Actions or Hand-Object Interactions? Human Inferior Frontal Cortex and Action Observation

Scott H. Johnson-Frey,1,* Farah R. Maloof,1 Roger Newman-Norlund,1 Chloe Farrer,1 Souheil Inati,2 and Scott T. Grafton1
1Center for Cognitive Neuroscience Dartmouth College Hanover, New Hampshire 03755
2Center for Neural Science and Department of Psychology New York University New York, New York 10003

Summary

Cells in macaque ventral premotor cortex (area F5c) respond to observation or production of specific hand-object interactions. Studies in humans associate the left inferior frontal gyrus, including putative F5 homolog pars opercularis, with observing hand actions. Are these responses related to the realized goal of a prehensile action or to the observation of dynamic hand movements? Rapid, event-related fMRI was used to address this question. Subjects watched static pictures of the same objects being grasped or touched while performing a 1-back orienting task. In all 17 subjects, bilateral inferior frontal cortex was differentially activated in response to realized goals of observed prehensile actions. Bilaterally, precentral gyrus was most frequently activated (82%) followed by pars triangularis (73%) and pars opercularis (65%).

Introduction

Single-unit electrophysiological studies identify cells within macaque ventral premotor cortex (area F5c) that code the goals of specific prehensile actions rather than the movements of which they are composed. For example, on the basis of their response preferences, these units can be categorized into those that represent holding, grasping, or tearing objects. Further, these units’ responses are context dependent: if the same hand movements are made as part of a different action, e.g., grooming instead of feeding, responses are weak or absent (Rizzolatti et al., 1988). This observation has led to the hypothesis that area F5c contains a “vocabulary” of hand actions (Rizzolatti et al., 1988), in which the goals of hand-object interactions are represented explicitly. A subclass of F5c units discharge not only when the monkey produces an action but also when it observes the experimenter perform a comparable behavior. Similar responses would be expected to actions performed by conspecifics; however, this is difficult to test in a controlled fashion because these cells only appear to respond to live actors. Like other F5c cells, these so-called “mirror neurons” also respond best to a specific type of prehensile action (Rizzolatti and Luppino, 2001). Critically, mirror neurons’ responses depend on the animal observing an interaction between the effector (hand and/or mouth) and the target object (Gallese et al., 1996). For instance, a mirror neuron that responds when the animal observes a hand grasping an object would respond weakly or not at all if the hand merely touched the object. This suggests that mirror neurons, like other cells in area F5c, are not responding to the observation of dynamic hand movements per se but rather to the perception of specific hand-object interactions, or goals. Mirror neurons have generated considerable interest because they provide a possible mechanism for matching observed and executed actions; a process that is central to the claim that organisms understand others’ actions via activation of their own internal motor representations (Konorski, 1967). Attempts to identify a homologous mirror system in humans have focused on the posterior portion of the inferior frontal gyrus known as the pars opercularis. Similar to area F5, this site has a distinctive agranular cytoarchitecture that may indicate that these two regions are structural homologs (Petrides, 1994; Preuss et al., 1996). In macaques (Fogassi et al., 2001) and humans (Binkofski et al., 1999; Ehrsson et al., 2000, 2001), the inferior frontal gyri in both hemispheres are associated with the production of prehensile actions. Evidence for activation of this region during observation of hand actions has therefore been sought in support of the existence of a mirror system in humans. Several functional neuroimaging investigations have shown activation within the inferior frontal gyrus when humans observe dynamic hand movements. Given the substantial individual variability in the morphology of this region (Amunts et al., 1999; Foundas et al., 2001; Tomaiuolo et al., 1999a), caution must be exercised in interpreting these claims of localization with respect to specific Brodmann areas. Early PET studies reported activation of BA45 during observation of grasping (Grafton et al., 1996; Rizzolatti et al., 1996) and meaningful hand actions (Grezes et al., 1998). More recently, fMRI has revealed activation in left BA44 during observation of finger movements (iacoboni et al., 1999) and grasping actions (Buccino et al., 2001). A recent MEG study also reported a source localized in left BA44 during the observation of grasping (Nishitani and Hari, 2000). Observation of static hand postures does not appear to elicit responses in inferior frontal cortex (Hermsdörfer et al., 2001).

