Lateral occipitotemporal cortex and action representation

Patricia Romaiguère [a,*], Bruno Nazarian [b], Muriel Roth [b], Jean-Luc Anton [b], Olivier Felician [c,d]

[a] Aix Marseille Université, CNRS, LNIA UMR 7260, 13331 Marseille, France
[b] Aix Marseille Université, CNRS, Centre IRMf, INT UMR 7289, 13385 Marseille, France
[c] Aix Marseille Université, INSERM, INS UMR-S 1106, 13385 Marseille, France
[d] APHM, Hôpital de la Timone, Service de Neurologie et Neuropsychologie, 13385 Marseille, France

ARTICLE INFO

ABSTRACT

Representation of body and body movements is essential for identifying others intentions or actions or for learning from them. Over the last 10 years, a large collection of research has demonstrated that body representations are distributed across a widely distributed brain network. In this functional magnetic resonance imaging study, we focus on lateral occipitotemporal cortex (LOTC), a recently identified brain region that could represent the body in a multisensory and dynamic manner. We addressed the question of LOTC involvement in visual processing of others' actions through a factorial analysis that manipulated the meaning of an observed action, completed by a psychophysiological interaction analysis. The results show that only left LOTC was significantly activated in relation to others' actions meaning. In addition, only left LOTC was activated during both action observation and action production but it was more dorsal than the activation related to the meaning of observed actions. Furthermore, the psychophysiological interaction analysis showed that when watching meaningless actions, the more dorsal part of the LOTC (the area active during both action production and action observation) had higher functional connectivity with primary visual areas while the more ventral part (that responded to action meaning) had higher correlation with anterior cingulate and medioprefrontal cortices. Taken together these results plead in favour of a strong implication of left LOTC in action observation and understanding, with a possible functional specialisation between the more ventral and the more dorsal parts of LOTC.

© 2014 Elsevier Ltd. All rights reserved.

1. Introduction

Much information on identities, emotions and intentions of our con-specifics can be gathered from visual processing of their faces. However, body postures and movements also carry substantial information to identify others' intentions and actions, or to learn from them.

Human and animal studies showed that perception of bodies and actions is distributed in several cortical areas. In animals, single-cell recordings in monkeys have identified neurons in the inferior temporal cortex which selectively discharge to the sight of human and monkey bodies and body parts (Desimone, Albright, Gross, & Bruce, 1984; Gross, Bender, & Rocha-Miranda, 1969). Neurons in the anterior upper bank of the superior temporal sulcus (STS) have been found to respond selectively to visually presented body posture and movements (Jellema & Perrett, 2003; Perrett et al., 1985). In addition, two populations of neurons (the so-called "mirror neurons") encoding object-directed actions performed by others and oneself have been described in the ventral part of the premotor cortex (area F5) and in area 7b of the inferior parietal lobule (Fogassi et al., 2005; Gallese, Fadiga, Fogassi, & Rizzolatti, 1996; Rizzolatti, Fadiga, Gallese, & Fogassi, 1996).

In humans, the STS is consistently activated when observing or imagining point-light displays depicting human biological motion (Bonda, Petrides, Ostry, & Evans, 1996; Grossman & Blake, 2001; Grossman et al., 2000). The STS has also been found activated in response to the observation of movements of body parts, such as hands (Decety et al., 1997; Grafton, Arbib, Fadiga, & Rizzolatti, 1996; Grezes, Costes, & Decety, 1999; Rizzolatti et al., 1996), or to images of implied motion (Kourtzi & Kanwisher, 2000). Several brain areas that could form the core of the "human mirror system" have also been identified in the rostral part of the inferior parietal lobule (BA 40), the lower part of the precentral gyrus (BA 6), and the posterior part of the inferior frontal gyrus (BA 44) (review in Rizzolatti & Craighero, 2004). The existence of a "human mirror system" does not entail the existence of mirror neurons in the human brain, as this is still to be clearly proven (Rizzolatti & Craighero, 2004).

