1 The Mirror Neuron System and Imitation

Giacomo Rizzolatti

1.1 Introduction

"Every one knows what attention is." This famous sentence by William James (1890, p. 403) appears to be appropriate also for imitation. Everyone knows what imitation is. Yet, as soon as imitation is more closely examined, this concept loses its simplicity; it appears to include different behaviors, some learned, some innate.

In this chapter, unless specified otherwise, I adopt Thorndike's definition of imitation. Imitation is learning to do an act from seeing it done (Thorndike, 1898). This definition includes two basic ideas: (1) imitation implies learning; and (2) during imitation the observer transforms an observed action into an executed action that is similar or even identical to the observed one.

How does imitation occur? The response to this question is obviously not easy. In the first place, why should an individual copy an action made by another individual? In everyday life, copying an action is typically useless and frequently dangerous. If an animal observing a conspecific eating some food imitates its movements, it will never get food. It will only aimlessly move its mouth. Imitation implies an understanding of what another individual is doing as well as the capacity to use this knowledge only in particular conditions.

Second, what information must the observer extract from an acting conspecific in order to imitate his behavior? Is it sufficient to understand the goal of the observed actions or must its details also be coded? Finally, there is the so-called "translation" problem. Sensory and motor systems are classically considered to be separate systems. Thus, how can the description of a visual event become a muscle excitation that faithfully replicates the observed event?

In this chapter, the following theoretical positions are defended:

1. Imitation is composed of two strictly related cognitive phenomena. The first is the capacity to make sense of others' actions. The second is the capacity, once the action is understood, to replicate it. According to the task and external contingencies, the imitated action can be structured differently. In some cases the observer replicates the goal of the observed action; in others the goal *and* the means used for achieving the goal are replicated.

2. The fundamental neurophysiological mechanism that underlies understanding of an action is a direct matching of the observed action with the motor representation of that action. This matching is made by the mirror neuron system. The matching of the observed action with its motor representation is a necessary prerequisite for imitation.

3. The matching mechanism by itself is not sufficient. It must be complemented by the activity of other mechanisms that modify and organize the mirror neuron system.

Here I summarize the properties of mirror neurons in monkeys, describe the properties of the mirror neuron system in humans, and finish by discussing the mechanisms that are necessary to achieve imitation.

1.2 The Monkey Mirror Neuron System: Motor and Visual Properties of F5 Neurons

Mirror neurons were originally discovered in area F5 of the monkey premotor cortex (di Pellegrino et al., 1992; Gallese et al., 1996; Rizzolatti et al., 1996a). This is a motor area that controls hand and mouth movements. A fundamental characteristic of this area is that many of its neurons discharge during specific goal-directed action (Rizzolatti et al., 1988). These neurons become active regardless of the effector (the right hand or the left hand or the mouth) used to achieve a specific goal (e.g., grasping an object). Conversely, they do not fire when a monkey uses the same effectors, but for another purpose (e.g., pushing objects away).

According to the action effective in triggering them, F5 neurons have been subdivided into various classes. Among them, the most represented are grasping, holding, tearing, and manipulating neurons.

A second fundamental characteristic of area F5 is that many of its neurons specify how a goal can be achieved. For example, the majority of grasping neurons discharge only if grasping is made using a particular type of prehension, such as a precision grip, finger prehension, and, more rarely, whole-hand prehension.

About 20% of F5 neurons respond to visual stimuli (Rizzolatti et al., 1988). One class of these visuomotor neurons is made up of *canonical neurons*, which discharge when a monkey sees an object that is congruent with the type of grip coded by the neuron (Murata et al., 1997). Visuomotor neurons in a second class do not discharge in response to the presentation of 3-D objects. The visual stimuli effective in triggering them are actions in which the experimenter (or a monkey) interacts with objects. Neurons with these properties are called mirror neurons (Gallese et al., 1996; Rizzolatti et al., 1996a).

Typically, in order to be triggered F5 mirror neurons require an interaction between hand and object. The sight of the object alone or of an agent mimicking an action is ineffective. The object's significance for the animal has no influence on mirror neuron response. Grasping a piece of food or a geometric solid produces responses of the same intensity.

A functional property of mirror neurons that is important for the issue of imitation is the relationship between their visual and motor properties. Most mirror neurons (93%) show a clear congruence between the visual actions they respond to and the motor response they code. According to the type of congruence they exhibit, mirror neurons were subdivided into strictly congruent and broadly congruent neurons (Gallese et al., 1996).

We labeled as strictly congruent those mirror neurons in which the effective observed and effective executed actions correspond both in terms of goal (e.g., grasping) and means, that is, the way the action is executed (e.g., precision grip). They represent about 30% of F5 mirror neurons.

We labeled as broadly congruent those mirror neurons that in order to be triggered do not require the observation of exactly the same action they code for motorically. Some of them discharge during the execution of a particular type of action (e.g., grasping) when executed using a particular grip type (e.g., precision grip). However, they respond to the observation of grasping made by another individual, regardless of the type of grip used (figure 1.1). Other broadly congruent neurons discharge in association with a single motor action (e.g., holding), but also respond to the observation of two actions (e.g., grasping and holding). Broadly congruent neurons are the largest class of mirror neurons (about 60%).

