

Representing actions through their sound

Salvatore M. Aglioti · Mariella Pazzaglia

Received: 8 March 2010 / Accepted: 18 June 2010
© Springer-Verlag 2010

Abstract Since the discovery of ‘mirror neurons’ in the monkey premotor and parietal cortex, an increasing body of evidence in animals and humans alike has supported the notion of the inextricable link between action execution and action perception. Although research originally focused on the relationship between performed and viewed actions, more recent studies highlight the importance of representing the actions of others through audition. In the first part of this article, we discuss animal studies, which provide direct evidence that action is inherently linked to multi-sensory cues, as well as the studies carried out on healthy subjects by using state-of-the-art cognitive neuroscience techniques such as functional magnetic resonance imaging (fMRI), event-related potentials (ERP), magnetoencephalography (MEG), and transcranial magnetic stimulation (TMS). In the second section, we review the lesion analysis studies in brain-damaged patients demonstrating the link between ‘resonant’ fronto-parieto-temporal networks and the ability to represent an action by hearing its sound. Moreover, we examine the evidence in favour of somatotopy as a possible representational rule underlying the auditory mapping of actions and consider the links between language and audio-motor action mapping. We conclude with a discussion of some outstanding questions for future research on the link between actions and the sounds they produce.

Keywords Fronto-parietal · Mirror circuits · Sound action · fMRI · TMS · Language

Introduction

Perception and representation of actions are a multimodal experience modulated by visual (Haslinger et al. 2005), auditory (Kaplan et al. 2007), somatic (Avenanti et al. 2007), and even olfactory inputs (Rossi et al. 2008). The discovery in the monkey brain of double-duty neurons involved in action execution and action observation (visuo-motor mirror neurons) provided a new insight into the possible parity between action perception and action execution (Buccino et al. 2001; Decety et al. 1997; Fogassi et al. 2005; Gallese et al. 1996; Gazzola and Keysers 2009; Rizzolatti et al. 1996). Studies indicate, for example, that mere observation of an action may strengthen the motor representation of the observed action (Stefan et al. 2005), and that mere motor experience of a particular action may improve its visual discrimination (Casile and Giese 2006). In view of these findings and possibly also of the notion that vision dominates the other senses and leads virtually any type of cross-modal and perceptuo-motor interaction, several studies have focused on the link between action observation and action execution (Buccino et al. 2004; Pazzaglia et al. 2008b; Rizzolatti et al. 1996; Vogt et al. 2007). It is noteworthy that our experience with the world is typically mediated by multiple sensorial experiences and that the sight of an action is usually influenced by the sound (Alaerts et al. 2009) and the somatic experience associated with it (Avenanti et al. 2007; Costantini et al. 2005; Romani et al. 2005). However, studies on cross-modal perception indicate that haptic perception (Gori et al. 2008) may be better than vision and that auditory perception

S. M. Aglioti (✉) · M. Pazzaglia
Dipartimento di Psicologia, Università di Roma ‘La Sapienza’,
Via dei Marsi 78, 00185 Rome, Italy
e-mail: salvatoremaria.aglioti@uniroma1.it

S. M. Aglioti · M. Pazzaglia
IRCCS Fondazione Santa Lucia, Via Ardeatina 306,
00179 Rome, Italy

(Shams et al. 2000) may predominate over vision in specific circumstances. More importantly, non-visual senses are fundamental for interacting with the environment when vision is not available. In addition to conditions when sight is absent (Amedi et al. 2005; Pietrini et al. 2004; Ricciardi et al. 2009), several daily life behaviours are guided by non-visual senses. Consider, for example, the importance of tactuo-motor interactions involved in searching for keys in a bag or audio-motor interactions involved when walking along a dark street and reacting to the sound of footsteps that signals someone's approach. Each action of animals and humans produces a characteristic sound that may permit its unequivocal recognition. For example, hearing hands clapping may allow an individual to draw several inferences about a given event. Similarly, hearing the footsteps of an individual may provide precise (e.g., specific identity) as well as general (e.g., sex or mood) information about an individual (Bidet-Caulet et al. 2005; Ekimov and Sabatier 2006; Saarela and Hari 2008). Of great importance to human interactions are oro-facial movements, fundamental for emotional and non-emotional vocalizations (Warren et al. 2006) as well as for language production. Therefore, in principle, hearing a particular emotional (Sauter and Eimer 2010; Van den Stock et al. 2008) or voice sound may trigger the representation of the individual producing the heard sound (Belin 2006; Campanella and Belin 2007).

Here, we review the studies revealing that it is not only the viewing of an action, but also the hearing of the sound that it produces, which triggers the representation of the same action in both animals and humans. Moreover, we also provide data from brain-damaged individuals; these data hint at the use of specific rules for representing actions via audition. Finally, we discuss outstanding questions for future research on the auditory mapping of actions.

Mechanisms and neural underpinnings of audio-motor action mapping

Direct evidence in animals

The classical view that posterior and anterior brain areas are involved in sensory and motor processes, respectively, has been challenged by pioneering single-cell recording studies on the brain of macaque monkeys who observed the actions performed by monkey or human agents ('mirror neurons', see review in Rizzolatti and Craighero 2004). These studies reveal that viewing the actions performed by another agent may trigger the activity of the frontal and parietal cortical neurons, which are also involved in both the execution and planning of goal-directed actions (di Pellegrino et al. 1992; Fogassi et al. 2005; Gallese et al. 1996). The existence of

double-duty 'mirror' neurons in classical motor areas indicated, for the first time, that a common neural substrate may underlie the perceptual and motor aspects of actions. Remarkably, some of these double-duty, frontal mirror neurons in fact turned out to be triple-duty cells, activated not only by action observation and execution but also by the acoustic perception of action-related sounds (Kohler et al. 2002). Moreover, these cells can display a multimodal additive response, suggesting they are involved in complex audio-visuo-motor integration functions (Keysers et al. 2003). Note, however, that some of these complex perceptuo-motor neurons respond strongly to action sounds (e.g., paper tearing), but not to the observation of actions (e.g., grasping of an object) performed by another human or monkey agents thus indicating their audio-motor specificity (Keysers et al. 2003). Although mirroring was originally thought of as an emergent property of the primate brain, recordings of neural activity from the telencephalic nucleus in the forebrain circuit that mediates singing (the caudal part of the ventral hyperstriatum, HVC) in swamp sparrows revealed that neurons whose frequency discharge increased during song production were the same as those whose frequency discharge increased when the birds heard a similar song produced by a conspecific (Prather et al. 2008). Swamp sparrows typically have several songs in their repertoire, and when they sing, a neuron may fire at different points in time, for example, at the onset of a certain sequence of notes. When the bird listens to its own song through a speaker, the same neuron will fire whenever this particular sequence appears. The same neuron may also fire in response to another swamp sparrow's song, but such firing is observed only when the other bird's song contains a similar sequence of notes. Therefore, there is a remarkable resemblance between the precise HVC auditory-vocal correspondence neurons in birds and 'audio-motor' mirror neurons in the fronto-parietal cortex in monkeys suggesting that mirroring may be a fundamental brain process that is not limited to the primate brain. It is noteworthy that the HVC nucleus is a crucial node in the circuit for song learning, suggesting that audio-motor mirroring may provide feedback for vocal learning. These studies may contribute to the understanding of the relationship between sounds, gestures, and language in humans, an issue that is considered later in this article.

