NeuroImage **11,** 644–655 (2000)
doi:10.1006/nimg.2000.0578, available online at http://www.idealibrary.com on IDEAL®

The Role of Lateral Occipitotemporal Junction and Area MT/V5 in the Visual Analysis of Upper-Limb Postures

Philippe Peigneux,*,†,‡ Eric Salmon,*,† Martial van der Linden,‡ Gaëtan Garraux,*,† Joël Aerts,*
Guy Delfiore,* Christian Degueldre,* André Luxen,* Guy Orban,§ and Georges Franck,*,†

**Centre de Recherches du Cyclotron, †Service de Neurologie, CHU Sart Tilman, and ‡Service de Neuropsychologie, Université de Liège, B-4000 Liège, Belgium; and §Laboratory for Neuro- and Psychophysiology, University of Leuven, B-3000 Leuven, Belgium*

Received July 26, 1999

Humans, like numerous other species, strongly rely on the observation of gestures of other individuals in their everyday life. It is hypothesized that the visual processing of human gestures is sustained by a specific functional architecture, even at an early prelexical cognitive stage, different from that required for the processing of other visual entities. In the present PET study, the neural basis of visual gesture analysis was investigated with functional neuroimaging of brain activity during naming and orientation tasks performed on pictures of either static gestures (upper-limb postures) or tridimensional objects. To prevent automatic object-related cerebral activation during the visual processing of postures, only intransitive postures were selected, i.e., symbolic or meaningless postures which do not imply the handling of objects. Conversely, only intransitive objects which cannot be handled were selected to prevent gesture-related activation during their visual processing. Results clearly demonstrate a significant functional segregation between the processing of static intransitive postures and the processing of intransitive tridimensional objects. Visual processing of objects elicited mainly occipital and fusiform gyrus activity, while visual processing of postures strongly activated the lateral occipitotemporal junction, encroaching upon area MT/V5, involved in motion analysis. These findings suggest that the lateral occipitotemporal junction, working in association with area MT/V5, plays a prominent role in the high-level perceptual analysis of gesture, namely the construction of its visual representation, available for subsequent recognition or imitation. © 2000 Academic Press

INTRODUCTION

Humans, like numerous species, need to observe gestures of other individuals in order to understand their underlying intentions and thoughts, to react appropriately to their behavior, or to imitate and initiate a new learning (Barresi and Moore, 1996). Moreover, even if human communication relies mainly on auditory signals produced by speech, it also relies on various limb movements, ranging from gesticulation, which accompanies and underlines speech, to autonomous communicative gestures. In the past decade, functional brain imaging studies have convincingly shown that the neural activity involved in the processing of visually presented stimuli may be both task- and stimulus-dependent. Hence, the utmost importance of a correct perception of limb movements of other individuals suggests that the visual processing of gestures might be functionally and anatomically segregated from the visual processing of other visual entities.

Several lines of evidence may support the hypothesis of a specific cognitive and structural architecture for the processing of gestures. Studies of human subjects observing grasping movements have highlighted regional cerebral blood flow (rCBF) increases in brain regions partly overlapping those activated during actual execution of hand actions (Decety *et al.*, 1994; Grafton *et al.*, 1996; Rizzolatti *et al.*, 1996), even when viewing was restricted to simulations of biological hand motion with point-light displays (Bonda *et al.*, 1996). In addition, the nature of the action and its meaning both may affect its neural processing. The viewing of American Sign Language, in contrast to the observation of meaningless "nonsigns," by normal-hearing bilingual (spoken and sign language) subjects elicited activation within left hemisphere structures classically linked to language processing (Neville *et al.*, 1998), as did the observation of meaningful gestures compared to the observation of meaningless ones (Decety *et al.*, 1997). Processing of meaningful gestures is not restricted to the left hemisphere only, since a bilateral occipitotemporal activation was also present in the Neville *et al.* (1998) study and during observation of sign language contrasted with the observation of spoken language (Soderfeldt *et al.*, 1997). Finally, the observation of an action, with the intent to recognize it later, strongly activated parahippocampal mem-

1053-8119/00 $35.00

ory structures, while the same observation, with the intent to imitate the action, elicited rCBF increases in brain regions involved in action generation (Decety *et al.*, 1997). Taken together, these findings favor the hypothesis of a partial functional and anatomical segregation of the processes involved in the processing of gestures, depending on the nature of the perceived action, the relationship between this action and the environment, and the purpose of the visual perception.

Nonetheless, the question of a specific brain locus dedicated to the visual analysis of human gestures remains virtually unanswered. Separate processes for the visual analysis of objects and gestures were posited by Rothi *et al.* (1997) in their cognitive model for limb apraxia, and Schwartz *et al.* (1998) reported recently the case of patient W.A., who sustained a lesion in bilateral ventral mesial occipital-temporal regions and could recognize visually presented gestures, but not visually presented objects, an interesting dissociation which may favor this hypothesis. Yet, to our knowledge no study has directly focused on the comparison between the visual processing of gestures and the visual processing of other classes of visual stimuli in normal subjects using brain-imaging methods. Despite the demonstration of the involvement of specific cerebral networks in action processing, this lack of direct comparison does not allow one to assert that a functional segregation of gesture processing may occur at the specific level of the visual analysis, before access to the lexical or semantic stages. In the present positron emission tomography (PET) experiment, we aimed to address this question assessing the rCBF variations elicited by the visual processing of photographs of upper-limb gestures in comparison to that of objects.

In this perspective, one can show concern about the fact that most studies concerned with action processing in visual modality have used transitive gestures as stimuli, i.e., gestures involving handling objects. This may be a confounding issue since using these transitive gestures does not allow the segregation of the cerebral activity elicited by an irrepressible, automatic, processing of the associated object from that related to gesture analysis per se. Indeed, neurons in the anterior intraparietal sulcus of the monkey parietal cortex are highly selectively activated during both manipulation and fixation of the same object (Murata *et al.*, 1996; Sakata *et al.*, 1995), and object-oriented action and object recognition both activate the anterior intraparietal sulcus in humans (Faillenot *et al.*, 1997), suggesting an interaction effect between the gesture itself and its related object. Conversely, seeing an object facilitates an action congruent with the visual properties of that object (Craighero *et al.*, 1996); it implies that the presentation of manipulable objects for identification is a task which does not prevent automatic activation of the representation of the hand-related posture when processing this object.

Given these concerns, we selected only intransitive postures, i.e., symbolic or meaningless postures which do not imply the handling of objects, in order to prevent possible automatic object-related activation during the visual processing of postures. For the object-processing condition, the same logic led us to avoid any gesture-related processing and to select meaningless and meaningful objects that could not be handled, which we called intransitive objects by analogy to the intransitive postures. Hence, we expected these two categories of intransitive stimuli (postures and objects) to be processed as independently as possible in the visual analysis system. Last, static presentation of gestures (i.e., postures) was preferred for this experiment since actual motion may act as a confounding effect in the differentiation between gestures and objects processing.