From this short review of the basic properties of F5 neurons, it appears that this area stores *potential* actions. The activation of F5 neurons does not necessarily determine an action; it evokes its representation. If other contingencies are met, this potential action becomes a *real* motor action (see Rizzolatti & Luppino, 2001). The potential actions associated with F5 neurons can be activated endogenously or exogenously. Exogenous (visual)
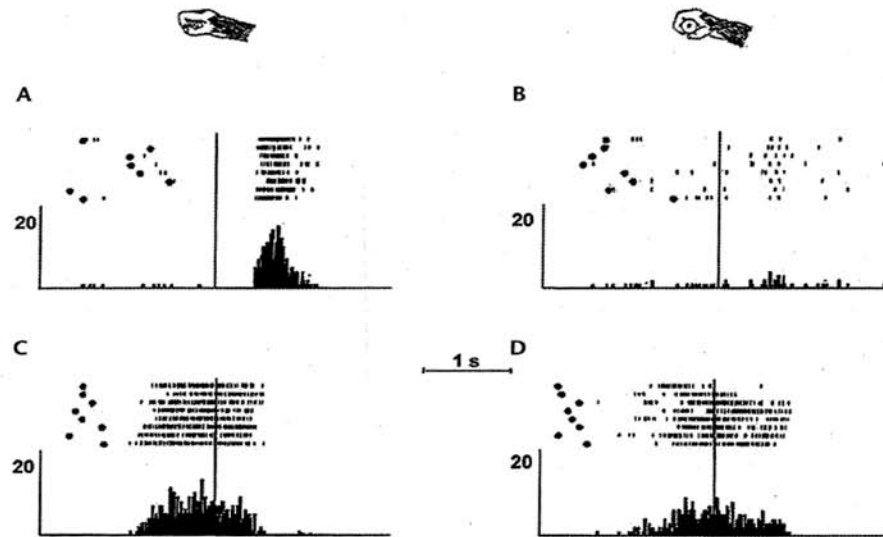
Figure 1.1
Example of a broadly congruent mirror neuron. (*A*) The monkey grasps a piece of food with a precision grip. (*B*) The monkey grasps a piece of food with whole-hand prehension. (*C*) The monkey observes an experimenter grasping a piece of food with a precision grip. (*D*) The monkey observes an experimenter grasping a piece of food with whole-hand prehension. In (*A*) and (*B*) the rasters are aligned with the moment in which the door of a testing box was opened (vertical line) and the monkey was allowed to grasp the objects. In (*C*) and (*D*) the rasters are aligned with the moment in which the experimenter touched the food (vertical line across trials). In the case of the monkey's active movements, the neuron showed a strong specificity for a precision grip. The filled circles indicate the beginning of the trials. Histogram bin width: 20 ms. Ordinates, spikes per bin; abscissas, time. (Modified from Gallese et al., 1996.)

activation is caused by the observation of objects (canonical neurons) or by the observation of actions made by others (mirror neurons).

1.3 Action Coding in the Temporal and Parietal Lobes of the Monkey

Neurons responding to the observation of actions made by others are not located only in area F5. In a brilliant series of studies, Perrett and his co-workers (Perrett et al., 1989; see for review Jellema & Perrett, 2002; Jellema et al., 2002) showed that neurons selectively responding to biological actions are present in the region of the superior temporal sulcus (STS). Actions effective in eliciting STS neuron responses are walking, turning the

head, bending the torso, moving the arms, and facial movements, as well as gaze direction. A small set of neurons discharge during the observation of goal-directed hand movements (Perrett et al., 1990b). The motor properties of STS neurons have not been specifically investigated. Motor-related activity, however, if present, should involve only a limited number of STS neurons.

Another cortical area where there are neurons that respond to action observation is area PF (Fogassi et al., 1998; Gallese et al., 2002). This area forms the rostral part of the inferior parietal lobule. PF receives input from STS and sends output to area F5. Conversely, F5 sends output to PF, which in turn sends projections to STS. Information is flowing, therefore, not only from STS to F5, but also from F5 to STS. Direct connections between STS and F5 have not been described.

Neurons in area PF are functionally heterogeneous. Most of them (about 90%) respond to sensory stimuli (Hyvarinen, 1982; Leinonen & Nyman, 1979; Fogassi et al., 1998; Gallese et al., 2002). About 50% of them also discharge in association with a monkey's active movements.

PF neurons responding to sensory stimuli can be subdivided into three categories: somatosensory neurons (33%); visual neurons (11%); and bimodal neurons, which respond to somatosensory and visual stimuli (56%). Among the neurons with visual responses (visual neurons and bimodal neurons), 41% respond to the observations of actions made by another individual. The effective actions most represented are grasping, holding, manipulating, and bimanual interactions. One third of PF neurons triggered by action observation do not appear to have motor-related activity. The other two-thirds discharge also during a monkey's movement and, in most cases, show the visuomotor congruence typical of mirror neurons (PF mirror neurons) (Gallese et al., 2002).

From these findings the following picture emerges. Visually described actions are first stored in STS. In this area many neurons "resonate" in response to the sight of specific actions. STS action description is then transferred to PF. In PF, some neurons are exclusively visual, but most of them also discharge during action execution.

If one considers that the repertoire of actions that each individual possesses is restricted in comparison with the richness and variety of visual representations of observed motor actions, the following tentative hypothesis about the organization of the STS-PF-F5 circuit can be advanced. Each PF neuron receives visual descriptions of those actions that have the same meaning, e.g., grasping in different ways or by different persons. The neurons that receive this information are bidirectionally connected with

the F5 neurons that code for grasping as a motor action. Thus the circuit on the one side "concentrates" the different visual descriptions of the same action on a restricted number of neurons, and on the other "labels" these neurons with a motor meaning. This convergence creates the basis for action understanding, regardless of the precise pictorial aspect of the action. I discuss later how this mechanism may also be involved in imitation.

1.4 Action Understanding: The Functional Role of the Mirror Neuron System

Since the discovery of mirror neurons, it has been proposed that they are involved in understanding actions. The core of this proposal is that an observed action acquires meaning for the observer when it activates motor schemas whose outcomes are known to the observer (see Rizzolatti et al., 2001).

There is an obvious objection to this proposal. Is motor activation really necessary to understand actions? In principle, an action could be understood in purely visual terms. Indeed, the data by Perrett and co-workers (see ch. 1.3) indicate that "prototypes" describing actions are present in STS. In addition, in humans, a rich description of body parts and body actions is present, not just in the STS region (see Allison et al., 2000), but also in the occipital cortex (Downing et al., 2001; Malach et al., 2002).

