Cortical Activation Related to Object Structure in Reach-to-Grasp Movements: An fMRI Study

Takeshi Sugio   Graduate School of Informatics, Kyoto University

To use a tool in an appropriate manner, we must correctly identify its functionality. In the present study, we used four types of objects (hammer, ruler, mug and bottle) that differed in either global property (elongation) or local feature ("handle") of object structure, and measured the cortical activation related to changes in object structure during imagined reach-to-grasp movements. As a result, we found that the peaks of activations in the posterior parietal cortex differed along with the change in object structure. Finally, we discussed the possible neural mechanisms of object structure processing related to both object recognition and object utilization.

Keywords: object structure, reach-to-grasp, intraparietal region, fusiform region

Introduction

It has been proposed that multiple levels of representation are necessary for successful object recognition. Humphreys et al. (1999) distinguished three processing stages that are involved in object naming. Knowledge about the object form is termed "structural description", and is most likely depicted from a canonical viewpoint specific to each object. In structural description theories, the image of an object is decomposed into regions corresponding to volumetric parts, and categorical spatial relations among those parts are represented explicitly (Biederman, 1987). As long as such structural descriptions are available from a given viewpoint, the structural identity of an object is recognized as constant over changes in its retinal image; therefore, object constancy can be achieved.

One of the main criticisms of structural description theories is related to the availability of such descriptions. Since local parts themselves suffer view-dependency (Hayward & Tarr, 1997), precise reconstructions of structural descriptions are by no means feasible. However, it seems quite probable that at least global representations of object structure are formed and employed in object recognition. Davidoff & Warrington (1999) reported that a patient with an apperceptive object agnosia who can recognize an object from canonical but not from noncanonical views utilizes an abstract object-centered representation for recognition from canonical views. By the term "global properties", we refer to the emergent properties derived from the interrelations among local parts. Presumably, such global representations are distinctive enough for semantic categorization; therefore, naming performance was preserved for such a patient. If such global representation can be used independent of the detailed structural descriptions, the parts of objects can be coded in relation to their global structure.

Recently, it has been illustrated that object images may automatically activate object motor representations (Tucker & Ellis, 1998). Learned actions to objects can be associated with the visual images of whole objects or local parts. In the latter case, local object parts are associated with a specific type of action ("handle" - "grasp"). However, in order to use an object appropriately, we need to know which part of the object corresponds to a handle and an effector. The output of such spatially localized visual processing can be a candidate for object motor representation.

Alternatively, the nature of object motor representation has also been investigated in the psychophysical measurement of kinematic properties of reach-to-grasp movements. Gentilucci (2002) demonstrated the involvement of a single object motor representation, in which the grasp kinematic implementation is based on the representation that codes all the object affordances. In one of his experiments, although the local parts to be grasped were identical between two objects, the maximum finger aperture was affected by the difference in global shape. Furthermore, the effect of object volume was stronger for familiar objects than simple geometrical solid (spheres). This may suggest that familiar objects automatically potentiate motor representation that is strongly associated with the types of interactions in everyday life.

Thus, the issue of object structure and its relation to action is still unresolved. Revealing the functional neuroanatomy between the representational levels of object structure and the stages in object-oriented actions might provide strong convergent evidence that object recognition and object-oriented action are closely related processes. In the present study, we explored the underlying neural mechanisms of reach-to-grasp movements using fMRI (functional magnetic resonance imaging). We employed four types of objects that differed in global elongation and local parts.

The task was to configure an appropriate hand shape for the visually presented object from four different viewpoints (the relative position of a handle with respect to the observer was different). Subjects were required to imagine themselves performing reach-to-grasp movements without actual movements. Since our major purpose was to investigate whether the difference in object structure recruits different activated regions in the brain, four different views were pooled among objects in the analysis in order to cancel out the effect of the graspability due to the egocentric location of handles.

Method

Participants

10 participants (20 - 29 years old, mean 24.5, 6 female, all right-handed) participated in this study. Written informed consent was obtained from each subject before testing.

Apparatus

A 1.5T MRI scanner (Magnex, Shimadzu) was used to acquire both T2 anatomical images and T2-weighted echo planar (EPI) images. The subjects were positioned in the scanner, with their heads immobilized. Stimuli were presented using a liquid crystal display (LCD) projected onto a custom viewing screen, and subjects viewed the screen through a mirror. The averaged size of each object image was a width of 8 ° and a height of 12 ° in visual angle. A total of 254 scans were acquired with a gradient echo EPI sequence (TR/TE 3000/49 ms, FA 90, FOV 256 mm, matrix 64 x 64, 30 axial slices, 5 mm slice thickness without gap). The first eight scans and the last six scans were discarded.