MATERIAL AND METHODS

Subjects

Twelve right-handed volunteers (five men and seven women, mean age 22.4 years, range 19–26 years) gave their informed consent to take part in this PET experiment, approved by the Ethical Committee of the University of Liège. None of them had any history of neurological or psychiatric illness, nor needed glasses to correct vision, nor had ever previously participated in a PET experiment. They were paid for their participation.

Apparatus and Scanning Procedure

PET acquisitions were obtained with a Siemens CTI 951 R 16/31 scanner (CTI, Knoxville, TN; in-plane resolution 8.7 mm) in 3D mode. A transmission scan was acquired for attenuation correction using three rotating sources filled with ^{68}Ge. Emission scans consisted of a 30-s background frame followed by a 90-s target frame. The rCBF was measured using 12 iterative infusions of oxygen-15-labeled water (6 mCi/222 MBq in 5 cc saline each), automatically infused for 20 s through a cannula into an antebrachial vein of the left arm. Participants lay in the supine position on the bed of the PET scanner, with a 17-in. monitor located in front of them at a mean distance of 80 cm from the eyes. The screen field of view was 42.1° horizontally and 31.6° vertically.

Cognitive Conditions

Participants were told that the purpose of the experiment was to study with PET the cerebral recruitment associated with semantic and spatial tasks. Twelve scans were obtained during the administration of four separate conditions repeated thrice (Meaningful Posture Naming, Meaningful Object Naming, Meaningless

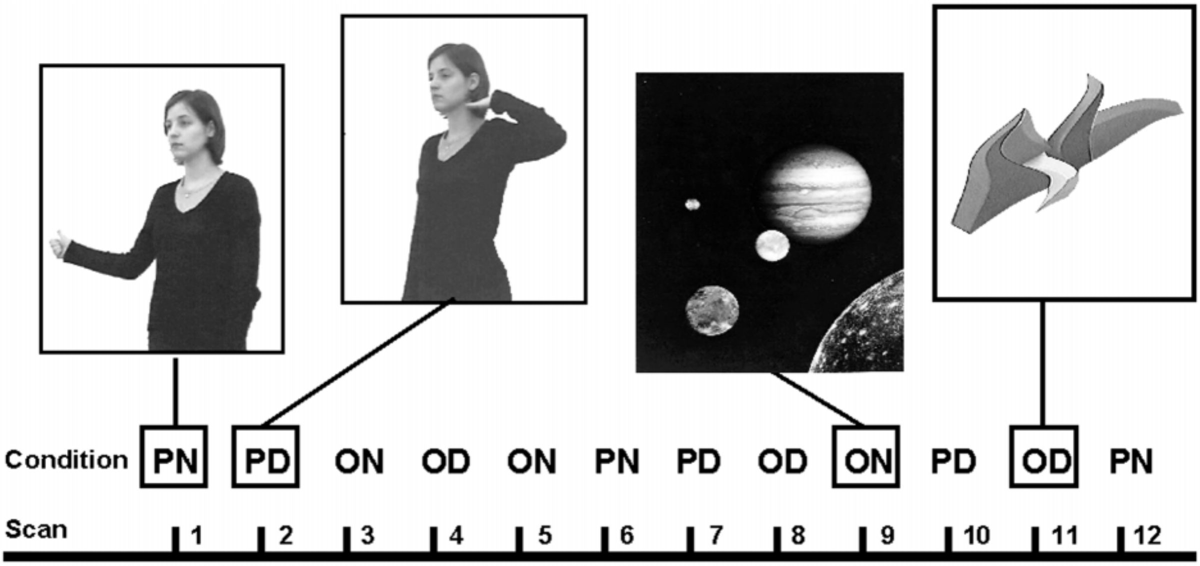

FIG. 1. Experimental design. Each PET scan acquisition (numbered 1 to 12) corresponds to one experimental condition, and each of the four conditions is repeated thrice, in a randomized order. Examples of each stimulus type are given for each respective experimental condition. Abbreviations: PN, meaningful posture naming; PD, meaningless posture orientation decision; ON, meaningful object naming; OD, meaningless object orientation decision.

Posture Orientation Decision, and Meaningless Object Orientation Decision), using two tasks (naming or orientation decision) crossed with two different paired stimuli, presented visually (intransitive postures or intransitive objects, either meaningful or meaningless). These four conditions were successively performed in a randomized order, different for each subject, and the entire four-conditions set was presented twice again in a counterbalanced randomized order (see an example of the timeline in Fig. 1). Prior to scanning and after general explanations and instructions, four examples of each condition were administered to ensure a good understanding of the instructions. These examples were different from the experimental stimuli.

In each experimental condition, 24 photographs were randomly displayed, in gray scale. Each picture was displayed for 5 s, during which time the subject was requested to give the appropriate response, and then the next picture was immediately displayed. No feedback was given during the experiment as to the correctness of the responses. The 24 displays in each condition were obtained using a selection of 12 digitized photographs and then creating a mirror image for each of them, the image being flipped horizontally. Instructions were given verbally 1 min prior to each scan, and the task was performed during the period starting 15 s before and ending 15 s after the 90-s target scan. Gray-scale pictures used as stimuli were displayed full screen in a 800 × 640-pixel format, using a 200-MHz PC system. During the object- and posture-naming conditions, subjects' verbal production was tape recorded through a microphone for subsequent verification, and motor key-press responses were computer stored during object- and posture-orientation decision.

Meaningful Posture Naming (PN)

Subjects were instructed to name aloud, in a brief sentence, visually presented postures. In this condition, only meaningful intransitive upper-limb postures were used as stimuli, i.e., symbolic or communicative postures which do not involve a direct interaction with an object (e.g., military salute, hitch-hiking). The selection of stimuli was made during a preliminary recognition test, administered to another validation group of 20 subjects. Twelve pictures of intransitive upper-limb postures which were recognized at least by 90% of the validation group were retained for the experiment. All postures were performed by the same actor, a young woman dressed in black, fitted clothes. Pictures were obtained in a half-profile perspective (vertical axis rotated 45°), which encompassed the body from the waist to the top of the head (see Fig. 1), with a constant size and a white background on display. The facial expression was always kept neutral.

Meaningful Object Naming (ON)

In exactly the same way as in the PN condition, subjects were instructed to name aloud, in a brief sentence, visually presented objects. By extension to the concept of intransitive postures, we selected here meaningful intransitive tridimensional objects, defined as objects which do not elicit (at least at first glance) the possibility of being handled by humans nor are the product of human technology. To fulfill these criteria, 12 pictures of easy-to-recognize natural objects were selected (e.g., a mountain, a planet, an iceberg).

