohler et al. 2002). These ‘‘audiovisual mirror neurons’’ are able to recognize an action either by hearing it or by seeing it. For example, Kohler et al. (2002) found neurons that were responsive to the sound of a peanut being crushed or to paper tearing, but were not responsive to non-action sounds. Most of these neurons increased in activity when the monkey saw an action (V), when it heard an action (A) or when it heard and saw an action (A + V), but a smaller set of neurons was selective either for V or for A. Keysers et al. (2003) further explored multimodal interactions in these neurons, and described three types of responses. Of the 22 neurons tested, 11 of them showed the same amount of response to the multimodal condition (A + V) as they did to the unimodal conditions. A set of eight neurons showed an additive property, such that the A + V condition was significantly stronger in activating the neuron compared with A or V alone. These neurons seem to be doing some kind of multimodal integration. The remaining three cells responded most strongly to the sound alone. Most of the audiovisual neurons, then, do not care about the modality of the stimulus, but some of them are able to integrate audio and visual information. The action representation in these neurons is special because it is abstracted across two dimensions: agency and sensory modality.

There is now growing evidence that hearing an action can activate motor cortex in humans as well as monkeys. It has been known for some time that human premotor cortex can respond to auditory stimuli. For example, Schubotz et al. (2003) found responses in premotor cortex to auditory pattern recognition, and Wilson et al. (2004) found premotor activations while listening to speech sounds. Direct evidence for a role of the human motor cortex in action recognition first came from transcranial magnetic stimulation (TMS) studies on motor facilitation. Aziz-Zadeh et al. (2004) showed increased motor excitability in primary motor cortex when listening to action sounds. In a repetition priming paradigm, Pizzamiglio et al. (2005) found that a positive-going ERP component was modulated by action sounds but not by non-action sounds. Source localization revealed two likely sources, one in the superior temporal sulcus (STS) and one in the left premotor cortex. Premotor cortex response to sound appears to depend upon the degree to which an action is associated with that sound. Lahav et al. (2007) trained subjects to produce musical sequences by pressing keys, then acquired fMRI data while subjects listened to musical segments they could produce in addition to novel sequences. Listening to producible sequences led to greater activations throughout the mirror neuron system, including the inferior frontal gyrus, premotor areas, and the inferior parietal lobe. Gazzola et al. (2006) also used fMRI to measure brain activity while subjects heard action sounds, and while they produced similar actions. They found a network of brain regions were active in both conditions, including the inferior frontal gryus, Brodmann’s Area 6, and the STS. These studies confirm that motor networks are recruited during the processing of auditory stimuli that are action-related.

One of the key features of action sounds is that they tend to occur simultaneously with the action that produces them. The association between the visual and auditory information may be key in the development of auditory-motor associations. The present experiment uses functional magnetic resonance imaging (fMRI) to investigate the visual and auditory responses of the human motor circuitry with particular attention to how the two modalities interact. Subjects in this experiment saw a visual (V), auditory (A), or multimodal (A + V) stimulus of a paper tearing action or a non-action control. We chose to use a single action, paper tearing, for several reasons. First, it is one of the actions that was used in the monkey experiments on auditory-visual mirror neurons so it is known to evoke activity in these neurons in monkeys. Second, it has an easily recognizable sound which has been shown to increase TMS-induced motor evoked potentials in humans (Aziz-Zadeh et al. 2004). Finally, it is a bimanual action, so any lateralized brain activations can’t be attributed to the hand performing the action. This is especially important in light of the considerations that link multimodality with language (Gallese and Lakoff 2005).

Materials and methods

Subjects

Ten right-handed subjects [five females, five males, mean age: 25 + (–5.2)] were recruited and compensated for their participation. Subjects gave informed consent according to the guidelines of the UCLA Institutional Review Board. All participants were screened to rule out medication use, head trauma, history of neurological or psychiatric disorders, substance abuse, or other serious medical conditions.

Image acquisition

Images were acquired using a Siemens Allegra 3.0 T MRI scanner. Two sets of high-resolution anatomical images were acquired for registration purposes. We acquired an MP-RAGE (TR = 2,300, TE = 2.93, flip angle = 8°) with 160 sagittal slices, each 1 mm thick with 0.5 mm gap and

1.33 mm × 1.33 mm in-plane resolution. We also acquired a T2-weighted co-planar image (TR = 5,000, TE = 33, flip angle = 90°) with 36 transverse slices covering the whole brain, each 3 mm thick with 1 mm gap, a 128 × 128 matrix and an in-plane resolution of 1.5 mm × 1.5 mm.

Each functional run involved the acquisition of 186 BOLD-weighted echo-planar volumes (TR = 2,000, TE = 25, flip angle = 90°), each with 36 transverse slices, 3 mm thick, 1 mm gap, and a 64 × 64 matrix yielding an in-plane resolution of 3 mm × 3 mm. A functional run lasted 6 min and 12 s, and each subject completed three functional runs.

Stimuli and task

Two videos were created for the experiment. The first, the *Action Clip*, was a color view of two female hands tearing a piece of paper into smaller and smaller pieces (see Fig. 1). The clip lasted 8 s, and was accompanied by the actual audio recorded with the video at 44 kHz. The second video, the *Control Clip*, consisted of an animation of a brown box on a gray background. The box moved up and down on the screen, and the movements of the box were matched with the video of the paper tearing so that each made the same number of movements with the same temporal pattern. The animation was accompanied by a soundtrack which consisted of a 1,000 kHz tone temporally paired with each movement the box made. The two audio tracks were equated for loudness matching their average RMS power, but due to the nature of the two sounds differed in their spectral frequency profile.

Each clip was presented with the video component alone (V), the audio component alone (A), or the video and audio together (A + V). When the audio component was played alone a fixation cross remained on the screen. Each functional run began with 12 s of rest, in which only a fixation cross appeared on the screen, followed by the stimuli in random order separated by 12 s rest periods. Each of the six stimuli was presented three times in each block in randomized order.