What remains unclear from these studies is whether, like mirror neurons, inferior frontal sites are responding to specific hand-object interactions or to the dynamics of hand actions irrespective of the behavioral goal. Put differently, is exposure to the realized goal of a prehensile action a sufficient condition for these responses, or is observation of dynamic hand movements necessary? We reasoned that if responses are coding hand-object interactions, then observation of static pictures of hands engaging objects in a prehensile manner should activate these regions. As illustrated in Figure 1, to evaluate this hypothesis we created static images that were identical in all ways except one: how the hand was contacting the object. Because F5c neurons in the monkey code...
Table 1. Coordinates of Group Mean Activations within the ROI Activated When Viewing Objects Being Grasped versus Touched

<table>
<thead>
<tr>
<th>t Value</th>
<th>p (unc)</th>
<th>x</th>
<th>y</th>
<th>z</th>
<th>Locus</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.82</td>
<td>0.001</td>
<td>-53</td>
<td>0</td>
<td>22</td>
<td>precentral gyrus</td>
</tr>
<tr>
<td>3.57</td>
<td>0.001</td>
<td>-47</td>
<td>8</td>
<td>6</td>
<td>pars opercularis</td>
</tr>
<tr>
<td>3.09</td>
<td>0.003</td>
<td>-38</td>
<td>8</td>
<td>16</td>
<td>pars opercularis</td>
</tr>
<tr>
<td>2.82</td>
<td>0.006</td>
<td>-48</td>
<td>34</td>
<td>8</td>
<td>pars triangularis</td>
</tr>
<tr>
<td>6.42</td>
<td>&lt; 0.0001</td>
<td>53</td>
<td>-1</td>
<td>15</td>
<td>precentral gyrus</td>
</tr>
<tr>
<td>4.67</td>
<td>&lt; 0.0001</td>
<td>34</td>
<td>36</td>
<td>4</td>
<td>pars triangularis</td>
</tr>
<tr>
<td>3.79</td>
<td>0.001</td>
<td>52</td>
<td>36</td>
<td>6</td>
<td>pars triangularis</td>
</tr>
<tr>
<td>3.42</td>
<td>0.002</td>
<td>36</td>
<td>28</td>
<td>8</td>
<td>pars triangularis</td>
</tr>
<tr>
<td>3.38</td>
<td>0.002</td>
<td>54</td>
<td>4</td>
<td>32</td>
<td>pars opercularis</td>
</tr>
</tbody>
</table>

Coordinates indicate locations of local maxima in standardized space (Talairach, 1988). Clusters contain a minimum of five voxels, and maxima are separated by at least 4 mm.

2B illustrates significant mean activations within inferior frontal cortex when subjects viewed objects being grasped versus touched. Table 1 summarizes the location of local maxima within these frontal areas. In the left hemisphere, the activation was observed in the inferior extent of the precentral gyrus extending ventrally into frontal operculum along the bank of the lateral sulcus and rostrally into the inferior frontal gyrus. Activation in left hemisphere inferior frontal gyrus is consistent with previous PET (Grafton et al., 1996; Rizzolatti et al., 1996), fMRI (Buccino et al., 2001; Iacoboni et al., 1999), and MEG (Nishitani and Hari, 2000) studies of the observation of dynamic hand actions. It has been suggested that these left-lateralized effects may reflect subvocalization (Grezes and Decety, 2001; Heyes, 2001). Because the same objects were observed in both grasp and touch conditions, this alternative cannot account for the present left hemisphere effects. In contrast to these earlier investigations, we also observed activations within infe-

**Results**

**Grasping versus Touching**

Figure 2A illustrates our inferior frontal region of interest (ROI), the rationale for which is detailed in the Experimental Procedures section below. As predicted, Figure 2B shows regions of interest and group results. The central sulcus is drawn in red.
rior frontal regions of the right hemisphere. Similar to the left, these included the inferior extent of the precentral gyrus, extending into the operculum along the bank of the lateral sulcus and rostrally into the inferior frontal gyrus (Figure 2B).