Stimuli

Four types of common objects (hammer, ruler, mug, bottle) were used (Figure. 1). These objects differed from each other in terms of global elongation and the existence of local parts ("handle"). Each object was presented in one of four different views (Figure. 2).

Design

Four experimental conditions with one null condition (only a black cross was presented) were conducted in random order for 36 trials each. This resulted in 180 trials in all. Inter-stimulus intervals were set between 3000 ms to 5000 ms (mean 4000 ms) randomly. The whole experiment took 720 s in all.

Procedure

Each trial began with a presentation of the image of an object on the screen (1 s). Subjects were told to imagine grasping the object with their right hands and configure their hand shape as if

Figure 1. Four objects used in the experiment: hammer, ruler, mug and bottle (the location of the handle is on the right, or compatible with a right-hand grasp).

Figure 2. Four possible views of an object presented in the experiment.

really grasping the object as soon as the object's image was presented. They were also told to keep in mind that grasp movements were performed to use the particular object, not simply to move it. After the image disappeared, a red fixation cross was presented on the screen for 1 s. Subjects were told to keep their hand shape until the next object image was presented or the color of the cross changed to black. When the cross changed to black, subjects brought back their right hand to the start position and spread their palm normally.

Results

Image preprocessing and statistical analysis were performed using SPM99 (Wellcome Department of Cognitive Neurology, http://www.fil.ion.ucl.ac.uk/spm). All functional images were resampled using the middle slice as a reference, realigned to adjust for motion-related artifacts, spatially normalized with the MNI template based on the spatial transformations derived from the coregistered anatomical images, and spatially smoothed with a Gaussian kernel of 8 x 8 x 10 mm full width at half-maximum (FWHM). Condition effects at each voxel were estimated according to the general linear model. Before the estimation, data were high-pass filtered, using a default cut-off period. Data analysis employed a two-stage procedure, implementing a mixed-effects analysis. In the first stage, the time series of each voxel was correlated with a time series generated by convolving a delta function at each stimulus onset with a canonical hemodynamic response function. Session-specific parameter estimates were calculated for each voxel, and images of contrasts of these parameter estimates were calculated for each subject. These contrast images were then entered into one sample t-test, treating subjects as a random variable. All the reported activations were thresholded at $p = 0.001$ (uncorrected for multiple comparisons) with a cluster extent threshold of 10 contiguous voxels. The foci of peak activation in terms of Brodmann areas (BA) were reported according to the Talairach brain system (Talairach & Tournoux, 1988).

Two types of contrasts were computed to reveal the activation related to global and local structural properties. First, the contrasts with the baseline condition (fixation only) were computed for each experimental condition (handle+ (H+), handle- (H-), elongation+ (E+), elongation- (E-)). For example, for the H+ condition, the set of coefficients was (hammer, ruler, mug, bottle) = (1, 0, 1, 0). The global activation patterns were quite similar among the four conditions. The typical pattern of activation is illustrated in Figure 3 (H+ condition). Next, in order to isolate effects specifically related to handle or global elongation, respectively, we performed two types of direct comparisons: the first was between H+ (hammer and mug) and H- (ruler and bottle), and the other was between E+ (hammer and ruler) and E- (mug and bottle). Further, to distinguish the increase in activation from the decrease in deactivation, only regions that were also observed in the activated regions contrasted with the baseline condition were adopted as ROIs (region of interest).

First, for the H+ > H- comparison, there were three regions of interest: the left middle occipitotemporal region including

middle occipital gyrus and middle temporal gyrus, the left anterior intraparietal region and right posterior intraparietal region. Second, for the H- > H+ comparison, there were four regions: the left mid-fusiform area, bilateral occipital regions (lingual gyrus, cuneus, inferior occipital gyrus and middle occipital gyrus) and the right middle frontal region that corresponds to the frontal eye field. Finally, for the E+ > E- comparison, the left parietooccipital region, the left medial parietal region including precuneus, the right posterior parietal region, the right superior parietal region and the left occipital region extending to the bilateral areas. Coordinates of the peaks of the activated regions are shown in Table 1. No noticeable increase in cortical activation was found for the E- > E+ comparison.