Action representation triggered by action-related sounds in humans

Transcranial magnetic stimulation studies

Although direct evidence of the existence of mirror neurons in humans is still lacking, increasing numbers of studies have provided indirect proof of motor mirroring in

healthy and brain-damaged individuals. In a seminal study on this issue, single-pulse transcranial magnetic stimulation (TMS) was delivered to the primary motor cortex of healthy subjects while they observed a model grasping an object. The study revealed that mere observation of the action induced an increase in motor evoked potentials (MEPs) recorded from the same muscles that would be active during actual execution of the same movement (Fadiga-effect, for short; Fadiga et al. 1995). This increase in MEP modulation during action observation was expanded by subsequent single-pulse TMS studies (Candidi et al. 2010; Gangitano et al. 2001; Romani et al. 2005; Strafella and Paus 2000; Urgesi et al. 2006, 2010) and by the more sophisticated combination of single pulse and repetitive TMS ('perturb-and-measure' approach), which allowed us to obtain information on the influence of different cortical regions on the motor mirroring contingent upon action viewing (Avenanti et al. 2007). In addition, listening to action-related sounds seems to increase cortico-spinal reactivity. Indeed, MEPs recorded from hand muscles during passive listening to sounds associated with bimanual actions produced greater cortico-spinal facilitation than control sounds or sounds associated with leg movements (Aziz-Zadeh et al. 2004). Interestingly, this effect was only observed when stimulating the left, language-dominant hemisphere, supporting the existence of a close link between action and language. While no TMS studies have explored cortico-spinal reactivity induced by the non-speech-related human action sounds produced by mouth movements, a clear motor activation when hearing speech sounds (e.g., phonemes) has been reported. For instance, Fadiga et al. (2002) reported an increase in the motor evoked potentials recorded from the tongue muscles of healthy individuals who heard words that involve tongue movements when uttered. In a similar vein, cortico-spinal activity in the lip muscles was recorded in subjects listening to speech and non-verbal sounds, viewing speech-related lip, and viewing eye and brow movements (Watkins et al. 2003). A clear increase in MEP amplitude was found only when hearing and viewing speech, and only after stimulation of the left hemisphere. Further, the lack of amplitude modulation of MEPs recorded from the right hand indicates the specificity of such an effect (Watkins et al. 2003). Moreover, a clear change in the cortico-spinal representation of the lips and tongue was found during the discrimination of lip- and tongue-articulated phonemes, respectively (D'Ausilio et al. 2009). This clear-cut neuro-functional double dissociation in speech sound discrimination supports the classical motor speech perception theory (Liberman and Mattingly 1985). Finally, causative evidence of the direct involvement of the motor cortex in speech perception was provided by a study where repetitive TMS was used to interfere with the representation of the lip

or the hand in the left primary motor cortex (Möttönen and Watkins 2009). The results reveal that the disruption of the lip motor representation specifically impaired the categorical perception of phonetic sounds that require lip movement for their articulation. No such effect was found for the disruption of hand representation, suggesting that speech production and speech perception may rely upon the same motor circuits.

Evoked potentials and Magnetoencephalography studies

The time course of action-related sound processing was explored using multi-channel event-related potentials (ERPs) in an elegant study using a visuo-auditory version of the repetition suppression paradigm, where written words could or could not be semantically related to sounds evoked by human hand (e.g., clapping), or mouth (e.g., whistling) actions, or by other non-human activities (e.g., the sound of the rain; Pizzamiglio et al. 2005). In this study, the left posterior superior temporal and premotor areas were selectively modulated (peak at 280 ms) by action-related sounds, while the left and right temporal poles were modulated (peak at 300 ms) by non-action-related sounds. This pattern of results clearly indicates that separate neural systems are used to represent sounds, which can or cannot be translated into human actions. Electrical neuroimaging analyses of auditory evoked potentials in response to sounds that typically cue a responsive action (e.g., a door bell), versus sounds that do not elicit automatic motor responses (e.g., a church bell), revealed the modulation of neural activity in the premotor and inferior (pre)frontal regions, mainly on the left side (De Lucia et al. 2009). In keeping with the study by Pizzamiglio et al. (2005), the effects peaked at about 300 ms, well after the general categorical object discrimination, thus supporting the notion that the reported electric modulation is associated with audio-motor action mapping. The dynamics of the sound-into-action translation process have also been explored by recording the mismatch negativity to deviant stimuli consisting of finger and tongue clicks as well as of sounds with comparable acoustic features but no association with actions (Hauk et al. 2006). Action-related sounds induced larger mismatch negativity than control sounds ~100 ms after the stimulus presentation. Topographical distribution analysis showed that hearing sounds related to finger and tongue actions induced higher neural activity in left hemisphere motor areas as well as in the more inferior regions of both hemispheres, respectively (Hauk et al. 2006). Therefore, an early frontal activation may underlie the process of the automatic recognition of actions from their sounds (Hauk et al. 2006). The application of the same experimental stimuli, and a paradigm of subdural

EEG recording in a 12-year-old girl undergoing intracranial monitoring of epileptic seizures, provided an accurate picture of the time course of the sound-into-action translation process (Lepage et al. 2010). Neural activity recorded from electrodes overlying the functionally defined hand representation of the motor cortex revealed both early (100 ms) and late (250–450 ms) modulation effects of natural finger-clicks when compared with control sounds. Although coming from a single subject, the results hint at the presence of two distinct time windows of M1 activation after action sound are heard and suggest that both early (Hauk et al. 2006) and late (De Lucia et al. 2009; Pizzamiglio et al. 2005) resonance may be triggered by hearing action sounds. The only magnetoencephalography (MEG) study that has specifically tested the effect of hearing action-related sounds on motor cortex activity also provides evidence of audio-motor mirroring in humans (Caetano et al. 2007). The authors of that study explored the modulation of the 20-Hz component of the mu rhythm. The 20-Hz activity is typically suppressed during movement performance and reappears soon after. This phenomenon is known as rebound-effect and reflects the stabilization of the primary motor cortex after a perturbation. A rebound of 20-Hz activity was found not only after execution and visual observation of actions, but also after hearing the sound associated with the same action, demonstrating the clear influence of vision and audition on action mapping (Caetano et al. 2007).