Data processing and statistical analysis

Analysis was carried out using FEAT (FMRI Expert Analysis Tool) Version 5.1, part of FSL (FMRIB's Software Library, http://www.fmrib.ox.ac.uk/fsl). After motion correction, images were smoothed using a 5 mm Gaussian FWHM algorithm in three dimensions, and temporally high-pass filtered with a cutoff period of 75 s. The BOLD response was modeled using a separate explanatory variable (EV) for each of the six stimulus types. For each stimulus type, the presentation design was convolved with a gamma function to produce an expected BOLD response. The temporal derivative of this timecourse was also included in the model for each EV. Data were then fitted to the model using FSL's implementation of the general linear model.

Each subject's statistical data was then warped into a standard space based on the MNI-152 atlas. We used FLIRT to register the functional data to the atlas space in three stages. First, functional images were aligned with the high-resolution co-planar T2-weighted image using a 6 degree of freedom rigid-body warping procedure. Next, the co-planar volume was registered to the T1-weighted MP-RAGE using a 6 degree of freedom rigid-body warp. Finally, the MP-RAGE was registered to the standard MNI atlas with a 12 degree of freedom affine transformation.

Higher-level analysis was carried out using FLAME (FMRIB's Local Analysis of Mixed Effects) (Behrens et al. 2003). Z (Gaussianised T/F) statistic images were thresholded using clusters determined by $Z > 2.3$ and a (corrected) cluster significance threshold of $p = 0.01$ (Worsley et al. 1992; Friston et al. 1994; Forman et al. 1995).

Results

Activity associated with hearing an action sound

We performed a contrast of brain regions more activated by the action audio clip compared with the control audio clip.

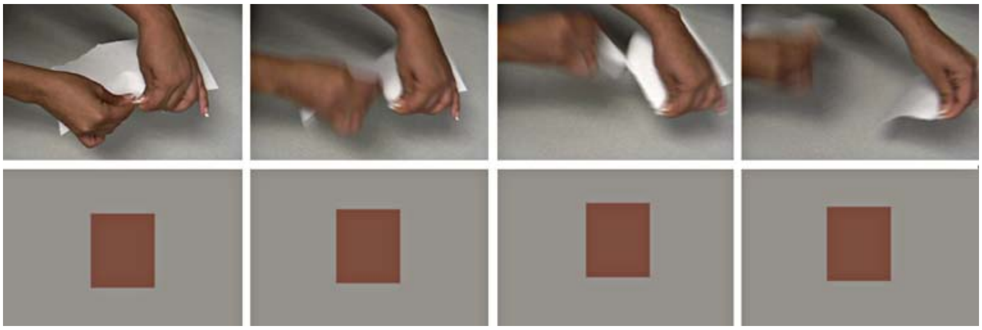


Fig. 1 Action clip (*top*) and control clip (*bottom*)

There were significant activations in the inferior frontal gyrus bilaterally, the middle frontal gyrus bilaterally, and the superior temporal gyrus bilaterally. There was also a cluster in the superior frontal gyrus of the right hemisphere, and on the medial wall of the right frontal lobe. Several thalamic foci were significantly active. One focus can be clearly identified as the medial geniculate nucleus on both sides (see Fig. 2), in accordance with probabilistic anatomical maps (Rademacher et al. 2002). The other focus is located more anteriorly in the thalamus at (13, –10, 3) and may be the ventrolateral nucleus, a thalamic motor nucleus. A similar activation (11, –12, 10) was found by Bestmann et al. (2004) during repetitive transcranial magnetic stimulation to the primary motor cortex. These activations are described in Table 1, and depicted in Fig. 2.

Activity associated with seeing an action

We next performed a contrast of brain regions more activated by the action video clip compared with the control video clip. These activations are detailed in Table 1 and depicted in Fig. 2. Significant signal changes were found in the occipital lobe, in the intraparietal region, in the precentral gyrus, and in the inferior frontal gyrus, as well as subcortical activations in the lateral geniculate nucleus.

Activity associated with hearing and seeing and action simultaneously

This contrast compared the activity during the audiovisual action clip to activity during the audiovisual control clip. There was significantly increased signal for the action stimulus throughout the occipital lobe, in the intraparietal region, in the postcentral gyrus bilaterally, and in the inferior frontal gyrus extending into the middle frontal gyrus bilaterally. These activations are detailed in Table 1 and depicted in Fig. 2.

Cross-modal interactions

To find areas activated by both vision and hearing, we looked for voxels that were active in all three stimulation conditions versus rest. Only one region was active in all three contrasts, the STS bilaterally (see Fig. 3). This same region was active for both the action and control stimuli.

We were also interested in identifying voxels which showed more activity in the A + V condition compared with the A and V conditions. To find this pattern, we performed an (A + V minus A) contrast, and an (A + V minus V) contrast for action and control stimuli separately. We then identified voxels that reached a z score of 2.3 or higher in both contrasts. Only one region survived this statistical threshold for the action stimuli, in the left ventral premotor cortex (see Figs. 4, 5). Activity in this region seems to be present only in the A + V condition, but not in the A condition or V condition. No voxels met these criteria for the control stimuli.

Discussion

Consistent with previous work, activations throughout the frontal and parietal sensorimotor networks were found to be greater for the action videos and action sounds compared with control stimuli. While we did not find greater signal for action sounds in the premotor cortex, this may be attributed to the reduced power of our small sample size.