Finally, two other regions have been associated with visual processing of human bodies and body parts. One region exhibiting

* Correspondence to: UMR 7260 - Case B, FR3C, Université d'Aix-Marseille, 3, Place Victor Hugo, 13331 Marseille cedex 3, France. Tel.: +33 4 13 55 08 40; fax: +33 4 13 55 08 44.
E-mail address: Patricia.Romaiguere@univ-amu.fr (P. Romaiguère).

strong responses to body pictures has been found in the fusiform gyrus, the "fusiform body area", in a region nearby but distinct from the so-called "fusiform face area" (Peelen & Downing, 2005a). The other region located in the lateral occipitotemporal visual cortex, and referred to as the "extrastriate body area" (EBA), has been specifically involved in the visual processing of human bodies and body parts (Downing, Jiang, Shuman, & Kanwisher, 2001). EBA responds strongly to rigid body representations, whether they are presented as photographs, line drawing, stick figure or silhouettes.

Astafiev, Stanley, Shulman, and Corbetta (2004) showed that EBA could also be activated during pointing movements performed without vision of the acting body part. The implication of EBA in both perception and production of gestures may therefore suggest mirror-like properties (Astafiev et al., 2004). Moreover, EBA activation is stronger when the visual feedback is incongruent to the executed movement (David et al., 2007) and in response to images of body parts presented from an allocentric perspective, rather than from an egocentric perspective (Saxe, Jamal, & Powell, 2006). In a previous experiment in which subjects were requested to point to body parts on drawings, photographs or videos of bodies, EBA activity was found to strongly correlate with the level of realism of body representations, suggesting an implication of EBA in the processing of dynamic and biologically relevant body stimuli (Felician et al., 2009). Taken together, these findings suggest that EBA is not only involved in the visual processing of static bodies or body parts, but could represent the body in a multisensory and dynamic manner as well (Jeannerod, 2004).

However, a debate arose as Peelen and Downing (2005b) showed that the part of the lateral occipitotemporal cortex (LOTC) that responded to unseen movements was adjacent but distinct from EBA, whereas Astafiev, Stanley, Shulman, and Corbetta (2005) argued that they were partly overlapping. Also adjacent or partly overlapping with these areas are motion-selective human homologue of macaque area MT (hMT) and object-form-selective lateral occipital complex (LOC). This anatomical convergence of body, action, motion and object selectivity hints at a strong implication of the LOTC in the processing of dynamic and biologically relevant body and action representations. More specifically, LOTC could be involved in object-directed and/or aimed actions with components endorsing mirror-like features. This proposition was supported by a multivoxel pattern analysis by Oosterhof, Wiggett, Diedrichsen, Tipper, and Downing (2010) who found crossmodal representation for vision and action in the bilateral LOTC.

If LOTC is indeed involved in visual processing of others' actions, how does it relate to understanding others' actions and/ or intentions? One of Oostherhof et al.'s experiments was designed to identify those areas that showed crossmodal representation that could also discriminate action goals and/or action effectors. Although they found a gradient of bias in parietal cortex, with the postcentral area being more sensitive to effectors and anterior parietal areas being more sensitive to goals, they found no such bias in the LOTC (Oostherhof et al, 2010). However the actions depicted in their study were reaching to grasp versus reaching to punch, that is very simple actions and straightforward aims. In the present study we address the question using more complex life-like actions. Subjects were presented a series of videos in which usual hand actions were performed. As we were interested in the relevance of the action for the observer but not in tool-use per se, those actions were performed with or without tools and were or not meaningful. We then performed a factorial analysis with factors Objects and Action Meaning. To complete this analysis we wanted to identify areas that could be activated during both action production and action observation. We thus ran a simple localizer trial to look for regions that would respond both during observation of dynamic tool manipulation (videos) and during pantomime by the subjects of the same tool manipulation. In a last step, because videos of tool manipulation could activate hMT and LOC, we also presented photographs of the same tool manipulation that were contrasted to photographs of the tools alone to identify areas that would respond to hands presented in an action context without actual movement nor actual objects.

