

# Chapter 5

## Mirror Neurons in Action: ERPs and Neuroimaging Evidence



Alice Mado Proverbio and Alberto Zani

**Abstract** According to V.S. Ramachandran (inaugural ‘Decade of the Brain’ lecture at Society for Neuroscience meeting), ‘mirror neurons are to neuroscience what DNA was to biology’. Their discovery (by Rizzolatti’s group) led to the understanding of how hominids rapidly evolved through imitation and cultural transmission in the last 100,000 years. In this chapter, we will review the role of human mirror neuron system (MNS) in several mental and brain functions including: interacting with the environment, grasping objects, empathy and compassion for others, empathizing, emulation and emotional contagion, observing and imitating, learning sports, motor skills and dance, motor rule understanding, understanding the intentions of others, understanding gestures and body language, lip reading, recognizing actions by their sounds, learning to play a musical instrument. The chapter is enriched with a discussion of possible criticalities and caveats.

**Keywords** Mirror neuron system · Empathy · Audio-visuomotor neurons

### Introduction

Many neuroimaging studies have searched for the human correlate of the monkey ‘mirror neuron system’ (MNS) and tried to isolate mirror neurons by using non-invasive imaging techniques such as fMRI (e.g. Dinstein, 2008; Iacoboni, 2005; Iacoboni et al., 2005; Schmidt et al., 2021). The results provided evidence of a strong activation of the anterior intraparietal area (AIP) and the ventral premotor areas (F5) when subjects passively observed others performing movements, actively executed movements themselves, or imitated movements made by others. In addition, AIP and F5 areas were frequently found engaged in tasks involving empathy,

---

A. M. Proverbio (✉)

Department of Psychology, University of Milano-Bicocca, Milan, Italy  
e-mail: [mado.proverbio@unimib.it](mailto:mado.proverbio@unimib.it)

A. Zani

School of Psychology, Vita-Salute San Raffaele University, Milan, Italy

© The Author(s) 2023

P. S. Boggio et al. (eds.), *Social and Affective Neuroscience of Everyday Human Interaction*, [https://doi.org/10.1007/978-3-031-08651-9\\_5](https://doi.org/10.1007/978-3-031-08651-9_5)

social cognition, and theory of mind, along with the inferior frontal gyrus, inferior parietal cortex, fusiform gyrus, posterior superior temporal sulcus, and amygdala. These data seem to suggest the existence of a shared neural mechanism for social cognition.

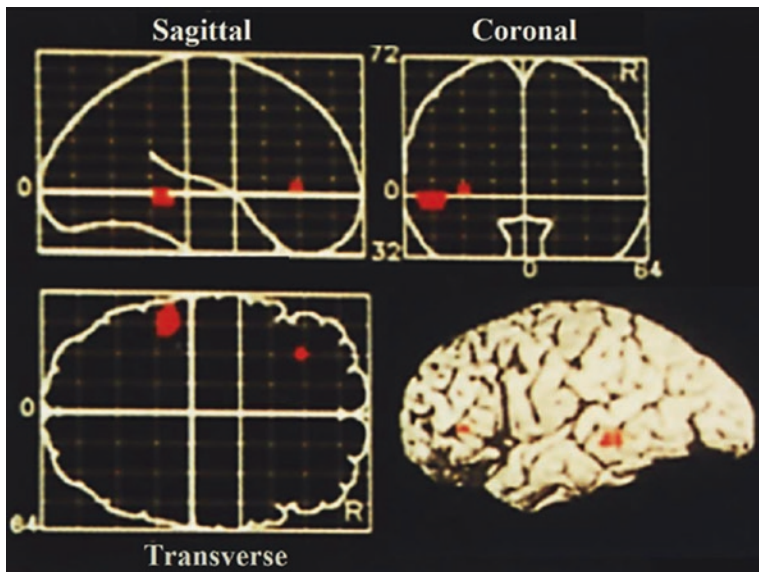
Despite the long line of research, studies on the human MNS still suffer from some severe methodological problems. Electrophysiological single unit recordings, which are required for a clear-cut demonstration of mirror neurons properties, are not feasible in humans. Therefore, the majority of studies approaching MNS in humans rely on methods with low temporal resolution (e.g., fMRI), which is an indirect method based on blood oxygenation signal (BOLD) and not directly measuring neuronal activity (see also the paragraph devoted to pitfalls at the end of the chapter). In this regard, ERPs can be excellent tools for providing the necessary temporal resolution for studying action and gesture recognition processes in healthy humans.

## Visuomotor Neurons and Action Encoding

Mirror neurons (MNs) were first discovered in the ventral premotor cortex (PMv cortex) of the macaque monkey (F5 area) by Rizzolatti and Luppino (2001). These neurons were activated both when the animal performed a specific motor action and when observed another simian or human individual performing that same action. The MNs do not respond to the simple presentation of food or other objects that also affect the animal, nor they are activated by the observation of a mimed action without the presence of the objects. In order for the MNs to activate (or 'to fire', i.e., to show an intense discharge frequency), an actual interaction of the hand with a target object of the action is essential. Despite being motor neurons, MNs are not activated by single movements (e.g. of the fingers) comprising a whole motor act, but, like all the other neurons in the PM cortex, are instead activated in association with goal-directed and purposeful motor actions. The MNs are stimulated by the execution/observation of motor actions performed with the hand, but also with the mouth. They are very sensitive to the type of grip (i.e. precision grip, power grip, grip of small or large objects, grip of little seeds, etc.) and encode the *actions goal*. For instance, the neural micro-population that encodes the gesture of taking an apple will not be the same if its purpose is to eat it (i.e. the animal takes the apple and then brings it to the mouth), or to throw it away (i.e. the monkey takes the apple and throws it). After the discovery of MNs in the premotor cortex (PM), other studies have shown their presence in the inferior parietal lobe (IPL), in particular in the rostral portion of this brain lobe. These neurons would be more involved in the representation of the actions associated with an object or a tool of which they process the motor properties (i.e. *affordances*), such as, for instance, its graspability and/or usability, while dealing with information coming from the fronto-parietal-occipital visual ventral stream (VVS).

