UNIVERSITÀ DEGLI STUDI DI TRENTO
CIMeC – Center for Mind/Brain Science

Doctoral school in Cognitive and Brain Science
XXIX cycle

Ph.D. Dissertation

WHERE SYMBOLS MEET MEANINGS: THE ORGANIZATION OF GESTURES AND WORDS IN THE MIDDLE TEMPORAL GYRUS

Beatrice Agostini

Advisor: Angelika Lingnau
Co-advisor: Liuba Papeo

A te, cara Bea, che a 20 anni sognavi questo momento…

Ringraziamenti

Innanzitutto desidero ringraziare Angelika. La tua attenzione, precisione e ampia veduta nell'affrontare tutte le fasi di un esperimento mi hanno insegnato l'importanza della rigorosità, ma allo stesso tempo della flessibilità mentale che è necessaria quando si vuole fare ricerca. Grazie per il tempo e la pazienza che mi hai dedicato in questi anni e per l'umanità dimostrata nei difficili momenti personali attraversati.

Con Angelika, ringrazio anche tutti i membri, passati e presenti, del mio lab group. Grazie per i consigli, le critiche, gli aiuti e le interessanti discussioni che hanno accompagnato, mese dopo mese, questi intensi tre anni di dottorato. In particolare grazie Elisa B., per aver condiviso con me l'avventura londinese. Conoscerti è stato bello e mi ha arricchito come persona ancor prima che come collega.

Grazie Arianna, the actress, per aver trascorso due lunghissime giornate a ripeterere gesti e pronunciare parole davanti alla telecamera. Il tuo aiuto è stato essenziale per la buona riuscita di questo lavoro.

Un ringraziamento a Seth, my American native-speaker. Grazie per aver letto la mia tesi a caccia di strafalcioni grammaticali che spero ti abbiano strappato un sorriso.

Un abbraccio e un grazie di cuore a Ceren e Stefania, mie fedeli compagne di dottorato e, ormai, carissime amiche. Grazie per aver condiviso con me i pensieri, i dubbi, le gioie e i dolori della vita universitaria e non.

Infine, grazie Fra, per esserci sempre. Ma soprattutto, grazie per aver tentato di leggere e comprendere i miei articoli di statistica, quando, in lacrime, pensavo che imparare il cinese sarebbe stato più facile!

Table of Contents

Chapter 1. Background 3

1.1 Introduction 3

1.2 Motor theories of action understanding 6

1.3 Cognitive theories of action understanding 10

1.4 Action and language 16

1.5 Conclusion 20

1.6 Current work 21

Chapter 2. A norming study of 228 high-quality video clips of pantomimes, emblems and meaningless gestures 23

2.1 Abstract 23

2.2 Introduction 24

2.3 Methods 26

2.4 Results 34

2.5 Discussion 39

Chapter 3. The role of temporal and precentral areas in action understanding: an fMRI-guided TMS study 41

3.1 Abstract 41

3.2 Introduction 42

3.3 Experiment 1 – fMRI study 46

3.3.1 Materials and Methods 46

3.3.2 Results 51

3.3.3 Discussion 56

3.4 Experiment 2 – fMRI-guided TMS study 59

3.4.1 Materials and Methods 59

3.4.2 Results 66

3.4.3 Discussion 68

Chapter 4. The organization of gestures and words in the middle temporal gyrus 71
4.1 Abstract 71
4.2 Introduction 72
4.3 Materials and Methods 73
4.4 Results 76
4.5 Discussion 83
Chapter 5. General Conclusions **87**
Chapter 6. Future directions **93**
Appendices **95**
References **119**

Chapter 1
Background

1.1 Introduction

Human beings are social animals able to process and understand each other's actions and intentions. These features constitute the basis of every fruitful interaction and cultural adaptation. For instance, while walking on the street, understanding whether a person coming towards us is raising her hand to greet us or to throw something at us, is fundamental in order to behave consequently and accordingly.

Early investigations on action understanding were performed in the 1990s on monkeys. In particular, in 1992 di Pellegrino and colleagues observed how neurons of the rostral part of the inferior premotor cortex of the monkey were discharging not only during goal-directed movements but also while observing the same movement performed by the experimenter. These neurons, located in the area F5 of the monkey's brain have been named "mirror neurons" (Rizzolatti et al., 1996), thanks to their capacity of *mirroring* a specific motor action in the observer's brain, as if he had been executing that action. This result has been interpreted as the basis of the process of action understanding: an observed action is understood because the observer knows what the outcome of that action would be if he were to perform it (Gallese et al., 1996; Umiltà et al., 2001; Fogassi et al., 2005). Further studies, using different methodologies, have shown that the observation of actions is linked to the activation of precentral regions also in humans (Gallese et al., 2004; Rizzolatti and Craighero, 2004; Keyser and Gazzola, 2009).

Researchers suggested the existence of a human mirror neuron system consisting of parietal, inferior frontal, and premotor areas that is responsible for action understanding (Rizzolatti et al., 2001). Therefore, it has been proposed that an action concept like *jumping* is understood thanks to the re-activation of the same motor program that would be active if we were to perform that action. The popularity of the idea that action concepts are embodied within our motor and sensory systems began in the mid-1900s, and many studies have been carried out following this main stream. In particular, the notion of *simulation* as the "process by which concepts re-evoke perceptual and motor states present when perceiving and acting the world" (Chatterjee, 2010 -p. 80) started to represent the focus of the research on action observation and understanding (Gallese and Goldman, 1998; Barsalou et al., 2008; Witt et al., 2010, for reviews and theoretical discussions, see Martin, 2007; Mahon and Caramazza, 2008; Kiefer and Pulvermüller, 2012). The same concept has been used to speculate about different cognitive domains, such as empathy and emotions recognition (Spaulding, 2012), theory of mind (Gallese and Goldman, 1998; Schulte-Rüther et al., 2007) as well as the nature of diseases such as the autism spectrum disorder (Dapretto et al., 2005; Oberman et al., 2005; Hadjikhani et al., 2006).

Nowadays, it is still debated whether the simulation process described above is necessary for understanding an action. The point of contention is linked to the importance and the role of motor information in the comprehension of meaningful actions (irrespective of whether the input modality is linguistic, visual, or auditory). More concretely, we might consider the following as the key question of the embodied/disembodied debate: does the understanding of an action (expressed through a gesture or a verb) depend on the activation of the network associated with the physical act? Alternatively, is it possible to understand the meaning of that action using only a symbolic (therefore disembodied) representation, without the contribution of the motor circuit necessary to perform that action?

Motor theories of action understanding, supporting the mirror neuron theory and therefore the embodied cognition hypothesis, claim that the ability to understand and recognize an action is grounded in the motor representation (i.e., the ability to mentally represent the actions seen). Therefore, action understanding requires an internal simulation of the observed action in the observer's motor system (Fischer and Zwaan, 2008; Rizzolatti and Sinigaglia, 2010; Meteyard et al., 2012; Kiefer and Pulvermüller, 2012). As already mentioned above, this means that the understanding of a concept like *jumping* depends on the successful re-activation of the motor program that is usually active when we perform the jump.

By contrast, *cognitive theories* state that we understand an action by having access to action knowledge stored in conceptual areas of the brain, and that the motor system is responding as a consequence of action understanding, through associative links (Mahon and Caramazza, 2008; Papeo et al., 2009; Hickok, 2009; Vannuscorps and Caramazza, 2016; Leshinskaya and Caramazza, 2016). In this context, an action concept like *jumping* is understood thanks to its access to areas representing the abstract concept of *jumping*.

In the next two sections, I will discuss the main findings typically taken as support for motor theories together with some limitations, and I will present the alternative approach offered by cognitive theories. Finally, I will discuss the neural evidence and relationships between actions, gestures, and language.

1.2 Motor theories of action understanding

The majority of behavioral studies available in the literature that are cited in support of motor theories try to show how motor and conceptual representations strictly interact and how the former can interfere with the latter. Earlier studies used priming paradigms, reaction times, and other behavioral measures to identify such interactions. For instance, Glover et al. (2004) used a priming experiment to demonstrate that words automatically activate motor representations. Subjects were instructed to read the name of a relatively large or small object (e.g., *apple* or *grape*), followed by a reaching and grasping movement on a target object. Results showed that the grip aperture was influenced by the prime: reading the name of relatively large objects lead to a larger grip aperture in comparison to reading the name of relatively small objects. This result was shown to be independent of the object size. Along the same lines, Glenberg and Kaschak (2002) reported a phenomenon called “action-sentence compatibility effect”; participants were presented with sense and non-sense sentences such as ‘Open the drawer’ and ‘Boil the air’. Sense sentences implied a direction (towards/away), while non-sense sentences did not. Participants were instructed to judge as fast as possible the sensibility of the sentence by pressing a button located either close or far from the body. Results showed that participants were faster in responding to sentences that implied a direction, when the movement toward the button to press was congruent. With a similar experiment, Bub et al. (2008) showed that both functional and volumetric knowledge of an object is automatically evoked by seeing the object or reading the word denoting that object. The interpretation of these studies goes in the same direction: sensory-motor variables modulate behavior and therefore, sensory and motor properties must play a role in the understanding of the action. Interestingly, these results, together with the results of other behavioral studies (Brass et al., 2001; Craighero et al., 2002; Tucker and Ellis, 2004), suggest that not only seeing an object but also reading the corresponding word automatically leads to a simulation of activity and, therefore, to the activation of the motor/premotor system.

Besides behavioral studies, transcranial magnetic stimulation (TMS) has emerged as a technique to investigate more directly the role of the motor system and precentral areas in action understanding. Delivering a TMS pulse is considered to be adding noise to the neural activity of a specific cortical region (for a critical assessment, see Perini et al., 2012). The result is the disruption of the normal activity of the stimulated region and, consequently, a behavioral change (Walsh and Rushworth, 1999; Harris et al., 2008 Rossi et al., 2009; Kammer et al., 2001; Groppa et al., 2012; Papeo et al., 2013; Sauve and Crowther, 2012; Rossini et al., 2015). Previous TMS studies on action observation and action understanding mainly focused on the behavioral and electrophysiological (motor evoked potential – MEP) consequences of the stimulation of motor and premotor cortex. The rationale of this approach is that TMS applied over the motor cortex (M1) elicits a contraction in the contralateral muscles (MEP) that can be recorded by applying an electrode on the distal muscle. The size of this MEP changes according to the cortical excitability. This means that the MEP is a non-invasive measure of the state of excitability of the cortico-spinal system and therefore a measure of the sensitivity of M1 (Bestmann and Krakauer, 2014). Using this method, it has been shown that the observation of hand movements elicits an increase of the MEP in the contralateral muscle involved in the observed movement (Fadiga and Rizzolatti, 1995; Strafella and Paus, 2000; Maeda et al., 2002). This result demonstrates that the activity of the motor system is modulated by the observation of actions.

Additional studies have shown that M1 is sensitive not only to the direct observation (and execution) of actions but also to action-related language. In a

TMS study by Buccino et al. (2005), single-pulse TMS was delivered to either the hand or the foot motor area in the left hemisphere while participants were listening to sentences containing hand or foot actions. MEPs induced by TMS were recorded. Results showed a decrease of the MEP amplitude recorded from hand muscles when listening to hand-related sentences. The same pattern was observed in the foot muscles when listening to foot-related sentences. Similar experiments have shown a sensitivity of M1 to language, in particular when such language is action-related (Fadiga et al., 2002; Watkins et al., 2003; Pulvermüller et al., 2005; Glenberg et al., 2008; Tremblay et al., 2012).

The sensitivity of M1 to visual actions and to action-related language was interpreted as evidence of its involvement in the semantic processing of actions and was used in support of the motor theories of action understanding. However, it is worth noting that the cited studies show a certain variability in terms of direction of the effect (increased/decreased activity), of the direction of the responses (facilitatory/inhibitory effect), and of the materials associated with the effect (for a review, see Papeo et al., 2013). These differences make the cited studies very difficult to compare and interpret.

Functional magnetic resonance imaging (fMRI) studies have been used to examine whether the processing of action-related stimuli leads to the activation of the precentral gyrus (for reviews, see Martin, 2007; Pulvermüller and Fadiga, 2010). For example, Tettamanti et al. (2005) showed that listening to action-related sentences, in comparison to abstract sentences, activates a fronto-parieto-temporal network in the left hemisphere, including the pars opercularis of the inferior frontal gyrus, premotor cortex, inferior parietal lobule, intraparietal sulcus, and posterior middle temporal gyrus (pMTG). This study is cited in support of the motor theories, as the authors highlight the activation found in the frontal and parietal areas, which “subserve action execution and observation" (p. 273), as claimed by the mirror neuron hypothesis. However, the importance of the pMTG was underestimated by the authors who relegated its function to a mere visuo-motor analysis of the stimulus.

Despite the apparent strength of the above-presented studies in showing a specific modulation of the motor and premotor system during conceptual processing of action-related stimuli, and despite the interesting and seemingly elegant interpretation that is given around the role of so-called "mirror neurons", we cannot conclude, based on the empirical results discussed, that the observed motor activity constitutes the semantic analysis of the stimulus. The demonstration of a sensitivity of the motor cortex to action-related stimuli does not provide a measure of action understanding. Alternatively, the motor activation might occur 1) in order to keep a state of readiness in the event that a motor act is needed (Negri et al., 2007), 2) as a consequence of feedback loops from other areas (Mahon and Caramazza, 2008), 3) in concert with automatic motor imagery of the action (Willems and Hagoort, 2007), 4) in order to enrich the processing of the action (Negri et al., 2007; Spaulding, 2012), 5) or even as an associative mechanisms to conceptual areas (Hickok, 2009).

Besides the different possible explanations regarding the role of the motor activation during action understanding, it is interesting to notice that motor and premotor regions are not the only areas that are recruited when attending to action-related stimuli. Numerous fMRI studies have shown that a wide network of temporal, parietal, and frontal areas is recruited when we observe an action (Tettamanti et al., 2005; Villareal et al., 2008, Xu et al., 2009; Keyser and Gazzola, 2009; Lingnau and Petris, 2009, Andric et al. 2013; for reviews, see e.g., Caspers et al., 2010; Turella et al., 2013; **Figure 1.1**). This network is called the Action-Observation Network (AON) and includes inferior frontal areas (BA44/BA45), lateral dorsal premotor cortex (BA6), supplementary motor area (BA6), rostral inferior parietal lobule (area PFt), primary somatosensory cortex (BA1/2), superior parietal (area 7A), intraparietal cortex (area hIP3), posterior middle temporal gyrus (pMTG) visual area V5, and fusiform face area/fusiform body area (FFA/FBA). The role of these other areas, in particular the role of the temporal cortex in action understanding, has often been neglected in previous studies.

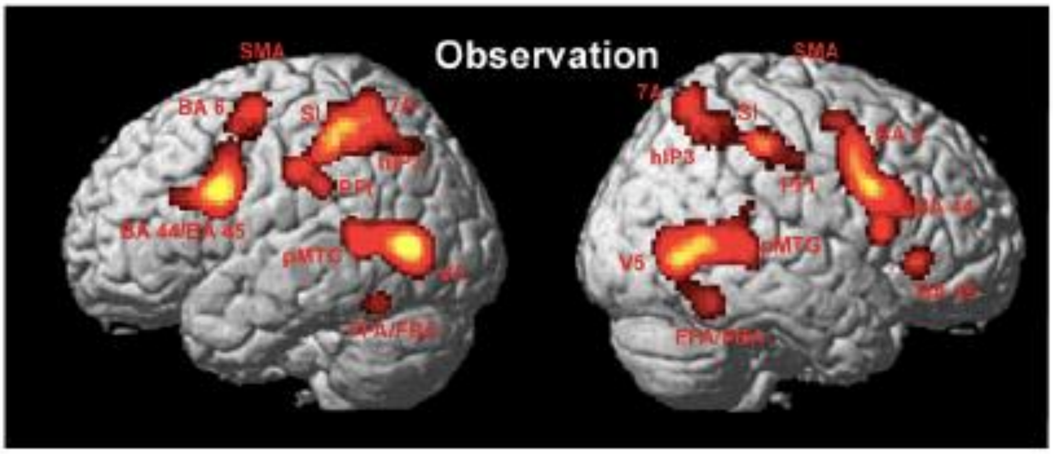


Figure 1.1. Brain regions comprising the Action-Observation network (AON). The network includes: inferior frontal areas (BA44/BA45), lateral dorsal premotor cortex (BA6), supplementary motor area (BA6), rostral inferior parietal lobule (area PFt), primary somatosensory cortex (BA1/2), superior parietal (area 7A), intraparietal cortex (area hIP3), posterior middle temporal gyrus (pMTG) visual area V5, and fusiform face area/fusiform body area (FFA/FBA). Figure from Caspers et al., 2010.

1.3 Cognitive theories of action understanding

Supporters of the so-called cognitive theories of action understanding offer an alternative view regarding how our brain associates meaning to an observed action, irrespective of whether we are observing an action or reading or listening to the corresponding word. A key point of this hypothesis is that conceptual representations are abstract and stored in conceptual areas outside the sensory and motor systems (Mahon and Caramazza, 2008; Hickok, 2009). In other words, semantic knowledge about actions is not dependent on the specific motor program necessary to perform them but rather is abstract and is housed in non-motor regions.

As we have seen in the previous section, motor regions are responding during different tasks involving action-related stimuli. On the one hand, Rizzolatti et al. (2001 -p. 6610) stated that an "action is understood when its observation causes the motor system of the observer to *resonance*". However, the *resonance* of the motor system while attending to action-related stimuli may be a consequence of associative linkage with proper conceptual areas (Mahon and Caramazza, 2008; Hickok, 2009; Spaulding, 2012). Recent neuropsychological studies examined the consequences of left hemisphere lesions on the performance of action-related tasks, in order to verify the necessity of motor processes in understanding the abstract meaning of actions. If action understanding and action execution rely on the same neural mechanism, as stated by motor theories, then both abilities should show impairment when these neural structures are damaged (Bauxbaum et al., 2005; Pazzaglia et al., 2008). Conversely, numerous studies provided evidence of double dissociations between action recognition and action production (Rumiati et al., 2001; Negri et al., 2007; Kalénine et al., 2010; Urgesi et al., 2014), challenging the basic assumption of the motor theories of action understanding.

In a recent study, individuals born with absent or severely shortened upper limbs (bilateral upper limb dysplasia) but cognitively and neurologically healthy, were tested in several tasks involving pantomiming, naming, learning new actions, predicting and anticipating observed actions, performed with the upper limbs (Vannuscorps and Caramazza, 2016). The logic behind this study is that if motor simulation is necessary for an efficient processing of an action, then an individual that has never developed any motor representation due to a pathological condition, as the one described above, should present some deficits in the interpretation of that actions. The authors found that participants affected by bilateral upper limb dysplasia were able to perform each task as efficiently as healthy control participants suggesting that a motor representation is not needed to understand actions. Neuropsychological studies have also offered interesting data regarding the location of the lesion and the corresponding deficits. As an example, Kalénine et al. (2010) examined the performance of 43 patients suffering from left hemisphere stroke in two different tasks: a semantic recognition task and a spatial recognition task. During the semantic recognition task, participants were instructed to indicate which of two gestures (depicting two different actions) corresponded to a previously presented verb; the spatial recognition task was identical, except that the two gestures, this time referring to the same action, differed by a perceptual component (hand posture, arm posture or amplitude/timing components). Researchers performed a whole-brain voxel-based lesion symptom mapping (VLSM) analysis. Thanks to this method, it is possible to assess whether behavioral deficits are reliably predicted by lesions of a particular brain region. The authors found that the ability to associate a meaningful label (semantic recognition task) was associated with the integrity of the posterior temporal lobe (BA21, BA22, BA37) but not IFG. Performance in the spatial recognition task was instead dependent on the integrity of the IPL. This study demonstrates on the one hand that IFG is not critical for gesture recognition, while on the other hand the posterior temporal lobe (as many fMRI study show) is fundamental for accessing the meaning of an action. Tarhan et al. (2015) reached a very similar conclusion after examining 131 left-hemisphere stroke patients (the largest lesion-based investigation to date in this topic). Participants were tested on two tasks involving action production and action recognition. In the action production task they were presented with some familiar tools and were asked to pantomime the use of each tool; the action recognition task was similar to the one used by Kalénine et al. (2010) described above. Researchers reported a proportion of patients showing double dissociations between action production and action recognition. Furthermore, impaired performance on action recognition was associated with lesions to the lateral temporo-occipital cortex, while impaired performance in action production was associated with lesions to sensory-motor cortex and IPL, suggesting a role of the pMTG in the association of actions and meanings and a role of IPL in encoding object-related movements.

Taken together, these results offer a new perspective on the role of temporal areas in action understanding. In order to investigate which kind of information is represented in these areas, recent fMRI studies have made use of a new method to analyze fMRI data, called multi-voxel pattern analysis (MVPA - for a review on the method, see Haxby, 2012). In contrast to classical univariate analyses that examine amplitude differences between experimental conditions while collapsing the signal across voxels (thus allowing claims about the *involvement* of a region in a task), MVPA examines patterns of activity across voxels to investigate the *representational content* contained in that region (Norman et al., 2006; Kriegeskorte et al., 2006; Mur et al., 2009). In particular, it has been suggested that if a region is involved in retrieving the meaning of an action, then this region should encode high-level abstract representation of actions (Hamilton and Grafton, 2008), irrespective of the low-level concrete features (e.g., how the action is performed, from which perspective, in which context,…). For example, the action *opening a box* might be performed with different kinematics, in different contexts, using different kind of boxes, but the concept of the action stays the same: opening a box. Different brain regions might encode either high-level abstract or low-level concrete representation. High-level abstract representations (Kiefer and Pulvermüller, 2012; Tulving, 1972) generalize across situations to create a concept that is disjoint from the specific low-level feature we are observing in that particular moment. Investigating which brain regions encode abstract representations of actions is fundamental to understand how a meaning is associated to an event.

In a recent fMRI study, Wurm and Lingnau (2015) used fMRI-based MVPA to identify action representations at different levels of abstraction. Inside the MR scanner, participants were asked to watch videos of actions involving different objects (bottles, containers) and different kinematics (i.e., opening and closing different types of bottles and containers, requiring different kinematics). The authors investigated a concrete level (e.g., opening and closing a specific box, using specific kinematics) and an abstract level that generalizes across objects and kinematics. The authors found that actions were represented both at a concrete and abstract level in LOTC, while PMv encoded actions at an abstract level only. Several experiments, using the same methodology, reached similar conclusions, indicating that occipito-temporal but not precentral areas encode abstract action concepts (Oosterhof et al., 2010, 2012, 2013; Wurm et al., 2015).

The studies described above suggest that the LOTC, and not the precentral gyrus, encode action information at an abstract level. However, the dynamics underlying concrete and abstract representations are not well understood. Researchers that argue in favor of motor theories of action understanding often claim that motor activation is *fast* and *automatic* and therefore *necessary* to understand an action (Pulvermüller et al., 2005, Kiefer and Pulvermüller, 2012). Magnetoencephelography (MEG) offers insight into this question by recording brain activity with millisecond resolution. Using MVPA of MEG data, Tucciarelli et al. (2015) examined the dynamics underlying action understanding. Participants were asked to observe short video clips of a reach-to-point or reach-to-grasp movement. The authors investigated which regions are able to distinguish between the two movements irrespective of the direction of the movement (towards the left or the right) and the effector used (left or right hand). They observed that the lateral occipitotemporal cortex (LOTC), in comparison to precentral regions, had access to abstract action representation significantly earlier and, most importantly, this activation reflected the moment in time when there was enough information to discriminate between the two hand actions.

The studies cited above do not provide insights into the question of whether the representations obtained in the LOTC are necessary for action understanding. Using TMS, Papeo et al. (2014) found that the perturbation of the lpMTG, a subportion of the LOTC, through repetitive TMS (rTMS) disrupts the semantic processing of action verbs in comparison to nouns. Moreover, by recording motor-evoked potentials (MEPs), they also showed that the rTMS on the lpMTG disrupts the verb-related activity of the motor cortex. With this study, they demonstrated that the motor activation during the semantic analysis of action is a consequence of the activation of the lpMTG, and that this area is therefore fundamental for the semantic representation of action verbs.

The studies discussed above suggest that the conceptual knowledge of actions is represented within the temporal lobe, and not in motor and premotor areas, as suggested by motor theories of action understanding. Moreover, they show that occipito-temporal areas, and the MTG in particular, are sensitive both to visually presented actions (Kable et al., 2002; Xu et al., 2009; Andric et al., 2013) and to linguistic stimuli (Kable et al., 2005; Peelen et al., 2012; Papeo et al., 2014). In the next section I will describe some of these studies in more detail, investigating the relationship between action and language in the brain.

1.4 Action and language

As we have seen in the previous two sections, many studies have investigated the neural mechanisms underlying the processing of meaningful actions using either visual actions (images, line drawing, gestures) or linguistic stimuli (verbs or action-related sentences). These studies aimed at defining the neural substrates involved in the recognition of an action. Moreover, results of these studies have been used to delineate two different theoretical frameworks (motor and cognitive theories, see Chapter 1.2, 1.3) that aim at explaining the role of different brain areas in action understanding.

The observation that the same areas that are sensitive to actions are also sensitive to linguistic material, also indicates a possible interplay between action and language in the brain (Willems and Hagoort, 2007; Papeo and Rumiati, 2013). Recent studies investigating the link between action and language in the brain, have focused on communicative stimuli (i.e., symbolic gestures such as pantomimes or emblems, and language) to see whether these forms of communication are processed by the same neural system (McNeill, 1992; Xu et al., 2009; Andric et al., 2013).

Pantomimes are mimic gestures that simulate actions or tool use, such as *playing guitar* or *playing basketball*. *Emblems* are symbolic conventional gestures, such as *thumb up*. They are usually used to communicate intentions or physiological/psychological states to other persons, especially when external circumstances prevent the use of speech (Ekman and Friesen, 1972; Molnar-Szakacs et al., 2007). *Speech* is the most common form of communication between human beings; it consists of symbols and is representative of a language. The commonality of these three categories is that all of them are communicative and symbolic. However, as I will describe in more detail in the following section, they present differences not only in terms of input modality (visual and auditory), but also in terms of the relationship between the stimulus and the linguistic properties, and in terms of arbitrariness of the symbol-meaning relationship.

The first difference was described by Kendon (1988; see also McNeill, 2005) who suggested the existence of a continuum between those stimuli (**Figure 1.2, top**) in terms of the presence or absence of linguistic properties. In particular, pantomimes do not seem to obey any system constraints in the sense that the relation between movement and meaning is not rigid: a slight variation in the movement kinematics does not affect the meaning of the gesture. By contrast, emblems can be considered a language of the body with a strict form-meaning relationship and rules of usage that are typical of words (Gullberg, 2006), but at the same time do not obey all the rules of a full linguistic system and do not have syntax.

The second difference refers to the arbitrariness of the relation between the symbol and its meaning (**Figure 1.2, bottom**). Pantomimes are understood by associating the gesture observed to the action to which it refers. For this reason, pantomimes present a concrete link between the sign (the gesture; e.g., mimicking *playing guitar*) and the meaning (the action it refers; e.g., *playing guitar*). By contrast, emblems are gestures with a higher level of abstractness. In fact, emblems do not faithfully reproduce actions, but embed a symbolic meaning that is socially learnt (Ekman and Friesen, 1969; Ekman, 2004; Gullberg, 2006). For some emblems the link between the sign (the gesture; e.g., *thumbs up*) and the meaning (e.g., *agreeing*) is abstract in the sense that gesture do not contain any aspect of the real-word referent. For other emblems, in particular those describing bodily actions such as *drinking*, the gesture symbolically recalls the object involved in the action, although it does not pantomime it. In this case, the link between the sign and the meaning is still abstract but to a lesser degree.

Spoken words consists of strings of sounds with an abstract relation to the object or action they refer.

***Continuum**: relationship to linguistic properties*

Pantomime	**Emblems**	**Language**
Linguistic properties absent	**Some linguistic properties present**	**Linguistic properties present**

***Continuum**: relationship to arbitrariness symbol-meaning*

Pantomime	**Emblems**	**Language**
Concrete link between stimulus and meaning	**Partially abstract link between stimulus and meaning**	**Abstract link between stimulus and meaning**

Figure 1.2. Continuum that shows the relationship between pantomimes, emblems and speech. (Revised from McNeill, 2005).

Xu et al. (2009) examined whether symbolic gestures (pantomimes and emblems) and spoken language are processed by a common neural system. Authors asked participants to observe clips of an actor performing emblems (such as *«thumb up»*), pantomimes (such as «*putting on a ring*»), and the corresponding spoken glosses. Results showed a left-lateralized network of inferior frontal and posterior temporal regions (**Figure 1.3**). They therefore suggested that these region contain the conceptual representation of symbolic gestures and spoken words, and they speculated on a broader role of the anterior and posterior perisylvian areas (by now considered the core of the language system) as a place where meaning and symbols are linked together, irrespective of the input modality. Similarly, Andric et al. (2013) investigated the brain responses to emblems, speech and grasping actions. In particular, the authors investigated which brain areas are sensitive to symbolic meaning irrespective of whether it is expressed through an emblem or through speech, and which brain areas are sensitive to hand actions regardless of the nature of the gesture (symbolic in the case of emblems, non-symbolic in the case of grasping actions). The authors showed that lateral temporal and inferior frontal regions are sensitive to symbolic meaning (emblems and speech), while parietal and premotor regions where recruited when observing either emblems or grasping actions.

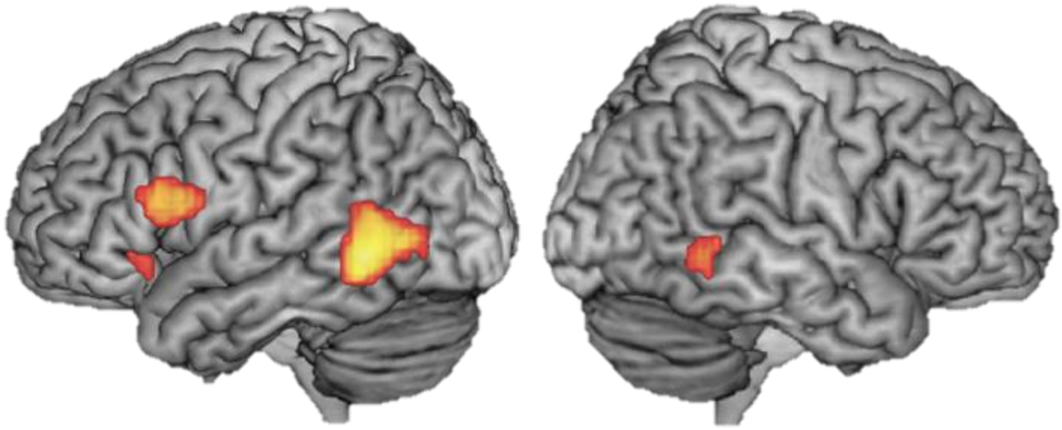


Figure 1.3. Common areas responding to symbolic gestures and spoken language. Random-effects conjunction: [(pantomime gestures + emblem gestures) - nonsense gesture controls] ∩ [(pantomime glosses + emblem glosses) - pseudo sound controls]. Figure adapted from Xu et al., 2009.

