# Brains in interaction

Schippers, Marleen Bernadette

# Brains in Interaction

Marleen Bernadette Schippers

rijksuniversiteit
groningen

# Brains in Interaction

Proefschrift

ter verkrijging van het doctoraat in de
Medische Wetenschappen
aan de Rijksuniversiteit Groningen
op gezag van de
Rector Magnificus, dr. E. Sterken,
in het openbaar te verdedigen op
woensdag 9 maart 2011
om 16:15 uur

door

Marleen Bernadette Schippers
geboren op 2 mei 1979
te Nijmegen

Promotor: Prof. dr. C. Keysers

Beoordelingscommissie: Prof. dr. U. Hasson (Princeton University)
Prof. dr. V. Lamme (Universiteit van Amsterdam)
Prof. dr. R. Malach (Weizmann Institute of Science)

ISBN printed version: 978-90-367-4699-1
ISBN digital version: 978-90-367-4700-4

# CONTENTS

# LIST OF FIGURES

# LIST OF TABLES

# 1

# INTRODUCTION

People involved in a social interaction can be thought of as being temporarily connected. This connection starts, for example, in the motor system of person A, which leads to observable behavior that is perceived and interpreted by the primary sensory and higher cortices of person B. This in turn can lead to activity in the motor system of person B, which leads to observable behavior that can be perceived and interpreted by person A, and so on. This connection is dynamic and dependent on events in the interaction, such as changes in direction of eye gaze, the words being said and the gestures being made. In this thesis, we investigate such a communicative connection on a neural level. This research is inspired both by theories explaining how humans understand each other and by the methodological advancements in connectivity analyses.

## 1.1 MIRRORING AND SIMULATION THEORY

Over the past few decades it has become clear that perception and action are closely linked in the human brain. They do not function as two independent modules, but instead represent a continuum. This relationship becomes clear when you, for example, try to execute an action while simultaneously observing someone else doing an action. The action you observe influences the action you want to do. If the observed action is similar to the one you want to execute, it can speed up reaction time. When the two actions are dissimilar, it can, however, slow down reaction time (Kilner et al., 2003; Brass et al., 2000; Craighero et al., 2002). Conversely, the way we perceive an action can be interfered by the concurrent planning of an action (Müsseler and Hommel, 1997; Hamilton et al., 2004). These bidirectional influences between perception and action led to the idea that these two processes are represented in a similar code (common coding theory, Prinz, 1990). This in turn inspired the core idea of simulation theory: we understand other people by transforming their actions into our own motor-representations of that action and thereby we internally simulate doing that action (Goldman, 1992; Gibson, 1986; Gallese, 2003).

The idea of simulation received neurophysiological support with the discovery of mirror neurons. Mirror neurons were originally discovered in area F5 (ventral premotor cortex) of the Macaque monkey and have the special property of firing both when executing and perceiving (viewing or listening to) a goal-directed hand

action (Fogassi et al., 2005; Gallese et al., 1996; Ferrari et al., 2003; Fujii et al., 2007; Keysers et al., 2003; Kohler et al., 2002; Umiltà et al., 2001; Rizzolatti et al., 1996). This discovery constitutes an important step in neuroscience, because it confirmed the idea that perception and action are inextricably linked on a neural level. Shortly after their discovery, mirror neurons were regarded as the neural basis of the simulation theory of action understanding (Gallese and Goldman, 1998).

In the *human* brain, there is strong evidence that five regions contain mirror neurons: the ventral and dorsal premotor cortex (Kilner et al., 2009; Lingnau et al., 2009), the supplementary motor area (Mukamel et al., 2010), the inferior parietal lobe (Chong et al., 2008), and the temporal lobe (Mukamel et al., 2010). Currently, the two methods that provide the strongest evidence for mirror neurons in humans are (cross-modal) repetition suppression paradigms (Kilner et al., 2009; Lingnau et al., 2009; Chong et al., 2008) and measurement of extracellular activity (Mukamel et al., 2010). These findings were preceded by experiments showing that perception of an action activates brain regions known in the literature or measured in the same participants to be involved in generating the same actions (Buccino et al., 2001; Grafton et al., 1996; Grèzes et al., 1998; Grèzes and Decety, 2001; Grèzes et al., 2003; Nishitani and Hari, 2000, 2002; Perani et al., 2001; Gazzola et al., 2007b,a; Gazzola and Keysers, 2008). All these studies culminated in the idea of a putative Mirror Neuron System in the human brain (Keysers and Gazzola, 2009) consisting of the ventral and dorsal premotor cortex, the inferior parietal lobe, and the middle temporal lobe (see Figure 1.1). Furthermore, several other brain areas show an overlap between the experience and observation of for example emotions (Bastiaansen et al., 2009; Wicker et al., 2003; Jabbi et al., 2007), and sensations (Singer et al., 2004; Keysers et al., 2004, 2010; Blakemore et al., 2005). In particular, there is accumulating evidence that BA 2, involved in sensing how our own body moves and in action observation and execution, represents a 'somatosensory' branch of the pMNS (Keysers et al., 2010).

Mirror neurons were originally conceptualized as *motor* neurons in the macaque monkey's brain that responded selectively to perceptual input. While it becomes more and more clear that mirroring is a more general mechanism of the human brain, it has been suggested to extend the definition of a mirror neuron to "any neuron involved in the execution of a motor action that shows significant vicarious activity to the observation of corresponding actions performed by others" (Keysers and Gazzola, 2009).

Simulation theory and the idea of mirroring have received much scientific attention and have led to the idea that it is one of the basic principles of human interaction. It has been implicated in many social skills, such as imitation (Iacoboni et al., 1999; Koski et al., 2003; Hurley, 2008), empathy (Fujii et al., 2007; Fogassi et al., 2005; Gazzola et al., 2006; Gazzola and Keysers, 2008; Keysers et al., 2003; Kohler et al., 2002; Umiltà et al., 2001), mind-reading (Gallese, 2003; Gallese and Goldman, 1998), facial expressions

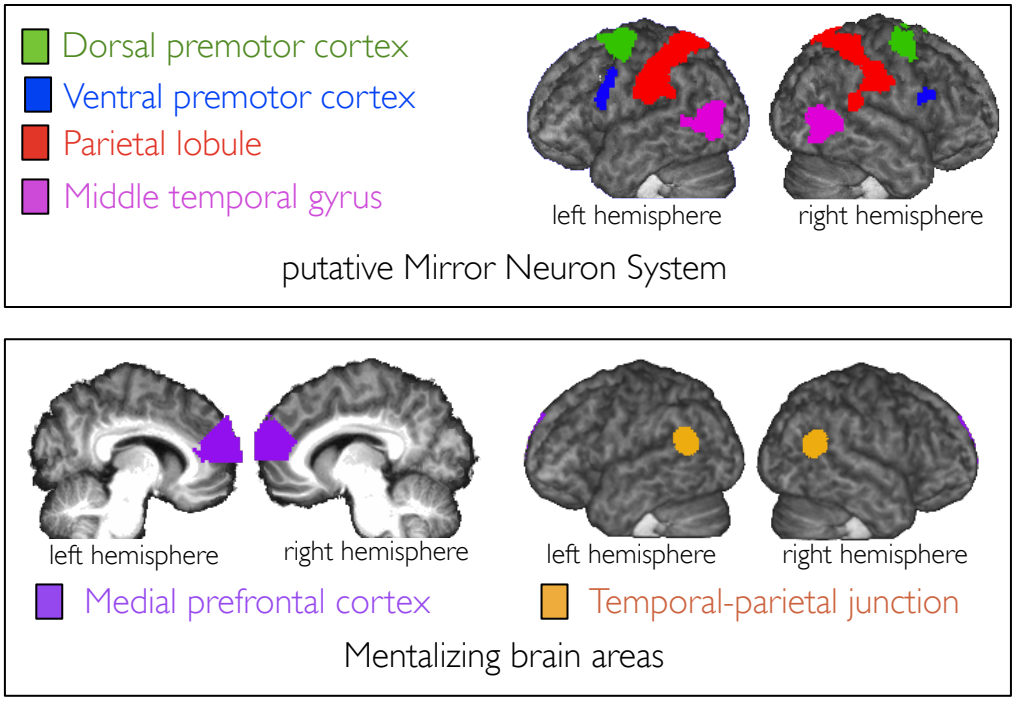



Figure 1.1: Areas constituting the putative Mirror Neuron System and the mentalizing system in the human brain.

(van der Gaag et al., 2007), joint action (Kokal et al., 2009; Newman-Norlund et al., 2008, 2007) and language (Rizzolatti and Arbib, 1998; Arbib, 2008).

## 1.2 MENTALIZING

Besides simulation as a mechanism to understand others, people have the capacity to think about and understand others on a more reflective level. This can be illustrated with a typical scene from a soap opera, such as the Bold and the Beautiful: Taylor and Ridge are about to get married. Without Taylor knowing this, Brooke is about to confess that she is pregnant with Ridge's baby. She hopes that by revealing this she can prevent this marriage from happening. To understand and appreciate such a situation, we have to be able to track what all characters involved know, what they do not know and predict what they will think when they will find out. This ability to attribute mental states, beliefs and desires to others is called having a Theory of Mind (Premack and Woodruff, 1978; Wimmer and Perner, 1983). The process of reasoning about other people's mental states is often referred to as mentalizing (Frith and Frith, 1999).

Neuroimaging studies have identified several brain areas that are associated with

mentalizing, such as the posterior superior temporal sulcus (pSTS), the temporal poles, the temporal-parietal junction (TPJ), precuneus and the medial prefrontal cortex (mPFC). The two most important of these seem to be the mPFC and the TPJ (see Figure 1.1) and together these are referred to as the mentalizing system (Overwalle and Baetens, 2009; Carrington and Bailey, 2009).

The mPFC is involved in a multitude of social cognitive tasks (Amodio and Frith, 2006), such as mentalizing (Fletcher et al., 1995; Gallagher et al., 2000; Vogeley et al., 2001; Mitchell et al., 2005; Grèzes et al., 2004), person perception (Mitchell et al., 2002; Bonda et al., 1996), but also self-reflection (Ochsner and Gross, 2004; van der Meer et al., 2010), emotional processing (Ochsner and Gross, 2004), intention attribution (Brunet et al., 2000; Castelli et al., 2000; Ciaramidaro et al., 2007; de Lange et al., 2008; Kampe et al., 2003), and autobiographical memory (Spreng et al., 2009). Damage to this area commonly leads to poor performance on mentalizing tasks (Rowe et al., 2001; Stuss et al., 2001; Gregory et al., 2002; Adenzato et al., 2010). One study, however, failed to find a deficit in ToM reasoning in a patient with medial frontal damage (Bird et al., 2004), suggesting that the mPFC is not the only area involved in mentalizing.

The TPJ (in particular in the right hemisphere, rTPJ) has also been associated with many different functions related to mentalizing (Saxe and Kanwisher, 2003; Apperly et al., 2004; Frith and Frith, 2003; Gallagher et al., 2000; Samson et al., 2004; Saxe, 2006; Sommer et al., 2007), intention attribution (Ciaramidaro et al., 2007; Noordzij et al., 2009), perspective-taking (Ruby and Decety, 2003), and sense of agency (Decety and Lamm, 2007; Blakemore and Frith, 2003). An unresolved question remains whether activity in the TPJ is unique to mentalizing processes or whether it can be attributed to lower-level computational processes such as attention re-orienting (Corbetta et al., 2008; Decety and Lamm, 2007; Mitchell, 2008; Young et al., 2010).

## 1.3 MIRRORING, MENTALIZING, AND COMMUNICATION

There has been a lot of debate about which of the two theories (theory of mind or simulation theory) can explain most of human interpersonal understanding (Gallagher, 2007; Hickok, 2009; Saxe and Wexler, 2005). Many now believe that mentalizing is a separate mechanism from the more basic, low-level motor simulation (Uddin et al., 2007; de Lange et al., 2008; Brass et al., 2007; Overwalle and Baetens, 2009). Thus the debate has shifted to the more fruitful question of how these two systems work together to achieve a full understanding of other people (Keysers and Gazzola, 2007). The charades experiment that is central to this thesis was set up in part to investigate this issue.

Charades is a social communicative game in which one participant has to use gestures, rather than verbal communication, to convey a concept to another participant.

This type of gestural communication essentially boils down to using a sequence of hand-actions to influence the mental state of another person. In this way, it provides a good means to study the involvement of both pMNS and mentalizing areas. Furthermore, gestural communication seems to be a primitive form of more sophisticated verbal communication and may have played a crucial role in language evolution (Rizzolatti and Arbib, 1998; Arbib, 2008; Gentilucci and Corballis, 2006). Interestingly, in the human brain, the production and perception of language seem to recruit similar brain regions, which also overlap partly with the premotor node of the pMNS. Gestural communication is therefore an interesting link between the mirror system as studied in non-human animals (such as the Macaque monkey and the swamp sparrow (Gallese et al., 1996; Rizzolatti et al., 1996; Prather et al., 2008)) and the unique language skills of humans.

## 1.4 THE CHARADES EXPERIMENT

In the charades experiment couples of participants played the game of charades by taking turns gesturing and guessing concepts in the MR-scanner. Participants were presented with a word on the screen (either an action or an object, for example nutcracker, knitting, shaving) and were instructed to convey the meaning of this word by gestures. Their gestures were recorded on video and presented to their partners, who went into the MR-scanner to guess what their partner had gestured. The partners had to push a button when they thought they knew what word was being portrayed. After a while, the two partners switched roles, with the former gesturer becoming the guesser and vice-versa. Both partners guessed and gestured 14 words in total. On a different day, they returned for the control condition in which they observed exactly the same movies of their gesturing partner, but now with the instruction to try not to interpret the gestures. In this way, we recorded the brain activity of both partners in the social interaction and during a control condition with the same gesture-recordings, but with a different instruction.

fMRI experiments are intrinsically rendered artificial by the fact that participants have to lie down on an MR-bed with their head strapped within a head-coil. This artificiality is worsened by the fact that stimuli have to be repeated more than ten times to result in a reliable fMRI-signal. Therefore, we wanted to create an experimental situation that is as close to the real world as possible. We did this by letting participants play the actual game of charades together. Both partners were present while scanning and they reported enjoying the game. They indicated to have tried their best to gesture the concepts as comprehensible as possible. This might not have been the case if we would have had the participants come on different days, for example, or if we would have used one set of gestured concepts and shown these to

different guessers. Furthermore, we did not impose a strict temporal structure on the timing of the experiment. Gestures could be made freely during a period of at least 50s dedicated to each concept and hardly any constraints: no restrictions on the amount of arm movements or eye gaze, no fixed number of repetitions for one word and no separation in planning and execution phase.

The charades experiment allowed us to investigate several research questions. First, we wanted to analyze the involvement of mirroring and mentalizing areas in gesturing and guessing concepts. Given that the pMNS maps goal-directed hand actions onto the motor programs for the execution of those actions, we hypothesized that parts of the areas involved in generating communicative actions would also become activated during the observation of these actions. Furthermore, we reasoned that activity during gesture production may reflect a theory of mind of how the partner might interpret the gestures, and activity during gesture interpretation may reflect a theory of mind of what the partner might have meant while generating the gestures. Thus, we hypothesized that pMNS areas would be involved in all three experimental conditions (gesturing, guessing and passive observation), while mentalizing areas would show activity during gesturing and guessing, but not during passive observation.

Second, besides studying activity patterns within either the gesturer or guesser's brain, we wanted to study the 'connectivity' between the brains of gesturer and guesser. Of course, no direct neural connectivity between two different brains exists. Here, 'connectivity' refers to the fact that the brain activity of the gesturer is linked to the brain activity of the guesser through an external chain of events: activity in the brain of the gesturer produces observable gestures which are video recorded and later observed by the guesser which again leads to brain activity. This research question was based on the hypothesis put forward by simulation theory that mirror neurons of an observer would resonate[1] with the actions of others. Our hypothesis was that pMNS areas of the gesturer would have a measurable influence on pMNS areas of the guesser.

Third, we wanted to investigate the influences from and to pMNS areas within the brain of the guesser. The inclusion of relatively long blocks of gesture observation in the experiment allowed us to compare hypotheses generated by different models of the pMNS. A classic account of the pMNS during action observation could be described as strictly feed-forward. This account is inherent to many descriptions of the anatomy of the pMNS, and applies the concept of an inverted forward model to action observation (Kilner et al., 2007). A forward model, derived from motor control theory sends a copy of motor commands to sensory areas to predict sensory and proprioceptive consequences of that action. This would predict that action observation leads to a predominantly temporal → parietal → premotor flow of information in which

1 The term 'resonance' is here used in a rather loose sense as opposed to a strict physical meaning of the word, see paragraph 4.1

a visual/auditory representation of an action is transformed into the corresponding motor-programs, which contributes to action understanding. An alternative account sees the pMNS as a dynamic feedback control system. This latter model is inspired by both forward and inverse models from motor control theory (Wolpert et al., 2003; Voss et al., 2006; Greenwald, 1970). An inverse model, in contrast to a forward model, calculates what motor commands should be executed to achieve a certain goal or end-state[2] (Wolpert et al., 2003). A dynamic feedback control system view of the pMNS incorporates both forward and inverse models and furthermore assumes that the forward model is inhibitory, while the inverse model is excitatory. Combined, this predicts that when we see a predictable chain of events, the beginning is fully represented in the visual cortex and triggers motor programs through the inverse model. These motor programs are then 'forwarded' to predict future visual (and somatosensory) stimuli. If these stimuli conform to the predictions, they will be inhibited. The visual → premotor stream of information is reduced and the premotor representations triggered will not be substantially updated. If the visual information violates the predictions, it will not be inhibited, leading to a renewed visual → premotor stream of information and an update of the motor representations. Thus, a dynamic feedback control system would predict that action observation leads to a predominantly premotor → parietal → temporal flow of information because of a combination of inhibitory forward and excitatory inverse models.

## 1.5 GRANGER CAUSALITY

To investigate directed influences between brain regions within and between brains, we used Granger causality mapping. Granger causality is a measure of directed influence between two time series. Originally conceptualized in the econometric field by Wiener and formalized by Granger (Wiener, 1956; Granger, 1969), it was introduced as a connectivity analysis for fMRI data in 2003 by Goebel et al. (2003) and Roebroeck et al. (2005). Clive Granger formalized causality between two time series using the intuitively appealing concept of temporal precedence: if a signal change in A is consistently followed by a signal change in B, A Granger-causes B. Mathematically, this is calculated by comparing two regression equations: one in which the current value of a time series $y_i$ is explained by its own past ($y_{i-j}$) with one in which the same time series $y_i$ is explained both by its own past and the past of another time series ($x_{i-j}$). This results in error variances, whose F-ratio quantifies the influence $x$ exerts on $y$. The converse influence of $y$ on $x$ is also calculated and the difference between

2 This might sound confusing, as the strictly feed-forward account that is described seems to have the same characteristics as an inverse model. However, Kilner et al. (2007) write: "[...] there is no separate inverse model or controller; a forward model is simply inverted by suppressing the prediction error generated by the forward model".

the two reveals the dominant direction of influence between the two time series. In this way, Granger causality provides a statistical measure of directed influences between brain regions and has the advantage that it does not require an underlying anatomical model. It maps influences from a certain seed region to the rest of that brain and vice versa.

When we apply Granger causality on data of different brains, the influences from a certain seed in the gesturer are mapped on the brain of the guesser. Thus, a map of the guesser will show the brain regions that are influenced by the seed of the gesturer more than the other way around. This result is thus the net influence between gesturer and guesser (gesturer → guesser − guesser → gesturer). In our experiment the guesser → gesturer influence immediately served as a control quantity, because the one-way nature of the information flow provided by the video-camera and screen ensured that information could only flow from the gesturer to the guesser.

## 1.6 GRANGER CAUSALITY AND THE BOLD RESPONSE

Results of Granger causality analyses of fMRI data are interpreted as indications of information flow on a neuronal level. This is an indirect inference, however, as fMRI measures BOLD responses rather than neuronal activity directly. The BOLD response (Blood Oxygen Level Dependent response) is essentially a measure of changes in deoxyhemoglobin level triggered by changes in neural activity. It is assumed to originate from neural activity, predominantly synaptic (Logothetis et al., 2001; Logothetis and Wandell, 2004), but unfolds later in time (~ 4 - 6s). The temporal characteristics of the hemodynamic response that links neural activity to changes in the BOLD signal is not equal across brain regions and participants (Rajapakse et al., 1998; Aguirre et al., 1998; Kruggel and von Cramon, 1999; Handwerker et al., 2004) and this variability is assumed to cause problems for Granger causality analyses (David et al., 2008; de Marco et al., 2009; Roebroeck et al., 2005; Friston, 2009; Chang et al., 2008). One fear is that a systematic difference in hemodynamic response between two regions might introduce temporal precedence where there was none, leading to the report of spurious Granger causality findings. Another fear is that a difference in hemodynamic response might invert the reported direction of Granger causality. The intuitive idea behind this is as follows: If region A causes neural activity changes in region B and region B has a faster hemodynamic response than region A, a Granger causality analysis of the BOLD signal might indicate a net influence going from B to A rather than the true underlying neural causality that goes from A to B.

In a between-brains analysis, these hemodynamic differences between brain regions and individuals pose less of a problem because the time lags between neural activity in the two brains, in the order of seconds, larger than differences in hemody-

namic delay, in the order of tenth of seconds (Handwerker et al., 2004). In a within-brain analysis, however, this issue is more pressing. In a recent study, Deshpande et al. (2009) has investigated the effect of the variability of hemodynamic responses between two brain regions of one subject on Granger causality analyses. They found that even when intra-subject differences in hemodynamic response function are present, Granger causality is still sensitive to influences in the order of a hundred milliseconds. However, the effect of inter- and intrasubject differences on a group level had never been investigated. We examined if Granger causality can indeed deduce the dominant flow of neural information even when differences in hemodynamic response are present between different brain regions and between participants, before we applied Granger causality on our data to investigate the connectivity from and to the pMNS.

## 1.7 OUTLINE OF THE THESIS

In Chapter 2 we review functional and lesion studies on the neural substrates of the perception and production of both language and action. In particular, Broca's area (here taken to indicate BA 44, BA 45 and the ventral part of BA 6) is important for perception and production of both language and action. Broca's area overlaps partly with the premotor node of the pMNS and in this chapter we discuss several theories about why language and action share common substrates. The remainder of this thesis focuses on non-verbal communication. Chapter 3, Chapter 4 and Chapter 6 describe results of the charades experiment. Chapter 3 describes both a whole-brain analysis and a region-of-interest analysis of the charades data. The main question is to what extent the pMNS and mentalizing areas are involved in gesturing and guessing concepts during the game charades. Furthermore, we analyzed how activity in these regions is dependent on the intention induced by the task, by comparing activity during guessing and passive observation. The aim of Chapter 4 is to investigate the information flow *between brains* during the game charades. We extend Granger causality, which is usually applied within one brain, to a between-brain Granger causality. Then, we used brain activity of the gesturer to map regions in the brain of the guesser, whose brain activity has a Granger causal relation with brain activity of the gesturer. Our hypothesis is that pMNS areas show a dependency between brains, following from the resonance property of the pMNS. In Chapter 5 we investigate a possible confound of Granger causality: inter- and intrasubject variability in hemodynamic responses. We performed simulations to systematically investigate the effect of hemodynamic response, neuronal delay and connectivity strength on the result of Granger causality analyses on group level. We first tested whether this variability could lead to false positive results. Then we investigated whether differences in hemodynamic response could lead to inverted directions on group level. In Chapter 6 we use Granger causal-

ity within one brain to investigate how areas of the pMNS influence each other and other areas of the brain during guessing and observation of gestures. We compare two different models of the pMNS, the pMNS as a dynamic feedback control system versus the pMNS as a strict feedforward system, by testing the conflicting hypotheses that these models generate. Chapter 7 concludes the thesis with an overview of the current work and some final remarks.

# BIBLIOGRAPHY


Adenzato, M., Cavallo, M., and Enrici, I. (2010). Theory of mind ability in the behavioural variant of frontotemporal dementia: an analysis of the neural, cognitive, and social levels. *Neuropsychology*, 48(1):2–12.

Aguirre, G. K., Zarahn, E., and D'esposito, M. (1998). The variability of human, bold hemodynamic responses. *Neuroimage*, 8(4):360–9.

Amodio, D. M. and Frith, C. D. (2006). Meeting of minds: the medial frontal cortex and social cognition. *Nature Reviews of Neuroscience*, 7(4):268–77.

Apperly, I. A., Samson, D., Chiavarino, C., and Humphreys, G. W. (2004). Frontal and temporo-parietal lobe contributions to theory of mind: neuropsychological evidence from a false-belief task with reduced language and executive demands. *Journal of Cognitive Neuroscience*, 16(10):1773–84.

Arbib, M. A. (2008). From grasp to language: embodied concepts and the challenge of abstraction. *Journal of Physiology - Paris*, 102(1-3):4–20.

Bastiaansen, J. A. C. J., Thioux, M., and Keysers, C. (2009). Evidence for mirror systems in emotions. *Philosophical transactions of the Royal Society of London Series B, Biological sciences*, 364(1528):2391–404.

Bird, C. M., Castelli, F., Malik, O., Frith, U., and Husain, M. (2004). The impact of extensive medial frontal lobe damage on 'theory of mind' and cognition. *Brain*, 127(Pt 4):914–28.

Blakemore, S. J., Bristow, D., Bird, G., Frith, C. D., and Ward, J. (2005). Somatosensory activations during the observation of touch and a case of visionÂŰtouch synaesthesia. *Brain*, 128(7):1571–1583.

Blakemore, S. J. and Frith, C. (2003). Self-awareness and action. *Current Opinion in Neurobiology*, 13(2):219–24.

Bonda, E., Petrides, M., Ostry, D., and Evans, A. C. (1996). Specific involvement of human parietal systems and the amygdala in the perception of biological motion. *The Journal of Neuroscience*, 16(11):3737–3744.

Brass, M., Bekkering, H., Wohlschläger, A., and Prinz, W. (2000). Compatibility between observed and executed finger movements: Comparing symbolic, spatial, and imitative cues. *Brain and Cognition*, 44(2):124–143.

Brass, M., Schmitt, R. M., Spengler, S., and Gergely, G. (2007). Investigating action understanding: inferential processes versus action simulation. *Current biology*, 17(24):2117–21.

Brunet, E., Sarfati, Y., Hardy-Baylé, M., and Decety, J. (2000). A pet investigation of the attribution of intentions with a nonverbal task. *Neuroimage*, 11(2):157–166.

Buccino, G., Binkofski, F., Fink, G. R., Fadiga, L., Fogassi, L., Gallese, V., Seitz, R. J., Zilles, K., Rizzolatti, G., and Freund, H.-J. (2001). Action observation activates premotor and parietal areas in a somatosopic manner: An fmri study. *The European Journal of Neuroscience*, 13:400–404.

Carrington, S. J. and Bailey, A. J. (2009). Are there theory of mind regions in the brain? a review of the neuroimaging literature. *Human brain mapping*, 30(8):2313–35.

Castelli, F., Happè, F., Frith, U., and Frith, C. D. (2000). Movement and mind: A functional imaging study of perception and interpretation of complex intentional movement patterns. *Neuroimage*, 12:314–325.

Chang, C., Thomason, M. E., and Glover, G. H. (2008). Mapping and correction of vascular hemodynamic latency in the bold signal. *Neuroimage*, 43(1):90–102.

Chong, T. T.-J., Cunnington, R., Williams, M. A., Kanwisher, N., and Mattingley, J. B. (2008). fmri adaptation reveals mirror neurons in human inferior parietal cortex. *Current biology*, 18(20):1576–80.

Ciaramidaro, A., Adenzato, M., Enrici, I., Erk, S., Pia, L., Bara, B., and Walter, H. (2007). The intentional network: how the brain reads varieties of intentions. *Neuropsychology*, 45(13):3105–13.

Corbetta, M., Patel, G., and Shulman, G. L. (2008). The reorienting system of the human brain: from environment to theory of mind. *Neuron*, 58(3):306–24.

Craighero, L., Bello, A., Fadiga, L., and Rizzolatti, G. (2002). Hand action preparation influences the responses to hand pictures. *Neuropsychology*, 40(5):492–502.

David, O., Guillemain, I., Saillet, S., Reyt, S., Deransart, C., Segebarth, C., and Depaulis, A. (2008). Identifying neural drivers with functional mri: an electrophysiological validation. *PLoS Biology*, 6(12):2683–97.

de Lange, F. P., Spronk, M., Willems, R. M., Toni, I., and Bekkering, H. (2008). Complementary systems for understanding action intentions. *Current biology*, 18(6):454–7.

de Marco, G., Devauchelle, B., and Berquin, P. (2009). Brain functional modeling, what do we measure with fmri data? *Neuroscience Research*, 64(1):12–9.

Decety, J. and Lamm, C. (2007). The role of the right temporoparietal junction in social interaction: how low-level computational processes contribute to meta-cognition. *The Neuroscientist*, 13(6):580–93.

Deshpande, G., Sathian, K., and Hu, X. (2009). Effect of hemodynamic variability on granger causality analysis of fmri. *Neuroimage*, 52(3):884–896.

Ferrari, P., Gallese, V., Rizzolatti, G., and Fogassi, L. (2003). Mirror neurons responding to the observation of ingestive and communicative mouth actions in the monkey ventral premotor cortex. *The European Journal of Neuroscience*, 17(8):1703–1714.

Fletcher, P. C., Happè, F., Frith, U., and Baker, S. (1995). Other minds in the brain: A functional neuroimaging study of 'theory of mind' in story comprehension. *Cognition*, 57(2):109–128.

Fogassi, L., Ferrari, P. F., Gesierich, B., Rozzi, S., Chersi, F., and Rizzolatti, G. (2005). Parietal lobe: from action organization to intention understanding. *Science*, 308(5722):662–7.

Friston, K. (2009). Causal modelling and brain connectivity in functional magnetic resonance imaging. *PLoS Biology*, 7(2):e33.

Frith, C. D. and Frith, U. (1999). Interacting minds–a biological basis. *Science*, 286(5445):1692–1695.

Frith, U. and Frith, C. D. (2003). Development and neurophysiology of mentalizing. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences*, 358(1431):459–473.

Fujii, N., Hihara, S., and Iriki, A. (2007). Social cognition in premotor and parietal cortex. *Social Neuroscience*, 3(3):250–260.

Gallagher, H. L., Happè, F., Brunswick, N., Fletcher, P. C., Frith, U., and Frith, C. D. (2000). Reading the mind in cartoons and stories: An fmri study of 'theory of mind' in verbal and nonverbal tasks. *Neuropsychology*, 38:11–21.

Gallagher, S. (2007). Simulation trouble. *Social Neuroscience*, 2(3-4):353–65.

Gallese, V. (2003). The manifold nature of interpersonal relations: The quest for a common mechanism. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences*, 358:517–528.

Gallese, V., Fadiga, L., Fogassi, L., and Rizzolatti, G. (1996). Action recognition in the premotor cortex. *Brain*, 119(2):593–609.

Gallese, V. and Goldman, A. (1998). Mirror neurons and the simulation theory of mind-reading. *Trends in Cognitive Sciences*, 12:493–501.

Gazzola, V., Aziz-Zadeh, L., and Keysers, C. (2006). Empathy and the somatotopic auditory mirror system in humans. *Current biology*, 16(18):1824–9.

Gazzola, V. and Keysers, C. (2008). The observation and execution of actions share motor and somatosensory voxels in all tested subjects: Single-subject analyses of unsmoothed fmri data. *Cerebral Cortex*, 19(6):1239–1255.

Gazzola, V., Rizzolatti, G., Wicker, B., and Keysers, C. (2007a). The anthropomorphic brain: The mirror neuron system responds to human and robotic actions. *Neuroimage*, 35:1674–1684.

Gazzola, V., van der Worp, H., Mulder, T., Wicker, B., Rizzolatti, G., and Keysers, C. (2007b). Aplasics born without hands mirror the goal of hand actions with their feet. *Current biology*, 17(14):1235–40.

Gentilucci, M. and Corballis, M. C. (2006). From manual gesture to speech: a gradual transition. *Neuroscience and Biobehavioral Reviews*, 30(7):949–60.

Gibson, J. (1986). The ecological approach to visual perception. *Boston, Houghton Mifflin*.

Goebel, R., Roebroeck, A., Kim, D.-S., and Formisano, E. (2003). Investigating directed cortical interactions in time-resolved fmri data using vector autoregressive modeling and granger causality mapping. *Magnetic resonance imaging*, 21:1251–1261.