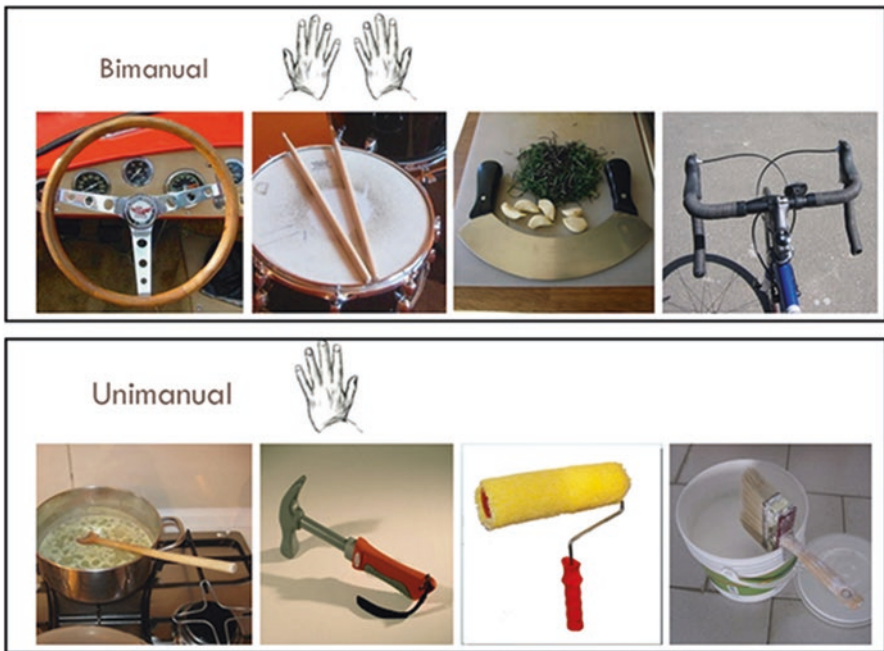
Many studies have shown that the mirror neuron system (MNS) is also present in humans (Rizzolatti & Craighero, 2004; Rizzolatti & Sinigaglia, 2016). Fine examples of these studies are the EEG investigations on the reactivity of brain rhythms during actions observation. Many studies have shown that the sight of actions performed by other individuals (with hands, legs, fingers, etc.) induces a block of observers' sensory-motor EEG rhythm (or so-called *mu* rhythm) recorded at scalp sites, which would reflect a state of relative inactivity in the Rolandic region (e.g. Lelord et al., 1998). An important PET study on human volunteers is the one carried out by Rizzolatti and colleagues (Rizzolatti et al., 1996), which allowed a first localization of the areas involved with the MNS during the observation of grasping movements. Volunteers were tested in three different conditions. In the first, they observed grasping gestures of common objects performed by the experimenter; in the second, they proceeded to reach and grasp the objects themselves, while in the third, they simply observed the objects. The results showed that only action observation activated significantly the inferior parietal lobule (IPL) and the ventral pre-motor area (PMv) together with the posterior portion of the inferior frontal gyrus (IFG) (Fig. 5.1).

Other studies have shown that the MNS is not only activated at the sight of gestures but also of manageable objects. By means of fMRI evidence, Creem-Regehr and Lee (2005) demonstrated that graspable tool shapes activated motor-related regions of the cortex, including the PMv area and the posterior parietal cortex



**Fig. 5.1** Adjusted mean regional cerebral blood flow recorded by Rizzolatti et al. (1996) during grasping observation. The data are displayed as statistical maps overlaid on three planar projections (sagittal, coronal, and transverse) frames and as cortical rendering of the lateral cortical surfaces of the left hemisphere. The pixel values significantly higher than  $p < 0.001$  are shown in red

(PPC). The event-related potentials (ERPs) study by Proverbio et al. (2011a, b) provided the possible time course of this activation showing that the earliest neural tool/non-tool discrimination was indexed by an increased anterior negativity in the 210–270 ms post-stimulus latency range in response to tools rather than to objects. Source reconstructions for these findings highlighted the contribution of left-sided brain premotor and somatosensory cortices, possibly including the anterior intraparietal sulcus (aIPS). Further studies demonstrated that the cortical representation of actions (especially tools manipulation and use) is asymmetrically represented over the left hemisphere. Indeed, a lesion of the left inferior parietal cortex (IPC, BA40) is often associated with apraxic deficits, whilst a right-sided lesion rarely causes these deficits (Goldenberg & Spatt, 2009). The question of whether this hemispheric asymmetry depends on right-hand use or a hemispheric functional specialization for fine-grained, precision movements has been explored in another ERP study by Proverbio et al. (2013). The authors recorded ERPs to pictures depicting unimanual (e.g. a hammer) or bimanual (e.g. a bicycle handlebar) tools, while participants were instructed to respond motorically to infrequent images of green plants (Fig. 5.2). A prefrontal N400 component (elicited by non-targets) was much larger over the left scalp sites to bimanual than unimanual tools. *swLORETA* (acronym for *standardized weighted LOW-Resolution electromagnetic TomogrAphy*) sources reconstruction revealed that besides the left and right parietal cortices (BA39,

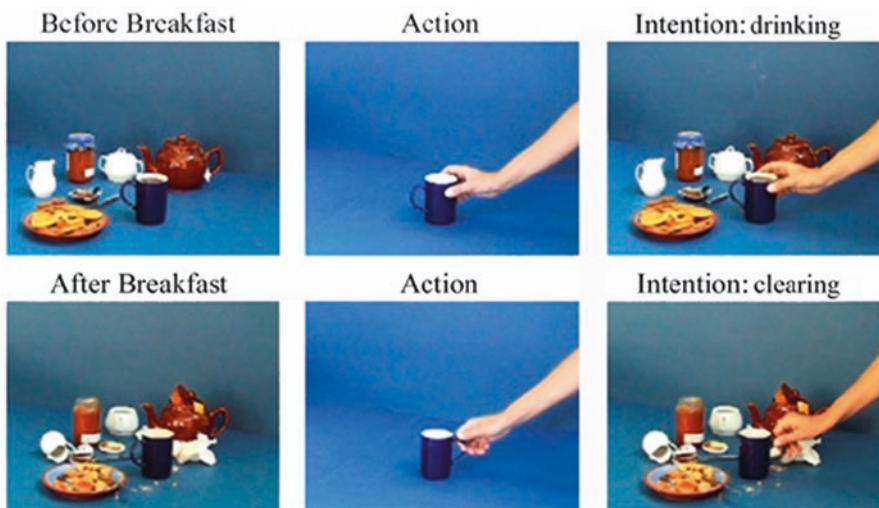


**Fig. 5.2** Examples of pictures depicting bimanual and unimanual tools used as stimuli in Proverbio's et al. (2013) ERP study

BA40), tools observation always activated the left premotor cortex (BA6) regardless of the hand involved in their manipulation/use. Overall, these data suggest that looking at tools automatically activates mental representations associated with their manipulation, with a left-sided hemispheric asymmetry for this brain activation.

## Mirror Neurons and Understanding the Intentions of Others: Empathy

An fMRI study by Iacoboni et al. (2005) has robustly demonstrated that the activation of visuomotor MNs makes it possible to share behavioral goals and to understand other people's intentions (a multifaceted capacity called *mentalizing* or *theory of mind*). In this famous experiment, participants observed three types of stimuli: grasping actions without context (the box in the middle in Fig. 5.3), the context without actions (the left box in Fig. 5.3), and manual actions performed in two different contexts (the right upper or lower boxes in Fig. 5.3). In this last condition, the context suggested the intention associated with the grasping action (i.e. drinking or clearing). Actions associated with the specific contexts produced a significant increase of the bold signals at the back of the IFG and in the PMv, being part of the MNS. Furthermore, the activation revealed to be greater for drinking (biologically more relevant) than for clearing. These data showed how these regions, active



**Fig. 5.3** Types of stimuli used in Iacoboni's et al. (2005) study. The same action (e.g. to take a mug) reveals an agent's different intention according to the context in which he/she is. Such an intention is encoded and inferred by means of fronto-parietal MNS activation. (Courtesy of Marco Iacoboni)