These results show that middle temporal gyrus is sensitive to communicative stimuli when presented in both the auditory (speech) and visual (gestures) modality and that there is a mapping between symbols that are presented either visually or aurally. Moreover, several studies have suggested the existence of a posterior-to-anterior gradient reflecting the concrete (gestural) – to – abstract (linguistic) properties of the stimuli (Kable et al., 2005; Chatterjee, 2008; Watson et al., 2013; Weiner and Grill-Spector, 2013; Tahran et al., 2015; Lingnau and Downing, 2015). However, none of these studies investigated how different categories of stimuli are organized in the MTG and which are the main peculiarity of this gradient. Because of their properties described above, pantomimes, emblems, and spoken words seem to be suitable candidates for such an investigation.

1.5 Conclusion

In the previous sections, I provided an overview on the current debate between motor and cognitive theories on action understanding, offering a theoretical description as well as empirical results typically cited in support of either of the two major opposing views. In light of the studies described above, motor theories of action understanding fail to stand up to the scientific evidence. On the one hand side, the results of behavioral and fMRI studies cited in favor of motor theories are often also in line with the alternative view put forward by cognitive theories. On the other hand, the finding of double dissociations between action production and action understanding in patient population contrast with one of the basic assumptions of the motor theories.

As empirical studies fail to support a purely embodied view of action understanding, alternative hypotheses started to arise regarding which brain regions encode abstract representation of actions. Recent studies identified the LOTC and the MTG as brain areas that encode knowledge about actions (Villareal et al., 2008; Xu et al., 2009; Lingnau and Petris, 2009; Andric et al., 2013; Papeo et al., 2014; Wurm and Lingnau, 2015; Lingnau and Downing, 2015; Tucciarelli et al., 2015).

The MTG, in particular, seems to be sensitive both to visually presented gestures (Xu et al., 2009; Andric et al., 2013), and to words (Bedny et al., 2008; Peelen et al., 2012; Kable et al., 2012; Papeo et al., 2014; Papeo and Lingnau, 2015). Researchers hypothesized the existence of one integrated system of communication (McNeill, 1992; Xu et al., 2009) and suggested the presence of an organizing principle that consists of a concrete-to-abstract gradient along the MTG (for a recent review, see Lingnau and Downing, 2015). However, differentiating the neural processes underlying different stimuli still needs to be investigated.

1.6 Current work

In the next Chapters, I illustrate three studies that I carried out during my time as a doctoral student. The first (Chapter 2) is a norming study that aimed at creating a standardized data set of video clips of meaningful and meaningless gestures. The goal was to provide researchers with a large number of well-controlled stimuli in order to promote replicability between studies. To our knowledge, no other dataset assessing different categories of gestures is present in the literature. The second study (Chapter 3) is divided into 2 experiments: an fMRI and an fMRI-guided TMS study. The first experiment aimed at characterizing the neural network associated with the understanding of meaningful gestures. The second experiment, aimed at contrasting the role of temporal and precentral areas in action understanding localized through the fMRI experiment. The main idea was to observe the consequences of applying rTMS to the left MTG and left PMv while processing semantic aspects of actions. In the third study (Chapter 4), I focused on how different kinds of gestures and words are represented in the middle temporal gyrus. The first experiment of the second study (Chapter 3 – Experiment 1) and the third study (Chapter 4) are based on different analyses carried out on the same dataset and, therefore, participant, stimuli, and design do not differ.

Chapter 2

A norming study of 228 high-quality video clips of pantomimes, emblems and meaningless gestures

Adapted from:

Agostini, B., Papeo, L., & Lingnau, A. (2016). *A norming study of 228 high-quality video clips of pantomimes, emblems and meaningless gestures.* Under review.

2.1 Abstract

Recent years have witnessed a growing interest in behavioural and neuroimaging studies on pantomimes and emblems, but well-controlled stimuli are scarce. This study describes a standardized data set of 228 video clips of an actress performing pantomimes (gestures that directly mimic actions or objects, e.g., *playing guitar*), emblems (symbolic communicative gestures, e.g., *thumb up*) and meaningless gestures. One hundred and thirty raters (sixty-two Italian- and sixty-eight non-Italian speakers) rated the meaningfulness of the stimuli, and provided names and descriptions for each action. Here we provide the results of those rating and norming measures. To the best of our knowledge, this is the first study providing a well-controlled set of normed meaningful and meaningless gestures. The stimuli are made available by the authors upon request.

2.2 Introduction

Gestures are complex movements that involve hands, fingers, and arms. Gestures occur every day, in a spontaneous or intentional manner, and may accompany speech but can also be used alone (McNeill, 2005; Molnar-Szakacs et al., 2007). A person walking on the street might wave her hand to greet a friend passing by on the other side. Moreover, a child might imitate an action using gestures while playing charades with her friends.

Over the last decade, there has been an increasing number of studies investigating the cognitive and neural mechanisms underlying the observation and recognition of gestures, in particular pantomimes and emblems (e.g., Molnar-Szakacs et al., 2007; Villareal et al., 2008; Xu et al., 2009; Papeo et al., 2010; Papeo and Rumiati, 2013; Andric et al., 2013; Kalénine et al., 2010; Vannuscorps and Caramazza, 2016; Tarhan et al., 2015). While all these studies have used video clips of actors performing gestures, the stimulus materials contain high variability. This variability is reflected in many features, such as the position of the actor (standing or sitting), the focus of the video (whole body or specific parts), the manner of dressing, the background as well as the duration of the video clips. Moreover, each author selected the set of stimuli based on different criteria (e.g., different rating/familiarization methods). All these sources of variability are obstacles when trying to replicate studies and compare results.

To the best of our knowledge, the only publicly available dataset assessing video clips of actions focuses on ballet dance movements (Christensen et al., 2014); thus there remains a lack of available databases of video clips of meaningful and meaningless gestures. The goal of the presented study is to promote replicability and to allow direct comparison between different categories of gestures by providing a large number of well-controlled stimuli from different categories that can be used for behavioural and neuroimaging studies. In particular, we included commonly used meaningful gestures (i.e., pantomimes and emblems), and meaningless gestures (i.e., movements that do not carry any shared meaning). Stimuli come with norms and detailed spreadsheets.

Pantomimes are mimic gestures that simulate actions or tool use, such as *playing guitar* or *playing basketball*.
Emblems are conventional gestures, such as *thumb up*. Emblems are usually used to communicate intentions or physiological/psychological states to other persons, especially when external circumstances prevent the use of speech (Ekman and Friesen, 1972; Molnar-Szakacs et al., 2007).
Meaningless gestures are movements that are not associated with any meaning.

Pantomimes and emblems are both communicative and symbolic gestures. However, the two classes present a different arbitrariness of sign-meaning relationship. Pantomimes are reproductions of common actions, the meaning of which is usually mediated by a general process of visual recognition (Poggi and Zomparelli, 1987; Ekman, 2004). They are understood by associating the observed gesture with the action to which it refers. For this reason, pantomimes present a concrete link between the sign (the gesture; e.g., mimicking *playing guitar*) and the meaning (the referential action; e.g., *playing guitar*). By contrast, emblems are gestures with a higher level of abstractness. In fact, emblems are gestures that do not faithfully reproduce an action, but embed a symbolic meaning that is socially learned (Ekman and Friesen, 1969; Ekman, 2004; Gullberg, 2006). Therefore, the sign (the gesture; e.g., *thumbs up*) and the meaning (e.g., *approving*) are linked together on a more abstract level. In summary, pantomimes and emblems can be seen as part of a continuum of communicative signs/symbols with a different degree of abstractness/arbitrariness.

The gestures we recorded for the purpose of this study were initially based on the Italian culture. However, they have been evaluated by a large group of Italian and non-Italian raters who provided responses about the meaningfulness of the gestures and, for gestures judged as meaningful, a naming and a description of their meaning. By including both Italian and non-Italian raters, we aimed to control for possible cultural aspects and to be able to generalize and standardize the dataset across cultures where possible.

Results showed that the meaning of pantomimes was largely shared by Italian and non-Italian raters. Emblems rather showed higher variability between Italians and non-Italians, both in terms of number of gestures judged as meaningful and in terms of the associated meaning. Because of these differences, we provide a unique data set of stimuli for the pantomimes and the meaningless gestures categories, combining the results from Italian and non-Italian raters, and two data sets of emblems, one specific for Italians and one that generalizes across other cultures (non-Italian raters were divided as follows: 89% American, 7% Indian, 1% English, 1% Chinese, 1% Russian).

As of today, these stimuli have been used to investigate the consequences of the impairment of the motor areas in individuals affected by dysplasia (Vannuscorps and Caramazza, 2016), and to investigate the organizing principles of the middle temporal gyrus (Agostini et al., in prep).

2.3 Methods

Video recording and editing

278 gestures were recorded using a Canon 5D MK2 camera using a temporal resolution of 23 frames per second and a spatial resolution of 1920x1088 pixels.

The camera was positioned on a tripod approximately 2 meters away from the actress. All videos were recorded with the same artificial light condition, on a white background, and the actress, a female Italian and native speaker, was wearing neutral clothes. The actress was instructed to keep both arms relaxed along the body and to return to the same neutral start position after completing the gesture. When performing the gesture, the actress made no noticeable facial movements and directed her gaze congruently with the performed action: toward the camera during emblems and toward the object during pantomimes. Gestures were divided into three main categories: pantomimes (iconic representations of object-directed actions; e.g., *wearing a ring*) emblems (symbolic communicative gestures, e.g., *thumb up)* and meaningless gestures (**Figure 2.1**). Specifically, we recorded 90 pantomimes, 108 emblems and 80 meaningless gestures. Meaningless gestures were created in two different ways: 39 were devised on the basis of a meaningful gesture (D = derived), whereas 41 were invented from scratch (ND = non-derived). In the first case, a meaningful gesture was taken and either the spatial feature or the direction of the movement were slightly changed to obtain a meaningless gesture. For example, the meaningless gesture derived from *using binocular* consisted of the actress pretending to have a binocular in her hands but instead of directing the gesture towards her eyes, she directed the gesture towards her shoulder. In the second case, the meaningless gesture had no relationship with any meaningful gesture.

In a postproduction phase, videos were cut to a duration of 2.5 seconds, cropped to 1920x1080 pixels, and converted to a frame rate of 30 frames per second. Moreover, sound was removed from the video clips.

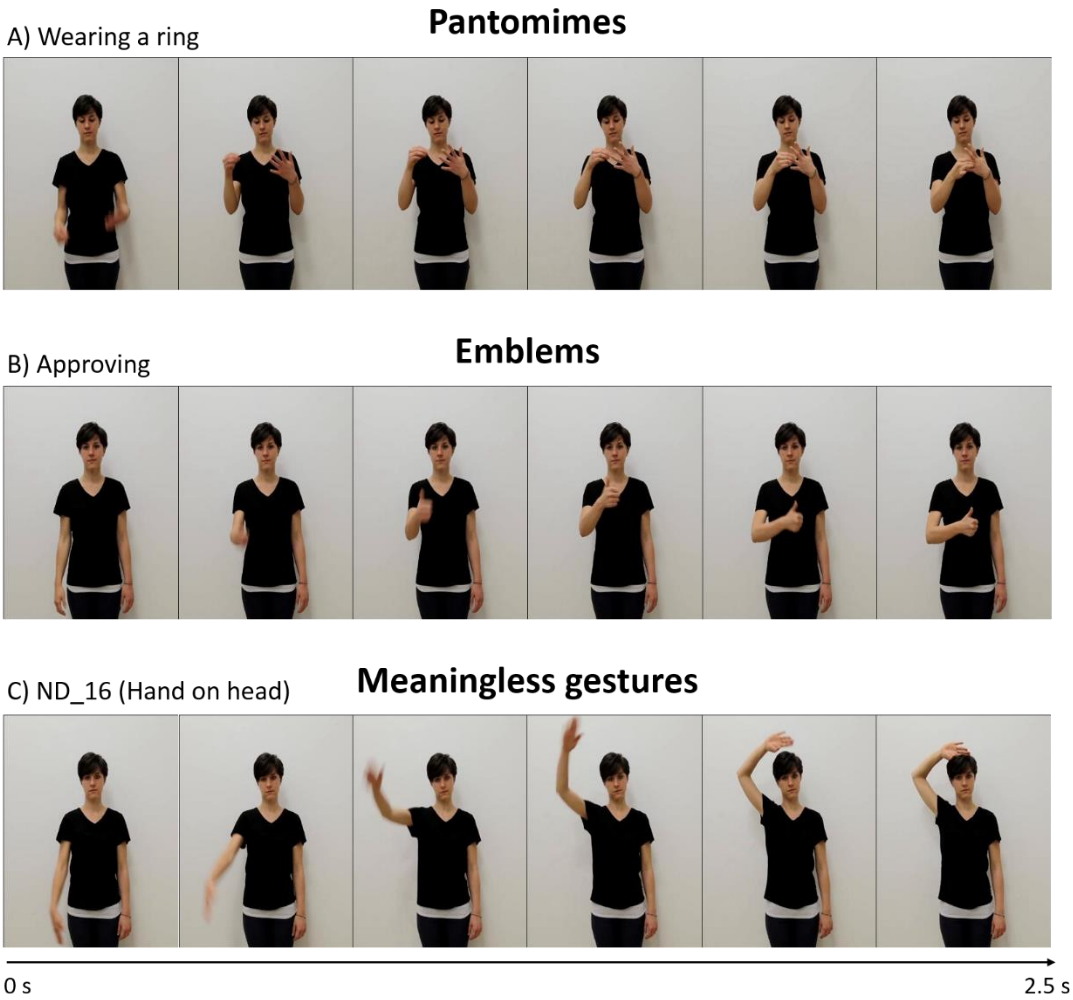


Figure 2.1. Example stimuli. Still frames from video clips showing pantomimes (panel A, i.e., *wearing a ring* - gestures that directly mimic actions or objects), emblems (panel B, i.e., *thumb up* - symbolic communicative gestures), and meaningless gestures (panel C, i.e., *hand on head* - gestures that do not carry any meaning; ND = non-derived). Each video clip lasted 2.5 seconds.

Raters

Videos were rated by 62 native Italian and 82 non-Italian raters, for a total of 144 raters. Of the 62 Italian raters, 40 were female and 22 were male with a mean age of 31.2 (SD = 12.4; range 18-70). Of the 82 non-Italian raters (73 American, 6 Indian, 1 English, 1 Chinese, 1 Russian), 42 were females and 40 were males with a mean age of 37.4 (SD = 10.7; range 18-70). Fourteen raters (13 American, 1 Indian) were excluded from the study due to unreliable rating responses (see section *Methods - Data Analysis – Meaningfulness*). Raters gave informed consent before participation in the study. The study was approved by the Ethics Committee for research involving human subjects of the University of Trento and conforms to the Declaration of Helsinki.

Procedure

Two hundred seventy-eight video clips were divided into four groups of stimuli, namely list_1, list_2, list_3, list_4. Pantomimes, emblems, and meaningless gestures were equally distributed across the four lists. Raters were randomly assigned to one of the four groups and were divided as follows: list_1) n=32; 18 females, 14 males (15 Italians: 10 females, 5 males); list_2) n=33; 18 females, 15 males (15 Italians: 10 females, 5 males); list_3) n=31; 18 females, 13 males (15 Italians: 11 females, 4 males); list_4) n=33; 19 females, 14 males (17 Italians: 9 females, 8 males). For each item in the group, raters were asked to (i) rate the meaningfulness of the gesture, (ii) name the gesture, and (iii) verbally describe the gesture. Data were collected using an online rating site called Qualtrics (http://www.qualtrics.com/), administered to the participants either via email or through Amazon Mechanical Turk (https://www.mturk.com/). Video clips were presented on a black background, on the top part of the web page followed by the three task demands (**Figure 2.2**). At the beginning of each survey, one example of video clips for each category was provided. The example videos were not part of the actual survey. Raters did not have any time restriction and were allowed to view the video as many times as they wanted. The average time needed to complete the survey was 36 minutes and raters received $3 for their participation.

Meaningfulness. Raters were asked to indicate on a 7-point Likert scale whether each gesture was meaningful or not (1 = not meaningful at all; 7 = very meaningful). When the gesture was rated as meaningless (i.e., rated as '1'), raters were instructed to skip the next two questions and to move to the next video clip.

Naming task. Raters were asked to name each gesture by typing its name. They were instructed to only use one or two words if possible.

Description. Raters were asked to describe the meaning of the gesture in more detail. The aim of this task was to examine whether there is a shared consistent meaning despite a lack of shared consistent lexical entry in the naming task.

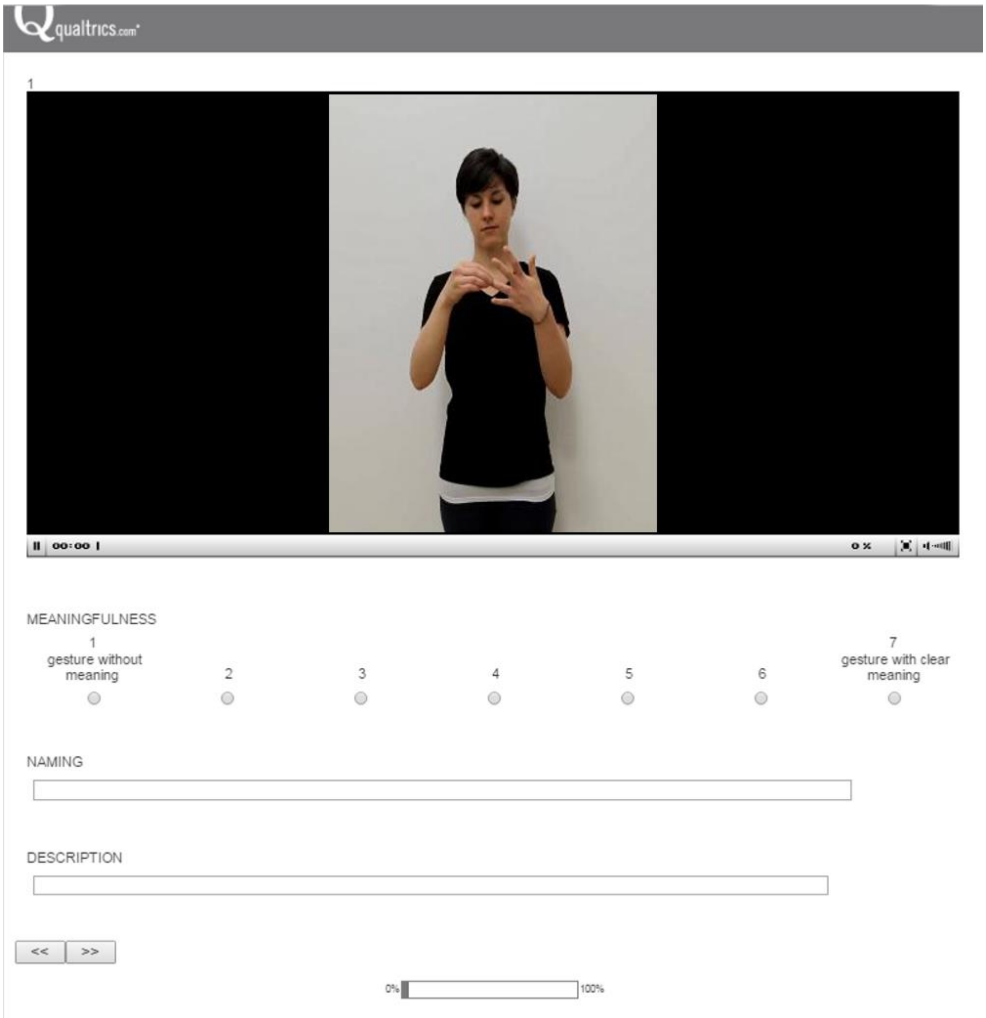


Figure 2.2. Example of a trial on the Qualtrics web page. The video clip was presented at the top of the page and raters provided responses to the three questions below before proceeding to the following video clip.

Data analysis

Gestures were recorded based on the Italian culture and, therefore, we expected Italian speaking raters to be more familiar with culturally driven video clips. Initially, data from the Italian and from the non-Italian speaking raters were kept separately to check for possible cultural differences. We found that both groups rated a similar number of pantomimes as meaningful and associated the same meaning to each gesture, making the results comparable (see *Results*). On the contrary, the number of emblems rated as meaningful, and the meaning associated to each gesture show a bigger variability between the two groups, meaning that emblems may be more susceptible to cultural differences. Therefore, for the analysis of the pantomimes and meaningless gestures categories, we collapsed data from Italian and non-Italian speaking raters, treating them as a unique *corpus*. By contrast, for the ratings of emblems, we kept the results from the Italian and non-Italian raters separate, obtaining two different sets.

Rating of Meaningfulness. As the instruction was to skip questions 2 and 3 if a video was considered as meaningless (i.e., meaningfulness rated as '1'), this might have encouraged some participants to rate many videos as meaningless to be able to move faster to the end of the survey and to obtain the payment. In each list, the proportion of gestures we considered to be meaningless was around the 35%. To prevent the inclusion of raters that provided unreliable responses, we therefore excluded results from raters that rated more than 60% of the video clips as meaningless (see section *Method – Raters*). For each video clip, we carried out the analysis in the following steps: 1) we performed an outlier analysis by calculating the *z*-value for each response and excluded any rating above or below two standard deviations from the mean; 2) we analysed the meaningfulness of each item by averaging the corresponding rating across raters; 4) we calculated mean and median of the meaningfulness; 3) meaningful gestures were considered those with a median >5, meaningless gestures were considered those with a median <3.

Naming and description. For each item we determined: 1) the most frequent meaning; 2) the percentage of raters providing an answer, calculated as the ratio between the number of *raters* that took part in the rating and the number of *respondents*, i.e., participants that rated the meaningfulness of the item as higher than 1 and therefore provided a name and a description; 3) two measures of meaning agreement: the percentage of raters that produced the canonical meaning (i.e., the most commonly recognized meaning) and the Shannon's diversity index (*H*). The first measure was calculated as the proportion between the number of raters providing the canonical meaning and the number of respondents for that particular item. The *H* index (also called Shannon entropy) was calculated based on the following formula:

$$H' = -\sum_{i=1}^{R} p_i \ln p_i$$

where R is the number of unique meanings provided and p_i is the proportion of raters that produced each unique meaning. This index provides the dispersion of the different meanings: *H* index is zero when there is maximum agreement, i.e., if only one meaning was provided by all respondents; it increases as the diversity of responses increases. Raters that judged an item as meaningless and thus did not provide a name or description of the meaning were not taken into account for the calculation of the meaning agreement.

Additional analysis on pantomimes. Pantomimes are transitive (i.e., object-directed) gestures. This means that the gesture involves the use of a manipulable object. We observed that many people, when asked to name the gesture, tended to report the name of the object. We therefore analysed and reported the proportion of raters that provided a verb, a noun, or another grammatical form in the naming description.

Additional analysis on emblems. Emblems are gestures that do not involve pantomiming the actual use of an object. An essential feature of emblems is their communicative intent. In particular, emblems can denote actions (e.g., *to clap*) or physical/psychological states (e.g., *thumb up*). Based on the naming and description, we distinguished between emblems that denote an action (emblems-event -EE), and emblems that denote a state (emblems-state -ES). It is important to note that the distinction between emblems-event and emblems-state was not set *a priori* but it was based on the responses given by the the two groups of raters (Italian and non-Italian). For this reason, depending on how Italian and non-Italian raters defined the gesture, the same video could denote an action or a state.

2.4 Results

Pantomimes. Of the 90 video clips representing pantomimes originally presented, 71 were rated as meaningful (median >5) by the Italian group and 73 by the non-Italian group of raters (**Figure 2.3A**). Both groups rated the same 62 video clips as meaningful and gave the same description to those videos.

Emblems. Of the 108 video clips representing emblems originally presented, 86 were rated as meaningful (median >5) by the Italian group, and 62 by the non-Italian group of raters. Both groups rated the same 56 video clips as meaningful. Of these 56 gestures, 46 were defined in a similar way by the two groups. The remaining ten gestures were associated with different meanings by the two groups. Interestingly, we noted that both Italian and non-Italian raters easily recognize emblems-event, the meaning of which is inferable by a bodily action (Italian raters rated 40 emblems-event as meaningful; non-Italian raters rated 34 emblems-event as meaningful). On the contrary, in comparison to Italian raters, non-Italian raters recognized far fewer emblems-state, the meanings of which cannot be inferred by the movement observed (Italian raters rated 46 emblems-state as meaningful; non-Italian raters 28 emblems-state rated as meaningful – **Figure 2.3B**).

Meaningless gestures. Of the 80 video clips representing meaningless gestures originally presented, 72 were rated as meaningless (median <3) by the Italian group, and 66 by the non-Italian group of raters (**Figure 2.3C**).

These data show a large variability in the number of emblems rated as meaningful by the two groups of raters, as well as in the description given around their meaning. The same variability was not observed for the rating of the pantomimes category: a similar number of pantomimes was rated by the two groups of raters, who also associated the same meaning to each gesture. This result allowed us to collapse data from Italian and non-Italian speaking raters when analyzing the pantomimes and meaningless gesture categories. By contrast, for the ratings of emblems, we kept the results from the Italian and non-Italian raters separate.

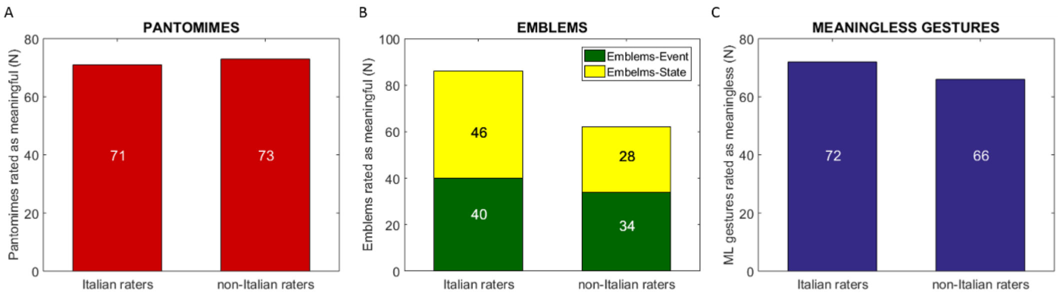


Figure 2.3. A) Number of pantomimes rated as meaningful by the Italian and the non-Italian group of raters. B) Number of emblems-event and emblems-state rated as meaningful by the Italian and the non-Italian group of raters. C) Number of meaningless gestures (ML) rated as meaningless by the Italian and the non-Italian group of raters.

Data sets. The ratings of the meaningfulness yielded a set of 67 pantomimes, 70 meaningless gestures, and 86 emblems rated by Italian raters, and 62 emblems rated by non-Italian raters.

The complete list of items with the corresponding statistics is reported in Appendix 1. **Table 2.1** provides a summary of the rating (for details see section *Methods – Data analysis*). For each item, the following information is provided: a) the most common English description of the item; b) the mean and median of the meaningfulness, rated on a scale from 1 to 7; c) the proportion (p) of raters providing an answer; d) the meaning agreement on the canonical name (expressed as both the proportion of raters providing that meaning and the H index); e) the proportion of raters that named the gesture by retrieving a verb, a noun, or another grammatical form (pantomimes only); f) mapping of the action onto an event (emblems-event -EE) or a state (emblems-state –ES; emblems only). In a few cases, the majority of raters provided a different meaning than the one we predicted while recording the videos. We indicate those cases with the symbol ** next to the gesture's name in the table. Moreover, we report both the meaning provided by the majority of raters and the expected meaning (e.g., the pantomime we recorded as *cutting with scissors* was instead identified as *stapling*; in this case, the name we reported is "*stapling** was: cutting with scissors*").

Table 2.1. Summary of the statistical analysis carried out on the stimuli. The first five items in alphabetic order, for each category are shown. The full dataset is provided in Appendix 1.

Number of order and Item	Meaningfulness (from 1 to 7) Mean	Median	Proportion of raters providing an answer	Canonical Meaning Agreement		Verb	Nouns	Other form
English name	*M*	*M*	*p*	*p*	*H*	*p*	*p*	*p*
1. PANTOMIMES								
1.1 applying make up	4.67	5	79	65	0.77	62	38	
1.2 archery	6.30	6.5	100	68	0.98	26	74	
1.3 blowing a whistle	5.45	6.5	88	54	1.06	39	61	
1.4 blowing nose	5.44	6	88	68	1.02	68	32	
1.5 calling on a telephone	6.88	7	97	100	0.00	33	63	4
			...					
2. MEANINGLESS GESTURES								
2.1 D_1 (peeling a banana)	1.31	1	19	33	1.56			
2.2 D_2 (using binoculars)	1.33	1	19	67	0.87			
2.3 D_3 (pouring from a bottle)	1.37	1	16	60	0.95			
2.4 D_4 (sweeping)	1.00	1	3	100	0.00			
2.5 D_5 (combing)	1.17	1	13	25	1.39			
			...					

Number of order and Item	Meaningfulness (from 1 to 7) Mean	Median	Proportion of raters providing an answer	Canonical Meaning Agreement		Emblems Event	Emblems State
English name	*M*	*M*	*p*	*p*	*H*	*EE*	*ES*
3. EMBLEMS (Italian raters)							
3.1 agreement** was: congratulation	6.17	7	80	67	0.98		X
3.2 anger	5.93	6.5	80	58	0.68		X
3.3 approving	6.43	7	93	100	0.00		X
3.4 arresting	6.64	7	93	100	0.00	X	
3.5 being late	6.86	7	100	73	0.50		X
			...				
4. EMBLEMS (non-Italian raters)							
4.1 anger	5.93	7	83	60	1.30		X
4.2 approving	6.20	6	94	100	0.00		X
4.3 being late	6.13	6.5	94	59	0.65		X
4.4 bowing	6.13	7	94	100	0.00	X	
4.5 calling	6.87	7	100	100	0.00	X	
			...				

Appendix 2 reports the proportion (in brackets) of canonical meanings and alternative meanings (see **Table 2.2**). 'NA' indicates the proportion of raters that did not provide an answer, i.e., those raters that judged the gesture as meaningless. In some cases, the same rater provided two or more different names for the same video (e.g., *"screwing or unscrewing"*). We considered both names as different answers and, therefore, the sum of the proportions of these particular items is sometimes more than 100. We indicate these cases with one asterisk (*).

Table 2.2. Summary of the naming analysis carried out on the stimuli. The first three items for each category are shown. The numbers in brackets refer to proportions. The full dataset is provided in Appendix 2.

Item	Canonical meaning	Alternative meanings	NA	*
1. PANTOMIMES				
1.1 applying make up	applying make-up (52)	combing hair (24), scratching the eye (3)	NA (21)	
1.2 archery	archery (68)	using a slingshot (19), using a rubber band (6), pulling (3), throwing something (3)		
1.3 blowing a whistle	blowing a whistle (47)	coughing (31), splitting (3), sneezing (3), asthma (3)	NA (13)	
	...			
2. MEANINGLESS GESTURES				
2.1 D_1 (peeling a banana)	peeling a banana upside down (6)	covering something (3), knotting (3), lifting two sides of a ribbon (3), winding up (3)	NA (81)	
2.2 D_2 (using binoculars)	carrying a burden (13)	indicate the size (3), explaining something (3)	NA (81)	
2.3 D_3 (pouring from a bottle)	tipping over/spilling (10)	turning over (3), coffee machine (3)	NA (84)	
	...			
3. EMBLEMS (Italian raters)				
3.1 agreement** was: congratulation	having an agreement (53)	greetings (13), congrats (7), being together (7)	NA (7)	
3.2 anger	indicate anger (47)	yawning (33)	NA (20)	
3.3 approving	approving (93)		NA (7)	
	...			
4. EMBLEMS (non-Italian raters)				
4.1 anger	yawning (50)	indicate nervousness (11), coughing (6), indicate a mistake (6), having bad news (6), indicate pain (6)	NA (17)	
4.2 approving	thumb up/approving (94)		NA (6)	
4.3 being late	asking/checking for the time (56)	inviting someone to be aware of the time (50)	NA (6)	*

...