Meaningless Posture Orientation Decision (PD)

Instructions were given to decide the orientation of a visually presented limb posture, by pressing on the key corresponding to the side toward which this posture was oriented. Twelve pictures of meaningless upper-limb postures were selected on the basis of a preliminary recognition test, keeping only those postures which were correctly classified as meaningless by at least 90% of a validation group of 20 other subjects. All postures were performed by the same actor as in condition PN and the pictures obtained in the same half-profile perspective encompassing the body from the waist to the top of the head. Size, background, and facial expression were also always kept constant. Participants' performance was evaluated by considering the responses for each particular stimulus and its own mirror image. For example, if a "left" response was given for a particular stimulus, then a "right" response was expected for the mirror image, because mirroring reverses the horizontal frame of reference. Participants were specifically instructed during the presentation of the stimuli examples to define their own criteria for deciding the orientation of the entire posture and to use these same criteria in all their subsequent decisions during the experiment.

Meaningless Object Orientation Decision (OD)

Exactly in the same way as in the PD condition, subjects were instructed to decide the orientation of a visually presented tridimensional object by pressing the key corresponding to the side toward which the object was oriented. Twelve meaningless tridimensional objects were created, using digitized deformations applied to pictures of wooden block puzzles, until they no longer evoked any recognizable familiar object. It is noteworthy that the three-dimensional structure of these objects was kept structurally possible in all cases.

PET Data Analysis

PET data were analyzed using the statistical parametric mapping software (SPM96; Wellcome Department of Cognitive Neurology, London; http://www.fil.ion.ucl.ac.uk/spm) implemented in MATLAB (Mathworks, Inc., Sherborn, MA). Data from each subject were realigned using a least-square approach and the first scan as a reference. Following realignment, all images were normalized in a stereotaxic standard space (Talairach and Tournoux, 1988) and smoothed using a 16-mm full-width half-maximum (FWHM) isotropic kernel. The covariance analysis subject-specific (SPM96) design matrix used in the subtractions and conjunctions performed in this study included the three replications in each of the four conditions and considered block effect and individual global flow as confounds. The resulting set of voxel values for each contrast constituted a map of the t statistic (SPM$\{t\}$), transformed to the unit normal distribution (SPM$\{Z\}$) and thresholded at $P < 0.001$ ($Z = 3.09$). Statistical inferences were then obtained at the voxel level (in terms of peak height at $P < 0.05$) corrected for multiple comparisons, using the conjoint test provided in SPM96, by reference to the theory of Gaussian random fields (Friston *et al.*, 1994). The SPM $\{Z\}$ had a volume of 187,921 voxels, with $11.7 \times 12.9 \times 13.3$-mm FWHM smoothness (750.3 RESELS) and 117 residual degrees of freedom. To be fair, we should mention here that an alternative analysis scheme would have been to use a three-way ANOVA model (Woods *et al.*, 1996). Compared to the two-way model evidencing areas of significant task- or stimulus-related changes in group analysis, the three-way model may offer the additional advantage of accounting for potential effects of task replications within multiple subjects and enhance reproducibility of results in a new population. The potential disadvantage is the reduction of the number of degrees of freedom. Since individual behavioral results were constant from scan to scan in each condition, we chose not to take the effect of subject-by-condition replication interaction into account in the present experiment. Hence, the statistical inferences described in this paper pertain to the population studied.

Our statistical design involves fixed-effect subtraction and conjunction analyses. In subtraction analyses, we looked at simple main effects between comparable conditions to emphasize stimulus-related rCBF distribution within each task separately. Four simple subtractions were included in the design matrix: (1) naming postures by comparison with naming objects (according to the formula PN − ON), (2) naming objects by comparison with naming postures (ON − PN), (3) orientation decision on postures by comparison with orientation decision on objects (PD − OD), and (4) orientation decision on objects by comparison with orientation decision on postures (OD − PD). However, subtraction designs require that the experimental conditions involved in the subtraction analyses differ by only one cognitive component (Friston *et al.*, 1996), a criterion which was difficult to fulfill in the present experiment. In this sense, conjunction designs involving two task pairs are much more powerful than simple one-factor subtraction designs and may overcome the problem of baseline selection (Price *et al.*, 1997), because there may be many processing differences in each pair of tasks, provided that the only common processing across differences in task pairs is the cog-

TABLE 1

Stimulus-Related Subtractions

Area of activation	BA	Naming task				Orientation decision task			
		x	*y*	*z*	*Z* score	*x*	*y*	*z*	*Z* score
Postures minus objects									
Inferior temporal/middle occipital gyrus	19/37	−50	−74	4	6.68	−52	−70	10	5.72
	19/37	52	−72	4	6.01	52	−72	4	5.65
Middle temporal gyrus	37	−52	−66	10	6.61	—	—	—	—
	39	56	−64	10	5.36	56	−64	10	5.40
	39	−56	−58	12	6.17	—	—	—	—
Superior temporal gyrus	22	−62	−54	16	5.35	—	—	—	—
	22	58	−58	16	4.67	52	−50	10	5.30
Inferior occipital gyrus	19	−46	−80	0	6.57	—	—	—	—
Objects minus postures									
Lingual gyrus	18	−10	−86	−2	8.27	4	−80	−8	5.89
	18	8	−76	−2	8.33	14	−76	−4	6.26
	18/19	−12	−62	6	7.05	—	—	—	—
Fusiform gyrus	37/19	−30	−46	−6	7.11	—	—	—	—
	37	−26	−44	−10	6.29	—	—	—	—
	18	—	—	—	—	−26	−74	−4	4.62
Middle occipital gyrus	19	−30	−80	20	7.57	−26	−88	14	6.78
	18/19	18	−86	14	7.56	20	−86	18	4.73
Cuneus	17	0	−92	6	8.15				
Calcarine sulcus	17	—	—	—	—	20	−82	8	5.73

Note. Posture-related (top) and object-related (bottom) rCBF peaks disclosed during subtraction analyses related to stimulus type during each task are displayed. Coordinates *x, y, z* refer to the standard Talairach and Tournoux (1988) stereotaxic space. Activation peaks' significance is reported at the voxel level, corrected for multiple comparisons (*P* corr. <0.05). Only the most representative voxels in each Brodmann area (BA) are displayed.

nitive component of interest. Hence, we used a subsequent conjunction analysis to highlight stimulus-specific related rCBF, irrespective of the task. In this conjunction, areas of common activation between the two task pairs were revealed by creating a statistical parametric map of the sum of both simple main effects (subtractions) and eliminating voxels where there were significant differences between activation in the first and the second task pairs (Price and Friston, 1997). A masking procedure ensured that the areas of common activation were already significantly activated ($P < 0.001$) during each of the two simple subtractions. Two conjunction analyses were designed to identify the neural correlates of (1) visually presented posture processing, by computing commonalities between (PN − ON) and (PD − OD) and (2) visually presented object processing, by computing commonalities between (ON − PN) and (OD − PD).

We did not report the results of rCBF task-related comparisons since they exceed the scope of this study. Also, deliberate differences between tasks in the response modalities (naming aloud versus key pressing) and the meaning of the stimuli will be confounding effects in such a comparison. Yet, we make these data available on request.

RESULTS

Behavioral results will be presented first. For the sake of simplification in the presentation of PET results, simple effects related to stimulus type within each task (i.e., stimuli-related subtractions) will be displayed in Table 1 without further comments since they are redundant of conjunction analyses which will be discussed.