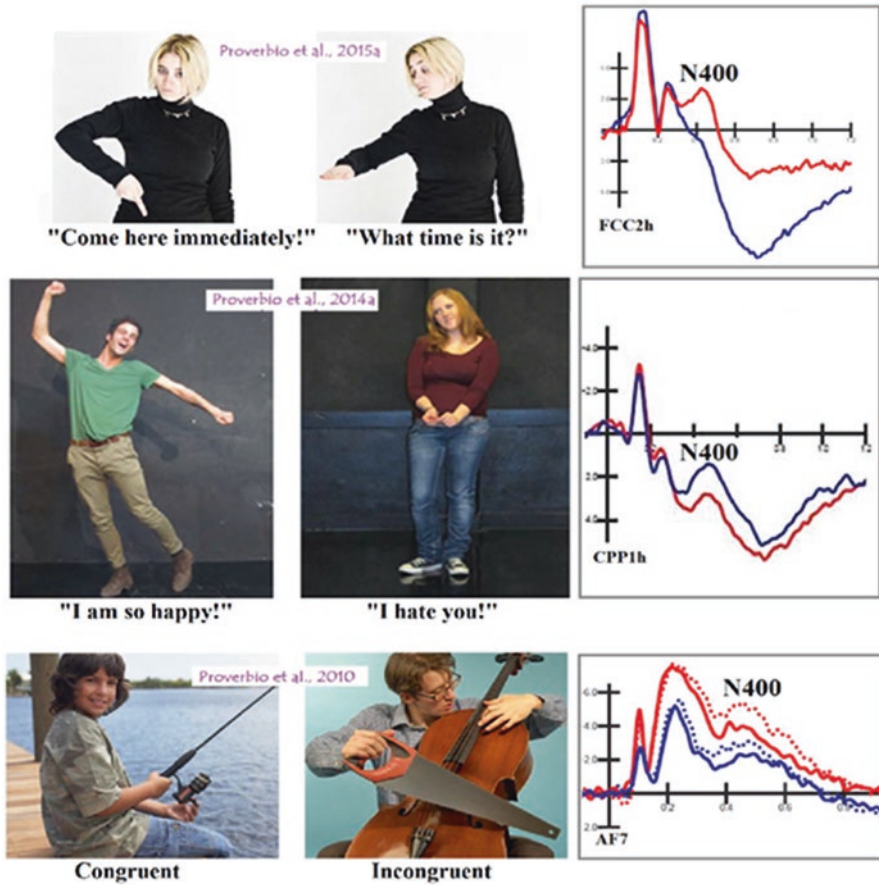
during the execution and observation of an action, were also involved in understanding the intentions of others.

For its ability of understanding visual gestures and their aims, the fronto-parietal MNS is involved in a multiplicity of mental functions including:

- (i) Understanding motor events
- (ii) Understanding actions and intentions of others
- (iii) Understanding mental/emotional states of others: empathy (adopting another person's point of view)
- (iv) Imitation, in yawning, scratching, crossing legs, posture
- (v) Learning: visuomotor processes (e.g. music, sport)
- (vi) Social cohesion, group behaviour, disgust (emulation, what the other feels, I also feel it: fear, embarrassment, shame or head emulation)
- (vii) Empathy for pain (when looking at someone who is suffering, feeling vicarious pain, compassion)

It has been shown how recognition of body language, both symbolic and affective, as well as the congruence of people's gestures are strongly related to the fronto-parietal MNS receiving and processing information from brain regions specialized in recognizing faces (i.e. fusiform face area or FFA), facial expressions (i.e. FFA and superior temporal sulcus, STS), and bodies (i.e. extrastriate body area, EBA). In a series of electrophysiological studies by Proverbio et al. (2010, 2014a, 2015a), visual ERPs were recorded in different samples of volunteers viewing hundreds of images depicting actors and actresses mimicking a symbolic gesture (iconic, deictic, or emblematic, such as, for instance, those in Fig. 5.4, top) or an emotional display of mood using body language (as shown in Fig. 5.4, middle), or using a tool (Fig. 5.4, bottom). In half the cases, the scene was incongruent with its verbal description and/or with respect to pragmatics or standard knowledge about tools use. In all cases, the perception of incongruent images (from the points of view of the gesture or of the action meaning and/or aim) elicited a wide negative response (i.e. N400) tending to be larger at anterior scalp sites. Applying swLORETA inverse solution to the N400 potential (within its time window of occurrence), it emerged that the incongruity between actions and their presumed intentions stimulated the activation of slightly different neural circuits in the three conditions (certainly more emotional in the case of body language; Fig. 5.4, middle), but invariably including the inferior regions of both the frontal premotor and parietal areas (i.e. the fronto-parietal MNS system, in addition to the anterior cingulate cortex (ACC), the superior temporal cortex (STC), and the visual (FFA and EBA areas).

All in all, these data suggest how the MNS underpins the ability to recognize the intentions of an agent, through the observation of a gesture and the motor simulation of that same gesture by an observer.



**Fig. 5.4** Examples of congruent (left column) and incongruent (right column) stimuli used in Proverbio's et al. studies (2010, 2014a, 2015a), associated with the N400 component electrophysiological effect (third column), reflecting the violation of an expectation related to the aim of the action or of the gesture expressed by the actors, as referred to a shared grammar of gestures or to the context and to the pre-established use of a tool. The N400 effect is drawn as a red continuous line in the upper waveforms, as a blue continuous line in the middle waves, and as a red dotted line in the lower waves (where ERPs are shown in red for women and in blue for men). (Reproduced and modified with the permission of the authors)

## Observation and Imitation

The ability to imitate the gestures of others, either unconsciously (e.g. as in yawning or in posture, such as crossing the legs) or consciously (e.g. when we imitate the master's gesture to successfully learn to play tennis) is strongly based on the MNS. The imitative ability of yawning, for example, has been investigated by Usui et al. (2013) in a study in which children with autistic spectrum disorder (ASD) and/

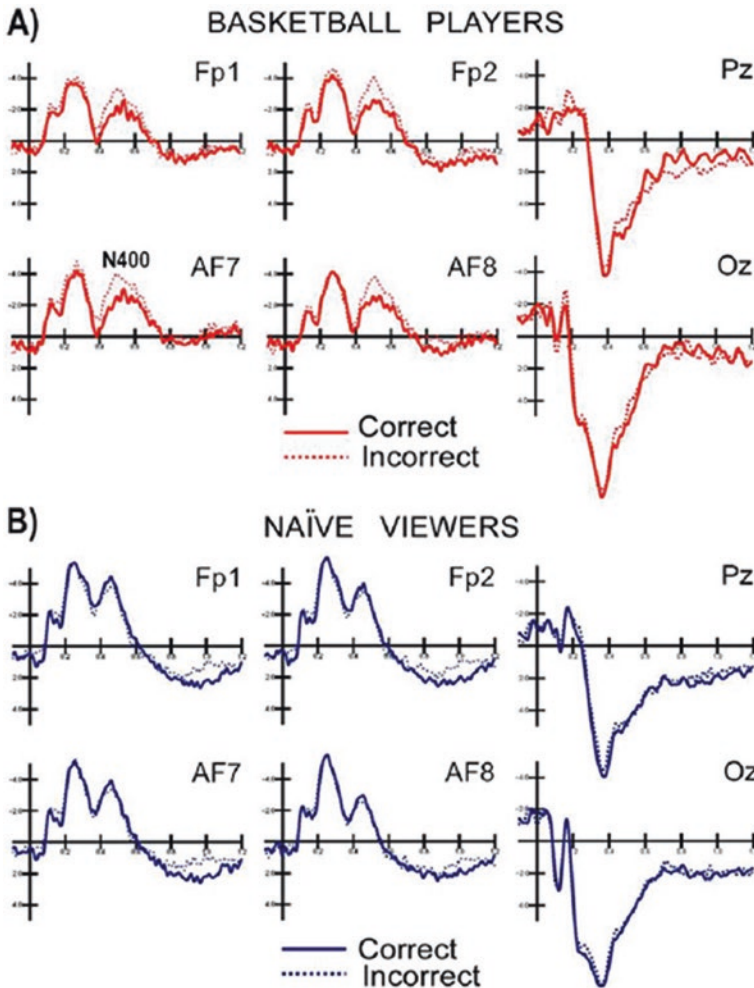
or typically developing children were shown yawning (i.e. the face of a yawning woman) vs. control frames (i.e. the face of a smiling woman) while watching a cartoon. To ensure participants' attention to the face, an eye tracker controlled the onset of the yawning and of the control stimuli. Results demonstrated that both ASD and control children yawned more frequently when they watched the yawning stimuli than the control stimuli (without any significant group differences). It was therefore suggested that the absence of contagious yawning in children with ASD, as reported in previous studies, might have been related to their weaker tendency to spontaneously attend to others' faces.