2.5 Discussion

The aim of this study was to develop a standardized library of video clips of gestures to be used for scientific research. To the best of our knowledge, this is the first study providing a large set of normed transitive (pantomimes), intransitive (emblems), and meaningless gestures. The advantage of this library is that it provides high quality videos, free from distracting and/or conflicting information such as facial expressions, together with detailed norms derived from large samples of Italian and non-Italian participants. Moreover, the library takes into account the cultural differences that characterize the emblem category and, therefore, offers separate ratings by Italian native speakers and by non-Italian native speakers for this category only.

Our results showed that the meaning of pantomimes is largely shared by members of different cultures. For this reason, the final database pools together the data from Italian and non-Italian raters, yielding a unique set of stimuli. Such set includes ratings from the following nationalities: Italian (48%), American (46%), Indian (4%), Chinese (1%), and Russian (1%).

The susceptibility of the two classes of gestures to culture is a debated aspect that has been mostly investigated from an anthropological perspective (Efron, 1941; 1972, Ekman and Friesen, 1969; Ekman, 1976; Poggi and Zomparelli, 1987; Kendon, 1992; Payrató, 1993). It has been suggested that learning emblems typical of a culture and learning words are achieved in a similar fashion, by learning an association between a symbol and its shared meaning (Ekman and Friesen, 1969; Gullberg, 2006). Because of their strict relation with language, and consequently, with the social group or culture that speaks that language (Ekman and Friesen, 1969; Ekman, 2004; Gullberg, 2006), emblems have been suggested to be more susceptible to cultural influence than pantomimes. In particular, Ekman and Friesen (1972) observed that emblems that appear to be cross-cultural are usually those describing bodily actions such as *drinking*, *sleeping* and *eating*. These movements are, for anatomical reasons, performed in a similar way by all human beings and, therefore, their meaning is recognized by members of different social/cultural groups. Other kinds of messages (e.g., gestures that indicate "being crazy" or "being hungry") are still represented by gestures, but the movements to express them appear more susceptible to cultural variations.

Our results showed a higher variability between Italians and non-Italians when judging emblems, both in terms of the number of gestures judged as meaningful and in terms of the associated meaning., Furthermore, we observed higher agreement between Italian and non-Italian raters when judging emblems-event (that usually correspond to bodily actions), than when judging emblems-state. On the contrary, non-Italian raters tended to describe emblems-state, whose meaning cannot be inferred by the movement observed, in a different way in comparison to Italians, with a higher internal variability, or not to recognize them at all. This observation is in line with the above-mentioned study carried out by Ekman and Friesen (1972), and suggests a general cultural specificity of emblems that communicate a state and a larger agreement on emblems that refer to actions.

This dataset is now available to investigate different aspects linked to gestures such as the neural and cognitive mechanisms of gesture processing, cultural differences, age of acquisition, or linguistic related aspects. This dataset, together with the detailed norms we provide, aims to promote replicability by allowing direct comparison between different categories of stimuli.

Chapter 3

The role of temporal and precentral areas in action understanding: an fMRI-guided TMS study

3.1 Abstract

Understanding the meaning of actions performed by other people is fundamental for successful social interactions. The neural basis underlying this ability is still unclear. In particular, it is debated whether simulation of the observed action in the observer's motor system is a necessary part of action understanding. The aim of this study was two-fold. Firstly, in Experiment 1, we aimed at investigating and further characterizing the neural activation associated with the processing of meaningful actions, comparing the activity related to meaningful and meaningless gestures. Secondly, in Experiment 2, we contrasted the role of the left posterior middle temporal gyrus (pMTG), a region considered to be crucial for the semantic processing of action verbs, with the role of the left premotor cortex (PMv). Offline rTMS was applied either to the left pMTG or to the left PMv, previously localised based on the functional data of Experiment 1. Participants performed either a semantic judgment task (experimental condition) or a perceptual judgment task (control condition). According to motor theories of action understanding, rTMS applied to the PMv, but not to the pMTG, should impair performance during the semantic judgment task. By contrast, according to cognitive theories of action understanding, rTMS applied to pMTG, but not to PMv, should impair performance in this task. Experiment 1 revealed a significant difference between meaningful and meaningless gestures in the MTG suggesting

a role of this temporal area in action understanding. Furthermore, our data suggest a contribution of other brain areas, when processing emblems or pantomimes specifically (such as the inferior frontal gyrus (IFG) in the case of emblems and the superior parietal lobe (SPL) and the precentral gyrus (PCG) in the case of pantomimes). In Experiment 2, we failed to obtain significant effects of rTMS in any condition. Possible limitations of the study are discussed.

3.2 Introduction

Understanding other people's movements and behaviors is fundamental for social adaptation, communication, and interaction. Neuroimaging studies revealed a large network of temporal, parietal, and frontal regions that is recruited during action observation (Casper et al., 2010; Turella et al., 2013). However, it is unclear which of these areas plays a causal role in action understanding and which of them responds to action stimuli as an epiphenomenon.

According to motor theories of action understanding, actions are understood via a simulation of the observed action in the observer's motor system, thereby providing access to our motor knowledge about the action (Rizzolatti and Sinigaglia, 2010; Meteyard et al., 2012; Kiefer and Pulvermüller, 2012). By contrast, cognitive theories state that we understand an action by having access to action knowledge stored in conceptual areas located outside the motor system, and that activation of the motor system occurs as an epiphenomenon (Mahon and Caramazza, 2008; Hickok, 2009; Caramazza et al., 2014).

Villareal et al. (2008) found that bilateral superior temporal sulcus (STS) and left MTG were sensitive to the difference between the recognition of transitive and intransitive gestures in comparison to meaningless gestures. However, no specific activation for transitive compared to intransitive actions was found, except for the left IFG, which showed a stronger response during the recognition of intransitive actions. The authors concluded that the left MTG is specifically involved in the retrieval of semantic components of actions, and they speculated on the role of the IFG suggesting a role in linking actions and implicit information (i.e., the social and symbolic meaning). Other studies investigating gestures have found similar results (Lotze et al., 2006; Xu et al., 2009; Andric et al., 2013; Möttönen et al., 2016).

Recent studies have reported the recruitment of the left middle temporal gyrus (MTG), seemingly involved in the representation of the abstract features of actions (Oosterhof et al., 2010, 2012, 2013; Wurm et al., 2015; Wurm and Lingnau, 2015).

Activation in the left MTG has also been found during the processing of action verbs compared to other words (Kable et al., 2005; Bedny et al., 2008; Crepaldi et al., 2011; Willms et al., 2011; Peelen et al., 2012; Peelen et al., 2012). Papeo et al. (2014) examined whether the meaning of a word is accessed directly by motor regions, or whether the motor representation is driven by the previous retrieval of conceptual representations stored in lpMTG. In the first part of the study, participants had to perform a synonym judgment task on verbs and nouns in three different conditions: 1) after rTMS delivered over the left pMTG (lpMTG-rTMS); 2) after rTMS delivered over the occipital cortex (control region- OCC-rTMS); 3) without any stimulation (no-rTMS). Participants had to indicate as fast as possible whether the two words presented on a screen were synonyms or not. Results showed that only the perturbation of the lpMTG through rTMS disrupted semantic processing of action verbs. The performance on nouns was consistent across all the conditions (**Figure 3.1A**). In the second part of the study, the authors investigated the consequences of rTMS stimulation on verb-related motor activity. For this purpose, motor evoked potentials (MEPs) derived from the stimulation of the motor area were recorded following the three different conditions described above (lpMTG-rTMS; OCC-rTMS; no-rTMS). Participants were presented with pairs of action or non-action verbs and were asked to think about whether the two verbs were synonyms or not. Without rTMS, MEPs were stronger for action- in comparison to non-action verbs. This distinction was eliminated following the perturbation of the left pMTG (**Figure 3.1B**).

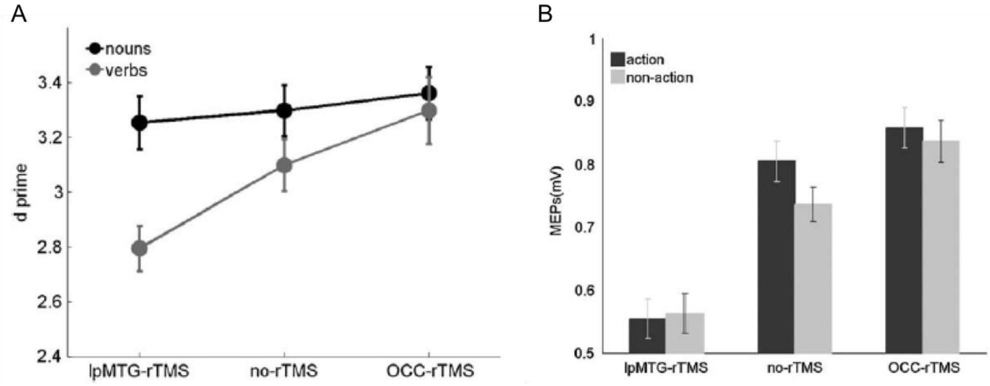


Figure 3.1. Results from Papeo et al., 2014. A) Subjects' performance on a semantic judgment task on verbs and nouns, after rTMS applied over the lpMTG, OCC and in a no-rTMS condition; B) MEPs amplitude for action and non-action verbs in the three experimental conditions.

These results support the idea that the motor activation that is observed when attending to action verbs is a consequence of the activation of the left pMTG, and that the precentral area *per se* is not sufficient for processing verbs. However, it has been argued that the left pMTG might play a specific role in verb processing (considered as a grammar category), and not in action understanding.

To address this criticism and to confirm a more general role of pMTG in understanding actions and not only in verb processing, we investigated the role of temporal and precentral areas in action understanding, using nonverbal material, in particular gestures and action line drawings.

In *Experiment 1*, we wanted to further characterize the network of areas associated with the observation of actions using fMRI. In particular, we aimed to examine which areas are sensitive to the difference between meaningful (pantomimes and emblems) and meaningless gestures. The comparison between meaningful and meaningless gestures is fundamental in order to identify areas that are specifically implicated in the processing of the conceptual/semantic information. In comparison to the study by Villareal et al. (2008) we decided to increase the number of unique gestures for each category and to avoid repetitions. Moreover, our data set was accurately rated by a large number of Italian-native speakers allowing a controlled selection of the stimuli and comparison between categories. Due to their different nature, we hypothesized that slightly different networks are recruited when observing pantomimes versus when observing emblems. Based on previous results reviewed above, we expected to observe a response in MTG for both conditions, in comparison to meaningless gestures. Furthermore, since emblems, in comparison to pantomimes, are symbolic gestures with a strong communicative component, we expected a stronger response in IFG during the observation of emblems in comparison to pantomimes, as shown in other studies (Lotze et al., 2006; Villareal et al., 2008). Motor and cognitive theories lead to different predictions regarding the activation of the precentral gyrus (PCG). If the PCG is sensitive to the semantic features of actions (as claimed by motor theories of action understanding), this area should show a difference between meaningful and meaningless gestures. By contrast, if PCG is sensitive to other low-level features of actions, rather than the their meaning, it should not show any difference in activation between meaningful and meaningless gestures.

Based on the results of Experiment 1, in *Experiment 2*, we aimed at directly contrasting the role of the lpMTG with the role of the premotor cortex while processing visual actions. To this aim, we applied offline rTMS over the action-preferring site in the left pMTG and the left PMv, previously localized. Participants performed both a semantic experimental task and a perceptual control task in three different conditions: after rTMS delivered over the left pMTG, after rTMS delivered on the left PMv, and in a no-rTMS condition (baseline). According to motor theories of action understanding, TMS applied to premotor cortex, but not to the pMTG, should impair the performance on the semantic task. On the contrary, according to cognitive theories of action understanding, TMS applied to pMTG but not to premotor cortex, should impair performance on the semantic task.

3.3 Experiment 1 – fMRI study

3.3.1 Materials and Methods

<u>Participants</u>

Seventeen native Italian-speaking participants (10 females; 7 males; mean age 24 years; age range 19-30 years) took part in the experiment. All participants were right-handed with normal or corrected-to-normal vision. None of them had a history of neurological or psychiatric diseases. Participants gave written informed consent before participation in the study.

<u>Stimuli</u>

The stimulus set consisted of 2.5 s long video clips of an actress performing silent gestures or pronouncing words in a still position (**Figure 3.2**). The actress was a female Italian-native speaker and the setting (background, illumination, clothing) was held constant across all videos.

Gestures belonged to 3 different categories: pantomimes (gestures that directly depict actions such as *playing violin*; **Figure 3.2.A**), emblems (gestures that express a symbolic goal-directed meaning, such as *listening*, *thumb up*; **Figure 3.2.B**), and meaningless gestures (**Figure 3.2.C**). Each category comprised 60 different gestures. The emblem category was further divided into two sub-categories: emblems-event, i.e., emblems whose meanings refer to an action (e.g., *to listen, to clap, to yawn*), and emblems-state, i.e., emblems whose meanings refer to a physical or psychological state (e.g., *thumb up, no, victory*). All gestures were selected from a larger standardized data set (Agostini et al., under review; see Chapter 2) and were rated by a group of 62 Italian-native participants.

Spoken word video clips (**Figure 3.2.E**) consisted in the same actress standing in front of the camera and pronouncing words. Two different categories of words were recorded: verbs and nouns. Each category comprised 60 different words: 30 verbs were action-related (e.g., *correre* – to run) and 30 were stative verbs (e.g., *pensare* – to think); 30 nouns were concrete (e.g., *casa* – house) and 30 were abstract (e.g., *talento* – talent). Verbs were presented in first-person singular form of the present tense (*io dipingo* – I paint), and nouns were presented in their singular form, preceded by the appropriate article (*la collina* –the hill). All words were matched for length (number of letters) and frequency (Dizionario di frequenza della lingua italiana, Consiglio Nazionale delle Ricerche, C.N.R.-I.L.C.). The distinction between emblems event and emblems state, as well as action-related and stative verbs and concrete and abstract nouns, was made to investigate the different levels of abstraction. For the full list of stimuli used, see Appendix 3.

Video-clips were back-projected inside the scanner onto a screen (frame rate; 60 Hz; screen resolution: 1024x768 pixels) via a liquid crystal projector (OC EMP 7900; Epson Nagano) and viewed through a mirror mounted on the head coil. Spoken words were presented via MR-compatible headphones (SereneSound Digital audio system). ASF (Schwarzbach, 2011) and the MATLAB Psychtoolbox-3 for Windows (Brainard, 1997) were used to control the stimulus presentation, response collection, and synchronisation with the scanner.

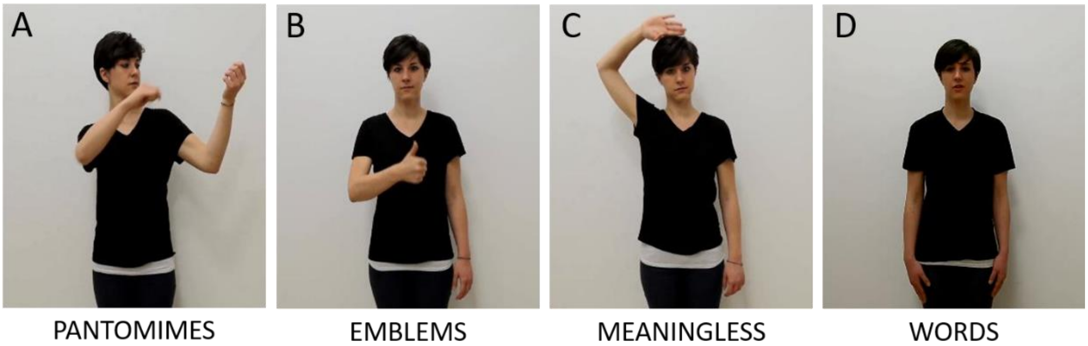


Figure 3.2. Still frame examples from video clips. A) Pantomime, a gesture that directly mimics actions or objects (in this example: *playing violin*). B) Emblems, symbolic communicative gesture (in this example: *thumb up*). D) Meaningless gesture. E) Words (spoken verbs or nouns).

Design and task

The experiment consisted of 6 runs of gestures, 2 runs of spoken words, and one motor localizer run (**Figure 3.3**). Each run consisted of 12 blocks and each block contained 5 video clips randomly selected from the same category. Each run started and ended with a 16.5 second fixation period. The duration of each run was approximately 6.9 minutes. Each block was separated by a 16.5 second fixation period. Videos within each block were separated by 1 second of fixation. During the 6 runs of gestures, participants were presented with blocks of pantomimes, emblems, or meaningless gestures. During the 2 runs of spoken words, participants were presented with blocks of verbs (action-related or stative) or nouns (concrete or abstract). Each run comprised 3 blocks of action-related verbs, 3 blocks of stative verbs, 3 blocks of concrete nouns, and 3 blocks of abstract nouns. The order of blocks was randomized across runs and across participants. Each video was presented only once. Participants were asked to watch the video clips and perform a *1-back* task, i.e., press a button with the left index finger whenever a gesture matched the previous gestures. To avoid a systematic association between the button press and one of the experimental condition, there was only one single repetition per run, randomly assigned to one of the conditions. The motor localizer run consisted of 15 blocks (16 s duration; 16.5 s fixation period between blocks), in which participants had to actively move either the right hand, right foot, or tongue, following written instructions presented on the screen.

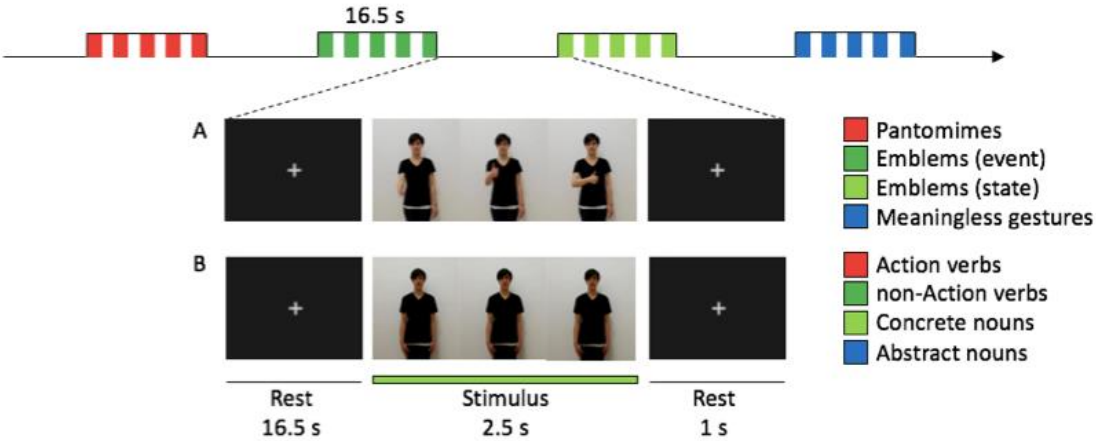


Figure 3.3. Experimental design. We used a block design with the conditions Pantomimes, Emblems-event, Emblems-state, and Meaningless gestures during gestures runs. Action-related and stative verbs and concrete and abstract nouns were used during the spoken word runs. A) Example of a trial from the gestures runs. B) Example of a trial from the spoken word runs.

Data acquisition

Functional and structural data were acquired using a 4T Bruker MedSpec Biospin MR scanner and an 8-channel birdcage head coil. Functional images were acquired with a T2*-weighted gradient EPI sequence. Acquisition parameters were the following: repetition time (RT) of 2000 ms; echo time (TE) of 33 ms; voxel resolution of 3x3x3; flip angle (FA) of 73°; field of view (FOV) of 192x192mm; gap size of 0.45 mm. Twenty-eight slices, acquired in ascending interleaved order were used. In each functional run, 207 images were acquired. Before each functional run, an additional scan was performed to measure the point-spread function (PSF) of the acquired sequence to correct the distortion expected with high-field imaging (Zaitsev et al., 2004). Structural T1-weighted anatomical scans were acquired with an MPRAGE sequence (176 sagittal slices; TR: 2700 ms; TE: 4.18 ms; voxel resolution 1x1x1 mm; FA: 7°; FOV: 256x224 mm; inversion time: 1020 ms) in order to be able to coregister low-resolution functional images to a high-resolution anatomical scan.

fMRI data analysis

Preprocessing

Data were preprocessed and analysed using BrainVoyager QX 2.8 (Brain Innovation) in combination with the BVQXtools/NeuroElf toolbox (by Jochen Weber: http://neuroelf.net/) and custom software written in MATLAB (The MathWorks). Distortion in geometry and intensity in the EPI images were corrected on the basis of the PSF data acquired before each functional run (Zeng and Constable, 2002). The first four volumes were removed to avoid T1 saturation. The first volume of the first functional run was aligned to the high-resolution anatomical scan using 6 rigid-body transformation parameters. Data were 3D motion correction (trilinear interpolation, using the first volume of the first run of each participant as reference), followed by slice time correction and high-pass filtering of 3 cycles per run. Spatial smoothing was applied with a Gaussian kernel of 5 mm FWHM. For group analyses, both functional and anatomical data were transformed into a common Talairach space, using trilinear interpolation.

Univariate analysis

To examine which areas are sensitive to differences between our main experimental conditions, we computed a group-level random-effects (RFX) GLM analysis. We used the following eight predictors: pantomimes, emblems-event, emblems-state, meaningless gestures, action-related verbs, stative verbs, concrete nouns, and abstract nouns. For the purpose of this study, we collapsed emblems-event and emblems-state (emblems), action-related and stative verbs (verbs) and concrete and abstract nouns (nouns). The final predictors were thus: pantomimes, emblems, meaningless gestures, verbs and nouns. Statistical maps were corrected using a false-discovery rate (FDR) of $q < 0.05$ (Genovese et al., 2002).

3.3.2 Results

Overall activity for observing gestures and words

Figure 3.4 shows the results of the RFX GLM contrast for all gestures (pantomimes + emblems + meaningless) against baseline and for all spoken words (verbs + nouns) against baseline ($q_{(FDR)} < 0.05$). We found that observing gestures recruits bilateral occipitotemporal cortex, the inferior temporal gyrus (ITG), the posterior middle temporal gyrus (pMTG), the middle part of the superior temporal cortex (pSTS), the superior and inferior parietal lobule (IPL) and a small cluster in the postcentral gyrus. Moreover, we found a recruitment of the bilateral middle frontal (MFG) and inferior frontal gyrus (IFG), with a larger extent in the left hemisphere (**Figure 3.4A**). During the processing of spoken words, the anterior portion of MTG, the superior temporal gyrus (STG), and the inferior and middle frontal gyrus were bilaterally recruited (**Figure 3.4B**).

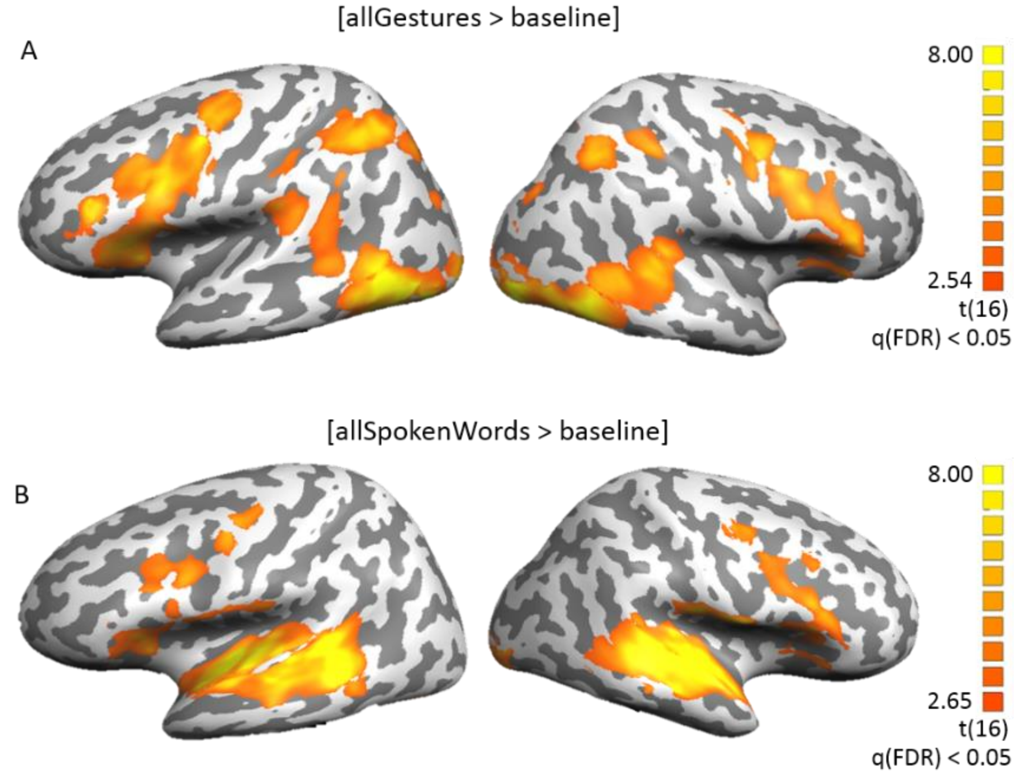


Figure 3.4. A) RFX-GLM [all Gestures (pantomimes + emblems + meaningless) > Baseline]; FDR <0.05; N=17; B) RFX-GLM [all Spoken Words (verbs + nouns) > Baseline]; $q_{(FDR)} < 0.05$; N=17.

<u>Contrast analysis</u>

Next, we sought to examine which areas showed stronger responses during the processing of pantomimes and emblems in comparison to meaningless gestures. The RFX-GLM contrast between meaningful (pantomimes + emblems) and meaningless gestures revealed bilateral recruitments of the posterior and anterior MTG (BA37/21), with a larger spatial extent in the left hemisphere (**Figure 3.5A**). Activations of the dorsal and ventral inferior frontal gyrus - IFG (BA44/45/47) and middle frontal gyrus (BA46) were found in the left hemisphere only. The comparison between pantomimes and meaningless gestures ($q_{(FDR)} < 0.05$) revealed the posterior MTG (BA37/posterior part of BA21) and the anterior STS (BA22) bilaterally (**Figure 3.5B**). Moreover, this contrast revealed a small cluster in the motor/premotor area (BA6) in the left hemisphere. Comparing emblems and meaningless gestures revealed bilateral activation of the middle and anterior MTG (BA21) and STS (BA22) as well as the left IFG. The latter consisted of a cluster including the pars triangularis (IFGTr – BA45) and pars opercularis (IFGOp – BA44, **Figure 3.5C**). Interestingly, precentral regions (BA4/BA6), partially overlapping with the activation revealed by the independent movement localizer task ([hand > foot], **Figure 3.5** in pink), responded only to pantomimes and meaningless gestures but not emblems. The activity in the motor/premotor area therefore seems to be specific for pantomimes and meaningless gestures but not for emblems.

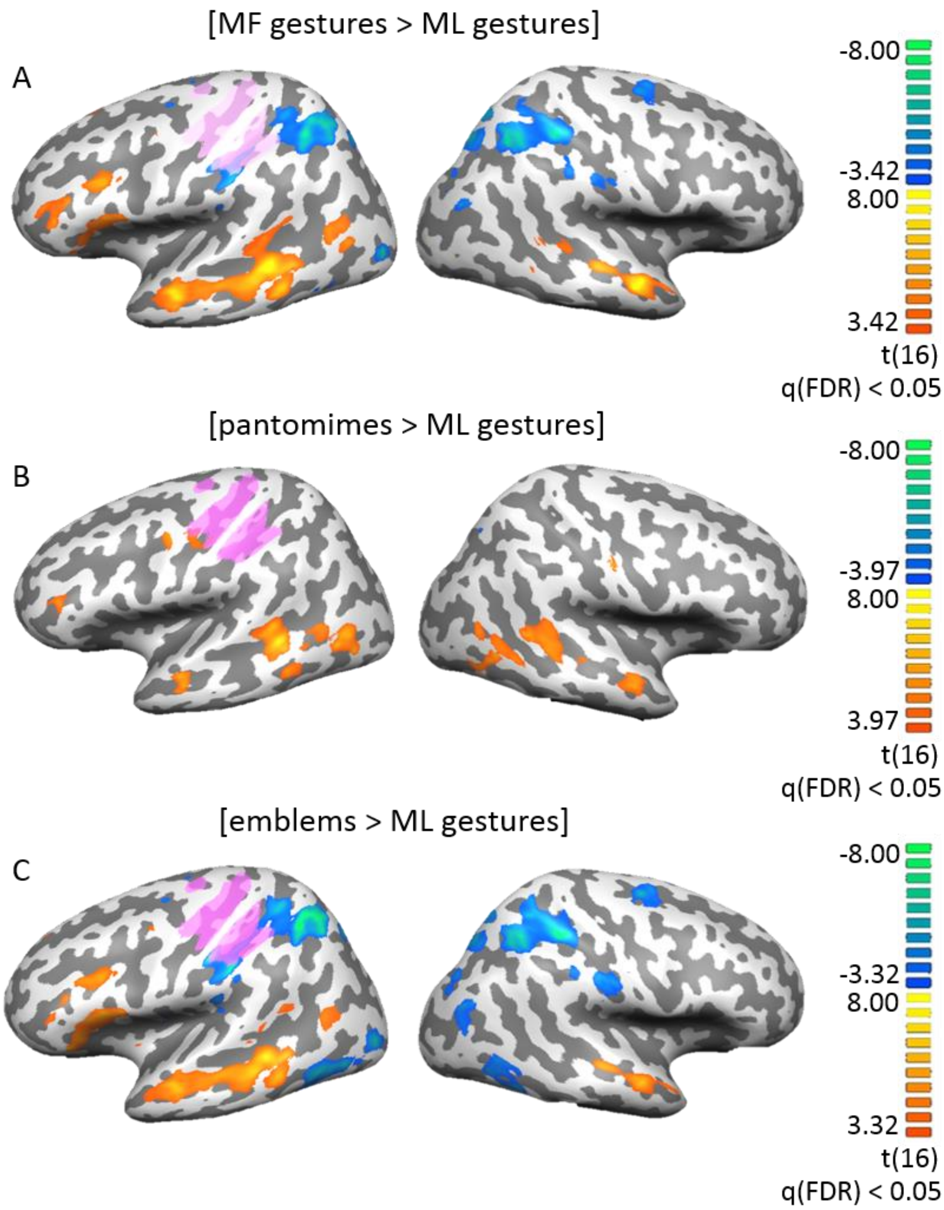


Figure 3.5. A) RFX-GLM [meaningful gestures (pantomimes + emblems; MF) > meaningless gestures (ML)]; $q_{(FDR)} < 0.05$; N=17; B) RFX-GLM [pantomimes > meaningless gestures]; $q_{(FDR)} < 0.05$; N=17; C) RFX-GLM [emblems > meaningless gestures]; $q_{(FDR)} < 0.05$; N=17. In pink: motor localizer [hand > foot].