A major aim of this study was to understand audiovisual interactions in the processing of action-related sensory information. We found that the posterior STS was the only region that was activated by both sound and video alone, and by the combination the two. The STS is involved in the perception of biological motion (Grossman et al. 2000; Jellema and Perrett 2003; Puce and Perrett 2003; Saygin et al. 2004; Vaina et al. 2001), and may participate in the process of matching the visual properties of observed actions onto the self (Iacoboni et al. 2007). This region has long been known to receive inputs from multiple senses in monkey (Hikosaka et al. 1988; Seltzer and Pandya 1989, 1994) and recent data have confirmed that STS neurons in the monkey are involved in cross-sensory integration for action stimuli (Barraclough et al. 2005). There is mounting evidence that STS also serves a multisensory integration function in humans. This has been investigated extensively in the domain of speech perception, where STS has been found to respond more when speech sounds are congruent with lip movements than when they are incongruent (Calvert et al. 2000; Macaluso et al. 2004), and more generally to the combination of hearing and seeing speech together (Wright et al. 2003). Recently, Beauchamp et al. (2004) investigated the audiovisual properties of STS using fMRI. Using videos and sounds of tool use, they found patches of the STS that responded exclusively to sound or vision, and some patches that responded to both. Both the Pizzamiglio et al. (2005) study and the Gazzola et al. (2006) study found activations in the STS to be greater for action sounds compared with non-action sounds.

In addition to its sensory properties, the STS may also be modulated by motor control. Gazzola et al. (2006) found STS (and the nearby middle temporal gyrus) to be activated during execution as well as observation of action sounds. Iacoboni et al. (2001) found that the posterior STS was activated not only during action observation, but also during action execution, and even more so during imitation. This pattern of activity suggests that the STS may be receiving reafferent motor signals. A ‘‘forward model’’ is

Table 1 Areas more active for action stimuli compared with control stimuli

MNI coordinates			Anatomical location	Brodmann's area	*Z* score
X	*Y*	*Z*			
Condition A: Action sounds minus control sounds					
–50	38	0	Left inferior frontal gyrus	45	3.75
52	36	–8	Right inferior frontal gyrus	45	3.83
–40	34	10	Left middle frontal gyrus	46	3.71
54	42	16	Right middle frontal gyrus	46	3.19
16	62	24	Right superior frontal gyrus	10	3.46
2	48	36	Right medial frontal gyrus	9	3.50
–44	–16	6	Left superior temporal gyrus	22, 41	4.79
52	–28	8	Right superior temporal gyrus	22, 41	5.21
–14	–25	0	Left medial geniculate, thalamus		3.70
16	–24	–4	Right medial geniculate, thalamus		3.40
13	–10	3	Right ventral lateral nucleus, thalamus		
Condition V: Action video minus control video					
48	–72	0	Right middle occipital gyrus	18, 19	7.24
8	–86	0	Midline cuneus	17, 18	6.82
34	–54	66	Right intraparietal sulcus	7	6.77
–32	–54	62	Left intraparietal sulcus	7	7.33
–26	–12	58	Left precentral gyrus	6	4.97
44	–8	58	Right precentral gyrus	6	4.58
64	8	24	Right inferior frontal gyrus	9	4.62
–60	4	32	Left inferior frontal gyrus	9	4.88
–60	–18	26	Left postcentral gyrus	3	5.76
72	–20	20	Right postcentral gyrus	3	4.34
22	–28	0	Right lateral geniculate nucleus		4.09
–18	–28	–2	Left lateral geniculate, thalamus		3.49
Condition A + V: AV action minus AV control					
54	–68	–2	Right middle occipital gyrus	19	7.03
–48	–80	0	Left middle occipital gyrus	29	6.52
0	–84	8	Midline cuneus	17, 18	6.59
–38	–46	64	Left intraparietal sulcus	7	6.34
40	–46	66	Right intraparietal sulcus	7	6.14
–66	–18	32	Left postcentral gyrus	3	4.86
70	–28	32	Right postcentral gyrus	3	3.96
64	10	28	Right inferior/middle frontal gyrus	46	5.00
–66	0	28	Left inferior/middle frontal gyrus	46	4.09

an internal simulation in which the sensory consequences of an action are predicted (Iacoboni et al. 2007; Iacoboni et al. 2001). The posterior STS may be the site at which the predicted sensory consequences of an action are matched with the observed action. Our present data suggest that the STS is capable of recognizing the sensory consequences of an action in both the visual and auditory domains, and thus may be involved in multimodal forward modeling.

Interestingly, in our study STS was activated not only by the paper tearing video, which involves the kind of hand-object interaction which easily activates STS cells in monkeys, but also by the control stimulus, which was an abstract shape moving rhythmically. This suggests that the activity of the STS, and its responsiveness to hearing and sound is not limited to purely biological stimuli. Consistent with this notion, Calvert et al. (2001) have shown cross-modal responses in STS to non-biological visual and auditory patterns. It may also be that the apparently self-propelled motion of the square in the control stimulus may have evoked the perception of animacy, a property which

Fig. 2 Brain regions significantly more active for action stimuli compared with control stimuli for the three conditions, A, V, and A + V. **a** Shows action sounds compared with control sounds, **b** shows action video compared with control video, and **c** shows simultaneous sound and video for action compared with control stimuli