Functional magnetic resonance imaging studies

Clear evidence for selective involvement of brain regions in sound-into-action translation processes has been provided in the past few years by the utilization of functional magnetic resonance imaging (fMRI) studies. The first evidence for the existence of an audio-motor mirror system in humans was obtained by asking healthy subjects to passively listen to hand- or mouth-related sounds (Gazzola et al. 2006). The dorsal premotor cortex and inferior parietal lobe in the left hemisphere showed a stronger response to the sound of a hand action than to the sound of a mouth action. In contrast, in the left ventral premotor cortex, the response to the sound of a mouth action was stronger than to the sound of a limb action. Interestingly, the audio-motor mirror system largely responded to the sight of similar actions, thus hinting at the multimodal nature of action simulation. Following up on their previous evoked potential study (Pizzamiglio et al. 2005), Galati et al. (2008) performed an event-related fMRI study in which a hand or mouth action-related sound (or a sound related to an environmental event) was or was not congruent with a previously presented written word that acted as an unconscious cross-modal priming stimulus. A clear

neural signature of prime congruency specific for action sound trials was found in the left inferior frontal and posterior temporal regions, supporting the notion of a fronto-parietal network underlying audio-motor mirroring. Interestingly, the fronto-parietal network is activated during the execution, viewing, and hearing of hand movements, supporting the inherently multimodal nature of motor mirroring (Galati et al. 2008; Gazzola et al. 2006). It is noteworthy that environmental sounds not associated with human actions (Galati et al. 2008), the phase-scrambled versions of action sounds (Gazzola et al. 2006), and silent events (Lewis et al. 2006), did not activate any of these areas. Because acting upon environmental elements (e.g., grasping rain, wind, or fire) is not possible, the motor system may not be involved in the perception of the sounds related to these elements. Therefore, environmental sounds do not generally qualify as sounds that are associated with embodiment. Interestingly, in the left fronto-parietal region, responses to animal sounds, the majority of which are classified as vocalizations (Lewis et al. 2005), were significantly reduced compared to responses to tool sounds, generated by actions of the user's hands. Moreover, the response of these regions was higher when an animal sound was incorrectly judged to be a tool sound, thereby suggesting that these regions are preferentially activated by sounds that are perceived as being produced by human hand and mouth actions. Furthermore, animal sounds that do not involve any vocalization but can be mimicked by human beings (e.g., flapping arms) produced left-lateralized activation in several regions, including the inferior frontal gyrus (IFG), the supplemental motor area (SMA), and the inferior parietal lobe (IPL; Engel et al. 2009). However, unlike the sounds produced by human sources, these sounds did not produce significant activation in the fronto-parietal network (Engel et al. 2009). Thus, the perception of a sound may evoke mental simulation of actions present in the subject's motor repertoire. Further, while tool sounds (hammering) associated with hand actions had a specific left motor cortical representation (Lewis et al. 2006), tool sounds that were not associated with hand actions (e.g., sounds linked to mechanical actions, like the sound of an airplane propeller) did not activate any motor-related networks (Engel et al. 2009). One factor influencing this difference may be related to the evolution of the sound. For example, modern mechanical sounds (e.g., washing machine), which have been heard by humans for less than 30 years, are considerably less common than the sound of a hammer. Importantly, the sounds produced by hand tools are associated with an agent. Evolutionary pressures may have also driven neural systems to selective, rapid, and efficient identification of human-specific sounds.

Lesion studies

The neural activation that correlates with audio-motor mapping, as demonstrated in most of the aforementioned electrophysiological and functional neuroanatomy studies, could be purely epiphenomenal. In contrast, lesion studies can reveal brain areas or circuits actively involved in the process of deriving actions from the sound they produce. Direct evidence of the anatomical and functional association between action execution and discrimination in the case of sound-associated actions was obtained in patients with apraxia, who typically show impairments in performing specific gestures (Pazzaglia et al. 2008a). Brain-damaged patients, with or without apraxia, were requested to match specific visual pictures to previously presented sounds, which may or may not be associated with hand- or mouth-related human actions. The results revealed a clear association between deficits in performing hand- or mouth-related actions and the ability to acoustically recognize the same sounds. Moreover, the results of a voxel-based lesion symptom mapping (VLSM) analysis demonstrated that deficits in performing and matching mouth and face-related actions are mainly associated with frontal and parietal lesions (Pazzaglia et al. 2008a). Importantly, deficits in the recognition of non-human action-related sounds were independent of apraxia and fronto-parietal lesions. This finding agrees with a combined fMRI and VLSM analyses of patients with brain damage, showing that right temporal areas, and not premotor and parietal regions, were necessary for the processing environmental sounds (Dick et al. 2007).

Representational rules underlying the auditory mapping of actions

As previously discussed, different cognitive neuroscience techniques and experimental protocols in healthy subjects and brain-damaged patients have provided convergent evidence for the existence of a fronto-parietal network involved in audio-motor mapping of human actions. Audio-motor mirroring is modulated by several factors including the intention to act (Knoblich and Repp 2009), the presence of body-object interactions (Lewis et al. 2005; De Lucia et al. 2009), and the body part involved in the action evoked by the heard sound (Fadiga et al. 2002, Pizzamiglio et al. 2005; Hauk et al. 2006; Gazzola et al. 2006; Pazzaglia et al. 2008a; Galati et al. 2008). The latter factor involves the specificity of the bodily instantiation of cognitive operations, which is known as ‘embodiment’ (Barsalou 2008). It is widely believed that performance of actions implies the activation of body parts according to a somatotopic, homuncular cortical representation (Penfield

and Rasmussen 1950). Studies of the visual observation of actions suggest that somatotopy may also be an important representational rule for mirroring processes. Indeed, neural activity in the dorsal and ventral parts of the premotor region was higher when viewing hand and mouth actions, respectively (Buccino et al. 2001; Aziz-Zadeh et al. 2006; Wheaton et al. 2004). Moreover, separate neural activities, specifically related to the observation of hand, mouth, and foot movements, were found in both frontal and parietal regions (Buccino et al. 2001). Likewise, imaging studies showed that the left hemisphere exhibits a somatotopic arrangement along the motor strip (Schubotz et al. 2003; Wheaton et al. 2004; Leslie et al. 2004) in response to action-related linguistic tasks, such as the passive reading of words (e.g., kick, pick, lick; Hauk et al. 2004; Pulvermüller 2001), or phrases (Aziz-Zadeh et al. 2006), or listening to sentences (Tettamanti et al. 2005). Whether or not somatotopy also applies to audio-motor mirroring is still somewhat controversial. High density EEG (Hauk et al. 2006) and fMRI (Gazzola et al. 2006) investigations of healthy subjects revealed that the motor mapping of heard actions follows a somatotopic pattern. Indeed, clear differences in the topographies of brain responses to the sounds produced by finger and tongue clicking movements were found (Hauk et al. 2006). Hearing the sound associated with mouth actions and executing such action lead to activation of the pars opercularis of the inferior frontal gyrus, which extends to the rostral-most sector of the ventral premotor cortex. In contrast, while the hearing and execution of hand actions lead to activation of the ventral premotor cortex, this activation partially extends to pars opercularis of the inferior frontal gyrus (Gazzola et al. 2006). In the left premotor cortex, a somatotopic pattern of activation was also observed with a dorsal cluster more involved when hearing and executing hand actions, and with a ventral cluster more involved when hearing and executing mouth actions (Gazzola et al. 2006).

These behavioural and neural double-dissociations indicate that a left-lateralized audio-motor network is actively involved in both the somatotopic motor mapping of sounds related to limb and mouth actions as well as in the execution of the very same actions. Note that the mouth sound region turned out to be represented ventrally to the hand sound area, in line with the well-known arrangement of the motor homunculus (Penfield and Rasmussen 1950). This finding is in accordance with the topographical representation of different effectors within the ventral premotor, supplementary motor (SMA), and the primary motor (M1) cortex in monkeys (Takada et al. 1998; Tokuno et al. 1997).