As an exploratory analysis, we compared verbs and nouns in the spoken word condition in order to see which clusters survive when processing actions presented as a linguistic input. **Figure 3.6A** shows the RFX-GLM contrast between verbs (action-related + stative) and nouns (concrete + abstract). Since no clusters survived FDR correction, we chose a more liberal threshold (uncorrected $p < 0.008$) for further exploration. This contrast revealed activations in the left hemisphere only. In particular, we observed two small clusters in the middle portion of MTG (BA21), STG (BA22), and IFGTr (BA45). These areas are similar to the clusters surviving when performing a conjunction analysis between meaningful gestures *versus* meaningless gestures and verbs *versus* nouns (i.e., conjunction between all the categories that entail the presence of meaningful actions, regardless of the modality of presentation – **Figure 3.6B**). Again, no activity in motor and premotor regions was observed.

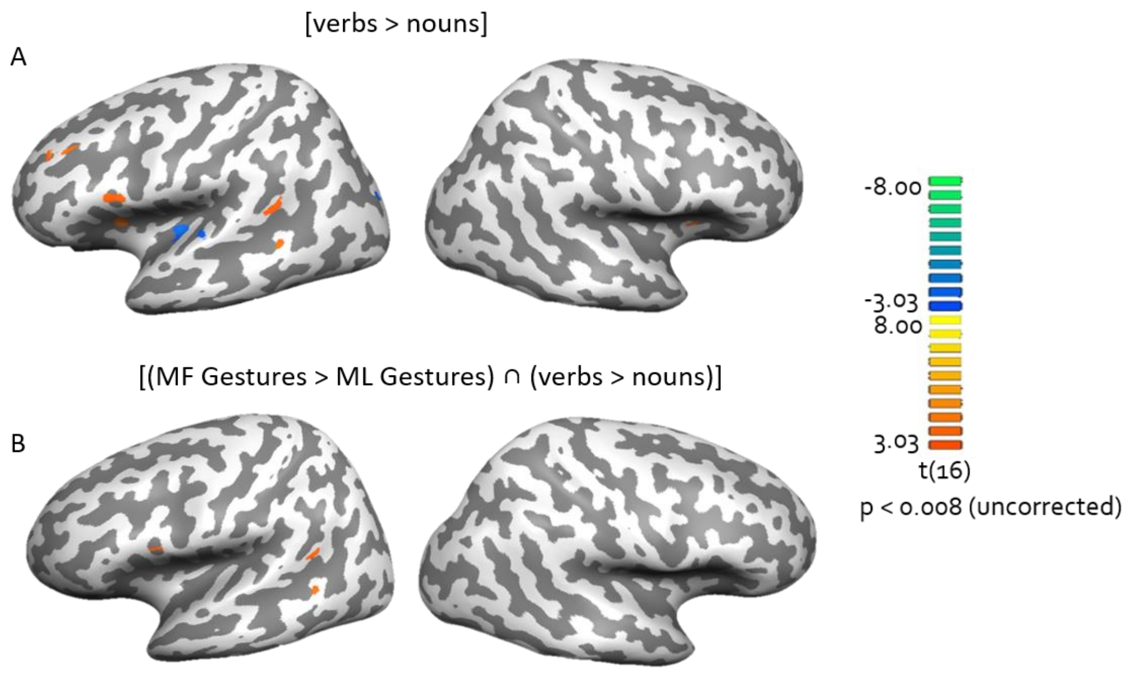


Figure 3.6. A) RFX-GLM [Verbs > Nouns]; *p*-uncorrected < 0.008; N=17; B) Conjunction analysis [(Meaningful (MF) gestures > Meaningless (ML) gestures) ∩ (Verbs > Nouns); *p*-uncorrected < 0.008; N=17.

3.3.3 Discussion

The aim of Experiment 1 was to investigate the neural responses to meaningful compared to meaningless actions, in order to establish which areas in the brain are sensitive to the semantic component of the action. Our results show that the left MTG is sensitive to the difference between meaningful and meaningless gestures. In particular, pantomimes in comparison to meaningless gestures were associated with the activation of the posterior part of MTG (in particular BA37), while emblems in comparison to meaningless gestures are associated with the activation of a more anterior part of MTG (BA21). We found a similar pattern, though on a slightly smaller scale, in the right MTG.

Despite this common MTG activation, two additional brain areas were specifically recruited when processing pantomimes or emblems in comparison to meaningless gestures: the left precentral gyrus (PCG) and the left IFG respectively.

PCG is known to be involved in the planning, control and execution of voluntarily movements. Several studies have observed activity in precentral areas during the performance of tasks which either involve the observation of motor actions (Fadiga et al., 1995) or the reading of action-related sentences (Olivieri et al., 2004; Glenberg et al., 2008). Authors suggested a role of the PCG in understanding the meaning of actions. We did find a very small cluster in PCG being sensitive to pantomimes. However, PCG also responded to meaningless gestures when compared to meaningful. We concluded that PCG is not sensible to the categorical distinction between meaningful and meaningless gestures. This result argues against the motor theories of action understanding. Moreover, we found involvement of motor regions when observing pantomimes and meaningless gestures but not when observing emblems. This result may be explained by considering the aspects that diversify pantomimes and emblems. Pantomimes and emblems are often referred to as transitive and intransitive actions, respectively (Villareal et al., 2008; Papeo and Rumiati, 2013). Pantomimes are, in fact, simulation of real actions that involve the use of a tool (e.g., *hammering*). Emblems are, instead, symbolic gestures that are not object-directed (e.g., *waving goodbye*). Even if the two kinds of gestures are perceptually very similar, they differ in terms of abstractness between the movement observed (the gesture) and the meaning to which it refers. When performing a pantomime, we faithfully reproduce the same movements that we would perform if we were to execute the action. Therefore there is a concrete link between the gesture and the action to which it refers. In the case of emblems, a conventional gesture is used to express something. This gesture does not reproduce an action but is symbolically linked to its meaning. Therefore, the link between the gesture and the meaning is more abstract.

When comparing emblems to meaningless gestures, we found activation in the left IFGTr, IFGOp and IFGOr. Xu et al. (2009) reported the recruitment of the same portions of IFG when participants were presented with emblems (dorsal portion of IFG) or speech (ventral portion of IFG). Other studies reported activation of this area during the processing of words (Peelen et al., 2012; Bedny et al., 2008; Xu et al., 2009), emblems (Villareal et al., 2008; Xu et al., 2009; Möttönen et al., 2016), pantomimes (Lewis et al., 2006; Emmorey et al., 2010), and manipulable objects (Kable et al., 2005; Lewis et al., 2006). We did find activation in the left IFG when processing emblems, but not during the processing of pantomimes. Our results suggest an involvement of the left IFG in the processing of linguistic aspects of the stimulus. In fact, emblems and speech, in comparison to pantomimes, share properties that are typical of the linguistic system because of their arbitrary relation between the form (gesture or word) and its meaning and because of the presence of language-like aspects (McNeill, 1992, 2005; Papeo and Rumiati, 2013). One question could be whether IFG is sensitive to the morphosyntactic aspects of the stimulus (e.g., the grammar class or other linguistic aspects), or whether it is involved in the communicative and semiotic aspects of it (i.e., the relation between symbol and meaning). A recent related study investigated whether the communicative intention is reflected in the IFG response (Möttönen et al., 2016). Hearing, non-signing participants were presented with sign language in three conditions: a) when the communicative intention was unknown (pre-training session), b) when the communicative intention was known but the meaning unknown, and c) when the communicative intention and the meaning were known (post-training sessions). The authors, comparing pre *versus* post training sessions, found that IFG was active only when participants knew that the gesture they were observing was communicative, irrespective of the fact of knowing meaning. The authors thus concluded that this area is sensitive to the communicative aspects of an action (see also Dick et al., 2009).

In conclusion, the robust activation found in MTG both for pantomimes and emblems in comparison to meaningless gestures seems to confirm a role of this area in representing the meaning of actions. Furthermore, our data suggest a contribution of other brain areas, when processing emblems or pantomimes specifically (such as IFG in the case of emblems and SPL and PCG in the case of pantomimes). Moreover, when comparing verbs (words that refer to actions) *versus* nouns, we observed activation of the posterior part of the left MTG (around position y = -50). For the latter comparison, we had to use an uncorrected p-value, due to the fact that in our design we had only 2 runs dedicated to the spoken words and, therefore, these data are less powerful. However, we are confident in accepting this result, as previous experiments on verbs have found activity in exactly the same location (Papeo et al., 2014; Peelen et al., 2012).

3.4 Experiment 2 – fMRI-guided TMS study

3.4.1 Materials and Methods

Participants

Twelve right-handed native Italian speakers (6 females; 6 males; mean age 24 years; age range 19-30 years) gave their consent for participating in the study. All of them took part in the previous fMRI experiment (see *Experiment 1*). Data from this fMRI session were used to localize the individual action-preferring site to be later stimulated with TMS. None of the participants presented any counter-indications to TMS (Rossi et al., 2009). The experimental procedures were approved by the Ethics Committee for research involving human subjects at the University of Trento, Italy.

fMRI data analysis: identification of the individual left pMTG and left PMv sites

To identify action-preferring clusters in left pMTG and left PMv, we computed a general linear model-based (GLM) analysis. Both a first level (single subject) and a second level random-effect (group) analysis were performed, using the following four predictors: pantomimes, emblems-event, emblems-state, and meaningless gestures. We initially identified clusters in the left pMTG and left PMv comparing pantomimes and meaningless gestures at the group level. Focusing on the left pMTG, this comparison revealed two separated clusters located at the same *y*-position (p-uncorrected < 0.001, **Figure 3.7A** top). Based on the results of a previous study carried out by Wurm and Lingnau (2015), we identified the most posterior active cluster in LOTC as a target region. Similarly, for the left PMv, we identified the action-preferring site comparing pantomimes and meaningless gestures at the group level (p-uncorrected < 0.01, **Figure 3.7A** bottom). The procedure to identify the individual sites in the left pMTG and PMv were the following: 1) we created a 12 mm sphere around the center of mass of the two clusters (pMTG - Talairach coordinates: -50/-67/6; PMv - Talairach coordinates: -46/-11/45, **Figure 3.7A**); 2) we performed a first level analysis, comparing pantomimes *versus* meaningless gestures for each participant and projected the group-sphere previously created on the individual surface (**Figure 3.7B**); 3) We then defined the ROI as all the voxels that fell into that sphere, starting from a threshold of $p < 0.001$ (uncorrected) until the first active voxels appeared (**Figure 3.7C**). Within each individual ROI we determined the peak coordinates (**Table 3.1**).

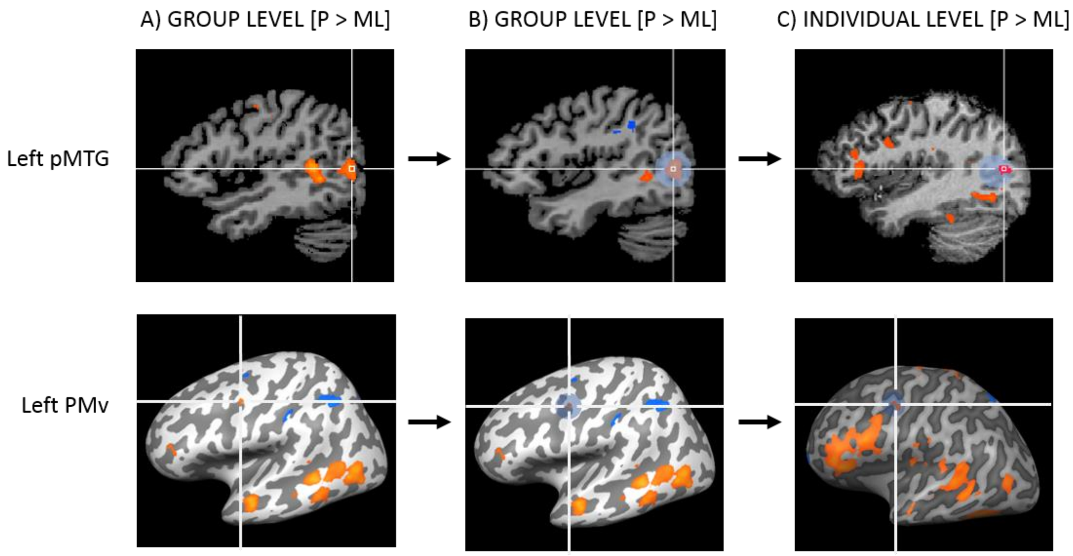


Figure 3.7. Procedure used to identify the individual site of stimulation in the left pMTG and left PMv. A) The region of interest was identified comparing pantomimes vs meaningless gestures at the group level; B) a 12-mm sphere was created around the center of gravity of the ROI; C) the sphere was projected on each individual brain and the first level comparison [pantomimes vs meaningless gestures] was computed; in pink, the individual active cluster that felt into the sphere.

For those participants where it was not possible to identify the left pMTG or PMv based on the functional data (i.e., the contrast meaningful gestures *versus* meaningless gestures did not reveal any cluster inside the sphere), the target site for stimulation was defined as the group-average Talairach coordinates. This was the case in 7 participants, only for the PMv site, and it is indicated by an asterisk in Table 3. All the Talairach coordinates were then converted into MNI coordinate to be used in the TMS Neuronavigation software (BrainSight 2.2.14).

Table 3.1. Individual Talairach coordinates in the left pMTG and PMv.

Individual coordinates: left pMTG					Individual coordinates: left PMv				
NrOfSubject	x	y	z	*P*	NrOfSubject	x	y	z	*P*
SUB01	-38	-70	9	0.005	SUB01	-39	-4	41	0.04
SUB02	-46	-71	2	0.006	SUB02 *	-46	-11	41	
SUB03	-46	-65	8	0.01	SUB03	-47	-8	45	0.04
SUB04	-45	-68	6	0.001	SUB04 *	-46	-11	41	
SUB05	-34	-65	4	0.005	SUB05 *	-46	-11	41	
SUB06	-40	-68	10	0.04	SUB06 *	-46	-11	41	
SUB07	-42	-68	9	0.001	SUB07	-46	-9	39	0.01
SUB08	-46	-74	9	0.001	SUB08 *	-46	-11	41	
SUB09	-53	-55	2	0.01	SUB09 *	-46	-11	41	
SUB10	-46	-71	0	0.005	SUB10	-46	-16	43	0.06
SUB11	-40	-68	9	0.04	SUB11	-47	-8	45	0.04
SUB12	-48	-68	2	0.001	SUB12	-51	-11	36	0.08

Peak Talairach coordinates of the participants in the fMRI localizer session. All *p* values are uncorrected.
* indicates when the mean group coordinates were used.

TMS study: stimuli and design

The stimuli were triads of black and white line drawings representing actions. The line drawings were selected from a standardized data set available on-line (http://crl.ucsd.edu/~aszekely/ipnp/, Szekely et al., 2004). Each triad consisted of a *target image*, presented in the upper part of the screen, and two images of choice (*alternatives*), presented in the lower part of the screen, one left and one right (see **Figure 3.8**). The experiment consisted of one experimental task and one control task. During the experimental task, participants were presented with three different pictures at a time and were instructed to indicate which of the two alternatives presented in the lower part of the screen was semantically related with the target image (semantic judgment task – **Figure 3.8A**). In order to improve the variability of the stimuli we also used horizontally flipped versions of them. In the control task, the three images were identical except for one of the two alternatives that differed by a small perceptual detail (e.g., in **Figure 3.8B** the colors on the palette are missing. Perceptual judgment task –). Participants were instructed to indicate which of the two alternatives was identical to the target image. Participants responded by pressing the right or left arrow on the keyboard with their left index and middle finger. Suitable triads were determined in a prior pilot study where a group of 14 participants performed both the semantic and the perceptual task on a larger set of triads. Only those triads that reached at least 85% of accuracy were included in the final set. None of the participants that took part in the pilot study participated in the TMS study. The two tasks were presented in two different blocks. Each block consisted of 86 unique triads; each triad was presented for 1000 ms and was preceded by a fixation cross that was presented for 500 ms. When the triad disappeared, a white screen appeared for a maximum of 4000 ms (**Figure 3.8C**). Participants were instructed to answer as fast and as accurate as possible. After pressing the button, a new trial began. Each block lasted around 5 minutes, and the entire experiment lasted no more than 10 minutes. The order of presentation of the two blocks, as well as the position on the screen (bottom left or bottom right) of the correct answer were randomized across sessions. Reaction times and accuracy were recorded. The trial duration was based on the same pilot experiments that were used to determine which triplet to use and that were run prior to the beginning of the study with a different group of participants.

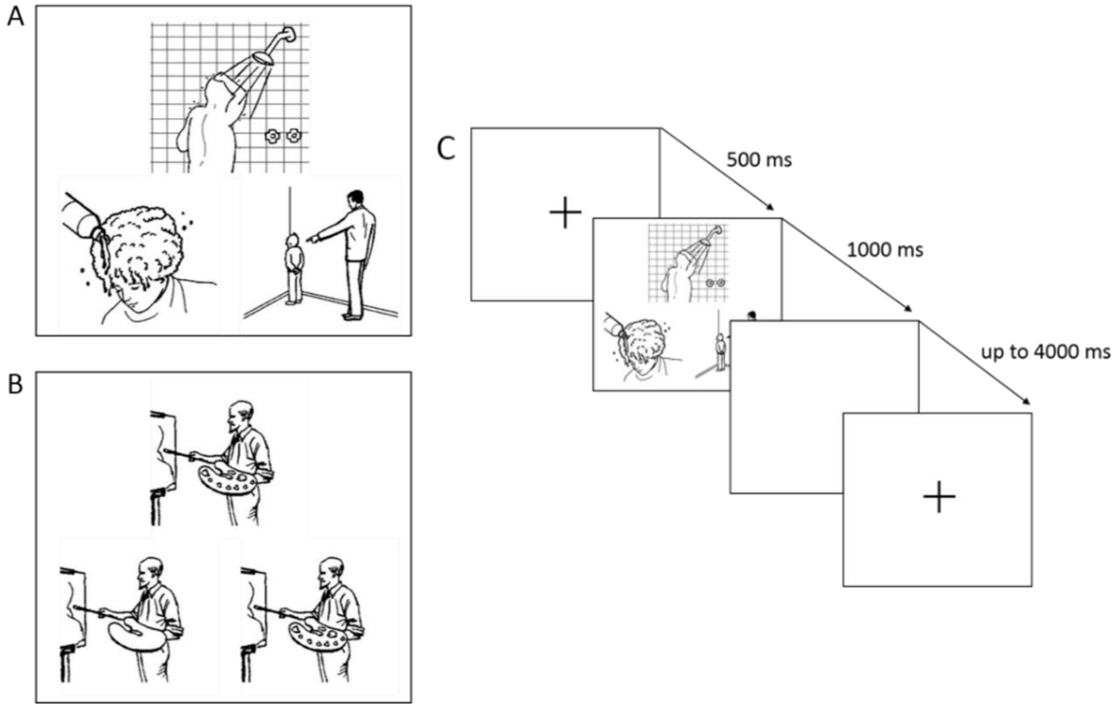


Figure 3.8. A) Example trial in the experimental task: participants were to indicate which of the two alternatives on the bottom is semantically related with the target image on the top. B) Example trial in the control task: participants were to indicate which of the two alternatives on the bottom is identical to the target image on the top. C) Timeline of an example trial: each experiment consisted of 86 trials.

Procedure

To test the effect of the perturbation of left pMTG and PMv on the semantic analysis of action images, participants performed both a semantic judgment task (experimental task) and a perceptual judgment task (control task) in three different conditions: after rTMS delivery over the left pMTG (pMTG-rTMS), after rTMS delivery on the left PMv (PMv-rTMS), and in a no-rTMS condition (baseline). This condition was necessary in order to interpret and sustain potential results. In fact, rTMS delivered to two sites that are located in different cortical positions, might have different and unpredictable effects due to the different scalp properties. Therefore, the interpretation of any effect seen would be difficult if we were to compare the effects of the rTMG at the two different sites only (Sack and Linden, 2003). The two rTMS conditions were separated by at least 24 hours, while the baseline condition was always performed at the beginning of one of the two sessions (i.e., before applying rTMS to one of the two sites – **Figure 3.9**). The order of the two conditions (left pMTG-rTMS and left PMv-rTMS) and the administering of the baseline were counterbalanced across participants. We adopted an *offline* approach where rTMS was delivered during rest, before administering the task.

Upon arrival, participants were provided with the necessary information about the TMS procedure and gave written informed consent. They were then presented with the instructions of the experiment and were asked to perform a brief familiarisation task consisting of 5 example trials. Each session started with the delivery of the rTMS on the appropriate site (offline protocol). Participants were sitting on a chair, wearing earplugs and were asked to lay the right part of their head on a pillow that was situated on the table. We asked them to find a comfortable position and to relax. rTMS stimulation was delivered for 20 minutes at a low-frequency (1 Hz). An infrared devise (Polaris, Northern Digital, Ontario, Canada) allowed us to co-register the participant's head with the anatomical MR image with frameless stereotaxy using the BrainSight system (BrainSight 2.2.14). Thanks to this co-registration system, we were able to position the TMS coil at the exact scalp position corresponding to the individual target site. Throughout the stimulation period, we were able to track the participant's head to make sure that the coil position stayed in the right position. After receiving rTMS stimulation, participants immediately started the experiment. It is assumed that aftereffects of a 1Hz protocol last as long as the duration of the stimulation (Robertson et al., 2003; Rossi et al., 2009). Therefore, delivering rTMS for 20 minutes should lead to a refractory period that last approximately 20 minutes beyond the stimulation. During this period, participants performed the two tasks.

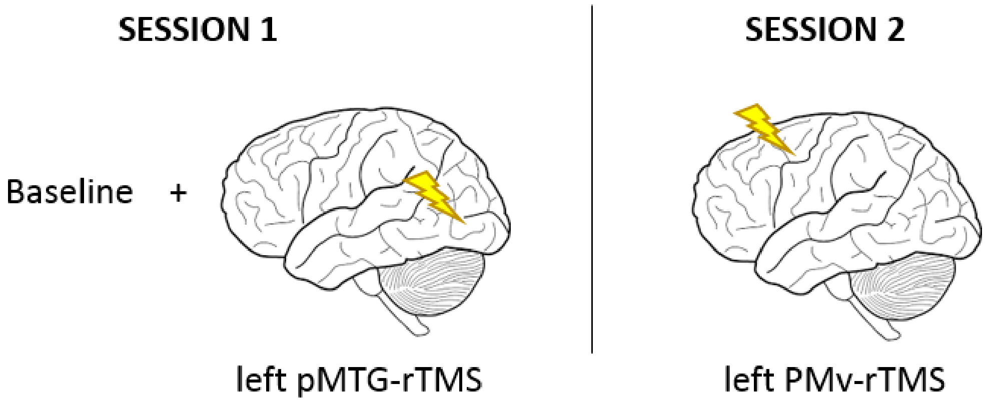


Figure 3.9. Example of one of the possible combination of experimental procedure. Participants' performance in a semantic judgment task (experimental condition) and in a perceptual judgment task (control condition) was measured after rTMS applied over the left pMTG and after rTMS applied over the left PMv. A no-rTMS condition (baseline) was always run before one of the two sessions. The order of the conditions, as well as the presentation of the baseline was randomized across participants.

<u>TMS protocol</u>

The stimulation was delivered through a figure-8 coil that was connected to a Magstim Rapid 2 stimulator. The coil was positioned on the scalp using individual ROI coordinates as determined in the fMRI session and with the neuronavigation system. The coil was kept in place using a mechanical arm. The experimenter monitored the correct position of the coil throughout the stimulation period.

pMTG stimulation. The left pMTG was stimulated with an intensity corresponding to 65% of the maximum stimulator output.

PMv stimulation. The intensity of stimulation of the left PMv was individually defined based on the individual motor threshold. In brief, we identified the participant's cortical motor threshold by stimulating the cortical motor area of the left hemisphere, and identifying the minimum intensity to evoke a visible twitch in the hand in at least 3 of 5 consecutive pulses (Rossini et al., 2015). The final intensity was adjusted to 120% of the individual resting motor threshold. At this intensity, rTMS over PMv did not induce any visible twitch in the contralateral muscles. The mean intensity of stimulation between subjects corresponded to the 73% of the maximum stimulator output (range: 56-88%).

Data analysis

We measured reaction time and accuracy and omitted the fastest and slowest 5% of trials in order to exclude possible outliers and only considered correct answers. Accuracy was converted to sensitivity and expressed as *d-prime* (Macmillan and Creelman, 2005). Correct responses were considered "hits" and incorrect responses as "false alarms". A 3 x 2 repeated-measure ANOVA was performed for both reaction times and sensitivity. The factors were TMS conditions (baseline, left pMTG-rTMG and left PMv-rTMS) and task (semantic, perceptual).

3.4.2 Results

Figure 3.10 shows the trimmed mean and sensitivity for both the semantic and the perceptual task for the 3 conditions: baseline, left pMTG-rTMS and left PMv-rTMS. The 3 x 2 ANOVA (3 conditions x 2 tasks) on reaction times revealed a significant effect of the task [$F_{1,66} = 9.97$, $p = 0.002$] with an overall faster performance in the perceptual judgment task compared to the semantic judgment task. No other significant effects were found in the reaction times analysis (effect of the condition [$F_{2,66} = 0.95$, $p = 0.39$]; interaction [$F_{2,66} = 0.02$, $p = 0.97$]). The *d-prime* analysis did not show any significant main effects or interactions (effect of the task [$F_{1,66} = 3.69$, $p = 0.06$]; effect of the condition [$F_{2,66} = 0.24$, $p = 0.79$]; interaction [$F_{2,66} = 0.98$, $p = 0.38$]), and the overall accuracy for all conditions and all tasks was very high (> 90%). Summarizing, no significant differences were found between TMS conditions.

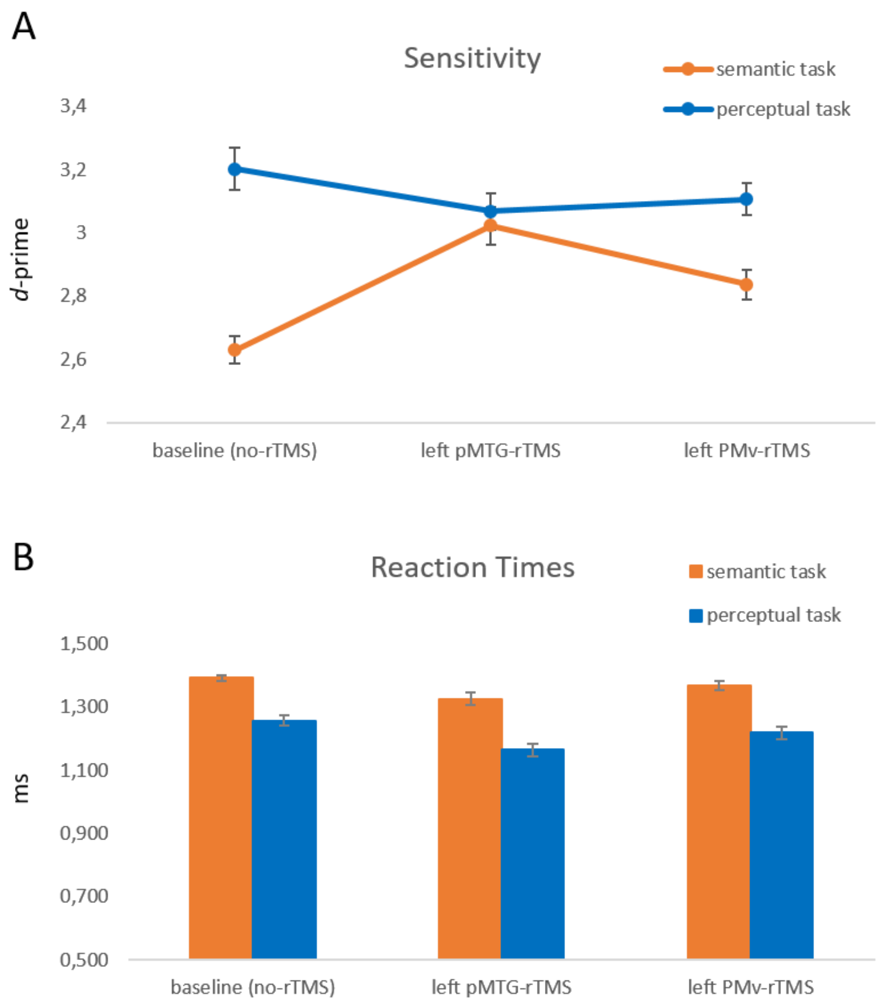


Figure 3.10. Behavioural results of the fMRI-guided TMS experiment. A) Performance, expressed as *d-prime*, in the semantic and perceptual task, in the three conditions (baseline, left pMTG-rTMS and left PMv-rTMS) for both the semantic and perceptual task. B) Mean reaction times in ms for the three conditions.

3.4.3 Discussion

In order to investigate the role of temporal and precentral areas in action understanding, we stimulated the individual site in the left pMTG and left PMv in two different sessions and asked participants to perform both a semantic and a perceptual task on visual actions. Based on previous studies, our prediction was that rTMS applied to the left pMTG but not to the left PMv should have led to an impoverished performance in the semantic task. The results reported above did not show any effect of the perturbation of the left pMTG in comparison to the left PMv during the semantic processing of action images.

What could have yielded this null result? First, we consider it likely that the presence of the baseline condition just before one of the two rTMS sessions induced unspecific learning effects. Even though we tried to prevent such learning effects by using a relatively high number of triads made of different combinations of images (every triad was repeated only once), and by introducing some variability in the stimuli (both in terms of position of the correct answer and the use of flipped images), the triads presented were exactly the same in the three conditions. In particular, during one of the two sessions, participants performed the task twice: first as a baseline and then again, around 40 minutes later, after receiving the rTMS stimulation. It is possible that the effect of rTMS was washed out by this learning effect. Another potential limitation of the study might be a ceiling effect: even though we conducted several pilot experiments before to determine the best combination of timing and error rate, the task was performed better than expected with accuracy rates higher than 90%. A more complex task or stimuli are undoubtedly needed in order to observe effects when stimulating an area with such a low intensity as the one we used. Alternatively, it might be advantageous to use a more powerful TMS protocol (e.g., theta burst stimulation). The last concern is related to the stimulation site chosen in the left pMTG. As described above, the GLM comparison between meaningful and meaningless gestures revealed two different clusters in the left pMTG (see **Figure 3.7A** top). We decided to consider the posterior cluster as a ROI because this area was found to encode abstract aspects of actions in a recent study by Wurm and Lingnau (2015). However, there is also evidence of the involvement of a slightly more anterior part of the pMTG in the representation of action knowledge (Xu et al., 2009; Papeo et al., 2014). The reason why we decided not to stimulate this anterior area is that previous experiments investigated actions using linguistic material, and we were concerned that this area might have been too language-related, whereas our stimuli were visual.