2. Material and method

2.1. Subjects

Fifteen healthy right-handed subjects (9 women, 6 men; mean age: 22 years, range: 19–41 years) gave their informed consent to participate to the fMRI experiment, which was approved by the local ethics committee. None of them had a history of neurological or psychiatric illness. Handedness was determined for each subject prior to inclusion on the basis of the Edinburgh Inventory (all above 75; mean 96.7; Oldfield, 1971). Subjects were paid for their participation.

2.2. Experimental protocol

2.2.1. Stimuli

Stimulus presentation was controlled by the Labview software package (National Instruments Corporation, Austin, TX, USA). Stimuli consisted of videos and photographs (352×288 pixels of size) or of audio files. Visual stimuli were delivered to a high luminance LCD projector, back projected onto a frosted screen positioned at the back end of the MRI tunnel, and viewed by the subjects through a mirror. Audio stimuli were delivered through an Optime 1 audio system, equipped with Peltor electrodynamic headphones (MR. confon GmbH, Magdeburg, Germany).

2.2.2. Experiment

Six perception conditions were presented to the subjects in four fMRI runs. In the meaningful action condition, subjects viewed videos of human hands manipulating objects (key+locked box, pen+paper, hammer+nail, screwdriver+screw, paint brush+paper, letter opener+envelope, bottle opener+bottle, coffee spoon+cup, washing up sponge+plate, shoe polishing brush+shoe) in their normal usage (opening a lock, writing, hammering a nail, etc.), whereas in the meaningless action condition, subjects viewed videos of usual objects manipulated in a meaningless way. Meaningless actions were constructed by associating movements from one action with two objects from two other different actions, i.e. hammering movements holding a pen and a shoe polishing brush (see Fig. 1). To minimise perceptual conflict in this condition, the two objects were chosen so that they were not usually paired together in any action, and the hand grip to hold them was very different from the hand grip on the "normal" objects and/or the kinematic of the movement was different (i.e., the hand grip on a shoe polishing brush is different from the hand grip on a nail, and because a pen is light with no eccentric mass, the kinematic of the up and down movement is different from the one with a hammer). In the meaningful mime condition, subjects viewed videos of hands miming the same usual tasks as in the meaningful action condition, and in the meaningless mime condition, subjects viewed videos of empty hands performing meaningless gestures (see Fig. 1). In the still action condition, subjects viewed photographs of the beginning of the same usual tasks (a hand holding a key in front of a lock, a hand holding a pen above a blank page of paper, etc.). In the still objects condition, subjects viewed photographs of the objects with no visible body part (a key in a lock, a pen next to a page of paper, etc.; see Fig. 1). Each trial (video or photograph) lasted 3 s. After each video there was a prompt requesting subjects to press one of two buttons with their right index or middle finger to indicate if they could see one or two hands on the videos. There were equal numbers of each. This response was only used to ensure subjects were actually watching the videos and was therefore modelled as a regressor of no interest in all statistical analyses. There were 30 trials per condition, amounting to 180 trials that were presented in a pseudo-random order and distributed in four runs. Within each run, trials were presented with a mean interstimulus interval of 4 s (jitter between 2 and 9 s).

To localise areas that were active during both action production and action observation, we ran two localisation conditions in two separate fMRI acquisition runs. In one run, subjects were required to passively watch videos similar to those used for the meaningful action task (visual localisation condition). Videos were presented for 3 s in a pseudo-random order with a mean interstimulus interval of 4 s (jitter between 2 and 9 s). In another run, subjects were required to mime the same usual actions on audio cues (motor localisation condition). Audio cues were single words (the name of the active object in the task to be mimed: key, pen, hammer, etc.) pronounced for half of them by a male voice and by a female voice for the other half. After 3 s an audio beep instructed the subjects to stop miming. Stimuli were presented in a pseudo-random order with a mean interstimulus interval of 4 s (jitter between 2 and 9 s). This run was always performed last.