However, not all the audio-motor mirroring studies report a somatotopic pattern of modulation. Galati et al. (2008), for example, found no body part specific

modulation of the premotor and prefrontal activity contingent upon semantic priming related to action sound recognition. Moreover, it has been suggested that any purported somatotopic pattern of activation is weak and coarse and is possibly linked to substantial variations in the localization of the activation peak and multiple representations of the same body part (Fernandino and Iacoboni 2010).

To further investigate the anatomical localization of the different effectors associated with performed and heard actions, we compared the locations of the selective deficits in limb- and face-sound perception in VLSM clusters in our previous lesion study (Pazzaglia et al. 2008a) to those of limb- and face-sound selective areas found in previous fMRI studies. We examined the fMRI studies showing localized somatotopy in the premotor cortex when subjects heard the sounds of a mouth or limb action and listened to phrases or words describing actions. Activation associated with sound-related actions was localized in only a limited number of studies. Furthermore, we only considered studies that provided the mean coordinates in standard space of peak activation of the left hemisphere. Coordinates in Talairach spaces (Talairach and Tournoux 1988) were transformed into MNI space by using Matlab (MathWorks, Natick, MA).

Table 1 lists the MNI coordinates reported by the fMRI studies that identified Talairach coordinates of premotor peaks for mouth and hand actions. These included five

studies that investigated the neural correlates of audio-motor mapping (de Zubicaray et al. 2010; Gazzola et al. 2006; Hauk et al. 2004; Kemmerer et al. 2008; Tettamanti et al. 2005). We also applied our analysis to three studies showing localized somatotopy when mouth and hand parts were viewed (Aziz-Zadeh et al. 2006; Buccino et al. 2001; Postle et al. 2008).

To identify hand- and face-selective functional regions of interest (ROIs) in fMRI studies considered for this analysis, we created a series of spherical 6-mm radius ROIs centred over these coordinates by using MRICron. Overlapping images were then created for sound/observation and mouth/hand activation, and then overlaid on the standard brain along with the VLSM clusters associated with selective deficits in buccofacial-, but not limb-related action sounds (Fig. 1).

Furthermore, we calculated the geometric distance between the centre of mass of the VLSM clusters selectively associated with face-related action sound discrimination deficits and the coordinates of mouth and hand activation in visual and auditory tasks, as reported in the fMRI studies selected for this analysis. A series of independent sample *t*-tests (two-tailed) was used to compare the geometric distances between the mouth VLSM cluster and hand and face-selective activation in fMRI studies (see Table 1). Mouth-related sound actions were localized in the left hemisphere by 5 fMRI studies (Table 1). The coordinates of the centre of mass of the left area associated

Table 1 The MNI coordinates of premotor activation peaks for mouth- and hand-related actions as inferred from 5 auditory- and 3 visuo-motor mapping studies

Mouth				Sound	Hand			
MNI coordinates					MNI coordinates			
<i>x</i>	<i>y</i>	<i>z</i>	<i>d</i>	Fmri studies	<i>d</i>	<i>x</i>	<i>y</i>	<i>z</i>
-56	6	8	2.23	Gazzola et al. (2006)	66.1	-22	-4	66
-50	10	20	12.4	Hauk et al. (2004)	63.7	-22	2	64
-56	12	12	7	Tettamanti et al. (2005)	53.4	-30	-2	56
-50	18	20	16.8	Kemmerer et al. (2008)	27.8	-46	28	24
-54	9	9	3	de Zubicaray et al. (2010)	16.4	-57	12	24
53.2	11	13.8	8	Mean	45.5	35.4	7.2	46.8
Mouth				Observation	Hand			
MNI coordinates					MNI coordinates			
<i>x</i>	<i>y</i>	<i>z</i>	<i>d</i>	Fmri studies	<i>d</i>	<i>x</i>	<i>y</i>	<i>z</i>
-54	4	26	17	Aziz-Zadeh et al. (2006)	45.7	-30	-6	46
-65	11	22	17.7	Buccino et al. (2001)	20.5	-65	3	26
-51	-6	51	43.8	Postle et al. (2008)	66.7	-33	-27	63
-57	3	33	26.2	Mean	44.3	-43	-10	45

The geometric distance between the coordinates of the fMRI studies and the centre of mass as reported in the VLSM study (Pazzaglia et al. 2008a) is represented in the table by 'd'

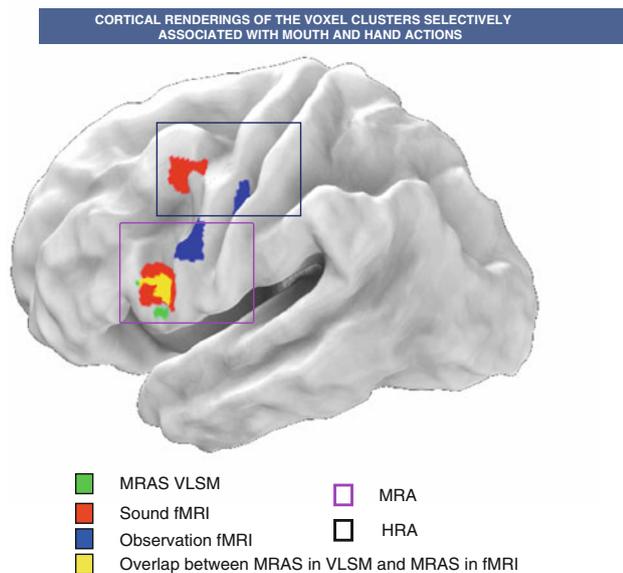


Fig. 1 Cortical renderings of the left-premotor voxel clusters selectively associated with deficits in processing mouth-related action sounds (MRAS; in green) but not hand-related action sounds and with the selective fMRI activations during mouth—(MRA; in the pink box) and hand-related action (HRA; in the black box) sound processing. The map for each functional area represents the ROIs around the activation peaks reported in the visual (in blue) and auditory (in red) fMRI studies. MNI coordinates of all activations are shown in Table 1

with mouth-related sound action perception deficits in our VLSM study ($x = -54$, $y = 6$, $z = 9$) were within the range of the coordinates reported in fMRI studies for left mouth-related sound action ($x =$ from -50 to -56 , $y =$ from 6 to 18 , $z =$ from 8 to 20). Furthermore, its geometric distance from left mouth-related sound action coordinates (mean = 8.27 mm, SD = 6.2) was smaller than the distance from left mouth-related visual action coordinates (mean = 26.2 mm, SD = 15.2 ; $t_6 = -2.42$, $P < 0.0517$), from hand action sound coordinates (mean = 45.49 mm, SD = 22.24 ; $t_8 = -3.60$, $P < 0.007$) and from hand action visual coordinates (mean = 45.28 mm, SD = 23.13 ; $t_8 = -3.45$, $P < 0.01$).

This analysis highlights that somatotopy may be an important representational rule for mapping actions not only on the basis of visual observation but also on the basis of auditory perception.