In conclusion, there are several possible reasons why this experiment led to non-significant results and future experiments should aim to create a stronger experimental design or to use a more powerful TMS protocol, in addition to investigating slightly different clusters along the left posterior MTG.

Chapter 4

The organization of gestures and words in the middle temporal gyrus

Adapted from:

Agostini, B., Papeo, L., & Lingnau, A. (2016). *The organization of gestures and words in the middle temporal gyrus.* Manuscript in preparation.

4.1 Abstract

The middle temporal gyrus (MTG) has been shown to be recruited during the processing of words. Likewise, this region is known to be recruited during action observation. Here we investigated how information related to words and gestures is organized in the MTG. To this aim, we measured the blood-oxygen level dependent (BOLD) response in MTG to video clips of gestures and spoken words in 17 participants. Gestures consisted of object-use pantomimes (iconic representations of object-directed actions; e.g., *playing guitar*), emblems (symbolic goal-directed gestures, e.g., *thumb up*), and meaningless gestures. Word stimuli (verbs, nouns) were presented in a different set of video clips, spoken by the same actress who performed the gestures. We found sensitivity to the difference between meaningful and meaningless gestures along the whole left and right MTG. Importantly, we observed a gradient, with posterior regions responding more strongly to gestures (pantomimes and emblems) than words, and anterior regions showing a stronger response to words than gestures. In an intermediate region in the left hemisphere, there was a significantly higher

response to words and emblems (i.e., items with a greater arbitrariness of the sign-to-meaning mapping) than to pantomimes. These results suggest that the organization of information in the MTG is driven by the input modality and also reflects the arbitrariness of the relationship between sign and meaning.

4.2 Introduction

Words and gestures play an important role in our daily interactions. In most situations, we produce words to communicate. Moreover, emblems such as *thumbs up* can be used to mean agreement or approval of something/someone. Finally, one can pantomime the action of drinking from a glass to invite someone else to have a drink. To understand the meaning of words and gestures, we need to map the different types of input onto the corresponding semantic representation. How are these different types of communication organized in the brain?

The middle temporal gyrus (MTG) is known to play a crucial role in the processing of semantics. As an example, it has been shown to be involved during the processing of words (Kable et al., 2005; Papeo and Lingnau, 2015; Peelen et al., 2012) and to be sensitive to the difference between meaningful and meaningless gestures (Villareal et al., 2008). Moreover, it has been shown that symbolic communicative gestures and spoken words are processed by the same common network of posterior temporal regions in the left hemisphere (Xu et al., 2009; Andric et al., 2013).

Pantomimes, emblems, and words differ with respect to the degree of abstractness/arbitrariness of the relationship between the sign (word or gesture) and its meaning. Pantomimes refer to gestures that faithfully reproduce object-related actions (e.g., *drinking*, *playing the guitar*). By contrast, emblems, similarly to words, are characterized by a more arbitrary relationship between the sign/gesture (e.g., *thumbs up*, *waving your hand*) and its meaning (agreeing, greeting). In the case of emblems that symbolically refer to objects (e.g., "inviting someone to cut short" represented by the index and middle finger moving repeatedly toward each other as to indicate the symbol of the scissor), the action no longer faithfully reproduces the real action, but rather uses an action in which a body part becomes the object (e.g., the hand becomes the scissor). In the case of spoken words, the relationship between sign and meaning is even more abstract and arbitrary.

Here we asked whether input modality and the relationship between sign and meaning contribute to the organization of the MTG. To this aim, we examined the blood-oxygen-level dependent (BOLD) response during the processing of pantomimes, emblems, and speech as a function of the position on the posterior-to-anterior axis. We found a posterior-to-anterior gradient, with posterior regions responding more strongly to gestures and anterior regions responding more strongly to words. Moreover, an intermediate region along the strip, only in the left hemisphere, showed stronger responses to words and emblems in comparison to pantomimes.

4.3 Materials and Methods

<u>Participants</u>

Seventeen native Italian-speaking participants (10 females; 7 males; mean age 24 years; age range 19-30 years) volunteered in the experiment. All participants were right-handed with normal or corrected-to-normal vision and no history of neurological or psychiatric diseases. Participants gave written informed consent before participation in the study.

Procedure

The stimulus set consisted of 2.5 s long video clips showing an actress performing gestures silently (pantomimes, emblems, or meaningless gestures) or speaking words (verbs or nouns). The experiment consisted of 6 runs of gestures and 2 runs of spoken word. Each run was made of 12 blocks. Each block consisted of 5 videos taken from the same category. Participants had to observe the videos and perform a *one-back* task, i.e., press a button every time a video was repeating the previous one (for details on the stimuli, design, data acquisition, and preprocessing see Chapter 3.3.1).

fMRI data analysis

Univariate analysis

To assess the sensitivity of the MTG to our stimuli, we contrasted each main condition (pantomimes, emblems, meaningless gestures, and words) against baseline using a random-effects (RFX) GLM analysis. The following eight predictors were used: pantomimes, emblems-event, emblems-state, meaningless gestures, action-related verbs, stative verbs, concrete nouns, and abstract nouns. While contrasting the conditions, we collapsed across emblems-event and emblems-state for the condition emblems, across action-related verbs, stative verbs, concrete nouns and abstract nouns for the word stimuli. Statistical maps were corrected using a false-discovery rate (FDR) $q < 0.05$ (Genovese et al., 2002).

Posterior-to-anterior organization

To investigate the organization of information in the MTG, we assessed how the amplitude of the blood-oxygen-level dependent (BOLD) response varied as a function of the stimulus category, for each of the *y*-coordinates along the MTG. To this end, we identified the left and right MTG on a standardized anatomical brain space (Talairach-transformed Colin template) based on the following anatomical landmarks: the superior and inferior borders corresponded to the superior temporal sulcus (STS) and the inferior temporal sulcus (ITS), respectively (**Figure 4.1 top**); the posterior end was delineated by the preoccipital notch and extended anteriorly along the sagittal plane for the entire length of the temporal lobe. The entire ROI encompassed 17.523 voxels.

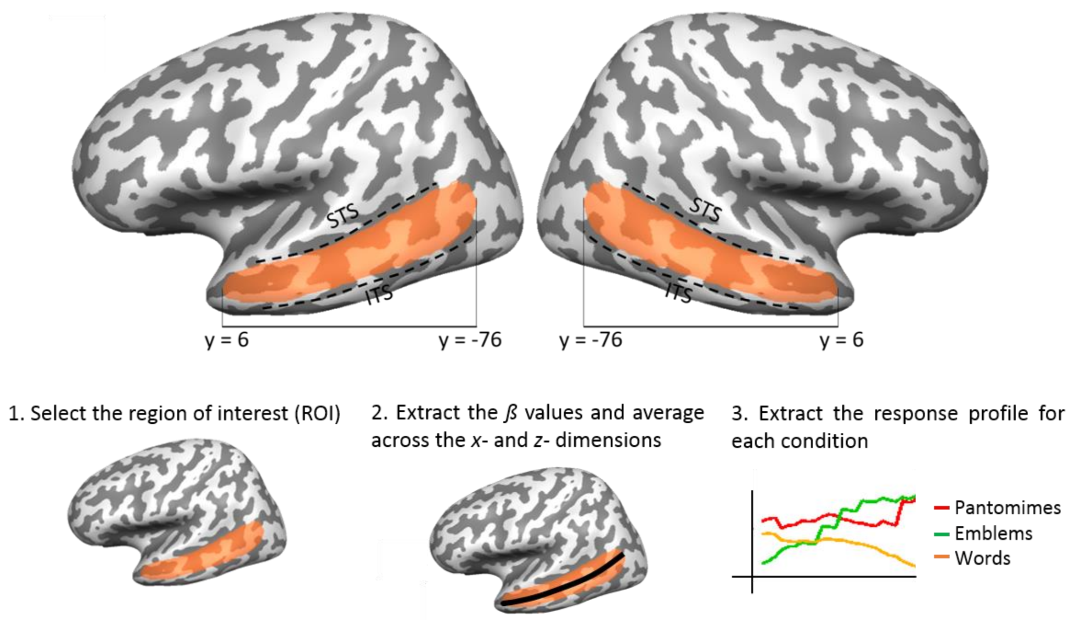


Figure 4.1. Top) The MTG ROI was defined as the cortical region between the STS and the IFG. Bottom) Procedure used for the MTG ROI analysis.

For each voxel in the MTG ROI, we extracted beta estimates for each experimental condition, separately for each participant. The design matrix and predictors were the same used for computing the RFX GLM for the univariate analysis (see *fMRI data analysis - Univariate analysis*). Beta estimates were averaged across the *x*- and *z*- dimensions to obtain one value for each *y*-coordinate (see **Figure 4.1, bottom**).

To statistically assess the difference between stimulus categories at different spatial positions, we computed paired t-tests for the contrasts of interest and corrected for multiple comparisons, using Threshold-Free Cluster Enhancement (TFCE; Smith and Nichols, 2009) with default implementation in CoSMoMVPA (Oosterhof et al., 2016).

4.4 Results

<u>Univariate analysis</u>

Figure 4.2 shows the contrasts [allGestures (pantomimes + emblems + meaningless gestures) > rest] (**Figure 4.2A**) and [words > rest] (**Figure 4.2B**). Both contrasts revealed bilateral activity in the temporal region, confirming the sensitivity of this area to our experimental stimuli. In particular, the contrast [gestures > rest] revealed activity in the posterior part of the MTG. Additionally, we observed the recruitment of a wide network of parietal and middle frontal areas (**Figure 4.2A**). The contrast [words > rest] revealed activity in the middle and anterior portion of MTG, in the superior temporal gyrus (STG), as well as in the inferior and middle frontal gyrus (**Figure 4.2B**).

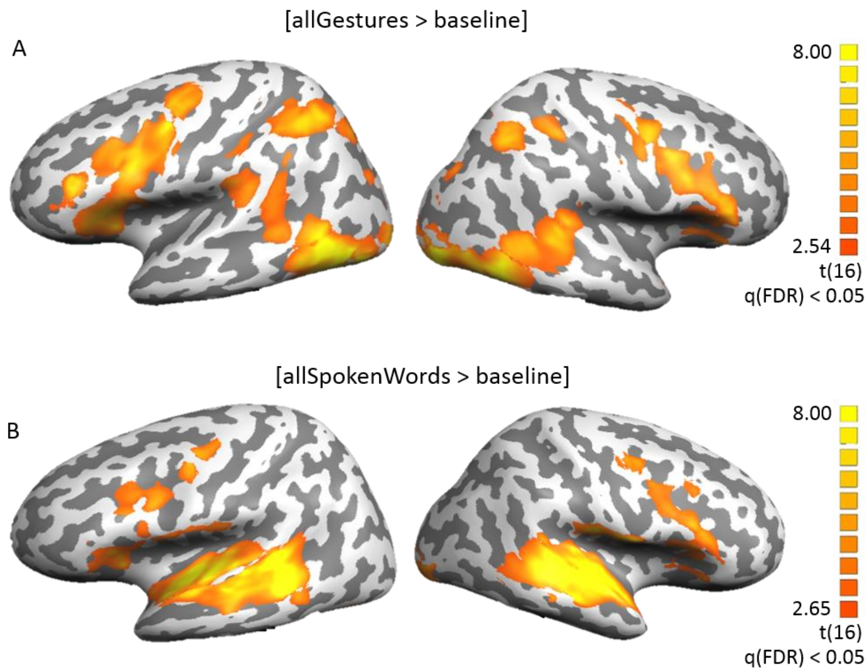


Figure 4.2. Activity against rest for gestures and words. A) RFX GLM [allGestures (pantomimes + emblems + meaningless gestures) > rest]; $q_{(FDR)}$ <0.05; N=17. B) RFX GLM [allSpokenWords (verbs + nouns) > rest]; $q_{(FDR)}$ <0.05; N=17.

<u>Posterior-to-anterior organization of the MTG</u>

Beta estimates for each category of interest were plotted as a function of the *y*-coordinate within the MTG ROI (for details, see *fMRI data analysis – Posterior-to-anterior organization*).

The comparison between meaningful (collapsing across pantomimes and emblems) and meaningless gestures, showed a statistically significant difference in the left hemisphere along the entire MTG (**Figure 4.3A**). This was true also when pantomimes (**Figure 4.3C**) and emblems were contrasted with meaningless gestures, separately (**Figure 4.3E**). In the right hemisphere, the comparison between meaningful and meaningless gestures was significant at anterior and posterior positions, but not at an intermediate position around –y = 20 (**Figure 4.3B**). This was the case also for pantomimes (**Figure 4.3D**). Emblems, in comparison to meaningless gestures, showed a higher activity only in the middle (around y = -30) and anterior (around y = 0) part of the ROI (**Figure 4.3F**).

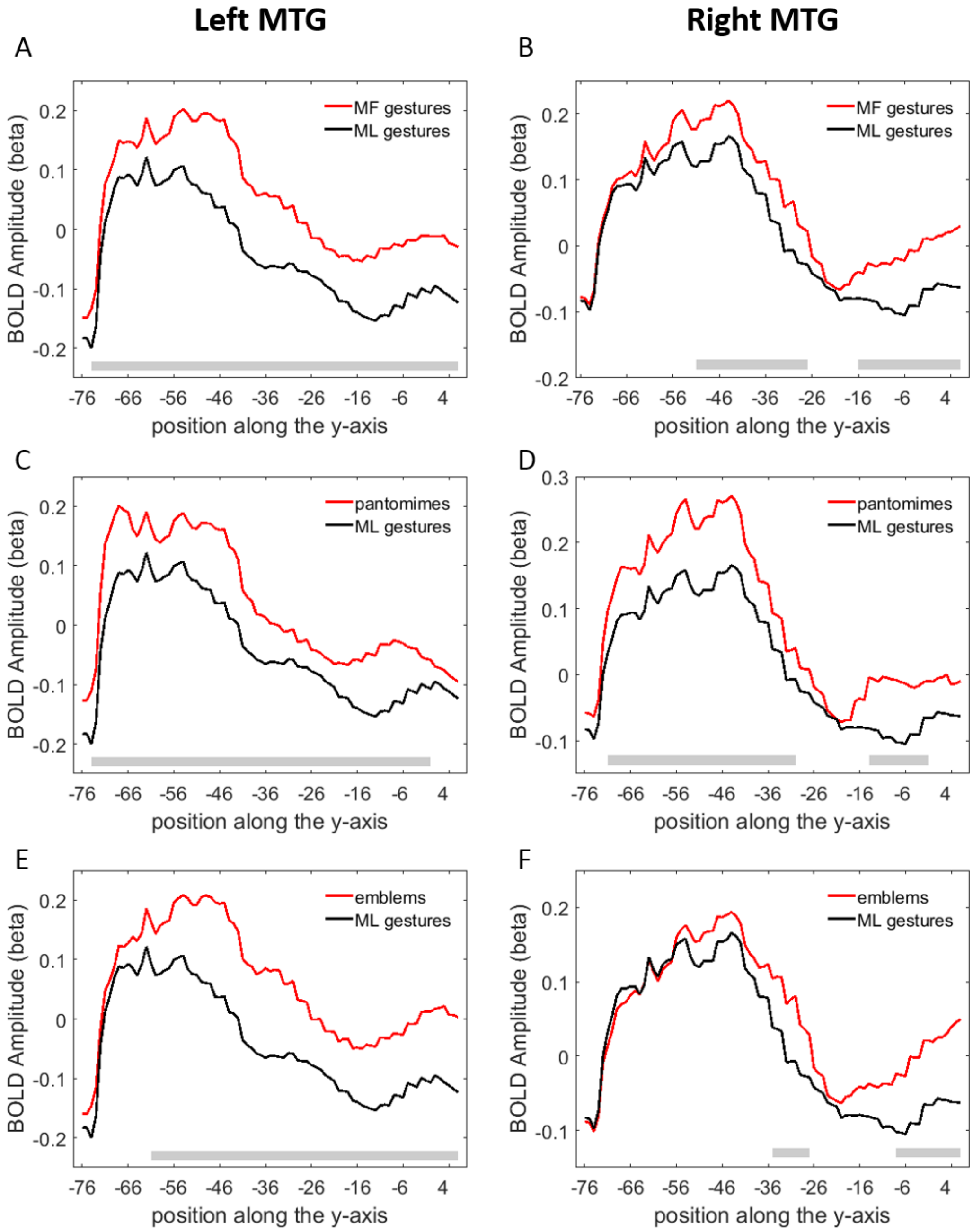


Figure 4.3. Response to meaningful (MF) and meaningless (ML) gestures as a function of the *y*-position along the left (left column) and right (right column) MTG. A, B) MTG response for meaningful and meaningless gestures. C, D) MTG response for pantomimes and meaningless gestures. E, F) MTG response for emblems and meaningless gestures. Grey bars indicate clusters that survived the correction for multiple comparisons using TFCE (see *Material and Methods – Gradient analysis* for details).

Once established that the MTG is sensitive to the categorical distinction between meaningful and meaningless gestures, we plotted the beta estimates for each meaningful category (pantomimes, emblems and words) as a function of the *y*-coordinate within the MTG ROI (**Figure 4.4**). We observed a representational structure along the posterior-to-anterior axis in the MTG, whereby posterior regions responded more strongly to gestures than words, while anterior regions showed a stronger response to words than gestures. This pattern held both in the left (**Figure 4.4A**) and the right (**Figure 4.4B**) hemisphere. In an intermediate region in the left hemisphere (around y = -30), emblems and words showed comparable activity that was stronger than the activity for pantomimes (**Figure 4.4A**).

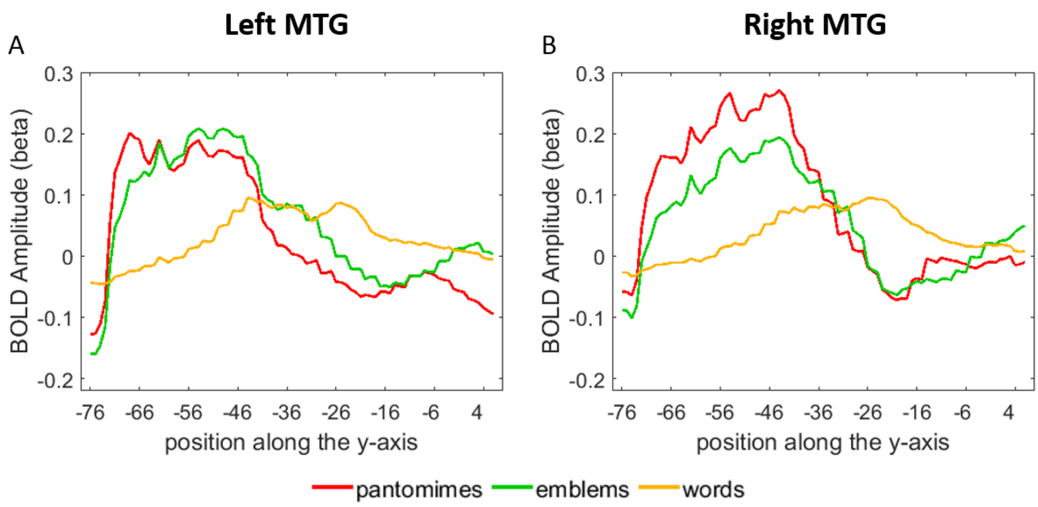


Figure 4.4. Beta estimates for pantomimes, emblems, words as a function of the *y*-position along the left (A) and right (B) MTG.

Statistical analysis supported this description (**Figures 4.5-8**). Gestures in comparison to words led to a significantly stronger activity in the posterior region of MTG (**Figure 4.5A, B**), both in the left and right hemisphere. By contrast, words in comparison to meaningful gestures led to a stronger response in anterior regions of the MTG bilaterally. Emblems showed a higher response than pantomimes in an intermediate region along the gradient (around y = -30) in the left hemisphere (**Figure 4.5G**). Pantomimes showed a higher response in comparison to emblems in the right posterior MTG **(Figure 4.5H)**. Emblems in comparison to words showed a stronger response in posterior and anterior regions bilaterally (**Figure 4.5E, F**). A similar pattern was observed when comparing pantomimes to words (**Figure 4.5C, D**). Note that in the same region that shows a significant difference between emblems and pantomimes (around y = -30; **Figure 4.5G**), emblems and words led to comparable responses **(Figure 4.5E)**.

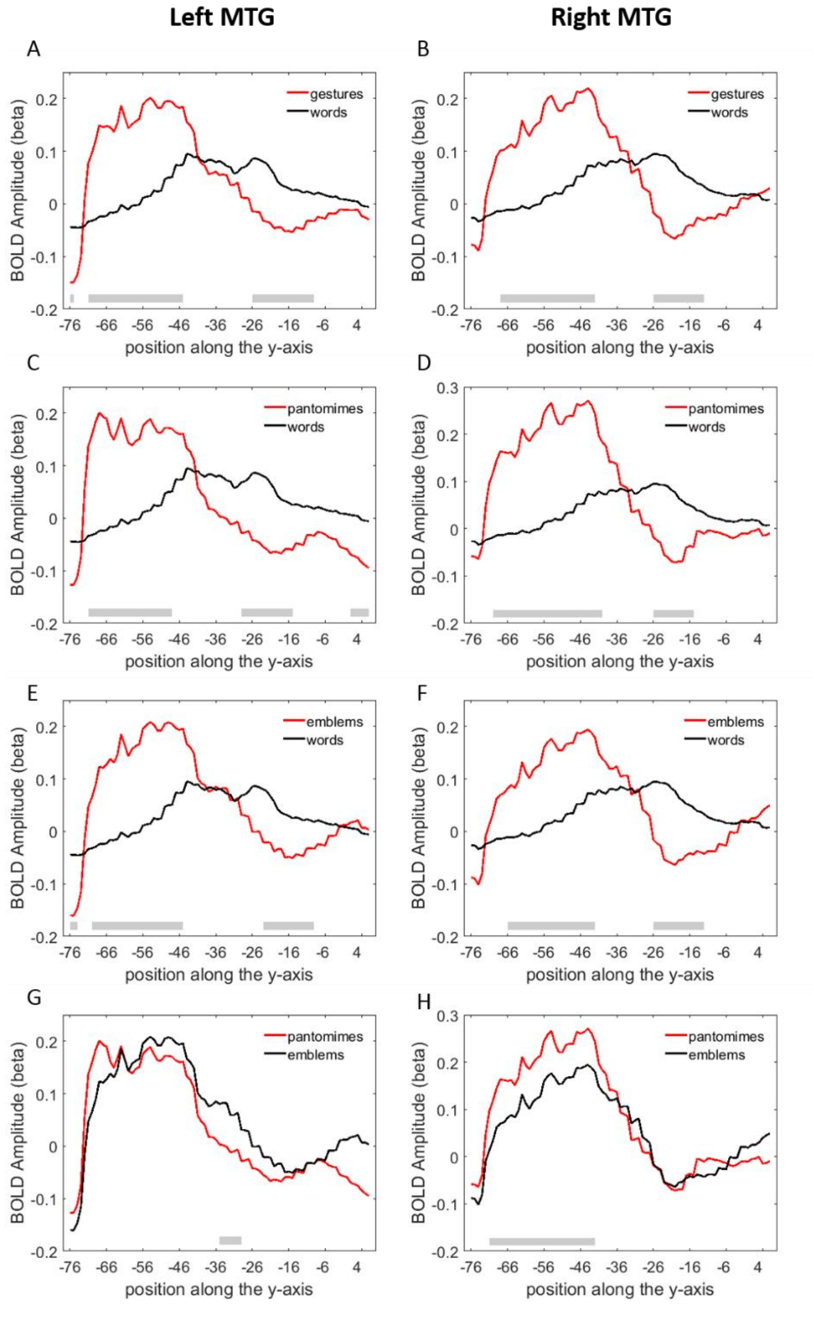


Figure 4.5. Response to different stimulus categories as a function of the *y*-coordinate along the left (left column) and right (right column) MTG. A, B) Meaningful gestures (collapsed across pantomimes and emblems -MF) in comparison to words (collapsed across nouns and verbs). C, D) Pantomimes in comparison to words. E, F) Emblems in comparison to words. G,H) Pantomimes in comparison to emblems. Grey bars indicate clusters that survived the correction for multiple comparisons using TFCE.

4.5 Discussion

MTG is consistently recruited during both action observation and word processing. Many studies have suggested that information in this part of the brain could be organized along a posterior-to-anterior gradient, reflecting a different sensitivity to different input modalities. In particular, it has been shown that posterior regions are preferentially recruited during the processing of visual motion, and anterior regions are specially recruited during the processing of abstract, symbolic stimuli such as words (Kable et al., 2005; Chatterjee, 2008; Watson et al., 2013; Weiner and Grill-Spector, 2013; Tarhan et al., 2015; Lingnau and Downing, 2015). Here we carried out a systematic investigation of the organization of information in the MTG, considering both the role of the modality (verbal or gestural) through which a meaning is expressed and of the relationship between the sign and its meaning. Our results show an internal organization along the posterior-to-anterior axis, whereby posterior regions responded more strongly to gestures (pantomimes and emblems) than words, while anterior regions showed a stronger response to words than gestures. In an intermediate region in the left hemisphere (around y = -30, BA21), we observed a significantly higher response to words and emblems (i.e., stimuli with an arbitrary sign-to-meaning mapping) in comparison to pantomimes. In the following section, we will discuss these results in more detail.

Meaningful vs. meaningless gestures

We found that the MTG is sensitive to the categorical distinction between meaningful and meaningless gestures (see also Villareal et al., 2008). This result is in line with the view that the MTG plays an important role in processing semantics (Kable et al., 2005; Noppeney et al., 2005; Villareal et al., 2008; Kalenine et al., 2010; Whitney et al., 2010; Visser et al., 2012; Peelen et al., 2012; Papeo et al., 2014).

We found that the MTG is sensitive to the difference between pantomimes and meaningless gestures in both hemispheres. By contrast, only the left MTG is sensitive to the difference between emblems and meaningless gestures. Whereas both pantomimes and emblems reflect a meaning in the absence of speech, emblems are symbols whose meaning consists of a more arbitrary relationship between the gesture and the meaning (Papeo and Rumiati, 2013). The arbitrariness between the form and the meaning is a typical aspect of language. This condition suggests a special relationship between emblems and language. Moreover, it has been suggested that emblems, in comparison to pantomimes, incorporate linguistic properties and can therefore be represented in a mental lexicon (McNeill, 2005). It is therefore likely that the preference of the left MTG for emblems in comparison to meaningless gestures is due to the language-like aspects that are typical of emblems (McNeill, 1992; Papeo and Rumiati, 2013).

Pantomimes, Emblems, and Words

Moving along the posterior-to-anterior axis, the neural response to pantomimes and emblems decreased, whereas the response to words increased and slightly decrease in the most anterior portions. At an intermediate position of the left MTG (around y = -30), we observed a plateau where emblems showed a stronger response to pantomimes (**Figure 4.5C**). In the same region, emblems and words did not differ in terms of their responses (**Figure 4.5E**). In a region more anterior to this region, words showed a stronger response than emblems **(Figure 4.5E, F)**. These results are in line with the view that pantomimes, emblems, and speech can be seen as part of a continuum of communicative symbols that differ in the input modality and in the degree of abstractness/arbitrariness of the relationship between the stimulus and the meaning (McNeill, 2005), with the lowest degree of abstractness/arbitrariness for pantomimes, the highest degree of abstractness/arbitrariness for speech, and emblems at an intermediate level. Our results suggest the existence of a continuum where gestures are represented posteriorly and words are represented anteriorly, bilaterally. Only in the left hemisphere, emblems group with pantomimes posteriorly because of their gestural properties. However, emblems also group with words in the middle of the gradient because of their language-like aspects and their arbitrary link between sign and meaning.

Conclusion

Taken together, our results support the view that MTG is sensitive to the difference between meaningful and meaningless stimuli and suggest a spatial organisation of this area, where pantomimes, emblems and speech are arranged along a posterior-to-anterior gradient. This gradient might reflect the sign-to-meaning relationship of the stimuli that is characterized by a gradual decrease of concrete gestural properties and an increase of abstract linguistic properties. This gradient is present in both the left and the right hemisphere, although the left MTG showed a general richer spatial structure.

Chapter 5
General Conclusions

Which brain activations allow us to understand gestures performed by others? Moreover, how does our brain associate symbols (gestures or speech) with meanings? Using fMRI we found that temporal and not precentral areas are sensitive to the categorical distinction between meaningful (pantomimes and emblems) and meaningless gestures. In particular, pantomimes recruited a more posterior region of the MTG, while emblems recruited a slightly more anterior region of the MTG. Furthermore, we observed that the anterior part of the MTG is sensitive to words. Therefore, we investigated how different stimuli (gestures and words) are represented in the MTG. We found that pantomimes, emblems, and speech are arranged along a posterior-to-anterior gradient that reflects the input modality as well as the arbitrariness of the relationship between the form and the meaning.

Unfortunately, our findings from the TMS study (Chapter 3, Experiment 2) were inconclusive regarding the role of temporal and precentral areas in action understanding. However, our results from the univariate analysis of the fMRI study (Chapter 3, Experiment 1) and the analysis of the posterior-to-anterior organization of the MTG (Chapter 4) demonstrated that MTG is indeed sensitive to the semantic content of actions. We showed that this area is sensitive to the categorical distinction between meaningful and meaningless gestures. By contrast, the precentral gyrus did respond to gestures but did not distinguish between meaningful and meaningless actions. Previous studies focused on the activation of motor regions as the trigger of the process of understanding, claiming that re-evoking perceptual and motor states is necessary to accomplish this ability (Rizzolatti and Sinigaglia, 2010; Kiefer and Pulvermüller, 2012). Our findings do not support this view. On the contrary, they are in in agreement with cognitive theories of action understanding, which claim that the content of actions is represented in conceptual areas outside the motor system, and in particular in the occipito-temporal area (Mahon and Caramazza, 2008, Hickok 2009).

Previous studies observed that different parts of the MTG are responding to different stimuli and in particular that visual action stimuli are processed in the posterior part of the gyrus, anterior to MT+, while words are processed in the middle/anterior part of the gyrus (Martin et al., 1995; Kable et al., 2005; Chatterjee, 2008; Watson et al., 2013; Tahran et al., 2015). However, our study (Chapter 4) is the first one that carefully analyzed the posterior-to-anterior organization of the MTG using stimuli that can be allocated along a concrete-to-abstract continuum (**Figure 5.1**).

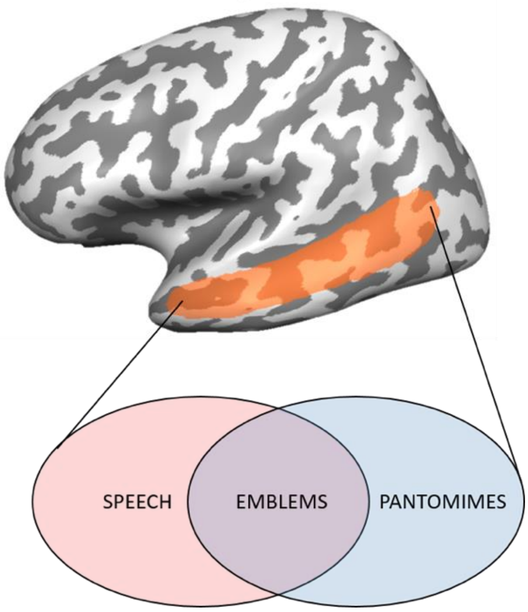


Figure 5.1. Schematic representation of the posterior-to-anterior gradient that characterizes the left MTG.