The inverse contrast of touching versus grasping revealed no significantly activated voxels within the inferior frontal cortex ROIs of either hemisphere at the same level of significance (i.e., $p < 0.01$, uncorrected) or at a more liberal threshold ($p < 0.05$, uncorrected).

As illustrated in Figure 3, the peristimulus hemodynamic changes in left (panel A) and right (panel B) inferior frontal areas also indicate differential responses to grasp and touch conditions. In both left and right inferior frontal regions, there is an increase in signal associated with the grasp condition and a decrease associated with touch.

Given the substantial individual variability in cortical topography in inferior frontal cortex, strong claims about localization on the basis of group averages must be interpreted with caution. In order to gain more precise knowledge of specific anatomical areas showing selectivity for hand-object interactions within inferior frontal cortex, we analyzed loci of activation peaks resulting from the contrast of the grasp versus touch conditions within the ROI in all 17 individuals. Figure 4 illustrates that all 17 individuals showed effects within inferior frontal cortex, and in 16 subjects activations were observed in both hemispheres. There was, however, individual variability in the locus of these activations. The overall pattern for the left and right hemispheres was highly symmetrical, with frequency of activation greatest in precentral gyrus (left, 82.35%; right, 82.35%) followed by pars triangularis (left, 70.5%; right, 76.47%), and finally the putative homolog of macaque F5, pars opercularis (left, 64.7%; right, 64.7%).

Having established the involvement of bilateral inferior frontal cortex in coding goal-specific representations of hand-object interactions, we next sought to determine whether these responses were modulated by familiarity of the object and/or the action.

Effect of Familiarity of the Object and Action

The relative magnitude of the independent variables in our statistical model ($\beta$ values) was compared at the locations of peak signal intensity for local maxima within the left ($–53, 0, 22$) and right ($53, –1, 15$) ROIs using repeated-measures ANOVA. As expected on the basis of the results described above, the type of hand-object interaction (grasp versus touch) was critical to responses in the left hemisphere, $F(16) = 25.4$, $MSE = 0.0009$, $p < 0.0001$. By contrast, neither the familiarity of the object (tool versus unfamiliar shape) nor the type of grasp used for tools (functional versus nonfunctional) made any difference, $p > 0.34$ in both cases.

Similarly, responses in the right inferior frontal region depended on the type of hand-object interaction, $F(16) = 59.8$, $MSE = 0.00001$, $p < 0.00001$. Again, neither
familiarity of the object nor the type of grasp used with familiar tools mattered, \( p > 0.30 \) in both cases.

**Discussion**

Results of our event-related fMRI investigation demonstrate that left and right precentral and inferior frontal (pars triangularis and pars opercularis) gyri are selectively activated when subjects passively observe the realized goals of hand-object interactions. Critically, we demonstrate that these responses occur in the absence of observing the dynamic actions involved in achieving these goals. Put differently, exposure to the realized goal of a prehensile action is a sufficient condition for these responses, while observation of active movements of the hand is unnecessary. In this sense, these areas behave similarly to mirror neurons in the macaque, which also appear to selectively code specific hand-object interactions (Gallese et al., 1996). However, to our knowledge, there are no published findings testing whether mirror cells respond selectively to hand-object interactions captured in still pictures. Importantly, our data show that although pars opercularis may be a structural homolog of macaque F5, selective coding of the goals of realized actions takes place throughout human inferior frontal cortex. Precise areas involved seem to vary across individuals, with left and right precentral gyri occurring most frequently, followed by pars triangularis and finally pars opercularis (Figure 4). Further, we demonstrate that these responses are not influenced by the familiarity of the object or the type of grasp with which familiar tools are engaged.