Importance of audio-motor mirroring for language

It has been postulated that language may have evolved from manual gestures, initially as a system of imitation and pantomime (Leroi-Gourhan 1964, 1965). Subsequent conventionalization of gestures and shifts to vocal gestures allowed the transition towards more symbolic and abstract forms of communication (Corballis 2009). In a similar

vein, increasing amount of evidence indicates that even complex aspects of language may be ‘embodied’ in ‘modal’ perceptuo-motor systems (Barsalou 2008; Rizzolatti and Sinigaglia 2008). The existence of mechanisms and neural machinery for the direct matching of auditory input with motor output has possible functional advantages and adaptive implications. For example, while speech sounds representations are highly variable, motor representations are comparatively stable. Therefore, acoustic mapping of speech sounds may facilitate their phonemic categorization and be important for reliable speech perception. Since language acquisition entails the constant matching of input–output processes (Rizzolatti and Arbib 1998; Gallese 2007; Arbib 2005), audio-motor mirroring may be crucial for language acquisition. Indeed, audio-motor mirror neurons with precise auditory-vocal correspondence are found in singing birds suggesting that a neural system for a fine-tuned matching of heard actions and action execution, facilitates imitation learning upon which speech and language development typically relies (Rizzolatti and Arbib 1998; Theoret and Pascual-Leone 2002; Gentilucci and Dalla Volta 2008).

It has been suggested that during evolution, the initial lateralization of vocalizations may have also influenced the lateralization of action sounds (Corballis 2002). The link between the sounds of body actions and language sounds is further supported by the large overlap of the neural systems involved in these two processes. It is noteworthy that both language and the auditory mirror system are left side lateralized (Keysers et al. 2003). This association is also supported by the finding that the lateralized neural system activated by listening to action sounds is also activated during listening or reading words describing the same action (Hauk et al. 2008; Pulvermüller 2005). Left-brain-damaged apraxic and aphasic patients have problems in auditorily or visually recognizing the actions they cannot perform, suggesting a close link between language, praxis, and audio-motor mapping of actions. Moreover, a close overlap of the networks underlying language and gesture has been demonstrated by lesion studies (Nelissen et al. 2010; Pazzaglia et al. 2008a, b; Saygin et al. 2003). Moreover, many of the regions activated in response to action sounds are located within or near the classically defined language area. In particular, left inferior frontal activations are consistent with the proposed role of Broca’s area in coding word-related thoughts into speech, irrespectively of whether the words are actually articulated or silently subvocalized by the subject. Thus, the neural structures underpinning action execution are also expected to play a role in understanding the semantic content of the verbally described actions or the heard action sounds. That the human vocal tract can utter speech and most of the non-linguistic mouth sounds, but cannot produce most of the

environmental sounds, would explain why the latter are not mapped on the audio-motor mirror system (Gazzola et al. 2006; Hauk et al. 2006; Pazzaglia et al. 2008a; Pizzamiglio et al. 2005).

Importance of audio-motor mirroring for social interactions

Human beings continuously obtain information about identities, actions, emotions, and intentions of other individuals from their actions. There is evidence to suggest that viewing the actions of another individual allows the onlooker to gain information about others' intentions and future behaviours that are fundamental to social interactions (Hari and Kujala 2009; Gallese et al. 2009). Fundamental aspects of social actions are acquired through hearing. In addition to the importance of audio-motor mapping for language, as outlined in the previous paragraph, it is worth mentioning here that efficient mechanisms for matching audition with action may allow the survival of all hearing individuals. For example, particularly in epochs when the nights were dark, the detection of potential danger (e.g., the footsteps of enemies) and the consequential implementation of fight-or-flight behaviour mainly relied on audition. Moreover, studies indicate that auditory detection of affective non-verbal vocalizations activates a premotor region, which is also activated during facial movements. These results suggest that the audio-motor mirror system may play an important role in reading the emotion of others, a process that is clearly fundamental for modulating interpersonal reactivity (Warren et al. 2006). This view is further supported by the positive correlation between empathy and neural activity in the fronto-parietal nodes of the audio-motor mirror system (Gazzola et al. 2006). It is worth noting that while an association between defective visuo-motor resonance and conditions of reduced empathy (e.g., in autistic spectrum disorders) has been found (Dapretto et al. 2006; Le Bel et al. 2009), no studies have explored whether the same association is present when deriving actions from sounds.

Future directions

We conclude by discussing some of the outstanding issues for future research on motor resonance in general and audio-motor mirroring in particular.

Unimodal and multimodal modulation of action mapping

The classical notion that vision predominates over the other senses in primates raises the question whether

audio-motor mirroring is necessarily mediated by visual perception and visual imagery transformations. The selective audio-motor mapping deficits reported by Pazzaglia et al. (2008a) may simply be an instance of the visuo-motor mapping deficits (Pazzaglia et al. 2008b). However, studies suggest that our perception of the world and our perception–action integration processes are inherently cross-modal (Driver and Spence 2000). Perceiving actions under daily life conditions, for example, typically implies the merging of visual and auditory information to achieve their best interpretation and optimally guide behaviour (Chen and Spence 2009). This is in accordance with the similar reactivity of the human premotor and parietal regions when viewing, hearing and performing hand movements (Galati et al. 2008; Gazzola et al. 2006). However, interesting research has attempted to distinguish the separate contribution of visual and auditory input on motor mirroring. Kaplan and Iacoboni (2007) compared to the BOLD signal when subjects observed a video of another individual tearing paper, viewed a silent video of the same action and heard the sound of paper tearing. Unimodal perception activated similar regions in the frontal cortex. Vision alone activated also the parietal cortex. Importantly, visual and auditory perception brought about an enhancement of neural activation in frontal regions thus suggesting that motor resonance is fed by cross-modal inputs. Similarly Alaerts et al. (2009) used single-pulse TMS to explore cortico-spinal motor reactivity while subjects were presented with unimodal or bimodal stimuli. The unimodal stimulation could be visual (presentation of a silent clip of a hand crushing a small drinking bottle) or auditory (presentation of the crush sound). The bimodal stimulation could be audio-visually congruent or incongruent (vision of a crushing hand coupled with a water pouring sound; or vision of a foot crushing a bottle coupled with a hand crushing sound). Maximal facilitation was found during the perception of congruent bimodal stimuli suggesting that while it is possible to elicit a modality specific mirroring, optimal tuning of the action system is typically multimodal (Alaerts et al. 2009). Crucial to the issue of the multimodal modulation of resonance with others' actions, is a recent fMRI study (Ricciardi et al. 2009) that assessed brain activity in congenitally blind individuals and in healthy blindfolded subjects who listened to hand-executed actions or environmental sounds, and performed manipulation tasks. The results showed that the same fronto-parietal network activated by viewing actions in healthy subjects was activated during perception of action sounds and action execution both in healthy and congenitally blind subjects. Moreover, in both sighted and blind individuals, higher fronto-parietal activity was elicited by familiar motor sounds than by unfamiliar action sounds. Expanding the results of this

study by exploring other conditions of unimodal perception defects (e.g., deaf children) may allow to tackle a fundamental question concerning the brain systems underpinning motor mirroring, namely their development and plasticity.