There is, however, a fundamental requirement that a description of actions must satisfy in order to provide meaning for the individual: It must link the external information to something that the individual knows. The visual system, like all sensory systems, is (by definition) a system that receives information. It does not generate it. In contrast, the motor system generates behavior and, on the basis of its consequences, is able also to "validate" the behavior produced. Thus, while the visual description of actions in STS is very useful for coding actions in a compressed way, this high-order visual information needs an additional mechanism to give it a meaning. F5 mirror neurons can effect this transformation. When the motor templates represented by mirror neurons resonate, the meaning of the observed action becomes transparent, because, when other contingencies are met, the activation of the same templates produces action.

The activation of representations of motor action is not the only way in which a visually described action may become meaningful. The observation that a certain visual event leads to consequences that the observer understands is another possibility. Note, however, that if the consequences of the observed actions do not directly concern the observer (such as a threatening gesture and its consequences), this type of understanding is different from that provided by motor mapping. It is a logical understanding, not a direct personal comprehension of what the other is doing.

An association between STS visual templates and subcortical centers also may give significance to an event. STS, besides sending information to PF, is part of a circuit that includes the amygdala and other centers related to emotions (Amaral et al., 1992). Activation of this circuit could give a personal significance to visual stimuli similar to that due to the activation of PF and F5 neurons. This, of course, assumes that there is a mirror neuron system for "hot," emotionally laden actions that is similar to that for the "cold actions" discussed earlier. Preliminary evidence suggests that this is the case (Wicker et al., 2003; see also Iacoboni, vol. 1, ch. 2).[1]
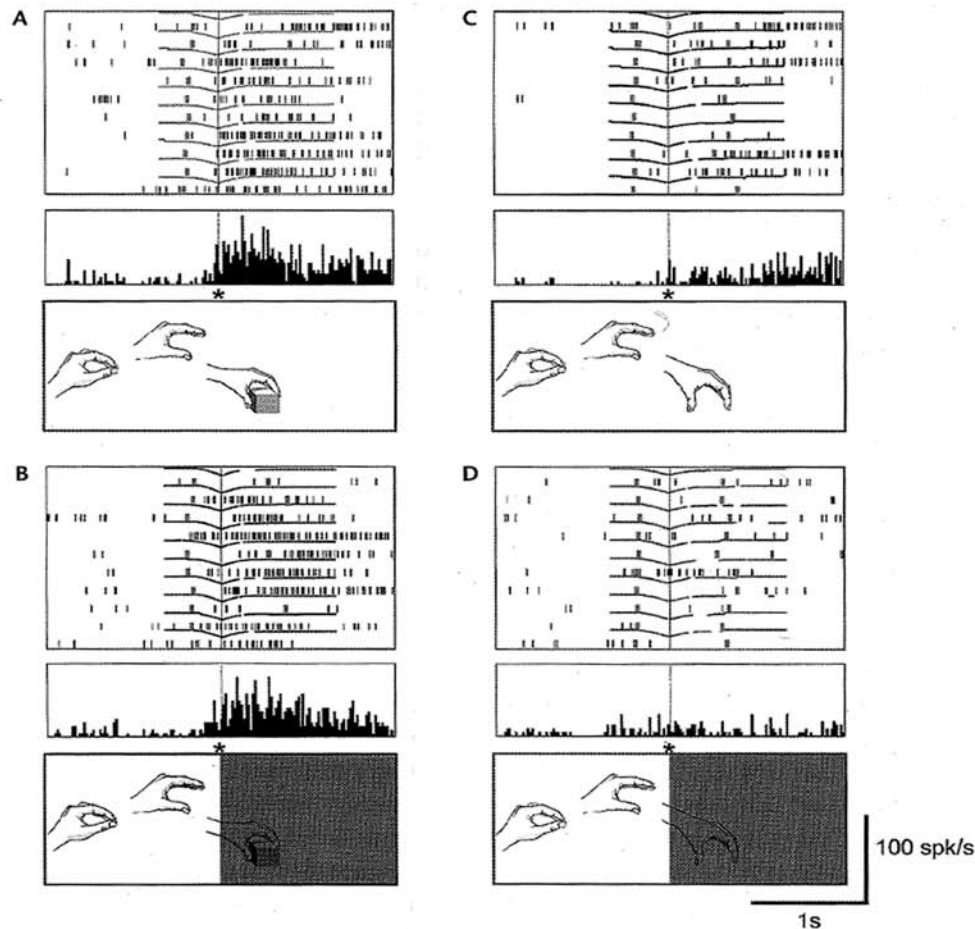
1.5 New Evidence of a Role for F5 Mirror Neurons in Action Understanding

The idea that the mirror neuron system is involved in action understanding can be tested by placing a monkey in situations in which the monkey is able to understand the meaning of an action, but the experimental sensory conditions are different from those that typically trigger mirror neurons. If mirror neurons are involved in action understanding, their activity should reflect the action meaning and not the sensory contingencies leading to action understanding.

A possible way to test this prediction is to present the monkey with auditory stimuli that evoke the idea of an action. This experiment was recently performed (Kohler et al., 2002). Activity in F5 mirror neurons was recorded while the monkey was observing a "noisy" action (e.g., ripping a piece of paper), or was presented with the same noise without seeing the action. The results showed that most mirror neurons that discharge on presentation of actions accompanied by sounds also discharge in response to the sound alone (audiovisual mirror neurons). Further testing showed that a large number of audiovisual mirror neurons respond selectively to a specific sound of an action. These results strongly support the notion that the discharge of F5 neurons correlates with the understanding of an action. The stimuli leading to action understanding are irrelevant. They could be visual or acoustical. Once the meaning of the action is specified, the neuron fires.

Another way to test whether action understanding triggers F5 mirror neurons is to prevent the monkey from seeing the action (and from

1. For discussions relevant to this section, see J. Prinz (vol. 2, ch. 13, p. 274ff), and the comments by Huesmann (vol. 2, ch. 19.6, p. 386).

Figure 1.2

Example of a neuron responding to action observation in full vision and in a hidden condition. The lower part of each panel schematically illustrates the experimenter's action as observed from the monkey's vantage point. In panels *B* and *D* the gray square inside the black frame represents an opaque sliding screen that prevented the monkey from seeing the action the experimenter performed behind it. The asterisk indicates the location of a stationary marker attached to the frame. In hidden conditions the experimenter's hand started to disappear from the monkey's vision when it crossed this marker.