The link between action production and observation has also been explored in 'automatic imitation' or 'visuomotor priming' paradigms, where participants perform an action that is either congruent or incongruent with an observed movement. If action observation and action production employed shared mechanisms (namely, mirror neurons, Iacoboni, 2005), performing an action that is compatible with the observed action should lead to facilitation, while performing an action that is incompatible with the observed action should result in an interference effect. This pattern of results has been widely documented. For example, Craighero et al. (1996) primed healthy subjects, while ready to execute a grasping movement, by visually presenting them with drawings irrelevant to the task to be executed. Drawings visually congruent with the object to be grasped markedly reduced the response times, thus facilitating grasping actions, and vice versa. This study provided one of the first evidences for the existence of a visuomotor priming.

When we observe others, the motor and sensorimotor systems are activated to process and simulate the observed gesture. This activation induces the desynchronization of EEG *mu* rhythm (i.e. an oscillation rhythm of 8–12 Hz with a central-parietal topographic distribution over the scalp) reflecting a state of relative inactivity of the Rolandic region, a kind of *stand-by* from the motor or somatosensory processing. Therefore, its desynchronization indicates an activation of the neurons of this same area, committed to coding an observed or performed action, and can be used to measure MNs activity in both human adults (Pfurtscheller et al., 2006) and infants (Nyström et al., 2011). For example, Proverbio (2012) provided evidence that watching manipulable objects automatically activates their motor properties as indexed by the EEG desynchronization of *mu* rhythm over centro-parietal scalp sites during perception of tools vs. non-manipulable objects. Other studies have shown a lack of reduction of event-related beta and mu desynchronization (ERD) in ASD children during perception of actions, as opposed to comparable ERD responses during action execution (Oberman et al., 2008). Interestingly, Van Elk et al. (2008) showed that the longer is the motor experience of infants with crawling, the stronger is the *mu* rhythm desynchronization during observation of other children's crawling. This piece of findings indicates that experience strongly modulates MNS responsivity. As proof of this, it has been shown that the skills acquired in a certain athletic or sporting discipline, or, for instance, in dance, strongly modulates MNS responsivity. Proverbio et al. (2012) compared EEG/ERP signals relative to the visual processing of actions that violated basketball rules (e.g. in defense, blocking, and shooting actions) with that of correct basketball actions in professional



basketball players and controls. They found that incorrect actions elicited anterior N400 responses reflecting the automatic detection of action incorrectness only in professional players (see ERP waveforms in Fig. 5.5). According to source reconstruction, N400 generators included the fronto-parietal MNS, the cerebellum, the EBA, and the STS. Similarly, the detection of incorrect dance gestures has been shown to elicit a response in the fronto/parietal MNS circuits in professional dancers vs. controls (Calvo-Merino et al., 2005; Orlandi et al., 2017).



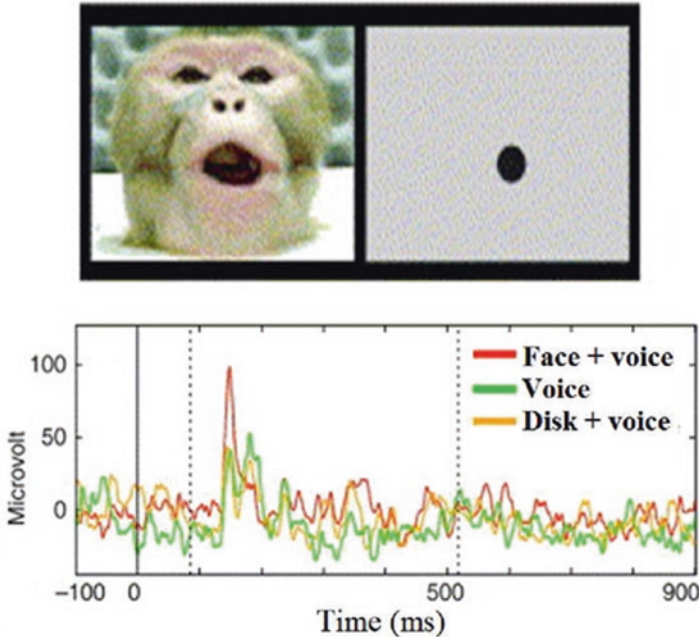
**Fig. 5.5** Grand-average ERPs recorded in professional basketball players (a) and naïve viewers (b) in response to correct and incorrect basketball actions at frontal, parietal, and occipital scalp sites. (Taken and redrawn from Proverbio et al., 2012)

## Audio-Visuomotor Neurons

The existence of multimodal audiovisual cortical regions has been demonstrated both for phonetic/articulatory language (i.e. verbal language) and for human and animal vocalizations (e.g. a chirp, a whinny, a cry, a laughter), as well as for encoding of noises typically produced using objects (e.g. the noise produced by crushing nuts, or by chewing). These multimodal neurons are a particular class of MNs that encode both visual and auditory information.

### *Audio-Visuomotor Neurons in Language and Vocalizations*

The existence of a link between motor and perceptual representations of language has been since long demonstrated. According to Liberman's theory (Liberman & Mattingly, 1985), knowing how to understand a phoneme would strictly correspond to how to pronounce it. For example, in a fMRI study on healthy subjects, Pulvermüller and Shtyrov (2006) found that, while listening to bilabial (/ p /) and dental occlusive phonemes (/ t /), simultaneous activations were observed of both auditory areas of the temporal lobe (for understanding) and of the precentral motor areas (for production), with a difference in the locus of activation depending on the processed phoneme: at the motor representation of the lips, for / p / e of the language for / t /. Fadiga et al. (2002) recorded motor evoked potentials (MEPs) from the muscles of the tongue in participants who had been asked to listen to acoustic stimuli. These stimuli consisted of words or pseudowords containing a double / f / (e.g. *bafoo* (i.e. *moustache*) in Italian) or a double / r / (e.g. *birra* (i.e. *beer*) in Italian) and bitonal sounds. The / f / is a labiodental consonant that, for being pronounced, does not require a particular involvement of the tongue, while the / r / is a linguo-palatal consonant that involves a marked involvement of the tongue for its pronunciation. The results of the experiment showed that listening to words and pseudowords containing the double / r / resulted in a significant increase of the MEPs, compared to the case of bitonal sounds, words, and pseudowords containing the double / f /. As a whole, these data demonstrated that, in humans, a MNS would exist dedicated to the comprehension of linguistic sounds (i.e. an *echo mirror system*): when an individual listens to verbal stimuli, an automatic activation would occur of motor centers responsible for the emission of the phonemes present in the words heard. These data are highly consistent with other findings deriving from fMRI investigations. Wright et al., (2003) evaluated whether speech accompanied by both auditory and visual information (as it normally does), induced a higher activation of STS, compared to speech associated only with mono-sensory information. In this study, the volunteers watched an actor speaking in three different conditions: audiovisual speech, auditory speech, and visual speech. The STS was strongly activated in all conditions, but, above all, and in a super additive way, in the audiovisual condition; apparently, these results confirmed the multisensory nature of the STS.