Development and plasticity of action mirroring: phenomenology and neural correlates

As previously mentioned, direct evidence for the existence of mirror neurons comes only from animal studies. However, there is a wealth of evidence that the adult human brain is equipped with neural systems and mechanisms that represent perception and action in common formats. Strikingly, little is known about the development of these systems and mechanisms. Classical behavioural studies demonstrate that, even minutes after birth, human neonates can imitate the oro-facial movements performed by adult models in front of them (Meltzoff and Moore 1977). Although indirect, this is perhaps the most compelling evidence of an innate presence of mirror neurons in the human brain. With the exception of EEG studies demonstrating the modulation of the mu rhythm during action observation (Lepage and Théoret 2006, 2007), evidence for neural correlates of visual, audio, or multimodal mirroring in children is meagre. Given the theoretical and practical importance of the topic, studies on it are somewhat overdue. The plastic nature of visuo-motor mirroring is suggested by studies showing that this process may be stronger (Calvo-Merino et al. 2006) or more finely tuned (Aglioti et al. 2008) in expert brains. Moreover, studies indicate that motor mapping of heard actions is highly sensitive to the individual's motor experience and can discriminate between the sounds of newly acquired actions and the sounds of actions that have no known motor representation. Learning based on repeated exposure to sounds, increases the listener's ability to associate sounds with the relevant actions in both novices (Lahav et al. 2007) and expert pianists (Drost et al. 2005), and also modulates neural activity in fronto-parietal regions. Given the bidirectional link between action observation and execution, it is highly plausible that re-training motor and perceptual components of actions may benefit from using perceptual and motor cues, respectively. While tasks based on visual observation of actions have been used for the rehabilitation of action execution disorders (Smania et al. 2006; Celnik et al. 2008), no studies have so far explored whether the ability to recognize human actions by their sound may be important for the planning of treatments for goal-directed action disorders in neurological patients. Therefore, deepening current knowledge on the plastic potentialities of action mirroring may be important for rehabilitation purposes.

References

- Aglioti SM, Cesari P, Romani M, Urgesi C (2008) Action anticipation and motor resonance in elite basketball players. *Nat Neurosci* 11:1109–1116
- Alaerts K, Swinnen SP, Wenderoth N (2009) Interaction of sound and sight during action perception: evidence for shared modality-dependent action representations. *Neuropsychologia* 47:2593–2599
- Amedi A, von Kriegstein K, van Atteveldt NM, Beauchamp MS, Naumer MJ (2005) Functional imaging of human crossmodal identification and object recognition. *Exp Brain Res* 166:559–571
- Arbib MA (2005) From monkey-like action recognition to human language: an evolutionary framework for neurolinguistics. *Behav Brain Sci* 28:105–124
- Avenanti A, Bolognini N, Maravita A, Aglioti SM (2007) Somatic and motor components of action simulation. *Curr Biol* 17:2129–2135
- Aziz-Zadeh L, Iacoboni M, Zaidel E, Wilson S, Mazziotta J (2004) Left hemisphere motor facilitation in response to manual action sounds. *Eur J Neurosci* 19:2609–2612
- Aziz-Zadeh L, Wilson SM, Rizzolatti G, Iacoboni M (2006) Congruent embodied representations for visually presented actions and linguistic phrases describing actions. *Curr Biol* 16:1818–1823
- Barsalou LW (2008) Grounded cognition. *Annu Rev Psychol* 59:617–645
- Belin P (2006) Voice processing in human and non-human primates. *Philos Trans R Soc Lond B Biol Sci* 361:2091–2107
- Bidet-Caulet A, Voisin J, Bertrand O, Fonlupt P (2005) Listening to a walking human activates the temporal biological motion area. *Neuroimage* 28:132–139
- Buccino G, Binkofski F, Fink GR, Fadiga L, Fogassi L, Gallese V, Seitz RJ, Zilles K, Rizzolatti G, Freund HJ (2001) Action observation activates premotor and parietal areas in a somatotopic manner: an fMRI study. *Eur J Neurosci* 13:400–404
- Buccino G, Vogt S, Ritzl A, Fink GR, Zilles K, Freund HJ, Rizzolatti G (2004) Neural circuits underlying imitation learning of hand actions: an event-related fMRI study. *Neuron* 42:323–334
- Caetano G, Jousmäki V, Hari R (2007) Actor's and observer's primary motor cortices stabilize similarly after seen or heard motor actions. *Proc Natl Acad Sci USA* 104:9058–9062
- Calvo-Merino B, Grèzes J, Glaser DE, Passingham RE, Haggard P (2006) Seeing or doing? Influence of visual and motor familiarity in action observation. *Curr Biol* 16:1905–1910
- Campanella S, Belin P (2007) Integrating face and voice in person perception. *Trends Cogn Sci* 11:535–543
- Candidi M, Vicario CM, Abreu AM, Aglioti SM (2010) Competing mechanisms for mapping action-related categorical knowledge and observed actions. *Cereb Cortex* [Epub ahead of print]
- Casile A, Giese MA (2006) Nonvisual motor training influences biological motion perception. *Curr Biol* 16:69–74
- Celnik P, Webster B, Glasser DM, Cohen LG (2008) Effects of action observation on physical training after stroke. *Stroke* 39:1814–1820. Epub 2008 Apr 10
- Chen YC, Spence C (2009) When hearing the bark helps to identify the dog: semantically-congruent sounds modulate the identification of masked pictures. *Cognition*
- Corballis MC (2002) From hand to mouth: the origins of language. Princeton University Press, Princeton
- Corballis MC (2009) The evolution of language. *Ann NY Acad Sci* 1156:19–43
- Costantini M, Galati G, Ferretti A, Caulo M, Tartaro A, Romani GL, Aglioti SM (2005) Neural systems underlying observation of