Goldman, A. (1992). In defence of the simulation theory. *Mind and Language*, 7:104–119.

Grafton, S. T., Arbib, M. A., Fadiga, L., and Rizzolatti, G. (1996). Localization of grasp representations in humans by positron emission tomography. 2. observation compared with imagination. *Experimental Brain Research*, 112(1):103–111.

Granger, C. (1969). Investigating causal relations by econometric models and cross-spectral methods. *Econometrica*, 37(3):424–438.

Greenwald, A. G. (1970). Sensory feedback mechanisms in performance control: with special reference to the ideo-motor mechanism. *Psychological Review*, 77(2):73–99.

Gregory, C., Lough, S., Stone, V., Erzinclioglu, S., Martin, L., Baron-Cohen, S., and Hodges, J. R. (2002). Theory of mind in patients with frontal variant frontotemporal dementia and alzheimer's disease: theoretical and practical implications. *Brain*, 125(Pt 4):752–64.

Grèzes, J., Armony, J., Rowe, J., and Passingham, R. E. (2003). Activations related to mirror and canonical neurones in the human brain: An fmri study. *Neuroimage*, 18:928–937.

Grèzes, J., Costes, N., and Decety, J. (1998). Top-down effect of strategy on the perception of human biological motion: A pet investigation. *Cognitive Neuropsychology*, 15:553–582.

Grèzes, J. and Decety, J. (2001). Functional anatomy of execution, mental simulation, observation, and verb generation of actions: a meta-analysis. *Human Brain Mapping*, 12(1):1–19.

Grèzes, J., Frith, C. D., and Passingham, R. E. (2004). Brain mechanisms for inferring deceit in the actions of others. *The Journal of Neuroscience*, 24(24):5500–5505.

Hamilton, A., Wolpert, D., and Frith, U. (2004). Your own action influences how you perceive another person's action. *Current biology*, 14(6):493–8.

Handwerker, D. A., Ollinger, J. M., and D'Esposito, M. (2004). Variation of bold hemodynamic responses across subjects and brain regions and their effects on statistical analyses. *Neuroimage*, 21(4):1639–51.

Hickok, G. (2009). Eight problems for the mirror neuron theory of action understanding in monkeys and humans. *Journal of Cognitive Neuroscience*, 21(7):1229–43.

Hurley, S. (2008). The shared circuits model (scm): how control, mirroring, and simulation can enable imitation, deliberation, and mindreading. *Behavioral and Brain Sciences*, 31(1):1–22; discussion 22–58.

Iacoboni, M., Woods, R., Brass, M., Bekkering, H., Mazziotta, J. C., and Rizzolatti, G. (1999). Cortical mechanisms of human imitation. *Science*, 286(5449):2526–2528.

Jabbi, M., Swart, M., and Keysers, C. (2007). Empathy for positive and negative emotions in the gustatory cortex. *Neuroimage*, 34(4):1744–53.

Kampe, K. K. W., Frith, C. D., and Frith, U. (2003). "hey john": signals conveying communicative intention toward the self activate brain regions associated with "mentalizing," regardless of modality. *The Journal of Neuroscience*, 23(12):5258–63.

Keysers, C. and Gazzola, V. (2007). Integrating simulation and theory of mind: from self to social cognition. *Trends in Cognitive Sciences*, 11(5):194–6.

Keysers, C. and Gazzola, V. (2009). Expanding the mirror: vicarious activity for actions, emotions, and sensations. *Current Opinion in Neurobiology*, 19(6):666–71.

Keysers, C., Kaas, J. H., and Gazzola, V. (2010). Somatosensation in social perception. *Nature Reviews of Neuroscience*, 11(6):417–28.

Keysers, C., Kohler, E., Umiltà, M., Nanetti, L., Fogassi, L., and Gallese, V. (2003). Audiovisual mirror neurons and action recognition. *Experimental Brain Research*, 153(4):628–636.

Keysers, C., Wicker, B., Gazzola, V., Anton, J., Fogassi, L., and Gallese, V. (2004). A touching sight: Sii/pv activation during the observation and experience of touch. *Neuron*, 42:335–346.

Kilner, J. M., Friston, K. J., and Frith, C. D. (2007). Predictive coding: an account of the mirror neuron system. *Cognitive Processes*, 8(3):159–166.

Kilner, J. M., Neal, A., Weiskopf, N., Friston, K. J., and Frith, C. D. (2009). Evidence of mirror neurons in human inferior frontal gyrus. *The Journal of Neuroscience*, 29(32):10153–9.

Kilner, J. M., Paulignan, Y., and Blakemore, S. J. (2003). An interference effect of observed biological movement on action. *Current biology*, 13(6):522–5.

Kohler, E., Keysers, C., Umiltà, M., Fogassi, L., Gallese, V., and Rizzolatti, G. (2002). Hearing sounds, understanding actions: Action representation in mirror neurons. *Science*, 297(5582):846–849.

Kokal, I., Gazzola, V., and Keysers, C. (2009). Acting together in and beyond the mirror neuron system. *Neuroimage*, 47(4):2046–56.

Koski, L., Iacoboni, M., Dubeau, M.-C., Woods, R., and Mazziotta, J. C. (2003). Modulation of cortical activity during different imitative behaviors. *Journal of Neurophysiology*, 89(1):460–471.

Kruggel, F. and von Cramon, D. Y. (1999). Temporal properties of the hemodynamic response in functional mri. *Human brain mapping*, 8(4):259–71.

Lingnau, A., Gesierich, B., and Caramazza, A. (2009). Asymmetric fmri adaptation reveals no evidence for mirror neurons in humans. *Proceedings of the National Academy of Sciences of the United States of America*, 106(24):9925–30.

Logothetis, N. K., Pauls, J., Augath, M., Trinath, T., and Oeltermann, A. (2001). Neurophysiological investigation of the basis of the fmri signal. *Nature*, 412(6843):150–7.

Logothetis, N. K. and Wandell, B. A. (2004). Interpreting the bold signal. *Annual Review of Physiology*, 66:735–69.

Mitchell, J. P. (2008). Activity in right temporo-parietal junction is not selective for theory-of-mind. *Cerebral Cortex*, 18(2):262–71.

Mitchell, J. P., Banaji, M., and Macrae, C. N. (2005). General and specific contributions of the medial prefrontal cortex to knowledge about mental states. *Neuroimage*, 28(4):757–762.

Mitchell, J. P., Heatherton, T. F., and Macrae, C. N. (2002). Distinct neural systems subserve person and object knowledge. *Proceedings of the National Academy of Sciences of the United States of America*, 99(23):15238–43.

Mukamel, R., Ekstrom, A. D., Kaplan, J., Iacoboni, M., and Fried, I. (2010). Single-neuron responses in humans during execution and observation of actions. *Current biology*, 20(8):750–756.

Müsseler, J. and Hommel, B. (1997). Blindness to response-compatible stimuli. *Journal of Experimental Psychology: Human perception and Performance*, 23(3):861–72.

Newman-Norlund, R. D., Bosga, J., Meulenbroek, R. G. J., and Bekkering, H. (2008). Anatomical substrates of cooperative joint-action in a continuous motor task: virtual lifting and balancing. *Neuroimage*, 41(1):169–77.

Newman-Norlund, R. D., van Schie, H. T., van Zuijlen, A. M. J., and Bekkering, H. (2007). The mirror neuron system is more active during complementary compared with imitative action. *Nat Neurosci*, 10(7):817–8.

Nishitani, N. and Hari, R. (2000). Temporal dynamics of cortical representation for action. *Proceedings of the National Academy of Sciences of the United States of America*, 97(2):913–8.

Nishitani, N. and Hari, R. (2002). Viewing lip forms: cortical dynamics. *Neuron*, 36(6):1211–20.

Noordzij, M. L., Newman-Norlund, S. E., de Ruiter, J. P., Hagoort, P., Levinson, S. C., and Toni, I. (2009). Brain mechanisms underlying human communication. *Frontiers in Human Neuroscience*, 3:1–13.

Ochsner, K. and Gross, J. (2004). Thinking makes it so: A social cognitive neuroscience approach to emotion regulation. *Handbook of self-regulation: Research*.

Overwalle, F. V. and Baetens, K. (2009). Understanding others' actions and goals by mirror and mentalizing systems: a meta-analysis. *Neuroimage*, 48(3):564–84.

Perani, D., Fazio, F., Borghese, N., Tettamanti, M., Ferrari, S., Decety, J., and Gilardi, M. (2001). Different brain correlates for watching real and virtual hand actions. *Neuroimage*, 14(3):749–758.

Prather, J. F., Peters, S., Nowicki, S., and Mooney, R. (2008). Precise auditory-vocal mirroring in neurons for learned vocal communication. *Nature*, 451(7176):305–10.

Premack, D. and Woodruff, G. (1978). Does the chimpanzee have a theory of mind? *Behavioral and Brain Sciences*, 1:515–526.

Prinz, W. (1990). A common coding approach to perception and action. *Relationships between perception and action. Berlin: Springer-Verlag*.

Rajapakse, J. C., Kruggel, F., Maisog, J. M., and von Cramon, D. Y. (1998). Modeling hemodynamic response for analysis of functional mri time-series. *Human brain mapping*, 6(4):283–300.

Rizzolatti, G. and Arbib, M. A. (1998). Language within our grasp. *Trends in Neurosciences*, 21(5):188–194.

Rizzolatti, G., Fadiga, L., Gallese, V., and Fogassi, L. (1996). Premotor cortex and the recognition of motor actions. *Cognitive Brain Research*, 3(18):131–141.

Roebroeck, A., Formisano, E., and Goebel, R. (2005). Mapping directed influence over the brain using granger causality. *Neuroimage*, 25:230–242.

Rowe, A., Bullock, P., Polkey, C., and Morris, R. (2001). 'theory of mind' impairments and their relationship to executive functioning following frontal lobe excisions. *Brain*, 124(3):600–616.

Ruby, P. and Decety, J. (2003). What you believe versus what you think they believe: a neuroimaging study of conceptual perspective-taking. *The European Journal of Neuroscience*, 17(11):2475–2480.

Samson, D., Apperly, I. A., Chiavarino, C., and Humphreys, G. W. (2004). Left temporoparietal junction is necessary for representing someone else's belief. *Nature Neuroscience*, 7(5):499–500.

Saxe, R. (2006). Uniquely human social cognition. *Current Opinion in Neurobiology*, 16(2):235–239.

Saxe, R. and Kanwisher, N. (2003). People thinking about thinking people. the role of the temporo-parietal junction in "theory of mind". *Neuroimage*, 19(4):1835–42.

Saxe, R. and Wexler, A. (2005). Making sense of another mind: the role of the right temporo-parietal junction. *Neuropsychology*, 43(10):1391–9.

Singer, T., Seymour, B., O'Doherty, J., Kaube, H., Dolan, R., and Frith, C. D. (2004). Empathy for pain involves the affective but not sensory components of pain. *Science*, 303(5661):1157–1162.

Sommer, M., Döhnel, K., Sodian, B., Meinhardt, J., Thoermer, C., and Hajak, G. (2007). Neural correlates of true and false belief reasoning. *Neuroimage*, 35(3):1378–84.

Spreng, R. N., Mar, R. A., and Kim, A. S. N. (2009). The common neural basis of autobiographical memory, prospection, navigation, theory of mind, and the default mode: a quantitative meta-analysis. *Journal of Cognitive Neuroscience*, 21(3):489–510.

Stuss, D., Gallup, G., and Alexander, M. (2001). The frontal lobes are necessary for 'theory of mind'. *Brain*, 124(Pt 2):279–286.

Uddin, L. Q., Iacoboni, M., Lange, C., and Keenan, J. P. (2007). The self and social cognition: the role of cortical midline structures and mirror neurons. *Trends in Cognitive Sciences*, 11(4):153–7.

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

van der Gaag, C., Minderaa, R. B., and Keysers, C. (2007). Facial expressions: what the mirror neuron system can and cannot tell us. *Social Neuroscience*, 2(3-4):179–222.

van der Meer, L., Costafreda, S., Aleman, A., and David, A. S. (2010). Self-reflection and the brain: a theoretical review and meta-analysis of neuroimaging studies with implications for schizophrenia. *Neuroscience and Biobehavioral Reviews*, 34(6):935–46.

Vogeley, K., Bussfeld, P., Newen, A., Herrmann, S., Happè, F., Falkai, P., Maier, W., Shah, N. J., Fink, G. R., and Zilles, K. (2001). Mind reading: neural mechanisms of theory of mind and self-perspective. *Neuroimage*, 14(1 Pt 1):170–181.

Voss, M., Ingram, J. N., Haggard, P., and Wolpert, D. M. (2006). Sensorimotor attenuation by central motor command signals in the absence of movement. *Nature Neuroscience*, 9(1):26–7.

Wicker, B., Keysers, C., Plailly, J., Royet, J., Gallese, V., and Rizzolatti, G. (2003). Both of us disgusted in my insula: The common neural basis of seeing and feeling disgust. *Neuron*, 40:655–664.

Wiener, N. (1956). Theory of prediction. *Modern Mathematics for Engineers, Series 1*.

Wimmer, H. and Perner, J. (1983). Beliefs about beliefs: Representation and constraining function of wrong beliefs in young children's understanding of deception. *Cognition*, 13(1):103–128.

Wolpert, D. M., Doya, K., and Kawato, M. (2003). A unifying computational framework for motor control and social interaction. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences*, 358(1431):593–602.

Young, L., Dodell-Feder, D., and Saxe, R. (2010). What gets the attention of the temporo-parietal junction? an fmri investigation of attention and theory of mind. *Neuropsychology*, 48(9):2658–64.

# 2

# BROCA'S AREA: LINKING PERCEPTION AND PRODUCTION IN LANGUAGE AND ACTIONS

*Published as:* Rossi, E.[1], Schippers, M.B.[1], & Keysers, C. (2011). Broca's area: linking perception and production in language and actions. In Han, S., editor, *Culture and Neural Frames of Cognition and Communication*, On Thinking. Vol. 3. Springer Berlin Heidelberg

## ABSTRACT

A distinction between action perception and production has always been emphasized by traditional accounts of brain function. The goal of this review is to show that this simple distinction seems no longer valid. Broca's area in particular seems to be important for both perception and production of language and action. Functional imaging studies suggest that Broca's area is active both when people produce and perceive syntactically complex sentences and while they produce and perceive complex actions. Lesions in this area disrupt the capacity to produce syntactically correct sentences and to perceive sentences in which syntax is essential. From an action-perspective, lesions to Broca's area disrupt the capacity to produce goal directed actions and to perceive the actions of others. Furthermore, the property and location of mirror neurons in the monkey might provide the reason why Broca's area in humans has a dual function in production and perception.

## 2.1 INTRODUCTION

Traditional accounts of brain function have often emphasized a distinction between action perception and production. For instance, the occipital and temporal lobe were considered to deal with visual and auditory perception while the frontal lobe was considered to deal with the production of goal directed actions. In the domain of language as well, the classic distinction between Wernicke and Broca's aphasia suggested that one part of the brain deals with perceiving what other people say while the other deals with producing speech.

1 Authors contributed equally to this bookchapter

The goal of this review is to show that this simple distinction seems no longer valid. In particular, we will show how both in the domain of language and action, a brain area, called Broca's Area, seems to be important for both perception and production. Finally, we will show how mirror neurons help us understand why and how a certain brain area can be important in both perception and production.

We will start by explaining where Broca's area is located in the brain. We will then show how lesions in this area disrupt the capacity to produce syntactically correct sentences and to perceive sentences in which syntax is essential. We will then show that lesions to Broca's area disrupt the capacity to produce goal directed actions and to perceive the actions of others. We will then review data from functional imaging studies that suggest that Broca's area is active both when people produce and perceive syntactically complex sentences and while they produce and perceive complex actions. Finally, we will review the property and location of mirror neurons in the monkey and suggest that they may be the reason why Broca's area in humans has a dual function in production and perception.

## 2.2 BROCA'S ANATOMY

The name 'Broca's area' comes from the French neurologist Pierre Paul Broca, who brought the inferior frontal gyrus into the spotlight as a possible location for the seat of language in the human brain. When he investigated the brains of his deceased patients who suffered from a "loss of speech", he found lesions located in the frontal lobe. He decided, however, not to dissect the brains and only described the lesions from the outside. He sent the brains to a museum in Paris giving neuroscientists nowadays the opportunity to use modern imaging techniques to investigate his original findings (Cabanis et al., 1994; Dronkers et al., 2007; Castaigne et al., 1980) From these images we know now that the lesions are not confined to parts of the inferior frontal gyrus only, but they extend medially into the arcuate/superior longitudinal fasciculus, which connects anterior and posterior language regions (Geschwind, 1972).

The original finding of Broca, however, has led to a substantial amount of research on Broca's area and these studies have made it further clear that Broca's area does not consist of one cytoarchitectonically well-defined area, but comprises several areas, including Brodmann areas (BA) 44 and 45 and the ventral part of Brodmann area 6. In the remainder of this chapter, we will use the term Broca's area to indicate BA 44 and 45 and the ventral part of BA 6.

## 2.3 LESION STUDIES OF LANGUAGE

### 2.3.1 *Language production*

As introduced, the term 'Broca's aphasia' has been minted after Paul Broca's description of patient 'Tan' to address production language impairment due to a damage in the posterior half of the left inferior frontal gyrus. The label 'Broca's aphasia', though, seemed very soon too broad and unspecific to describe the complex pattern of linguistic deficits related to a brain damage in the 'language production areas'. After the fundamental studies of Arnold Pick (Pick, 1898, 1913) the term 'agrammatism' has been used to denote a type of Broca's aphasia, which involves a specific impairment at the grammatical level. Since then, the production deficits seen in Broca's aphasia and agrammatism have been described in a great number of lesion studies across languages (see Bates et al., 1991, for a review), studies which all converge in showing that lesions in Broca's area result in a production deficit with a particular involvement of grammar. The production deficits observed in Broca's aphasia are generally charac-

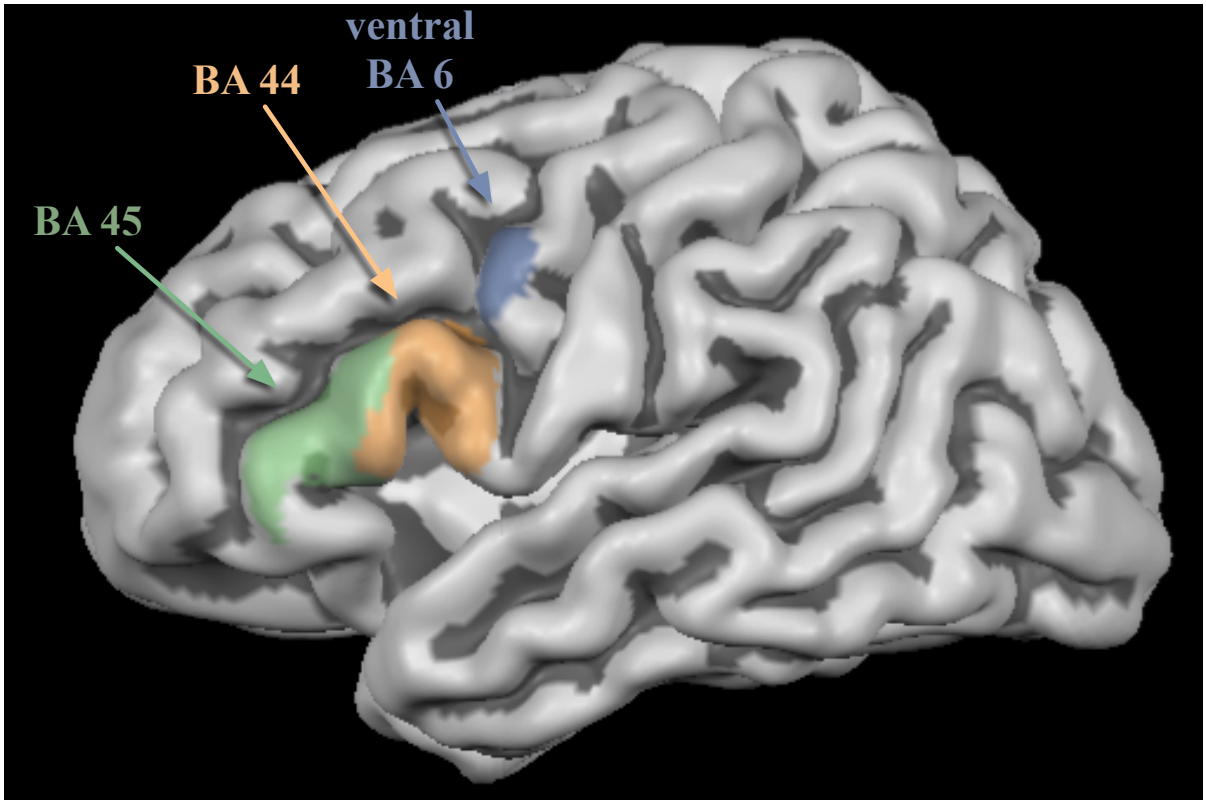



Figure 2.1: Broca's area shown on the cortical surface of the brain. Broca's area comprises several areas, including Brodmann areas (BA) 44 and 45 and the ventral part of Brodmann area 6.

terized by a decrease in speech rate, omission and/or substitution of function words (articles, prepositions, pronouns), a frequent use of uninflected verbs (particularly detectable in morphologically rich languages in which the non-finite form of verbs is morphologically marked, like Italian 'mangiare' - to eat), and a reduced use in the use of coordinated and subordinated sentences together with a loss of comprehension of complex syntactic structures. Miceli et al. (1984) showed that Italian aphasic speakers with a lesion in Broca's area omit free morphemes (articles and prepositions) and substitute bound morphemes (verb inflections for Tense and Person agreement). The same deficits were observed for German by DeBleser et al. (1996) and Luzzatti and Bleser (1996) who showed that agrammatic aphasic speakers were impaired in 'syntax-dependent morphology', i.e. morphological processes which are more directly linked to syntactic operations (like verbal inflection).

As far as language production goes, the data provided by lesion studies seem to converge in showing that lesions in Broca's area produce an agrammatic speech output, which suggests that these brain areas are the locus in which syntax is stored. But if Broca's area indeed represents the anatomical locus for syntax, once a damage is located in that area, a language impairment at the syntactic level should not only be detectable in production but also in comprehension.

### 2.3.2 *Language comprehension*

Thanks to a series of pioneer lesion studies on agrammatism (Caramazza and Zurif, 1976; Grodzinsky, 2000; Luzzatti et al., 2001; Shapiro and Levine, 1990), it became clear that people with lesions in Broca's area (which had been addressed for a long time as the specific locus for language production) not only show deficits in production but they also show similar impairments in the comprehension of complex grammatical structures. Caramazza and Zurif (1976) were among the first to describe that patients with a lesion in Broca's area show deficits in comprehension of complex syntactic structures. The authors showed that agrammatic patients performed at chance when interpreting semantically reversible sentences, i.e. sentences that require a syntactic analysis to be correctly understood like: "The girl was kicked by the boy". In contrast, they performed above chance with sentences that could be interpreted using semantic rules as a disentangling strategy, such as: "The ball was kicked by the girl". Further studies on comprehension deficits in Broca's aphasia (e.g. Grodzinsky, 1995; Luzzatti et al., 2001) report that other complex syntactic structures, like passive and subordinate sentences as well as sentences with pronominal clitic pronouns, are difficult to interpret for aphasic speakers.

Lesion studies have provided (and still provide) fundamental information on the role played by Broca's area in language processing, providing converging evidence that

Broca's area is involved both in language production and in language comprehension, disconfirming the initial thought that Broca's area was exclusively involved in language production. Second, the evidence is strong in favor of Broca's area being involved in syntactic processing, supporting the view that this area is the locus of grammatical encoding.

## 2.4 LESION STUDIES OF ACTIONS

Lesions in and around Broca's area are relatively well known for their association with Broca's aphasia, which we have discussed in the previous paragraphs. In this section, we will focus on how lesions in this part of the brain can cause disruptions in executing action and observing the actions of others.

### 2.4.1 *Action execution*

Deficits in producing actions are known as apraxia. Patients who suffer from apraxia have difficulties in executing learned movements even though they are physically able to perform these movements. The disorders aphasia and apraxia very often occur together (Renzi et al., 1980).

Lesions in Broca's area play an important role in apraxia (Leiguarda and Marsden, 2000). They can lead to weakness of muscles that control oro-facial, laryngeal and tongue movements (Förster, 1936). Furthermore, lesions in Broca's area can lead to difficulties in sequencing of actions (Harrington et al., 1998), to a loss of regularity of exploratory finger movements during manipulation of objects (Binkofski et al., 2001), to deficits in visuomotor associative learning (Binkofski and Buccino, 2004), and to deficits in grasping (Dettmers et al., 2003).

Natural lesions that are restricted to one cytoarchitectonically well-defined brain area are rare. Fortunately, with the technique of repetitive transcranial magnetic stimulation (rTMS), it is possible to create temporary 'lesions' in a particular brain area to investigate the necessity of this brain area during a certain task. To investigate the role of Broca's area during imitation of finger movements, Heiser et al. (2003) delivered rTMS over right and left Broca's area while participants had to perform this task. They found that these stimulations caused a disruption in the imitation task, while the control rTMS over the occipital cortex did not show such a disruption. This indicates that Broca's area could be the place in which a matching between observed and to- be-executed actions takes place. Since rTMS did not cause a disruption in a motor control task, the deficits in imitation could not have been caused by a disruption in planning or selection.

The design of this experiment was, however, criticized by Makuuchi (2005), who

argued that the movement that had to be imitated was too simple. This has as a consequence that only in the first few trials has the observed movement really to be transformed into one's own motor representation, but in later trials, the observed movement merely functions as a visual cue to perform a learned movement. Makuuchi (2005) performed a new experiment from which he concludes that it is not imitation that Broca's area is essential for, but it is the delayed execution of actions.

Damage to Broca's area thus does not only have an influence on the production and perception of language, but also disrupts action production.

### 2.4.2 *Action perception*

Besides disrupting action production, damage in Broca's area also has a profound influence on the perception of actions performed by other people. Damage to Broca's area leads to an impairment in conceptual knowledge about actions (Tranel et al., 2003).

Aphasic patients with lesions in the inferior frontal gyrus and the ventral premotor cortex have trouble with action understanding (Saygin et al., 2004). This is independent of whether the action is presented linguistically (i.e. a written description) or non-linguistically (i.e. visual presentation). Furthermore, the linguistic and non-linguistic deficits are correlated with each other in the mild and relatively fluent aphasics, which implies a common underlying cause of the deficits. Artificially lesioning this area with TMS disrupts action understanding of other people's actions (Pobric and de C Hamilton, 2006). These results are in accordance with results of Aziz-Zadeh et al. (2006), who (in non brain damaged subjects) found overlapping activations for action observation and reading sentences about actions in Broca's area.

Pazzaglia et al. (2008) provide evidence that indicate that the deficits in action production are related to deficits in action perception. First, they showed that brain-damaged patients who suffer from limb apraxia also show a greater impairment in recognition of gestures than brain-damaged patients who do not suffer from this disorder. Second, premotor and parietal lesions that impair hand action execution (as compared to mouth action execution) also selectively impair the recognition of hand gestures and their sounds (Pazzaglia et al., 2008).

All these studies show that deficits in action comprehension and action production are very much interrelated and associated with Broca's area.

## 2.5 FUNCTIONAL STUDIES ON LANGUAGE

If lesion studies inform us on the specific linguistic deficits that Broca's patients show in production and comprehension, the growing use in linguistic research of

neuroimaging techniques such as fMRI (functional Magnetic Resonance Imaging), ERP (Event-Related Potentials), MEG (Magnetoencephalography) and PET (Positron Emission Tomography) has made possible the investigation of language processing in healthy people. fMRI and PET, with their high spatial resolution permit the localization of particular brain areas which are involved during specific language tasks. PET is used more often for production studies given that it is less sensitive to movement and muscular artifacts, whereas fMRI has been extensively used in comprehension paradigms.

### 2.5.1 *Language production*

In a PET stud Indefrey et al. (2001) elicited the production of sentences, which were increasingly grammatically complex. Participants were asked to describe short animated movies, which involved non-animated participants (a circle, an ellipse and a square). Results showed that the production of syntactically more complex sentences resulted in an activation of the left anterior operculum, caudally adjacent to BA44. Haller et al. (2005) performed an fMRI study involving open speech production. Participants were required to generate sentences given bare syntactic constituents (for example using the constituents: 'child', 'throw' and 'ball', participants should produce sentences like: 'The child throws the ball'). The activation resulting from the sentence generation task was compared with a sentence reading and a word repetition tasks. Both contrasts revealed that BA44/45 and BA6 were activated. Word level production studies involving syntactic processing have also been performed. Jaeger et al. (1996) and Indefrey et al. (1997) investigated which are the neurocorrelates of regular and irregular past verb formation. According to linguistic accounts, regular past verbs are formed using morpho/syntactic rules, for example in English, affixing the -ed morpheme to the verbal root. Irregular past verbs, however, cannot be 'blindly' formed applying a morpho/syntactic rule but their specific forms need to be stored in the lexicon. The two studies report that producing the past tense of regular verbs activates inferior frontal regions (regions that have been found to be active in morpho/syntactic processes), whereas producing the past tense of irregular verbs activates middle temporal regions (more involved with lexical processes).

### 2.5.2 *Language comprehension*

Ben-Shachar et al. (2003) performed an fMRI study to check which were the areas involved in a specific syntactic operation, i.e. syntactic movement. Participants had to listen to sentences and after this make a grammatical judgment about them. Results show that Broca's area was activated when sentences contained a moved element.

Broca's area (together with Wernicke's regions in both hemispheres) was activated in another fMRI study by Ben-Shachar et al. (2004). In this study, the task consisted of a comprehension test with two other types of grammatical structures involving syntactic movement, i.e. topicalization and embedded questions. These authors consider the activation of the left Broca's area crucial for syntactic processes. Meyer et al. (2000) found an activation of the left Broca's area during auditory presentation of grammatically correct and incorrect sentences. Fiebach et al. (2001) conducted a study aimed at detecting the areas involved in syntactic transformation or in the detection of syntactic anomalies. The results revealed that BA 44/45 were active in sentences with syntactic transformations, and BA 44/6 were active while detecting syntactic anomalies. Area BA 44 was activated in a study by Dapretto and Bookheimer (1999), when participants had to focus their attention towards more syntactic aspects of sentences compared to more semantic ones.

Summarizing the results from these studies, it is possible to speak about a network of regions within Broca's area, which support syntactic processing both for production and language comprehension. More specifically, the left inferior frontal gyrus with areas BA 44/45 are actively involved in more complex syntactic processing, while the frontal operculum seems to support the detection of whether a structure is grammatical or not.

## 2.6 FUNCTIONAL STUDIES ON ACTIONS

### 2.6.1 *Action execution*

Activation in Broca's area is found during the programming of object directed action execution, particularly when the action is a complex motor act which requires a high degree of sensorimotor control (Binkofski and Buccino, 2004). In our laboratory, activation in Broca's area is always found when comparing object- directed action execution against rest (Gazzola et al., 2006; Gazzola and Keysers, 2008; Gazzola et al., 2007a,b).

Further evidence for the fact that Broca's area is involved in the motor programming of actions comes from a study by Haslinger et al. (2002) in which participants have to perform increasingly complex finger movements. Results show that the more complex the sequence of movements is, the more Broca's area is involved. Other studies by Schubotz & Cramon (2003; 2001; 2002a; 2002b; 2002c) have shown similar results in that the ventral premotor cortex part of Broca's area is engaged when a sequential-based prediction of the action has to be made (for example, to predict the end state of a sequence of movements).

Examples of other kinds of motor acts that involve Broca's area are grasping actions

(Decety et al., 1994; Ehrsson et al., 2000; Grafton et al., 1996), manipulation of objects (Binkofski et al., 1999), finger movements (Krams et al., 1998; Seitz and Roland, 1992), and gesturing (Fridman et al., 2006).

Summarizing, parts of Broca's area seems to be involved in action execution, particularly when the action is complex (both in terms of movement and sequencing) and is object-directed.