**Fig. 5.6** (Above) Examples of stimuli used for the study on neurons sensory preference (i.e. the face of a conspecific emitting a vocalization vs. the opening and closing of a disc without any facial stimulus). (Below) Bioelectrical responses displayed by a multisensory cell of the associative auditory cortex of the macaque monkey. Note that the response to the combined voice and face conditions (red line) is far superior than the uni-sensory stimulation (in this case, the response to the incongruous coupling between disk and voice that did not stimulate the cell enough is also drawn as a yellow line). (Adapted from Ghazanfar and Schroeder (2006). Courtesy of the authors)

A very similar, but more direct demonstration of the existence of audio visuomotor neurons derives from a single-cell recording, neurophysiological study carried out by Ghazanfar and Schroeder (2006). The authors identified neurons in the STS that not only responded to faces or voices but also exhibited a far greater responsiveness to the audiovisual association, thus demonstrating their multisensory specialization (Fig. 5.6).

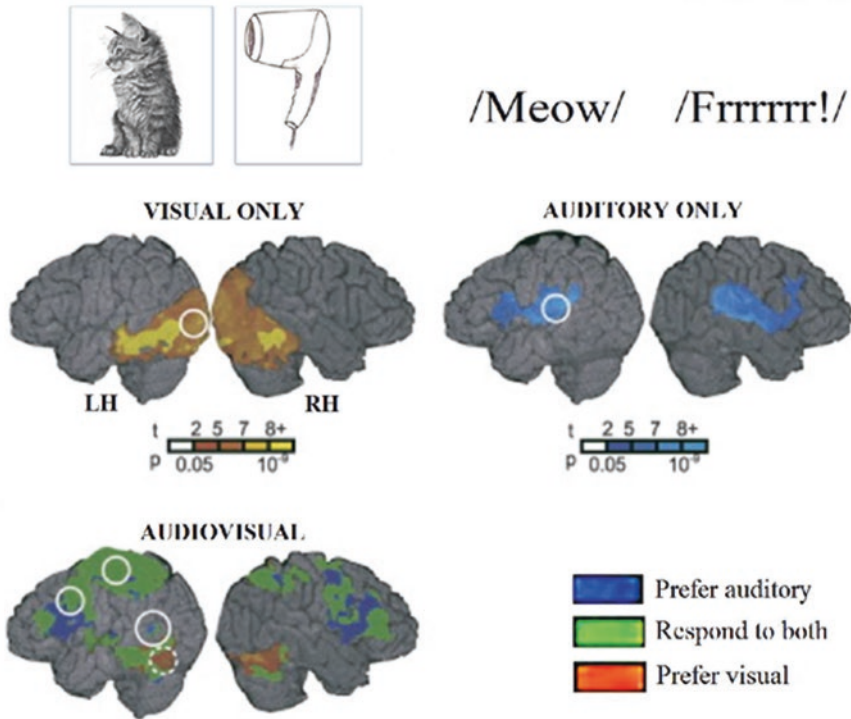
## Audio-Visuomotor Neurons and the Sound of Objects

In a famous study by Kohler et al. (2002), published in *Science*, it was demonstrated that the brain retains specific neural representations of the actions performed on objects (e.g. beating eggs, hammering) and of the sounds typically produced by their use. Congruently, the research group coordinated by Giacomo Rizzolatti discovered neurons in the PMv of the macaque monkey that ‘fired’ both when the animal performed a specific action and when it only heard its sound. Most neurons

also fired when the monkey simply watched an action. These audiovisual MNs encoded the actions regardless of whether they were performed, listened to, or simply seen; altogether, these observations led to the discovery of the audio-visuo-motor MNs. Besides the PMv cortex, hosting the audio-visuo-motor MNs, there are interesting audiovisual neurons that conjointly encode the objects and the sound they produce (which, of course, reveal of fundamental importance for music learning and for the regulation of sensory feedback). Many neuroimaging studies have long shown the existence of multisensory audiomotor neurons in the posterior region of the STS and in the middle temporal gyrus (MTG) that respond to the sounds and visual images of objects and animals. The data showed how these regions are activated more strongly by audiovisual stimuli than by uni-sensory stimuli, thus suggesting the crucial role of these regions in the multi-sensory integration of inputs coming from the two modalities (see, for instance, Beauchamp et al., 2004a, b, and Tranel et al., 2003). For instance, Beauchamp et al. (2004a, b) explored how the brain integrated visual and auditory information related to familiar animals and objects, presenting them individually or in association with each other, by means of fMRI scanning of cerebral activity in a sample of participants. Their findings clearly showed the existence of multisensory systems simultaneously encoding visual and auditory features linked to an action, such as a phonatory gesture of an animal or the manipulation of tools (see Fig. 5.7).

Because of the repeated association between an object and its typical sound, and of the fact that the brain represents the so-called *object-sound knowledge*, we can activate the image of a sound based on the object's view. It is for this reason that a musician can visually recognize the sound associated with a gesture or knows how to predict the sound that will be emitted, before it is played, observing, for instance, the tension of the hair of a bow, the position of the fingers on a keyboard, or the key pressed down.

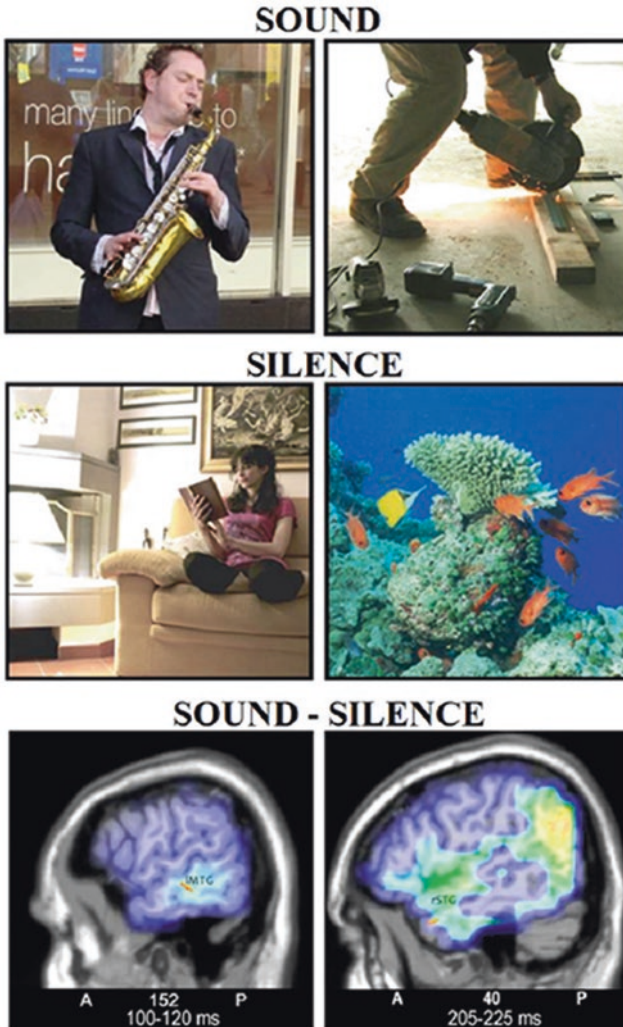
In an electrophysiological study by Proverbio et al. (2011b), it was shown that the only view of objects or actions associated with a sound can activate brain temporal cortex, a region overseeing auditory perception. In this study, high-density ERPs were recorded in 15 students who were required to look at hundreds of images associated with a given sound or to silence (see Fig. 5.8 for some examples of stimuli). ERP signals analysis showed that, despite stimulation being only visual, sound-related stimuli were distinguished from non-sound-related stimuli already after only 110 milliseconds post-stimulus processing. According to the authors, this happened because perception and recognition of objects, agents, and stimulus-contexts stimulated the access to conjoined auditory information. Indeed, as it was well known to silent movies filmmakers, there is no need for a real auditory stimulus to activate the sensation of hearing sounds typically associated with what we are seeing: This is how in a silent movie you will almost hear the whistle of the steam train or his rattling on the tracks.