This work represents an important departure from previous functional neuroimaging investigations of action observation in several key ways. First, because static pictures were used in all conditions, we rule out the possibility that inferior frontal responses depend on the observation of dynamic movements irrespective of the behavioral goal; an alternative interpretation of all previous studies that have reported inferior frontal activation when comparing moving displays with static control stimuli (Buccino et al., 2001; Grafton et al., 1996; Grezes et al., 1998; Iacoboni et al., 1999; Rizzolatti et al., 1996). Second, by creating stimuli that are identical except for the manner in which the hand was interacting with the objects, we eliminated the possibility that inferior frontal responses are related to other stimulus differences between conditions, such as the presence or absence of a goal object (Buccino et al., 2001). Third, because the same objects were viewed in both the grasp and touch conditions, our findings are unlikely to reflect subvocalization. Because the role of left inferior frontal gyrus in overt and covert speech production is long-established (Dejerine, 1914), this is a major criticism of previous
Hand-Object Interactions
1057

studies reporting involvement of these lateralized sites in action observation (Greze and Decety, 2001; Heyes, 2001). Finally, our use of a rapid, event-related paradigm eliminates confounds associated with cognitive set inherent in previous investigations of action observation where stimuli from different categories were blocked together. Along with our use of a simple 1-back orienting task, this design minimizes the influence of higher-level cognitive processing on our findings.

In spite of these substantial differences, these results are partially consistent with several previous action observation studies in reporting left inferior frontal gyrus involvement: BA44 (Buccino et al., 2001; Iacoboni et al., 1999; Nishitani and Hari, 2000) or BA45 (Grafton et al., 1996; Grezes et al., 1998; Rizzolatti et al., 1996).

In conclusion, it is important to note that human inferior frontal areas appear to play a variety of roles in processing action-relevant perceptual information. Observing and naming pictures of tools relative to nonmanipulable familiar objects (houses) is associated with activity in left ventral precentral gyrus (Chao and Martin, 2000). A recent study by Kellenbach et al. indicates that inferior frontal involvement is task independent and therefore may reflect automatic activation of action representations by the perception of tools and to a lesser degree by nonmanipulable objects (Kellenbach et al., 2003). Greater activation in precentral sulcus during observation of moving tools and human forms relative to radial gratings has also been reported recently (Beauchamp et al., 2002), and responses in human inferior frontal cortex can be driven by abstract, dynamic visual properties (Schubotz and Yves von Cramon, 2002). These findings together suggest that inferior frontal areas are not only important for manual and oral movements but also represent action-relevant perceptual properties at a variety of different levels. An important goal for future work is to explore the possible role of these representations in deciphering the behaviors of conspecifics and choosing appropriate behavioral responses.

Experimental Procedures

Subjects

Eighteen adult volunteers (eight females, nine males) participated in a 1 hr testing session approved by the Committee for the Protection of Human Subjects at Dartmouth College. One subject was eliminated due to excessive head motion. None of the participants had a history of psychiatric or neurological disease, and all provided written informed consent. All subjects were identified as right hand dominant according to the Edinburgh Handedness Inventory (Oldfield, 1971).

Stimuli

As illustrated in Figure 1, stimuli consisted of 100 black and white digital photographs of 20 familiar tools and 20 unfamiliar 3D shapes. Each object was photographed being grasped and touched by a right hand. Further, familiar tools were photographed being grasped both in a manner consistent and inconsistent with their most common usage.

In this rapid, event-related study, each subject completed five experimental runs in counterbalanced order. Each run lasted 6 min and 22 s, began with 10 s of fixation, and ended with 20 s of fixation. Within each run, 100 pictorial stimuli were intermixed with 40 null events consisting of a blank screen. As illustrated in Figure 1, a picture or null event appeared for 1500 ms followed by a 1000 ms blank screen. Excepting the initial and final fixation periods, a pictorial stimulus appeared every 3.5 s on average during each functional run. A fixation cross was present in the center of the screen throughout the entire run. The counterbalancing of stimuli in each run was optimized so that there was an equal probability of stimuli from one condition following stimuli from any of the other conditions. Following sequence optimization, two pictorial stimuli from each stimulus category (10% of stimuli per run) were randomly chosen and duplicated within each run. Subjects were instructed to press a button with their right index finger whenever the same picture appeared twice in a row.