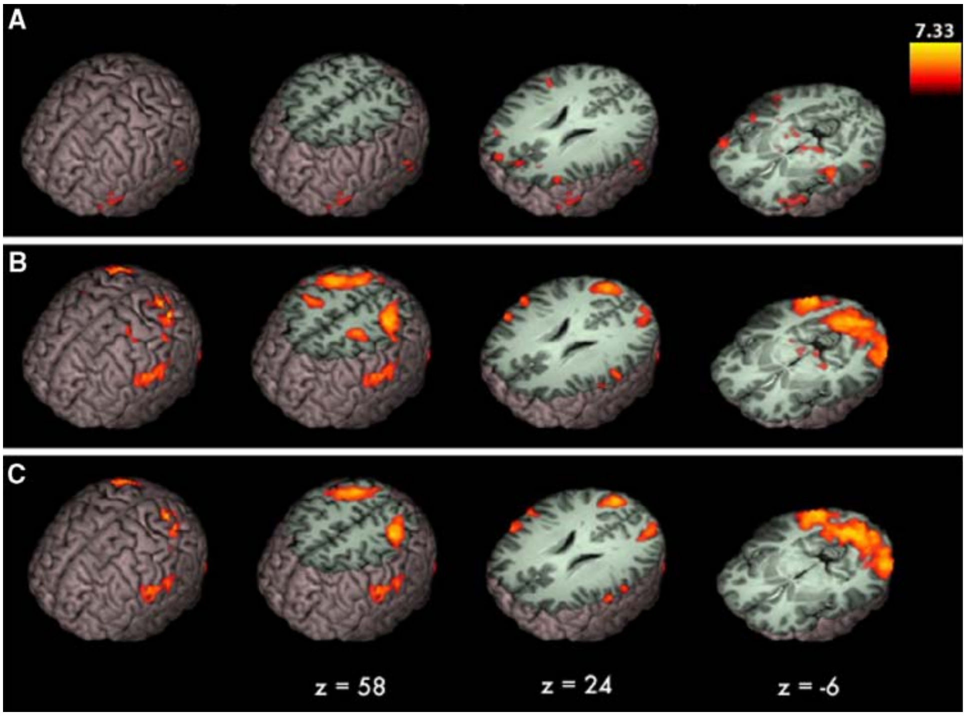
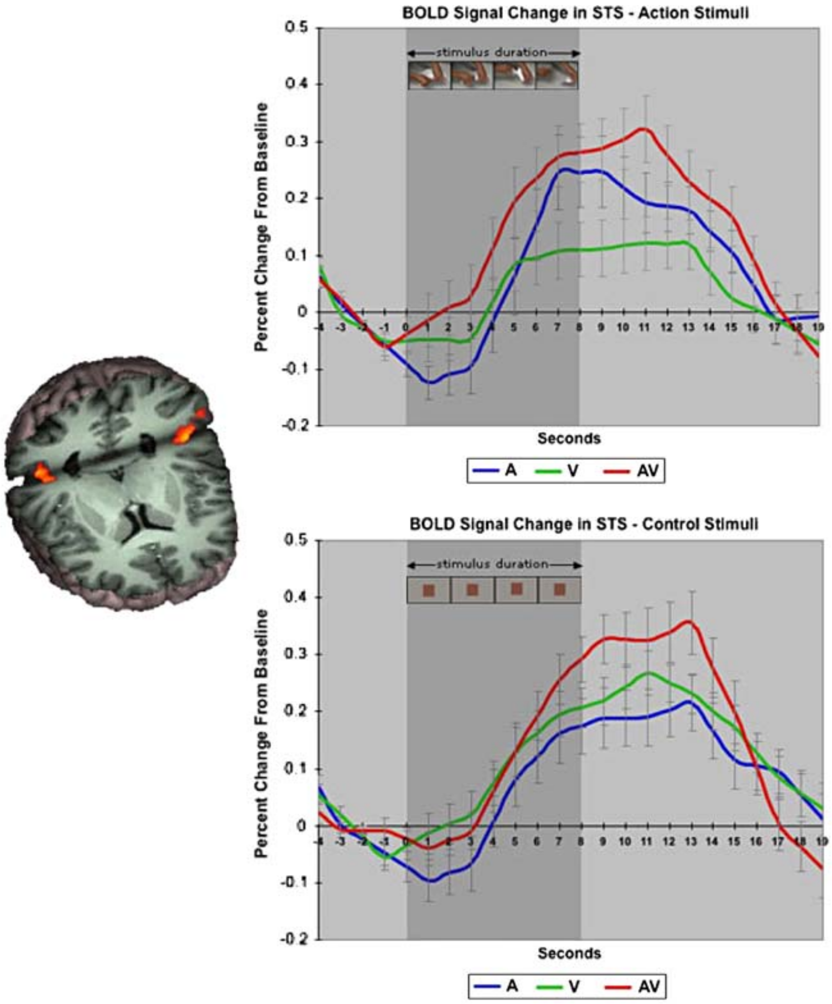


Fig. 3 The posterior superior temporal sulcus was active in A, V, and A + V conditions for action stimuli (*top*) and control stimuli (*bottom*). The *graphs* show the average timecourses for a window of time surrounding each event of interest, linearly interpolated to 1 s resolution