**Fig. 5.7** Visual stimulation consisted in the silent presentation of pictures of animals and tools while the auditory stimulation consisted of the blind presentation of their verse or typical sound. The audiovisual stimulation involved the integration between the two modes. Brain images show the BOLD signals of neurometabolic activation obtained by fMRI in the various stimulation conditions. Note that the audiovisual condition activated the multimodal prefrontal regions, as well as the motor and premotor cortices, the posterior region of the STS, and the MTG. (Drawn and modified by Beauchamp et al. (2004a, b). Courtesy of the authors)

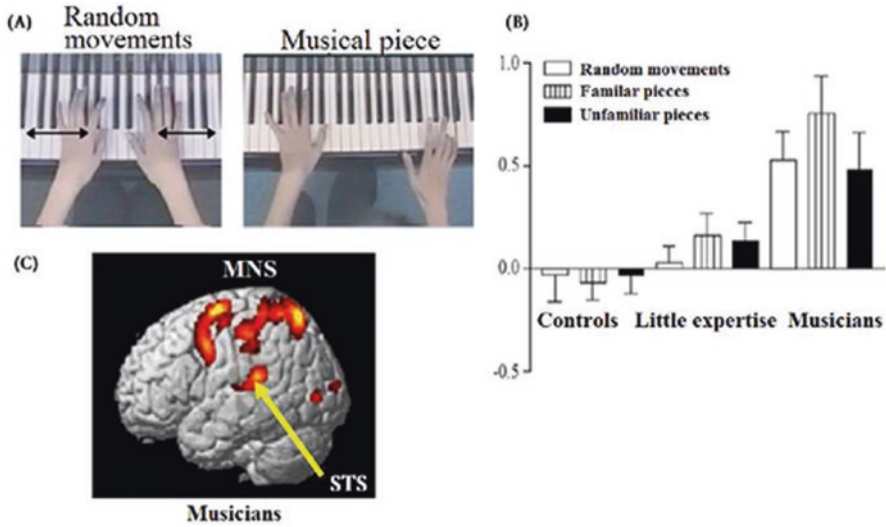
### *Audio-Visuomotor Neurons in the Coding of Musical Actions and Sounds*

While investigating how professional pianists could identify the musical piece performed in silent scenes by looking at the movements of the musicians' hands on the keys (i.e. looking at actions performed on objects), Hasegawa et al. (2004) hypothesized that visuomotor representation of musical gestures was strictly associated with the auditory representation following a specific learning. In this study, seven participants without any musical experience (control group), ten participants with some experience of the piano (not very experienced), and nine professional pianists were tested. During fMRI scanning, the participants observed silent videos showing bimanual movements of a pianist pressing the keys of a piano keyboard (Fig. 5.9a: Right), or, in a basic condition, only random, sliding across keyboard, key touches



**Fig. 5.8** Some examples of 'sound' (top) and "silent' (centre) visual stimuli presented together with other hundreds of stimuli to unaware observers, instructed to detect and respond to infrequent images of cycling races. The analysis of ERP peaks, together with the reconstruction of their intracerebral generators by means of the swLORETA technique, demonstrated the activation of the left medial temporal cortex after only 110 ms from the presentation of the image. The extraction of sound information associated with the use of familiar tools after ~200 ms activated the primary (BA38) and secondary (BA41) auditory cortices. This information is responsible, for example, for auditory hallucinations, which, in this case, refer, in a dim way, to the call of the specific sound produced by the tool (in the figure, the sounds produced by the sax or by the infernal chainsaw). (Taken from Proverbio et al. (2011b). Courtesy of the authors)

(Fig. 5.9a: Left). Pressure movements could be completely random, that is, not at all combined with a musical piece or related to the execution of a more or less famous piece. Professional pianists were able to identify these pieces, but, above all, the



**Fig. 5.9** (a) Examples of visual stimuli used in the study by Hasegawa et al. (2004) (b) Activation of the left temporal region as a function of musical performance in the three groups of participants. (c) fMRI activations in response to an exclusively visual stimulation in the brain of professional pianists. (Courtesy of the authors)

view of the musical performance – regardless of the piece – activated their fronto-parietal MNS (i.e. motor simulation) and STS, thus demonstrating that seeing familiar musical gestures activates the stored memory of the associated sounds, but only in those who actually know how to perform them. This study clearly demonstrated the role of audio-visuomotor neurons in musical learning (Paraskevopoulos et al., 2012; Schulz et al., 2003).

A similarly interesting study on audio-visuomotor coding is the one carried out by Lahav et al. (2007). In this study, naïve participants (i.e. non-musicians) were trained to play a short musical sequence by ear. Their cerebral activity was then tested by means of fMRI while they listened to the newly learned piece. The authors found that, despite the participants not making any kind of movement while listening, both motor and mirror regions were activated, including the bilateral frontoparietal motor circuit, along with the IFG and the PMv, the IPS and the IPG. Moreover, the presentation of the same musical notes organized in a different order, activated in a much less measure the same regions, whereas listening to a familiar musical sequence whose motor program was unknown, did not activate these regions at all. These data supported the hypothesis of the existence of a “*hearing-doing*” (or “*hearing-action*”) system, strongly dependent on the individual’s motor’s repertoire. In this regard, with a study combining Transcranial Magnetic Stimulation (TMS) and MEP recordings, Candidi et al. (2014) showed that, in expert pianists, the observation of a piano fingering error – a visual gesture shown without any audio – induced a significant motor effect, and in particular a somatotopic cortico-spinal facilitation concerning the finger of the hand engaged in the fingering error.

Together, the studies described above demonstrated how learning of skilled gestures characterized by a complex timing applied to a given musical instrument (or to a vocal performance) occurs through the progressive and long-term association between motor, somatosensory, and auditory functional patterns, namely through a substantial audio-visuomotor coding of the musical gesture, which takes many years.