Behavioral Analysis

The mean accuracy scores for each task and each stimulus type are shown in Table 2. A two-way ANOVA with Stimulus type (posture vs object) and Task (naming vs orientation decision) as within-subject variables failed to show any significant effect of the interaction or of the stimulus type ($P > 0.8$), but disclosed that the mean accuracy score differs significantly between tasks ($P < 0.05$). Actually, the main point is that the stimulus effect is not affected; there are no significant score differences related to stimulus type within each of these tasks. Thus, the possible relative difficulty of the task is not a confounding factor during analysis related to the stimulus type, which is our main purpose.

TABLE 2

Mean Accuracy Scores

	Mean (%)	Standard error (%)
PN	97	1.2
ON	96	1.3
PD	82	5
OD	81	6

Note. Mean accuracy scores and standard error expressed in percentage for each stimulus type (P, posture or O, object) during each task (N, naming or D, orientation decision).

Stimulus-Related rCBF

In the conjunction analyses related to the stimulus type (see Table 3), meaningless and meaningful stimuli were combined. Upper-limb posture processing, irrespective of the task and of stimulus meaningfulness, was associated with a bilateral activation in the inferior temporal gyrus close to the middle occipital gyrus (BA 19/37) and, in the right hemisphere, in the posterior middle and superior temporal gyri (BA 22 and 39). Main activation sites fall in the lateral occipital-temporal junction, encroaching upon the location of area MT/V5 (see Fig. 2). Conversely, significant activation sites related to object processing, independent of the task, were found more medially located, bilaterally in the middle occipital (BA 18, 19) and fusiform (BA 18, 19, 37) gyri, the cuneus (BA 17, 18/31), and the cerebellum and, in the right hemisphere, in the calcarine sulcus (BA 17), the superior occipital (BA 19), and the lingual (BA 18, 19) gyri.

DISCUSSION

Dissociations between Posture and Object Processing

The most prominent finding of this study is that distinct patterns of neural activity were shown to be differentially related to the processing of visually presented postures and objects, irrespective of the task to be executed on these stimuli and of their meaning. As clearly shown in Fig. 2, upper-limb posture processing elicited activation mainly in the posterior brain at the lateral occipitotemporal junction, encroaching upon

TABLE 3

Stimulus-Related Conjunctions

			Coordinates				
Area of activation	Hemisphere	BA	*x*	*y*	*z*	*Z* score	*P* corr. value
			Postures minus objects				
Middle occipital gyrus	R	19	50	−78	0	8.65	<0.001
Inferior temporal/middle occipital gyrus	L	19/37	−52	−70	10	8.64	<0.001
	R	19/37	52	−66	6	7.81	<0.001
Middle temporal gyrus	R	39	52	−58	10	6.87	<0.001
	L	39	−56	−62	14	7.41	<0.001
Superior temporal gyrus	R	22	52	−50	12	6.34	<0.001
			Objects minus postures				
Fusiform gyrus	L	18	−24	−74	−6	8.29	<0.001
	L	19	−16	−64	−8	7.44	<0.001
	L	37	−22	−56	−10	6.63	<0.001
	R	37	30	−52	−8	7.37	<0.001
	R	19/37	28	−62	−8	7.08	<0.001
Middle occipital gyrus	L	18/19	−22	−88	16	10.24	<0.001
	R	18	22	−88	12	8.79	<0.001
	R	19	34	−84	18	7.75	<0.001
Superior occipital gyrus	R	19	32	−82	26	6.46	<0.001
Lingual gyrus	R	18	14	−76	−4	10.15	<0.001
Cuneus	L	18/31	−12	−72	16	6.34	<0.001
	R	17	4	−72	14	6.39	<0.001
Cerebellum	L	Cb	−4	−78	−16	7.75	<0.001
	R	Cb	26	−44	−14	6.66	<0.001
Calcarine sulcus	R	17	16	−80	8	9.46	<0.001

Note. Posture-related (top) and object-related (bottom) rCBF peaks disclosed during conjunction analysis of stimulus-related subtraction results. Coordinates *x*, *y*, *z* refer to the standard Talairach and Tournoux (1988) stereotaxic space. Activation peaks' significance is reported at the voxel level, corrected for multiple comparisons (*P* corr. <0.05), with a mask on each member of the conjunction (*P* corr. <0.05). Only the most representative voxels in each Brodmann area (BA) are displayed. L, left; R, right.