- humanly impossible movements: an fMRI study. *Cereb Cortex* 15:1761–1767
- D'Ausilio A, Pulvermüller F, Salmas P, Bufalari I, Begliomini C, Fadiga L (2009) The motor somatotopy of speech perception. *Curr Biol* 19:381–385 (Epub 2009 Feb 12)
- Dapretto M, Davies MS, Pfeifer JH, Scott AA, Sigman M, Bookheimer SY, Iacoboni M (2006) Understanding emotions in others: mirror neuron dysfunction in children with autism spectrum disorders. *Nat Neurosci* 9:28–30
- De Lucia M, Camen C, Clarke S, Murray MM (2009) The role of actions in auditory object discrimination. *Neuroimage* 48:475–485 (Epub 2009 Jun 24)
- de Zubicaray G, Postle N, McMahon K, Meredith M, Ashton R (2010) Mirror neurons, the representation of word meaning, and the foot of the third left frontal convolution. *Brain Lang* 112:77–84 (Epub 2008 Nov 1)
- Decety J, Grèzes J, Costes N, Perani D, Jeannerod M, Procyk E, Grassi F, Fazio F (1997) Brain activity during observation of actions. Influence of action content and subject's strategy. *Brain* 120:1763–1777
- di Pellegrino G, Fadiga L, Fogassi L, Gallese V, Rizzolatti G (1992) Understanding motor events: a neurophysiological study. *Exp Brain Res* 91:176–180
- Dick F, Saygin AP, Galati G, Pitzalis S, Bentrovato S et al (2007) What is involved and what is necessary for complex linguistic and nonlinguistic auditory processing: evidence from functional magnetic resonance imaging and lesion data. *J Cogn Neurosci* 19:799–816
- Driver J, Spence C (2000) Multisensory perception: beyond modularity and convergence. *Curr Biol* 10:R731–R735
- Drost UC, Rieger M, Brass M, Gunter TC, Prinz W (2005) When hearing turns into playing: movement induction by auditory stimuli in pianists. *Q J Exp Psychol A* 58:1376–1389
- Ekimov A, Sabatier JM (2006) Vibration and sound signatures of human footsteps in buildings. *J Acoust Soc Am* 120:762–768
- Engel LR, Frum C, Puce A, Walker NA, Lewis JW (2009) Different categories of living and non-living sound-sources activate distinct cortical networks. *Neuroimage* 47:1778–1791 (Epub 2009 May 22)
- Fadiga L, Fogassi L, Pavesi G, Rizzolatti G (1995) Motor facilitation during action observation: a magnetic stimulation study. *J Neurophysiol* 73:2608–2611
- Fadiga L, Craighero L, Buccino G, Rizzolatti G (2002) Speech listening specifically modulates the excitability of tongue muscles: a TMS study. *Eur J Neurosci* 15:399–402
- Fernandino L, Iacoboni M (2010) Are cortical motor maps based on body parts or coordinated actions? Implications for embodied semantics. *Brain Lang* 112:44–53 (Epub 2009 Apr 3)
- Fogassi L, Ferrari PF, Gesierich B, Rozzi S, Chersi F, Rizzolatti G et al (2005) Parietal lobe: from action organization to intention understanding. *Science* 308:662–667
- Galati G, Committeri G, Spitoni G, Aprile T, Pitzalis S et al (2008) A selective representation of the meaning of actions in the auditory mirror system. *Neuroimage* 40:1274–1286
- Gallese V (2007) Before and below 'theory of mind': embodied simulation and the neural correlates of social cognition. *Philos Trans R Soc Lond B Biol Sci* 362:659–669
- Gallese V, Fadiga L, Fogassi L, Rizzolatti G (1996) Action recognition in the premotor cortex. *Brain* 119:593–609
- Gallese V, Rochat M, Cossu G, Sinigaglia C (2009) Motor cognition and its role in the phylogeny and ontogeny of action understanding. *Dev Psychol* 45:103–113
- Gangitano M, Mottaghy FM, Pascual-Leone A (2001) Phase-specific modulation of cortical motor output during movement observation. *Neuroreport* 12:1489–1492
- Gazzola V, Keysers C (2009) The observation and execution of actions share motor and somatosensory voxels in all tested subjects: single-subject analyses of unsmoothed fMRI data. *Cereb Cortex* 19:1239–1255
- Gazzola V, Aziz-Zadeh L, Keysers C (2006) Empathy and the somatotopic auditory mirror system in humans. *Curr Biol* 16:1824–1829
- Gentilucci M, Dalla Volta R (2008) Spoken language and arm gestures are controlled by the same motor control system. *Q J Exp Psychol (Colchester)* 61:944–957
- Gori M, Del Viva M, Sandini G, Burr DC (2008) Young children do not integrate visual and haptic form information. *Curr Biol* 18:694–698
- Hari R, Kujala MV (2009) Brain basis of human social interaction: from concepts to brain imaging. *Physiol Rev* 89:453–479
- Haslinger B, Erhard P, Altenmüller E, Schroeder U, Boecker H, Ceballos-Baumann AO (2005) Transmodal sensorimotor networks during action observation in professional pianists. *J Cogn Neurosci* 17:282–293
- Hauk O, Johnsrude I, Pulvermüller F (2004) Somatotopic representation of action words in human motor and premotor cortex. *Neuron* 41:301–307
- Hauk O, Shtyrov Y, Pulvermüller F (2006) The sound of actions as reflected by mismatch negativity: rapid activation of cortical sensory-motor networks by sounds associated with finger and tongue movements. *Eur J Neurosci* 23:811–821
- Hauk O, Shtyrov Y, Pulvermüller F (2008) The time course of action and action-word comprehension in the human brain as revealed by neurophysiology. *J Physiol Paris* 102:50–58 (Epub Apr 1)
- Kaplan JT, Iacoboni M (2007) Multimodal action representation in human left ventral premotor cortex. *Cogn Process* 8:103–113
- Kemmerer D, Castillo JG, Talavage T, Patterson S, Wiley C (2008) Neuroanatomical distribution of five semantic components of verbs: evidence from fMRI. *Brain Lang* 107:16–43
- Keysers C, Kohler E, Umiltà MA, Nanetti L, Fogassi L et al (2003) Audiovisual mirror neurons and action recognition. *Exp Brain Res* 153:628–636
- Knoblich G, Repp BH (2009) Inferring agency from sound. *Cognition* 111:248–262
- Kohler E, Keysers C, Umiltà MA, Fogassi L, Gallese V et al (2002) Hearing sounds, understanding actions: action representation in mirror neurons. *Science* 297:846–848
- Lahav A, Saltzman E, Schlaug G (2007) Action representation of sound: audiomotor recognition network while listening to newly acquired actions. *J Neurosci* 27:308–314
- Le Bel RM, Pineda JA, Sharma A (2009) Motor-auditory-visual integration: the role of the human mirror neuron system in communication and communication disorders. *J Commun Disord* 42:299–304
- Lepage JF, Théoret H (2006) EEG evidence for the presence of an action observation-execution matching system in children. *Eur J Neurosci* 23:2505–2510
- Lepage JF, Théoret H (2007) The mirror neuron system: grasping others' actions from birth. *Dev Sci* 10:513–523
- Lepage JF, Tremblay S, Nguyen DK, Champoux F, Lassonde M, Théoret H (2010) Action related sounds induce early and late modulations of motor cortex activity. *Neuroreport* 21:250–253
- Leroi-Gourhan Le Geste et la Parole I (1964) *Technique et Langage*, Parigi, Albin Michel (coll. Sciences d'aujourd'hui), p 323
- Leroi-Gourhan Le Geste et la Parole II (1965) *La Mémoire et les rythmes*, Parigi, Albin Michel (coll. Sciences d'aujourd'hui), p 285
- Leslie KR, Johnson-Frey S, Grafton ST (2004) Functional imaging of face and hand imitation: towards a motor theory of empathy. *NeuroImage* 21:601–607