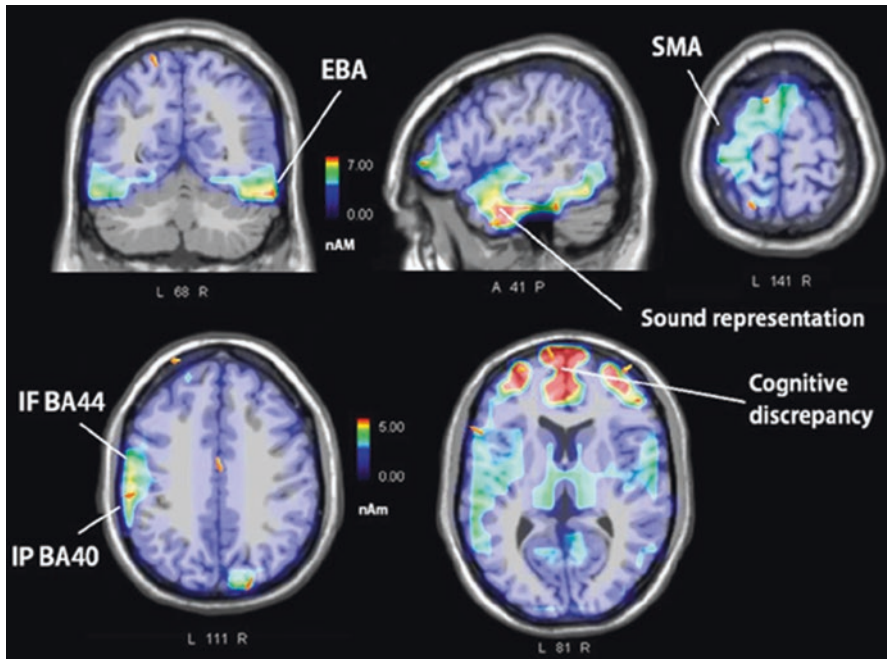
A cross-sectional study by Proverbio et al. (2015b) investigated how the representation of musical sounds changed as a function of the years of study in relation to the motor gesture necessary to produce these sounds. This study considered the development of audio-visuomotor mirror systems in young students going from the second year of study course up to the master and beyond. In all, 19 music students were tested: 10 violinists and 9 clarinetists. Their chronological age ranged from 14 to 24 years, while their academic practicing of their instrument ranged from 2 to 18 years. These students (recruited in their instrument classes while waiting to attend a lesson) watched – on a PC screen – and listened – by means of headphones – a total of 400 video clips of professional violinists and clarinetists who played non-melodically 200 totally new combinations of double or single notes that covered all sound heights. Their task was simply to indicate the possible congruence between the gesture and the sound reproduced in each video clip on the basis of their senses. Half of the time, in fact, the sounds were not congruent with the motor gestures but were mounted onto the video track in an incongruous although perfectly synchronized way. The data showed that the actual years of study at the Conservatory correlated directly with the performance in the task. It was as if the more advanced students had so firmly internalized the connection between sound, gesture, and image that they automatically perceived a possible incongruity, with a percentage of error that decreased linearly as the years of practice increased. This happened thanks to the ability of multimodal neurons to create audio-visuomotor correlations that increased with the years of study and practice, regardless of the talent and age of the individual. The first effects of cerebral modification were observable after 4–6 years of intensive study and progressively continued after graduation and master's degree. Up to three years of study, the percentage of error was close to 50%, while only after obtaining the diploma (and about 10,000 h of study), the percentage fell below 10% for music teachers. This research highlighted the crucial role of exercise in shaping brain musical functions, regardless of musical talent.

The same stimuli of the study described above were shown to 12 professional musicians and 12 naïve university students to study in real-life neural mechanisms of audio-visuomotor coding of the musical gesture for their instrument and/or for an unfamiliar instrument (Proverbio et al., 2014). While the musicians watched the stimuli, they had to decide whether the note played was double or single – an easily resolvable task – not only for their instrument but even for an unfamiliar musical instrument. Throughout the task duration, their EEG was recorded in continuous mode by means of 128 sensors placed all over their scalp. Averaged ERPs indicated that audiovisual incongruity generated a prominent N400 mismatch response for the musicians' own instrument only, since it appeared almost impossible for these subjects to reach robust decisions for the unfamiliar instrument. The swLORETA



applied to the N400 response identified the areas mediating multimodal motor processing: the prefrontal cortex (PFC: attention, cognitive discrepancy), the superior and middle temporal gyri (STG and MTG: auditory coding of sound), the premotor cortex (PM: motor programming, simulation), the inferior frontal and parietal areas (IF and IP, mirror system), the extrastriate region for coding of body parts (EBA), the somatosensory cortex (maps of the fingers and the hand), the cerebellum (motor coordination), and the supplementary motor area (SMA), which encodes the learned motor sequences (Fig. 5.10). In conclusion, these data indicate the existence of audio-visuomotor MNs that respond to both visual and auditory incongruent information, thus suggesting that they can encode multimodal learned motor skill representations of musical gestures and sounds.

In summary, we have reviewed a wide neuroimaging and electrophysiological literature reporting the involvement of visuomotor MNs in many mental functions including the comprehension of actions and action intentions, understanding the others' emotional and mental state, action imitation and learning, processing of visuomotor aspects of speech, vocalizations and music, developing motor or musical skills, and many others. Some criticalities still challenge the concept that the human MNs can be viewed as roughly correspondent to the monkey's MNs, for which we have direct neurophysiological recording. First of all, MNs are not always



**Fig. 5.10** Coronal, sagittal, and axial views of the standardized and weighted LOW Resolution electromagnetic Tomography (swLORETA) applied to the N400 bioelectric response generated only for one's own musical instrument. (Taken from Proverbio et al. (2014) and redrawn)

observed while recording from the fronto/parietal areas of the monkey's brain, and their incidence can be very variable, ranging from 8.9% for ventral intra-parietal areas (VIP) to 60% for premotor dorsal areas (PMD). Other criticalities concern the fact that cell-recording studies are not very numerous (also for ethical reasons) and that in humans, evidences are relatively indirect (not based on intracranial recordings). It should be also borne in mind that MNs are only indirectly involved in social and affective processes, such as empathy, contributing for the visuomotor recognition of body language and gestures only.

## References

- Beauchamp, M. S., Argall, B. D., Bodurka, J., Duyn, J. H., & Martin, A. (2004a). Unraveling multisensory integration: Patchy organization within human STS multisensory cortex. *Nature Neuroscience*, *7*(11), 1190–1192.
- Beauchamp, M. S., Lee, K. E., Argall, B. D., & Martin, A. (2004b). Integration of auditory and visual information about objects in superior temporal sulcus. *Neuron*, *41*(5), 809–823.
- Calvo-Merino, B., Glaser, D. E., Grèzes, J., Passingham, R. E., & Haggard, P. (2005). Action observation and acquired motor skills: An fMRI study with expert dancers. *Cerebral Cortex*, *15*, 1243–1249.
- Candidi, M., Sachelì, L. M., Mega, I., & Aglioti, S. M. (2014). Somatotopic mapping of piano fingering errors in sensorimotor experts: TMS studies in pianists and visually trained musically naïves. *Cerebral Cortex*, *24*(2), 435–443.
- Craighero, L., Fadiga, L., Umiltà, C. A., & Rizzolatti, G. (1996). Evidence for visuomotor priming effect. *Neuroreport*, *8*, 347–349.
- Creem-Regehr, S. H., & Lee, J. N. (2005). Neural representations of graspable objects: Are tools special? *Cognitive Brain Research*, *22*, 457–469.
- Dinstein, I. (2008). Human cortex: Reflections of mirror neurons. *Current Biology*, *18*(20), 956–959.
- Fadiga, L., Craighero, L., Buccino, G., & Rizzolatti, G. (2002). Speech listening specifically modulates the excitability of tongue muscles: A TMS study. *The European Journal of Neuroscience*, *2*, 399–402.
- Ghazanfar, A. A., & Schroeder, C. E. (2006). Is neocortex essentially multisensory? *Trends in Cognitive Sciences*, *10*(6), 278–285.
- Goldenberg, G., & Spatt, J. (2009). The neural basis of tool use. *Brain*, *132*(Pt 6), 1645–1655.
- Hasegawa, T., Matsuki, K., Ueno, T., Maeda, Y., Matsue, Y., Konishi, Y., & Sadato, N. (2004). Learned audio-visual cross-modal associations in observed piano playing activate the left planum temporale. An fMRI study. *Brain Res Cogn*, *20*, 510–518.
- Iacoboni, M. (2005). Understanding others: Imitation, language, empathy. In S. Hurley & N. Chater (Eds.), *Perspectives on imitation: From cognitive neuroscience to social science* (Vol. 1, pp. 77–99). MIT Press.
- Iacoboni, M., Molnar-Szakacs, I., Gallese, V., Buccino, G., Mazziotta, J. C., et al. (2005). Grasping the intentions of others with one's own mirror neuron system. *PLoS Biology*, *3*(3), e79.
- Kohler, E., Keysers, C., Umiltà, M. A., Fogassi, L., Gallese, V., & Rizzolatti, G. (2002). Hearing sounds, understanding actions: Action representation in mirror neurons. *Science*, *297*(5582), 846–848.
- Lahav, A., Saltzman, E., & Schlaug, G. (2007). Action representation of sound: Audiomotor recognition network while listening to newly acquired actions. *The Journal of Neuroscience*, *27*, 308–314.