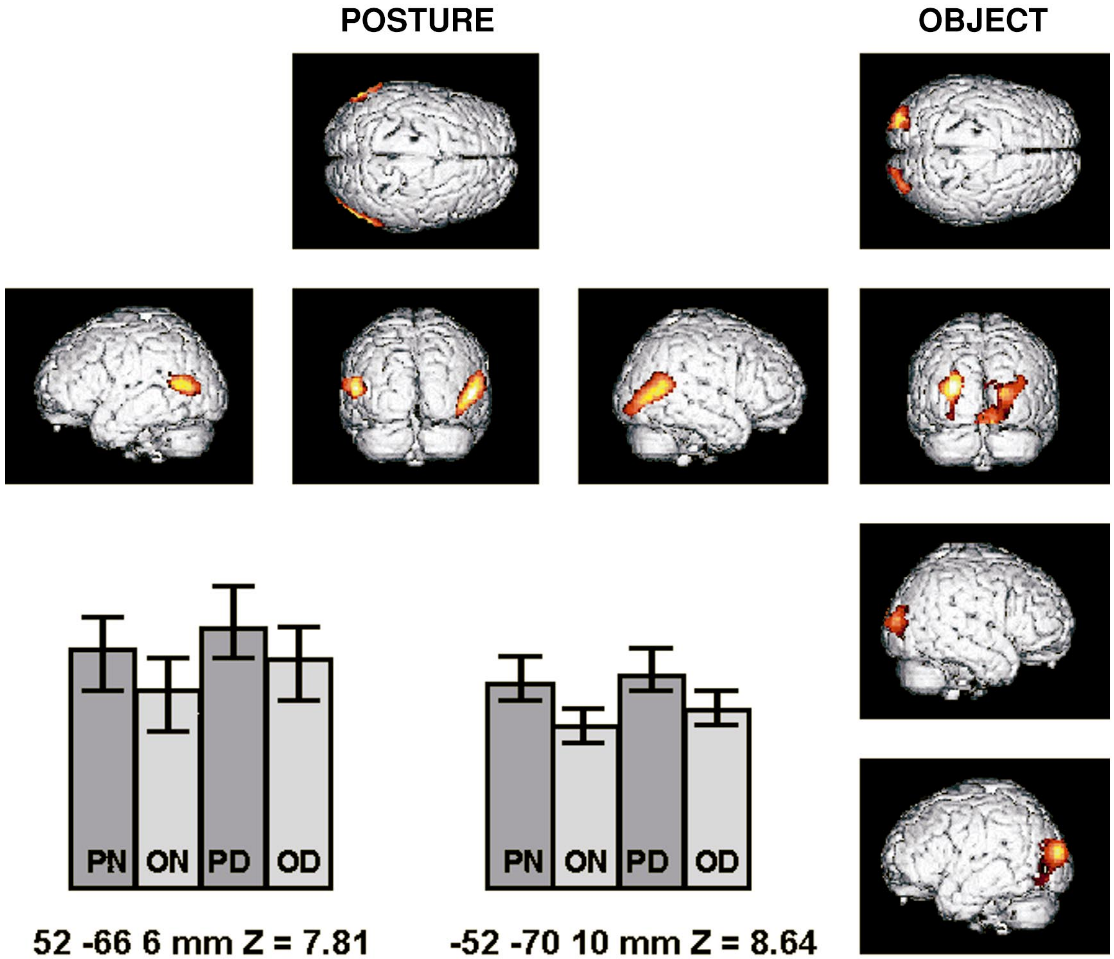

FIG. 2. Brain areas disclosed during posture and object visual processing. Brain areas where rCBF was significantly associated with visual processing of intransitive upper-limb postures (left) and intransitive objects (right) during conjunction analysis, rendered on a template MRI image at the level of peak activation (voxel-level corrected for multiple comparisons, P corr. <0.05) in top, rear, and lateral (right and left) perspective. The mean rCBF is illustrated for each experimental condition in the two highest significant voxels of the lateral occipitotemporal junction, activated in posture-related conditions, one in each hemisphere. Mean rCBF is displayed with minimal and maximal rCBF values measured across all subjects. Condition abbreviations: PN, meaningful posture naming; PD, meaningless posture orientation decision; ON, meaningful object naming; OD, meaningless object orientation decision.

area MT/V5. More medial activity in the posterior and fusiform areas characterized object processing.

Regarding the activation sites reported to be associated with object-related conditions in the present study, our results share commonalities with those of numerous studies focusing on the processing of visually presented objects. Significant rCBF increases in the occipital and fusiform gyri were already found bilaterally during visual object naming (Kiyosawa *et al.*, 1996) and matching (Kohler *et al.*, 1995); left-sided during living–nonliving object categorization (Sergent *et al.*, 1992), overt or covert object naming (Zelkowicz *et al.*, 1998), or matching object to name (Kosslyn *et al.*, 1995); and in the right hemisphere during object identity or location retrieval (Moscovitch *et al.*, 1995) and object recognition (Kanwisher *et al.*, 1997). Perceptual learning during exposure to degraded images of objects before and after exposure to the corresponding undegraded versions was associated with higher activity in the left fusiform gyrus, while rCBF increase occurred in the right fusiform gyrus when images of faces were presented in the same way (Dolan *et al.*, 1997). Moreover, bilateral fusiform activity was higher during object than during spatial (Kohler *et al.*, 1998) or face (Kanwisher *et al.*, 1996) processing. The lingual gyrus was also activated in the left hemisphere (Kiyosawa *et al.*, 1996) or bilaterally (Kohler *et al.*, 1995) during object-related processing tasks (Kohler *et al.*, 1998), along with the right cuneus (Moscovitch *et al.*, 1995). With regard to postures or gestures, reports of activation studies directly concerning their visual processing are rare, and none of them have directly compared the visual processing of postures to that of other categorical entities, hence precluding direct comparison with our study. However, we found significant activation differences between the processing of these two categories of stimuli, with posture processing being associated with rCBF recruitment of more lateral occipitotemporal regions than object processing. How can we account for these differences?

It is noteworthy that we chose to use intransitive postures as stimuli, i.e., gestures that do not involve any handling of objects or interaction with them, as well as so-called intransitive objects, that is objects which do not, in principle, imply handling or manipulation by humans. Actually, one would argue that at least our meaningless objects (see an example in Fig. 1) could nonetheless be grasped, rotated, tilted, even used as a tool, hence that there remains some residual overlap between objects and gestures. While we do not conceal this overlap, it does not preclude the fact that these “intransitive” objects clearly differentiate from tools or typically handled objects. Our concern was to avoid an automatic, irrepressible, activation of the associated gesture if we had used usually manipulable objects. In that case, it is likely that less noticeable differences could have been found between the activation sites disclosed during posture and object processing. In keeping with this idea, it is worth mentioning an inferotemporal cortex activation during observation of hand grasping movements (Decety *et al.*, 1994) and a left middle temporal gyrus activation during observation of meaningful transitive versus meaningless (and consequently intransitive) gestures (Decety *et al.*, 1997). Hand grasping, as well as transitive actions, implies object manipulation; thus an inferotemporal involvement during gesture processing may be possibly due to the interaction with the processing of the objects related to these actions. Conversely, when activation related to the observation of meaningful transitive gestures was subtracted from that related to the observation of meaningless (and intransitive) gestures, a right inferotemporal locus was reported (Decety *et al.*, 1997), more lateral and only slightly inferior to our gesture-related activation site.

Cognitivofunctional Interpretation: Levels of Processing

Considering the level of differentiation between posture and object processing, what remains common through the four conditions in the present study is the visual stage, and what differentiates the neural activity elicited during this visual stage is the stimulus type. Regarding the visual stage, most theorists agree to consider that at least three main components may account for visual stimulus recognition: (1) extraction of the visual features of the image, (2) high-level perceptual analysis of the visual stimulus, i.e., construction of a representation of the stimulus configuration, and (3) matching of that perceptual description to a stored visual representation in memory.

Considering the present experiment within this framework, the extraction of the visual elementary features of a stimulus image is a basic process which is not considered to be dependent on categorical variations between visual stimuli. Nevertheless, one would argue that, in the gesture condition, the stimuli always entail the same woman with relatively subtle changes in arm and hand position, while in the object condition, the stimuli always entail a different visual stimulus. Hence, posture conditions would require a more detailed analysis of the visual input than the object conditions, but there would be more visual novelty in the object condition than in the posture condition, which is a confounding factor in our analyses. While the latter assertion may be true for meaningful stimuli processing, it is noteworthy that the meaningless objects are actually largely akin to each other, due to similar digitized deformations applied to the pictures. Therefore, there is no more visual novelty during meaningless-object processing than during posture processing, and the use of a mask in the conjunction analysis ensures that only those voxels significantly activated during both meaningless- and meaningful-object visual processing are highlighted, thus it is not likely that visual novelty may have significantly influenced the activations.

On the other hand, we observed differences between posture and object processing in the same locations both with meaningless and with meaningful stimuli. This precludes also the hypothesis that the common processing across tasks is a matching one with stored visual representations, because meaningless objects or postures are, by definition, not stored in memory. Instead, meaningless and meaningful postures and objects in this study have in common the fact that they are three-dimensional structured entities, hence we suggest that the differential rCBF activation observed between postures and objects is compatible with the processes involved in the construction of a representation of stimulus configuration, which is the second level of visual analysis in this general framework for visual object recognition.