In each panel above the illustration of the experimenter's hand, the raster displays and histograms of ten consecutive trials recorded are shown. Above each raster, the continuous line represents the kinematics of the experimenter's hand movements

hearing its sound), but to provide it with clues to what the action may be. If mirror neurons are involved in action understanding, they should also discharge in this condition.

An experiment testing this hypothesis was recently carried out by Umiltà et al. (2001). The experimental paradigm consisted of two basic conditions (figure 1.2). In one, the monkey was shown a fully visible action directed toward an object ("full vision" condition). In the other, the monkey saw the same action, but with its final critical part hidden ("hidden" condition). Before each trial, the experimenter placed a piece of food behind the screen so that the monkey knew that there was an object behind it. Only those mirror neurons were studied that discharged at the observation of the final part of a grasping movement and/or holding.

Figure 1.2 shows the main result of the experiment. The neuron illustrated in the figure responded to the observation of grasping and holding (A, full vision). The neuron also discharged when the stimulus triggering features (a hand approaching the stimulus and subsequently holding it) were hidden from the monkey's vision (B, hidden condition). As is the case for most mirror neurons, the observation of a mimed action did not activate the neuron (C, full vision and D, hidden condition). Note that from a physical point of view, B and D are identical. It was therefore the understanding of the meaning of the observed actions that determined the discharge in the hidden condition.

In total, more than half of the tested neurons discharged in the hidden condition. Out of them ($n = 19$), 7 did not show any difference between the hidden and full vision conditions, while 9 responded more strongly in the full vision condition. Of the remaining 3, the response was either more pronounced in the hidden condition than in full vision (1 neuron) or showed a temporal shift in response intensity.

In conclusion, both experiments in which the stimulus conditions were altered showed that F5 mirror neuron activation correlates with action understanding rather than with the stimulus properties leading to it. This finding strongly supports the notion that F5 activity plays a fundamental role in this function.

Figure 1.2

(continued)

expressed as the distance between the hand of the experimenter and the stationary marker over time. The rasters and histograms are aligned with the moment when the experimenter's hand was closest to the fixed marker (vertical line). Histogram bin width = 20 ms. The ordinates are in spikes per second. (Modified from Umiltà et al., 2001.)

1.6   The Mirror System in Humans

There is rich evidence that a mirror system exists also in humans. Evidence for this comes from electroencephalography (EEG), magneto-encephalography (MEG), transcranial magnetic stimulation (TMS), and brain imaging studies (e.g., Fadiga et al., 1995; Rizzolatti et al., 1996b; Grafton et al., 1996; Decety et al., 1997; Hari et al., 1998; Cochin et al., 1999). Many of these studies have been reviewed recently (Rizzolatti et al., 2001). Here only those particularly relevant for imitation are examined.

1.6.1   Transcranial Magnetic Stimulation Studies

The rationale of TMS studies of the mirror neuron system is the following. If there is a system endowed with mirror properties, the observation of an action performed by another individual should increase the motor-evoked potentials (MEPs) recorded from the observer's muscles involved in producing that action.

Fadiga et al. (1995) demonstrated that this is the case. Normal volunteers were required to observe an experimenter grasping different objects (transitive hand movements) or performing meaningless arm gestures in the air (intransitive arm gestures). As control conditions, detection of the dimming of a small spot of light or the presentation of 3-D objects was used. The results showed that observation of both transitive and intransitive actions produced an increase in the motor-evoked potentials recorded from the observers' hand and arm muscles. The increase was found in those muscles that the subjects would use to produce the movements observed.

Subsequent experiments confirmed the selectivity of the muscle excitation and described various cortical and spinal cord excitability changes caused by the observation of actions made by others (Baldissera et al., 2001; Gangitano et al., 2001; Maeda et al., 2002). Of these studies, the last two are of particular interest for imitation.

Gangitano et al. (2001) recorded MEPs from the hand muscles of normal subjects while they were observing grasping movements made by another individual. The MEPs were recorded at different intervals following onset of the movement. The results showed that cortical motor excitability faithfully followed the phases of the observed grasping movement (figure 1.3). This finding indicates that in humans the mirror neuron system codes for the temporal aspects of the observed movements and not only the meaning of the observed action.

Maeda et al. (2002) also recorded MEPs from two hand muscles of normal volunteers. The recordings were made while they observed video clips of
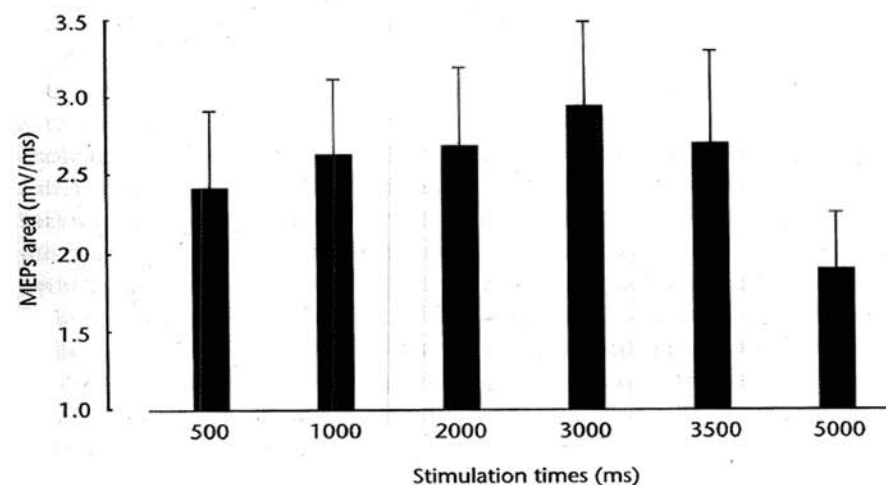
Figure 1.3
Averaged values of motor-evoked potentials (MEPs) of a hand muscle (first dorsal interosseus) collected at different times during the observation of a hand approaching a ball and grasping it. 500 ms, hand at the starting position (time value refers to the onset of the video clip showing the action); 3000 ms, hand maximum aperture. (Modified from Gangitano et al., 2001.)

different finger movements, such as thumb abduction or adduction. The finger movements were presented in two hand orientations: as if the actor were sitting next to the observer (hand "away" position) and as if the actor were in front of the observer (hand "toward" position). The results showed that the degree of cortical motor modulation depended on the orientation of the hand. Modulation was greater when the observed movement was performed in the hand away position (i.e., when the actor and the observer were in the same position) than in the hand toward position.