### 2.6.2 *Action perception*

Studies using movies of simple hand actions show that Broca's area is consistently activated when observing these simple hand or mouth actions (Buccino et al., 2001; Gazzola and Keysers, 2008; Gazzola et al., 2007b,a; Grafton et al., 1996; Rizzolatti et al., 1996b). Broca's area is particularly involved when the action is goal-directed and includes an object, for example grasping a little cup, biting and chewing an apple (Buccino et al., 2001). Not only the visual perception of an action involves Broca's area, but the mere sound of actions also elicits a response in this area (Gazzola et al., 2006).

It is argued that, for Broca's area to respond to observed action, the action needs to be part of the motor repertoire of the observer. For example, Broca's area responds to the observation of mouth actions of humans (speech pronunciation) and monkeys (lip smacking), but not to mouth actions of a dog (barking) (Buccino et al., 2004). Gazzola et al. (2007a) extend this finding by showing that Broca's area is also involved in actions the kinematics of which we cannot match onto our own motor repertoire, but of which we do understand the goal (e.g., 'human' actions performed by an industrial robot).

The fact that Broca's area was historically linked to language processing, raised the question whether activation in this region is truly due to the processing of the action or to a form of inner verbalization of the action (Decety et al., 1997; Grèzes and Decety, 2001). There is now, however, relatively wide agreement about the fact that the idea of 'silent speech' cannot account for the activation in Broca's area. If activation in Broca's area would be due to inner speech, then one would expect that imitation with the left or right hand would activate this area similarly; however, Koski et al. (2003) found a difference in activation due to imitation with one hand or the other. Second, inner speech would predict that hearing and performing hand and mouth actions should cause similar patterns of activity in premotor regions. Gazzola et al. (2006) and Etzel et al. (2008), however, showed that hand and mouth actions determine different patterns of activity, which are, however, similar during execution and perception. In another study, rTMS was applied over left and right BA 44, causing a disruption in the imitation process (Heiser et al., 2003). Could it be a disturbance in the silent verbalization of the action that disrupted the imitation? The authors

note that this is a highly unlikely explanation, since pre-verbal little children cannot verbalize actions but can imitate them.

Summarizing these functional studies, we can say that Broca's area is highly important for the perception of other people's actions and for programming the execution of complex actions of hands and mouth, given that these are not too repetitive.

## 2.7 MIRROR NEURONS & THE PUTATIVE MIRROR NEURON SYSTEM

In the previous section, we have seen that Broca's area is involved in both perception and production of complex actions. Is it truly the same neural substrate that is responsible for these different tasks? In the monkey's brain there is evidence for the idea that production and perception depend on the same neurons, so-called mirror neurons.

Mirror neurons were first discovered in Italy (Gallese et al., 1996; Pellegrino et al., 1992; Rizzolatti et al., 1996a). Activity from single neurons in the macaque monkey's brain had been recorded when the monkey was performing an action (i.e. grasping a peanut, shelling a peanut). The researchers discovered that some neurons in this area not only showed activity during action execution, but also when the monkey observed the researcher grasping a peanut or shelling it. Later, the same laboratory would show that some mirror neurons also respond to the sound of a similar action (Keysers et al., 2003; Kohler et al., 2002). These neurons thus have the special property of firing not only when the monkey performs an action but also when a similar action is perceived. Mirror neurons therefore show a direct connection between perception and action. The areas in which mirror neurons have been recorded from in the monkey are the rostral part of inferior area 6 (area F5) (Pellegrino et al., 1992; Gallese et al., 1996; Keysers et al., 2003; Kohler et al., 2002) and the rostral part of the inferior parietal lobule (area 7b) (Fogassi et al., 2005).

Since the moment of discovery of mirror neurons, the question arose whether such neurons would be present in the human brain. Indeed, evidence for a mirror neuron system in humans has been derived from neuroimaging and transcranial magnetic stimulation studies, with the former showing that a network of areas is active both while people perform actions and while they view or hear other people's actions (Gazzola et al., 2006; Keysers and Gazzola, 2006; Rizzolatti and Craighero, 2004). In humans, this system seems to include the dorsal premotor, somatosensory, cerebellar and posterior temporal cortex in addition to BA 44 and 6 and the inferior parietal lobule (Gazzola and Keysers, 2008).

But is it also true for humans that it is the same population of neurons that respond both to the observation and execution of actions? We cannot say anything about

individual neurons, but on the level of individual voxels[2], we can affirm that this is the case. Gazzola and Keysers (2008) have shown that within individual subjects, they are truly the same voxels that respond both to the perception and production of complex actions. In addition, in a recent study, Etzel et al. (2008), using an analysis technique known as multivariate classification, could show that the perception and execution of actions not only both recruit Broca's area, but that they indeed determine similar patterns of activity in Broca's area, a finding most compatible with the presence of mirror neurons in the human Broca's area.

Mirror neurons show activation both in response to the execution of an action and to the observation of an action. In the human brain, we have seen that Broca's area is part of the putative mirror neuron system and has similar properties: it is active during perception and production of complex actions. Could it be that these two areas have a common evolutionary ancestor? Probably yes: there is a wide agreement that area F5 finds its homologue either in BA 44, 45 or 6. There is, however, less agreement about where exactly in these three areas it is (Amunts et al., 1999; von Bonin and Bailey, 1947, 1961; Campbell, 1905; Grèzes et al., 2003; Grèzes and Decety, 2001; Morin and Grèzes, 2008; Passingham, 1993, 1981; Petrides, 2006; Petrides and Pandya, 1994; Rizzolatti and Arbib, 1998).

The discovery of mirror neurons has led to the idea that we understand, at least in part, the goal-directed actions of others such as grasping and manipulating objects by activating our own motor and somatosensory representations of similar actions (Buccino et al., 2001, 2004; Gallese and Goldman, 1998; Gazzola et al., 2006, 2007a; Hamzei et al., 2003; Heiser et al., 2003; Iacoboni et al., 2005; Keysers and Gazzola, 2006; Keysers et al., 2003; Kilner et al., 2007; Nishitani and Hari, 2000).

## 2.8 BROCA'S AREA: BETWEEN LANGUAGE AND ACTION

In the previous paragraphs, we reported a series of studies (both lesion and functional) that show that BA 44/45 and BA 6 are critical brain areas underlying language production and comprehension as well as action execution and perception. Several studies addressed the question whether action and language share a common functional architecture in the brain.

Within the linguistic domain, one well-known phenomenon observed in Broca's aphasia is the noun-verb dissociation. Broca's aphasic speakers are reported to comprehend (and produce) nouns better than verbs (Miceli et al., 1984; Zingeser and Berndt, 1990; Rossi and Bastiaanse, 2008). This observation gave rise to a large number of studies investigating the causes of this dissociation. The first careful anatomical

2 A voxel is a volume element that constitutes the building blocks of a 3D MR image of the brain. A voxel is analogous to a pixel in a 2D image

study in an agrammatic patient who showed a selective impairment in action naming was conducted by Damasio and Tranel (1993), who described that the patient presented a lesion in the left pre-motor frontal cortex. Saygin et al. (2004) reported an action comprehension study with 29 aphasic patients. Patients were tested with a comprehension task involving the process of actions presented visually (with a relevant drawing - pantomime) and linguistically (with a sentence). Results showed that patients were impaired in the comprehension of both modes of presentation. Arévalo et al. (2007) tested 21 aphasic speakers and a control group. Participants had to name, read or repeat single words, which were nouns or verbs. Behavioral results showed that both aphasics and non-brain-damaged speakers were less accurate in naming verbs, but a cross-item analysis revealed that the crucial factor that influenced the performance was 'manipulability', and this was true across category (both for verbs and for nouns).

These studies importantly confirm that people with damage in language areas (BA 44/45) show difficulties in the comprehension of both language and pantomime actions which indicates that similar brain areas are recruited for both tasks, bringing evidence for a convergence between the areas that are important for language and pantomime processing.

Hamzei et al. (2003) explicitly addressed the question of whether action recognition and language production share a common functional architecture. They performed an experiment in which they instructed participants to either recognize an action shown in a picture or to silently verbalize an action verb written on the screen. They found an overlap between activation of the language and the action task in the Broca's area on a group level. On a single subject level, however, no overlap was found and no consistent spatial pattern could be detected between the two activation peaks. This indicates that there seems to be no functional subdivision for language and action in Broca's area.

The observation that language and action share common neural substrates opens the question of whether this occurs as a coincidence or whether this is the base for advocating a closer relation between the two systems.

There exist a number of speculative ideas about this. The first one is represented by the 'Motor Theory of Speech Perception' (Liberman and Mattingly, 1985) which states that we understand speech by perceiving the phonetic information as intended gestures of the sender, represented in the brain as motor commands. Galantucci et al. (2006) reviewed this theory and gathered evidence for the claims that perceiving speech is perceiving gestures and that the motor system is recruited for this. Another idea for why both action and language perception and production have overlapping brain substrates is the hypothesis of 'embodied semantics', which claims that language comprehension stems from the internal referring to the actions that are conveyed by the language. This theory is supported by studies showing that listening to action-

related sentences activates the motor-related areas in the brain (Aziz-Zadeh et al., 2006; Hauk and Pulvermüller, 2004; Tettamanti et al., 2005). The third idea argues that language evolution originated from hand gestures, which is the reason they are represented in the same region in the brain (Rizzolatti and Arbib, 1998).

*Acknowledgements*

Eleonora Rossi and Marleen Schippers wrote the manuscript jointly, with equal contribution. Christian Keysers contributed to the outline of the chapter and provided critical comments.

The work was supported by a VIDI grant of the Dutch science foundation (N.W.O.), a Marie Currie Excellence grant of the European commission to Christian Keysers, and a Marica De Vincenzi grant to Eleonora Rossi.

# BIBLIOGRAPHY


Amunts, K., Schleicher, A., Bürgel, U., Mohlberg, H., Uylings, H. B., and Zilles, K. (1999). Broca's region revisited: cytoarchitecture and intersubject variability. *The Journal of comparative neurology*, 412(2):319–41.

Arévalo, A., Perani, D., Cappa, S., Butler, A., and Bates, E. (2007). Action and object processing in aphasia: From nouns and verbs to the effect of manipulability. *Brain and Language*, 100(1):79–94.

Aziz-Zadeh, L., Wilson, S. M., Rizzolatti, G., and Iacoboni, M. (2006). Congruent embodied representations for visually presented actions and linguistic phrases describing actions. *Current biology*, 16(18):1818–23.

Bates, E., Wulfeck, B., and MacWhinney, B. (1991). Cross-linguistic research in aphasia: an overview. *Brain and Language*, 41(2):123–48.

Ben-Shachar, M., Hendler, T., Kahn, I., Ben-Bashat, D., and Grodzinsky, Y. (2003). The neural reality of syntactic transformations: evidence from functional magnetic resonance imaging. *Psychological Science*, 14(5):433–40.

Ben-Shachar, M., Palti, D., and Grodzinsky, Y. (2004). Neural correlates of syntactic movement: converging evidence from two fmri experiments. *Neuroimage*, 21(4):1320–36.

Binkofski, F. and Buccino, G. (2004). Motor functions of the broca's region. *Brain and Language*, 89(2):362–369.

Binkofski, F., Buccino, G., Stephan, K. E., Rizzolatti, G., Seitz, R. J., and Freund, H.-J. (1999). A parieto-premotor network for object manipulation: Evidence from neuroimaging. *Experimental Brain Research*, 128(1-2):210–213.

Binkofski, F., Kunesch, E., Classen, J., Seitz, R. J., and Freund, H. J. (2001). Tactile apraxia: unimodal apractic disorder of tactile object exploration associated with parietal lobe lesions. *Brain*, 124(Pt 1):132–44.

Buccino, G., Binkofski, F., Fink, G. R., Fadiga, L., Fogassi, L., Gallese, V., Seitz, R. J., Zilles, K., Rizzolatti, G., and Freund, H.-J. (2001). Action observation activates premotor and parietal areas in a somatosopic manner: An fmri study. *The European Journal of Neuroscience*, 13:400–404.

Buccino, G., Binkofski, F., and Riggio, L. (2004). The mirror neuron system and action recognition. *Brain and Language*, 89(2):370–376.

Cabanis, Iba-Zizen, M., Abelanet, R., Monod-Broca, P., and Signoret, J. L. (1994). "tan-tan" the first paul broca's patient with "aphemia" (1861): Ct (1979), & mri (1994) of the brain. *4th refresher course of the ESNR: language and the aphasias. Nancy: European Society of Neuroradiology*, 4:9–22.

Campbell, A. (1905). Histological studies on the localisation of cerebral function. *Cambridge University Press*.

Caramazza, A. and Zurif, E. B. (1976). Dissociation of algorithmic and heuristic processes in language comprehension: evidence from aphasia. *Brain and Language*, 3(4):572–82.

Castaigne, P., Lhermitte, F., Signoret, J. L., and Abelanet, R. (1980). Description and scanographic study of leborgne's brain. broca's discovery. *Revue neurologique*, 136(10):563–83.

Damasio, A. R. and Tranel, D. (1993). Nouns and verbs are retrieved with differently distributed neural systems. *Proceedings of the National Academy of Sciences of the United States of America*, 90(11):4957–60.

Dapretto, M. and Bookheimer, S. Y. (1999). Form and content: dissociating syntax and semantics in sentence comprehension. *Neuron*, 24(2):427–32.

DeBleser, R., Bayer, J., and Luzzatti, C. (1996). Linguistic theory and morphosyntactic impairments in german and italian aphasics. *Journal of Neurolinguistics*, 9(3):175–185.

Decety, J., Grèzes, J., Costes, N., Perani, D., Jeannerod, M., Procyk, E., Grassi, F., and Fazio, F. (1997). Brain activity during observation of actions. influence of action content and subject's strategy. *Brain*, 120 (Pt 10):1763–1777.

Decety, J., Perani, D., Jeannerod, M., Bettinardi, V., Tadary, B., Woods, R., Mazziotta, J. C., and Fazio, F. (1994). Mapping motor representations with positron emission tomography. *Nature*, 371(6498):600–2.

Dettmers, C., Liepert, J., Hamzei, F., Binkofski, F., and Weiller, C. (2003). Läsion im ventrolateralen prämotorischen kortex beeinträchtigt die greiffunktion. *Aktuelle Neurologie*, 30:247–255.

Dronkers, N. F., Plaisant, O., Iba-Zizen, M. T., and Cabanis, E. A. (2007). Paul broca's historic cases: high resolution mr imaging of the brains of leborgne and lelong. *Brain*, 130(Pt 5):1432–41.

Ehrsson, H. H., Fagergren, A., Jonsson, T., Westling, G., Johansson, R. S., and Forssberg, H. (2000). Cortical activity in precision- versus power-grip tasks: an fmri study. *Journal of Neurophysiology*, 83(1):528–36.

Etzel, J. A., Gazzola, V., and Keysers, C. (2008). Testing simulation theory with cross-modal multivariate classification of fmri data. *PLoS ONE*, 3(11):e3690.

Fiebach, C. J., Schlesewsky, M., and Friederici, A. D. (2001). Syntactic working memory and the establishment of filler-gap dependencies: insights from erps and fmri. *Journal of Psycholinguistic Research*, 30(3):321–38.

Fogassi, L., Ferrari, P. F., Gesierich, B., Rozzi, S., Chersi, F., and Rizzolatti, G. (2005). Parietal lobe: from action organization to intention understanding. *Science*, 308(5722):662–7.

Förster, O. (1936). The motor cortex in man in the light of hughlings jackson's doctrines. *Brain*, 59:135–159.

Fridman, E. A., Immisch, I., Hanakawa, T., Bohlhalter, S., Waldvogel, D., Kansaku, K., Wheaton, L., Wu, T., and Hallett, M. (2006). The role of the dorsal stream for gesture production. *Neuroimage*, 29(2):417–28.

Galantucci, B., Fowler, C. A., and Turvey, M. T. (2006). The motor theory of speech perception reviewed. *Psychonomic Bulletin & Review*, 13(3):361–77.

Gallese, V., Fadiga, L., Fogassi, L., and Rizzolatti, G. (1996). Action recognition in the premotor cortex. *Brain*, 119(2):593–609.

Gallese, V. and Goldman, A. (1998). Mirror neurons and the simulation theory of mind-reading. *Trends in Cognitive Sciences*, 12:493–501.

Gazzola, V., Aziz-Zadeh, L., and Keysers, C. (2006). Empathy and the somatotopic auditory mirror system in humans. *Current biology*, 16(18):1824–9.

Gazzola, V. and Keysers, C. (2008). The observation and execution of actions share motor and somatosensory voxels in all tested subjects: Single-subject analyses of unsmoothed fmri data. *Cerebral Cortex*, 19(6):1239–1255.

Gazzola, V., Rizzolatti, G., Wicker, B., and Keysers, C. (2007a). The anthropomorphic brain: The mirror neuron system responds to human and robotic actions. *Neuroimage*, 35:1674–1684.

Gazzola, V., van der Worp, H., Mulder, T., Wicker, B., Rizzolatti, G., and Keysers, C. (2007b). Aplasics born without hands mirror the goal of hand actions with their feet. *Current biology*, 17(14):1235–40.

Geschwind, N. (1972). Language and the brain. *Sci Am*, 226(4):76–83.

Grafton, S. T., Arbib, M. A., Fadiga, L., and Rizzolatti, G. (1996). Localization of grasp representations in humans by positron emission tomography. 2. observation compared with imagination. *Experimental Brain Research*, 112(1):103–111.

Grèzes, J., Armony, J., Rowe, J., and Passingham, R. E. (2003). Activations related to mirror and canonical neurones in the human brain: An fmri study. *Neuroimage*, 18:928–937.

Grèzes, J. and Decety, J. (2001). Functional anatomy of execution, mental simulation, observation, and verb generation of actions: a meta-analysis. *Human Brain Mapping*, 12(1):1–19.

Grodzinsky, Y. (1995). Trace deletion, theta-roles, and cognitive strategies. *Brain and Language*, 51(3):469–97.

Grodzinsky, Y. (2000). The neurology of syntax: language use without broca's area. *Behavioral and Brain Sciences*, 23(1):1–21; discussion 21–71.

Haller, S., Radue, E. W., Erb, M., Grodd, W., and Kircher, T. (2005). Overt sentence production in event-related fmri. *Neuropsychology*, 43(5):807–14.

Hamzei, F., Rijntjes, M., Dettmers, C., Glauche, V., Weiller, C., and Büchel, C. (2003). The human action recognition system and its relationship to broca's area: an fmri study. *Neuroimage*, 19(3):637–44.

Harrington, D. L., Haaland, K. Y., and Knight, R. T. (1998). Cortical networks underlying mechanisms of time perception. *The Journal of Neuroscience*, 18(3):1085–95.

Haslinger, B., Erhard, P., Weilke, F., Ceballos-Baumann, A. O., Bartenstein, P., von Einsiedel, H. G., Schwaiger, M., Conrad, B., and Boecker, H. (2002). The role of lateral premotor-cerebellar-parietal circuits in motor sequence control: a parametric fmri study. *Cognitive Brain Research*, 13(2):159–68.

Hauk, O. and Pulvermüller, F. (2004). Neurophysiological distinction of action words in the fronto-central cortex. *Human Brain Mapping*, 21:191–201.

Heiser, M., Iacoboni, M., Maeda, F., Marcus, J., and Mazziotta, J. C. (2003). The essential role of broca's area in imitation. *The European Journal of Neuroscience*, 17(5):1123–8.

Iacoboni, M., Molnar-Szakacs, I., Gallese, V., Buccino, G., Mazziotta, J. C., and Rizzolatti, G. (2005). Grasping the intentions of others with one's own mirror neuron system. *PLoS Biology*, 3(3):e79.

Indefrey, P., Brown, C., Hagoort, P., Herzog, H., and Sach, M. (1997). A pet study of cerebral activation patterns induced by verb inflection. *Neuroimage*, 5:S548.

Indefrey, P., Brown, C. M., Hellwig, F., Amunts, K., Herzog, H., Seitz, R. J., and Hagoort, P. (2001). A neural correlate of syntactic encoding during speech production. *Proceedings of the National Academy of Sciences of the United States of America*, 98(10):5933–6.

Jaeger, J., Lockwood, A., Kemmerer, D., van Valin, R., Murphy, B., and Khalak, H. (1996). A positron emission tomography study or regular and irregular verb morphology in english. *Language*, 72:451–497.

Keysers, C. and Gazzola, V. (2006). Towards a unifying neural theory of social cognition. *Progress in Brain Research*, 156:379–401.

Keysers, C., Kohler, E., Umiltà, M., Nanetti, L., Fogassi, L., and Gallese, V. (2003). Audiovisual mirror neurons and action recognition. *Experimental Brain Research*, 153(4):628–636.

Kilner, J. M., Friston, K. J., and Frith, C. D. (2007). The mirror-neuron system: a bayesian perspective. *Neuroreport*, 18(6):619–23.

Kohler, E., Keysers, C., Umiltà, M., Fogassi, L., Gallese, V., and Rizzolatti, G. (2002). Hearing sounds, understanding actions: Action representation in mirror neurons. *Science*, 297(5582):846–849.

Koski, L., Iacoboni, M., Dubeau, M.-C., Woods, R., and Mazziotta, J. C. (2003). Modulation of cortical activity during different imitative behaviors. *Journal of Neurophysiology*, 89(1):460–471.

Krams, M., Rushworth, M., Deiber, M., Frackowiak, R. S., and Passingham, R. E. (1998). The preparation, execution and suppression of copied movements in the human brain. *Experimental Brain Research*, 120:386–398.

Leiguarda, R. C. and Marsden, C. D. (2000). Limb apraxias: higher-order disorders of sensorimotor integration. *Brain*, 123 (Pt 5):860–79.

Liberman, A. and Mattingly, I. (1985). The motor theory of speech perception revised*. *Cognition*, 21:1–36.

Luzzatti, C. and Bleser, R. D. (1996). Morphological processing in italian agrammatic speakers: eight experiments in lexical morphology. *Brain and Language*, 54(1):26–74.

Luzzatti, C., Toraldo, A., Guasti, M., Ghirardi, G., Lorenzi, L., and Guarnaschelli, C. (2001). Comprehension of reversible active and passive sentences in agrammatism. *Aphasiology*, 15(5):419–441.

Makuuchi, M. (2005). Is broca's area crucial for imitation? *Cerebral Cortex*, 15(5):563–70.

Meyer, M., Friederici, A. D., and von Cramon, D. Y. (2000). Neurocognition of auditory sentence comprehension: event related fmri reveals sensitivity to syntactic violations and task demands. *Cognitive Brain Research*, 9(1):19–33.

Miceli, G., Silveri, M. C., Villa, G., and Caramazza, A. (1984). On the basis for the agrammatic's difficulty in producing main verbs. *Cortex*, 20(2):207–20.

Morin, O. and Grèzes, J. (2008). What is "mirror" in the premotor cortex? a review. *Neurophysiologie clinique = Clinical neurophysiology*, 38(3):189–95.

Nishitani, N. and Hari, R. (2000). Temporal dynamics of cortical representation for action. *Proceedings of the National Academy of Sciences of the United States of America*, 97(2):913–8.

Passingham, R. (1993). The frontal lobes and voluntary action. *Oxford University Press*.

Passingham, R. E. (1981). Broca's area and the origins of human vocal skill. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences*, 292(1057):167–75.

Pazzaglia, M., Smania, N., Corato, E., and Aglioti, S. M. (2008). Neural underpinnings of gesture discrimination in patients with limb apraxia. *The Journal of Neuroscience*, 28(12):3030–41.

Pellegrino, G., Fadiga, L., Fogassi, L., Gallese, V., and Rizzolatti, G. (1992). Understanding motor events: A neurophysiological study. *Experimental Brain Research*, 91:176–180.

Petrides, M. (2006). Broca's area in the human and nonhuman primate brain. In Grodzinsky, Y., editor, *Broca's region*. Oxford University Press.

Petrides, M. and Pandya, D. (1994). Comparative cytoarchitectonic analysis of the human and the macaque frontal cortex. *Handbook of neuropsychology*, 9:17–58.

Pick, A. (1898). Über agrammatismus als folge cerebraler herderkrankungen; ein beitrag zur lehre vom verhältnis der worttaubheit. *Beiträge zur Pathologie und pathologischen Anatomie des ZentralNervensystems*, 9:123–133.

Pick, A. (1913). *Die agrammatischen Sprachstörungen. Studien zur psychologischen Grundlegung der Aphasielehre. Teil I.* Berlino: Springer.

Pobric, G. and de C Hamilton, A. F. (2006). Action understanding requires the left inferior frontal cortex. *Current biology*, 16(5):524–9.

Renzi, E. D., Motti, F., and Nichelli, P. (1980). Imitating gestures - quantitative approach to ideomotor apraxia. *Archives of Neurology*, 37(1):6–10.

Rizzolatti, G. and Arbib, M. A. (1998). Language within our grasp. *Trends in Neurosciences*, 21(5):188–194.

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

Rizzolatti, G., Fadiga, L., Gallese, V., and Fogassi, L. (1996a). Premotor cortex and the recognition of motor actions. *Cognitive Brain Research*, 3(18):131–141.

Rizzolatti, G., Fadiga, L., Matelli, M., Bettinardi, V., Paulesu, E., Perani, D., and Fazio, F. (1996b). Localization of grasp representations in humans by pet: 1. observation versus execution. *Experimental Brain Research*, 111(2):246–252.

Rossi, E. and Bastiaanse, R. (2008). Spontaneous speech in italian agrammatic aphasia: A focus on variability and verb production. *Aphasiology*, 22:347–362.

Saygin, A., Wilson, S. M., Dronkers, N. F., and Bates, E. (2004). Action comprehension in aphasia: linguistic and non-linguistic deficits and their lesion correlates. *Neuropsychology*, 42(13):1788–804.

Schubotz, R. I. and von Cramon, D. Y. (2001). Functional organization of the lateral premotor cortex: fmri reveals different regions activated by anticipation of object properties, location and speed. *Cognitive Brain Research*, 11(1):97–112.

Schubotz, R. I. and von Cramon, D. Y. (2002a). A blueprint for target motion: fmri reveals perceived sequential complexity to modulate premotor cortex. *Neuroimage*, 16(4):920–35.

Schubotz, R. I. and von Cramon, D. Y. (2002b). Dynamic patterns make the premotor cortex interested in objects: influence of stimulus and task revealed by fmri. *Cognitive Brain Research*, 14(3):357–69.

Schubotz, R. I. and von Cramon, D. Y. (2002c). Predicting perceptual events activates corresponding motor schemes in lateral premotor cortex: an fmri study. *Neuroimage*, 15(4):787–96.

Schubotz, R. I. and von Cramon, D. Y. (2003). Functional-anatomical concepts of human premotor cortex: evidence from fmri and pet studies. *Neuroimage*, 20 Suppl 1:S120–31.

Seitz, R. J. and Roland, P. (1992). Learning of sequential finger movements in man: A combined kinematic and positron emission tomography (pet) study. *The European Journal of Neuroscience*, 4(2):154–165.

Shapiro, L. and Levine, B. (1990). Verb processing during sentence comprehension in aphasia. *Brain and Language*, 38:21–47.

Tettamanti, M., Buccino, G., Saccuman, M. C., Gallese, V., Danna, M., Scifo, P., Fazio, F., Rizzolatti, G., Cappa, S. F., and Perani, D. (2005). Listening to action-related sentences activates fronto-parietal motor circuits. *Journal of Cognitive Neuroscience*, 17(2):273–81.

Tranel, D., Kemmerer, D., Adolphs, R., and Damasio, H. (2003). Neural correlates of conceptual knowledge for actions. *Cognitive Neuropsychology*, 20:409–432.

von Bonin, G. and Bailey, P. (1947). *The neocortex of Macaca mulatta*. Illinois: The University of Illinois Press,.

von Bonin, G. and Bailey, P. (1961). Pattern of the cerebral isocortex. In Hofer, H., Schulz, A., and Starck, D., editors, *Primatologia:*, pages 1–42 Part 2, Lieferung 10. Karger, Basel.

Zingeser, L. B. and Berndt, R. S. (1990). Retrieval of nouns and verbs in agrammatism and anomia. *Brain and Language*, 39(1):14–32.

# 3

# PLAYING CHARADES IN THE FMRI: ARE MIRROR AND/OR MENTALIZING AREAS INVOLVED IN GESTURAL COMMUNICATION?

*Published as:* Schippers MB, Gazzola V, Goebel R, Keysers C (2009) Playing Charades in the fMRI: Are Mirror and/or Mentalizing Areas Involved in Gestural Communication? *PLoS ONE* 4(8): e6801. doi:10.1371/journal.pone.0006801

ABSTRACT

Communication is an important aspect of human life, allowing us to powerfully coordinate our behaviour with that of others. Boiled down to its mere essentials, communication entails transferring a mental content from one brain to another. Spoken language obviously plays an important role in communication between human individuals. Manual gestures however often aid the semantic interpretation of the spoken message, and gestures may have played a central role in the earlier evolution of communication. Here we used the social game of charades to investigate the neural basis of gestural communication by having participants produce and interpret meaningful gestures while their brain activity was measured using functional magnetic resonance imaging. While participants decoded observed gestures, the putative mirror neuron system (pMNS: premotor, parietal and posterior mid-temporal cortex), associated with motor simulation, and the temporo-parietal junction (TPJ), associated with mentalizing and agency attribution, were significantly recruited. Of these areas only the pMNS was recruited during the production of gestures. This suggests that gestural communication relies on a combination of simulation and, during decoding, mentalizing/agency attribution brain areas. Comparing the decoding of gestures with a condition in which participants viewed the same gestures with an instruction not to interpret the gestures showed that although parts of the pMNS responded more strongly during active decoding, most of the pMNS and the TPJ did not show such significant task effects. This suggests that the mere observation of gestures recruits most of the system involved in voluntary interpretation.

## 3.1 INTRODUCTION

Communication is an important aspect of human life, allowing us to powerfully coordinate our behaviour with that of others. Boiled down to its mere essentials,

communication entails transferring a mental content from one brain to another. Spoken language obviously plays an important role in communication between human individuals. Manual gestures however often aid the semantic interpretation of the spoken message (Iverson and Goldin-Meadow, 1998; Kendon, 1994; McNeill, 1992; Melinger and Levelt, 2004; Willems and Hagoort, 2007), and gestures may have played a central role in the earlier evolution of communication (Arbib, 2008; Gentilucci and Corballis, 2006; Rizzolatti and Arbib, 1998). Therefore we will examine here the neural substrates of gestural communication in humans. Although this question has received less attention in the field of neuroscience than spoken language, two potentially complementary processes have been implicated in the perception and/or production of gestures: simulation and mentalizing (de Lange et al., 2008; Keysers and Gazzola, 2007; Thioux et al., 2008).

The concept of simulation has received a surge of popularity since the discovery of mirror neurons in macaque monkeys (Ferrari et al., 2003; Fogassi et al., 2005; Fujii et al., 2007; Gallese et al., 1996; Keysers et al., 2003; Kohler et al., 2002; Rizzolatti et al., 1996; Umiltà et al., 2001). These neurons are active not only while the monkey performs an action (e.g. shelling a peanut), but also while the monkey sees or hears a similar action. Mirror neurons have been found in the ventral premotor and inferior parietal cortex of the monkey. However, it remains unclear whether other regions of the monkey brain contain mirror neurons for actions, because extensive single cell recording during both action execution and observation have so far not been performed outside of the premotor and inferior parietal lobule. Evidence for a similar system in humans has been derived from neuroimaging and transcranial magnetic stimulation studies (Blakemore and Decety, 2001; Buccino et al., 2001; Decety et al., 1997; Fadiga et al., 1995; Gazzola and Keysers, 2008; Grafton et al., 1996; Iacoboni et al., 1999; Jeannerod, 2001; Rizzolatti and Craighero, 2004), with the former showing that a network of areas is active both while people perform actions in the scanner and while they view or hear other people's actions. In humans, this system seems to include the dorsal premotor, somatosensory, cerebellar and posterior temporal cortex in addition to the ventral premotor, inferior frontal gyrus and inferior parietal lobule (Chong et al., 2008; Gazzola and Keysers, 2008). These are the likely homologues of the aforementioned regions of the monkey (Nelissen et al., 2005; Petrides et al., 2005). This extended set of areas can be called the putative Mirror Neuron System (pMNS) in order to emphasize that if a voxel in an fMRI experiment is involved in both execution and observation, the neurons within these voxels can, but do not have to, be mirror neurons (Gazzola and Keysers, 2008; Gazzola et al., 2007b): different populations of neurons within the same voxel could play the lead role during observation and execution. This caveat means that functional neuroimaging findings have to be interpreted with care: the fact that a region involved in action observation and execution is recruited during the processing of stimuli X might be suggestive

of the fact that processing X involves 'simulation' (i.e. the recruitment of motor programs 'as if' the participant were producing these gestures him/herself) but it is not a guarantee that processing X truly depends on mirror neurons or simulation (Hickok, 2009). Neuroimaging therefore needs to ask questions in terms of brain regions (are regions of the pMNS involved?), and not in terms of cognitive processes involved (is simulation involved?): the former can be empirically measured using neuroimaging, the latter only tentatively suggested (Poldrack, 2006).