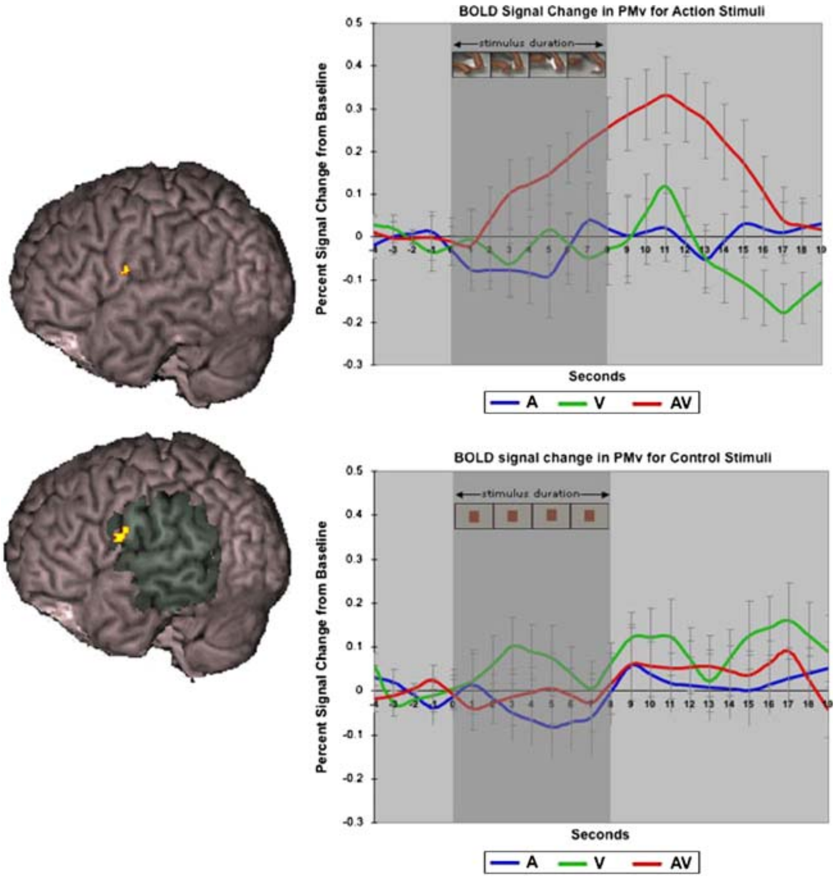


Fig. 4 The ventral premotor cortex was more active in the A + V condition compared with both A and V conditions for action stimuli (*top*). It was not activated by control stimuli (*bottom*). The center of this cluster is at (–64, 0, 18). The *graphs* show the average timecourses for a window of time surrounding each event of interest, interpolated to 1 s resolution

has been associated with STS activity (Blakemore et al. 2003). Recent imaging studies have found activation in STS when subjects watch abstract geometric shapes when those shapes are perceived as acting intentionally (Castelli et al. 2000; Schultz et al. 2003).

While STS seemed to be involved in sensory integration in both our action and non-action stimuli, we found a small region in the ventral premotor cortex with a response pattern suggesting a role in cross-modal integration that is specific to actions. This region was more active for the A + V action video than it was for either modality alone, and it was not activated by any of the control stimuli. While we attempted to match our control and action stimuli, our claim that this pattern reflects action-related processing should be tempered by the fact that we had only one kind of action and control stimulus in this experiment. Particularly in the auditory domain, the control stimulus may differ from the action stimulus in meaningful ways. This spot lies at (–64, 0, 18) on the inferior part of the left precentral gyrus. According to probabilistic cytoarchitectonic maps (Amunts et al. 1999; Geyer 2004; Geyer et al. 1999, 2000b), this putative multimodal integration site is situated posterior to area 44, inferior to area 6, lateral to area 3a, and anterior to area 3b (see Fig. 5). The location of this region at the border of several cytoarchitectonically distinct patches of cortex may relate to its role in cross-sensory integration. Rademacher et al. (2001) have suggested that cytoarchitectonic ''border zones'' are places where integration takes place because they can draw on processing from both regions. Our data suggest that this region of ventral premotor cortex may be such an area that integrates sensory information about meaningful actions. However, rather than lying at the junction of areas 44, 6, 3a and 3b, this sensory integration activation may lie in a yet-to-be-defined region which is found between these three areas. According to the probabilistic cytoarchitectonic maps, the activation found here has approximately 10% probability of falling within either areas 44, 6, 3a, or 3b (see Fig. 5). We speculate that a small cytoarchitectonically distinct patch of cortex may lie in the ventral part of the precentral gyrus between these areas. Anatomical mapping of this critical region is indeed undergoing (K. Amunts, personal communication). Such a region would be in a position to integrate processing from the surrounding cortex.

While it may seem unusual that this region shows activation for the conjunction of audio and visual stimuli but not for either modality alone, this pattern of activity has been found in the monkey brain and elsewhere in the human nervous system. Meredith and Stein (1986) studied

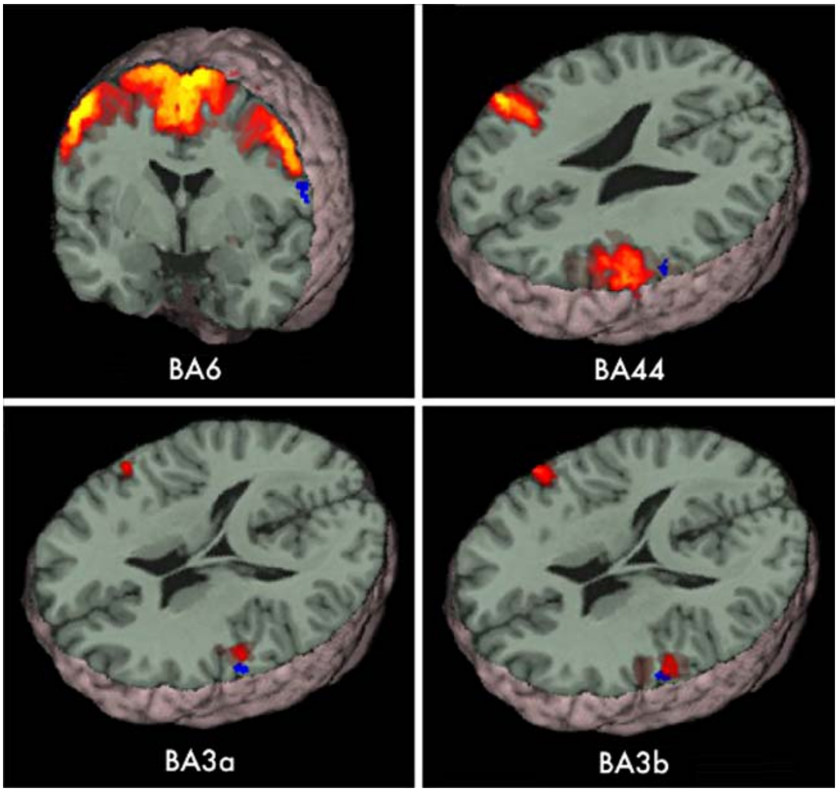


Fig. 5 The activation in ventral premotor cortex (*blue*) lies ventral to BA6, posterior to BA44, lateral to BA3a and anterior to BA3b. Cytoarchitectural probability maps are from Amunts et al. (1999), Geyer (2004), Geyer et al. (1999, 2000a)

multimodal interactions in the superior colliculus of the cat. They found that many neurons showed a little to no response to an auditory or visual stimulation on its own, but a large response when both were presented together. In fact, there was an inverse relationship between the size of the multimodal facilitation effect and the degree to which a neuron was stimulated by each modality alone. This kind of processing may serve to enhance recognition when perception by each modality on its own is insufficient to elicit the activity necessary for recognition.