Summing up, TMS studies have shown two important properties of human mirror systems that have not been observed in the monkey. First, intransitive meaningless movements produce mirror neuron activation (Fadiga et al., 1995; Strafella & Paus, 2000; Maeda et al., 2002). Second, the correlation between the time course of the observed movements and the MEPs facilitation suggests a mirror mechanism that also codes for the movements forming an action. I previously referred (see Rizzolatti et al., 2002) to the movement-related mirror mechanism as the "low-level resonance mechanism," contrasting it with the "high-level resonance mechanism" of F5 where the coded element is the action. These properties of the

human mirror neuron system, which may explain the great human capacity for imitation, are discussed in ch. 1.10.

1.6.2 Brain Imaging Studies

Early brain imaging studies showed that the observation of hand actions activates (besides various occipital visual areas) the STS region, the inferior parietal lobe, and the ventral premotor cortex, including Broca's area (see Rizzolatti et al., 2001). The finding of activation of Broca's area during observation of hand action was rather unexpected. Although comparative cytoarchitectonic studies indicate that the *pars orbicularis* of Broca's area (area 44) is the human homologue of area F5 (see Petrides & Pandya, 1994), the traditional view is that area 44 is the speech motor area.

In recent years, however, rich evidence has been accumulating that, in addition to speech representation, area 44 contains, similarly to monkey area F5, a hand motor representation (Krams et al., 1998; Binkofski et al., 1999a; Iacoboni et al., 1999; Gerardin et al., 2000; Ehrsson et al., 2000; Schubotz & Von Cramon, 2001). The hand motor representation, albeit greatly overlapping with that of mouth, is situated dorsally to the latter, sometimes invading the adjacent ventral area 6, where proximal arm movements are located. It is interesting to note that precision grip is richly represented in area 44 (Ehrsson et al., 2000). The same overrepresentation of precision grip is found in the monkey area F5 (Rizzolatti et al., 1988).

This activation of area 44 gave rise to some speculation about a possible exclusive role for this area in functions mediated by the mirror neuron system, with the explicitly stated doubt that in humans, verbal mediation rather than the mirror neuron system plays a fundamental role in these functions (see Heyes, 2001a). New experiments on the functional organization of the mirror system have shown that this view is wrong. Buccino et al. (2001) examined the general organization of the mirror neuron system using as stimuli mouth, hand, and foot actions. Transitive actions (directed toward an object) and intransitive actions were used. The following stimuli were presented: biting an apple or chewing; grasping a cup, grasping an apple or miming these actions; kicking a ball, and pushing a brake or miming these actions. Observation of an action was contrasted with the observation of a static face, hand, and foot, respectively.

The observation of object-related mouth movements resulted in activation of areas 6 and 44 bilaterally. In addition, two activation foci were present in the parietal lobe. The rostral focus was located in area PF (BA 40), while the caudal one was (most likely) in area PG (BA 39). The observation of intransitive actions produced activation of the same premotor areas as the observation of transitive actions, but there was no parietal lobe activation.

Observation of object-related hand and arm movements resulted in two areas of activation in the premotor cortex, one corresponding to area 44 and the other more dorsal in ventral area 6. Considering the motor organization of this region, it is likely that the former activation was caused by observation of grasping hand movements, while that of area 6 was caused by observation of reaching. As for mouth movements, there were two activation foci in the parietal lobe. The rostral focus was still in PF, but was more posteriorly located than the focus observed during mouth actions, while the caudal focus was essentially in the same location as that for mouth actions. During the observation of intransitive movements, the premotor activations were present, but not the parietal ones.

Finally, the observation of object-related foot actions resulted in activation of a dorsal sector of area 6 and activation of the posterior parietal lobe, in part overlapping with that seen during mouth and hand actions (BA 39), in part extending more dorsally. Nonobject-related foot actions produced the area 6 activations, but not the parietal ones.

The results of this study are important for several reasons. First, they demonstrate that the mirror system includes a large part of premotor cortex and the inferior parietal lobule. It is not limited to Broca's area. Second, they show that the activation map obtained during observation of actions made with different effectors is similar to the motor map (the so-called "homunculus") obtained with electrical stimulation of the same region. Finally, they allow one to rule out the idea advanced by some authors (see Grèzes & Decety, 2001; see also Heyes, 2001a) that the activation of area 44 is due to internal verbalization. Verbalization cannot be present during the observation of hand movements and magically disappear during the observation of foot movements.

In conclusion, the human mirror system is widespread and centered on the inferior parietal lobule and the premotor cortex, including area 44. The next section examines how this system is involved in imitation.

1.7 The Mirror Neuron System and Imitation

Imitation (as defined in ch. 1.1) is based on two distinct but related mechanisms: the capacity to understand actions done by others and the capacity to replicate those actions. The data reviewed in the previous sections strongly suggest that the mirror neuron system plays a fundamental role in

understanding actions peformed by others. In this section I examine how this system is also involved in replicating the observed action.

Before doing this, however, it is important to stress the different ways in which the term "imitation" is used in psychological and in ethological studies. In psychological studies, "imitation" refers to the behavior of subjects instructed to replicate an action that is already in their motor repertoire. The aim is to discover the rules that the subjects use in copying others' behavior and to compare them with the rules used in acting on the basis of other biological and nonbiological stimuli (e.g., Brass et al., 2000; Bekkering & Wohlschläger, 2002; W. Prinz, 2002). In ethological studies, the stress is on learning. Imitation is the capacity to acquire a motor behavior previously not present in the observer's motor repertoire (see R. Byrne, 1995, 2002c; Tomasello & Call, 1997; Visalberghi & Fragaszy, 2002). Particular emphasis is often given here to the precise motor details of the imitated action (Tomasello & Call, 1997).