The discovery of mirror neurons has lead to the idea that we understand, at least in part, the goal-directed actions of others such as grasping and manipulating objects by activating our own motor and somatosensory representations of similar actions (Buccino et al., 2001, 2004; Buxbaum et al., 2005; Gallese et al., 1996; Gallese and Goldman, 1998; Gazzola et al., 2006, 2007a; Hamzei et al., 2003; Heiser et al., 2003; Iacoboni et al., 2005; Keysers and Gazzola, 2006; Keysers et al., 2003; Kilner et al., 2007; Nishitani and Hari, 2000; Rizzolatti et al., 1996; Umiltà et al., 2001) as if we had performed similar actions. This 'as if' component is why this process is called simulation. It seems that simulation occurs simultaneously at different levels of representations (Thioux et al., 2008): strictly and broadly congruent mirror neurons in the monkey for instance represent details of an action and the goal of an action, respectively and simultaneously (Gallese et al., 1996), and experiments in human support the notion that both the details (TMS) and goals (Gazzola et al., 2007a,b) of actions are simulated. Whether the same system is involved in perceiving communicative gestures has been much less investigated.

Several lesion studies have investigated the neural basis of gesture production and perception in the context of apraxia. This is a disorder in which patients have difficulty with the control of action, including impairment in the production of gestures. In ideational apraxia, patients have preserved basic motor skills, but if asked to mimic the use of tools (e.g. show me how you would use a hammer to hammer a nail), they fail to produce the correct actions (Ochipa et al., 1989). The ability to mimic is therefore traditionally used as a localizer for areas related to apraxia (Mozaz et al., 2002). These studies have shown that the normal production of gestures requires an intact left posterior parietal lobe, including the parietal node of the pMNS (Choi et al., 2001; Fridman et al., 2006; Hermsdörfer et al., 2001; Higuchi et al., 2007; Lotze et al., 2006; Moll et al., 2000; Nair et al., 2003; Ohgami et al., 2004). More recently, Montgomery et al. (2007) use a functional neuroimaging study to show that observing and producing communicative hand gestures activated the superior temporal sulcus, inferior parietal lobule and frontal operculum - a set of regions that corresponds to those of the pMNS. A limitation of this well controlled study is the fact that the participants had no genuine communicative intent: they produced pre-trained gestures in response to words (e.g. "thumbs up") in the production condition, and passively observed stereotyped short movie clips of hand gestures in the observation condition. In

addition, the authors intermixed imitation trials with passive observation trials. This may have lead to activations in motor production areas during gesture observation trials simply as a covert rehearsal of the motor programs that will later be needed for imitation. Overall, this task may therefore differ in important ways from the real life processes involved. For example, if one is in a foreign country, does not speak the language, and has only gestures to ask where to find a good restaurant. Would such a situation also primarily recruit the pMNS? Would other regions become important, including those involved in asking yourself what the other person is thinking, i.e. mentalizing areas?

A set of brain regions has been implicated in such reflection about the mental state of others. These areas include the medial prefrontal cortex (mPFC, in particular the paracingulate gyrus) and the temporo-parietal junction (TPJ) (Brunet et al., 2000; Castelli et al., 2000; Ciaramidaro et al., 2007; Fletcher et al., 1995; Frith and Frith, 2006; Gallagher and Frith, 2003; Gallagher et al., 2000, 2002; Hampton et al., 2008; McCabe et al., 2001; Saxe and Kanwisher, 2003; Sommer et al., 2007; Vogeley et al., 2001; Walter et al., 2004).Gallagher and Frith (2004) compared the recognition of hand gestures expressing internal states (e.g. I feel cold) with those expressing a command (e.g. come here!). They additionally contrasted a recognition condition (was the gesture positive?) against an observation condition (which hand moved higher in the movie?). In particular, they report in the results and their Table 4 that the left anterior paracingulate cortex (putative BA32), thought to be a key node of the putative 'theory of mind' network (pToM area) appeared in an interaction contrast (recognizing expressive gestures - observing expressive gestures - recognizing orders + observing orders), and interpreted this finding as evidence for ToM involvement in interpreting gestures that express inner states. From the evidence presented in the report however, this interpretation is problematic, as they also report in the results and their Table 3, that the left anterior paracingulate cortex (putative BA32) is more active while observing gestures compared to recognizing them. While it is uncertain from the tables alone whether overlapping regions of the paracingulate cortex were present in these two contrasts, the paracingulate cortex was absent from the contrast recognizing - observing. This would be difficult to reconcile with the area being responsible for recognition. The involvement of ToM regions in gesture recognition therefore remains uncertain. In addition, although the TPJ is reliably recruited by tasks requiring mentalizing (Fletcher et al., 1995; Gallagher and Frith, 2003; Saxe and Kanwisher, 2003; Sommer et al., 2007), it is unlikely that this region specializes in attributing mental states to others: it is likely that it serves domain general functions relating to attention (Mitchell, 2008) and/or comparing sensory input with motor commands (Decety and Lamm, 2007) which happen also to be important during mental state attribution.

The study described here explicitly investigates the role of both the pMNS and

pToM areas by pioneering the use of a well-established gestural communication task into the field of neuroscience: the game of 'charades'. We recorded brain activity while (romantically involved) couples played this game with each other. One partner would first be scanned while gesturing an action or object into a camera in the knowledge that his partner would later need to guess the action/object based on his recorded gestures. The other partner was to be scanned while decoding the gestures. The roles were then reversed. This allowed us to measure brain activity while people invent and execute gestures suitable to communicate a complex concept to another person, and while another person is decoding these gestures to guess the original concept. In addition, we examined if the brain activity recorded during this natural form of communication was specific for a communicative setting. We replayed the movies of their partner's gestures to each participant on a separate day, but this time, did not ask them to guess what their partner was trying to tell them. All participants reported finding the game very motivating, and experienced the experiment as a genuine and spontaneous form of communication.

Based on the idea that the pMNS might map the communicative actions of others onto the programs for producing similar actions, we hypothesized that parts of the areas involved in generating gestures would also become activated during the observation of communicative actions. To examine if this system overlaps with the pMNS for goal-directed actions, we examined if the pMNS as defined in previous experiments (Gazzola et al., 2007a) becomes active both during gesture production and observation. Furthermore, several studies have shown the involvement of the TPJ and mPFC in tasks where people have to explicitly infer the mental states of another person. We therefore examined whether these pToM areas are involved during the charades game. Activity during gesture production may reflect a theory-of-mind of how the partner might interpret the gestures, and activity during gesture interpretation may reflect a theory-of-mind of what the partner might have meant while generating the gestures. pMNS and pToM areas could complement each other during the charades task (de Lange et al., 2008; Keysers and Gazzola, 2007; Thioux et al., 2008). The pMNS areas have been shown to be relatively stimulus driven independent of the task (e.g. de Lange et al., 2008; van der Gaag et al., 2007), while pToM areas seem more recruited during tasks that explicitly direct peoples minds to the mental states of others (de Lange et al., 2008). This line of reasoning would predict that pMNS areas would respond during the charades game and the control condition because they involved similar stimuli and motor actions. However, the pToM areas might respond during the charades game because this encourages mental state attribution but not during the control condition, which does not.

## 3.2 MATERIALS AND METHODS

### *Participants*

Twelve couples (total: 24 participants) were scanned while playing the game charades. The mean age of the participants was 27.5 ± 3.8 years. Each couple consisted of a man and a woman involved in a romantic relationship for at least 6 months. As in previous studies on emotional empathy (Singer et al., 2004), we included this criterion not to study romantic relations specifically but to maximise the social relevance of this experiment because we expected couples to be more motivated, more at ease, and to have a better or faster understanding of each other's gestures than a strangers do. Participants were asked to fill out a questionnaire about their neurological and medical history including whether they had metal objects in their body. This is a standard procedure to ensure the safety of the participants whilst in the scanner. Participants were also asked not to drink coffee before scanning commenced. The participants freely consented to participating in the study by signing an informed consent form and were scaled for their right-handedness on the Edinburgh Righthandedness scale (Oldfield, 1971). This entire study was approved by the Medical Ethics Committee of the University Medical Center Groningen (2007/080).

### *Task / Experimental Design*

The experiment consisted of two separate sessions on different days. In the first session, the couple was required to play the game of charades. In the second, detailed anatomical scans and a control condition were acquired. For the game of charades, participants took turns going into the scanner, alternating gesturing and guessing of words. Words were either objects (for example nutcracker, watch, pencil sharpener) or actions (for example painting, knitting, shaving, see Tab. 3.1). Each participant performed two gesture and two guess runs in which they gestured and guessed 14 words in total (7 per run). The set of words used was the same for each couple, but word order was randomized between participants. After the last gesture-session, a T1 image was acquired.

### *Gesture run*

During a gesture run, the participant was presented with a word on the screen and was instructed to communicate this word to his or her partner by means of gestures. Every word had to be gestured for 90 seconds. Prior to scanning participants were trained not to repeat the same gesture over and over again, but to keep generating

new gestures to provide their partner with multiple sources of information. The participant could see how much time he/she needed to keep gesturing by a progress bar on the screen. A fixation cross was presented for 20 s after each word, which served as our baseline. The gestures were recorded from the control room of the MR-scanner with a video camera (Sony DSR-PDX10P). After the participant had gestured seven words, he/she was taken out of the scanner and went into the waiting room, while his/her partner went into the scanner to guess what he/she had gestured. During this changeover, the experimenter cut the recording of the gestures into movies of 90s in which the participant gestured a word (see online supplementary information for an example of a gesture recording, movie S1). To ensure that the movies were cut at exactly the moment the word was presented to the gesturing participant, the stimulus computer's sound card emitted a sound at the beginning of word presentation. The output of the sound card was connected to the audio input of the video camera, thus allowing the auditory signal to serve as a marker for cutting. To minimize the amount of head motion in the participants, the upper arms of the participant were fixed to the bed by means of a Velcro strap band. This left the participants free to gesture with their lower arms, hand, and fingers, which was sufficient to ensure 86% percent correct gesture recognition.

*Guess run*

During a guess run, the participant was shown the movies that were recorded in the gesture run of their partner. The task they had to perform was to guess what their partner was trying to gesture to them. Participants were asked to consider the gestures for at least 50 seconds before committing to a specific interpretation of the gestures. This was done to ensure at least 50 seconds of data in each trial to examine the time course of activity (i.e. is brain activity in region X sustained for as long as participants are interpreting the gestures?). This was done by showing a progress

| Actions | | Objects | |
|---|---|---|---|
| peel fruit | fold | nutcracker | telephone |
| ride a bike | drive a car | pencil sharpener | winding stairs |
| shuffle cards | play the piano | pistol | ashtray |
| polish nails | squeeze fruit | electric eel | bow |
| juggle | paint | watch | handcuffs |
| knit | light fireworks | board game | glove |
| throw a snowball | shave | canoe | cork screw |

Table 3.1: Action and object words used in the charades

bar under the movie, changing from red to green after 50 seconds, indicating the beginning of the period (50-90s post stimulus onset) during which participants could decide on their interpretation of the gestures, whenever they felt confident. After the button press with which the participants indicated to be ready to respond, a multiple choice menu was presented. In this menu they had to choose the correct word from five alternatives. One of the alternatives was always 'none of the above' and the correct answer was always present in the multiple-choice menu. The correct answer was never the option 'none of the above'. This marked the end of a trial. Two consecutive trials were separated by 20 seconds of a white fixation cross against a black background, which served as our baseline.

*Passive observation run*

As a control condition for the guess run, the participants watched the movies again which they had seen during the guessing condition. This time, they were instructed not to guess what was gestured, but only to passively view them. To keep the run exactly the same as the original guess run, the movie would stop at the moment the participant during the original run had pushed the button. The same multiple-choice menu would appear and the participant had to answer again. This time, however, they had to select the word written in green letters. The green word was the correct answer. A fixation cross was presented between two consecutive trials for 20 seconds and served as our baseline.

*Data Acquisition*

Functional imaging data was recorded with a Philips 3.0T MR scanner, using gradient echo planar imaging (EPI). T2* weighted images revealed changes in blood oxygen level. Repetition time was 1.33 seconds. The whole brain was scanned in 28 (axial) slices with a thickness of 4.5mm. In the first session, a fast structural image ("fast anatomy") was acquired of the participant's brain, while in the second session an additional structural image of higher resolution was acquired. Both were structural, T1-weighted images.

*Data Analysis / Statistical Analysis*

Data were analyzed using the Statistical Parametric Mapping Software, version 2 (SPM2). EPI data were corrected for slice timing and realigned. The T1 image was co-registered to the mean EPI and segmented, the normalization parameters to normalize the gray-matter segment onto the MNI gray-matter template were determined,

and applied to all the EPI images. Normalized EPI images were then smoothed with a Gaussian kernel of 10mm. Three general linear models were estimated: one for the gesturing, one for the guessing and one for the passive observation sessions. All words, whether they were actions or objects, guessed correctly or incorrectly, were modelled together in one condition. The predictor in the gesture run consisted of the whole period during which the gesture was executed (90s). In the active guessing and passive observation runs two predictors were included in the general linear model: (a) the period in which the movie was shown until button press and (b) from button-press until the participant had given the answer. All predictors were convolved with the hemodynamic response function. Each participant's mean parameter estimates were then tested at the second level (one-sampled t-test). Activations are displayed on a mean anatomical image of all participants (see Fig. 3.1). To examine differences between object words and action words, the data was also modelled using separate predictors for the two categories but the contrasts 'guessing objects-guessing actions', and the reverse contrast, were not significant at $p<0.05$ (FDR corrected) in any voxel. Therefore only analyses using a single predictor are reported here. The same applies to the gesture analyses. To control for head motion, we included six motion parameters as covariates of no interest (translation and rotation in x, y and z directions) and excluded four participants, who moved more than the voxel size (3.5x3.5x4.5mm). Thus, the analyses and results presented in this paper are based on 20 participants.

*Comparisons Guessing vs Passive Observation*

Given that passive observation always had to be acquired after guessing, differences between these conditions could in theory be linked, amongst others, to systematic differences in the MR-signal across sessions. We examined this possibility by calculating average global maps for each participant (i.e. a contrast with ones in the last columns of the SPM design matrix for the two sessions). These maps were compared in a paired t-test. There were no significant differences at $p<0.05$ (FDR corrected).

*Localizing shared circuits*

We define shared circuits as those voxels that are active both during an execution and an observation condition. This was done by thresholding the group-level analysis of the gesturing condition (vs. passive baseline) at $p<0.001$ (uncorrected) to create a binary map (all above-threshold voxels have the value 1 and all the other have the value 0) and applying this image as a mask in the second level analysis of guessing or passive observation.

*Putative Mirror Neuron System ROIs*

The areas which together form the mirror neuron system were defined based on a previous study done in our lab with 16 participants (Gazzola et al., 2007a). In this study, healthy participants observed and performed goal-directed hand actions. The subset of areas that are active both during the execution and the observation condition form the pMNS. The areas included a section of the ventral-and dorsal premotor cortex, the parietal lobe (including Brodmann Area (BA) 2 and the cortex along the intraparietal sulcus and the supramarginal gyrus) and the middle temporal gyrus (see Fig. 3.3 for location and size of the rois).

*Putative Theory of Mind areas ROIs*

The medial prefrontal cortex and the temporo-parietal junction are considered typical theory-of-mind areas. We included both these areas in our analyses. We based the ROIs in the medial prefrontal cortex on the review article of Amodio and Frith (2006) in which different tasks are outlined that lead to activation in this area. Based on this meta-analysis, we drew our ROI in the anterior rostral medial frontal cortex. Activations in this region are associated with mentalizing, person-perception and self-knowledge. This roughly corresponds to Brodmann area 10. We used the Talairach coordinates from that article to hand-draw a quadrilateral ROI (from (-2,34,5) and (-2,26,15) to (-2,71,5) and (-2,55,44) respectively). This triangular shape started medially (at X=±2) and extended laterally 13 mm to cover the grey matter (until X=±15). To fit the ROI in the best possible way to our participants' data, we multiplied this hand drawn image with a thresheld mask ($> 0.3$) of the mean grey matter segment that was obtained through segmenting the brain of each individual participant.

In a similar fashion we defined the temporal parietal junction on the basis of coordinates mentioned in Mitchell (2008). Mitchell (2008) gives an overview of all different peak coordinates associated with the temporal parietal junction. To construct our ROI, we calculated the mean of these three coordinate-pairs ((54,-51,18), (54,-54,24), (60,-57,15)) and used this as the centre point of a sphere with a radius of 10 mm sphere. Again, we multiplied this with the mean grey matter segment to exclude out-of-brain voxels as much as possible. For the location and sizes of these regions of interest, see Figure 3.4.

*Calculating the finite impulse response for the ROIs*

For each ROI, we extracted the average BOLD response around two events of interest: the onset of a gesture and the moment the button was pushed when the word was

guessed. During guessing and passive observation 28 peri-stimulus timebins were extracted, in which each bin had the same length as the repetition time (1.33s). The signal was extracted from the period commencing 8 bins before gesture onset and continuing until 20 bins following it. The same was done for the button press, including 20 bins before and 8 bins after. During gesturing, the average BOLD response was extracted for the whole period in which the gesture was performed, starting at 8 bins before the onset and lasting for 84 bins. The MarsBar toolbox in SPM2 was used for this extraction (Brett et al., 2002). This modeling resulted in para-stimulus time histograms, which show the development of brain activity over time (see Fig. 3.3-3.4).

*Thresholding*

All final whole brain analysis results are thresheld at p<0.001 (uncorrected). Only clusters that additionally survived a false discovery rate correction at p<0.05 are reported. This means that all whole brain results presented in this manuscript survive fdr correction at p<0.05, but are presented at p<0.001 (uncorrected) because this turned out to be the most stringent of the two. Note that in the case of masking, the correction is only applied after the masking. Given that the mPFC failed to show significant activation at these thresholds, we additionally performed a small volume corrected analysis at p<0.05 within the volume defined as our mPFC ROI to challenge our negative findings. For the regions of interest analysis, we specify the significance of any difference with p<0.05. This was done for the reader to have the freedom to challenge negative findings at a permissive threshold (p<0.05), while at the same time providing more stringent evidence for the key positive results.

## 3.3 RESULTS

### 3.3.1 *Behavioural Results*

During guessing the participants were asked to consider each movie for at least 50 seconds after which they could push the button when they thought they knew what was being gestured to enter the multiple-choice menu. The average latency to response was 58 seconds. Participants were equally accurate on both categories: 82.5% of the object words were guessed correctly against 86.5% of the action words ($t(17) = -1$, $p > .33$). We did not find a significant difference between the two types of gestures, neither in terms of latency to respond (58.7s ± 6.5s for action and 60.8s ± 6.8s for object words, $t(17) = -1.16$, $p > .26$) nor in terms of accuracy (6.06 ± 0.73sd correct out of 7 action and 5.78 ± 1.11sd correct our of 7 object words, $t(17) = -1$, $p > .33$). Words that were guessed incorrectly were watched significantly longer than words

that were guessed correctly: 58s ± 5s for the 289 correct guesses versus 68s ± 12s for the 47 incorrect guesses (t(16) = -4.18, p < .0005).

### 3.3.2 *Whole Brain fMRI Results*

*Main effects of guessing*

Activation clusters during guessing compared to baseline are shown in Table 3.2 and Figure 3.1-A. Of particular interest were the clusters of activity found along the precentral gyrus (BA 6) and extending into the inferior frontal gyrus (BA 44 and 45), in the middle and superior temporal areas (including the TPJ), the primary somatosensory cortex (BA 2 in particular) and the supramarginal gyri. Inspection of the medial wall (see Figure 3.2) revealed activations in the superior medial gyrus in what Amodio and Frith (Amodio and Frith, 2006) call the posterior section of the rostral medial frontal cortex but not in the anterior section associated with theory-of-mind (our mPFC ROI). During this condition, reductions in the BOLD signal were found in the precuneus, right insula, and bilaterally the angular gyrus and the operculum (OP 1 to 4). There were no differences in activation when object words are compared with action words or vice versa (not shown).

*Main effects of passive observation*

Table 3.3 and Figure 3.1-B show activation clusters during passive observation compared to passive baseline. Clusters of activity were found in locations very similar to those during active guessing, including BA 6, 44, 45, 2, middle and superior temporal areas (including the TPJ), and supramarginal gyri. Inspection of the medial wall (see Figure 3.2) revealed activations in the superior medial gyrus and adjacent middle cingulate gyrus in what Amodio and Frith (Amodio and Frith, 2006) call the posterior section of the rostral medial frontal cortex but not in the anterior section associated with theory-of-mind (our mPFC ROI). Reductions in the BOLD signal were found in the precuneus, the caudate nucleus and two small clusters in the cerebellum.

*Main effects of gesturing*

All activation clusters during gesturing compared to a passive baseline are shown in Table 3.4 and Figure 3.1-C. Notably, clusters of activity were found in the primary, pre- and supplementary motor areas (BA 4a/p and 6), BA 44 and 45. Both inferior and superior parietal lobules were involved, together with somatosensory cortices and the middle and superior temporal gyri (including the TPJ). Inspection of the medial

wall (see Figure 3.2) revealed activations in the superior medial gyrus and adjacent middle cingulate cortex in what Amodio and Frith (Amodio and Frith, 2006) call the posterior section of the rostral medial frontal cortex but not in the anterior section associated with theory-of-mind (our mPFC ROI). Instead, the most anterior sections show evidence of reduced BOLD relative to baseline. Extensive clusters were found in the precuneus, the angular gyrus bilateral, the medial prefrontal cortex and the left temporal pole, which were more active during the baseline than during gesturing. Additional reductions in BOLD signal were found in the more posterior superior parts of BA 17 and 18 and in the right hippocampus and amygdala.

*Similarities and differences between guessing and passive observation*

The comparison of activity between guessing and passive observation is rendered more difficult by the fact that they were acquired in separate sessions, and results should be considered with care. Counterbalancing the order of acquisition would however have interfered with the aims of the experiments for two reasons. First, an instruction not to engage in active guessing would be even more difficult during a passive observation trial if participants would know that they later need to guess the meaning of the same movie. Second, capturing the neural processes involved in interpreting gestures in an ecologically plausible way would be disturbed by 'passively' viewing the movies before. Using different movies for passive observation and active guessing would not be a solution either because the stimuli might differ in important ways.

To exclude the possibility that differences in brain activity between guessing and passive observation could simply be due to systematic differences in the state of the scanner, we additionally compared the mean fMRI signal between the two sessions (using a two-sample t-test comparing the globals in the two sessions, see Methods). No region in the brain showed such an effect under a threshold of $p < 0.05$ (FDR corrected). This means that functional differences cannot be due to differences in the mean signal alone.

Two analyses were then performed to compare brain activity during the processing of the same movies during active guessing versus passive observation: one to map differences and one to map similarities between the two conditions. Areas, which were recruited to a greater extent during guessing than during passive observation were as follows: the inferior and middle temporal gyri and areas V5/MT+ bilaterally, and more anterior in the brain a cluster in BA 44. Again, inspection of the medial wall (see Figure 3.2) showed no clusters of activation in the mPFC ROI associated with theory-of-mind. Differences due to a greater involvement during passive observation than during guessing were located in the angular gyrus and the precuneus. These were

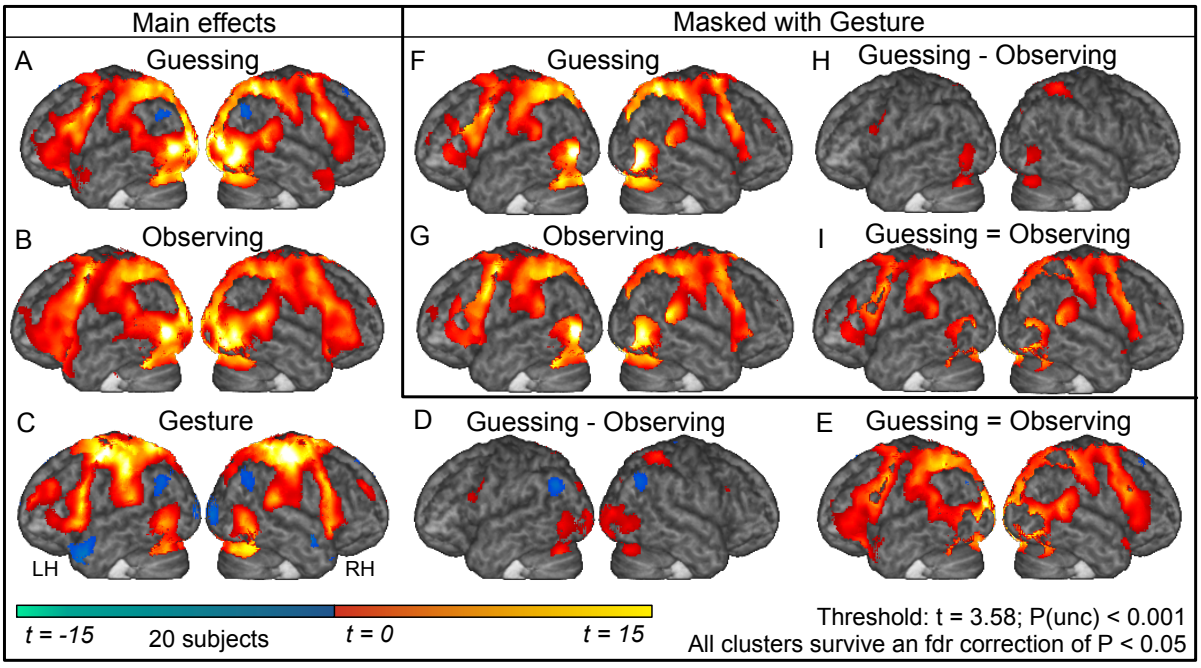



Figure 3.1: Activation maps rendered on the mean anatomy of all 20 subjects. (A-D) Main effects guessing-baseline, passive observation-baseline, gesture-baseline, guessing-passive observation. (E) Areas similarly activated during guessing and passive observation (i.e. guessing-baseline $p<0.001$ & passive observation-baseline $p<0.001$ & guessing-passive observation $p>0.001$). (F-I) A, B, D and E, each masked inclusively with C. All images are thresholded at $t = 3.58$ which corresponds to an uncorrected $p \leq 0.001$. All voxels also survive a false discovery rate at $p<0.05$.

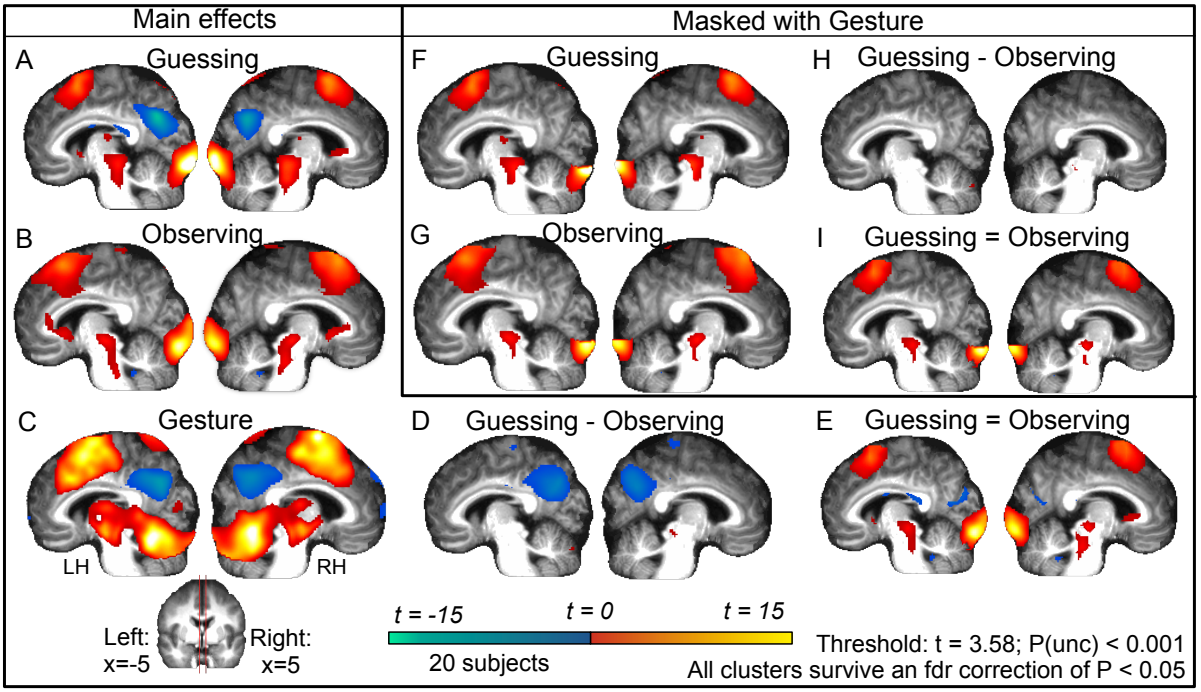



Figure 3.2: Same as Fig. 3.1, but activations are now shown on the left (x=-5) and right (x=5) medial wall of the mean anatomy of the 20 subjects

areas that were deactivated compared to the passive baseline in the main effects. A full description and visualization of the areas can be found in Table 3.5 and Figure 3.1-D. In contrast, much larger areas were recruited during both active guessing and passive observation without significant difference between these conditions. These included the precentral gyrus (BA 6) and BA 44 and 45, the somatosensory cortex (BA2), the inferior parietal lobule, and the middle and superior temporal areas. For a full description and visualization of the areas, see Table 3.6 and Figure 3.1-E.

*Guessing masked with gesturing, passive observation masked with gesturing (shared circuits)*

We defined shared circuits as voxels recruited both during the execution and the observation of gestures. Masking the activity during guessing with the activity during gesturing shows, among others, shared recruitment of the following areas: the precentral gyrus (BA 6) extending into the inferior frontal gyrus (BA 44 and 45), the primary somatosensory cortex (BA2 in particular), the middle and superior temporal areas and the supramarginal gyri. Roughly the same pattern emerges when the activity during observing is masked with the activity during gesturing. Figures 3.1-F and 3.1-G detail these activations.