- Lelord, G., Cochin, S., Adrien, J. L., Barthelemy, C., & Martineau, J. (1998). Latent imitation of human movements presented on a videoscopic screen, disclosed by electroencephalographic mapping in the spectator. *Bulletin de L'Academie Nationale de Medecine*, *182*(4), 833–42, Discussion 843–4.
- Liberman, A. M., & Mattingly, I. G. (1985). The motor theory of speech perception revised. *Cognition*, *21*, 1–36.
- Nyström, P., Ljunghammar, T., Rosander, K., & von Hofsten, C. (2011). Using mu rhythm desynchronization to measure mirror neuron activity in infants. *Developmental Science*, *14*, 327–335.
- Oberman, L. M., Ramachandran, V. S., & Pineda, J. A. (2008). Modulation of mu suppression in children with autism spectrum disorders in response to familiar or unfamiliar stimuli: The mirror neuron hypothesis. *Neuropsychologia*, *46*(5), 1558–1565.
- Orlandi, A., Zani, A., & Proverbio, A. M. (2017). Dance expertise modulates visual sensitivity to complex biological movements. *Neuropsychologia*, *104*, 168–181.
- Paraskevopoulos, E., Kuchenbuch, A., Herholz, S. C., & Pantev, C. (2012). Musical expertise induces audiovisual integration of abstract congruency rules. *The Journal of Neuroscience*, *32*, 18196–18203.
- Pfurtscheller, G., Brunner, C., Schlögl, A., & Lopes da Silva, F. H. (2006). Mu rhythm (de)synchronization and EEG single-trial classification of different motor imagery tasks. *Neuro Image*, *31*, 153–159.
- Proverbio, A. M. (2012). Tool perception suppresses 10–12Hz  $\mu$  rhythm of EEG over the somatosensory area. *Biological Psychology*, *91*(1), 1–7.
- Proverbio, A. M., Riva, F., & Zani, A. (2010). When neurons do not mirror the agent's intentions: Sex differences in neural coding of goal-directed actions. *Neuropsychologia*, *48*(5), 1454–1463.
- Proverbio, A. M., Adorni, R., & D'Aniello, G. E. (2011a). 250 ms to code for action affordance during observation of manipulable objects. *Neuropsychologia*, *49*(9), 2711–2717.
- Proverbio, A. M., D'Aniello, G. E., Adorni, R., & Zani, A. (2011b). When a photograph can be heard: Vision activates the auditory cortex within 110 ms. *Scientific Reports*, *1*, 54.
- Proverbio, A. M., Crotti, N., Manfredi, M., Adorni, R., & Zani, A. (2012). Who needs a referee? How incorrect basketball actions are automatically detected by basketball players' brain. *Scientific Reports. Nature*, *2*, 883; <https://doi.org/10.1038/srep00883>.
- Proverbio, A. M., Azzari, R., & Adorni, R. (2013). Is there a left hemispheric asymmetry for tool affordance processing? *Neuropsychologia*, *51*(13), 2690–2701.
- Proverbio, A. M., Calbi, M., Manfredi, M., & Zani, A. (2014). Audio-visuomotor processing in the musician's brain: An ERP study on professional violinists and clarinetists. *Scientific Reports*, *29*(4), 5866.
- Proverbio, A. M., Calbi, M., Manfredi, M., & Zani, A. (2014a). Comprehending body language and mimics: an ERP and neuroimaging study on Italian actors and viewers. *PLoS One*, *9*(3), e91294.
- Proverbio, A. M., Gabaro, V., Orlandi, A., & Zani, A. (2015a). Semantic brain areas are involved in gesture comprehension: An electrical neuroimaging study. *Brain and Language*, *147*, 30–40.
- Proverbio, A. M., Attardo, L., Cozzi, M., & Zani, A. (2015b). The effect of musical practice on gesture/sound pairing. *Frontiers in Psychology, Auditory Cognitive Neuroscience*, *6*, 376.
- Pulvermüller, F., & Shtyrov, Y. (2006). Language outside the focus of attention: The mismatch negativity as a tool for studying higher cognitive processes. *Progress in Neurobiology*, *79*, 49–71.
- Rizzolatti, G., & Craighero, L. (2004). The mirror-neuron system. *Annual Review of Neuroscience*, *27*, 169–192.
- Rizzolatti, G., & Luppino, G. (2001). The cortical motor system. *Neuron*, *31*(6), 889–901.
- Rizzolatti, G., & Sinigaglia, C. (2016). The mirror mechanism: A basic principle of brain function. *Nature Reviews. Neuroscience*, *17*(12), 757–765.
- Rizzolatti, G., Fadiga, L., Matelli, M., Bettinardi, V., Paulesu, E., Perani, D., & Fazio, F. (1996). Localization of grasp representations in humans by PET: 1. Observation versus execution. *Experimental Brain Research*, *111*, 246–252.

- Schmidt, S. N. L., Hass, J., Kirsch, P., & Mier, D. (2021). The human mirror neuron system – a common neural basis for social cognition? *Psychophysiology*, *58*, e13781.
- Schulz, M., Ross, B., & Pantev, C. (2003). Evidence for training-induced crossmodal reorganization of cortical functions in trumpet players. *Neuroreport*, *14*(1), 157–161.
- Tranel, D., Damasio, H., Eichhorn, G. R., Grabowski, T., Ponto, L. L., & Hichwa, R. D. (2003). Neural correlates of naming animals from their characteristic sounds. *Neuropsychologia*, *41*(7), 847–854.
- Usui, S., Senju, A., Kikuchi, Y., Akechi, H., Tojo, Y., Osanai, H., & Hasegawa, T. (2013). Presence of contagious yawning in children with autism spectrum disorder. *Autism Research and Treatment*, *2013*, 971686.
- van Elk, M., van Schie, H. T., Hunnius, S., Vesper, C., & Bekkering, H. (2008). You'll never crawl alone: Neurophysiological evidence for experience-dependent motor resonance in infancy. *NeuroImage*, *43*(4), 808–814.
- Wright, T. M., Pelphrey, K. A., Allison, T., McKeown, M. J., & McCarthy, G. (2003). Polysensory interactions along lateral temporal regions evoked by audiovisual speech. *Cerebral Cortex*, *13*, 1034–1043.

**Open Access** This chapter is licensed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.