1.8 Imitation of Actions Present in the Observer's Repertoire: Brain Imaging Experiments

An important role in the renewal of interest in imitation in psychology has been played by the reconsideration of the concept of ideomotor compatibility (R. Lotze, 1852; James, 1890; Greenwald, 1970) by Prinz and his colleagues (see W. Prinz, 2002). According to these authors, stimuli and responses are represented in the cognitive system as events and coded in a commensurable format. Thus the perception of a stimulus event that shares features with a similar motor event tends to induce it (W. Prinz, 2002). The greater the similarity between the stimulus event and the motor event, the stronger will be the induction of the observed action.

These theoretical ideas and the finding that mirror neurons directly match the observed actions in their corresponding motor representations prompted brain imaging experiments aimed at finding the neural substrate that is specifically activated during imitation (Iacoboni et al., 1999, 2001; Nishitani & Hari, 2000).

Using functional magnetic resonance imaging (fMRI), Iacoboni et al. (1999) studied normal human subjects under two basic conditions: "observation-only" and "observation-execution." In the observation-only condition, the subjects were shown a moving finger, a cross on a stationary finger, or a cross on an empty background. The instruction was to observe the stimuli. In the observation-execution condition, the same stimuli were presented, but this time the instruction was to lift the right finger as fast as possible in response to them.

The fundamental comparison of the study was the one between the trials of the observation-execution condition in which the volunteers made the movement in response to an observed action ("imitation") and the trials of the same condition in which the movement was triggered by the cross projected on a finger or an empty background. The results showed that there were four areas in which the activity was stronger during imitation trials than during other trials: left area 44, the right anterior parietal region, the right parietal operculum, and the right STS region (see for this last activation, Iacoboni et al., 2001).

In all trials of the observation-execution condition, the motor action (finger lifting) made by the subject was identical. The fact that activation of area 44 (an area also active in the observation-only condition) was stronger during imitation trials than during the other two observation-execution trials strongly suggests that a direct mapping between the observed and the executed act occurs in this area. Area 44, therefore, appears to contain a mirror mechanism.

Remarkably similar results were obtained by Nishitani and Hari (2000) by using the event-related neuromagnetic technique. This technique is inferior to fMRI in terms of spatial resolution, but allows one to obtain an excellent time resolution. In their experiment, Nishitani and Hari asked normal human volunteers to grasp a manipulandum, or to observe the same movement performed by an experimenter, or to observe and replicate the observed action. The results showed that during an active grasping condition, there was an early activation in the left inferior frontal cortex (area 44), with a response peak appearing approximately 250 ms before the touching of the target. This activation was followed within 100–200 ms by activation of the left precentral motor area and 150–250 ms later by activation of the right one. During imitation, the pattern and sequence of frontal activation were similar to those found during execution, but activation of area 44 was preceded by an occipital activation that was due to visual stimulation present in the imitation condition.

As far as the other activations described by Iacoboni et al. (1999) are concerned, the parietal activation could reflect a mirror mechanism similar to that proposed for area 44. This interpretation, however, is in contrast with the finding that superior parietal lobule activation is typically not present in experiments in which the subjects are instructed to observe actions only in order to understand them (see Buccino et al., 2001).

Furthermore, in the monkey, the superior parietal lobule, in contrast to the inferior one, does not receive input from STS, where visual templates of biological actions are coded (G. Luppino and M. Matelli, personal communication).

A possible alternative interpretation may be that during imitation the activation of motor representations of the intended actions produces, through backward projections, sensory copies of the intended actions. In the monkey, the superior parietal lobule and especially its rostral part (area PE) contains neurons that are active during proprioceptive as well as during active arm movements (Mountcastle et al., 1975; Kalaska et al., 1983; Lacquaniti et al., 1995). These properties suggest that the observed superior parietal activation may represent a kinesthetic copy of the intended movements. This interpretation fits well with positron emission tomography (PET) data by Decety and his co-workers (Decety et al., 1997; Grèzes et al., 1998), who also showed an increase in superior parietal activations when the subjects' task was to memorize actions in order to repeat them.

An interpretation in terms of sensory copies of the intended action may also explain the activations observed by Iacoboni et al. (1999) in the parietal operculum. It is known from monkey studies that there are several sensory areas located in this sector of the parietal lobe, among them areas PV and SII (Robinson & Burton, 1980a,b; Krubitzer et al., 1995). Brain imaging data have shown a similar organization in the human brain (Disbrow et al., 2000). Thus, by analogy with the interpretation of the parietal activation, one may hypothesize that the observed activation represents a tactile copy of the intended action. Interestingly enough, the pure observation of hand manipulation actions decreases signals evoked in the SII region by median nerve stimulation (Avikainen et al., 2002). In accord with these findings, the experiments of Iacoboni et al. (1999) found no activation in the parietal operculum during the observation-only condition.

The activation in STS is particularly intriguing. This activation, which is located in a caudal part of the STS region, rostral to and slightly dorsal to the motion area V5/MT, was close to significance only in the experiment by Iacoboni et al. (1999). Considering, however, the theoretical importance of a visual copy of the intended action, this activation was further investigated in a new experiment in which, as in the previous experiment, volunteers observed ("observation-only") or executed ("observation-execution") a finger movement with their right hand. The hand whose movement they observed was this time either the right or the left one and not the left hand only, as in the previous experiment. In half