A secondary aim of this study was to understand how the brain recognizes actions by sound alone. Listening to the sound of paper tearing (A) produced significantly greater signal changes in several prefrontal clusters compared with listening to the control sound. These activations were bilateral, and included a cluster in the pars triangularis of the inferior frontal gyrus extending dorsally into the medial frontal gyrus. The ventrolateral prefrontal cortex (particularly areas 45 and 12) is known to show responses to auditory stimuli and is anatomically connected with auditory cortex (Hackett et al. 1999; Romanski et al. 1999a, b; Romanski and Goldman-Rakic 2002). However, this location probably does not correspond to area F5 where the audiovisual mirror neurons were found. The human analog to area F5 is considered to be the pars opercularis of the inferior frontal gyrus (area 44) (Geyer et al. 2000a; Rizzolatti and Arbib 1998) which is more ventral and posterior compared with this activation. Visual action observation does activate the pars triangularis (area 45), but this activation too tends to be more posterior than what we are observing here with action sounds. A recent meta-analysis of 58 subjects observing hand actions (Molnar-Szakacs et al. 2005) found a peak in pars triangularis at (50, 26, 4) about a centimeter posterior to the peak of the activation found here with action sounds (52, 36, –8).

The anterior IFG, activated here more by an action sound than a non-action sound, is known to be associated with semantic processing in language tasks (see Bookheimer 2002 for review), and may be more generally involved in associating sounds with their meanings. Recent neuroimaging studies have found this region activated when subjects listen to environmental sounds (Adams and Janata 2002; Engelien et al. 1995; Maeder et al. 2001; Zatorre et al. 2004). For example, Adams and Janata (2002) found a similar activation to ours when subjects were required to recognize objects based on the sounds they make. Based on this, they argue that the inferior frontal gyrus plays a role in associating sounds with their conceptual semantic representations. Our data suggest that this is true not only for objects, but for actions as well. Subjects in the present study were under instructions to simply listen to the sounds, and so these activations are probably related to the passive recognition of the meaning of the paper tearing sound.

Signal changes related to seeing an action (V) were consistent with previous data showing activation of motor control networks during action observation. There were greater activations in the intraparietal region, premotor cortex, inferior frontal cortex, and secondary somatosensory cortex during action observation compared with watching the non-action control. The inferior frontal activation was located in the dorsal part of the pars opercularis (area 44). Molnar-Szakacs et al. (2005) argue that this dorsal sector of the pars opercularis represents the true ''mirror'' region of the human motor system, as it is active during both observation but even more so during imitation of motor tasks. We found a similar activation of mirror networks activated in the multimodal (A + V) action video compared with the multimodal control video. The anterior/dorsal prefrontal activation found in the audio only condition (A) was not seen in this analysis. This suggests that the visual modality dominates processing when both auditory and visual information are present.

In summary, our data confirm previous reports that the STS is involved in cross-sensory processing, and show that its involvement extends to action-related stimuli. We also show that a cytoarchitectonically undefined region of the ventral premotor cortex responds specifically to the conjunction of visual and auditory action-related stimuli. Neurons which respond to the combination of sight and sound are interesting because they may produce a representation of the action that is modality-independent. The left-lateralization of this region is consistent with data showing left lateralization of action sound representations (Aziz-Zadeh et al. 2004) and with the hypothesis that multimodality in action perception is an important precursor of language. The hypothesis that sensorimotor representations are crucial to language is strengthened by recent data showing a correspondence between the somatotopic activations of the premotor cortex and the semantic content of a sentence which refers to the body. Aziz-Zadeh et al. (2006) showed that hearing sentences which involve a body part activate the same regions in premotor cortex that are activated by the observation of actions performed with that same effector. We suggest that this region in the ventral premotor cortex may contribute to conceptual representations of actions that abstract across both agency (self and other) and sensory modality (hearing and seeing) (Kaplan and Iacoboni 2005).

Acknowledgments For generous support the authors also wish to thank the Brain Mapping Medical Research Organization, Brain Mapping Support Foundation, Pierson-Lovelace Foundation, The Ahmanson Foundation, Tamkin Foundation, Jennifer Jones-Simon Foundation, Capital Group Companies Charitable Foundation, Robson Family, William M. and Linda R. Dietel Philanthropic Fund at the Northern Piedmont Community Foundation, Northstar Fund, the National Center for Research Resources grants RR12169, RR13642 and RR08655, and NIH grant MH63680. The authors would also like to thank Lisa Aziz-Zadeh for her contributions to this work.

References

Adams RB, Janata P (2002) A comparison of neural circuits underlying auditory and visual object categorization. Neuroimage 16(2):361–377

Amunts K, Schleicher A, Burgel U, Mohlberg H, Uylings HB, Zilles K (1999) Broca's region revisited: Cytoarchitecture and intersubject variability. J Comp Neurol 412(2):319–341

Aziz-Zadeh L, Iacoboni M, Zaidel E, Wilson S, Mazziotta J (2004) Left hemisphere motor facilitation in response to manual action sounds. Eur J Neurosci 19(9):2609–2612

Aziz-Zadeh L, Wilson SM, Rizzolatti G, Iacoboni M (2006) Congruent embodied representations for visually presented actions and linguistic phrases describing actions. Curr Biol 16(18):1818–1823

Barraclough NE, Xiao D, Baker CI, Oram MW, Perrett DI (2005) Integration of visual and auditory information by superior temporal sulcus neurons responsive to the sight of actions. J Cogn Neurosci 17(3):377–391

Beauchamp MS, Argall BD, Bodurka J, Duyn JH, Martin A (2004) Unraveling multisensory integration: Patchy organization within human sts multisensory cortex. Nat Neurosci 7(11):1190–1192