Pantomimes, emblems, and words are, in fact, all symbols that require a mapping between the form and the mental representation (i.e., the meaning). However, the arbitrariness, and therefore the abstractness of the relationship between the form and the meaning differs gradually. Pantomimes are gestures that faithfully represent the action they refer; emblems are conventional gestures that retain fewer aspects of the real-word referent; words are completely abstract symbols. Interestingly, we found an effect of the input modality in both the right and the left MTG: a posterior preference for gestural stimuli and an anterior preference for linguistic stimuli. However, the left MTG showed a more articulate spatial structure that seemed to reflect the different degrees of abstractness of the relationship between symbols and meaning described above. In particular, emblems group with pantomimes in the left posterior MTG possibly because of their gestural nature. However, emblems also group with words in the middle of the gradient possibly because of their language-like aspects and their arbitrary link between sign and meaning.

I would like to mention and comment on a few criticisms that might be raised in light of the results of the posterior-to-anterior organization of the MTG.

1. The observed posterior-to-anterior organization in MTG might be due to the input modality, visual (gestures) *versus* auditory (spoken word).

Differences in the input modality between gestures and spoken words could indeed partially explain the observed effects. However, it is worth noting that all our stimuli were visually presented. In the case of the spoken words, the actress was keeping a still position and pronouncing words. Moreover, if the posterior-to-anterior gradient was reflecting only the input modality, we should not observe any difference between pantomimes and emblems and any similarity between emblems and words (as we observe in the middle portion of the gradient).

2. The observed posterior-to-anterior gradient in MTG might reflect the movement *versus* static features of the stimuli. In particular, when performing gestures the actress occupies and uses the space around her to move, while when pronouncing words she is standing still without moving.

MTG might be sensitive to the amount of motion observed. However, if this was the only explanation to the effect seen, we should obtain no difference between meaningful and meaningless gestures (as they both contain a similar amount of motion), and between pantomimes and emblems in more anterior portions of MTG. Instead, we found a significant difference between meaningful and meaningless gestures along the entire MTG, suggesting that this area is indeed sensitive to the semantic content of the gesture observed.

3. Were the spoken glosses corresponding to the emblems seen? If so, this could explain the plateau observed in the middle of the gradient in the left hemisphere for emblems and words.

When designing our experiment we took care to avoid a consistent overlap between words pronounced in the spoken word condition and gestures. In the final set of stimuli, there are only 3 cases out of 60 of correspondence between emblems and spoken glosses. Therefore, I would not attribute any effect to this.

In summary, our results indicate that MTG, and not PCG, is sensitive to the semantic content of gestures and that different stimuli (gestures and word) are represented in the MTG following a posterior-to-anterior gradient that reflects the input modality (visual to linguistic) and the arbitrariness of the relationship between the symbol and the meaning (concrete to abstract). Additionally, we offer a standardized dataset made of well-controlled stimuli comprising pantomimes, emblems, and meaningless gestures, complete with detailed norms derived from a large sample of raters, that will help in designing high-quality experiments and will promote replicability.

Chapter 6
Future directions

In Chapter 3.5, I discussed the possible reasons why the fMRI-guided TMS study led to non-significant results. I strongly believe that TMS is a suitable tool to directly compare the role of different brain areas and the behavioral consequences of the stimulation of different sites during the same task. In particular, observing the effect of the stimulation of the temporal and precentral areas on semantic tasks might be the key to shed light on the debate between motor and cognitive theories. For this reason, I think that the approach we developed in the study presented in Chapter 3, Experiment 2 was strong and consistent with the scientific question we were trying to assess. As a future direction, I would work on developing an improved paradigm in order to facilitate the investigation on the effect of the perturbation of the pMTG and PMC during action recognition. One possibility would be to modify the stimuli used in order to make the task more difficult thereby increasing the possibility of observing an effect of the stimulation. Another possibility would be to opt for a different stimulation protocol, such as single pulse online protocol or theta burst stimulation, which might lead to larger effects.

Regarding the posterior-to-anterior organization of the MTG (Chapter 4), it would be interesting to investigate the main effects we discovered in a deeper manner. For example, one could further investigate the word and emblem category observing how the activation of abstract and concrete words, action-related and stative verbs, and emblems-event and emblems-state changes along the gradient. In our study, we had only a few runs dedicated to each sub-category, therefore, we were not able to detect any effect. Adding these categories would allow us to better understand the nature of the gradient observed and to determine links between different categories of stimuli. Moreover, in a follow up study, I would add a meaningless/non-sense word condition in order to confirm the specificity of the middle temporal gyrus in encoding meanings. In our experiment, in fact, meaningful gestures were always compared to meaningless gestures, but we missed a proper contrasting condition for the spoken word category.

Appendices

Appendix 1.

Number of order and Item	Meaningfulness (from 1 to 7) Mean	Median	Proportion of raters providing an answer	Canonical Meaning Agreement		Verbs	Nouns	Other forms
English name	***M***	***M***	***p***	***p***	***H***	***p***	***p***	***p***
1. PANTOMIMES								
1.1 applying make up	4.67	5	79	65	0.77	62	38	
1.2 archery	6.30	6.5	100	68	0.98	26	74	
1.3 blowing a whistle	5.45	6.5	88	54	1.06	39	61	
1.4 blowing nose	5.44	6	88	68	1.02	68	32	
1.5 calling on a telephone	6.88	7	97	100	0.00	33	63	4
1.6 changing a light bulb	6.08	6.5	97	91	0.31	59	41	
1.7 combing 1	6.38	7	100	100	0.00	59	22	19
1.8 combing 2	6.71	7	100	100	0.00	100		
1.9 cutting with a knife	4.40	5	77	96	0.17	83	17	
1.10 drinking	6.96	7	97	100	0.00	83	17	
1.11 driving	6.94	7	100	100	0.00	79	21	
1.12 driving a motorbike	6.68	7	97	100	0.00	58	42	
1.13 eating with a fork	5.04	6	81	100	0.00	88	12	
1.14 eating with a spoon	6.23	7	100	97	0.14	3	97	
1.15 knotting	6.06	6.75	97	100	0.00	60	40	
1.16 nail filing	6.67	7	100	97	0.14	71	23	6
1.17 opening a jar 1	6.66	7	100	94	0.35	78	22	
1.18 opening a jar 2** was: grinding pepper	5.95	6.25	94	55	1.21	79	21	
1.19 opening a lighter	5.15	6	81	92	0.27	27	73	
1.20 opening an umbrella	5.03	7	72	87	0.47	35	65	
1.21 painting a wall	5.66	6	94	90	0.39	83	17	
1.22 painting nails	4.76	5.25	76	96	0.17	44	56	
1.23 peeling a banana	5.99	6.75	94	84	0.61	77	20	3
1.24 playing basketball	6.80	7	100	100	0.00	39	61	
1.25 playing cards	6.39	6.5	94	86	0.50	76	24	
1.26 playing cello	6.59	7	100	66	1.01	47	53	
1.27 playing drums	5.77	6.5	88	89	0.46	21	75	4
1.28 playing flute	6.90	7	100	100	0.00	36	64	
1.29 playing golf	6.50	7	100	100	0.00	30	70	
1.30 playing guitar	6.79	7	100	100	0.00	34	66	
1.31 playing violin	6.89	7	94	100	0.00	41	59	
1.32 pouring from a bottle	6.83	7	100	100	0.00	94	6	
1.33 pushing a button	6.15	7	100	68	0.63	55	35	10
1.34 scribbling** was: erasing	4.93	5.75	84	67	1.11	93	7	
1.35 seasoning	4.89	5	88	90	0.33	76	24	

1.36 sewing	5.15	5.75	84	96	0.16	88	12	
1.37 smoking a cigar	6.34	7	91	90	0.45	79	21	
1.38 smoking a cigarette	6.94	7	100	97	0.14	82	15	3
1.39 spraying deodorant	5.81	6	97	81	0.69	35	62	3
1.40 spraying perfume	6.41	7	94	100	0.00	62	38	
1.41 stapling** was: cutting with scissors	5.54	6.5	87	59	0.87	78	15	7
1.42 stirring	5.00	5	84	92	0.32	96	4	
1.43 sweeping	5.78	6.25	91	87	0.47	93	7	
1.44 taking pictures	6.93	7	90	100	0.00	36	61	4
1.45 throwing	6.32	7	91	100	0.00	90	10	
1.46 turning a key	5.96	5.75	97	69	0.88	80	20	
1.47 turning a key 2** was: screwing	5.25	6	91	45	1.33	72	28	
1.48 turning pages	6.74	7	100	100	0.00	84	16	
1.49 turning pages 2	6.02	6.5	91	93	0.25	80	17	3
1.50 typing on a keyboard	6.63	7	100	69	0.63	75	25	
1.51 uncorking a bottle	4.72	5.75	78	76	0.66	76	24	
1.52 unscrewing	6.44	6.75	97	83	0.49	80	20	
1.53 using a corkscrew	6.42	6.75	94	87	0.47	70	27	3
1.54 using a mobile phone	5.91	6	100	97	0.14	45	55	
1.55 using a remote control	5.07	5.5	81	84	0.61	46	50	4
1.56 using a toothpick	4.38	5	77	63	1.12	63	33	4
1.57 using binoculars	6.29	6.5	100	100	0.00	76	24	
1.58 washing hands	6.17	6.5	94	90	0.39	80	20	
1.59 wearing a necklace	6.39	7	94	100	0.00	42	58	
1.60 wearing a ring	6.84	7	94	100	0.00	43	57	
1.61 wearing a security belt	4.86	5.5	71	59	1.02	50	50	
1.62 wearing earrings	6.67	7	100	100	0.00	33	67	
1.63 wearing glasses	6.57	7	100	88	0.54	42	58	
1.64 whipping	6.33	6.5	100	100	0.00	94	6	
1.65 wiping the mouth	4.95	5.25	88	93	0.26	86	14	
1.66 writing	6.21	6.25	100	97	0.14	84	16	
1.67 zipping up a jacket	5.91	6.5	91	97	0.15	66	27	7
2. MEANINGLESS GESTURES								
2.1 D_1 (peeling a banana)	1.31	1	19	33	1.56			
2.2 D_2 (using binoculars)	1.33	1	19	67	0.87			
2.3 D_3 (pouring from a bottle)	1.37	1	16	60	0.95			
2.4 D_4 (sweeping)	1.00	1	3	100	0.00			
2.5 D_5 (combing)	1.17	1	13	25	1.39			
2.6 D_6 (eating)	1.17	1	10	67	0.64			
2.7 D_7 (erasing)	2.55	2.5	39	67	1.24			
2.8 D_8 (playing flute)	1.25	1	19	50	1.24			
2.9 D_9 (wearing glasses)	1.29	1	25	63	0.90			
2.10 D_10 (using a hairdryer)	1.00	1	9	67	0.64			

2.11 D_11 (waving hello)	1.67	1	24	50	1.21
2.12 D_12 (turning a key)	1.70	1	31	40	1.42
2.13 D_13 (cutting with a knife)	1.53	1	28	44	1.43
2.14 D_14 (opening a lighter)	1.27	1	25	38	1.49
2.15 D_15 (applying lipstick)	1.47	1	30	70	0.61
2.16 D_16 (listening)	1.46	1	23	71	0.80
2.17 D_17 (painting nails)	2.46	1	34	64	1.16
2.18 D_18 (sewing)	1.21	1	23	43	1.56
2.19 D_19 (grinding pepper)	1.86	1	39	38	1.27
2.20 D_20 (praying)	1.52	1	25	13	2.08
2.21 D_21 (praying 2)	1.46	1	21	43	1.48
2.22 D_22 (cleaning)	1.67	1	23	43	1.28
2.23 D_23 (screwing)	1.20	1	10	33	1.10
2.24 D_24 (sleeping)	2.04	1	47	53	1.51
2.25 D_25 (smoking)	1.83	1	21	71	0.80
2.26 D_26 (eating with a spoon)	2.25	1.25	38	58	1.08
2.27 D_27 (calling on a telephone)	1.90	1	26	50	1.21
2.28 D_28 (calling on a telephone)	3.50	2.75	56	61	0.98
2.29 D_29 (using a toothbrush)	2.40	2	32	30	1.83
2.30 D_30 (victory)	1.79	1	30	40	1.33
2.31 D_31 (walking)	1.95	1	36	42	1.70
2.32 D_32 (driving)	1.69	1.5	39	85	0.09
2.33 D_33 (writing)	1.62	1	18	33	1.33
2.34 D_34 (writing 2)	1.30	1	23	71	0.80
2.35 ND_1	1.00	1	13	100	0.00
2.36 ND_2	1.65	1	23	43	1.08
2.37 ND_3	3.09	2.75	61	100	0.00
2.38 ND_4	1.27	1	10	67	0.64
2.39 ND_5	2.665	1.5	24	50	0.97
2.40 ND_6	1.145	1	18	67	0.87
2.41 ND_7	1.37	1	18	50	1.24
2.42 ND_8	2.56	1	44	86	0.41
2.43 ND_9	2.54	2	44	64	0.99
2.44 ND_10	1.24	1	9	100	0.00
2.45 ND_11	1.34	1	30	30	1.97
2.46 ND_12	1.46	1	25	38	1.49
2.47 ND_13	1.00	1	6	50	0.69
2.48 ND_14	1.07	1	13	25	1.39
2.49 ND_15	2.10	1	31	10	2.30
2.50 ND_16	1.13	1	12	100	0.00
2.51 ND_17	1.00	1	6	50	0.69
2.52 ND_18	1.18	1	13	25	1.39
2.53 ND_19	3.07	3	13	50	1.04
2.54 ND_20	1.27	1	16	80	0.50
2.55 ND_21	1.00	1	16	100	0.00
2.56 ND_22	2.29	1	41	54	1.30

2.57 ND_23	1.10	1	10	100	0.00		
2.58 ND_24	2.27	2.25	31	50	1.36		
2.59 ND_25	2.72	1.75	48	73	0.76		
2.60 ND_26	1.03	1	9	67	0.64		
2.61 ND_27	1.16	1	13	25	1.39		
2.62 ND_28	3.08	2	53	53	1.28		
2.63 ND_29	1.76	1	6	50	0.69		
2.64 ND_30	2.29	1	29	44	1.43		
2.65 ND_31	1.35	1	31	30	1.83		
2.66 ND_32	2.13	1	30	40	1.70		
2.67 ND_33	2.21	1	31	50	1.50		
2.68 ND_34	1.00	1	0	0	0.00		
2.69 ND_35	1.12	1	13	75	0.56		
2.70 ND_36	1.66	1	26	50	1.39		

Number of order and Item	**Meaningfulness (from 1 to 7)**		**Proportion of raters providing an answer**	**Canonical Meaning Agreement**		**Emblems Event**	**Emblems State**
	Mean	**Median**					
English name	***M***	***M***	***p***	***p***	***H***	***EE***	***ES***
3. EMBLEMS (Italian raters)							
3.1 agreement** was: congratulation	6.17	7	80	67	0.98		X
3.2 anger	5.93	6.5	80	58	0.68		X
3.3 approving	6.43	7	93	100	0.00		X
3.4 arresting	6.64	7	93	100	0.00	X	
3.5 being late	6.86	7	100	73	0.50		X
3.6 bowing	5.06	6	82	100	0.00	X	
3.7 calling	7	7	100	100	0.00	X	
3.8 caress	4.6	5	73	100	0.00	X	
3.9 clapping	6.71	7	100	100	0.00	X	
3.10 closer	4.13	5	67	90	0.33	X	
3.11 come here	7	7	100	100	0.00	X	
3.12 counting	7	7	100	100	0.00	X	
3.13 coupling	5.93	6	93	100	0.00		X
3.14 cracking the fingers	5.27	6	87	46	1.38		X
3.15 crazy	6.92	7	100	100	0.00		X
3.16 crazy 2	6.71	7	94	100	0.00		X
3.17 crazy 3	6.47	7	100	100	0.00		X
3.18 cunning	6.15	7	80	100	0.00		X
3.19 cutting	6.79	7	100	87	0.39		X
3.20 cutting throat	6.85	7	93	100	0.00		X
3.21 disapproving	5.92	6	87	100	0.00		X
3.22 drinking	6.5	7	100	100	0.00	X	
3.23 eating	7	7	100	100	0.00	X	
3.24 fear	7	7	100	100	0.00		X
3.25 fingers crossed	7	7	100	93	0.24		X

3.26 finished	6.5	7	100	87	0.49	X	
3.27 forgetting	6.92	7	100	87	0.49		X
3.28 four	6.73	7	94	100	0.00		X
3.29 gross smell	6.57	7	93	100	0.00		X
3.30 hitchhiking	5.8	7	93	79	0.52	X	
3.31 horns	5.64	6	80	42	1.10		X
3.32 hungry	5.13	6	87	54	1.26		X
3.33 I beat you	6.58	7	93	100	0.00	X	
3.34 I don't care	6.71	7	100	87	0.39		X
3.35 I don't know	5.77	6	87	100	0.00		X
3.36 I kill myself	6.92	7	100	100	0.00	X	
3.37 I shoot at you	6.92	7	87	100	0.00	X	
3.38 it wasn't me	6.73	7	69	54	0.69		X
3.39 kiss	6.64	7	93	100	0.00	X	
3.40 later	4.73	5	80	92	0.29		X
3.41 listening	7	7	100	47	0.72	X	
3.42 long time ago** was: forget about it	5.86	6	93	57	0.68		X
3.43 looking at	6.25	6.5	100	71	0.75	X	
3.44 looking far	6.21	7	87	100	0.00	X	
3.45 money	6.92	7	100	100	0.00		X
3.46 no	7	7	93	100	0.00		X
3.47 ok	6.8	7	88	100	0.00		X
3.48 one** was: one moment	4.93	6	80	50	1.06		X
3.49 over** was: dying	6.36	7	93	71	0.76		X
3.50 pay attention** was: smart	6	6	100	33	1.36		X
3.51 pointing	6.85	7	100	100	0.00	X	
3.52 praying	5.92	6	80	92	0.29	X	
3.53 praying 2	5.07	6	87	54	0.67	X	
3.54 relief** was: getting tired	6.33	7	94	50	0.69		X
3.55 rising something** was: standing up	5.79	6	93	43	1.20	X	
3.56 rub one's hands	5.92	7	87	69	0.81		X
3.57 run away	7	7	100	87	0.39	X	
3.58 salute	6.71	7	100	100	0.00	X	
3.59 sated	6.27	7	100	53	1.02		X
3.60 sending away	6.3	6	100	100	0.00	X	
3.61 shouting	6.54	7	100	100	0.00	X	
3.62 silence	7	7	100	100	0.00		X
3.63 sleeping	7	7	100	100	0.00	X	
3.64 slowing down	6.64	7	94	100	0.00	X	
3.65 slowly	7	7	100	93	0.24	X	
3.66 smelling	4.87	5	93	64	0.83	X	
3.67 so-so	6	7	100	100	0.00		X
3.68 speaking	5.73	6	100	80	0.68	X	
3.69 stealing	4.53	6	60	100	0.00	X	
3.70 stomach ache	4.87	6	93	86	0.41		X

3.71 stopping	6	6	93	100	0.00	X	
3.72 strangling	6.8	7	100	100	0.00	X	
3.73 swearing	5.87	6	100	100	0.00	X	
3.74 tasty	7	7	100	100	0.00		X
3.75 three	7	7	88	100	0.00		X
3.76 to bear a grudge	4.64	6.5	60	89	0.35		X
3.77 together	4.2	5	67	100	0.00		X
3.78 triumphing	4.94	6	88	53	0.97		X
3.79 unpleasant person	5.92	6	87	100	0.00		X
3.80 victory** was: two	6.57	7	100	67	0.64		X
3.81 walking	6.46	7	93	100	0.00	X	
3.82 waving hello	7	7	100	100	0.00	X	
3.83 waving hello 2	7	7	100	100	0.00	X	
3.84 what do you want	6.6	7	100	100	0.00		X
3.85 winning	5.64	6	87	77	0.93		X
3.86 yawning	6.13	7	88	100	0.00	X	
4. EMBLEMS (non-Italian raters)							
4.1 anger	5.93	7	83	60	1.30		X
4.2 approving	6.20	6	94	100	0.00		X
4.3 being late	6.13	6.5	94	59	0.65		X
4.4 bowing	6.13	7	94	100	0.00	X	
4.5 calling	6.87	7	100	100	0.00	X	
4.6 checking** was: smart	4.38	5	81	38	1.46	X	
4.7 clapping	7.00	7	94	100	0.00	X	
4.8 come here	7.00	7	100	100	0.00	X	
4.9 counting	6.40	7	100	100	0.00	X	
4.10 cutting	6.71	7	100	100	0.00	X	
4.11 cutting throat	6.80	7	100	100	0.00		X
4.12 disapproving	6.40	7	100	100	0.00		X
4.13 eating	6.56	7	100	78	0.65	X	
4.14 exploding	4.63	5.5	75	100	0.00	X	
4.15 fingers crossed	6.29	7	94	100	0.00		X
4.16 four	6.73	7	100	100	0.00		X
4.17 giving** was: begging	4.69	5	88	71	0.76	X	
4.18 good job	4.50	5	100	88	0.38		X
4.19 gross smell	6.87	7	100	88	0.46		X
4.20 handshake** was: congratulating	4.94	6	88	40	1.31		X
4.21 hitchhiking	5.87	6	94	33	1.43	X	
4.22 hungry** was: sated	6.19	6.5	100	53	1.17		X
4.23 I don't know	6.93	7	100	100	0.00		X
4.24 I kill you	7.00	7	100	100	0.00	X	
4.25 I shoot at you	6.47	7	100	89	0.35	X	
4.26 idea	4.69	5	83	73	0.86		X

4.27 it wasn't me** was: I don't know anything	4.56	5	81	46	1.27		X
4.28 kiss	7	7	100	100	0.00	X	
4.29 listening	6.88	7	94	63	0.66	X	
4.30 looking at	6.79	7	100	83	0.56	X	
4.31 looking far	6.13	7	100	81	0.69	X	
4.32 loser	5.56	7	88	100	0.00		X
4.33 mistake** was: forgetting	6.60	7	94	75	0.82		X
4.34 money	6.13	6	94	81	0.69		X
4.35 no	7	7	100	100	0.00		X
4.36 ok	6.79	7	100	100	0.00		X
4.37 one moment	6.00	6.5	94	76	0.71		X
4.38 peace** was: two	5.17	6	89	56	0.86		X
4.39 peace** was: winning	6.53	7	100	56	1.01		X
4.40 pointing	6.79	7	100	100	0.00	X	
4.41 praying 1	6.79	7	100	89	0.35	X	
4.42 praying 2	5.44	5	89	75	0.70	X	
4.43 relief** was: getting tired	7	7	100	56	0.94		X
4.44 salute	6.87	7	100	100	0.00	X	
4.45 shouting	6.75	7	100	53	0.69	X	
4.46 silence	7	7	100	100	0.00		X
4.47 sleeping	7	7	100	100	0.00	X	
4.48 slowing down	5.31	5	100	67	0.93	X	
4.49 so-so	5.18	6	82	71	0.76		X
4.50 speaking	6.12	7	100	88	0.44	X	
4.51 standing up	4.71	5	82	64	0.99	X	
4.52 stomach ache	5.25	5.5	94	88	0.44		X
4.53 stop** was: slowly	5.76	6	100	83	0.79	X	
4.54 stopping	6.93	7	100	81	0.60	X	
4.55 strangling	6.75	7	100	100	0.00	X	
4.56 thinking** was: crazy 2	5.88	6	100	72	0.73	X	
4.57 three	5.60	6	83	100	0.00		X
4.58 triumphing	4.44	5	78	71	0.89		X
4.59 walking	6.73	7	94	100	0.00	X	
4.60 waving hello	6.88	7	100	88	0.36	X	
4.61 waving hello 2	7	7	100	100	0.17	X	
4.62 yawning	6.00	6	89	69	0.83	X	

Note: the *H* value increase as the name agreement decrease. *H*=0 means that only one meaning was given for that gesture.

** indicates video clips that were named differently than the original label.

EE = emblems – event.

ES = emblems – state.

D (derived) = meaningless action that was created starting from a meaningful one. In brackets, the name of the meaningful action from which it was derived.

ND (non-derived) = meaningless action created from scratch.

Appendix 2.

Item	Canonical meaning	Alternative meanings	NA	*
1. PANTOMIMES				
1.1 applying make up	applying make-up (52)	combing hair (24), scratching the eye (3)	NA (21)	
1.2 archery	archery (68)	using a slingshot (20), using a rubber band (6), pulling (3), throwing something (3)		
1.3 blowing a whistle	blowing a whistle (47)	coughing (32), splitting (3), sneezing (3), asthma (3)	NA (13)	
1.4 blowing nose	blowing nose (59)	gross smell (13), tweaking (9), sneezing (3), pretending not to have heard something (3)	NA (13)	
1.5 calling on a telephone	calling on a telephone (97)		NA (3)	
1.6 changing a light bulb	changing a light bulb (88)	screwing/twisting something (9)	NA (3)	
1.7 combing 1	combing hair (100)			
1.8 combing 2	combing hair (100)			
1.9 cutting with a knife	cutting/chopping (74)	knitting (3)	NA (23)	
1.10 drinking	drinking (97)		NA (3)	
1.11 driving	driving (100)			
1.12 driving a motorbike	driving/starting a motorcycle (97)		NA (3)	
1.13 eating with a fork	eating (81)		NA (19)	
1.14 eating with a spoon	eating (97)	cooking (3)		
1.15 knotting	knotting (97)		NA (3)	
1.16 nail filing	nail filing (97)	playing violin (3)		
1.17 opening a jar 1	opening a jar (94)	closing a jar (16)		*
1.18 opening a jar 2** was: grinding pepper	opening a jar (52)	grinding pepper (19), closing (19), squeezing (3), grating (3)	NA (6)	*
1.19 opening a lighter	opening a lighter (75)	using a remote control (6)	NA (19)	
1.20 opening an umbrella	opening an umbrella (63)	fishing (6), knotting (3)	NA (28)	
1.21 painting a wall	painting (85)	writing on a board (6), zipping (3)	NA (6)	
1.22 painting nails	painting nails (73)	scratching (3)	NA (24)	
1.23 peeling a banana	peeling a banana (79)	opening/unwrapping (6), flipping (6), indicate a game (3)	NA (6)	
1.24 playing basketball	playing basketball (100)			
1.25 playing cards	dealing cards (82)	handing out papers (6), pay/count money (6)	NA (6)	
1.26 playing cello	playing cello (66)	playing violin (19), playing bass (19), ironing (3)		*
1.27 playing drums	playing drums (78)	hammering (3), ringing bells (3), joining (3)	NA (13)	
1.28 playing flute	playing flute (100)			
1.29 playing golf	playing golf (100)			
1.30 playing guitar	playing guitar (100)			
1.31 playing violin	playing violin (94)		NA (6)	
1.32 pouring from a bottle	pouring (100)			
1.33 pushing a button	pushing a button (68)	pointing at someone or something (32)		
1.34 scribbling** was: erasing	scribbling (56)	erasing (13), cleaning (6), painting (3), scrapping (3), drizzling (3)	NA (16)	

1.35 seasoning	seasoning (79)	shaking something (9)	NA (12)	
1.36 sewing	sewing (81)	pouring (3)	NA (6)	
1.37 smoking a cigar	smoking (81)	drinking (3), streamer (3), removing something from the lip (3)	NA(9)	
1.38 smoking a cigarette	smoking (97)	shushing (3)		
1.39 spraying deodorant	spraying (78)	pointing the arm/armpit (9), showing the muscle (6), scratching (3)	NA (3)	
1.40 spraying perfume	spraying perfume (94)		NA (6)	
1.41 stapling** was: cutting with scissors	stapling (52)	cutting with scissors (29), other (6)	NA (3)	
1.42 stirring	stirring/mixing (78)	opening (3), churning (3)	NA (16)	
1.43 sweeping	sweeping (79)	rowing (9), hoeing (3)	NA (9)	
1.44 taking pictures	taking a picture (90)		NA (10)	
1.45 throwing	throwing something (91)		NA (9)	
1.46 turning a key	turning a key (71)	turning/twisting something (20), winding (3), igniting (3)	NA(3)	
1.47 turning a key 2** was: screwing	turning a key (63)	screwing (19), unscrewing (6), opening gas (3)	NA (9)	
1.48 turning pages	turning pages (100)			
1.49 turning pages 2	turning pages/reading (85)	petting an animal (6)	NA (9)	
1.50 typing on a keyboard	typing on a keyboard (69)	playing piano (38)		*
1.51 uncorking a bottle	opening a bottle (59)	removing something (16), pouring (3)	NA (22)	
1.52 unscrewing	unscrewing (81)	screwing (23)	NA (3)	*
1.53 using a corkscrew	removing a cork (82)	pulling out (9), pineapple peeling (3)	NA (6)	
1.54 using a mobile phone	using a mobile phone (97)	using the calculator (3)		
1.55 using a remote control	using a remote control (69)	giving (6), shooting (3), taking (3)	NA (19)	
1.56 using a toothpick	using a toothpick (48)	picking nose (13), secret (10), hiding (3), quiet (3)	NA (23)	
1.57 using binoculars	looking through binoculars (100)			
1.58 washing hands	opening the tap to wash the hands (85)	unscrewing (6), applying lotion (3)	NA (6)	
1.59 wearing a necklace	putting on a necklace (94)		NA (6)	
1.60 wearing a ring	putting on a ring (94)		NA (6)	
1.61 wearing a security belt	wearing a seatbelt (42)	putting something in a bag/pocket (19), holster/sheath (6), picking (3)	NA (30)	
1.62 wearing earrings	wearing earrings (100)			
1.63 wearing glasses	putting on glasses (88)	putting on a hat (9), wearing headphones (3), putting something over the eyes (3)		*
1.64 whipping	whipping/stirring (100)			
1.65 wiping the mouth	wiping the mouth (81)	silence (6)	NA (13)	
1.66 writing	writing (97)	drawing (3)		
1.67 zipping up a jacket	zipping up a jacket (88)	indicate nervousness (3)	NA (9)	
2. MEANINGLESS GESTURES				
2.1 D_1 (peeling a banana)	peeling a banana upside down (6)	covering something (3), knotting (3), lifting two sides of a ribbon (3), winding up (3)	NA (82)	
2.2 D_2 (using binoculars)	carrying a burden (13)	indicate the size (3), explaining something (3)	NA (81)	
2.3 D_3 (pouring from a bottle)	tipping over/spilling (10)	turning over (3), coffee machine (3)	NA (84)	