Functional evidence supporting the hypothesis of the involvement of these brain areas in high-level perceptual analysis is provided by several studies which have associated high-level perceptual processing during visual object recognition with bilateral extrastriate activations in regions similar to those related to object processing in the present study. Indeed, a significant increase in rCBF was reported using PET in the inferior occipitotemporal (fusiform) junction, in response to stimuli supporting clear three-dimensional shape interpretations compared to stimuli containing similar low-level features but without obvious shape interpretation (Kanwisher *et al.*, 1997). This brain area was also thought to be involved in the bottom-up construction of shape descriptions from simple visual features, but not with processes associated with memory-matching, naming, or accessing semantic memory information (Menard *et al.*, 1996; Vandenberghe *et al.*, 1996). Similarly, line drawings of familiar or novel objects compared to visual noise patterns elicited an increased rCBF activity in a slightly more posterior region (Martin *et al.*, 1996). A bilateral fusiform activation was also associated with object decision for possible, but not for impossible, objects, suggesting a selective involvement of these structures in the representation of structurally coherent visual objects (Schacter *et al.*, 1995). Finally, using the fMRI technique, both photographs of familiar objects, famous faces, and unfamiliar three-dimensional abstract sculptures elicited strong responses in the extrastriate area LO, namely the lateral occipital complex (Malach *et al.*, 1995), located just posterior to area MT/V5. This lateral posterior area was proposed to be involved in an intermediate stage of visual object recognition, but others have criticized the reference condition (i.e., simple textures) and considered that LO would be involved not only in visual shape analysis, but also in feature extraction (Kanwisher *et al.*, 1997), explaining its large spatial extent.

In keeping with the idea of a particular sensitivity of the occipitotemporal junction area to structured stimuli, a hypothesis to account for location differences between posture and object processing in the present study may be that the perceptual salience of postures differs from that of other objects. Indeed, postures are the product of a limited set of combinations between different body components (head, torso, arm, forearm, hand, . . .), that is to say body elements with limited shape variance, whereas objects can exhibit virtually unlimited shapes. If the body itself is kept perceptually constant because minor changes occur in its elements, then attention may be easily attracted to the proper configuration of the posture, i.e., the combination between body elements. In this sense, upper-limb postures may be thought to have a clearer three-dimensional shape interpretation than other objects; consider for example Kanwisher's perspective (1997).

Visual Analysis of Gesture and Motion Area MT/V5

The peaks of significant lateral occipitotemporal junction activity, which we report when participants processed visually presented postures, encroaches on the visual area MT/V5, and one may be surprised that a related activity was observed near this site, found to be activated mainly when human subjects visually perceive motion (Dupont *et al.*, 1994; Tootell *et al.*, 1995; Watson *et al.*, 1993; Zeki *et al.*, 1991) or even imagine it (Goebel *et al.*, 1998), whereas in the present experiment only photographs of gestures were displayed, i.e., motionless gestures or postures. However, looking at the entire set of possible human gestures, one can see that most gestures are associated with movement. It would then be understandable that area MT/V5, dedicated to motion analysis, partly intervenes in the processing of this particular category. The fact that the postures presented here were static does not preclude this proposal, because they are only a subset of a more general class in which movement is one of the prototypical features. Alternatively, it may be that subjects simply "fill in" the gaps between sequentially displayed gestures as arm movements and therefore activate this motion-sensitive region. It is well known that actual motion is not necessary since stationary stimuli inducing illusory motion are sufficient to elicit activity in area MT/V5 (Goebel *et al.*, 1998; Tootell *et al.*, 1995; Zeki *et al.*, 1993). Hence, the involvement of the motion analysis area MT/V5 in the visual processing of postures could be due either to this motion reconstruction or to the fact that movement is one of the prototypical attributes of the specific category of gestures. In either case, area MT/V5 interacts with the lateral occipitotemporal junction because it is a necessary mediator to analyze one of the main attributes of gestures, namely motion, but it is likely that the construction of the tridimensional representation of the gesture remains mainly the prerogative of the lateral occipitotemporal junction.

Another hypothesis for MT/V5 involvement is suggested by Soderfeldt *et al.* (1997), who reported bilateral neural activity in area MT/V5 in bilingual volunteers with normal hearing during observation of sign language, compared to videotaped spoken language. The authors claimed that the activation in motion area MT/V5 was due to attention mechanisms which assign importance to the element of movements (signing hands) related to sign language rather than to the simple recording of passive movements, suggesting that area MT/V5 is not only involved in motion analysis, but also in higher order visual language processes. However, area MT/V5 was active in our study during the naming of meaningful postures as well as during the orientation-decision task performed on meaningless postures, the processing of the latter kind of stimuli being considered to bypass the semantic and gestural lexical processing stages in current cognitive models of limb praxis (Rothi *et al.*, 1997; Roy and Hall, 1992). Also, our experimental conditions were designed to ensure that most of the task-specific activity would be removed during statistical analyses, thus the cerebral activity elicited here is not likely to rely on naming or spatial-specific processes. It suggests that the importance assigned to the analysis of movements in area MT/V5 is not related to the semantic processing of the information in a "visual language" route, but rather favors the proposal that MT/V5 is involved in the construction of representations of visually presented posture shapes in association with other neural structures.

Lesion Studies and Visuoimitative Apraxia

Several accounts of limb apraxia have proposed that the "movement formulas" or "visuokinesthetic engrams" are stored in the dominant parietal lobe, especially in the supramarginal and angular gyrus (Heilman *et al.*, 1982). If the proposal that the attributes of an object are stored close to the regions of the cortex which mediate perception of those attributes (Martin *et al.*, 1995) can be generalized to gestures, it is plausible that the stored representations of gestures are close to the locus involved in the construction of the representation of the gesture's configuration. In this context, one should observe specific behavioral deficits in patients with focal brain damage: selective injury in the lateral occipitotemporal junction would be associated with difficulties related to the visual analysis of gestures, while deficits in the generation of meaningful gestures from memory would be related to parietal dysfunction.

Accordingly, Merians *et al.* (1997) have described patient SS, suffering from a left occipital and inferior temporal lesion which spared the parietal cortex. This patient exhibited significant deficits when imitating a movement that he performed close to normal when requested to execute it on verbal command, a pattern of limb praxis performance called visuoimitative apraxia. On the other hand, the assessment of limb praxis in the same study in other patients with selective left parietal damage disclosed pronounced deficits for gestures executed to verbal command, with a substantial improvement of the performance during imitation of the same gestures, in other words the reverse pattern. Patient SS' occipitotemporal infarct included the occipitotemporal junction, and his behavioral deficit in imitating gestures while being able to execute them on verbal command suggests a visuogestural analysis deficit, which is compatible with the hypothesis that the lateral occipitotemporal junction area is involved in the construction of a representation of the gesture's shape necessary for subsequent imitation. Conversely, SS' preservation of the movement formulas, thought to be stored in the parietal areas, enables the possibility of activating the motor areas involved in the execution of gestures during pantomime to verbal command, which is not the case for the parietal-damaged patients showing the reverse performance pattern.