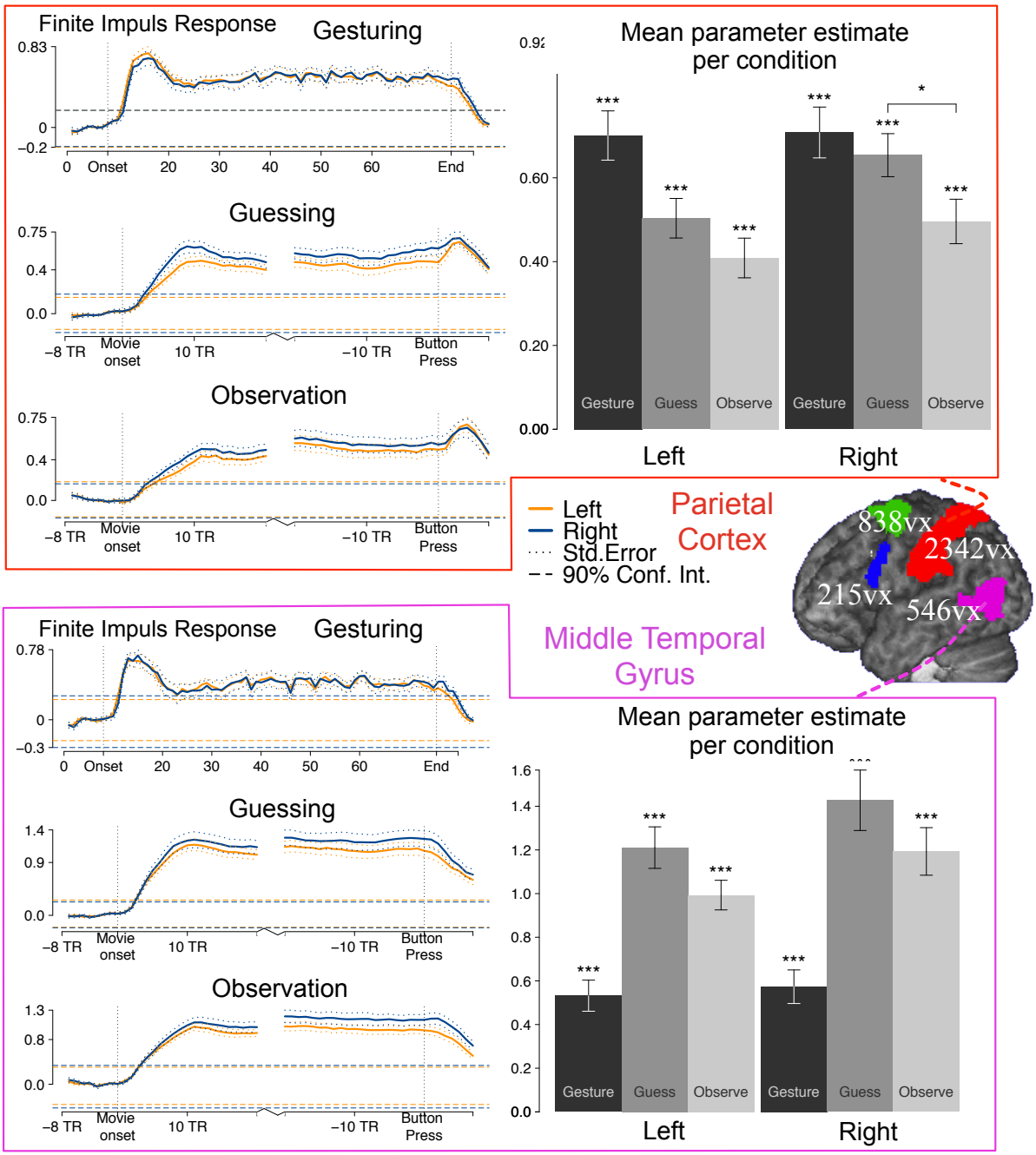
*Similarities and differences between guessing and passive observation masked with gesturing*

Contrasting active guessing with passive observation and masking this with the activation during gesturing shows noticeable peaks in the right inferior parietal lobule and in the left BA 44 (Fig. 3.1-H). Substantially larger areas remain when the activity that is present during both active guessing and passive observation is masked with activity during gesturing, without there being a significant difference between these conditions. These include much of the somatosensory, premotor, middle temporal- and supramarginal cortex (Fig. 3.1-I).

### 3.3.3 *Regions of Interest fMRI Results*

*Putative mirror neuron system (Figure 3.3)*

The bar plot of the parameter estimates during the different conditions show that all conditions activate all putative mirror neuron areas significantly even at an uncorrected threshold of $P < 0.001$. The time courses show further that all areas are substantially activated during the whole period of each condition (as evidenced by




the mean activity exceeding the confidence interval (dashed line) of the mean activity during the 5 volumes prior to stimulus onset). Two of these areas make a significant distinction between guessing and passive observation, but only under an uncorrected threshold of $P < 0.05$. These areas are the right parietal cortex and the left ventral premotor cortex.

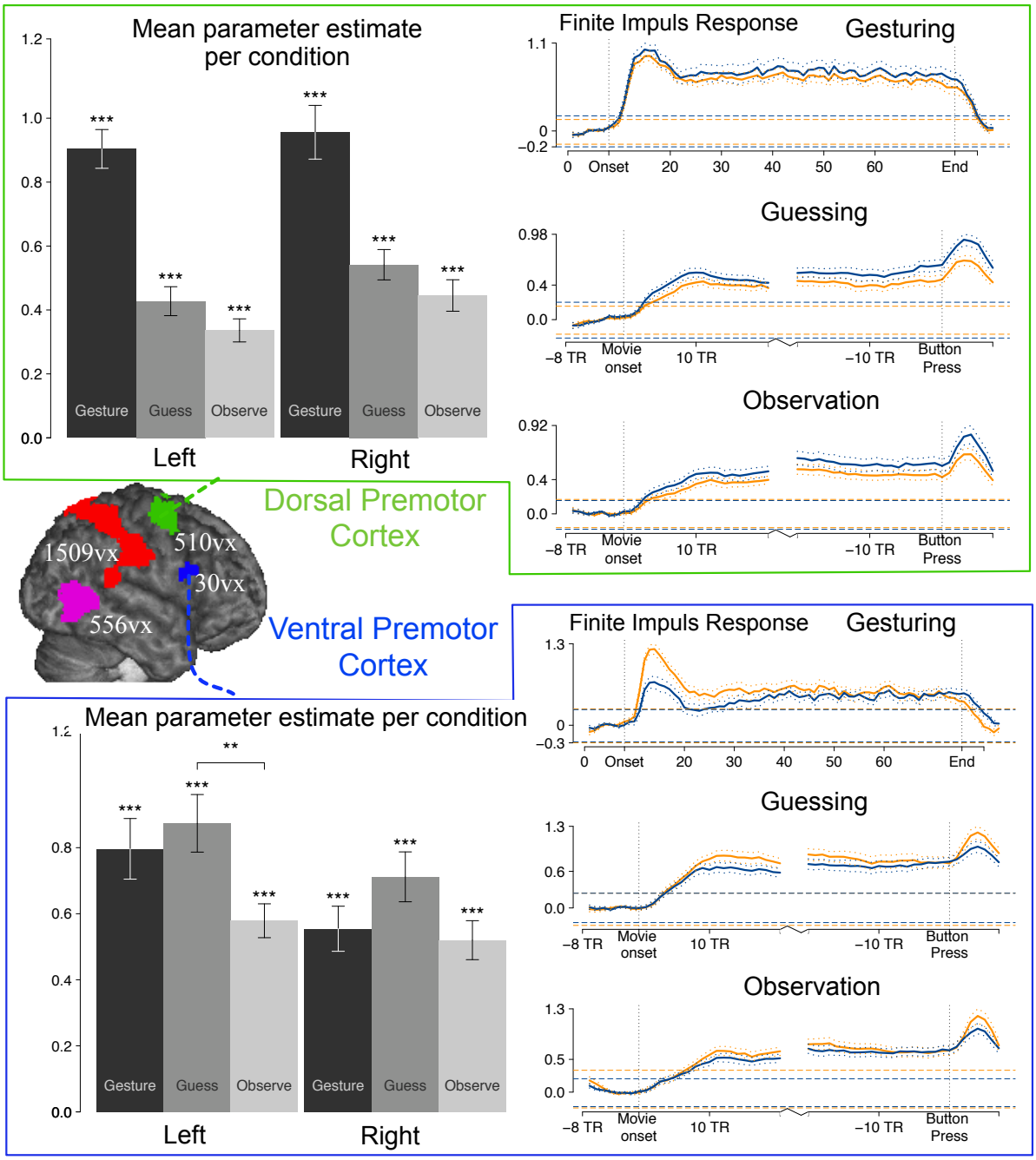



Figure 3.3: Locations and sizes of the pMNS ROI (center) together with their parameter estimates for each condition (bar graphs). Curves show the peri-stimulus time histogram for each condition in each ROI. For gesturing, the whole period of gesturing is plotted, from 8 volumes before the onset of the gesture until 8 volumes after the gesture has stopped. During both guessing and passive observation, the begin period (8 volumes before onset of the movie of the gesture until 20 volumes after) and the end period (20 volumes before button press until 8 volumes after) are plotted in the same graph, with the interruption due to the participants responding after variable amounts of time. See center legend for further details.

*Putative theory-of-mind areas (Figure 3.4)*

The medial prefrontal cortex shows no significant response to any of the conditions when applying an uncorrected threshold of P < 0.001, in contrast to the temporo-

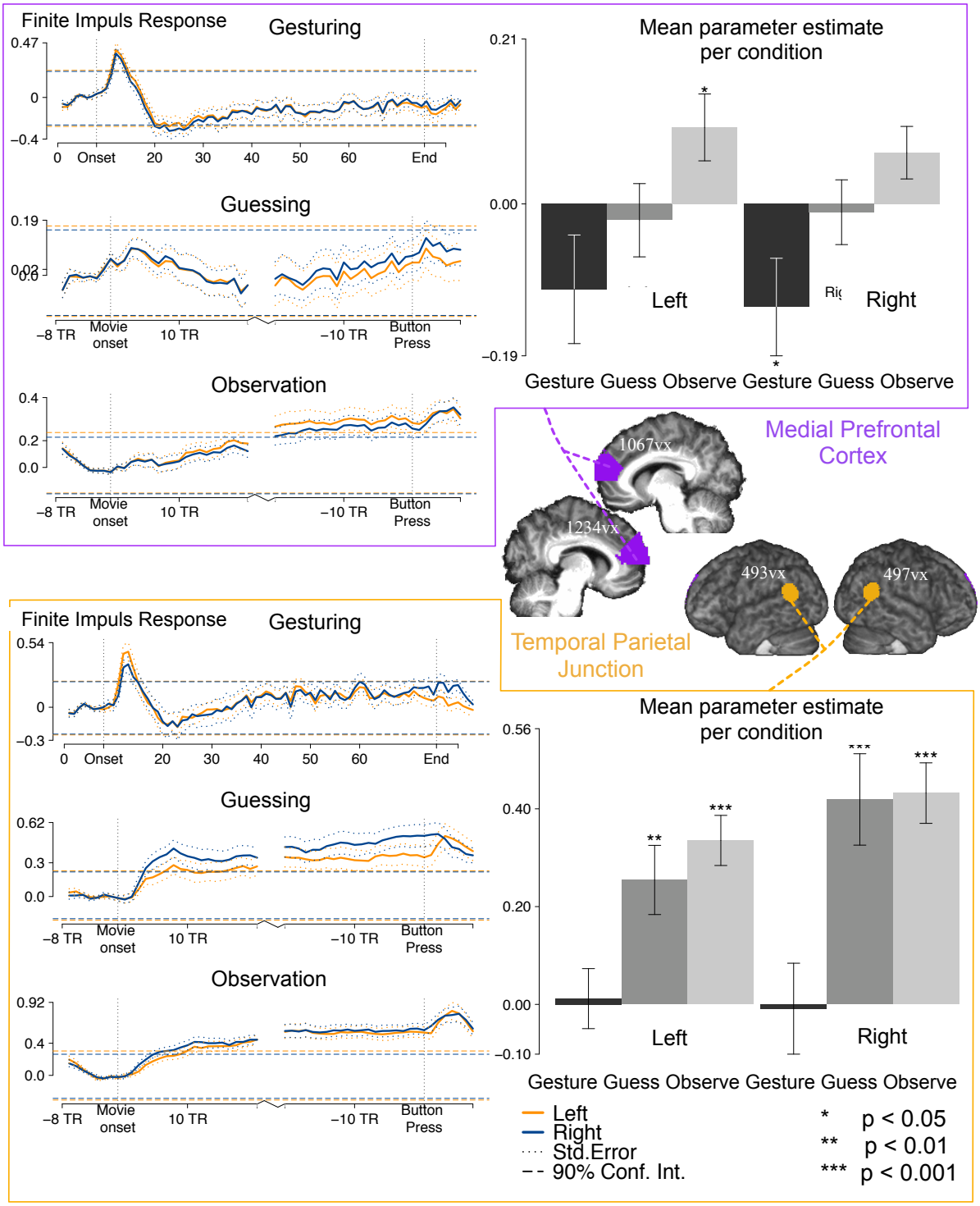



Figure 3.4: Same as Fig. 3.3 for pToM areas.

parietal junction. The time courses confirm this observation: activation almost never reaches significantly above the baseline activity, except at the end of a movie during the passive observation condition. The temporo-parietal junction is recruited significantly during both guessing and passive observation, but not during gesturing. This is also confirmed in its time courses.

## 3.4 DISCUSSION

In this experiment romantically involved couples played the game of charades in the scanner, taking turns as either the sender (gesturing) or receiver (guessing) of gestures. In this motivating context, they very naturally generated and decoded novel gestures with a communicative intention. The main goals of the study were to investigate to what extent (a) the pMNS for transitive hand actions and (b) pToM areas are involved in deliberate communication through gestures, and (c) how dependent the activity in these areas is on the communicative intention induced by the task. We analyzed the involvement of these two networks in two ways: through a whole-brain and a region-of-interest (ROI) analysis. Both analyses gave similar results. The pMNS does indeed become activated during communication through gestures, with highly overlapping brain areas involved in sending and receiving the gestural message. In contrast, the most typical of pToM areas, the anterior rostral medial frontal cortex associated with theory-of-mind (Amodio and Frith, 2006) (which we will refer to as mPFC) was not recruited beyond baseline levels during either sending or receiving gestural messages; the TPJ was engaged during observation but not during gesturing. The pMNS and TPJ were significantly activated both during guessing and passive viewing. The hypothesis that the TPJ would only be activated during the guessing conditions that explicitly encourages decoding the mental states (i.e. what is he trying to tell me?) but not the control condition (passive viewing), was not confirmed.

### *Involvement of the putative mirror neuron system*

Our study shows that brain regions associated with the pMNS for goal-directed, transitive actions were recruited during gestural communication - even when physical objects are not being present. A whole-brain analysis, in which the execution of gestures is used to mask the guessing or passive observation of gestures, shows a large overlap between the areas recruited in the three conditions (Fig. 3.1-F, 3.1-G). Furthermore, the ROI analysis of the pMNS, as defined using actions directed at objects (Gazzola et al., 2007a), shows sustained activity in these areas during the whole period of gesturing, guessing and passive observation (Fig. 3.3). Combining the study of Gazzola et al. (2007a) with the results of the current study show that

the same set of voxels in the brain is therefore involved in (a) mapping the object-directed hand actions of others onto the neural substrates involved in executing similar object-directed hand actions and (b) mapping the gestures of others onto the neural substrates involved in executing similar gestures. This extends previous findings (Montgomery et al., 2007) by showing that even in the absence of imitation trials, and during a genuinely communicative task, the brain regions associated with the pMNS for goal-directed actions are consistently activated. See Supplementary Information 1 for a discussion of how this finding relates to the question of whether the pMNS requires objects to be activated.

To maintain the flow of the game, control conditions involving the static vision of hands or meaningless hand actions were not included in this study. One might therefore question whether the activity found in the ROIs during gesture viewing (guessing or passive observation) is specific to actions or whether it reflects unspecific attentional resources. The ROIs used to extract the signal in the pMNS have been extensively examined in our laboratory using the same scanner and analysis software (Gazzola and Keysers, 2008; Gazzola et al., 2007a,b). Supplementary Information 2 illustrates the peak percent signal changes of the time courses measured in Gazzola et al. (2007a) (their Fig. S3) and those observed during the same time period of the gesture condition in the present experiment. Doing so revealed that activations in the guessing condition here exceeded those of the control conditions of Gazzola et al. (2007a) in all but the right ventral premotor ROI. Indeed, in the same ROIs, the activity in the present experiment often exceeded even the vision of goal directed actions in all but the right ventral premotor ROI. Although comparisons across experiments are problematic and should be interpreted with caution, this does suggest that the activity during the viewing of gestures in the present experiment reflects genuine action processing that exceeds that during the sight of mere movements.

Interestingly, the brain activity induced while engaged in active guessing overlapped considerably with that obtained during the second showing of the exact same visual stimuli but without the task (Fig. 3.1-F,3.1-G). As noted in the results, quantitative comparisons across different sessions are problematic, and conclusions drawn from these comparisons have to be considered with care. A quantitative comparison between activity in the two conditions within the confines of regions involved in gesture production however did reveal significantly higher BOLD during active guessing compared to passive viewing. The areas particularly involved were BA44 and the MTG (Fig. 3.1-H). These differences are unlikely to be due to systematic differences in the sensitivity of the scanner, as there were no significant differences in these areas between the globals extracted by the general linear model on the two scanning days (see methods). These differences were also marginal compared to the much more extensive network of premotor, parietal and temporal regions of the pMNS that did not show a significant difference between the two tasks (Fig. 3.1-I). This finding is in

line with a previous study which showed that the pMNS for facial movements is only marginally affected by task (van der Gaag et al., 2007). A number of studies (Brass et al., 2000; Kilner et al., 2007) have shown that observing other people's behaviour interferes with the observer's own movements even if it would be beneficial for the observer to ignore the movements of the other person. We believe that the similarity between the activity in passive viewing and active guessing, and the fact that both significantly activate the pMNS, highlights the tendency of the pMNS and/or the subjects to process the actions of others even if the experimenter's instructions do not explicitly encourage them to do so. With 'and/or the subject' we refer to the fact that upon debriefing, some of our participants reported finding it hard to refrain entirely from interpreting the gestures in the passive viewing condition. They did report however, that they interpreted the actions more during the guessing condition.

It should be noted that activation of the pMNS regions during gesture observation and production can, but does not have to reflect activity in mirror neurons within these voxels. This is because a voxel involved in two tasks could contain a population of neurons involved in both, as has been shown in the monkey (Gallese et al., 1996; Keysers et al., 2003; Kohler et al., 2002) and/or two distinct populations, each of which being involved in only one of the two tasks, interdigitated within the volume of the voxel (Gazzola and Keysers, 2008).

*Involvement of Theory-of-Mind areas*

Because playing charades could require the explicit guessing of the communicative mental state of the gesturer ("what was he trying to tell me?"), our second experimental question was whether pToM areas, including the mPFC and the TPJ, would be significantly recruited during the gesturing, active guessing and/or passive viewing.

*Medial Prefrontal Cortex*

Previous studies have shown that mentalizing is associated with activity in the mPFC (Brunet et al., 2000; Castelli et al., 2000; Gallagher and Frith, 2003; Gallagher et al., 2000, 2002; Iacoboni et al., 2004; Siegal and Varley, 2002; Vogeley and Fink, 2003; Walter et al., 2004). More specifically, Sommer et al. (2007) showed that true belief reasoning (which might be closer to what participants need to do here compared to false-belief reasoning) involves the mPFC. Furthermore, Kampe et al. (2003), as well as Walter et al. (2004), and Ciaramidaro et al. (2007) found the anterior paracingulate cortex to be recruited while recognizing the communicative intentions of others (for reviews see Amodio and Frith, 2006; Frith and Frith, 2006). In our experiment, neither the ROI nor the whole brain analysis revealed activations above baseline in the mPFC during

any of the conditions. This was true using a threshold of $p<0.001$, and for the ROI analysis at using $p<0.01$ (see Fig. 3.4). This negative finding suggests that the mPFC may not play an active role in gestural communication. This finding seems different from Gallagher and Frith (2004) conclusions that the left anterior paracingulate cortex was selectively more involved in recognizing gestures expressing inner states. This difference may be due to the fact that our gestures referred to objects (nutcracker) and object-directed actions (riding a bicycle) while Gallagher and Frith's expressive gestures referred to inner states (I feel cold). Thinking about the inner states of others is indeed known to be particularly effective at triggering mPFC activity (Amodio and Frith, 2006).

We asked participants to consider the movies of their partner's actions for at least 50 seconds before reporting their interpretation of the gestures. This requirement was established to ascertain sufficient data points to examine the time course of activity. A consequence of this requirement, however, is the participants may have guessed the meaning of the gestures early in the trial, and before they gave their answer. Could the lack of mPFC activity in the whole-brain and ROI analysis be due to these trials? We believe not. If this were the case, the time course extracted from the mPFC ROI during the guessing condition should exceed the baseline activity or that during observation condition at least early in the trial. Our data (Fig. 3.4) does not support this hypothesis.

It should be note however, that all conditions in our experiment were compared against a passive baseline. It has been argued that a seemingly passive baseline actually goes hand-in-hand with increased metabolism in the mPFC (Raichle and Snyder, 2007), possibly because of self referential processing. Such default, self-referential activity would have been suspended by our tasks, leading to a decrease in mPFC activity that may have masked mentalizing processes of comparatively smaller metabolic demands.

### *Temporal Parietal Junction*

We found that the TPJ was significantly activated during guessing and passive observation but not gesturing. The TPJ has been associated with the ability to mentalize (Pelphrey et al., 2004; Saxe and Kanwisher, 2003; Saxe and Wexler, 2005; Saxe et al., 2004), but other studies suggest that this involvement might reflect attentional reorientation necessary for mentalizing rather than mentalizing per se (Decety and Lamm, 2007; Mitchell, 2008). It therefore remains unclear what can be deduced from its activation in some of our conditions. It might be that activity truly reflects mentalizing (Scholz et al., 2009), suggesting that the decoding of gestures but not their generation requires mentalizing. What sheds doubt on this interpretation is that during mentalizing tasks, the TPJ typically coactivates with mPFC, and this coactivation may be more

unique for mentalizing than the activity of either region taken alone. Alternatively, activity in the TPJ may reflect attentional reorienting (Decety and Lamm, 2007; Mitchell, 2008) (for instance between the gestures as an outer stimulus and the hypothesis about their meaning as an inner stimulus), which gesture interpretation may share with mentalizing. Finally, some have interpreted TPJ activity during the attribution of agency (Decety and Lamm, 2007), an interpretation that would match our finding TPJ activity only during to the third person conditions (guessing and passive observation) Further experiments are needed to disentangle these alternatives.

## 3.5 CONCLUSIONS

The putative mirror neuron system (pMNS) is recruited by observing communicative gestures (both with and without an instruction to interpret) and by the production of similar gestures. In contrast, the mPFC, which is often associated with mentalizing and ToM, was not recruited above baseline during gestural communication. Finally the TPJ, which is associated with mentalizing but also attention reorienting and the attribution of agency, was recruited during both passive observation and guessing. This suggests that observing gestures recruits a combination of TPJ and pMNS both when participants actively decode gestures and when they passively watch them. The pMNS - but not the TPJ - is recruited during the generation of similar gestures. These findings are in accordance with the idea that gestural communication could build upon a pMNS for goal-directed hand actions (Gentilucci and Corballis, 2006; Rizzolatti and Arbib, 1998). The pMNS could create a simulated first person perspective of the gestures through a combination of forward and reverse models in the somatosensory and motor domain (Gazzola and Keysers, 2008). This simulation could then provide additional information for associating the vision of gestures to their meaning. Evidence for mentalizing during gestural communication in this experiment is weak however. During gesture interpretation, TPJ activity could reflect the fact that information from the pMNS could feed into pToM components (the TPJ) (de Lange et al., 2008; Keysers and Gazzola, 2007; Thioux et al., 2008), but it is unclear why the mPFC would not have been active if activity truly reflects mentalizing. During gesture generation, neither the TPJ nor the mPFC were active above baseline. Alternatively, TPJ activity during gestural interpretation may reflect the attribution of agency to the action representations in the pMNS (Decety and Lamm, 2007).

We have introduced the game of charades in neuroimaging research as a motivating social game to study gestural communication. This provides a new tool to study the involvement of pMNS in a genuinely communicational context. By extending this method to study virtual or neurological lesions it can be determined whether these regions play a necessary role in understanding and generating communicative ges-

tures. A number of studies using gesturing tasks have found impairments in gesture recognition following motor skill impairment (Cubelli et al., 2006; Pazzaglia et al., 2008; Rothi et al., 1985). This suggests that the pMNS may indeed play a critical role. A recent study (Pazzaglia et al., 2008) shows that premotor and parietal lesions that impair hand action execution (as compared to mouth action execution) selectively impair the recognition of hand gestures (and their sounds). This confirms that lesions in the pMNS can selectively affect the production and perception of particular motor programs. This finding would be expected if simulation were important in gestural communication given that the pMNS is roughly somatotopically organized (Buccino et al., 2001; Gazzola et al., 2006; Thompson et al., 2007; Wheaton et al., 2004). Nevertheless, although gesture recognition is impaired in apraxic patients, performance typically remains substantially above chance level, suggesting that the pMNS cannot be the only route to associate gestures with meaning. Understanding the complementary nature of various sources of information within the brain during gestural communication will be an important focus of future research (de Lange et al., 2008; Keysers and Gazzola, 2007; Thioux et al., 2008).

# BIBLIOGRAPHY


Amodio, D. M. and Frith, C. D. (2006). Meeting of minds: the medial frontal cortex and social cognition. *Nature Reviews of Neuroscience*, 7(4):268–77.

Arbib, M. A. (2008). From grasp to language: embodied concepts and the challenge of abstraction. *Journal of Physiology - Paris*, 102(1-3):4–20.

Blakemore, S. J. and Decety, J. (2001). From the perception of action to the understanding of intention. *Nature Reviews of Neuroscience*, 2(8):561–7.

Brass, M., Bekkering, H., Wohlschläger, A., and Prinz, W. (2000). Compatibility between observed and executed finger movements: Comparing symbolic, spatial, and imitative cues. *Brain and Cognition*, 44(2):124–143.

Brett, M., Anton, J., Valabregue, R., and Poline, J.-B. (2002). Region of interest analysis using the marsbar toolbox for spm 99. *Neuroimage*.

Brunet, E., Sarfati, Y., Hardy-Baylé, M., and Decety, J. (2000). A pet investigation of the attribution of intentions with a nonverbal task. *Neuroimage*, 11(2):157–166.

Buccino, G., Binkofski, F., Fink, G. R., Fadiga, L., Fogassi, L., Gallese, V., Seitz, R. J., Zilles, K., Rizzolatti, G., and Freund, H.-J. (2001). Action observation activates premotor and parietal areas in a somatosopic manner: An fmri study. *The European Journal of Neuroscience*, 13:400–404.

Buccino, G., Binkofski, F., and Riggio, L. (2004). The mirror neuron system and action recognition. *Brain and Language*, 89(2):370–376.

Buxbaum, L. J., Kyle, K. M., and Menon, R. (2005). On beyond mirror neurons: internal representations subserving imitation and recognition of skilled object-related actions in humans. *Cognitive Brain Research*, 25(1):226–39.

Castelli, F., Happè, F., Frith, U., and Frith, C. D. (2000). Movement and mind: A functional imaging study of perception and interpretation of complex intentional movement patterns. *Neuroimage*, 12:314–325.

Choi, S. H., Na, D. L., Kang, E., Lee, K., Lee, S. W., and Na, D. G. (2001). Functional magnetic resonance imaging during pantomiming tool-use gestures. *Experimental Brain Research*, 139(3):311–7.

Chong, T. T.-J., Cunnington, R., Williams, M. A., Kanwisher, N., and Mattingley, J. B. (2008). fmri adaptation reveals mirror neurons in human inferior parietal cortex. *Current biology*, 18(20):1576–80.

Ciaramidaro, A., Adenzato, M., Enrici, I., Erk, S., Pia, L., Bara, B., and Walter, H. (2007). The intentional network: how the brain reads varieties of intentions. *Neuropsychology*, 45(13):3105–13.

Cubelli, R., Bartolo, A., Nichelli, P., and Sala, S. D. (2006). List effect in apraxia assessment. *Neuroscience Letters*, 407(2):118–20.

de Lange, F. P., Spronk, M., Willems, R. M., Toni, I., and Bekkering, H. (2008). Complementary systems for understanding action intentions. *Current biology*, 18(6):454–7.

Decety, J., Grèzes, J., Costes, N., Perani, D., Jeannerod, M., Procyk, E., Grassi, F., and Fazio, F. (1997). Brain activity during observation of actions. influence of action content and subject's strategy. *Brain*, 120 (Pt 10):1763–1777.

Decety, J. and Lamm, C. (2007). The role of the right temporoparietal junction in social interaction: how low-level computational processes contribute to meta-cognition. *The Neuroscientist*, 13(6):580–93.

Eickhoff, S., Stephan, K. E., Mohlberg, H., Grefkes, C., Fink, G. R., Amunts, K., and Zilles, K. (2005). A new spm toolbox for combining probabilistic cytoarchitectonic maps and functional imaging data. *Neuroimage*, 25(4):1325–1335.

Fadiga, L., Fogassi, L., Pavesi, G., and Rizzolatti, G. (1995). Motor facilitation during action observation: A magnetic stimulation study. *Journal of Neurophysiology*, 73(6):2608–2611.

Fein, G. (1981). Pretend play in childhood: An integrative review. *Child Development*.

Ferrari, P., Gallese, V., Rizzolatti, G., and Fogassi, L. (2003). Mirror neurons responding to the observation of ingestive and communicative mouth actions in the monkey ventral premotor cortex. *The European Journal of Neuroscience*, 17(8):1703–1714.

Fletcher, P. C., Happè, F., Frith, U., and Baker, S. (1995). Other minds in the brain: A functional neuroimaging study of 'theory of mind' in story comprehension. *Cognition*, 57(2):109–128.

Fogassi, L., Ferrari, P. F., Gesierich, B., Rozzi, S., Chersi, F., and Rizzolatti, G. (2005). Parietal lobe: from action organization to intention understanding. *Science*, 308(5722):662–7.

Fridman, E. A., Immisch, I., Hanakawa, T., Bohlhalter, S., Waldvogel, D., Kansaku, K., Wheaton, L., Wu, T., and Hallett, M. (2006). The role of the dorsal stream for gesture production. *Neuroimage*, 29(2):417–28.

Frith, C. D. and Frith, U. (2006). The neural basis of mentalizing. *Neuron*, 50(4):531–534.

Fujii, N., Hihara, S., and Iriki, A. (2007). Social cognition in premotor and parietal cortex. *Social Neuroscience*, 3(3):250–260.

Gallagher, H. L. and Frith, C. D. (2003). Functional imaging of 'theory of mind'. *Trends in Cognitive Sciences*, 7(2):77–83.

Gallagher, H. L. and Frith, C. D. (2004). Dissociable neural pathways for the perception and recognition of expressive and instrumental gestures. *Neuropsychology*, 42(13):1725–36.

Gallagher, H. L., Happè, F., Brunswick, N., Fletcher, P. C., Frith, U., and Frith, C. D. (2000). Reading the mind in cartoons and stories: An fmri study of 'theory of mind' in verbal and nonverbal tasks. *Neuropsychology*, 38:11–21.

Gallagher, H. L., Jack, A., Roepstorff, A., and Frith, C. D. (2002). Imaging the intentional stance in a competitive game. *Neuroimage*, 16(3 Pt 1):814–821.

Gallese, V., Fadiga, L., Fogassi, L., and Rizzolatti, G. (1996). Action recognition in the premotor cortex. *Brain*, 119(2):593–609.

Gallese, V. and Goldman, A. (1998). Mirror neurons and the simulation theory of mind-reading. *Trends in Cognitive Sciences*, 12:493–501.

Gazzola, V., Aziz-Zadeh, L., and Keysers, C. (2006). Empathy and the somatotopic auditory mirror system in humans. *Current biology*, 16(18):1824–9.

Gazzola, V. and Keysers, C. (2008). The observation and execution of actions share motor and somatosensory voxels in all tested subjects: Single-subject analyses of unsmoothed fmri data. *Cerebral Cortex*, 19(6):1239–1255.

Gazzola, V., Rizzolatti, G., Wicker, B., and Keysers, C. (2007a). The anthropomorphic brain: The mirror neuron system responds to human and robotic actions. *Neuroimage*, 35:1674–1684.

Gazzola, V., van der Worp, H., Mulder, T., Wicker, B., Rizzolatti, G., and Keysers, C. (2007b). Aplasics born without hands mirror the goal of hand actions with their feet. *Current biology*, 17(14):1235–40.

Gentilucci, M. and Corballis, M. C. (2006). From manual gesture to speech: a gradual transition. *Neuroscience and Biobehavioral Reviews*, 30(7):949–60.

Grafton, S. T., Arbib, M. A., Fadiga, L., and Rizzolatti, G. (1996). Localization of grasp representations in humans by positron emission tomography. 2. observation compared with imagination. *Experimental Brain Research*, 112(1):103–111.

Hampton, A. N., Bossaerts, P., and O'Doherty, J. P. (2008). Neural correlates of mentalizing-related computations during strategic interactions in humans. *Proceedings of the National Academy of Sciences of the United States of America*, 105(18):6741–6.

Hamzei, F., Rijntjes, M., Dettmers, C., Glauche, V., Weiller, C., and Büchel, C. (2003). The human action recognition system and its relationship to broca's area: an fmri study. *Neuroimage*, 19(3):637–44.

Heiser, M., Iacoboni, M., Maeda, F., Marcus, J., and Mazziotta, J. C. (2003). The essential role of broca's area in imitation. *The European Journal of Neuroscience*, 17(5):1123–8.

Hermsdörfer, J., Goldenberg, G., Wachsmuth, C., Conrad, B., Ceballos-Baumann, A. O., Bartenstein, P., Schwaiger, M., and Boecker, H. (2001). Cortical correlates of gesture processing: clues to the cerebral mechanisms underlying apraxia during the imitation of meaningless gestures. *Neuroimage*, 14(1 Pt 1):149–61.