Discussion

Only one region was in common among H+ > H- and E+ > E- comparisons: the right posterior intraparietal region. This region is referred to as cIPS (caudal intraparietal sulcus) area in the
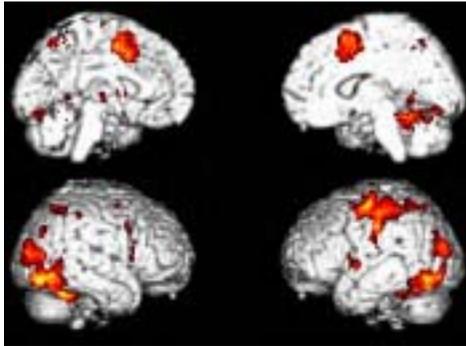
Figure 3. The activation in the handle+ condition mapped onto the cortical surface.

Table 1. Coordinates of the peak anatomical regions for the direct comparisons (L, left; R, right)

	Coordinates (mm)		
Peak anatomical region	x	y	z
contrast: Handle+ > Handle-			
L middle occipital gyrus	-36	-83	17
L superior parietal lobule	-28	-46	56
R precuneus	20	-63	51
contrast: Handle- > Handle+			
L fusiform gyrus	-38	-48	-20
L inferior occipital gyrus	-28	-88	-4
R lingual gyrus	22	-88	-2
R middle frontal gyrus	50	4	40
contrast: Elongation+ > Elongation-			
L cuneus	-24	-80	30
L precuneus	-12	-44	56
R precuneus	20	-67	49
R precuneus	26	-54	52
L lingual gyrus	-10	-72	-5

neurophysiological literature. In cIPS area, various types of neurons sensitive to the three-dimensional properties of objects such as axis, surface orientation and volumetric structure are found in the monkey (Sakata et al., 1997). Our present results showed that both global and local shape properties of an object are processed in this region, and the degree of the activation is related to the task relevancy of the properties.

On the other hand, it has been demonstrated that the anterior intraparietal region (AIP) is involved in visually guided hand movements that are relevant to grasping (Sakata et al., 1997). It has been known that one of the major sources of visual input to AIP area is cIPS area. Along with the input from the inferotemporal region, AIP area receives three-dimensional shape properties of an object and its semantic knowledge independently. However, the increase in the activation of AIP area was observed only for objects with handles. This means that although both global and local shape properties are processed in the cIPS area, only properties that are directly related to grasping, that is, the shape of a graspable part, has an effect on AIP area.

For elongated objects, three parietal areas other than the right cIPS area showed increased activation. First, it has been shown that the left parietooccipital region is a portion of a neural circuit mediating visually guided reaching (Chapman et al., 2002). Moreover, precuneus activation along the medial parietal region may partly correspond to the region known as the parietal reach region (Astafiev et al., 2003). In the monkey, the parietooccipital region is activated in the task that requires reaching movements for nonfoveated targets. This suggests that potentially graspable locations of an objects is processed in the parietooccipital region together with the right cIPS area, and its output is conveyed to the medial parietal region, where movement intention for reaching is represented. Consequently, the increased parietal activations for elongated objects comprise the internal reach simulation mechanism (Johnson et al., 2002).

The differential increased activations in the left parietal regions related to handle and global elongation were clear-cut. Local graspable features are related to visually guided grasp regions, whereas global elongated structure is related to visually guided reach regions. It has already been suggested that the arm transportation and hand manipulation components of reach-to-grasp movements involve separate mechanisms (Jeannerod et al., 1995). However, as far as we know, the present study is the first report showing that different object properties affect the activation strength of the related regions.

Occipitotemporal activations were also different for three comparisons. First, the increased activations in the E+ > E- comparison were restricted to the bilateral lingual gyri and cuneus. On the other hand, activated regions were located on the lateral side in the H- > H+ comparison. Specifically, activations of the middle occipital gyrus were more prominent for this comparison. Middle occipital gyri are known as the portion of the lateral occipital complex (LOC), which is sensitive to the perception of three-dimensional volumes (Moore & Engel, 2001). However, the function of LOC is different from that of cIPS area in that the former is not related to the object-oriented movements. The left mid-fusiform region also showed increased

activation for the H- > H+ comparison. This region is sensitive to the degrees of overlap in global structure of objects (Joseph & Gathers, 2003), which implies its importance in object processing postulated in computational accounts of object recognition (Biederman, 1987). When the object structure does not provide sufficiently salient information with respect to task demands (reach-to-grasp), it is possible that the ventral-temporal region is involved in the retrieval of object-specific grasping information from semantic knowledge.

The peak location of the increased activation of the left occipitotemporal area in the H+ > H- comparison was at a more superior region. Senior et al. (2000) found that the posterior region of V5 is related to the perception of scenes that imply motion, such as an object dropping off a shelf. This finding is in accordance with the present results in that grasping movements can easily induce the motor imagery of object usage, whereas reaching movements themselves are not directly related to object motion.