At variance, another study of visuoimitative apraxia in two patients with selective damage to the inferior portion of the left angular gyrus shows a striking contrast between severely defective imitation of meaningless gestures and preserved performance on imitation and pantomime to verbal command of meaningful gestures (Goldenberg and Hagmann, 1997). The studies of Merians *et al.* (1997) and Goldenberg and Hagmann (1997) are difficult to compare because of large methodological and scoring differences, but the latter does not support the idea that gesture representations are stored per se in the left angular gyrus, because the injury in this structure did not impair pantomime on command. On the other hand, the occipitotemporal junction seems to be spared in the Goldenberg and Hagmann study, suggesting that a disconnection would account for the selective impairment of imitation of meaningless gestures. According to this hypothesis, after correct completion of the construction of the representation of posture configuration in the lateral occipitotemporal junction, a differentiation would occur in the subsequent stages of imitation: while imitation of meaningful gestures might be supported by an activation of the corresponding stored representation (as when performing on verbal command), imitation of meaningless gestures could depend only upon spatiotemporal transformation. Now, it has been shown that the angular gyrus is part of a more distributed dorsal system involved in gesture spatiotemporal layout transformation (Decety *et al.*, 1997), from visual to motor stages. Thus, an injury in an early part of the dorsal stream may severely impair imitation of meaningless gestures, while other structures may accurately mediate performance of imitation of meaningful gestures. Consistent with the Decety *et al.* (1997) point of view, Ochipa *et al.* (1997) brought to light an association between focal atrophy of the parietal lobes and parallel spatial deficits in limb gesture imagery and production, which might confirm the importance of these parietal structures in visuospatial transformation processes.

CONCLUSION

To summarize, the present study clearly disclosed stimuli-dependent effects during naming and orientation decision tasks performed on meaningless and meaningful intransitive, visually presented, upper-limb postures and objects. Conjunction analyses have shown that when meaningless and meaningful stimuli were combined irrespective of the task to be executed, a significant functional segregation between the processing of static intransitive gestures (postures) and the processing of intransitive tridimensional objects was demonstrated. Visual processing of objects elicited mainly a fusiform gyrus activity, while visual processing of postures activated the lateral occipitotemporal junction encroaching upon the motion analysis area MT/V5. Activity in the inferior occipitotemporal junction region was previously associated with the high-level perceptual processing of visual stimuli, an intermediate stage during visual object recognition. In this sense, our results suggest that the lateral occipitotemporal junction, working in association with area MT/V5, is particularly sensitive to human posture configurations and is involved at the same level in the construction of the visual representation of gesture, which can be used for subsequent recognition or imitation. Still, further studies will be necessary to replicate these findings and ensure the generalizability of these results. In this respect, ongoing experiments in our laboratory will hopefully extend our comprehension of the human cerebral network which supports the visual analysis of gestures.

ACKNOWLEDGMENTS

The authors acknowledge two anonymous reviewers for their insightful comments on a previous version of this article. Many thanks to Xavier Delbeuck for his kind help in the acquisition, selection, and transformation of digitized pictures; Caroline Scuvee for being our "lab model" for postures; Pierre Maquet; and Karl Friston for helpful statistical advice in the SPM list. This work was supported by the Interuniversity Pole of Attraction (IAP) Programme P4/22, Belgian State, Prime Minister's Office, Federal Office for Scientific, Technical, and Cultural Affairs, the Belgian National Fund for Scientific Research (FNRS), the Research Fund of ULg, and the Fondation Médicale Reine Elisabeth (FMRE). P.P. is supported by IAP.

REFERENCES

Barresi, J., and Moore, C. 1996. Intentional relations and social understanding. *Behav. Brain Sci.* **19:**107–154.

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

Craighero, L., Fadiga, L., Umilta, C. A., and Rizzolatti, G. 1996. Evidence for visuomotor priming effect. *NeuroReport* **8:**347–349.

Decety, J., Grezes, J., Costes, N., Perani, D., Jeannerod, M., Procyk, E., Grassi, F., and Fazio, F. 1997. Brain activity during observation of actions—Influence of action content and subject's strategy. *Brain* **120:**1763–1777.

Decety, J., Perani, D., Jeannerod, M., Bettinardi, V., Tadary, B., Woods, R., Mazziotta, J. C., and Fazio, F. 1994. Mapping motor representations with positron emission tomography. *Nature* **371:** 600–602.

Dolan, R. J., Fink, G. R., Rolls, E., Booth, M., Holmes, A., Frackowiak, R. S., and Friston, K. J. 1997. How the brain learns to see objects and faces in an impoverished context. *Nature* **389:**596–599.

Dupont, P., Orban, G. A., De Bruyn, B., Verbruggen, A., and Mortelmans, L. 1994. Many areas in the human brain respond to visual motion. *J. Neurophysiol.* **72:**1420–1424.

Faillenot, I., Toni, I., Decety, J., Gregoire, M. C., and Jeannerod, M. 1997. Visual pathways for object-oriented action and object recognition: Functional anatomy with PET. *Cereb. Cortex* **7:**77–85.

Friston, K. J., Worsley, K. J., Frackowiak, R. S., Mazziotta, J. C., and Evans, A. C. 1994. Assessing the significance of focal activations using their spatial extent. *Hum. Brain Mapp.* **1:**214–220.

Friston, K. J., Price, C. J., Fletcher, P., Moore, C., Frackowiak, R. S., and Dolan, R. J. 1996. The trouble with cognitive subtraction. *NeuroImage* **4:**97–104.

Goebel, R., Khorram-Sefat, D., Muckli, L., Hacker, H., and Singer, W. 1998. The constructive nature of vision: Direct evidence from functional magnetic resonance imaging studies of apparent motion and motion imagery. *Eur. J. Neurosci.* **10:**1563–1573.

Goldenberg, G., and Hagmann, S. 1997. The meaning of meaningless gestures: A study of visuo-imitative apraxia. *Neuropsychologia* **35:**333–341.

Grafton, S. T., Arbib, M. A., Fadiga, L., and Rizzolati, G. 1996. Localization of grasp representation in humans by positron emission tomography. *Exp. Brain Res.* **112:**103–111.

Heilman, K. M., Rothi, L. J., and Valenstein, E. 1982. Two forms of ideomotor apraxia. *Neurology* **32:**342–346.

Kanwisher, N., Chun, M. F., McDermott, M., and Ledden, P. J. 1996. Functional imaging of human visual recognition. *Cognit. Brain Res.* **5:**55–67.

Kanwisher, N., Woods, R. P., Iacoboni, M., and Mazziotta, J. C. 1997. A locus in human extrastriate cortex for visual shape analysis. *J. Cognit. Neurosci.* **9:**133–142.

Kiyosawa, M., Inoue, C., Kawasaki, T., Tokoro, T., Ishii, K., Ohyama, M., Senda, M., and Soma, Y. 1996. Functional neuroanatomy of visual object naming: A PET study. *Graefes Arch. Clin. Exp. Ophthalmol.* **234:**110–115.