Behrens T, Woolrich MW, Smith S (2003) Multi-subject null hypothesis testing using a fully bayesian framework: theory. In: Hum brain mapping meeting, New York

Bestmann S, Baudewig J, Siebner HR, Rothwell JC, Frahm J (2004) Functional mri of the immediate impact of transcranial magnetic stimulation on cortical and subcortical motor circuits. Eur J Neurosci 19(7):1950–1962

Blakemore SJ, Frith C (2005) The role of motor contagion in the prediction of action. Neuropsychologia 43(2):260–267

Blakemore SJ, Boyer P, Pachot-Clouard M, Meltzoff A, Segebarth C, Decety J (2003) The detection of contingency and animacy from simple animations in the human brain. Cereb Cortex 13(8):837–844

Bookheimer S (2002) Functional mri of language: new approaches to understanding the cortical organization of semantic processing. Annu Rev Neurosci 25:151–188

Buccino G, Binkofski F, Fink GR, Fadiga L, Fogassi L, Gallese V et al (2001) Action observation activates premotor and parietal areas in a somatotopic manner: an fmri study. Eur J Neurosci 13(2):400–404

Calvert GA, Campbell R, Brammer MJ (2000) Evidence from functional magnetic resonance imaging of crossmodal binding in the human heteromodal cortex. Curr Biol 10(11):649–657

Calvert GA, Hansen PC, Iversen SD, Brammer MJ (2001) Detection of audio-visual integration sites in humans by application of electrophysiological criteria to the bold effect. Neuroimage 14(2):427–438

Castelli F, Happe F, Frith U, Frith C (2000) Movement and mind: a functional imaging study of perception and interpretation of complex intentional movement patterns. Neuroimage 12(3):314–325

Engelien A, Silbersweig D, Stern E, Huber W, Doring W, Frith C et al (1995) The functional anatomy of recovery from auditory agnosia. A pet study of sound categorization in a neurological patient and normal controls. Brain 118(Pt 6):1395–1409

Forman SD, Cohen JD, Fitzgerald M, Eddy WF, Mintun MA, Noll DC (1995) Improved assessment of significant activation in functional magnetic resonance imaging (fmri): use of cluster-size threshold. Magn Reson Med 33:636–647

Friston KJ, Worsley KJ, Frakowiak RSJ, Mazziotta JC, Evans AC (1994) Assessing the significance of focal activations using their spatial extent. Hum Brain Mapp 1:214–220

Gallese V, Lakoff G (2005) Brain's concepts: the role of the sensory-motor system in conceptual knowledge. Cogn Neuropsychol 22:455–479

Gallese V, Fadiga L, Fogassi L, Rizzolatti G (1996) Action recognition in the premotor cortex. Brain 119(Pt 2):593–609

Gazzola V, Aziz-Zadeh L, Keysers C (2006) Empathy and the somatotopic auditory mirror system in humans. Curr Biol 16(18):1824–1829

Geyer S (2004) The microstructural border between the motor and cognitive domain in the human cerebral cortex. Adv Anat Embryol Cell Biol 174:1–89

Geyer S, Schleicher A, Zilles K (1999) Areas 3a, 3b, and 1 of human primary somatosensory cortex. Neuroimage 10(1):63–83

Geyer S, Matelli M, Luppino G, Zilles K (2000a) Functional neuroanatomy of the primate isocortical motor system. Anat Embryol (Berl) 202(6):443–474

Geyer S, Schormann T, Mohlberg H, Zilles K (2000b) Areas 3a, 3b, and 1 of human primary somatosensory cortex. Part 2. Spatial normalization to standard anatomical space. Neuroimage 11(6 Pt 1):684–696

Grezes J, Decety J (2001) Functional anatomy of execution, mental simulation, observation, and verb generation of actions: a meta-analysis. Hum Brain Mapp 12(1):1–19

Grezes J, Armony JL, Rowe J, Passingham RE (2003) Activations related to ''mirror'' and ''canonical'' neurones in the human brain: an fmri study. Neuroimage 18(4):928–937

Grossman E, Donnelly M, Price R, Pickens D, Morgan V, Neighbor G et al (2000) Brain areas involved in perception of biological motion. J Cogn Neurosci 12(5):711–720

Hackett TA, Stepniewska I, Kaas JH (1999) Prefrontal connections of the parabelt auditory cortex in macaque monkeys. Brain Res 817(1–2):45–58

Hikosaka K, Iwai E, Saito H, Tanaka K (1988) Polysensory properties of neurons in the anterior bank of the caudal superior temporal sulcus of the macaque monkey. J Neurophysiol 60(5):1615–1637

Iacoboni M, Woods RP, Brass M, Bekkering H, Mazziotta JC, Rizzolatti G (1999) Cortical mechanisms of human imitation. Science 286(5449):2526–2528

Iacoboni M, Koski LM, Brass M, Bekkering H, Woods RP, Dubeau MC et al (2001) Reafferent copies of imitated actions in the right superior temporal cortex. Proc Natl Acad Sci USA 98(24):13995–13999

Iacoboni M, Molnar-Szakacs I, Gallese V, Buccino G, Mazziotta JC, Rizzolatti G (2005) Grasping the intentions of others with one's own mirror neuron system. PLoS Biol 3(3):e79

Iacoboni M, Kaplan JT, Wilson S (2007) A neural architecture for imitation. In: Nehaniv CL, Dautenhahn K (eds) Models and mechanisms of imitation and social learning in robots, humans and animals: behavioural, social, and communicative dimension. Cambridge University Press, London, pp 71–87

Jellema T, Perrett DI (2003) Cells in monkey sts responsive to articulated body motions and consequent static posture: a case of implied motion? Neuropsychologia 41(13):1728–1737