Hickok, G. (2009). Eight problems for the mirror neuron theory of action understanding in monkeys and humans. *Journal of Cognitive Neuroscience*, 21(7):1229–43.

Higuchi, S., Imamizu, H., and Kawato, M. (2007). Cerebellar activity evoked by common tool-use execution and imagery tasks: an fmri study. *Cortex; a journal devoted to the study of the nervous system and behavior*, 43(3):350–8.

Iacoboni, M., Lieberman, M. D., Knowlton, B. J., Molnar-Szakacs, I., Moritz, M., Throop, C. J., and Fiske, A. P. (2004). Watching social interactions produces dorsomedial prefrontal and medial parietal bold fmri signal increases compared to a resting baseline. *Neuroimage*, 21(3):1167–73.

Iacoboni, M., Molnar-Szakacs, I., Gallese, V., Buccino, G., Mazziotta, J. C., and Rizzolatti, G. (2005). Grasping the intentions of others with one's own mirror neuron system. *PLoS Biology*, 3(3):e79.

Iacoboni, M., Woods, R., Brass, M., Bekkering, H., Mazziotta, J. C., and Rizzolatti, G. (1999). Cortical mechanisms of human imitation. *Science*, 286(5449):2526–2528.

Iverson, J. and Goldin-Meadow, S. (1998). Why people gesture when they speak. *Nature*, 396(6708):228.

Jeannerod, M. (2001). Neural simulation of action: A unifying mechanism for motor cognition. *Neuroimage*, 14(1):S103–S109.

Kampe, K. K. W., Frith, C. D., and Frith, U. (2003). "hey john": signals conveying communicative intention toward the self activate brain regions associated with "mentalizing," regardless of modality. *The Journal of Neuroscience*, 23(12):5258–63.

Kendon, A. (1994). Do gestures communicate? a review. *Research on Language and Social Interaction*, 27(3):175 – 200.

Keysers, C. and Gazzola, V. (2006). Towards a unifying neural theory of social cognition. *Progress in Brain Research*, 156:379–401.

Keysers, C. and Gazzola, V. (2007). Integrating simulation and theory of mind: from self to social cognition. *Trends in Cognitive Sciences*, 11(5):194–6.

Keysers, C., Kohler, E., Umiltà, M., Nanetti, L., Fogassi, L., and Gallese, V. (2003). Audiovisual mirror neurons and action recognition. *Experimental Brain Research*, 153(4):628–636.

Kilner, J. M., Friston, K. J., and Frith, C. D. (2007). Predictive coding: an account of the mirror neuron system. *Cognitive Processes*, 8(3):159–166.

Kohler, E., Keysers, C., Umiltà, M., Fogassi, L., Gallese, V., and Rizzolatti, G. (2002). Hearing sounds, understanding actions: Action representation in mirror neurons. *Science*, 297(5582):846–849.

Lotze, M., Heymans, U., Birbaumer, N., Veit, R., Erb, M., Flor, H., and Halsband, U. (2006). Differential cerebral activation during observation of expressive gestures and motor acts. *Neuropsychology*, 44(10):1787–95.

McCabe, K., Houser, D., Ryan, L., Smith, V., and Trouard, T. (2001). A functional imaging study of cooperation in two-person reciprocal exchange. *Proceedings of the National Academy of Sciences of the United States of America*, 98(20):11832–11835.

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

Melinger, A. and Levelt, W. (2004). Gesture and the communicative intention of the speaker. *Gesture*, 4(2):119–141.

Mitchell, J. P. (2008). Activity in right temporo-parietal junction is not selective for theory-of-mind. *Cerebral Cortex*, 18(2):262–71.

Moll, J., de Oliveira-Souza, R., Passman, L. J., Cunha, F. C., Souza-Lima, F., and Andreiuolo, P. A. (2000). Functional mri correlates of real and imagined tool-use pantomimes. *Neurology*, 54(6):1331–6.

Montgomery, K. J., N.Isenberg, and Haxby, J. V. (2007). Communicative hand gestures and object-directed hand movements activated the mirror neuron system. *Social Cognitive and Affective Neuroscience*, 2(2):114–122.

Mozaz, M., Rothi, L. J. G., Anderson, J. M., Crucian, G. P., and Heilman, K. M. (2002). Postural knowledge of transitive pantomimes and intransitive gestures. *Journal of the International Neuropsychological Society*, 8(7):958–62.

Nair, D. G., Purcott, K. L., Fuchs, A., Steinberg, F., and Kelso, J. A. S. (2003). Cortical and cerebellar activity of the human brain during imagined and executed unimanual and bimanual action sequences: a functional mri study. *Cognitive Brain Research*, 15(3):250–60.

Nelissen, K., Luppino, G., Vanduffel, W., Rizzolatti, G., and Orban, G. (2005). Observing others: multiple action representation in the frontal lobe. *Science*, 310(5746):332–336.

Nishitani, N. and Hari, R. (2000). Temporal dynamics of cortical representation for action. *Proceedings of the National Academy of Sciences of the United States of America*, 97(2):913–8.

Ochipa, C., Rothi, L. J. G., and Heilman, K. M. (1989). Ideational apraxia: a deficit in tool selection and use. *Annals of Neurology*, 25(2):190–3.

Ohgami, Y., Matsuo, K., Uchida, N., and Nakai, T. (2004). An fmri study of tool-use gestures: body part as object and pantomime. *Neuroreport*, 15(12):1903–6.

Oldfield, R. (1971). The assessment and analysis of handedness: the edinburgh inventory. *Neuropsychology*, 9(1):97–113.

Pazzaglia, M., Smania, N., Corato, E., and Aglioti, S. M. (2008). Neural underpinnings of gesture discrimination in patients with limb apraxia. *The Journal of Neuroscience*, 28(12):3030–41.

Pelphrey, K., Morris, J., and McCarthy, G. (2004). Grasping the intentions of others: The perceived intentionality of an action influences activity in the superior temporal

sulcus during social perception. *Journal of Cognitive Neuroscience*, 16(10):1706–1717.

Petrides, M., Cadoret, G., and Mackey, S. (2005). Orofacial somatomotor responses in the macaque monkey homologue of broca's area. *Nature*, 435(7046):1235–8.

Poldrack, R. A. (2006). Can cognitive processes be inferred from neuroimaging data? *Trends in Cognitive Sciences*, 10(2):59–63.

Raichle, M. E. and Snyder, A. Z. (2007). A default mode of brain function: a brief history of an evolving idea. *Neuroimage*, 37(4):1083–90; discussion 1097–9.

Rizzolatti, G. and Arbib, M. A. (1998). Language within our grasp. *Trends in Neurosciences*, 21(5):188–194.

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

Rizzolatti, G., Fadiga, L., Gallese, V., and Fogassi, L. (1996). Premotor cortex and the recognition of motor actions. *Cognitive Brain Research*, 3(18):131–141.

Rothi, L. J. G., Heilman, K. M., and Watson, R. T. (1985). Pantomime comprehension and ideomotor apraxia. *Journal of Neurology, Neurosurgery, and Psychiatry*, 48(3):207–10.

Saxe, R. and Kanwisher, N. (2003). People thinking about thinking people. the role of the temporo-parietal junction in "theory of mind". *Neuroimage*, 19(4):1835–42.

Saxe, R. and Wexler, A. (2005). Making sense of another mind: the role of the right temporo-parietal junction. *Neuropsychology*, 43(10):1391–9.

Saxe, R., Xiao, D., Kovacs, G., Perrett, D., and Kanwisher, N. (2004). A region of right posterior superior temporal sulcus responds to observed intentional actions. *Neuropsychology*, 42:1435–1446.

Scholz, J., Triantafyllou, C., Whitfield-Gabrieli, S., Brown, E. N., and Saxe, R. (2009). Distinct regions of right temporo-parietal junction are selective for theory of mind and exogenous attention. *PLoS ONE*, 4(3):e4869.

Siegal, M. and Varley, R. (2002). Neural systems involved in theory of mind. *Nature Reviews of Neuroscience*, 3(6):463–471.

Singer, T., Seymour, B., O'Doherty, J., Kaube, H., Dolan, R., and Frith, C. D. (2004). Empathy for pain involves the affective but not sensory components of pain. *Science*, 303(5661):1157–1162.

Sommer, M., Döhnel, K., Sodian, B., Meinhardt, J., Thoermer, C., and Hajak, G. (2007). Neural correlates of true and false belief reasoning. *Neuroimage*, 35(3):1378–84.

Thioux, M., Gazzola, V., and Keysers, C. (2008). Action understanding: how, what and why. *Current biology*, 18(10):R431–4.

Thompson, J. C., Hardee, J. E., Panayiotou, A., Crewther, D., and Puce, A. (2007). Common and distinct brain activation to viewing dynamic sequences of face and hand movements. *Neuroimage*, 37(3):966–73.

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

van der Gaag, C., Minderaa, R. B., and Keysers, C. (2007). Facial expressions: what the mirror neuron system can and cannot tell us. *Social Neuroscience*, 2(3-4):179–222.

Vogeley, K., Bussfeld, P., Newen, A., Herrmann, S., Happè, F., Falkai, P., Maier, W., Shah, N. J., Fink, G. R., and Zilles, K. (2001). Mind reading: neural mechanisms of theory of mind and self-perspective. *Neuroimage*, 14(1 Pt 1):170–181.

Vogeley, K. and Fink, G. R. (2003). Neural correlates of the first-person-perspective. *Trends in Cognitive Sciences*, 7(1):38–42.

Walter, H., Adenzato, M., Ciaramidaro, A., Enrici, I., Pia, L., and Bara, B. (2004). Understanding intentions in social interaction: the role of the anterior paracingulate cortex. *Journal of Cognitive Neuroscience*, 16(10):1854–1863.

Wheaton, K. J., Thompson, J. C., Syngeniotis, A., Abbott, D. F., and Puce, A. (2004). Viewing the motion of human body parts activates different regions of premotor, temporal, and parietal cortex. *Neuroimage*, 22(1):277–88.

Willems, R. M. and Hagoort, P. (2007). Neural evidence for the interplay between language, gesture, and action: a review. *Brain and Language*, 101(3):278–89.

Each of the following tables specify for each supra-threshold cluster of activation during the indicated contrast, the t-value, location, anatomical description and, when available, probabilistically determined Brodmann area according to the anatomy toolbox (Eickhoff et al., 2005).

| Main effects of guessing | | | | | P (unc) < 0.001; T > 3.58; N = 20; K > 10 | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| Clustersize (K) | T | Talairach X | Y | Z | MNI X | Y | Z | Hem. | Area | BA |
| 65367 | 17.30 | 50 | -72 | 3 | 50 | -68 | -2 | Bilateral | Middle Temporal Gyrus | |
| | 15.87 | 28 | -54 | 57 | 28 | -50 | 52 | Bilateral | Inferior Parietal Lobule | |
| | 14.77 | 56 | -46 | 21 | 56 | -42 | 16 | Bilateral | Superior Temporal Gyrus | |
| | 13.90 | 48 | -32 | 41 | 48 | -28 | 36 | Bilateral | SupraMarginal Gyrus | |
| | 13.15 | 34 | -40 | 57 | 34 | -36 | 52 | Bilateral | Postcentral Gyrus | 2 |
| | 10.03 | 60 | 14 | 33 | 60 | 18 | 28 | Bilateral | Inferior Frontal Gyrus | 44, 45 |
| | 8.19 | -40 | 18 | -17 | -40 | 22 | -22 | Bilateral | Temporal Pole | |
| | 7.69 | -40 | -30 | 69 | -40 | -26 | 64 | Bilateral | Precentral Gyrus | 4a, 6 |
| Reductions in BOLD signal during Guessing | | | | | | | | | | |
| 4719 | 13.3 | -26 | -58 | 13 | -26 | -54 | 8 | Bilateral | Precuneus | 18 |
| 489 | 6.09 | 42 | -14 | 17 | 42 | -10 | 12 | R | Insula Lobe (OP 3) | |
| 404 | 7.23 | -52 | -68 | 45 | -52 | -64 | 40 | Bilateral | Angular Gyrus | |

Table 3.2: Activation Table for the contrast Guessing - Baseline.

| Main effects of passive observation | | | | | P (unc) < 0.001; T > 3.58; | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| Clustersize (K) | T-value | Talairach X | Y | Z | MNI X | Y | Z | Hem. | Area | BA |
| 74505 | 15.20 | 50 | -70 | 3 | 50 | -66 | -2 | Bilateral | Middle / Superior Temporal Gyrus | |
| | 12.32 | -30 | -52 | 55 | -30 | -48 | 50 | Bilateral | Inferior Parietal Lobule | |
| | 11.99 | 32 | -36 | 51 | 32 | -32 | 46 | Bilateral | Postcentral Gyrus | 3a |
| | 10.34 | -38 | -40 | 57 | -38 | -36 | 52 | Bilateral | Postcentral Gyrus | 2 |
| | 9.62 | 60 | 4 | 43 | 60 | 8 | 38 | Bilateral | Precentral Gyrus | 6 |
| | 9.31 | 10 | 8 | 73 | 10 | 12 | 68 | Bilateral | SMA | 6 |
| | 9.14 | 38 | 4 | 41 | 38 | 8 | 36 | Bilateral | Middle Frontal Gyrus | |
| | 9.10 | 60 | 12 | 33 | 60 | 16 | 28 | Bilateral | Inferior Frontal Gyrus | 44 |
| | 8.96 | -60 | -46 | 29 | -60 | -42 | 24 | Bilateral | SupraMarginal Gyrus | |
| | 8.42 | -54 | -28 | 57 | -54 | -24 | 52 | Bilateral | Postcentral Gyrus | 1 |
| | 8.14 | 54 | 12 | -9 | 54 | 16 | -14 | Bilateral | Temporal Pole | |
| Reductions in BOLD signal during Observation | | | | | | | | | | |
| 1287 | 9.16 | -24 | -56 | 17 | -24 | -52 | 12 | L | Precuneus | |
| 97 | 4.21 | -4 | -52 | -29 | -4 | -48 | -34 | L | Cerebelum (IX) | |
| 26 | 4.49 | -16 | 14 | 21 | -16 | 18 | 16 | L | Caudate Nucleus | |

Table 3.3: Activation Table for the contrast Passive Observation - Baseline.

| Main effects of gesturing | | | | | P (unc) < 0.001; T > 3.58; N = 20; K > 10 | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| Clustersize (K) | T | Talairach X | Y | Z | MNI X | Y | Z | Hem. | Area | BA |
| | 14.55 | -28 | -58 | 71 | -28 | -54 | 66 | Bilateral | Superior Parietal Lobule | |
| | 13.81 | 44 | -24 | 59 | 44 | -20 | 54 | Bilateral | Postcentral Gyrus | 3b |
| | 13.46 | -36 | -18 | 67 | -36 | -14 | 62 | Bilateral | Precentral Gyrus | 6 |
| | 13.41 | 4 | -4 | 73 | 4 | 0 | 68 | Bilateral | SMA | 6 |
| | 12.04 | -46 | -36 | 43 | -46 | -32 | 38 | Bilateral | Inferior Parietal Lobule | |
| | 11.98 | -52 | 2 | 15 | -52 | 6 | 10 | Bilateral | Inferior Frontal Gyrus | 44 |
| | 10.61 | -52 | -66 | 3 | -52 | -62 | -2 | Bilateral | Middle Temporal Gyrus | |
| | 10.03 | 66 | -38 | 27 | 66 | -34 | 22 | Bilateral | Superior Temporal Gyrus | |
| | 5.38 | 26 | -76 | 45 | 26 | -72 | 40 | Bilateral | Superior Occipital Gyrus | |
| 615 | 6.26 | 32 | 42 | 33 | 32 | 46 | 28 | Bilateral | Middle Frontal Gyrus | |

| Reductions in BOLD signal during Gesturing | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| 3228 | 9.03 | 4 | -64 | 37 | 4 | -60 | 32 | R | Precuneus | |
| 675 | 6.58 | 50 | 10 | -19 | 50 | 14 | -24 | R | Temporal Pole | |
| 610 | 7.15 | -54 | -66 | 43 | -54 | -62 | 38 | L | Angular Gyrus | |
| 588 | 7.83 | 52 | -68 | 47 | 52 | -64 | 42 | R | Angular Gyrus | |
| 21 | 4.17 | -42 | 10 | -29 | -42 | 14 | -34 | L | Medial Temporal Pole | |

Table 3.4: Activation Table for the contrast Gesturing - Baseline.

| Guessing - Observation | | | | | P (unc) < 0.001; T > 3.58; N = 20; K > 10 | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| Clustersize (K) | T | Talairach X | Y | Z | MNI X | Y | Z | Hem. | Area | BA |
| 3035 | 5.82 | 44 | -72 | 3 | 44 | -68 | -2 | R | Middle Temporal Gyrus | |
| | 5.30 | 26 | -92 | -3 | 26 | -88 | -8 | R | Lingual Gyrus | 18 |
| 2304 | 6.18 | -42 | -72 | 1 | -42 | -68 | -4 | L | Inferior Occipital Gyrus (V5/MT+) | |
| | 6.14 | -44 | -70 | 3 | -44 | -66 | -2 | L | Middle Occipital Gyrus | |
| 1213 | 6.50 | 36 | -48 | 65 | 36 | -44 | 60 | R | Superior Parietal Lobule | 2 |
| 280 | 5.83 | 26 | -30 | 1 | 26 | -26 | -4 | R | Hippocampus | |
| 201 | 4.69 | -48 | 10 | 27 | -48 | 14 | 22 | L | Inferior Frontal Gyrus | 44 / 45 |

| Observation - Guessing | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| 3116 | 7.68 | -8 | -66 | 37 | -8 | -62 | 32 | Bilateral | Precuneus | |
| | 3.76 | 4 | -34 | 41 | 4 | -30 | 36 | R | Middle Cingulate Cortex | |
| 618 | 6.38 | -46 | -66 | 37 | -46 | -62 | 32 | L | Angular Gyrus | |
| 355 | 5.64 | 52 | -68 | 45 | 52 | -64 | 40 | R | Angular Gyrus | |
| | 5.63 | 54 | -66 | 43 | 54 | -62 | 38 | R | Inferior Parietal Lobule | |
| 265 | 4.22 | 10 | -30 | 73 | 10 | -26 | 68 | Bilateral | Paracentral Lobule | 6 |

Table 3.5: Activation Table for the contrast Guessing - Passive Observation.

| Guessing = Observation | | | | | | | | P (unc) < 0.001; T > 3.58; N = 20; K > 10 | | |
|---|---|---|---|---|---|---|---|---|---|---|
| Clustersize (K) | T | Talairach X | Y | Z | MNI X | Y | Z | Hem. | Area | BA |
| 51164 | 17.12 | 36 | -78 | -3 | 36 | -74 | -8 | R | Inferior Occipital Gyrus | |
| | 15.47 | 14 | -88 | -9 | 14 | -84 | -14 | R | Lingual Gyrus | 18 |
| | 14.58 | -4 | -90 | -7 | -4 | -86 | -12 | L | Calcarine Gyrus | 18 |
| | 14.02 | -30 | -92 | 21 | -30 | -88 | 16 | L | Middle Occipital Gyrus | |
| | 13.79 | -28 | -84 | -5 | -28 | -80 | -10 | L | Inferior Occipital Gyrus | |
| | 13.36 | 58 | -46 | 21 | 58 | -42 | 16 | R | Superior Temporal Gyrus | |
| | 12.92 | 18 | -96 | 15 | 18 | -92 | 10 | R | Cuneus | |
| | 12.75 | 42 | -60 | -9 | 42 | -56 | -14 | R | Inferior Temporal Gyrus | |
| | 12.32 | 26 | -82 | -5 | 26 | -78 | -10 | R | Fusiform Gyrus | |
| | 11.57 | -44 | -58 | -11 | -44 | -54 | -16 | L | Fusiform Gyrus | |
| | 11.20 | 32 | -38 | 59 | 32 | -34 | 54 | R | Postcentral Gyrus | 3b |
| | 11.09 | 32 | -88 | 21 | 32 | -84 | 16 | R | Middle Occipital Gyrus | |
| | 11.00 | -50 | -4 | 43 | -50 | 0 | 38 | L | Precentral Gyrus | |
| | 10.73 | 54 | -72 | 13 | 54 | -68 | 8 | R | Middle Temporal Gyrus (hOC5 V5/MT+) | |
| | 10.67 | 36 | -34 | 49 | 36 | -30 | 44 | R | Postcentral Gyrus | 2 |
| | 10.41 | -32 | -56 | 65 | -32 | -52 | 60 | L | Superior Parietal Lobule | |
| | 10.10 | 32 | -66 | -17 | 32 | -62 | -22 | R | Cerebelum VI | |
| | 9.63 | 26 | -58 | 57 | 26 | -54 | 52 | R | Inferior Parietal Lobule | |
| | 9.62 | 60 | 4 | 43 | 60 | 8 | 38 | R | Precentral Gyrus | 6 |
| | 9.51 | -46 | -70 | 15 | -46 | -66 | 10 | L | Middle Temporal Gyrus | |
| | 9.46 | 52 | 12 | 33 | 52 | 16 | 28 | R | Inferior Frontal Gyrus (p. Opercularis) | 44 |
| | 9.36 | 22 | -8 | 69 | 22 | -4 | 64 | R | Superior Frontal Gyrus | 6 |
| | 9.00 | -52 | 8 | 21 | -52 | 12 | 16 | L | Inferior Frontal Gyrus (p. Opercularis) | 44 |
| | 8.90 | -16 | -78 | -17 | -16 | -74 | -22 | L | Cerebelum VI | |
| | 8.88 | -26 | -8 | 63 | -26 | -4 | 58 | L | Middle Frontal Gyrus | |
| | 8.72 | -52 | -30 | 45 | -52 | -26 | 40 | L | Inferior Parietal Lobule | 2 |
| | 8.65 | 40 | -4 | 67 | 40 | 0 | 62 | R | Middle Frontal Gyrus | |
| | 8.63 | -44 | -40 | -13 | -44 | -36 | -18 | L | Inferior Temporal Gyrus | |
| | 8.17 | -32 | 26 | 1 | -32 | 30 | -4 | L | Inferior Frontal Gyrus (p. Orbitalis) | |
| | 8.09 | -4 | 8 | 63 | -4 | 12 | 58 | L | SMA | 6 |
| | 8.09 | 24 | -60 | 67 | 24 | -56 | 62 | R | Superior Parietal Lobule | |
| | 8.06 | -56 | -26 | 53 | -56 | -22 | 48 | L | Inferior Parietal Lobule | 1 |
| | 7.99 | 40 | -44 | 53 | 40 | -40 | 48 | R | Inferior Parietal Lobule (hIP2) | |
| | 7.89 | 20 | -28 | 7 | 20 | -24 | 2 | R | Thalamus | |
| | 7.82 | -58 | -46 | 27 | -58 | -42 | 22 | L | Superior Temporal Gyrus | |
| | 7.82 | 4 | 8 | 67 | 4 | 12 | 62 | R | SMA | 6 |
| | 7.77 | -22 | -88 | 43 | -22 | -84 | 38 | L | Superior Occipital Gyrus | 45 |
| | 7.39 | -52 | 24 | 27 | -52 | 28 | 22 | L | Inferior Frontal Gyrus (p. Triangularis) | |
| | 7.34 | -42 | 18 | -15 | -42 | 22 | -20 | L | Temporal Pole | |
| | 6.93 | 56 | 22 | 29 | 56 | 26 | 24 | R | Inferior Frontal Gyrus (p. Triangularis) | 45 |
| | 6.80 | 36 | -26 | 63 | 36 | -22 | 58 | R | Precentral Gyrus | 4a |
| | 6.55 | -46 | -30 | 65 | -46 | -26 | 60 | L | Postcentral Gyrus | 1 |
| | 6.30 | 46 | -38 | 45 | 46 | -34 | 40 | R | SupraMarginal Gyrus | |
| | 6.08 | 44 | -74 | -21 | 44 | -70 | -26 | R | Cerebelum Crus 1 | |
| | 5.81 | -12 | -72 | 69 | -12 | -68 | 64 | L | Precuneus | |
| | 5.77 | -10 | 14 | 47 | -10 | 18 | 42 | L | Superior Medial Gyrus | |
| | 5.65 | -14 | -4 | 5 | -14 | 0 | 0 | L | Pallidum | |
| | 5.43 | -18 | -72 | -25 | -18 | -68 | -30 | L | Cerebelum Crus 1 | |
| | 5.29 | 54 | 10 | -13 | 54 | 14 | -18 | R | Temporal Pole | |
| | 5.15 | -20 | 14 | 1 | -20 | 18 | -4 | L | Putamen | |
| | 5.13 | -60 | -22 | 27 | -60 | -18 | 22 | L | Postcentral Gyrus (OP 1) | |
| | 4.98 | 26 | 8 | -15 | 26 | 12 | -20 | R | Inferior Frontal Gyrus (p. Orbitalis) | |
| | 4.85 | 26 | 16 | 9 | 26 | 20 | 4 | R | Putamen | |
| | 4.46 | -42 | 48 | 3 | -42 | 52 | -2 | L | Middle Orbital Gyrus | |
| | 4.46 | -12 | -24 | 9 | -12 | -20 | 4 | L | Thalamus | |
| | 4.45 | 12 | 8 | 7 | 12 | 12 | 2 | R | Caudate Nucleus | |
| | 3.73 | -52 | 4 | -27 | -52 | 8 | -32 | L | Medial Temporal Pole | |
| 28 | 4.36 | -32 | -10 | -17 | -32 | -6 | -22 | L | Hippocampus (Amyg. LB) | |
| 13 | 3.81 | -14 | -26 | 51 | -14 | -22 | 46 | L | Middle Cingulate Cortex | |

Table 3.6: Activation Table for the contrast Guessing equals Passive Observation.

## SUPPLEMENTARY INFORMATION 1

### DOES THE MNS NEED OBJECTS TO BE ACTIVATED?

Some studies have investigated whether the MNS can respond to actions not directed at objects (Buccino et al., 2001; Umiltà et al., 2001, e.g.). Can the current study provide further insights into this question? While it is true that none of our stimuli had an object physically present, all of the words people had to mime referred to objects (Tab. 3.1 in main article). Accordingly, the production and observation of our gestures may have involved mentally filling in objects that were implied by the gestures. Single-cell recordings in the monkey (Umiltà et al., 2001) support this view: implying the presence of objects can make mirror neurons selective to viewing actions for which the object is not physically present. In monkeys, implying the presence of an object cannot be done by miming the action without the object (which did not trigger activity in Umiltà et al. (2001)), but can be achieved by placing an occluding screen in front of the object. Humans, unlike monkeys, routinely engage in "let's pretend" play as children (Fein, 1981). This raises the question of whether miming might be more effective at implying the presence of objects to the MNS. (Buccino et al., 2001) showed that most of the pMNS seems to respond to the sight of grasping a cup but not when the cup was absent. Our data however shows activity in those parietal nodes of the pMNS in which (Buccino et al., 2001), found activity only for the movies including the object. This may suggest that a simple mimed grasp may fail to conjure up a mental object, but that within the context of our experiment, more elaborate gestures simulating the presence of an object may be more effective. In conclusion, the apparently simple question of whether the pMNS responds to actions not directed to physically present objects becomes more complex if one considers that mimed actions may differ in their effectiveness at conjuring mental representations of objects. A similar difficulty applies to communicative gestures (e.g. Montgomery et al., 2007, "come here!" gesture), which always imply another person as a target. However, what our experiment does show is that in the context of deliberate gestural communication, gestures can recruit the pMNS even if their object is not physically present.

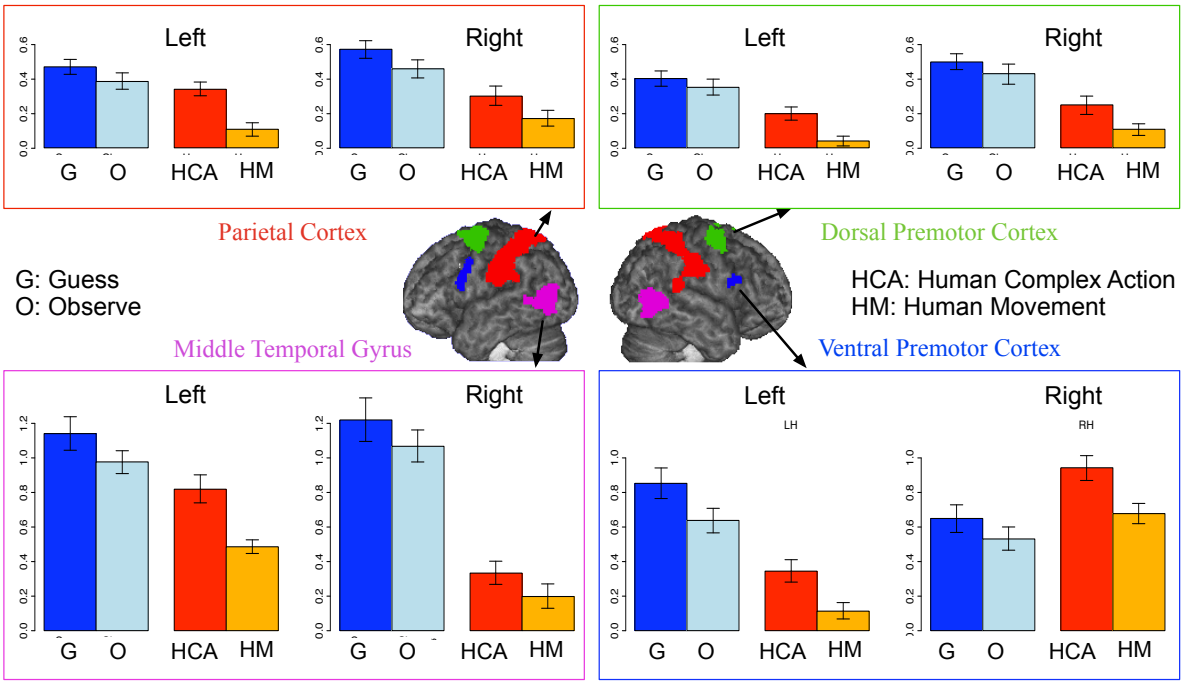



Figure 3.5: Peak percent signal changes of the time courses measured in Gazzola et al. (2007a, their Fig. S3) and those observed during the same time period of the gesture condition in the present experiment.

# 4

# MAPPING THE INFORMATION FLOW FROM ONE BRAIN TO ANOTHER DURING GESTURAL COMMUNICATION

*Published as:* Schippers, M.B., Roebroeck, A., Renken, R., Nanetti, L., Keysers, C. (2010) Mapping the information flow from one brain to another during gestural communication. *Proc. Nat. Acad. Sci. USA.* 107 (20) 9388-9393. doi:10.1073/pnas.1001791107

ABSTRACT

Both the putative Mirror Neuron System (pMNS) and the ventral medial prefrontal cortex (vmPFC) are deemed important for social interaction: the pMNS because it supposedly 'resonates' with the actions of others, the vmPFC because it is involved in mentalizing. Strictly speaking the resonance property of the pMNS has never been investigated. Classical fMRI experiments have only investigated whether pMNS regions augment their activity when an action is seen or executed. Resonance, however, entails more than only 'going on and off together': activity in the pMNS of an observer should continuously follow the more subtle changes over time in activity of the pMNS of the actor. Here we directly explore if such resonance indeed occurs during continuous streams of actions. We let participants play the game of charades, while we measure brain activity of both gesturer and guesser. We then apply a new method to localize directed influences between the brains of the participants: between brains Granger Causality Mapping. Results show that a guesser's brain activity in regions involved in mentalizing and mirroring echo the temporal structure of a gesturer's brain activity. This provides evidence for resonance theories and indicates a fine-grained temporal interplay between regions involved in motor planning and regions involved in thinking about the mental states of others. Furthermore, this new method enables experiments to be more ecologically valid by providing the opportunity to leave social interaction unconstrained. This, in turn, would allow us to tap into the neural substrates of social deficits, such as autism spectrum disorder.