of the trials, the stimulus was a finger movement, in half a small cross presented on the finger (which was still). The hand of the subjects was hidden from their vision (Iacoboni et al., 2001).
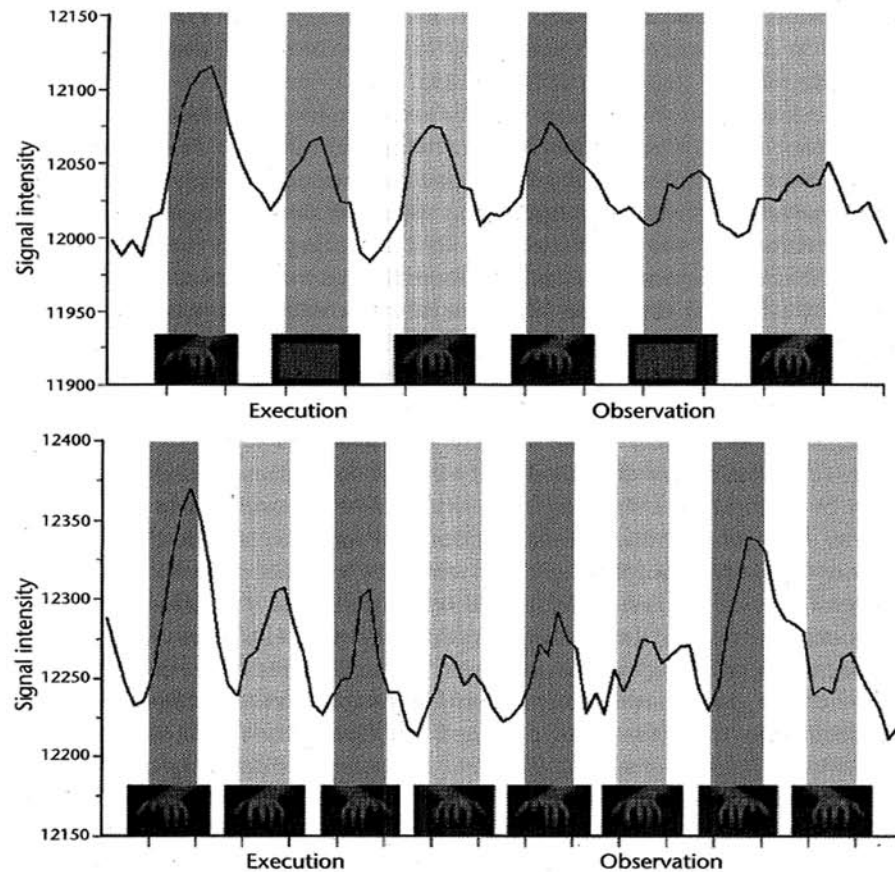
The most interesting result of this experiment was the difference in activation of STS during observation (observation only) versus imitation (observation-execution) of the finger movements, according to which hand was *observed*. During observation, the strongest activation was caused by the movement of the hand *anatomically* corresponding to that used by the subjects in the experiment (i.e., the right hand). In contrast, during imitation, the strongest activation was seen in the condition in which the hand *spatially* corresponding to that of the subjects triggered the movement (figure 1.4). In other words, during observation the anatomical congruence was favored, while when imitation was required, the space common to the acting hand and to the observed hand was favored.

This reversal of activation intensity in STS during imitation is consistent with behavioral data showing that when an individual has to imitate the action of another individual facing him, he tends to imitate the observed action in spatial coordinates. This tendency can be easily demonstrated by touching one's own cheek and telling a person, "You have something here on your cheek." Invariably the addressed person touches the cheek of his/her own that is spatially congruent with that indicated by the speaker, and not the one anatomically congruent. Experimental evidence for this mirror-image imitation was provided by Shofield (1976a,b) and more recently by Bekkering et al. (2000). For an in-depth discussion of the effect, see Gattis et al. (2002).

From the neurophysiological point of view, it is likely that the reversal in activation from the anatomically congruent to the spatially congruent effector during imitation is caused by an efferent discharge coming from PF and priming the STS hand prototypes that are spatially congruent with the observed ones. This is most likely an adaptive behavior determined by the experience that there is an advantage in sharing space when two individuals face one another. Thus, the hand prototypes spatially congruent with the hand action to be imitated prevail over the prototypes representing the hand anatomically congruent to the observed one.

Taken together, these experiments strongly support the idea that the mirror system plays a central role in the imitation of actions that are already in the motor repertoire of the individuals. The mirror system matches the observed action with motor responses stored in the premotor cortex and allows a fast, efficient response to that action. In addition, these experiments suggest that sensory copies of actions to be imitated are formed

Figure 1.4
Time series of the active superior temporal sulcus (STS) site during a finger movement (lifting the right index finger—execution) and during observation of a similar finger movement made by another individual (observation). The finger movement in execution was triggered by the observation of finger lifting by another individual or by the presentation of a black cross on a finger. The small pictures correspond to the type of stimulus presented. Nine rest periods that were alternated with the eight active task periods are presented. See the text for other explanations. (Modified from Iacoboni et al., 2001.)

in various areas. This important theoretical point, however, requires further experiments in order to be definitively proved.

1.9   Imitation of Actions Not Present in the Observer's Repertoire: Brain Imaging Experiments

Unlike psychologists, ethologists typically stress the learning aspects of imitation. Many consider "true" imitation to require the precise repetition of an observed action previously not present in the observer's motor repertoire; the learning of actions with effects on the environment that are similar to the observed ones is not sufficient (A. Spence, 1937; Galef, 1988; Tomasello, 1990; R. Byrne & Tomasello, 1995). This view in large part relates to the need to exclude from imitation motor behaviors apparently learned by observation of action but in fact triggered by the meaning of the stimulus (A. Spence, 1937) or by its affordances (Tomasello, 1990).

Two different ways of learning a new motor behavior should be distinguished. One is substitution for the motor pattern spontaneously used by the observer in response to a given stimulus of another motor pattern, more adequate to reach the intended goal, on the basis of observation of the behavior of another individual. Examples could include the correct way to hold a tennis racket or to place a finger on a guitar's neck (action adjustment). The second way is learning, by observation, a new motor sequence that is useful to reach a certain goal (sequence learning). The ability to open a box only if a certain action sequence is followed could be an example of this second type of imitation learning (see also the artificial fruit of Whiten and Custance, 1996).

There are no experiments that I am aware of that have studied these two types of motor behavior from the perspective of mirror neurons. So in this section I discuss the issue of acquisition of new motor behaviors following observation of actions made by others mostly in terms of possible mechanisms that may explain them rather than on the basis of empirical studies.