Kohler, S., Kapur, S., Moscovitch, M., Winocur, G., and Houle, S. 1995. Dissociation of pathways for object and spatial vision: A PET study in humans. *NeuroReport* **6:**1865–1868.

Kohler, S., Moscovitch, M., Winocur, G., Houle, S., and McIntosh, A. R. 1998. Networks of domain-specific and general regions involved in episodic memory for spatial location and object identity. *Neuropsychologia* **36:**129–142.

Kosslyn, S. M., Thompson, W. L., Kim, I. J., and Alpert, N. 1995. Topographical representations of mental images in primary visual cortex. *Nature* **378:**496–498.

Malach, R., Reppas, J. B., Kwong, K. K., Jiang, H., Kennedy, W. A., Ledden, P. J., Brady, T. J., Rosen, B. R., and Tootell, R. B. 1995. Object-related activity revealed by functional magnetic resonance imaging in human occipital cortex. *Proc. Natl. Acad. Sci. USA* **92:**8135–8138.

Martin, A., Haxby, J. V., Lalonde, F. M., Wiggs, C. L., and Ungerleider, L. G. 1995. Discrete cortical regions associated with knowledge of color and knowledge of action. *Science* **270:**102–105.

Martin, A., Wiggs, C. L., Ungerleider, L. G., and Haxby, J. V. 1996. Neural correlates of category-specific knowledge. *Nature* **379:**649–652.

Menard, M. T., Kosslyn, S. M., Thompson, W. L., Alpert, N. M., and Rauch, S. L. 1996. Encoding words and pictures: A positron emission tomography study. *Neuropsychologia* **34:**185–194.

Merians, A. S., Clark, M., Poizner, H., Macauley, B., Rothi, L. G., and Heilman, K. M. 1997. Visual-imitative dissociation apraxia. *Neuropsychologia* **35:**1483–1490.

Moscovitch, M., Kapur, N., Kohler, S., and Houle, S. 1995. Distinct neural correlates of visual long-term memory for spatial location and object identity: A positron emission tomography study in humans. *Proc. Natl. Acad. Sci. USA* **92:**3721–3725.

Murata, A., Gallese, V., Kaseda, M., and Sakata, H. 1996. Parietal neurons related to memory-guided hand manipulation. *J. Neurophysiol.* **75:**2180–2186.

Neville, H. J., Bavelier, D., Corina, D., Rauschecker, J., Karni, A., Lalwani, A., Braun, A., Clark, V., Jezzard, P., and Turner, R. 1998. Cerebral organization for language in deaf and hearing subjects: Biological constraints and effects of experience. *Proc. Natl. Acad. Sci. USA* **95:**922–929.

Ochipa, C., Rapcsak, S. Z., Maher, L. M., Rothi, L. J., Bowers, D., and Heilman, K. M. 1997. Selective deficit of praxis imagery in ideomotor apraxia. *Neurology* **49:**474–480.

Price, C. J., and Friston, K. J. 1997. Cognitive conjunction: A new approach to brain activation experiments. *NeuroImage* **5:**261–270.

Price, C. J., Moore, C. J., and Friston, K. J. 1997. Subtractions, conjunctions, and interactions in experimental design of activation studies. *Hum. Brain Mapp.* **5:**264–272.

Rizzolatti, G., Fadiga, L., Matelli, M., Bettinardi, V., Paulesu, E., Perani, D., and Fazio, F. 1996. Localization of grasp representations in humans by PET. 1. Observations versus execution. *Exp. Brain Res.* **111:**246–252.

Rothi, L. J., Ochipa, C., and Heilman, K. M. 1997. A cognitive neuropsychological model of limb praxis. In *Apraxia: The Neuropsychology of Action* (Rothi and Heilman, Eds.), pp. 29–50. Psychology Press, Hove, UK.

Roy, E. A., and Hall, C. 1992. Limb apraxia: A process approach. In *Vision and Motor Control* (P. Luc and E. Digby, Eds.), pp. 261–282. North-Holland, Amsterdam.

Sakata, H., Taira, M., Murata, A., and Mine, S. 1995. Neural mechanisms of visual guidance of hand action in the parietal cortex of the monkey. *Cereb. Cortex* **5:**429–438.

Schacter, D. L., Reiman, E., Uecker, A., Polster, M. R., Yun, L. S., and Cooper, L. A. 1995. Brain regions associated with retrieval of structurally coherent visual information. *Nature* **376:**587–590.

Schwartz, R. L., Barrett, A. M., Crucian, G. P., and Heilman, K. M. 1998. Dissociation of gesture and object recognition. *Neurology* **50:**1186–1188.

Sergent, J., Ohta, S., and MacDonald, B. 1992. Functional neuroanatomy of face and object processing. A positron emission tomography study. *Brain* **115:**15–36.

Soderfeldt, B., Ingvar, M., Ronnberg, J., Eriksson, L., Serrander, M., and Stone-Elander, S. 1997. Signed and spoken language perception studied by positron emission tomography. *Neurology* **49:**82–87.

Talairach, J., and Tournoux, P. 1988. *Co-planar Stereotaxic Atlas of the Human Brain*. Thieme, New York.

Tootell, R. B., Reppas, J. B., Dale, A. M., Look, R. B., Sereno, M. I., Malach, R., Brady, T. J., and Rosen, B. R. 1995. Visual motion aftereffect in human cortical area MT revealed by functional magnetic resonance imaging. *Nature* **375:**139–141.

Vandenberghe, R., Price, C., Wise, R., Josephs, O., and Frackowiak, R. S. 1996. Functional anatomy of a common semantic system for words and pictures. *Nature* **383:**254–256.

Watson, J. D., Myers, R., Frackowiak, R. S., Hajnal, J. V., Woods, R. P., Mazziotta, J. C., Shipp, S., and Zeki, S. 1993. Area V5 of the human brain: Evidence from a combined study using positron emission tomography and magnetic resonance imaging. *Cereb. Cortex* **3:**79–94.

Woods, R. P., Iacobini, M., Grafton, S. T., and Mazziotta, J. C. 1996. Improved analysis of functional activation studies involving within-subject replications using a three-way ANOVA model. In *Quantification of Brain Function Using PET* (R. Myers, V. Cunningham, D. Bailey, and T. Jones, Eds.), pp. 353–358. Academic Press, San Diego.

Zeki, S., Watson, J. D., and Frackowiak, R. S. 1993. Going beyond the information given: The relation of illusory visual motion to brain activity. *Proc. R. Soc. London Biol.* **252:**215–222.

Zeki, S., Watson, J. D., Lueck, C. J., Friston, K., Kennard, C., and Frackowiak, R. 1991. A direct demonstration of functional specialization in human visual cortex. *J. Neurosci.* **11:**641–649.

Zelkowicz, B., Herbster, A. N., Nebes, R. D., Mintun, M. A., and Becker, J. T. 1998. An examination of regional cerebral blood flow during object naming tasks. *J. Int. Neuropsychol. Soc.* **4:**160–166.