## 4.1 INTRODUCTION

How do humans understand each other? In the last decade two parallel lines of research have investigated this question. On the one hand, the finding that some brain regions and neurons involved in performing an action are also active while viewing

the actions of others (jointly referred to as the Mirror Neuron System, MNS, (Aziz-Zadeh et al., 2006; Chong et al., 2008; Dinstein et al., 2007; Fadiga et al., 1995; Filimon et al., 2007; Gallese et al., 1996; Gazzola et al., 2006; Grèzes et al., 2003; Iacoboni et al., 1999; Keysers and Gazzola, 2009; Kilner et al., 2009; Ricciardi et al., 2009; Turella et al., 2009) has lead to the idea that we understand the actions of others in part by transforming them into the motor vocabulary of our own actions. On the other hand, reflecting on other people's thoughts and beliefs is mediated by another part of the brain, the ventromedial prefrontal cortex (vmPFC, including the anterior cingulate and paracingulate gyrus). This area is consistently activated when we think about other people's mental states (Amodio and Frith, 2006; Frith and Frith, 2006; Gallagher and Frith, 2003; Sommer et al., 2007). Given that we often deduce the beliefs and attitude of others through their actions, it is intuitively appealing to believe that these two networks would work together to achieve a coherent representation of the mental states of others (Keysers and Gazzola, 2007; Brass et al., 2007). However, a recent meta-analysis has shown that these two networks are often found to be dissociated (Overwalle and Baetens, 2009). In what follows, we will first briefly describe some key issues that have limited our understanding of how these two systems contribute to reading the mental states of other individuals during naturalistic situations, and then present a new experimental paradigm to explore their role and interaction in a naturalistic communicative situation.

In humans, the dorsal and ventral premotor, somatosensory cortex, anterior inferior parietal lobule and mid-temporal gyrus have the peculiar property of being active not only when we perform an action but also when we witness similar actions of others (Aziz-Zadeh et al., 2006; Chong et al., 2008; Dinstein et al., 2007; Filimon et al., 2007; Gazzola et al., 2006; Grèzes et al., 2003; Iacoboni et al., 1999; Keysers and Gazzola, 2009; Kilner et al., 2009; Ricciardi et al., 2009; Turella et al., 2009; Keysers, 2009). This set of brain regions has therefore jointly been referred to as the putative MNS (pMNS) (Keysers and Gazzola, 2009). It has been proposed that through this system, the brain of two interacting individuals 'resonate' with each other: "other people's mental states are represented by [...] tracking [...] their states with resonant states of one's own" (Gallese and Goldman, 1998). In this context, the term'resonance' is used rather loosely and metaphorically; not in a strict physical sense but rather to suggest that the ups and downs in the activity of one person's motor system lead to sequences of actions and rest, which trigger similar ups and downs in the activity of the observer's (Gallese and Goldman, 1998; Gallese et al., 2004; Rizzolatti et al., 2001). This concept of resonance is very influential, however the only case in which this proposed temporal 'tracking through resonant states' has really been tested is for viewing repetitive cyclic ups and downs of the wrist (Borroni et al., 2005). Whether it applies to the natural streams of actions that typically lead us to read the minds of others, e.g. the sight of two gesticulating individuals on the side of the road, remains untested. This is

because experimental designs so far have merely tested whether the pMNS becomes active at the transition between a control condition and the sight of a single complex action. This shows that the pMNS of the observer is indeed triggered by the sight of an action. Single cell recordings (Keysers et al., 2003; Mukamel et al., 2010) and magnetoencephalography (Caetano et al., 2007) show that the temporal profile of this activity is indeed similar during action observation and execution, potentially providing a neural basis for resonance. The concept of resonance, however, entails more than only 'going on and off together' at the beginning and end of a single action: it involves a continuous tracking of the more subtle changes in activity during the execution and observation of entire streams of action. Natural social interactions are composed of complex sequences of actions where it is often difficult to know when one action ends and another starts. Here we will directly explore if such resonance indeed occurs within the pMNS during such continuous streams of actions.

The literature on the role of the vmPFC in social interaction suffers from another problem. The vmPFC not only seems to be involved in reflecting on the mental states of others (Amodio and Frith, 2006), it is also one of the brain regions that systematically decreases its activity whenever participants process external stimuli (the 'default network') (Raichle and Snyder, 2007). Studies investigating mentalizing and the default network show strongly overlapping results while exploring seemingly very different functions (Spreng et al., 2009). The fact that the vmPFC is not typically found to be active while people observe the behaviors of others (Overwalle and Baetens, 2009) is therefore difficult to interpret: is activity due to interpreting another mind masked by the fact that its overall level of activity is reduced compared to baseline because of attention to external stimuli? A powerful way to examine this possibility would be to look at the activity of the vmPFC during the observation of longer streams of actions, as the overall level of activity in the default network might then be decreased but the subtle ups-and-downs could still reflect the mentalizing activity of this region in response to the sequence of actions.

Here, participants played the game of charades in the MR-scanner to allow us to examine brain activity during longer streams of gestures. The game of charades was chosen because its success as a commercial game shows how powerfully it triggers the naturalistic motivation to communicate a mental state to a partner through hand actions. It also has the advantage of making the participants generate and observe streams of actions that are naturalistic both in duration and complexity. Given that the type of gestures involved in this game are hand actions, charades can serve to examine the unresolved issue of whether the pMNS would make two individuals' brains resonate during longer streams of actions. Furthermore, since the aim of charades is also to make one player guess a concept that is in the mind of the other player, it is also a powerful instrument to check if fluctuations in the activity of the vmPFC during longer streams of gestures could reflect mentalizing processes triggered

by the behavior of another individual. Indeed both the pMNS (Montgomery et al., 2007; Pazzaglia et al., 2008; Schippers et al., 2009) and the vmPFC (Montgomery et al., 2007) have been implicated in the observation of single gestures, maximizing our chances to examine the yet unexplored issue of whether the activity of these regions during longer streams of actions would indeed resonate with the activity of the brain of the gesturer.

We therefore asked couples to take turns in the fMRI scanner while we measured their brain activity. Each partner knew that on half of the trials, they would see a word on a screen and would have to gesture this word into a video camera for their partner later to guess; and on the other half, they would see a video of their partner's gestures and have to guess what the word had been. Using this manipulation, a single fMRI scanner was enough to measure the brain activity both when one person generates gestures and (later) when another person decodes these very gestures. By aligning the time courses of the two brains' activity, as measured using fMRI, relative to the video recording, we can then directly investigate the temporal coupling of the two brains activity during gestural communication. In order to do this quantitatively, we introduce a new analysis method: we extend Granger Causality Mapping originally used to track information flow within a brain (Roebroeck et al., 2005) to a between brain Granger Causality Mapping (bbGCM, Fig. 4.1 and Supplementary Information 2). BbGCM quantifies the influence from a selected seed region Y in the gesturer's brain to all voxels Xi of the guesser's brain by statistically comparing the G-causalities in both directions, i.e. $(Y \rightarrow X_i) - (X_i \rightarrow Y)$ (Roebroeck et al., 2005). A preliminary analysis of the same data using a traditional general linear model approach ignoring the temporal relationship between the brain activities of each couple has been published previously (Schippers et al., 2009) and shows involvement of pMNS areas, but not the vmPFC. Using bbGCM, however, we will show that even if one ignores the beginning and end of a gesture, activity in both the pMNS and the vmPFC of the observer does carry fine grained information about the time course of the activity in the brain of the gesturer, providing a powerful demonstration of resonance across brains during gestural communication.

## 4.2 MATERIALS AND METHODS

### *Participants*

Twelve couples (total: 24 participants) were scanned while playing the game charades. Four participants had moved more than the voxel-size during the gesturing phase, which lead us to exclude 3 couples from the data analysis that contained these participants. All the analyses in this paper are performed on 18 participants. The mean age

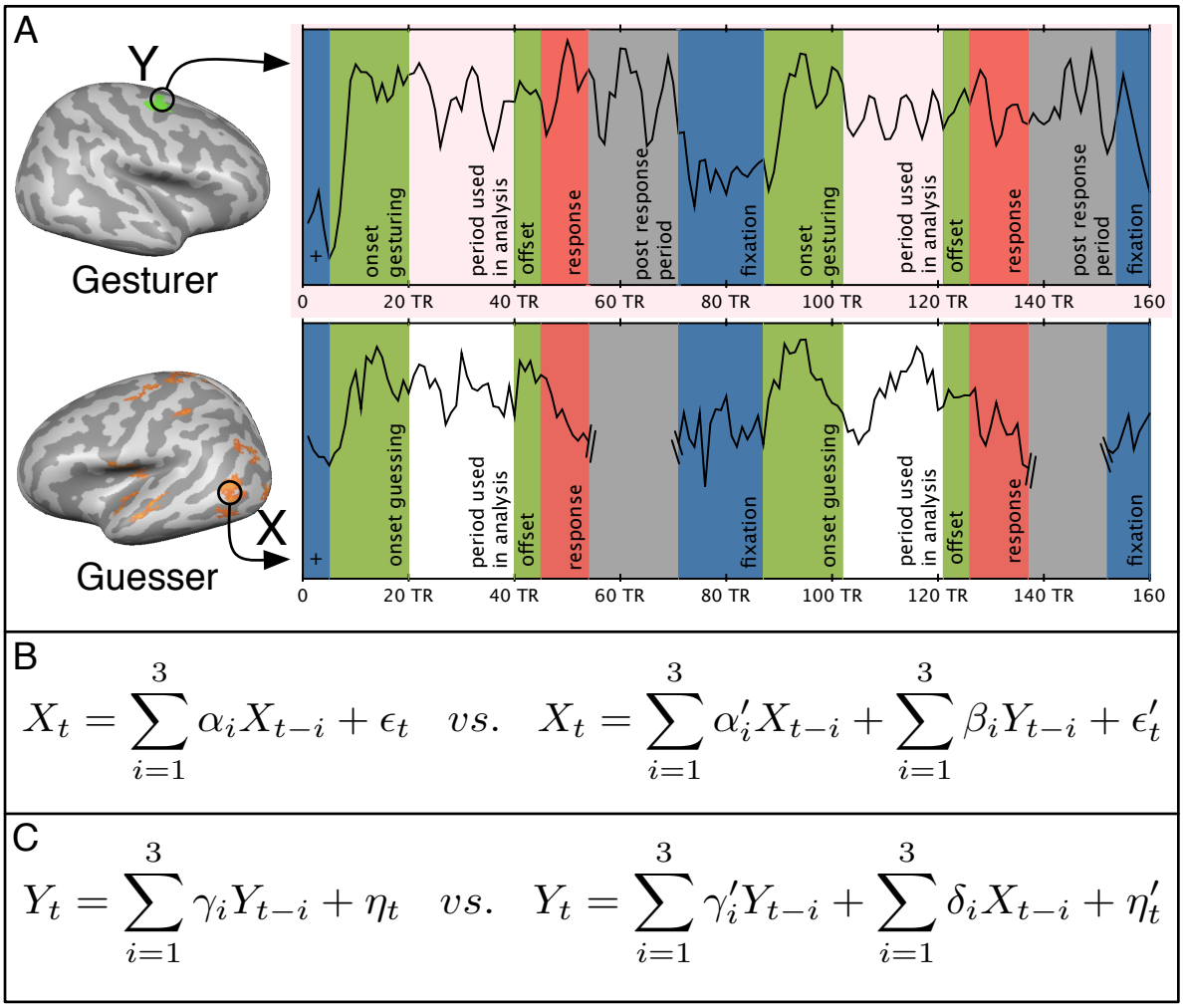



Figure 4.1: Between brains Granger Causality during active guessing. (A) Time series X in the guesser's brain is stimulus aligned with Y in the gesturer's brain. The fixation periods between words (blue) are discarded. Guesser typically responded before the recorded 90s. Response (red) and post-response (grey) periods were thus removed. 15TR at onset and 5TR at offset of each guessing period were trimmed to remove transients. (B) Two regressions are compared at autoregressive order three (see Methods): one including only the past of X itself, and one additionally considering the past of Y. The residual error (variance) reduction from $\sigma^2(\epsilon)$ to $\sigma^2(\epsilon')$ quantifies how much Y G-causes X. (C) Reverse regressions are compared, and differential, directed influences exists if one variable helps more in predicting the other (see also Supplementary Information 2).

of the participants was 27.5 ± 3.8 years. Each couple consisted of a man and a woman involved in a romantic relationship for at least 6 months. More details are described in Supplementary Information 1.

*Task / Experimental Design*

The experiment consisted of two separate sessions on different days. In the first session, the couple was required to play the game of charades. In the second, detailed anatomical scans and a passive observation control condition were acquired. For the game of charades, participants took turns going into the scanner, alternating gesturing and guessing of words. Words were either objects (for example nutcracker, watch, pencil sharpener) or actions (for example painting, knitting, shaving, see Table 4.1 in Supplementary Information 1). Each participant performed two gesture and two guess runs in which they gestured 14 words and guessed 14 words in total (7 per run).

During a gesture run, the participant was presented with a word on the screen and was instructed to communicate this word to his or her partner by means of gestures. Every word had to be gestured for 90 seconds and was then followed by a 20 seconds fixation cross. During a guess run, the participant was shown the movies that were recorded in the gesture run of their partner. The task they had to perform was to guess what their partner was trying to gesture to them. Participants were asked to consider the gestures for at least 50 seconds before committing to a specific interpretation of the gestures. This was done to ensure at least 50 seconds of data in each trial to examine the time course of activity using between brains Granger causality. As a control condition for the guess run, the participants watched the movies they had seen during the guessing condition again. This time, they were instructed not to guess what was gestured, but only to passively view them. More details are described in Supplementary Information 1.

*Granger causality analyses*

Granger causality analyses were performed as described in Roebroeck et al. (2005) but applied here to data from different brains (see Supplementary Information 2). In short, given two time-series (for a seed and another point on the cortical surface), autoregressive models are estimated that quantify G-causality. Given a seed, maps are created that specify G-causal influence from the seed in the gesturer to all of the guesser's brain, as well as influence in the reverse direction, i.e. from anywhere in the guesser's brain to the seed in the gesturer's brain. These two directions of G-causality are then subtracted from each other to generate differential G-causality maps, such

that positive values indicate more G-causality from the gesturer to the guesser than from the guesser to the gesturer. A separate differential G-map was calculated for each of the 8 seed regions (see below) for each participant. These differential G-causality maps were then taken, separately for each seen region map, to the second level (see below) and thresholded for multiple comparisons at $p<0.05$ using a cluster threshold determined by a Monte Carlo simulation method (Forman et al., 1995; Hagler et al., 2006). The order of the estimated autoregressive models was 3, i.e. the 3 preceding time points are taken into account to predict the current activity, corresponding to ~4 seconds (3TR).

*Seed ROIs*

The ROIs that were used as seeds in the between brains Granger causality analysis were defined as those 'mirror' areas that were active both during gesturing and guessing using a traditional GLM analysis on the same data (Schippers et al., 2009).

*Instantaneous Motion Energy GLM*

We extracted motion energy from the gesture movies using Matlab. For two consecutive frames of the recorded movies, motion energy was quantified in every pixel as the sum of the squared differences in the red, green and blue channels and then summed over all pixels. This time course was then mean corrected, convolved with the hemodynamic response function and sampled at the acquisition rate of the fMRI signal (TR = 1.33s).

## 4.3 RESULTS

### 4.3.1 *BbGCM: gesturer to guesser*

The resulting bbGCMs are shown in Figure 4.2 separately for each ROI (top 4 rows), as well as summarized over all ROIs (bottom row). They show that activity in the pMNS of the gesturer indeed predicts brain activity in the brain of the guesser more than the other way around (warm colors). Given that we used only the very recent past (4 seconds) of the gesturer's brain activity in the analysis, this provides the first evidence, to our knowledge, that the moment-to-moment activity in the guesser's pMNS indeed mirrors the close past of the gesturer's pMNS activity during gestural communication. Notably, regions of the vmPFC cortex (including the anterior cingulate and paracingulate gyrus) are also G-caused by activity in the pMNS. The opposite directionality (guesser to gesturer) is much rarer (only found in a small region on

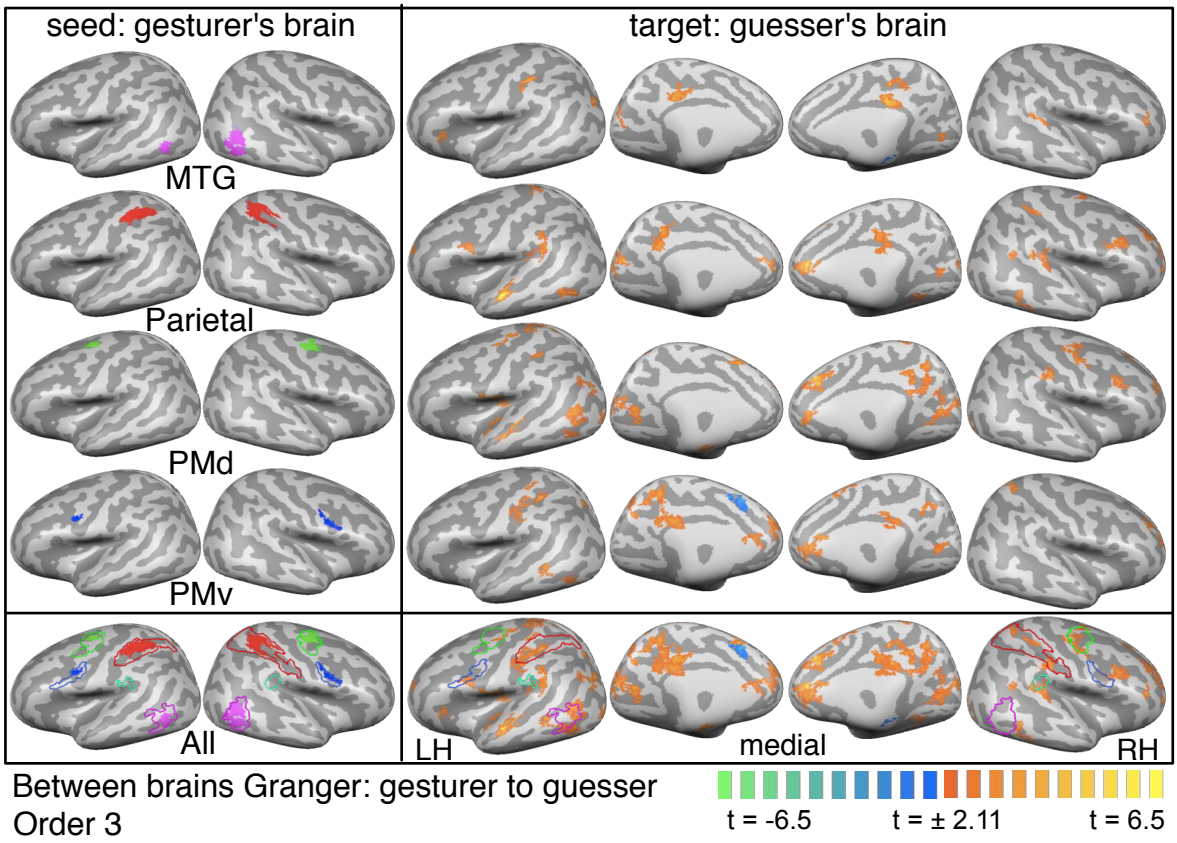



Figure 4.2: Results of second-level bbGCM. Granger analyses executed separately for the left and right seed are shown together. The right side represents the guesser's brain showing t-values of the paired t-test between gesturer → guesser versus guesser → gesturer G-causality (Random Effects, n=18). Upper four rows: differential G-causality originating from the seeds on the left. Bottom row: summary of all seeds (solid colors, left) and bbGCMs (right) with an outline of the pMNS according to a traditional GLM (Schippers et al., 2009) for visual orientation. BbGCM maps are statistically thresholded at $p<0.05$ corrected for multiple comparisons by using a Monte Carlo simulation-based cluster-size threshold adjustment (Forman et al., 1995; Hagler et al., 2006)

the mesial wall, cold colors). This suggests that bbGCM is indeed able to track the prevalent direction of information flow between brains which has to be in the gesturer to guesser direction given that the guesser could see the gesturer in the video but not vice versa.

### 4.3.2 *BbGCM: gesturer to passive observation*

To examine whether these results were dependent on the observer actively trying to guess the meaning of the observed gestures, we had participants watch the same gesture movies again but with an instruction not to actively interpret the movies (see Methods). We computed between-brain influences between the gesturer's brain activity while generating the gestures and their partner's during this control condition (Fig. 4.3A and Fig. 4.4). We then directly contrasted these results with those during active guessing (Fig. 4.2 and Fig. 4.3B). Directed influence was significantly reduced within the ventral premotor and parietal regions associated with the pMNS when directly comparing the two situations (Fig. 4.3C and Fig. 4.5), and differential G-causality was more consistent during deliberate guessing than passive viewing (Fig. 4.3A vs. Fig. 4.3B). This suggests that a task to decode gestures does influence the consistency (and therefore statistical significance) with which an observer's brain time locks onto the gestures, and thereby pMNS activity, of the gesturer. An instruction not to interpret the gestures however cannot ensure that participants indeed refrained entirely from interpretation and a traditional GLM analyses of the same data set (Schippers et al., 2009) showed that overall activity during active guessing and passive observation is indeed similar. Accordingly, contrasting bbGCM during active guessing and passive viewing is a very conservative approach to localizing the neural basis of gesture interpretation that would exclude all neural processes that are triggered automatically by the vision of gestures. However, at debriefing, participants reported having interpreted the gestures at least less consistently during passive viewing than active guessing, and we did find differences in bbGCM results. This shows that an instruction not to interpret the gestures of a partner does seem to partially decouple the observer's brain regions from the pMNS of the gesturer.

### 4.3.3 *BbGCM: gesturer to a random guesser*

To further test whether bbGCM is indeed identifying information flow based on the fine-grained temporal chain of behaviors that makes each social interaction unique, we recalculated bbGCMs while pairing each gesturer's brain activity with that of a randomly selected guesser that had viewed different gestures of another gesturer (Fig. 4.3D and Fig. 4.6). Virtually no vertices (vertices refer to the nodes on the cortical surface, and are therefore similar to voxels except that they are on a cortical surface instead of a brain volume) demonstrated significant differential G-causality in this control analysis and there was significantly more G-causality from gesturer to guesser than from gesturer to random guesser (Fig. 4.3E and Fig. 4.7). Because the sequence of words used for each couple was randomized, the original guesser and his/her

randomly selected control guesser saw a different sequence of words being gestured to them. As a yet stricter control analysis, we therefore repeated this control analysis by substituting 'word by word' the time series of the original guesser with that of a randomly selected control viewing the same word being gestured to him/her but by someone else than the original gesturer (Fig. 4.8). Differential G-causality was again significantly stronger for the original gesturer-guesser pair. This provides direct evidence that the brain activity of the guesser was indeed an echo of the unique way in which his/her particular partner generated these gestures and suggests that bbGCM can indeed track the unique way in which two brains time lock onto one another during communication.

### 4.3.4 *GLM: using instantaneous motion energy of the gestures as predictor*

Here we used bbGCM to indentify brain regions involved in tracking the gestures of others. To test whether this new technique can unravel the involvement of brain regions that more traditional techniques do not, we compared bbGCM with a classical GLM in which we enter two regressors. One contains the timing of the gesture movies (as a boxcar function), and the other the fluctuation of instantaneous motion energy within each movie, both of which were convolved with the hemodynamic response function. This analysis shows that while the brain is strongly reacting to the on- and offset of the movies (see Schippers et al., 2009, their Figure 1a), it does not show a correlation between the fluctuations in the movement of the gesturer (as approximated using instantaneous motion energy in the movie) and fluctuations in brain activity of the guesser as the parameter estimates for the predictor motion-energy was not significantly above zero in any cluster (See Fig. 4.9).

## 4.4 DISCUSSION

In the present study, we introduced bbGCM to investigate to which degree two brains 'resonate' during gestural communication. We show that activity in the pMNS and the vmPFC of the guesser is Granger-caused by fluctuations in activity in the pMNS of the gesturer. These findings have three sets of implications. First, they show that pMNS regions indeed 'resonate' across brains, thereby providing evidence for resonance theories (Gallese and Goldman, 1998; Gallese et al., 2004; Rizzolatti et al., 2001). Second, they extend our understanding of the neural basis of gestural communication by providing evidence for a fine grained temporal interplay between regions involved in motor planning (pMNS) and regions involved in thinking about the mental states of others (vmPFC). Third, they demonstrate more generally that G-causality can be used to map directional information flow across brains during social interactions

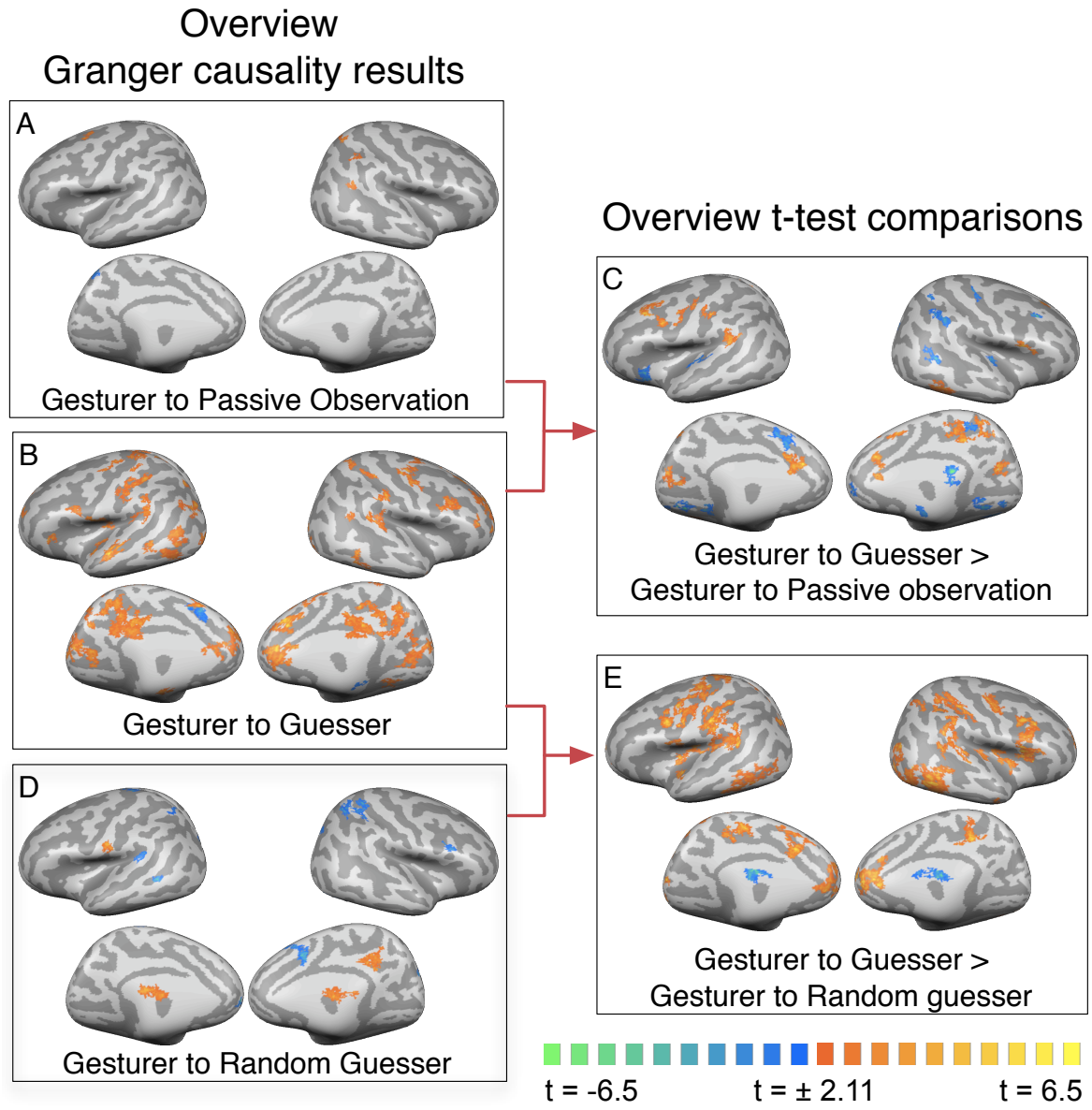



Figure 4.3: Specificity of G-Causality. (A) t-values for (gesturer → passive observer) - (passive observer → gesturer) differential G-causality. (B) Same as (A) but for the original active guessing condition as in Fig. 4.2. (C) Paired t-test of (B)-(A). (D) t-values for (gesturer →random guesser) - (random guesser → gesturer) differential G-causality. (E) Paired t-test for (B)-(D). All maps are statistically thresholded at $p<0.05$ corrected for multiple comparisons by using a Monte Carlo simulation-based cluster-size threshold adjustment (Forman et al., 1995; Hagler et al., 2006), and represent the summary of the results for the 8 seed regions. See Figures 4.4-4.8 for similar maps separately for each seed.

without the need for the experimenter to impose temporal structure on the social interaction.

Before going into details of each of these, we would like to discuss how such G-causality should be interpreted (see also Supplementary Information 2). Directed G-causality between brains has to be mediated by what the guesser can perceive: the observable gestures of the gesturer. BbGCM as we apply it is therefore not a method to determine direct causal interactions across brains, but a method to map brain regions that are at opposite ends of a longer indirect chain of causality that goes through the external world: neural activity in the gesturer causes both the execution of the gestures and the BOLD signal that we measure with fMRI. The video-taped gestures are seen by the guesser, which causes brain activity and an observable BOLD response that we again measure with fMRI. Keeping the indirect nature of the causal pathway in mind, our bbGCM maps brain regions in the receiver that echo the brain activity of the sender. By 'echo' in this context, we mean providing temporal information about the state of the other person's brain region. In general, the directed bbGCM should be interpreted with a further issue in mind. Because it is calculated by contrasting bbGCM in two directions (Brain A → Brain B minus Brain B → Brain A), a value larger than zero is evidence for an influence of brain A on brain B, whereas the opposite is not evidence for a lack of influence from A → B. This is because, in addition to a potential lack of statistical power, negative findings with directed bbGCM could originate from two very different scenarios: because there is no significant information flow in either direction, or because the influence is significant but equally strong in both directions. In our experiment, the latter possibility is unlikely, as the guesser can view the gesturer but the gesturer cannot see the guesser. It is therefore unlikely that the guesser's brain could influence the gesturer's. The second scenario will, however, be more likely in future experiments in which two interacting partners might be able to mutually observe each other in real time.

We still know relatively little about the neural basis of mind reading and gestural communication. However, two sets of brain regions could play a role: pMNS areas and the vmPFC. Our findings support the idea that both play a role: we show that BOLD activity in functionally defined regions of the pMNS and in the vmPFC are G-caused by BOLD activity in the pMNS of the gesturer. This strengthens the idea that simulation and mentalizing both contribute to our interpretation of other people's gestures (Keysers and Gazzola, 2007; de Lange et al., 2008; Thioux et al., 2008). Because of the constraints of traditional data analysis (e.g. GLM analysis), the degree to which the sequences of complex actions composing a gestural phrase (McNeill, 1992) would be parsed (Byrne, 2003) similarly by the gesturer and guesser has never been explored. Our finding of significant G-causality between the pMNS in the two brains using a temporal window of 3 (i.e. regressing the brain activity of the guesser onto the activity in the past 3 volumes = 4 seconds of the gesturer) provides the first evidence,

to our knowledge, that the two brain activities go up and down together in naturalistic gestural communication. Reducing the order to 1 (i.e. considering only the past 1.33 seconds of the gesturer) much reduces this differential G-causality (See Fig. 4.10). This shows that the pMNS of communicative partners indeed resonates with each other (in the loose sense used in this literature) as had been suggested by simulation accounts of mind reading and communication (Gallese and Goldman, 1998; Gallese et al., 2004; Rizzolatti et al., 2001; Rizzolatti and Craighero, 2004). Moreover, it informs us that this resonance is not most evident at the second-to-second time scale of time window 1, but at a more moderate time scale of several (~4s) seconds that is commensurable with the time it takes to plan, generate and perceive a gestural element (McNeill, 1992). BbGCM is therefore a powerful tool to test the temporal resonance phenomenon at the core of simulation accounts of communication. It furthermore shows that the pMNS indeed provides the time-resolved information about the state of the gesturer's motor system that would be required for motor simulation to be useful for communication in a naturalistic context. This adds to the evidence that the excitability of an observer's motor system fluctuates in synch with the repetitive wrist flexion of another individual by showing that the concept of resonance indeed applies to the complex, non repetitive and non rhythmical streams of gestures that are more typical of real mind-reading situations and may have been essential in the early state of language evolution (Rizzolatti and Arbib, 1998). Additionally, this finding dovetails with the observation that brain activity in both these regions predicts the accuracy with which participants can judge the moment to moment emotional state of another individual (Zaki et al., 2009).