The neurophysiological network that should intervene in action is that formed by STS, PF, and F5. As discussed earlier, this circuit stores many visually described actions in its visual node, STS. STS neurons send information to PF, where there are neurons that receive, in addition to STS input, backward connections from F5. The way an individual interacts with an object before learning is established by F5 canonical neurons that specify which type of movement (e.g., a specific type of grip) has to be used on the basis of the object's affordances (see Jeannerod et al., 1995). When the observer sees that another grip is more efficient than the one previously used

to reach the goal of the action, this new grip is coded in STS. The learning process consists of the production of a motor pattern that activates, via backward connections, those PF neurons that receive the sensory copy of the desired action from STS. The comparison between the visual aspect of the performed action and the sensory copy of it will allow a modification of the internal motor pattern until this pattern produces an action similar to the observed one.

This model is basically an internal forward model (see Wolpert et al., 1995; Wolpert, 1997; Kawato, 1999; Arbib & Rizzolatti, 1999). Its main biological constraints are, on the motor side, the motor repertoire present in PF and F5 and, on the sensory side, the variety of action prototypes coded in STS and their plasticity. The presence in humans of a rich representation of intransitive motor acts, shown by TMS studies, renders the human mirror neuron system much more apt for imitation than the analogous monkey system, where the poor representation of intransitive actions (or even its absence) and the apparent poverty (on the basis of available evidence) of mirror neurons coding for precise details of actions present serious limits to the capacity for imitation. Without the storage of intransitive actions to complement basic object-related actions and precise copies of actions, the capacity of the monkey system to imitate the behavior of others should be rather limited.

Logically, the mechanism that is the basis of learning a sequence by imitation ought to be different. Here, unlike the case of action adjustment, the essential achievement is not the substitution of an action determined by an object's affordances with a more effective action, but rather the capacity to replicate a series of actions previously never executed.

An interesting hypothesis to explain how this type of imitation may occur has been recently advanced by Byrne (see R. Byrne, 2002c and chapter 9). According to Byrne, sequence learning by imitation is based on two operations. The first is the capacity to segment the perceived action into smaller units and to match them to "motor acts" already present in the motor repertoire of the observer. Mirror neurons are the elements that perform this matching.

The other essential operation ("string-parsing") consists of extracting the statistical regularities that characterize an action's sequence. This operation imposes high-order organization on the observed action sequence and, if successful, mirrors the original planning structure that produced the behavior. On the basis of neurophysiological data indicating a role for the mesial cortical area in sequence learning and execution (see Hikosaka et al., 1995, 2000; Tanji, 1996; Tanji et al., 1996; Shima & Tanji, 2000), Byrne

proposed that these areas also play a role in string parsing. An additional possible neural substrate for this operation is the basal ganglia, which also appears to play a role in sequence learning.

Obviously, at present the proposed mechanisms for action adjustment and for sequential learning by imitation are both merely hypothetical. However, they suggest a series of brain imaging experiments that may be easily performed using the available technology.

1.10 Concluding Remarks

A point central to this chapter's attempt to give imitation a neurophysiological basis is that an understanding of actions preceded imitation in evolution. The mirror system evolved as a system whose main aim was to match sensory information to personal motor knowledge of action meaning. This system became progressively richer and more complex and, in humans, came to include intransitive actions and detailed specifications of how an observed action is executed. This evolved mirror system became the basis for reproducing actions performed by others; that is, for imitation.

A possible criticism of this view is that some actions produce imitation without any evidence that they have been understood. There are several examples of this type of behavior. In many species of animals, for instance, the observation of a movement made by one individual is a signal for the rest of the group to start a similar movement (e.g., the behavior of shorebirds studied by Thorpe, 1963). Imitation of this type, that is, imitation without understanding the meaning of an action, is present in humans. A well-known example is the capacity of newborns, first described by Meltzoff and Moore (1977), to imitate buccal gestures. Other examples are laughing, yawning, crying, and, as shown by Dimberg et al. (2000), involuntary mimicking of facial expressions. It is likely that the main purpose of these behaviors is to create a link between individuals by facilitating affiliative behaviors and inhibiting aggressive behaviors.

Is such imitation without understanding also dependent on mirror neurons? In the absence of empirical data, a response to this question can be, obviously, only hypothetical. It is tempting, however, to think that the same mechanism underlies these behaviors and action understanding. At this point an obvious conceptual difficulty arises. It is difficult to accept that relatively simple behaviors such the escape behavior of shorebirds mentioned earlier developed after action understanding and requires this understanding as its prerequisite. The interpretation given by Thorpe (1963, see also Tinbergen, 1953) in terms of releasing signals appears to be

more parsimonious and convincing. A possible solution of this paradox might lie in the distinction between high-level and low-level resonance mechanisms (Rizzolatti et al., 2002).[2]

According to this proposal, there are neurons endowed with motor properties (motor neurons in a broad sense) that resonate when an appropriate stimulus is presented. The effect of this resonance is radically different according to the role that these neurons play in motor control. If they are close to the effectors, their low-level resonance elicits an actual motor action, with little if any cognitive effects. In contrast, if the neurons represent the action internally without necessarily causing motor effects (e.g., F5 mirror neurons), their high-level resonance would produce mostly cognitive rather than motor phenomena, such as action understanding.

This view, although hypothetical, has some interesting consequences. First, it allows one to give a unitary explanation of the different types of imitative behaviors, those accompanied by action understanding and those without it. Second, assuming that a mirror mechanism underlies both these phenomena, the unitary interpretation avoids the paradoxical notion that a cognitive function such as action understanding preceded in evolution capacities that can be explained without invoking high-level cognitive processes. Third, but not least, it provides clear, testable hypotheses about the mechanisms underlying imitation.[3]

Acknowledgments

2. For a relevant discussion see Hurley (vol. 1, ch. 7, p. 184).
3. See the comments on this chapter by Rawlins (vol. 1, ch. 8.1, p. 195).