2.4 D_4 (sweeping)	chopping something with an axe (3)		NA (97)	
2.5 D_5 (combing)	swirling (3)	finishing (3), lasso (3), indicate something magical (3)	NA (88)	
2.6 D_6 (eating)	eating (6)	putting something away (3)	NA (91)	
2.7 D_7 (erasing)	wiping/cleaning a surface (26)	sanding an object to make it smooth (10), levelling (3), fulling the tank (3), handwashing (3)	NA (61)	*
2.8 D_8 (playing flute)	playing an instrument (9)	spider (3), camera (3), other (3)	NA (82)	
2.9 D_9 (wearing glasses)	putting on a mask (16)	wearing something (6), measuring the length of the head (3)	NA (75)	
2.10 D_10 (using a hairdryer)	doing hair (6)	cutting something (3)	NA (91)	
2.11 D_11 (waving hello)	indicate that you can't see me (12)	catching the attention (6), crazy (3), 'five' as a quantity (3)	NA (76)	
2.12 D_12 (turning a key)	screwing (13)	unscrewing (9), rotating (3), snapping fingers (3), opening a lamp (3)	NA (69)	
2.13 D_13 (cutting with a knife)	twirling a banner (13)	western lasso (6), zipping (3), cracking fingers (3), handler (3)	NA (69)	
2.14 D_14 (opening a lighter)	perfume (9)	taking a picture (6), indicate silence (3), indication (3), indicate up and down (3)	NA (76)	
2.15 D_15 (applying lipstick)	applying make-up (21)	writing/painting on the face (9)	NA (70)	
2.16 D_16 (listening)	listening (16)	asking to speak louder (3), other (3)	NA (78)	
2.17 D_17 (painting nails)	writing/painting on the hand (22)	cutting the skin (3), pinching (3), itching (3), finger nailing (3)	NA (66)	
2.18 D_18 (sewing)	sewing (10)	looping (6), wrapping (3), indicate 'over' and 'under' (3), indicate the shape of a spring or coil (3)	NA (77)	*
2.19 D_19 (grinding pepper)	grinding pepper or salt (15)	unscrewing (12), twisting something (9), loosening something (3)	NA (61)	
2.20 D_20 (praying)	praying (3)	joined hands (3), symbol of shark (3), preparing for a plunge (3), yoga position (3), scooping together (3), pushing together (3), joining (3)	NA (76)	
2.21 D_21 (praying 2)	indicate togetherness (9)	pressing (3), yoga position (3), making a sandwich (3), putting something on top of something (3)	NA (79)	
2.22 D_22 (cleaning)	cleaning (10)	petting an animal (6), assuring to someone (3), indicate that something is closed (3)	NA (78)	
2.23 D_23 (screwing)	adjusting item on the body (3)	scratching (3), indicate a valve (3)	NA (91)	
2.24 D_24 (sleeping)	indicate it is bedtime/being sleepy (25)	stretching (6), protecting yourself (30, tilting (3), resting (3), fainting (3), other (3)	NA (54)	
2.25 D_25 (smoking)	smoking (15)	special thank (3), swearing (3)	NA (79)	
2.26 D_26 (eating with a spoon)	pouring something on the head (22)	wearing a hat (9), scooping (3), zipping (3)	NA (63)	
2.27 D_27 (calling on a telephone)	twisting/spinning (13)	asking to finish whatever you are doing (6), repeating an activity (3), winding something up (3)	NA (75)	
4.28 D_28 (calling on a telephone)	slang: to greet or cool/hang loose (34)	calling (16), indicate the length of something (3), showing agreement (3)	NA (44)	
2.29 D_29 (using a toothbrush)	hitting (10)	brushing teeth (6), hammering (3), smelling the perfume from the wrist (3), pulling up and down (3), chopping something (3), telling someone to chisel something (3)	NA (69)	
2.30 D_30 (victory)	pointing at someone (12)	indicate 'you two' (6), keeping an eye on someone (3), blinding (3)	NA (70)	

2.31 D_31 (walking)	walking upside down (15)	flying (6), running (3), indicate that it's going to rain (3), indicate a slope (3), having a haircut (3), catching the attention (3)	NA (64)	
2.32 D_32 (driving)	opening something over the head -trapdoor, valve, submarine door (33)	closing something (6)	NA (61)	
2.33 D_33 (writing)	dealing cards or something (6)	sprinkling something (6), scribbling (3), handing or shelling things out (3)	NA (82)	
2.34 D_34 (writing 2)	itching (16)	stitches (3), rubbing (3)	NA (78)	
2.35 ND_1	joining/meeting (13)		NA (87)	
2.36 ND_2	whisper (10)	hiding from the sight (6), sleeping (6)	NA (78)	
2.37 ND_3	covering the eyes not to see something (61)		NA (39)	
2.38 ND_4	showing the height (6)	indicate something is done (3)	NA (91)	
2.39 ND_5	lying (12)	make fun of someone (9), indicate a drunk person (3)	NA (76)	
2.40 ND_6	crashing (12)	indicate togetherness (3), pushing (3)	NA (82)	
2.41 ND_7	grabbing something (9)	hiding something in the hand (3), wrapping (3), closing something (3)	NA (82)	
2.42 ND_8	togetherness (38)	others (6)	NA (56)	
2.43 ND_9	drawing a spiral (28)	symbol of internet '@' (9), indicate zero (3), turning (3)	NA (57)	
2.44 ND_10	crying/sadness (9)		NA (91)	
2.45 ND_11	attached together/linked (9)	stacked (3), indicate an agreement (3), indicate submission (3), indicate 'go to jail' (3), crashing fists (3), putting fists together (3), indicate being strong (3)	NA (70)	
2.46 ND_12	recalling (9)	indicate headache (6), indicate disappointment (3), indicate craziness (3), thinking (3)	NA (76)	
2.47 ND_13	thinking (3)	shushing (3)	NA (94)	
2.48 ND_14	showing the level (3)	showing something (3), dividing in half (3), indicate someone unpleasant (3)	NA (88)	
2.49 ND_15	crashing on the face (3)	crazy (3), proceeding (3), indicate the centre of the face (3), position of defence (3) measuring approximately (3), salute (3), loser (3), indicate the sound of the train (3), two-face (3)	NA (70)	
2.50 ND_16	salute (12)		NA (88)	
2.51 ND_17	to thank (3)	indicate 'head up' (3)	NA (94)	
2.52 ND_18	spreading (3)	stopping (3), dropping (3), driving (3)	NA (88)	
2.53 ND_19	pointing the top of the nose (6)	indicate equality (3), touching the nose (3)	NA (88)	
2.54 ND_20	crossing fingers to express good luck (13)	other (3)	NA (84)	
2.55 ND_21	indicate 'down' (16)		NA(84)	
2.56 ND_22	indicate two sides (22)	stopping (6), flipping (6), indicate cards (3), checking something (3)	NA (60)	
2.57 ND_23	consoling (10)		NA (90)	
2.58 ND_24	indicates to come closer (16)	asking to kiss the hand (6), indicate 'down' (3), indicate the ring (3)	NA (72)	
2.59 ND_25	indicate fever/feeling unwell (35)	forgetting (13), thinking (3)	NA (52)	*
2.60 ND_26	indicate the back of the head (6)	adjusting a hat (3)	NA (91)	

2.61 ND_27	indicate a pledge (3)	hitting someone (3), salute (3), indicate respect (3)	NA (88)	
2.62 ND_28	stopping (28)	doing physical exercise (9), going away (9), lifting (3), parking (3)	NA (48)	
2.63 ND_29	sent to death -roman empire (3)	hitchhike (3)	NA (94)	
2.64 ND_30	indicate respect/honesty (13)	salute (6), indicate agreement (3), leaving someone (3), indicate myself (3)	NA (72)	
2.65 ND_31	antler horn (9)	three' as a number (6), transmitting an information (3), bullying (3), saying bye (3), knowing (3), giving hints (3)	NA (70)	
2.66 ND_32	ok/approved (12)	indicate a hole (6), indicate 'zero' (6), indicate the diameter (3), surrounding (3), 'three' as a number (3)	NA (70)	*
2.67 ND_33	ok, perfect (16)	deer (3), indicating 'small' (3), posture (3), symbol of eye (3), other (3)	NA (69)	
2.68 ND_34			NA (100)	
2.69 ND_35	asking for a pause/time out (10)	being subdued (3)	NA (87)	
2.70 ND_36	so-so/half and half (13)	indicate different situations (3), indicate a direction (3), shaking (3), indicate something worthless (3)	NA (75)	
3. EMBLEMS (Italian raters)				
3.1 agreement** was: congratulation	having an agreement (53)	greetings (13), congrats (7), being together (7)	NA (20)	
3.2 anger	indicate anger (47)	yawning (33)	NA (20)	
3.3 approving	approving (93)		NA (7)	
3.4 arresting	indicate the detention of someone (93)		NA (7)	
3.5 being late	indicate that is late (73)	asking "what's the time" (67)		*
3.6 bowing	bowing (82)		NA (18)	
3.7 calling	calling (100)			
3.8 caress	caressing (73)		NA (27)	
3.9 clapping	clapping (100)			
3.10 closer	asking someone to come closer (60)	asking to kiss the hand (7)	NA (33)	
3.11 come here	asking someone to come closer (100)			
3.12 counting	counting to five (100)			
3.13 coupling	indicate union between persons or objects (93)		NA (7)	
3.14 cracking the fingers	cracking the fingers (40)	catching the attention (20), indicate that something is almost done (13), having an idea (7), keeping the rhythm (7)	NA (13)	
3.15 crazy	indicate that a person is crazy (100)			
3.16 crazy 2	indicate that someone is crazy (94)		NA (6)	
3.17 crazy 3	“are you crazy??” (100)			
3.18 cunning	indicate astuteness (80)		NA (20)	
3.19 cutting	cutting someone short (87)	symbol of scissors (13)		
3.20 cutting throat	indicate the intention of killing (93)		NA (7)	

3.21 disapproving	disapproving (87)		NA (13)	
3.22 drinking	asking for a drink (100)			
3.23 eating	being hungry and requiring some food (100)			
3.24 fear	indicate fear (100)			
3.25 fingers crossed	fingers crossed/express good luck (93)	indicate agreement (7)		
3.26 finished	indicate that something is finished/there is nothing anymore (86)	more or less (7), displeasure (7)		
3.27 forgetting	realizing something that was forgotten (86)	something was obvious (7), indicate regret (7)		
3.28 four	‘four’ as a quantity (94)		NA (6)	
3.29 gross smell	indicate a gross smell (93)		NA (7)	
3.30 hitchhiking	hitchhiking (73)	indicate a direction (20)	NA (7)	
3.31 horns	offending someone (33)	approving/rock (27), sign of antler (27)	NA (20)	*
3.32 hungry	being hungry (46)	being full (20), indicate an unpleasant person (7), smoothing (7), hand on hip (7)	NA (13)	
3.33 I beat you	intimidating someone (93)		NA (7)	
3.34 I don't care	“I don't care”/indicate indifference (87)	denying (13)		
3.35 I don't know	“I don't know” (69)	“I can't do anything” (25)	NA (31)	*
3.36 I kill myself	shoot me! (100)			
3.37 I shoot at you	indicate the intention of killing (87)		NA (13)	
3.38 it wasn't me	“it wasn't me” (47)	“I don't know” (40)	NA (13)	
3.39 kiss	blowing a kiss (93)		NA (7)	
3.40 later	postponing something (73)	going on (7)	NA (20)	
3.41 listening	to listen with more attention (47)	indicate that I did not understand (40)	NA (13)	
3.42 long time ago** was: forget about it	indicate something that happened long time ago (53)	“forget about it” (40)	NA (7)	
3.43 looking at	checking/keeping an eye on (70)	watching (24), other (6)		
3.44 looking far	looking far away (100)			
3.45 money	indicate money (100)			
3.46 no	no (100)			
3.47 ok	ok, perfect (88)		NA (12)	
3.48 one** was: one moment	‘one’ as a quantity (40)	asking for a moment (27), asking for speaking (20)	NA (20)	*
3.49 over** was: dying	indicate that something is over (66)	indicate that someone is dead (20), cutting a half of something (7)	NA (7)	
3.50 pay attention** was: smart	paying attention (33)	keeping an eye on someone (27), looking (20), indicate astuteness (20)		
3.51 pointing	pointing at a person (100)			
3.52 praying	to pray someone (73)	thank you in Japanese (7)	NA (20)	
3.53 praying 2	begging someone (47)	praying (47)	NA (13)	*
3.54 relief** was: getting tired	expressing relief (47)	expressing effort in doing something (47)	NA (6)	
3.55 raising something** was: standing up	to raise something (40)	invite someone to stand up (33), to bounce (13), to lift (7)	NA (7)	
3.56 rub one's hands	foretaste of victory or something (60)	washing hands (27), I don't care (7)	NA (13)	*
3.57 run away	driving away/shoo someone (87)	threatening (13)		

3.58 salute	salute (100)			
3.59 sated	being full in the stomach (53)	being hungry (33), massaging the belly (20)		*
3.60 sending away	kicking someone away (100)			
3.61 shouting	to shot/call aloud (100)			
3.62 silence	“silence!” (100)			
3.63 sleeping	being sleepy/needing sleep (100)			
3.64 slowing down	slowing down/keeping calm (94)		NA (6)	
3.65 slowly	indicate to wait (93)	stop (7)		
3.66 smelling	indicate bad smell (59)	indicate astuteness (27), suspecting something (7)	NA (7)	
3.67 so-so	so-so/indicate uncertainty (100)			
3.68 speaking	indicate that someone is speaking too much (80)	indicate to open the light of the car (20), indicate the presence of something (7)		*
3.69 stealing	indicate stealing something (60)		NA (40)	
3.70 stomach ache	having a stomach ache (80)	bowing (13)	NA (7)	*
3.71 stopping	stopping (93)		NA (7)	
3.72 strangling	strangling metaphorically (100)			
3.73 swearing	swearing not to talk (100)			
3.74 tasty	indicate that a food is good (100)			
3.75 three	‘three’ as a quantity (88)		NA (12)	
3.76 to bear a grudge	Indicate that I will remember something bad that happened (53)	substituting (7)	NA (40)	
3.77 together	indicate agreement or union (67)		NA (33)	
3.78 triumphing	triumphing (47)	gesture of communism (29), other (12)	NA (12)	
3.79 unpleasant person	finding someone unpleasant (87)		NA (13)	
3.80 victory** was: two	indicate victory (67)	"two" of something (40)		*
3.81 walking	walking (93)		NA (7)	
3.82 waving hello	saying hi (100)			
3.83 waving hello 2	saying hi (100)			
3.84 what do you want	“what do you want?” (100)			
3.85 winning	indicate victory (67)	‘two’ as a quantity (20), indicate peace (7), scout gesture (7)	NA (13)	*
3.86 yawning	yawning because of tiredness or boredom (88)		NA (12)	
4. EMBLEMS (non-Italian raters)				
4.1 anger	yawning (49)	indicate nervousness (11), coughing (6), indicate a mistake (6), having bad news (6), indicate pain (6)	NA (16)	
4.2 approving	thumb up/approving (94)		NA (6)	
4.3 being late	asking/checking for the time (56)	inviting someone to be aware of the time (50)	NA (6)	*
4.4 bowing	bowing to salute or thank (94)		NA (6)	
4.5 calling	calling (100)			

4.6 checking** was: smart	checking (31)	crying/being sad (25), watching (19), wondering what happened (6), keeping a suspicious eye on someone (6)	NA (19)	*
4.7 clapping	clapping/applause (94)		NA (6)	
4.8 come here	asking to come here (100)			
4.9 counting	counting to five (100)			
4.10 cutting	cut something with scissors (100)			
4.11 cutting throat	cutting throat (100)			
4.12 disapproving	thumb down/disagreement (100)			
4.13 eating	eating something (77)	being hungry (17), indicate a pinch (6)		
4.14 exploding	breaking apart (75)		NA (25)	
4.15 fingers crossed	indicate good luck (94)		NA (6)	
4.16 four	‘four' as a quantity (100)			
4.17 giving** was: begging	giving someone something (63)	asking for something (19), showing something (6)	NA (12)	
4.18 good job	congratulating myself/good job (88)	indicate shoulder pain (12)		
4.19 gross smell	indicate something stinky (88)	holding the breath (6), breathing Indian exercises (6)		
4.20 handshake** was: congratulating	handshake (35)	indicate an agreement (24), being friends (18), holding hands (12)	NA (11)	
4.21 hitchhiking	hitchhiking (32)	pointing (25), getting out (25), highlighting (6), mentioning (6)	NA (6)	*
4.22 hungry** was: sated	being hungry (53)	having a full stomach (29), indicate that food tastes good (18), indicate a part of the body (6)		*
4.23 I don't know	“I don't know”/”whatever” (100)			
4.24 I kill you	shooting at yourself (100)			
4.25 I shoot at you	shooting/threatening (89)	indicate 'you' (11)		
4.26 idea	getting an idea (61)	thinking (11), forgetting (6), other (6)	NA (17)	
4.27 it wasn't me** was: I don't know anything	“it wasn't me” (38)	indicate hands off (19), feeling crowded (13), indicate being undecided (13)	NA (19)	
4.28 kiss	blowing a kiss (100)			
4.29 listening	listening (59)	asking to speak up (35)	NA (6)	
4.30 looking at	looking/watching (83)	showing eyes (11), demonstrating awareness (6)		
4.31 looking far	looking (82)	salute (6), shading eyes from the sun (6), showing the height (6)		
4.32 loser	loser (88)		NA (12)	
4.33 mistake** was: forgetting	indicate a dumb mistake (71)	being upset (12), remembering something (6), hitting (6)	NA (5)	
4.34 money	indicate money (76)	a pinch of something (6), checking the structure (6), twisting (6)	NA (6)	
4.35 no	no (100)			
4.36 ok	ok, fantastic (100)			
4.37 one moment	indicate to wait one minute (72)	warning (11), 'one' as a quantity (11)	NA (6)	
4.38 peace** was: two	sign of peace (50)	two' as a quantity (33), indicate to wait two minutes (6)	NA (11)	
4.39 peace** was: winning	sign of peace (56)	two' as a quantity (38), sign of victory (19)		*
4.40 pointing	pointing at something/someone (100)			

4.41 praying 1	praying (89)	greeting (11)		
4.42 praying 2	praying (67)	indicate union (17), shaking hands (6)	NA (10)	
4.43 relief** was: getting tired	sign of relief (56)	wiping sweat (33), being tired (11)		
4.44 salute	salute (100)			
4.45 shouting	yelling (53)	calling (47)		
4.46 silence	being quiet/silent (100)			
4.47 sleeping	sleeping (100)			
4.48 slowing down	indicate calm down (66)	pushing down (22), staying (6), flattering something (6)		
4.49 so-so	so-so/not sure (59)	indicate that everything is all right (18), toasting (6)	NA (17)	
4.50 speaking	indicate that someone is talking too much (88)	blinking (6), fishing (6)		
4.51 standing up	asking to stand up (53)	raising up something (18), asking for more (6), bouncing (6)	NA (17)	
4.52 stomach ache	indicate a stomach ache (83)	being full (6), bowing (6)	NA (6)	
4.53 stop** was: slowly	stop (83)	slowing down (6), take five (6), honking horn (6), pushing (6)		*
4.54 stopping	stopping (81)	pushing (13), indicate welcome (6)		
4.55 strangling	choking (100)			
4.56 thinking** was: crazy 2	thinking (72)	indicate knowledge (22), indicate craziness (6)		
4.57 three	'three' as a quantity (83)		NA (17)	
4.58 triumphing	sign of victory (56)	indicate power (28), indicate solidarity (11)	NA (22)	*
4.59 walking	walking (94)		NA (6)	
4.60 waving hello	waving goodbye (88)	asking someone to come closer (12)		
4.61 waving hello 2	waving goodbye (100)	disagreeing (6)		*
4.62 yawning	yawning (61)	indicate silence (17), covering the mouth (11)	NA (11)	

** indicates video clips that were named differently than the original label.

NA = no answers (proportion of participants that did not provide any answer).

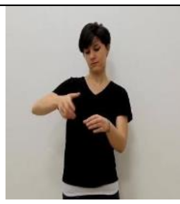
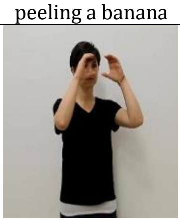
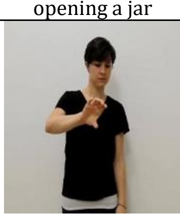
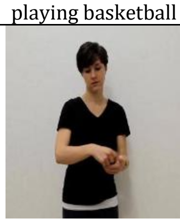
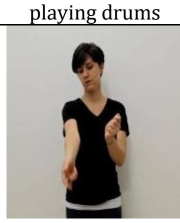
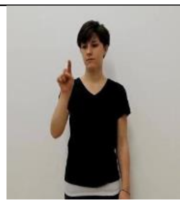
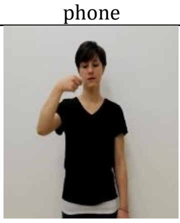
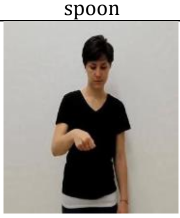
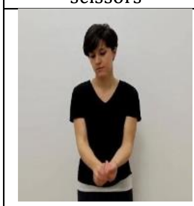
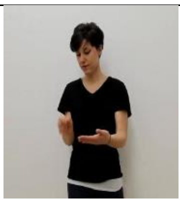
* indicates when one or more participants provided more than one meaning.

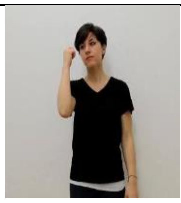
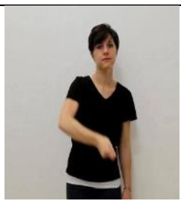
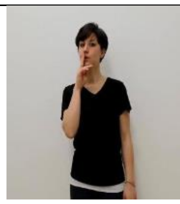
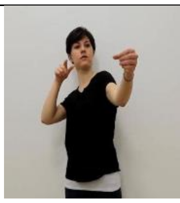
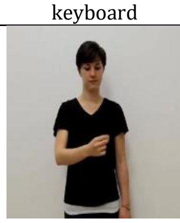
D (derived) = meaningless action that was created starting from a meaningful one. In brackets, the name of the meaningful action from which it was derived.

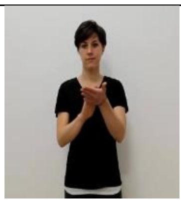
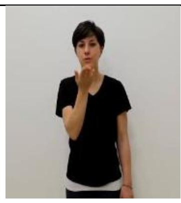
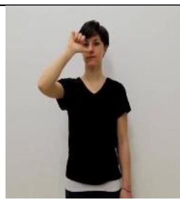
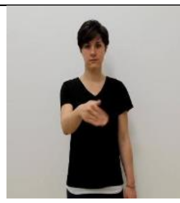
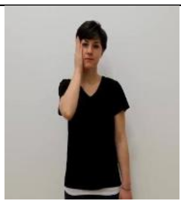
ND (non-derived) = meaningless action created from scratch.

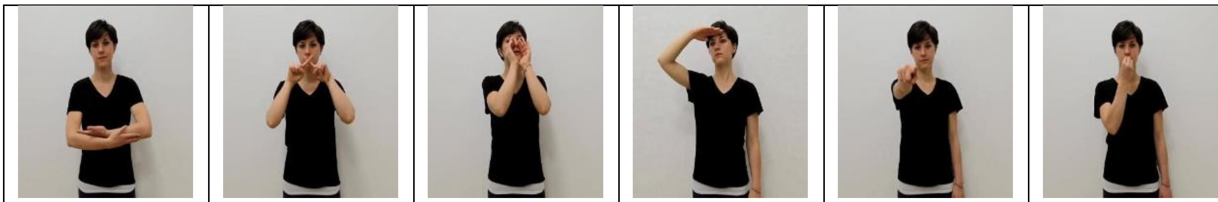
Appendix 3.

Gestures

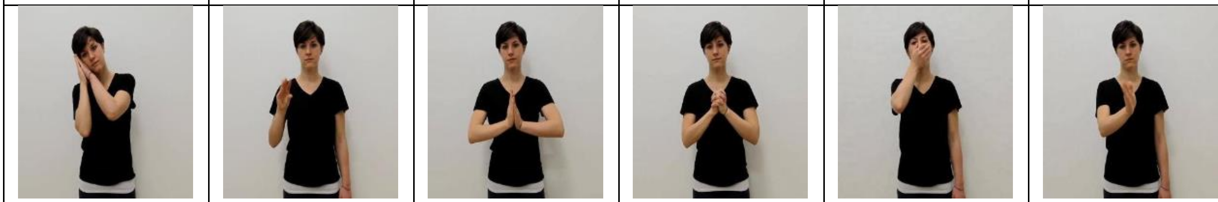
Pantomimes					
					
wearing a ring	peeling a banana	opening a jar	closing a jar	playing basketball	playing drums
					
drinking	using binoculars	pouring from a bottle	pushing a button	playing cards 1	playing cards 2
					
uncorking a bottle	using a mobile phone	turning a key	playing guitar	wearing a belt	ringing the bell
					
wearing a necklace	eating with a spoon	spraying deodorant	blowing nose	blowing a whistle	playing flute
					
cutting with scissors	eating with a fork	stirring	zipping up a jacket	erasing	changing a light bulb
					
washing hands	reading	turning pages	nail filing	taking pictures	playing golf

					
driving a motorbike	knotting	wearing glasses	applying make up	wearing earrings	throwing
					
writing	painting a wall	grinding pepper	combing	spraying perfume	smoking a cigarette
					
smoking a cigar	sweeping	brushing	driving	using a toothpick	typing on a keyboard
					
using a remote control	calling on a telephone	archery	playing violin	playing cello	sewing

Emblems					
					
stand up	clapping	listening	blowing a kiss	drinking	run away
					
walking	caress	waving hello 1	waving hello 2	slowly	counting

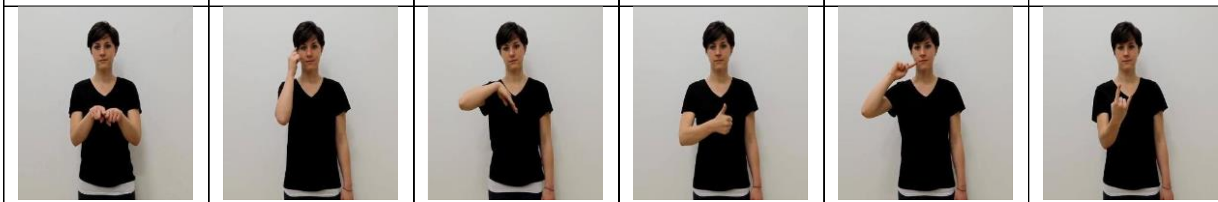


cradling	swearing	shouting	looking far	pointing	eating

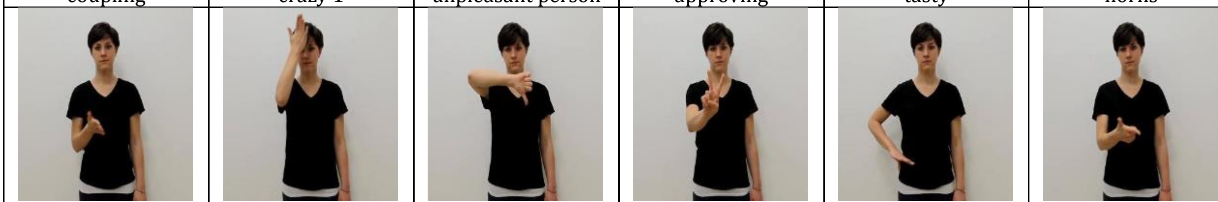


sleeping	speaking	praying 1	praying 2	yawning	sending away

writing	stopping	strangling	I beat you	I shoot at you	come here



coupling	crazy 1	unpleasant person	approving	tasty	horns

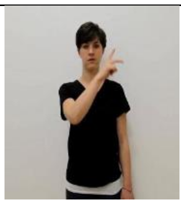
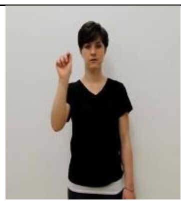
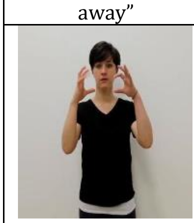
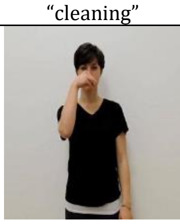
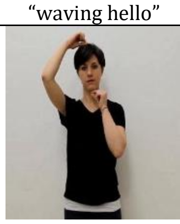
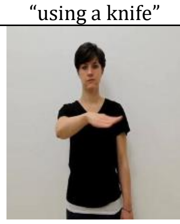
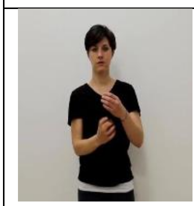
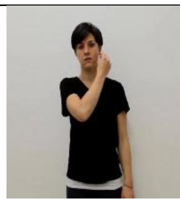
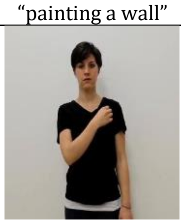
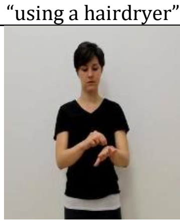
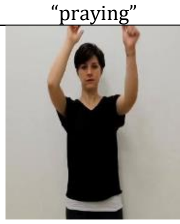


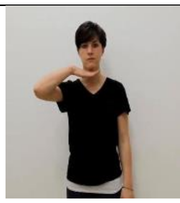
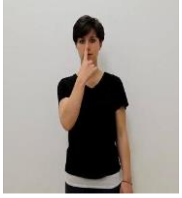
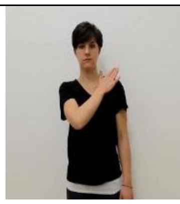
so-so	forgetting	disapproving	two	hungry	it’s over

pain in the belly	crazy 2	to bear a grudge	no	I don't care	It wasn’t me

I don't know	ok	fear	four	anger	crazy 3

					
cunning	money	being late	there	linked	victory

Meaningless gestures					
					
derived from "opening a lighter"	Derived from "sewing"	Derived from "listening"	Derived from "banana"	Derived from "using binoculars"	Derived from "pouring from a bottle"
					
Derived from "run-away"	Derived from "walking"	Derived from "cleaning"	Derived from "turning a key"	Derived from "waving hello"	Derived from "using a knife"
					
Derived from "playing flute"	Derived from "smoking a cigarette"	Derived from "eating"	Derived from "sleeping"	Derived from "wearing glasses"	Derived from "erasing"
					
Derived from "grinding pepper"	Derived from "painting a wall"	Derived from "using a hairdryer"	Derived from "praying"	Derived from "praying"	Derived from "applying lipstick"
					
Derived from "writing"	Derived from "writing"	Derived from "painting nails"	Derived from "using a toothbrush"	Derived from "driving"	Derived from "calling on a telephone"

					
Derived from “screwing”	Derived from “victory”	Derived from “sweeping”	Non-derived	Non-derived	Non-derived
					
Non-derived	Non-derived	Non-derived	Non-derived	Non-derived	Non-derived
					
Non-derived	Non-derived	Non-derived	Non-derived	Non-derived	Non-derived
					
Non-derived	Non-derived	Non-derived	Non-derived	Non-derived	Non-derived
					
Non-derived	Non-derived	Non-derived	Non-derived	Non-derived	Non-derived

Spoken words

Action verbs	Non-action verbs
io accarezzo [I caress]	io adoro [I adore]
io acchiappo [I grab]	io affascino [I fascinate]
io afferro [I grasp]	io aggiorno [I update]
io ammanetto [I handcuff]	io alludo [I allude]
io annodo [I knot]	io ammiro [I admire]

io annoto [I take note]	io annullo [I cancel]
io applaudo [I clap]	io aspetto [I wait]
io avvito [I screw]	io auguro [I wish]
io bastono [I beat]	io contagio [I infect]
io batto [I hit]	io deludo [I disappoint]
io busso [I knock]	io desidero [I desire]
io clicco [I click]	io detesto [I hate]
io colpisco [I hit]	io dimentico [I forget]
io depilo [I shave]	io dispero [I despair]
io digito [I type]	io distinguo [I distinguish]
io dipingo [I paint]	io dubito [I doubt]
io gratto [I scratch]	io esordisco [I begin]
io impugno [I hold]	io fallisco [I fail]
io incateno [I chain up]	io gioisco [I rejoice]
io inchiodo [I nail down]	io giuro [I swear]
io manometto [I tamper]	io gradisco [I appreciate]
io mescolo [I stir]	io illudo [I deceive]
io mungo [I milk]	io immagino [I imagine]
io riscrivo [I rewrite]	io imparo [I learn]
io saluto [I greet]	io memorizzo [I memorise]
io scavo [I dig]	io moltiplico [I multiply]
io sparo [I shoot]	io rallegro [I cheer up]
io spingo [I push]	io sopporto [I bear]
io stiro [I iron]	io spero [I hope]
io strofino [I rub]	io taccio [I fall silent]

Concrete nouns	**Abstract nouns**
il binario [the rail]	il dogma [the gospel]
il cacciavite [the screwdriver]	il fato [the fate]
il cammello [the camel]	il fattore [the factor]
il cancello [the gate]	il pettegolezzo [the gossip]
il cigno [the swan]	il reame [the realm]
il coniglio [the rabbit]	il rimpianto [the regret]
il diamante [the diamond]	il sostentamento [the sustenance]
il divano [the couch]	la bramosia [the yearning]
il leopardo [the leopard]	la congettura [the conjecture]
il pavimento [the floor]	la disonestà [the dishonesty]
il rene [the kidney]	la dote [the talent]
il soffitto [the ceiling]	la falsità [the falsity]
il tappeto [the carpet]	la farsa [the farce]
la collina [the hill]	la magia [the magic]
la colonna [the column]	la malizia [the malice]
la cupola [the dome]	la proporzione [the proportion]

la forchetta [the fork]	la serietà [the seriousness]
la giraffa [the giraffe]	la sottrazione [the subtraction]
la mucca [the cow]	la supplica [the plea]
la nuvola [the cloud]	la supposizione [the supposition]
la pinzetta [the tweezers]	la viltà [the cowardice]
la racchetta [the racket]	l'addizione [the addition]
la tastiera [the keyboard]	l'attinenza [the relevance]
la tigre [the tiger]	l'avversione [the aversion]
la torre [the tower]	l'enigma [the mystery]
l'asfalto [the asphalt]	l'ignoranza [the ignorance]
l'attico [the mansard]	l'illusione [the illusion]
l'edificio [the building]	l'ingratitudine [the ingratitude]
lo scaffale [the shelf]	lo schema [the tactic]
lo spazzolino [the toothbrush]	l'usanza [the custom]

References

Agostini, B., Papeo, L., & Lingnau, A. (2016). *A norming study of 228 high-quality video clips of pantomimes, emblems and meaningless gestures.* Under review.