Recently, Binkofski et al. (1999) demonstrated that a neural circuit specifically involved in object manipulation exists in humans. The ventral premotor region and anterior intraparietal region form the circuit that corresponds to the F5-AIP circuit found in the monkey (Rizzolatti & Luppino, 2001). In the present study, both regions were activated in all four objects. However, only AIP area was sensitive to the change in the object properties (e.g. the existence of local graspable features). This finding may suggest that the visual function served in the AIP area is to extract three-dimensional features suitable for the hand-object interaction. Conversely, the global elongation of an object is also important for the object-oriented action, especially for reaching movements. These two different levels of object properties influence the activity of the neural mechanism involved in reach-to-grasp action in a selective manner.

Conclusion

Cortical activations related to object structure in reach-to-grasp movements were observed in both the ventral and the dorsal stream. Local features ("handle") directly linked to the specific type of action (i.e. grasping) are useful for the appropriate guidance of reach-to-grasp movements, whereas global properties of object structure (i.e. elongation) are employed to determine where and how the object can be grasped (grip selection). Differential activities in the ventral stream might suggest that object recognition is an essential component in the reach-to-grasp cortical network, and structural descriptions might serve as the intermediate representation between recognition and action.

References

Astafiev, S. V., Shulman, G. L., Stanley, C. M., Snyder, A. Z., Van Essen, D. C., & Corbetta, M. 2003 Functional organization of human intraparietal and frontal cortex for attending, looking, and pointing. *Journal of Neuroscience*, 23, 4689-4699.

Biederman, I. 1987 Recognition-by-components: A theory of human image understanding. *Psychological Review*, 94, 115-145.

Binkofski, F., Buccino, G., Posse, S., Seitz, R. J., Rizzolatti, G., Freund, H. J. 1999 A frontal-parietal circuit for object manipulation in man: evidence from an fMRI-study. *European Journal of Neuroscience*, 11, 3276-3286.

Chapman, H., Gavrilescu, M., Wang, H., Kean, M., Egan, G., & Castiello, U. 2002 Posterior parietal cortex control of reach-to-grasp movements in humans. *European Journal of Neuroscience*, 15, 2037-2042.

Davidoff, J., & Warrington, E. K. 1999 The bare bones of object recognition: implications from a case of object recognition impairment. *Neuropsychologia*, 37, 279-292.

Gentilucci, M. 2002 Object motor representation and reaching-grasping control. *Neuropsychologia*, 40, 1139-1153.

Hayward, W. G., & Tarr, M. J. 1997 Testing conditions for viewpoint invariance in object recognition. *Journal of Experimental Psychology: Human Perception and Performance*, 23, 1511-1521.

Humphreys, G. W., Price, C. J., & Riddoch, M. J. 1999 From objects to names: A cognitive neuroscience approach. *Psychological Research*, 62, 118-130.

Jeannerod, M., Arbib, M. A., Rizzolatti, G., & Sakata, H. 1995 Grasping objects: the cortical mechanisms of visuomotor transformation. *Trends in Neurosciences*, 18, 314-320.

Johnson, S. H., Rotte, M., Grafton, S. T., Hinrichs, H., Gazzaniga, M. S., & Heinze, H-J. 2002 Selective activation of a parietofrontal circuit during implicitly imagined prehension. *NeuroImage*, 17, 1693-1704.

Joseph, J. E., & Gathers, A. D. 2003 Effects of structural similarity on neural substrates for object recognition. *Cognitive, Affective, & Behavioral Neuroscience*, 3, 1-16.

Moore, C., & Engel, S. A. 2001 Neural response to perception of volume in the lateral occipital complex. *Neuron*, 29, 277-286.

Rizzolatti, G., & Luppino, G. 2001 The cortical motor system. *Neuron*, 31, 889-901.

Sakata, H., Taira, M., Kusunoki, M., Tanaka, Y. 1997 The parietal association cortex in depth perception and visual control of the hand action. *Trends in Neurosciences*, 20, 350-357.

Senior, C., Barnes, J., GIampietro, V., Simmons, A., Bullmore, E. T., Brammer, M., & David, A. S. 2000 The functional neuroanatomy of implicit-motion perception or 'representational momentum'. *Current Biology*, 10, 16-22.

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

Tucker, M., & Ellis, R. 1998 On the relations between seen objects and components of potential actions. *Journal of Experimental Psychology: Human Perception and Performance*, 24, 830-846.