Magnetic Resonance Imaging

Imaging was performed with a General Electric Horizon whole body 1.5T MRI scanner using a standard birdcage head coil. Head movements were minimized by the use of a foam pillow and padding. Prior to each functional run, four images were acquired and discarded to allow for longitudinal magnetization to approach equilibrium. Within each functional run, an ultrafast echo planar gradient echo imaging sequence sensitive to blood oxygenation level-dependent (BOLD) contrast was used to acquire 25 slices per TR (4.5 mm thickness, 1 mm gap, in-plane resolution [3.125 × 3.125 mm]). The following parameters were used: TR, 2500 ms; TE, 35 ms; flip angle, 90°. A high-resolution, T1-weighted, axial fast spin echo sequence was used to acquire 25 contiguous slices (4.5 mm slice thickness with 1.0 mm gap) coplanar to BOLD images: TE, Min full; TR, 650 ms; Echo Train, 2; FOV, 24 cm. High-resolution (0.94 × 0.94 × 1.2 mm), whole-brain, T1-weighted structural images were also acquired using a standard GE SPGR 3D sequence.

Regions of Interest

We define human inferior frontal cortex as including the precentral gyrus and inferior frontal gyrus (pars opercularis and pars triangularis) in both cerebral hemispheres. We operationalize this region as including those frontal areas falling within a 30 mm radius sphere centered at standardized coordinates x = ± 48, y = −18, z = 8 (Figure 2A). These coordinates were chosen on the basis of a probabilistic map such that the centroid of each sphere had the highest possible likelihood (50%–75%) of falling in pars opercularis as defined morphologically in 108 healthy adults (Tomaiuolo et al., 1999b). The radius was selected so as to include precentral gyrus, pars triangularis, and pars opercularis in each of our 17 subjects (Figure 2B).

Image Processing

Structural and functional images were preprocessed and analyzed using SPM99 (http://www.fil.ion.ucl.ac.uk/spm). Functional data for each individual subject were corrected for differences in time of slice acquisition and head motion. Functional and structural images were coregistered and transformed into a standardized stereotactic space. This resulted in 25 axial slices of isotropic 3.125 mm³ voxels. Data were smoothed with an 8 mm FWHM isotropic Gaussian kernel. Within the ROI, fixed-effects analyses were performed on individual subjects’ data with session as the random variable. Results of these analyses were then submitted to a second-level, random-effects analysis, with subjects as the random variable. Statistical activation maps were constructed based on differences between trial types using a t statistic. Clusters consisting of at least five voxels, separated by a minimum of 4 mm, and having t values equal to or greater than 2.57 (p < 0.01, uncorrected for multiple comparisons), were considered statistically significant. Results were converted to the standardized coordinate system used by the Talairach Atlas (Talairach, 1988) using a nonlinear transformation (http://www.mrc-cbu.cam.ac.uk/Imaging/mnispace.html). Surface renderings were created using MRICRO software (http://www.cla.sc.edu/psyc/faculty/orden/render.html).

Acknowledgments

Scott H. Johnson-Frey was formerly known as Scott H. Johnson. This paper is dedicated to the memory of Georgianne Frey and was supported by grants from NIH (K01 MH00222-01) and the James S. McDonnell Foundation to the first author. The authors wish to acknowledge Dr. Petr Janata and Jed Dobson for technical assis-
tance; and three anonymous reviewers for their constructive comments on an earlier draft.

Received: December 19, 2002
Revised: June 26, 2003
Accepted: August 1, 2003
Published: September 10, 2003

References