Agostini, B., Papeo, L., & Lingnau, A. (2016). *The organization of gestures and words in the middle temporal gyrus.* Manuscript in preparation.

Andric, M., Solodkin, A., Buccino, G., Goldin-Meadow, S., Rizzolatti, G., & Small, S.L. (2013). Brain function overlaps when people observe emblems, speech, and grasping. *Neuropsychologia, 51*, 1619-1629.

Barsalou, L.W., Santos, A., Simmons, W.K., & Wilson, C.D. (2008). Language and simulation in conceptual processing. In M. De Vega, A.M. Glenberg, & A.C. Graesser (Eds.), *Symbols, Embodiment, and Meaning*. Oxford: Oxford University Press.

Bauxbaum, L.J., Kyle, K.M., & Menon, R. (2005). On beyond mirror neurons: internal representations subserving imitation and recognition of skilled object-related actions in humans. *Cognitive Brain Research, 25*, 226-239.

Bedny, M., Caramazza, A., Grossman, E., Pascual-Leone, A., & Saxe, R. (2008). Concepts are more than percepts: The case of action verbs. *The Journal of Neuroscience, 28,* 11347–11353.

Bestmann, S., Krakauer, J.W. (2014). The uses and interpretations of the motor-evoked potential for understanding behaviour. *Experimental Brain Research, 233(3)*, 679-89.

Brainard, D.H. (1997). The Psychophysics Toolbox. *Spatial Vision, 10*, 433– 436.

Brass, M., Bekkering, H., & Prinz, W. (2001). Movement observation affects movement execution in a simple response task. *Acta Psychologica, 106*, 3-22.

Bub, D.N., Masson, M.E.J., & Cree, G.S. (2008) Evocation of functional and volumetric gestural knowledge by objects and words. *Cognition, 106(1)*, 27–58.

Buccino, G., Riggio, L., Melli, G., Binkofski, F., Gallese, V., & Rizzolatti, G. (2005). Listening to action-related sentences modulates the activity of the motor system: A combined TMS and behavioral study. *Cognitive Brain Research, 24(3)*, 355–363.

Caramazza, A., Anzellotti, S., Strnad, L., & Lingnau, A. (2014). Embodied cognition and mirror neurons: a critical assessment. *Annual Review of Neuroscience, 37*, 1-15.

Caspers, S., Zilles, K., Laird, A.R., & Eickhoff, S.B. (2010). *NeuroImage, 50*, 1148-1167.

Chatterjee, A. (2008). The neural organization of spatial thought and language. *Seminars in Speech and Language, 29(3)*, 226-238.

Chatterjee, A. (2010). Disembodying cognition. *Language and Cogntion, 2(1)*, 79-116.

Christensen, J.F., Nadal, M., Cela-Conde, C.J., & Gomila, A. (2014). A norming study and library of 203 dance movements. *Perception, 43*, 178-206.

Craighero, L., Bello, A., Fadiga, L., & Rizzolatti, G. (2002). Hand action preparation influences the responses to hand pictures. *Neuropsychologia, 40*, 492-502,

Crepaldi, D., Berlingeri, M., Paulesu, E., & Luzzatti, C. (2011). A place for nouns and a place for verbs? A critical review of neurocognitive data on grammatical-class effect. *Brain and Language, 116*, 33-49.

Dapretto, M., Davies, M.S., Pfeifer, J.H., Scott, A.A., Sigman, M., Bookheimer, S.Y., & Iacoboni, M. (2005). Understanding emotions in others: mirror neuron dysfunction in children with autism spectrum disporders. *Nature Neuroscience, 9*, 28-30.

di Pellegrino, G., Fadiga, L., Fogassi, L., Gallese, V. & Rizzolatti, G. (1992). Understanding motor events: a neurophysiological study. *Experimental Brain Research, 91*, 176-180.

Dick, A.S., Goldin-Meadow, S., Hasson, U., Skipper, J.I., & Small, S.L. (2009). Co-speech gestures influence neural activity in brain regions associated with processing semantic information. *Human Brain Mapping, 30(11)*, 3509-3526.

Efron, D. (1941). *Gesture and Environment*. New York: King's Crown Press.

Efron, D. (1972). *Gesture, Race and Culture*. The Hague: Mouton & Co.

Ekman, P. (1976). Movements with precise meanings. *The Journal of Communication, 26(3)*, 14-26.

Ekman, P. (2004). Emotional and conversational nonverbal signals. *Philosophical Studies Series*, 39-50.

Ekman, P., & Friesen, W.V. (1969). The repertoire of non-verbal behaviour: categories, origins, usage and coding. *Semiotica, 1(1)*, 49-98.

Ekman, P., & Friesen, W.V. (1972). Hand Movements. *The Journal of Communication, 22*, 353-374.

Emmorey, K., Xu, J., Gannon, P., Goldin-Meadow, S., & Braun, A.R. (2010). CNS activation and regional connectivity during pantomime observation: No engagement of the mirror neuron system for deaf signers. *Neuroimage 49*, 994–1005.

Fadiga, L., & Rizzolati, G. (1995). Motor facilitation during action observation: a magnetic simulation study. *Journal of Neurophysiology, 73*, 2608-2611.

Fadiga, L., & Rizzolati, G. (1995). Motor facilitation during action observation: a magnetic simulation study. *Journal of Neurophysiology, 73*, 2608-2611.

Fadiga, L., Craighero, L., Buccino, G., & Rizzolatti, G. (2002). Speech listening specifically modulates the excitability of tongue muscles: a TMS study. *European Journal of Neuroscience, 15(2)*, 399-402.

Fischer, M.H., & Zwaan, Q. (2008). Embodied language: a review of the role of the motor system in language comprehension. *The Quarterly Journal of Experimental Psychology, 61 (60)*, 825-850.

Fogassi, L., Ferrari, P.F., Gesierich, B., Rozzi, S., Chersi, F., & Rizzolatti, G. (2005). Parietal lobe: from action organization to intention understanding. *Science, 308*, 662-667.

Gallese, V., & Goldman, A. (1998). Mirror neurons and the simulation theory of mind-reading. *Trends in Cognitive Sciences, 2*, 493–501.

Gallese, V., Fadiga, L., Fogassi, L., & Rizzolatti, G. (1996). Action recognition in the premotor cortex. *Brain, 119*, 593-609.

Gallese, V., Keysers, C., & Rizzolatti, G. (2004). A unifying view of the basis of social cognition. *Trends in Cognitive Neuroscience, 8(9)*, 396-403.

Genovese, C.R., Lazar, N.A., & Nichols, T. (2002). Thresholding of statistical maps in functional neuroimaging using the false discovery rate. *NeuroImage, 15(4)*, 870-878.

Glenber, A.M., Sato, M., Cattaneo, L., Riggio, L. Palumbo, D., & Buccino, G. (2008). Processing abstract language modulates motor system activity. *The Quarterly Journal of Experimental Psychology, 61(6)*, 905-919.

Glenberg, A.M., & Kaschak, M.P. (2002). Grounding language in action. *Psychonomica Bulleting & Review, 9(3)*, 558-565.

Glover, S., Rosenbaum, D., Graham, J. & Dixon, P. (2004) Grasping the meaning of words. *Experimental Brain Research, 154(1)*, 103–108.

Groppa, S., Oliviero, A., Eisen, A., Quartarone, A., Cohen, L.G., Mall, V., Kaelin-Lang, A., Mima, T., Rossi, S., Thickbroom, G.W., Rossini, P.M., Ziemann, U., Valls-Solé, J., & Siebner, H.R. (2012). A practical guide to diagnostic transcranial magnetic stimulation: report of an IFCN committee. *Clinical Neurophysiology, 123*, 858–882.

Gullberg, M. (2006). Some reasons for studying gesture and second language acquisition (Hommage à Adam Kendon). *International Review of Applied Linguistic in Language Teaching, 44(2)*, 103-124.

Hadjikhani, N., Joseph, R.M., Snyder, J., & Tager-Flusberg, H. (2006). Anatomical differences in the mirror neuron system and social cognition network in autism. *Cerebral Cortex, 16(9)*, 1276-1282.

Hamilton, A.F. de C., & Grafton, S.T. (2008). Action outcomes are represented in human inferior frontoparietal cortex. *Cerebral Cortex, 18*, 1160-1168.

Harris, J.A., Clifford, C.W., & Miniussi, C. (2008). The functional effect of transcranial magnetic stimulation: signal suppression or neural noise generation? *Journal of Cognitive Neuroscience, 20*, 734–740.

Haxby, J. V. (2012). Multivariate pattern analysis of fMRI: the early beginnings. *NeuroImage, 62(2)*, 852–5.

Hickok, G. (2009). Eight problems for the mirror neuron theory of action understanding in monkeys and humans. *Journal of Cognitive Neuroscience, 21(7)*, 1229-1243.

Kable, J.W., Kan, I.P., Wilson, A., Thompson-Schill, S.L., & Chatterjee, A. (2005). Conceptual representations of action in the lateral temporal cortex. *Journal of Cognitive Neuroscience, 17(12)*, 1855-1870.

Kable, J.W., Lease-Spellmeyer, J., & Chatterjee, A. (2002). Neural substrates of action event knowledge. *Journal of Cognitive Neuroscience, 14*, 795–804.

Kalénine S., Buxbaum, L.J., & Coslett, H.B. (2010). Critical brain regions for action recognition: lesion symptom mapping in left hemisphere stroke. *Brain, 133(11)*, 3269-3280.

Kammer, T., Beck. S., Thielscher, A., Laubis-Herrmann, U., & Topka, H. (2001). Motor thresholds in humans: a transcranial magnetic stimulation study comparing different pulse waveforms, current directions and stimulator types. *Clinical Neurophysiology,112*, 250–258.

Kendon, A. (1988). How gestures can become like words. In F. Poyatos (Ed.), *Cross-cultural perspectives in nonverbal communication* (pp. 207-227). The Hague: Mounton and Co.

Kendon, A. (1992). Some recent work from Italy on “Quotable Gestures (Emblems)”. *Journal of Linguistic Anthropology, 2(1)*, 92-108.

Keyser, C., & Gazzola, V. (2009). Unifyig social cognition. In J.A. Pineda (Ed.) *Mirror neuron system* (pp. 1-35). New York. Humana Press.

Kiefer, M., & Pulvermüller, F. (2012). Conceptual representations in mind and brain: Theoretical developments, current evidence and future directions. *Cortex, 48(7)*, 805-825.

Kriegeskorte, N., Goebel, R., & Bandettini, P. (2006). Information-based functional brain mapping. *PNAS, 103*, 3863-3868.

Leshinskaya, A., & Caramazza, A. (2016). For a cognitive neuroscience of concepts: moving beyond the grounding issue. *Psychonomic Bulletin & Review*. doi:10.3758/s13423-015-0870-z.

Lewis, J.W. (2006). Cortical networksrelated to human use of tools. *The Neuroscientist, 12*, 211-231.

Lingnau, A., & Downing, P.E. (2015). The lateral occipitotemporal cortex in action. *Trends in Cognitive Neuroscience, 19(5)*, 268-277.

Lingnau, A., & Petris, S., (2012). Action understanding within and outside the motor system: the role of task difficulty. *Cerebral Cortex*. doi:10.1093/cercor/bhs112.

Lotze, M., Heymans, U., Birbaumer, N., Veit, R., Erb, M., Flor, H., & Halsband, U. (2006). Differential cerebral activation during observation of expressive gestures and motor acts. *Neuropsychologia, 44*, 1787-1795.

Macmillan, N.A., & Creelman, C.D. (2005). *Detection theory: a user’s guide* (2nd ed.). Mahwah, New Jersey: Lawrence erlbaum associates.

Maeda, F., Kleiner-Fisman, G., & Pascual-Leone, A. (2002). Motor facilitation while observing hand actions: specificity of the effect and role of observer’s orientation. *Journal of Neurophysiology, 87*, 1329-1335.

Mahon, B.Z., & Caramazza, A. (2008). A critical look at the embodied cognition hypothesis and a new proposal for grounding conceptual content. *Journal of Psychology, 102*, 59-70.

Martin, A. (2007). The representation of object concepts in the brain. *Annual Review of Psychology, 58*, 25-45.

Martin, A., Haxby, J.V., Lalonde, F.M., Wiggs, C.L., & Ungerleider, L.G. (1995). Discrete cortical regions associated with knowledge of color and knowledge of action. *Science, 270*, 102-105.

McNeill, A. (2005). *Gestures and thoughts*. Chicago: University of Chicago Press.

McNeill, D. (1992). *Hand and mind: what gestures reveal about thought*. Chicago: University of Chicago Press.

Meteyard, L., Rodriguez-Cuadrado, S., Bahrami, B., & Vigliocco, G. (2012). Coming of age: a review of embodiment and the neuroscience of semantics. *Cortex, 48(7)*, 788-804.

Molnar-Szakacs, I., Wu, A.D., Robles, F.J., & Iacoboni, M. (2007). Do you see what I mean? Corticospianl excitability during observation of cultural-specific gestures. *PLoS ONE, 7,* doi: http://dx.doi.org/10.1371/journal.pone.0000626.

Möttönen, R., Farmer, H., & Watkins, K.E. (2016). Neural basis of understanding communicative actions: changes associated with knowing the actor's intention and the meanings of the actions. *Neuropsychologia, 81*, 230-237.

Mur, M., Bandettini, P.A., & Kriegeskorte, N. (2009). Revealing representational content with pattern-information fMRI-and introductory guide. *Social Cognitive and Affective Neuroscience*. doi: 10.1093/scan/nsn044.

Negri, G.A.L., Rumiati, R.I., Zadini, A., Ukmar, M., Mahon, B.Z., & Caramazza, A. (2007). What is the role of motor simulation in action and object recognition? Evidence from apraxia. *Cognitive Neuropsychology, 24(8)*, 795-816.

Noppeney, U., Josephs, O., Kiebel, S., Friston, K.J., & Price, C.J. (2005). Action selectivity in parietal and temporal cortex. *Cognitive Brain Research, 25*, 641-649.

Norman, K.A., Polyn, S.M., Detre, G.J., & Haxby, J.V. (2006). Beyond mind-reading: multi-voxel pattern analysis of fMRI data. *Trends in Cognitive Neuroscience, 10(9)*, 424-430.

Oberman, L.M., Hubbard, E.M., McCleery, J.P., Altschuler, E.L., Ramachandran, V.S., & Pineda, J.A. (2005). EEG evidence for mirror neuron dysfunction in autism spectrum disorders. *Cognitive Brain Research, 24(2)*, 190-198.

Oliveri, M., Finocchiaro, C., Shapiro, K., Gangitano, M., Caramazza, A., & Pascual-Leone A. (2004). All talk and no action: a transcranial magnetic stimulation study of motor cortex activation during action word production. *Journal of Cognitive Neuroscience, 16*, 347–381.

Oosterhof, N.N, Tipper, S.P., & Downing P.E. (2013). Crossmodal and action-specific: neuroimaging the human mirror neuron system. *Trends in Cognitive Neuroscience, 17(7)*, 311-318.

Oosterhof, N.N., Tipper, S.P., & Downing, P.E. (2012). Viewpoint (In)dependence of action representations: an MVPA study. *Journal of Cognitive Neuroscience, 24(4)*, 975-989.

Oosterhof, N.N., Wiggett, A.J., Diedrichsen, J., Tipper, S.P., & Downing, P.E. (2010). Surface-based information mapping reveals crossmodal vision-action representations in human parietal and occipitotemporal cortex. *Journal of Neurophysiology, 104(2)*, 1077-1089.

Papeo, L., & Lingnau, A. (2015). First-person and third-person verbs in visual motion-perception regions. *Brain and Language, 141*, 135-141.

Papeo, L., & Rumiati, R.I. (2013). Lexical and gestural symbols in left-damaged patients. *Cortex, 49*, 1668-1678.

Papeo, L., Cecchetto, C., Mazzon, G., Granello, G., Cattaruzza, T., Verriello, L., Eleopra, R., & Rumiati, R.I. (2015). The processing of actions and action-words in amyotrophic lateral sclerosis patients. *Cortex, 64,* 136-147.

Papeo, L., Lingnau, A., Agosta, S., Pascual-Leone, A., Battelli, L., & Caramazza, A. (2014). The origin of word-related motor activity. *Cerebral Cortex*. doi:10.1093/cercor/bht423.

Papeo, L., Negri, G.A.L., Zadini, A., & Rumiati, R.I. (2010). Action performance and action-word understanding: evidence of double dissociations in left-damaged patients. *Cognitive Neuropsychology, 27(5)*, 428-461.

Papeo, L., Pasucal-Leone, A., & Caramazza, A. (2013). Disrupting the brain to validate hypotheses on the neurobiology of language. *Frontiers in Human Neuroscience, 7(148)*.

Papeo, L., Vallesi, A., Isaja, A., & Rumiati, R.I. (2009). Effect of TMS on different stages of motor and non-motor verb processing in the primary motor cortex. *PLoS ONE, 4 (2)*, 1-11.

Payrató, L. (1993). A pragmatic view on autonomous gestures: a first repertoire of Catalan emblems. *Journal of Pragmatics, 20*, 193-216.

Pazzaglia, M., Smania, N., Corato, E., & Aglioti, S.M. (2008). Neural underpinnings of gesture discrimination in patients with limb apraxia. *The Journal of Neuroscience, 28(12)*, 3030-3041.

Peelen, M.V., Romagno, D., & Caramazza, A. (2012). Independent representations of verbs and actions in left lateral temporal cortex. *The Journal of Cognitive Neuroscience, 24*, 2096–2107.

Perini, F., Cattaneo, L., Carrasco, M., & Schwarzbach, J.V. (2012). Occipital transcranial magnetic stimulation has an activity-dependent suppressive effect. *The Journal of Neuroscience, 32(36)*, 12361-12365.

Poggi, I., & Zomparelli, M. (1987). Lessico e grammatical nei gesti e nelle parole. In I. Poggi (Ed.), *Le parole nella testa: guida a un'educazione linguistica cognitivista* (pp. 329-338). Bologna: Il Mulino.

Pulvermüller, F., & Fadica, L. (2010). Active perception: sensorimotor circuits as a cortical basis for language. *Nature Reviews Neuroscience, 11*, 351-360.

Pulvermüller, F., Hauk, O., & Nikulin, V.V. (2005) Functional links between motor and language systems. *European Journal of Neuroscience 2005, 21*, 793–797.

Rizzolatti, G., & Craighero, L. (2004). The mirror-neuron system. *Annual Review of Neuroscience, 27(1)*, 169-192.

Rizzolatti, G., & Sinigaglia, C. (2010). The functional role of the parieto-frontal mirror circuit: interpretations and misinterpretations. *Nature Review Neuroscience, 11*, 264-274.

Rizzolatti, G., Fogassi, L., & Gallese, V. (2001). Neurophysiological mechanisms underlying the understanding and imitation of action. *Nature Reviews Neuroscience, 2*, 661-670.

Rizzolatti, G., Fadiga, L., Gallese, V., & Fogassi, L. (1996). Premotor cortex and the recognition of motor actions. *Cognitive Brain Research, 3(2)*, 131-141.

Robertson, E.M, Theoret, H., & Pascual-Leone, A. (2003). Studies in cognition: the problems solved and created by transcranial magnetic stimulation. *Journal of Cognitive Neuroscience, 15(7)*, 948-960.

Rossi, S., Hallett, M., Rossini, P.M., Pascual-Leone, A., & The Safety of TMS Consensus Group (2009). Safety, ethical consideration, and application guidelines for the use of transcranial magnetic stimulation in clinical practice and research. *Clincal Neuropsychology, 120*, 2008-2039.

Rossini P.M, Burke, D., Chen, R., Cohen, L.G., Daskalakis, Z., Di Iorio, R., Di Lazzaro, V., Ferreri, F., Fitzgerald, P.B., George, M.S., Hallett, M., Lefaucheur, J.P., Langguth, B., Matsumoto, H., Miniussi, C., Nitsche, M.A., Pascual-Leone, A., Paulus, W., Rossi, S., Rothwell, J.C., Siebner, H.R., Ugawa, Y., Walsh, V., & Ziemann, U. (2015). Non-invasive electrical and magnetic stimulation of the brain, spinal cord, roots and peripheral nerves: Basic principles and procedures for routine clinical and research application. An updated report from an I.F.C.N. Committee. *Clinical Neuropsychology, 126*, 1071-1107.

Rumiati, R.I., Zanini, S., Vorano, L., & Shallice, T. (2001). A form of ideational apraxia as a selective deficit of contention scheduling. *Cognitive Neuropsychology, 18*, 617-642.

Sack, A.T., & Linden, D.E.J. (2003). Combining transcranial magnetic stimulation and functional imaging in cognitive brain research: possibilities and limitations. *Brain Research Reviews, 43*, 41-56.

Sauve, W.M., & Crowther, L.J.. The science of transcranial magnetic stimulation. *Psychiatric Annals, 44*, 279–283.

Schulte-Rüther, M., Markowitsch, H.J., Fink, G.R., & Piefke, M. (2007). Mirror neuron and theory of mind mechanisms involved in face-to-face interactions: a functional magnetic resonance imaging approach to empathy. *Journal of Cognitive Neuroscience, 19(9)*, 134-1372.

Schwarzbach, J. (2011). A simple framework (ASF) for behavioral and neuroimaging experiments based on the psychophysics toolbox for MATLAB. *Behavioral Research Methods*, *43*, 1194 –1201.

Smith, S. M., & Nichols, T. (2009). Threshold-free cluster enhancement: Addressing problems of smoothing, threshold dependence and localisation in cluster inference. *Neuroimage, 44*, 83–98.

Spaulding, S. (2012). Mirror neurons are not evidence for the simulation theory. *Synthese, 189*, 515-534.

Strafella, A., & Paus, Tomáš. (2000). Modulation of cortical excitability during action observation: a transcranial magnetic stimulation study. *Neuroreport, 11(10)*, 2289-2292.

Szekely, A., Jacobsen, T., D'Amico, S., Devescovi, A., Andonova, E., Herron, D., Lu, C. C., Pechmann, T., Pleh, C., Wicha, N., Federmeier, K., Gerdjikova, I., Gutierrez, G., Hung, D., Hsu, J., Iyer, G., Kohnert, K., Mehotcheva, T., Orozco-Figueroa, A., Tzeng, A., Tzeng, O., Arevalo, A., Vargha, A., Butler, A. C., Buffington, R., & Bates, E. (2004). A new on-line resource for psycholinguistic studies. *Journal of Memory and Language, 51(2)*, 247-250.

Tahran, L.Y., Watson, C.E., & Buxbaum, L.J. (2015). Shared and distinct neuroanatomic regions critical for tool-related action production and recognition: evidence from 131 left-hemisphere stroke patients. *Journal of Cognitive Neuroscience, 27(12)*, 2491-2511.

Tettamanti, M., Buccino, G., Saccuman, M., Gallese, V., Danna, M., Scifo, P., Fazio, F., Rizzolatti, G., Cappa, S.F., & Perani, D. (2005). Listening to action-related sentences activates fronto-parietal motor circuits. *Journal of Cognitive Neuroscience, 17(2)*, 273–281.

Tremblay, P., Sato, M., & Small, S.L. (2012). TMS-induced modulation of action sentence priming in the ventral premotor cortex. *Neuropsychologia, 50*, 319-326.

Tucciarelli, R., Turella, L., Oosterhof, N.N., Weisz, N., & Lingnau, A. (2015). MEG multivariate analysis reveals early abstract action representations in the lateral occipitotemporal cortex. *The Journal of Neuroscience, 35(49)*, 16034-16045.

Tucker, M., & Ellis, R. (2004) Action priming by briefly presented objects. *Acta Psychologica, 116(2)*, 185–203.

Tulving, E. (1972). Episodic and semantic memory. In E. Tulving, & W. Donaldson (Eds.), *Organization of memory* (pp.381-402). New York: Academic Press.

Turella, L., Wurm, M.F., Tucciarelli, R., & Lingnau, A. (2013). Expertise in action observation: recent neuroimaging findings and future perspectives. *Frontiers in human neuroscience.* doi: 10.3389/fnhum.2013.00637.

Umiltà, M.A., Kohler, E., Gallese, V., Fogassi, L., Fadiga, L., Keysers, C., & Rizzolatti, G. (2001). I know what you are doing: a neurophysiological study. *Neuron, 31,* 155-165.

Urgesi, C., Candidi, M., & Avenanti, A. (2014). Neuroanatomical substrates of action perception and understanding: an anatomic likelihood estimation meta-analysis of lesion-symptom mappin studies in brain injured patients. *Frontiers in human neuroscience, 8,* doi: 10.3389/fnhum.2014.00344.

Vannuscorps, G., & Caramazza, A. (2016). Typical action perception and interpretation without motor simulation. *PNAS, 113(1)*. 86-91.

Villareal, M., Fridman, E., Amengual, A., Falasco, G., Roldan Gerscovich, E., Ulloa, E.R., & Leiguarda, R.C. (2008). The neural substrate of gesture recognition. *Neuropsychologia, 46*, 2371-2382.

Visser, M., Jefferies, E., Embleton, K.V., & Lambon Ralph, M.A. (2012). Both the middle Temporal gyrus and the ventral anterior temporal area are crucial for multimodal semantic processing: distortion-corrected fMRI evidence for a double gradient of information convergence in the temporal lobes. *Journal of Cognitive Neuroscience, 24(8)*, 1766–1778.

Walsh, V., & Rushworth, M. (1999). A primer of magnetic stimulation as a tool for neuropsychology. *Neuropsychologia, 37*, 125-135.

Watking, K.E., Strafella, A.P., & Paus, T. (2003). Seeing and hearing speech excites the motor system involved in speech production. *Neuropsychologia, 41(8)*, 989-994.

Watson, C.E., Cardillo, E.R., Ianni, G.R., & Chatterjee, A. (2013). Action concepts in the brain: an activation likelihood estimation meta-analysis. *The Journal of Cognitive Neuroscience, 25(8)*, 1191-1205.

Weiner, K.S., & Grill-Spector, K. (2013). Neural representations of faces and limbs neighbour in human high-level visual cortex: evidence for a new organization principle. *Psychological Research, 77*, 74-97.

Whitney, C., Kirk, M., O'Sullivan, J., Lambon Ralph, M.A., & Jefferies, E. (2010). The neural organization of semantic control: TMS evidences for a distributed network in left inferior frontal and posterior middle temporal gyrus. *Cerebral Cortex*, doi:10.1093/cercor/bhq180.

Willems, R.M., & Hagoort, P. (2007). Neural evidence for the interplay between language, gestures, and action: a review. *Brain and Language, 101*, 278-289.

Willms, J.L., Shapiro, K.A., Peelen, M.V., Pajtas, P.E., Costa, A., Moo, L.R., & Caramazza, A. (2011). Language-invariant verb processing regions in Spanish-English bilinguals. *NeuroImage, 57(1)*, 251-261.

Witt, J.K., Kemmerer, D., Linkenauger, S.A., & Culham, J. (2010). A functional role for motor simulation in identifying tools. *Psychological Science, 21(9)*, 1215-1219.

Wurm, M.F., & Lingnau, A. (2015). Decoding actions at different levels of abstraction. *The journal of Neuroscience, 35(20)*, 7727-7735.

Wurm, M.F., Ariani, G., Greenlee, M.W., & Lingnau, A. (2015). Decoding concrete and abstract action representations during explicit and implicit conceptual processing. *Cerebral Cortex*. doi: 10.1093/cercor/bhv169.

Xu, J., Gannon, P.J., Emmorey, K., Smith, J.F., & Braun, A. (2009). Symbolic gestures and spoken language are processed by a common neural system. *PNAS, 106(49)*, 20664-20669.

Zaitsev, M., Hennig, J., & Speck, O. (2004). Point spread function mapping with parallel imaging techniques and high acceleration factors: fast, robust, and flexible method for echo-planar imaging distortion correction. *Magnetic Resonance in Medicine, 52*, 1156 –1166.

Zeng, H., & Constable, R.T. (2002). Image distortion correction in EPI: comparison of field mapping with point spread function mapping. *Magnetic Resonance in Medicine, 48*, 137–146.