- Lewis JW, Brefczynski JA, Phinney RE, Janik JJ, DeYoe EA (2005) Distinct cortical pathways for processing tool versus animal sounds. *J Neurosci* 25:5148–5158
- Lewis JW, Phinney RE, Brefczynski-Lewis JA, DeYoe EA (2006) Lefties get it “right” when hearing tool sounds. *J Cogn Neurosci* 18:1313–1314
- Lieberman AM, Mattingly IG (1985) The motor theory of speech perception revised. *Cognition* 21:1–36
- Meltzoff AN, Moore MK (1977) Imitation of facial and manual gestures by human neonates. *Science* 198:74–78
- Möttönen R, Watkins KE (2009) Motor representations of articulators contribute to categorical perception of speech sounds. *J Neurosci* 29:9819–9825
- Nelissen N, Pazzaglia M, Vandenbulcke M, Sunaert S, Fannes K, Dupont P, Aglioti SM, Vandenberghe R (2010) Gesture discrimination in primary progressive aphasia: the intersection between gesture and language processing pathways. *J Neurosci* (in press)
- Pazzaglia M, Pizzamiglio L, Pes E, Aglioti SM (2008a) The sound of actions in apraxia. *Curr Biol* 18:1766–1772 (Epub 2008 Nov 13)
- Pazzaglia M, Smania N, Corato E, Aglioti SM (2008b) Neural underpinnings of gesture discrimination in patients with limb apraxia. *J Neurosci* 28:3030–3041
- Penfield W, Rasmussen T (1950) *The cerebral cortex of man*. MacMillan, New York
- Pietrini P, Furey ML, Ricciardi E, Gobbi MI, Wu WH, Cohen L, Guazzelli M, Haxby JV (2004) Beyond sensory images: object-based representation in the human ventral pathway. *Proc Natl Acad Sci USA* 101:5658–5663
- Pizzamiglio L, Aprile T, Spitoni G, Pitzalis S, Bates E, D’Amico S, Di Russo F (2005) Separate neural systems for processing action- or non-action-related sounds. *Neuroimage* 24:852–861
- Postle N, McMahon K, Ashton R, Meredith M, de Zubicaray G (2008) Action word meaning representations in cytoarchitecturally defined primary and premotor cortices. *NeuroImage* 43:634–644
- Prather JF, Peters S, Nowicki S, Mooney R (2008) Precise auditory-vocal mirroring in neurons for learned vocal communication. *Nature* 451:305–310
- Pulvermüller F (2001) Brain reflections of words and their meaning. *Trends Cogn Sci* 5:517–524
- Pulvermüller F (2005) Brain mechanisms linking language and action. *Nat Rev Neurosci* 6:576–682
- Ricciardi E, Bonino D, Sani L, Vecchi T, Guazzelli M, Haxby JV, Fadiga L, Pietrini P (2009) Do we really need vision? How blind people “see” the actions of others. *J Neurosci* 29:9719–9724
- Rizzolatti G, Arbib MA (1998) Language within our grasp. *Trends Neurosci* 21:188–194
- Rizzolatti G, Craighero L (2004) The mirror-neuron system. *Annu Rev Neurosci* 27:169–192
- Rizzolatti G, Sinigaglia C (2008) Further reflections on how we interpret the actions of others. *Nature* 455:589
- 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. *Exp Brain Res* 111:246–252
- Romani M, Cesari P, Urgesi C, Facchini S, Aglioti SM (2005) Motor facilitation of the human cortico-spinal system during observation of bio-mechanically impossible movements. *Neuroimage* 26:755–763
- Rossi S, De Capua A, Pasqualetti P, Ulivelli M, Fadiga L, Falzarano V, Bartalini S, Passero S, Nuti D, Rossini PM (2008) Distinct olfactory cross-modal effects on the human motor system. *PLoS One* 27(3):e1702
- Saarela MV, Hari R (2008) Listening to humans walking together activates the social brain circuitry. *Soc Neurosci* 3:401–409
- Sauter DA, Eimer M (2010) Rapid detection of emotion from human vocalizations. *J Cogn Neurosci* 22:474–481
- Saygin AP, Dick F, Wilson SM, Dronkers NF, Bates E (2003) Neural resources for processing language and environmental sounds: evidence from aphasia. *Brain* 126:928–945
- Schubotz RI, von Cramon DY, Lohmann G (2003) Auditory what, where, and when: a sensory somatotopy in lateral premotor cortex. *Neuroimage* 20:173–185
- Shams L, Kamitani Y, Shimojo S (2000) Illusions. What you see is what you hear. *Nature* 408:788
- Smania N, Aglioti SM, Girardi F, Tinazzi M, Fiaschi A, Cosentino A, Corato E (2006) Rehabilitation of limb apraxia improves daily life activities in patients with stroke. *Neurology* 67:2050–2052
- Stefan K, Cohen LG, Duque J, Mazzocchio R, Celnik P, Sawaki L, Ungerleider L, Classen J (2005) Formation of a motor memory by action observation. *J Neurosci* 25:9339–9346
- Strafella AP, Paus T (2000) Modulation of cortical excitability during action observation: a transcranial magnetic stimulation study. *Neuroreport* 11:2289–2292
- Takada M, Tokuno H, Nambu A, Inase M (1998) Corticostriatal projections from the somatic motor areas of the frontal cortex in the macaque monkey: segregation versus overlap of input zones from the primary motor cortex, the supplementary motor area, and the premotor cortex. *Exp Brain Res* 120:114–128
- Talairach J, Tournoux P (1988) *Co-planar stereotaxic atlas of the human brain*. Thieme, New York
- Tettamanti M, Buccino G, Saccuman MC, Gallese V, Danna M et al (2005) Listening to action-related sentences activates frontoparietal motor circuits. *J Cogn Neurosci* 17:273–281
- Theoret H, Pascual-Leone A (2002) Language acquisition: do as you hear. *Curr Biol* 12:R736–R737
- Tokuno H, Takada M, Nambu A, Inase M (1997) Reevaluation of ipsilateral corticocortical inputs to the orofacial region of the primary motor cortex in the macaque monkey. *J Comp Neurol* 389:34–48
- Urgesi C, Candidi M, Fabbro F, Romani M, Aglioti SM (2006) Motor facilitation during action observation: topographic mapping of the target muscle and influence of the onlooker’s posture. *Eur J Neurosci* 23:2522–2530
- Urgesi C, Maieron M, Avenanti A, Tidoni E, Fabbro F, Aglioti SM (2010) Simulating the future of actions in the human cortico-spinal system. *Cereb Cortex* [Epub ahead of print]
- Van den Stock J, Grèzes J, de Gelder B (2008) Human and animal sounds influence recognition of body language. *Brain Res* 1242:185–190
- Vogt S, Buccino G, Wohlschläger AM, Canessa N, Shah NJ, Zilles K, Eickhoff SB, Freund HJ, Rizzolatti G, Fink GR (2007) Prefrontal involvement in imitation learning of hand actions: effects of practice and expertise. *Neuroimage* 37:1371–1383
- Warren JE, Sauter DA, Eisner F, Wiland J, Dresner MA et al (2006) Positive emotions preferentially engage an auditory-motor “mirror” system. *J Neurosci* 26:13067–13075
- Watkins KE, Strafella AP, Paus T (2003) Seeing and hearing speech excites the motor system involved in speech production. *Neuropsychologia* 41:989–994
- Wheaton KJ, Thompson JC, Syngieniotis A, Abbott DF, Puce A (2004) Viewing the motion of human body parts activates different regions of premotor, temporal, and parietal cortex. *NeuroImage* 22:277–288