Kaplan JT, Iacoboni M (2005) Listen to my actions!Behav Brain Sci 28(2):135

Kaplan JT, Iacoboni M (2006) Getting a grip on other minds: mirror neurons, intention understanding, and cognitive empathy. Soc Neurosci 1(3–4):175–183

Keysers C, Kohler E, Umilta MA, Nanetti L, Fogassi L, Gallese V (2003) Audiovisual mirror neurons and action recognition. Exp Brain Res 153(4):628–636

Kohler E, Keysers C, Umilta MA, Fogassi L, Gallese V, Rizzolatti G (2002) Hearing sounds, understanding actions: action representation in mirror neurons. Science 297(5582):846–848

Koski L, Wohlschlager A, Bekkering H, Woods RP, Dubeau MC, Mazziotta JC et al (2002) Modulation of motor and premotor activity during imitation of target-directed actions. Cereb Cortex 12(8):847–855

Lahav A, Saltzman E, Schlaug G (2007) Action representation of sound: audiomotor recognition network while listening to newly acquired actions. J Neurosci 27(2):308–314

Macaluso E, George N, Dolan R, Spence C, Driver J (2004) Spatial and temporal factors during processing of audiovisual speech: a pet study. Neuroimage 21(2):725–732

Maeder PP, Meuli RA, Adriani M, Bellmann A, Fornari E, Thiran JP et al (2001) Distinct pathways involved in sound recognition and localization: a human fmri study. Neuroimage 14(4):802–816

Meredith MA, Stein BE (1986) Visual, auditory, and somatosensory convergence on cells in superior colliculus results in multisensory integration. J Neurophysiol 56(3):640–662

Molnar-Szakacs I, Iacoboni M, Koski L, Mazziotta JC (2005) Functional segregation within pars opercularis of the inferior frontal gyrus: evidence from fmri studies of imitation and action observation. Cereb Cortex 15(7):986–994

Pizzamiglio L, Aprile T, Spitoni G, Pitzalis S, Bates E, D'Amico S et al (2005) Separate neural systems for processing action- or non-action-related sounds. Neuroimage 24(3):852–861

Puce A, Perrett D (2003) Electrophysiology and brain imaging of biological motion. Philos Trans R Soc Lond B Biol Sci 358(1431):435–445

Rademacher J, Burgel U, Geyer S, Schormann T, Schleicher A, Freund HJ et al (2001) Variability and asymmetry in the human precentral motor system. A cytoarchitectonic and myeloarchitectonic brain mapping study. Brain 124(Pt 11):2232–2258

Rademacher J, Burgel U, Zilles K (2002) Stereotaxic localization, intersubject variability, and interhemispheric differences of the human auditory thalamocortical system. Neuroimage 17(1):142–160

Rizzolatti G, Arbib MA (1998) Language within our grasp. Trends Neurosci 21(5):188–194

Rizzolatti G, Craighero L (2004) The mirror-neuron system. Annu Rev Neurosci 27:169–192

Rizzolatti G, Fadiga L, Gallese V, Fogassi L (1996) Premotor cortex and the recognition of motor actions. Brain Res Cogn Brain Res 3(2):131–141

Romanski LM, Goldman-Rakic PS (2002) An auditory domain in primate prefrontal cortex. Nat Neurosci 5(1):15–16

Romanski LM, Bates JF, Goldman-Rakic PS (1999a) Auditory belt and parabelt projections to the prefrontal cortex in the rhesus monkey. J Comp Neurol 403(2):141–157

Romanski LM, Tian B, Fritz J, Mishkin M, Goldman-Rakic PS, Rauschecker JP (1999b) Dual streams of auditory afferents target multiple domains in the primate prefrontal cortex. Nat Neurosci 2(12):1131–1136

Saygin AP, Wilson SM, Hagler DJ Jr, Bates E, Sereno MI (2004) Point-light biological motion perception activates human premotor cortex. J Neurosci 24(27):6181–6188

Schubotz RI, von Cramon DY (2003) Functional-anatomical concepts of human premotor cortex: evidence from fmri and pet studies. Neuroimage 20(Suppl 1):S120–S131

Schubotz RI, von Cramon DY, Lohmann G (2003) Auditory what, where, and when: a sensory somatotopy in lateral premotor cortex. Neuroimage 20(1):173–185

Schultz RT, Grelotti DJ, Klin A, Kleinman J, Van der Gaag C, Marois R et al (2003) The role of the fusiform face area in social cognition: implications for the pathobiology of autism. Philos Trans R Soc Lond B Biol Sci 358(1430):415–427

Seltzer B, Pandya DN (1989) Frontal lobe connections of the superior temporal sulcus in the rhesus monkey. J Comp Neurol 281(1):97–113

Seltzer B, Pandya DN (1994) Parietal, temporal, and occipital projections to cortex of the superior temporal sulcus in the rhesus monkey: a retrograde tracer study. J Comp Neurol 343(3):445–463

Vaina LM, Solomon J, Chowdhury S, Sinha P, Belliveau JW (2001) Functional neuroanatomy of biological motion perception in humans. Proc Natl Acad Sci USA 98(20):11656–11661

Wilson SM, Saygin AP, Sereno MI, Iacoboni M (2004) Listening to speech activates motor areas involved in speech production. Nat Neurosci 7(7):701–702

Wolpert DM, Doya K, Kawato M (2003) A unifying computational framework for motor control and social interaction. Philos Trans R Soc Lond B Biol Sci 358(1431):593–602

Worsley KJ, Evans AC, Marrett S, Neelin P (1992) A three-dimensional statistical analysis for cbf activation studies in human brain. J Cereb Blood Flow Metab 12(6):900–918

Wright TM, Pelphrey KA, Allison T, McKeown MJ, McCarthy G (2003) Polysensory interactions along lateral temporal regions evoked by audiovisual speech. Cereb Cortex 13(10):1034–1043

Zatorre RJ, Bouffard M, Belin P (2004) Sensitivity to auditory object features in human temporal neocortex. J Neurosci 24(14):3637–3642