The fact that vmPFC activity was also G-caused by pMNS activity in the gesturer is surprising. This region is well known to play a role in inferring mental states from written stories or cartoons (Amodio and Frith, 2006) and activity in this region is increased while participants try to interpret certain gestures (Montgomery et al., 2007). These findings suggest that the vmPFC might be involved in attributing mental states to others, and could do so during gestures and actions. However, a previously published preliminary GLM analysis of our data revealed that this region does not demonstrate more brain activity while guessing gestures than while fixating a cross (Schippers et al., 2009). In addition, during the free-viewing of a Hollywood movie, the vmPFC does not seem to synchronize across viewers (Hasson et al., 2004). The inferential nature of these processes seems to detach brain activity in this region from the exact timing of the stimulus, leading it not to synchronize across viewers and therefore also making it difficult to link activity in this region directly to the stimulus itself. Indeed activity in this region also does not simply correlate with the low level motion contained in the stimuli (see Fig.4.9). Here, on the other hand, we show that activity in this brain region in a guesser does contain information (in the sense of Supplementary Information 2) about the time course of activity in the

regions involved in planning and executing gestures in the gesturer. This supports the idea that the pMNS and mentalizing brain areas may work in concert to derive mental states from observed actions (Keysers and Gazzola, 2007; Thioux et al., 2008). A traditional GLM analysis of the same data (Schippers et al., 2009) may have been unable to detect the involvement of the vmPFC because this region is also part of the default network (Raichle and Snyder, 2007; Spreng et al., 2009). The default network is a set of brain regions that demonstrate augmented metabolism during passive baseline conditions, supposedly because they contain neurons that are involved in the self-referential processes we engage in while not performing a particular task (Raichle and Snyder, 2007; Spreng et al., 2009). During guessing, these self-referential processes would have been suspended, lowering the BOLD in this brain region below baseline. The activity, in this region, of a smaller number of neurons engaging in the mental state attribution required by the game of charades would then have been masked by the concurrent reduction of self-referential activity. Our bbGCM can detect such activity nevertheless, because it examines not whether activity overall goes up or down relative to a baseline condition, but rather whether fluctuations in activity during the stream of gestures covaries with the past activity of the gesture execution system of the gesturer. This changes the interpretation of the same data set compared to a classical GLM analysis which showed that only the pMNS but not the vmPFC demonstrate augmented activity compared to baseline during the active decoding of gestures (Schippers et al., 2009).

A number of control analyses served to establish that bbGCM indeed tracks a specific information flow between two communicating participants. When pairing the time course of a given gesturer with that of a randomly selected guesser, instead of the one that had actually observed the gestures, significantly less between brain influences were observed compared to the analysis with the active guessing condition (Fig. 4.3E). This suggests that bbGCM indeed revealed the specific effect of a particular pattern of gestures on the brain activity of the guesser. Furthermore, we found that active guessing, but not passive viewing of the same gestures leads to significant bbGCM of the pMNS and vmPFC (Fig. 4.3A). This shows that an instruction to actively decode the gestures increases the coherence between the activities in the two brains.

Given that the brain activity of our guesser is not directly caused by brain activity in the gesturer but by the pre-recorded movie of his/her gestures, one might argue that measuring the brain activity of the gesturer is not necessary to map brain regions involved in social information transfer. Instead, quantifying what is in the stimulus would suffice to localize those brain regions in the guesser's brain that respond to that stimulus. We tested this approach by using instantaneous motion energy from the gesture movies as a predictor in a GLM analysis. Results show no correlation between their activity and the instantaneous motion energy from the movies, indicating that using this particular measure with a traditional GLM approach does not provide any

extra information. While alternative approaches to quantifying the content of the stimulus and introducing time-lagged versions of these predictors into a GLM may help, the fundamental problem with such a stimulus centered approach is that quantifying the relevant dimensions of a naturalistic stream of gestures is far from trivial: it is a highly multidimensional stimulus, and transforming it into univariate time series for a GLM requires knowledge of what aspects of the stimulus are relevant for the brain of the observer - a knowledge that we often lack. BbGCM has the elegant property of circumventing this problem altogether and thereby directly testing those theories, like the pMNS 'resonance' theory of mind reading (Gallese and Goldman, 1998; Rizzolatti and Arbib, 1998), which are formulated not as a link between a stimulus and a neural state, but between the neural states of two individuals.

As an alternative for this method, one might show the same gestures to many participants, and examine what brain regions synchronize across participants (Hasson et al., 2004). Between-viewer correlation also has the elegant property of circumventing the problem of quantifying the stimulus, and is conceptually related to our approach. It however has other limitations: it requires many viewers of the same stimulus and, in its standard form, only examines instantaneous dependencies between brain regions (i.e. it does not allow for time shifts between the brain activity of different viewers). BbGCM overcomes these limitations: It can be applied to pairs of interacting partners, with only one participant viewing any particular communicative episode, and is therefore more suited for studying dyadic communication. Additionally, it allows examination of dependencies between brain activity over a longer time period (the G-causality order, in our case 4s), which is more appropriate for the analysis of brain activity during communication, where several seconds can separate the planning of a gesture from its execution and perception. BbGCM and between-viewer correlation could be combined to a between viewer GCM: if the brain activity of one viewer contains information about the brain activity of another viewer, in the absence of direct communication between the viewers, this shared information has to be information about the stimulus. Between-viewer GCM would thus map regions containing information about the stimulus while allowing for slight time shifts across viewers.

More generally, for the field of social neuroscience, our findings show that it is possible to map the brain regions involved in the flow of information across individuals with fMRI without imposing a temporal structure on the social interaction and without depending on certain choices for the quantification of the information in a stimulus. A similar approach seems to be suited for analyzing EEG data during social interactions (Babiloni et al., 2006). Demonstrating bbGCM between one brain region in partner 1 and another region in partner 2 then allows a data driven identification of brain regions that could play a role in the information flow across participants. Much as for other data analysis techniques (Hasson et al., 2004), further experiments that control the content of the stimulus seen by participants are then needed to iso-

late which aspects of the complex interaction were encoded in the brain activity in these brain regions, and virtual lesions using transcranial magnetic stimulation will be needed to examine whether these brain regions are necessary for normal social interactions.

In conclusion, bbGCM has advantages over existing techniques: it can map the information transfer from brain to brain in pairs of participants without having to impose temporal structure on social interactions and without requiring knowledge about the relevant dimensions of the complex social stimulus. Here we used this approach to show for the first time that even for naturalistic streams of gestural communication, the core prediction of MNS theories of communication and mind reading hold: the pMNS of the guesser does indeed reflect moment to moment information about the state of the motor system of the gesturer. In addition, using this method, we narrow the gap between literatures exploring the pMNS and that exploring mentalizing by show that the vmPFC of the observer could add to this mirroring by also resonating with the motor system of the gesturer. More generally, we hope that this technique will enable and inspire the investigation of one of the most defining feature of human beings: their capacity to transfer knowledge from one person to another. In particular, the opportunity to leave the social interaction unconstrained will enable experiments to be more ecologically valid. This, in turn, could allow us to tap into the neural substrates of social deficits and ask questions like: Which neural substrates are responsible for the difficulty autistic individuals have in taking turns during communication?

## ACKNOWLEDGEMENTS

The research was supported by a VIDI grant of the Dutch science foundation (N.W.O.) and a Marie Curie Excellence Grant of the European Commission to CK. MS and CK designed the study with help of AR. MS analyzed the data with guidance from CK and AR and help from LN and RR. CK, MS, and AR wrote the manuscript. We thank V. Gazzola and R. Goebel for help in designing the experiment, and the latter for providing BrainVoyager; P. Toffanin, M. Spezio, and D. Arnstein for critical comments on the manuscript.

# BIBLIOGRAPHY


Amodio, D. M. and Frith, C. D. (2006). Meeting of minds: the medial frontal cortex and social cognition. *Nature Reviews of Neuroscience*, 7(4):268–77.

Aziz-Zadeh, L., Wilson, S. M., Rizzolatti, G., and Iacoboni, M. (2006). Congruent embodied representations for visually presented actions and linguistic phrases describing actions. *Current biology*, 16(18):1818–23.

Babiloni, F., Cincotti, F., Mattia, D., Mattiocco, M., Fallani, F. D. V., Tocci, A., Bianchi, L., Marciani, M. G., and Astolfi, L. (2006). Hypermethods for eeg hyperscanning. *Conference Proceedings: Annual International Conference of the IEEE Engineering in Medicine and Biology Society IEEE Engineering in Medicine and Biology Society Conference*, 1:3666–9.

Barnett, L., Barrett, A. B., and Seth, A. K. (2009). Granger causality and transfer entropy are equivalent for gaussian variables. *Physical Review Letters*, 103(23):238701.

Borroni, P., Montagna, M., Cerri, G., and Baldissera, F. (2005). Cyclic time course of motor excitability modulation during the observation of a cyclic hand movement. *Brain Research*, 1065(1-2):115–24.

Brass, M., Schmitt, R. M., Spengler, S., and Gergely, G. (2007). Investigating action understanding: inferential processes versus action simulation. *Current biology*, 17(24):2117–21.

Byrne, R. W. (2003). Imitation as behaviour parsing. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences*, 358(1431):529–36.

Caetano, G., Jousmäki, V., and Hari, R. (2007). Actor's and observer's primary motor cortices stabilize similarly after seen or heard motor actions. *Proceedings of the National Academy of Sciences of the United States of America*, 104(21):9058–62.

Chong, T. T.-J., Cunnington, R., Williams, M. A., Kanwisher, N., and Mattingley, J. B. (2008). fmri adaptation reveals mirror neurons in human inferior parietal cortex. *Current biology*, 18(20):1576–80.

Cover, T. and Thomas, J. (1991). *Elements of Information Theory*. John Wiley and Sons, New York.

de Lange, F. P., Spronk, M., Willems, R. M., Toni, I., and Bekkering, H. (2008). Complementary systems for understanding action intentions. *Current biology*, 18(6):454–7.

Dinstein, I., Hasson, U., Rubin, N., and Heeger, D. J. (2007). Brain areas selective for both observed and executed movements. *Journal of Neurophysiology*, 98(3):1415–27.

Eickhoff, S., Stephan, K. E., Mohlberg, H., Grefkes, C., Fink, G. R., Amunts, K., and Zilles, K. (2005). A new spm toolbox for combining probabilistic cytoarchitectonic maps and functional imaging data. *Neuroimage*, 25(4):1325–1335.

Fadiga, L., Fogassi, L., Pavesi, G., and Rizzolatti, G. (1995). Motor facilitation during action observation: A magnetic stimulation study. *Journal of Neurophysiology*, 73(6):2608–2611.

Filimon, F., Nelson, J. D., Hagler, D. J., and Sereno, M. I. (2007). Human cortical representations for reaching: mirror neurons for execution, observation, and imagery. *Neuroimage*, 37(4):1315–28.

Fischl, B., Sereno, M. I., and Dale, A. M. (1999). Cortical surface-based analysis. ii: Inflation, flattening, and a surface-based coordinate system. *Neuroimage*, 9(2):195–207.

Forman, S. D., Cohen, J. D., Fitzgerald, M., Eddy, W. F., Mintun, M. A., and Noll, D. C. (1995). Improved assessment of significant activation in functional magnetic resonance imaging (fmri): use of a cluster-size threshold. *Magnetic resonance in medicine*, 33(5):636–47.

Frith, C. D. and Frith, U. (2006). The neural basis of mentalizing. *Neuron*, 50(4):531–534.

Gallagher, H. L. and Frith, C. D. (2003). Functional imaging of 'theory of mind'. *Trends in Cognitive Sciences*, 7(2):77–83.

Gallese, V., Fadiga, L., Fogassi, L., and Rizzolatti, G. (1996). Action recognition in the premotor cortex. *Brain*, 119(2):593–609.

Gallese, V. and Goldman, A. (1998). Mirror neurons and the simulation theory of mind-reading. *Trends in Cognitive Sciences*, 12:493–501.

Gallese, V., Keysers, C., and Rizzolatti, G. (2004). A unifying view of the basis of social cognition. *Trends in Cognitive Sciences*, 8(9):396–403.

Gazzola, V., Aziz-Zadeh, L., and Keysers, C. (2006). Empathy and the somatotopic auditory mirror system in humans. *Current biology*, 16(18):1824–9.

Gazzola, V. and Keysers, C. (2008). The observation and execution of actions share motor and somatosensory voxels in all tested subjects: Single-subject analyses of unsmoothed fmri data. *Cerebral Cortex*, 19(6):1239–1255.

Granger, C. (1969). Investigating causal relations by econometric models and cross-spectral methods. *Econometrica*, 37(3):424–438.

Granger, C. (1980). Testing for causality: A personal viewpoint. *Journal of Economic Dynamics and Control*, 2:329–352.

Grèzes, J., Armony, J., Rowe, J., and Passingham, R. E. (2003). Activations related to mirror and canonical neurones in the human brain: An fmri study. *Neuroimage*, 18:928–937.

Hagler, D. J., Saygin, A. P., and Sereno, M. I. (2006). Smoothing and cluster thresholding for cortical surface-based group analysis of fmri data. *Neuroimage*, 33(4):1093–103.

Hasson, U., Nir, Y., Levy, I., Fuhrmann, G., and Malach, R. (2004). Intersubject synchronization of cortical activity during natural vision. *Science*, 303(5664):1634–40.

Holmes, C. J., Hoge, R., Collins, L., Woods, R., Toga, A. W., and Evans, A. C. (1998). Enhancement of mr images using registration for signal averaging. *Journal of computer assisted tomography*, 22(2):324–33.

Iacoboni, M., Woods, R., Brass, M., Bekkering, H., Mazziotta, J. C., and Rizzolatti, G. (1999). Cortical mechanisms of human imitation. *Science*, 286(5449):2526–2528.

Keysers, C. (2009). Mirror neurons. *Current Biology*, 19(21):R971–3.

Keysers, C. and Gazzola, V. (2007). Integrating simulation and theory of mind: from self to social cognition. *Trends in Cognitive Sciences*, 11(5):194–6.

Keysers, C. and Gazzola, V. (2009). Expanding the mirror: vicarious activity for actions, emotions, and sensations. *Current Opinion in Neurobiology*, 19(6):666–71.

Keysers, C., Kohler, E., Umiltà, M., Nanetti, L., Fogassi, L., and Gallese, V. (2003). Audiovisual mirror neurons and action recognition. *Experimental Brain Research*, 153(4):628–636.

Kilner, J. M., Neal, A., Weiskopf, N., Friston, K. J., and Frith, C. D. (2009). Evidence of mirror neurons in human inferior frontal gyrus. *The Journal of Neuroscience*, 29(32):10153–9.

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

Montgomery, K. J., N.Isenberg, and Haxby, J. V. (2007). Communicative hand gestures and object-directed hand movements activated the mirror neuron system. *Social Cognitive and Affective Neuroscience*, 2(2):114–122.

Mukamel, R., Ekstrom, A. D., Kaplan, J., Iacoboni, M., and Fried, I. (2010). Single-neuron responses in humans during execution and observation of actions. *Current biology*, 20(8):750–756.

Oldfield, R. (1971). The assessment and analysis of handedness: the edinburgh inventory. *Neuropsychology*, 9(1):97–113.

Overwalle, F. V. and Baetens, K. (2009). Understanding others' actions and goals by mirror and mentalizing systems: a meta-analysis. *Neuroimage*, 48(3):564–84.

Pazzaglia, M., Smania, N., Corato, E., and Aglioti, S. M. (2008). Neural underpinnings of gesture discrimination in patients with limb apraxia. *The Journal of Neuroscience*, 28(12):3030–41.

Pearl, J. (2000). *Causality: Models, Reasoning and Inference*. Cambridge Univ Press, Cambridge, UK and New York.

Raichle, M. E. and Snyder, A. Z. (2007). A default mode of brain function: a brief history of an evolving idea. *Neuroimage*, 37(4):1083–90; discussion 1097–9.

Ricciardi, E., Bonino, D., Sani, L., Vecchi, T., Guazzelli, M., Haxby, J. V., Fadiga, L., and Pietrini, P. (2009). Do we really need vision? how blind people "see" the actions of others. *Journal of Neuroscience*, 29(31):9719–24.

Rizzolatti, G. and Arbib, M. A. (1998). Language within our grasp. *Trends in Neurosciences*, 21(5):188–194.

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

Rizzolatti, G., Fogassi, L., and Gallese, V. (2001). Neurophysiological mechanisms underlying the understanding and imitation of action. *Nature Reviews of Neuroscience*, 2(9):661–70.

Roebroeck, A., Formisano, E., and Goebel, R. (2005). Mapping directed influence over the brain using granger causality. *Neuroimage*, 25:230–242.

Schippers, M. B., Gazzola, V., Goebel, R., and Keysers, C. (2009). Playing charades in the fmri: are mirror and/or mentalizing areas involved in gestural communication? *PLoS ONE*, 4(8):e6801.

Schreiber, T. (2000). Measuring information transfer. *Physical Review Letters*, 85(2):461–4.

Shannon, C. and Weaver, W. (1949). *The mathematical theory of communication*. University of Illinois Press, Urbana, Illinois.

Sommer, M., Döhnel, K., Sodian, B., Meinhardt, J., Thoermer, C., and Hajak, G. (2007). Neural correlates of true and false belief reasoning. *Neuroimage*, 35(3):1378–84.

Spreng, R. N., Mar, R. A., and Kim, A. S. N. (2009). The common neural basis of autobiographical memory, prospection, navigation, theory of mind, and the default mode: a quantitative meta-analysis. *Journal of Cognitive Neuroscience*, 21(3):489–510.

Thioux, M., Gazzola, V., and Keysers, C. (2008). Action understanding: how, what and why. *Current biology*, 18(10):R431–4.

Turella, L., Erb, M., Grodd, W., and Castiello, U. (2009). Visual features of an observed agent do not modulate human brain activity during action observation. *Neuroimage*, 46(3):844–53.

Wiener, N. (1956). Theory of prediction. *Modern Mathematics for Engineers, Series 1*.

Zaki, J., Weber, J., Bolger, N., and Ochsner, K. (2009). The neural bases of empathic accuracy. *Proceedings of the National Academy of Sciences of the United States of America*, 106(27):11382–7.

## SUPPLEMENTARY INFORMATION 1

### MATERIALS AND METHODS IN DETAIL

#### *Participants*

Twelve couples (total: 24 participants) were scanned while playing the game charades. Four participants had moved more than the voxel-size during the gesturing phase, which lead us to exclude 3 couples from the data analysis that contained these participants. All the analyses in this paper are performed on 18 participants. The mean age of the participants was 27.5 ± 3.8 years. Each couple consisted of a man and a woman involved in a romantic relationship for at least 6 months. We included this criterion for two reasons. First, we expected that the participants would be more motivated and more at ease during gesturing if they knew it was their partner who had to interpret their gestures. Second, we expected them to have a better or faster understanding of each other's gestures since they knew each other better than a stranger would. Our aim was not to study specifically romantic processes, but simply to let the participants feel as comfortable as possible during the game. Participants were prescreened to exclude those with a history of neurological or psychiatric illness. Participants were also asked not to drink coffee before scanning commenced. The participants freely consented to participating in the study by signing an informed consent form and were scaled for their right-handedness on the Edinburgh Right-handedness scale (Oldfield, 1971). This entire study was approved by the Medical Ethics Committee of the University Medical Centre Groningen (2007/080).

#### *Task / Experimental Design*

The experiment consisted of two separate sessions on different days. In the first session, the couple was required to play the game of charades. In the second, detailed anatomical scans and a passive observation control condition were acquired. For the game of charades, participants took turns going into the scanner, alternating gesturing and guessing of words. Words were either objects (for example nutcracker, watch, pencil sharpener) or actions (for example painting, knitting, shaving, see Table 4.1). Each participant performed two gesture and two guess runs in which they gestured 14 words and guessed 14 words in total (7 per run). The set of words used was the same for each couple, but word order was randomized between participants.

After the last gesture-session, a T1-weighted anatomical image was acquired. See Fig. 4.11 for a schematic overview of the experiment.

*Gesture run*: during a gesture run, the participant was presented with a word on the screen and was instructed to communicate this word to his or her partner by means of gestures (see Table 4.1 for an overview of the words). Every word had to be gestured for 90 seconds. Prior to scanning participants were trained not to repeat the same gesture over and over again, but to keep generating new gestures to provide their partner with multiple sources of information. The participant could see how much time he/she needed to keep gesturing by a progress bar on the screen. A fixation cross was presented for 20 s after each word, which served as our baseline (see Fig. 4.11 for a detailed overview). The gestures were recorded from the control room of the MR-scanner with a video camera (Sony DSR-PDX10P). After the participant had gestured seven words, he/she was taken out of the scanner and went into the waiting room, while his/her partner went into the scanner to guess what he/she had gestured. During this changeover, the experimenter cut the recording of the gestures into movies of 90s in which the participant gestured a word. To ensure that the movies were cut at exactly the moment the word was presented to the gesturing participant, the stimulus computer's sound card emitted a sound at the beginning of word presentation. The output of the sound card was connected to the audio input of the video camera, thus allowing the auditory signal to serve as a marker for cutting. To minimize the amount of head motion in the participants, the upper arms of the participant were fixed to the bed by means of a Velcro strap band. This left the participant free to gesture with his lower arms and fingers, which still allowed 86% percent correct gesture recognition.

*Guess run*: during a guess run, the participant was shown the movies that were recorded in the gesture run of their partner. The task they had to perform was to guess what their partner was trying to gesture to them. Participants were asked to consider the gestures for at least 50 seconds before committing to a specific interpretation

| Actions | | Objects | |
|---|---|---|---|
| peel fruit | fold | nutcracker | telephone |
| ride a bike | drive a car | pencil sharpener | winding stairs |
| shuffle cards | play the piano | pistol | ashtray |
| polish nails | squeeze fruit | electric eel | bow |
| juggle | paint | watch | handcuffs |
| knit | light fireworks | board game | glove |
| throw a snowball | shave | canoe | cork screw |

Table 4.1: Action and object words used in the charades

of the gestures. This was done to ensure at least 50 seconds of data in each trial to examine the time course of activity using between brains Granger causality. This was done by showing a progress bar below the movie, changing from red to green after 50 seconds, indicating the beginning of the period (50-90s post stimulus onset) during which participants could decide on their interpretation of the gestures, whenever they felt confident, by pressing a button on their 4-button button-box, triggering the appearance of a multiple choice screen. In the multiple-choice menu they had to choose the correct word from five alternatives. One of the alternatives was always 'none of the above' and the correct answer was always present in the multiple-choice menu. The correct answer was never the option 'none of the above'. This marked the end of a trial. Two consecutive trials were separated by 20s of a white fixation cross against a black background, which served as our baseline (Fig. 4.11).

*Passive observation run*: As a control condition for the guess run, the participants watched the movies they had seen during the guessing condition again. This time, they were instructed not to guess what was gestured, but only to passively view them. We are aware that such instructions cannot ensure that participants entirely stopped to interpret the gestures, but at debriefing, participants reported having interpreted the gestures at least less consistently than during the guess run. To keep the run exactly the same as the original guess run, the movie stopped at the moment the participant during the original run had pushed the button. The same multiple-choice menu then appeared and the participant had to answer again. This time, however, they had to select the word written in green letters. The green word was the correct answer. A fixation cross was presented between two consecutive trials for 20 seconds and served as our baseline.

*Data Acquisition*

Functional imaging data was recorded with a Philips 3.0T MR scanner, using gradient echo planar imaging (EPI) and an 8-channel head coil using SENSE technology. T2* weighted images revealed changes in blood oxygen level. Volume repetition time (TR) was 1.33 seconds. The whole brain was scanned in 28 (axial) slices with a thickness of 4.5mm. Further imaging parameters include echo time (TE) 28 ms, field of view 224 x 224 mm, 64 x 62 matrix, SENSE acceleration factor 2.4, ensuing 3D voxel size 3.5 x 3.5 x 4.5mm. This set of imaging parameters were chosen to cover the entire neo-cortex while at the same time providing a TR short enough to expect sufficient power in a Granger Causality analysis (Roebroeck et al., 2005). In the first session, a fast structural image ("fast anatomy") was acquired of the participant's brain, while in the second session an additional structural image of higher resolution was acquired. Both were structural, T1-weighted images acquired with a T1TFE sequence (echo

time 3.5 ms, repetition time 7.6 ms, 224 x 160 x 256 matrix, 1 x 1 x 1 mm3 voxels).

*Data pre-processing*

All analyses and preprocessing was performed in BrainVoyager QX 1.10 along with custom written C++ code for the between brains Granger causality analyses. The pre-processing steps included slice scan time correction, 3D motion correction and temporal filtering (consisting of linear trend removal and a high pass filter with a cut-off at 0.004Hz). The images were not smoothed spatially. Functional images were co registered with the structural images and morphed into Talairach space.

The structural images were corrected for inhomogeneity (to improve segmentation results), normalized into Talairach space after which the cortical grey matter/white matter boundary was segmented into a topologically correct surface representation. After segmentation, the cortical surface representations of all subjects were aligned using a cortical curvature based alignment procedure. This procedure aligns the sulci and the gyri of the different brains using their cortical curvature-maps (Fischl et al., 1999). Included in the cortex-based alignment was also the Colin (27) brain (Holmes et al., 1998). The fMRI time courses were resampled on the curvature aligned cortical surface representations.

*Behavioral Results*

During guessing the participants were asked to consider each movie for at least 50 seconds. After the 50s they could push the button when they thought they knew what was being gestured to enter the multiple-choice menu. The average latency to response was 58 seconds. Participants were equally accurate on both categories: 82.5% of the object words were guessed correctly against 86.5% of the action words ($t(17)=-1$, $p>.33$). We did not find a significant difference between the two types of gestures, neither in terms of latency to respond (58.7s ± 6.5s for action and 60.8s ± 6.8s for object words, $t(17)=1.16$, $p>.26$) nor in terms of accuracy (6.06 ± 0.73sd correct out of 7 action and 5.78 ± 1.11sd correct our of 7 object words, $t(17)=-1$, $p>.33$). Words that were guessed incorrectly were watched significantly longer than words that were guessed correctly: 58s ± 5s for the 289 correct guesses versus 68s ± 12s for the 47 incorrect guesses ($t(16)=-4.41$, $p<.0005$).

*Granger causality analyses*

Granger causality analyses were performed as described in Roebroeck et al. (2005) but applied here to data from different brains Supplementary Information 2. In

short, given two time-series (for a seed and another point on the cortical surface), autoregressive models are estimated that quantify G-causality. Given a seed in the gesturer, maps are created that specify G-causal influence from the seed in the gesturer to all of the guesser's brain, as well as influence in the reverse direction, i.e. from anywhere in the guesser's brain to the seed in the gesturer's brain. These two directions of G-causality are then subtracted from each other to generate differential G-causality maps, such that positive values indicate more G-causality from the gesturer to the guesser than from the guesser to the gesturer.

This differential G-causality measure was used for three reasons. First, it generates values that are approximately normally distributed, with a mean of zero under the null hypothesis of an absence of (indirect) causal relationship at the neural level between the two brains, and are thus suitable for parametric testing at the second level (see below and Roebroeck et al., 2005). Second, given that the guesser saw the gesturer but not the other way around (one-way video feed), we know that there should be more information flow gesturer → guesser than guesser → gesturer (the latter could only be due to anticipatory neural computations), allowing a directed hypothesis testing. Thirdly, it has been shown that when the BOLD signal is used to estimate G-causality, the differential G-causality (i.e. X → Y - Y → X) is more robust than testing the individual components due to the filtering properties of the hemodynamic response and the relatively low sampling rate of fMRI (Roebroeck et al., 2005). A separate differential G-map was calculated for each of the 8 seed regions (see below) for each gesturer. These differential G-causality maps were then taken, separately for each seed region map, to the second level (see below) and thresholded for multiple comparisons at $p<0.05$ using a cluster threshold determined by a Monte Carlo simulation method (Forman et al., 1995; Hagler et al., 2006). The order of the estimated autoregressive models was 3, i.e. the 3 preceding time points are taken into account to predict the current activity, corresponding to ~4 seconds (3TR). This interval was chosen a priori because it roughly covers the time it takes for a typical gestural phrase to unfold (McNeill, 1992), and would therefore permit the analysis to include time points involved in planning a gestural phrase in the regression of time points involved in perceiving the end of the phrase. In addition, an early exploratory analysis confirmed this expectation. Performing the analysis with orders of 1TR or 5TR revealed similar but weaker effects (Fig. 4.10, due to too little relevant history used in the order-1 models and power loss due to the increased amount of estimated parameters in the order-5 models. It is likely that other experimental paradigms may need different bbGCM orders, depending on the time scale of the semantic units involved.

*Seed ROIs*

The ROIs that were used as seeds in the between brains Granger causality analysis were defined as those 'mirror' areas that were active both during gesturing and guessing using a traditional GLM analysis on the same data (Schippers et al., 2009). The GLM for gesturing was estimated using the entire period in which the gesture was executed as the only predictor. The GLM for the guess runs included two predictors: 1) the period from onset of the movie in which the gesturer was shown until the time of button press and 2) from button press until the participant had given a response. All predictors were convolved with canonical hemodynamic response functions. The mean parameter estimates of the contrasts gesturing versus baseline and guessing versus baseline were tested at the second level using a one-sample t-test. Both results of the second-level random effects analysis of gesturing and guessing versus baseline were thresholded at $p<0.0001$. We used this stringent criterion to reduce the size of our seeds. The resulting maps were binarized (i.e. contained value 1 at above-threshold vertices and 0 at below-threshold vertices). The binary maps were multiplied and the resulting clusters contained the middle temporal gyrus (LH: 69, RH: 256 vertices), the ventral (LH: 42, RH: 123 vertices) and dorsal premotor cortex (LH: 69, RH 147 vertices) and a larger cluster in the parietal lobe. To reduce the size of our parietal lobe seed, we used only that part of the ROI that overlapped with area BA2 (LH: 529, RH: 552 vertices), because BA2 was found to be the most consistent location of mirror voxels in a previous analysis of the execution and observation of goal directed behavior (Gazzola and Keysers, 2008). The location of BA2 was defined by projecting the maximum probability map of BA2 from the anatomy toolbox (Eickhoff et al., 2005) onto a cortical surface segmentation of the Colin brain (Holmes et al., 1998). These particular ROIs were chosen because our primary aim was to examine the resonance theory of mind-reading which suggests that the pMNS of to people resonate with each other, with the pMNS in the gesturer involved in triggering the gestures, which are then viewed by the guesser, triggering activity in the pMNS of the guesser.

*Time series*

The input of the Granger causality analysis consisted of the average time course of the seed of the gesturer during gesturing as well as all the corresponding time courses of the vertices of the guesser's brain during guessing. The time courses were truncated to contain only those parts that reflected the steady-state part of either the gesturing or the guessing. We excluded 1) 15TR from the beginning and 5TR from the end of both the gesturing and the observation, thereby removing the on- and offset transients (Fig. 4.1, green blocks), 2) the period from the button press of the guesser until the

onset of the next gesture (Fig. 4.1, red and grey blocks), and 3) the baseline fixation cross period between two trials (Fig. 4.1, blue blocks). Additionally, information about the beginning and end of each separate gesture-part was taken into account in the Granger causality analysis, such that autoregressive model estimation was pooled over calculations on separate blocks rather than calculated over a single time-course with all blocks concatenated. On average, participants watched the gesture movies for 58 seconds, which corresponds to ±43 TR. This means that on average 43 - 15 - 5 = 23TR per trial were included in the Granger analyses, cumulating to 23 x 14 = 322TR per participant.

*Second level Granger analysis*

Random effects' testing was performed by t-tests at the second level. Two-tailed one-sample t-tests were computed for a single differential G-causality maps, with a null-hypothesis of a zero value (i.e. there is as much G-causality from Gesturer → Guesser than from Guesser → Gesturer). Two-tailed paired t-tests were computed when differential G-causality maps were compared between conditions/situations, with a null-hypothesis of a zero value for the difference of the maps. Random effects t-maps were then statistically thresholded at $p<0.05$ and corrected for multiple comparisons by using a Monte Carlo simulation-based cluster-size threshold adjustment (Hagler et al., 2006).

*Motion Energy GLM*

We extracted motion energy from the gesture movies using Matlab. For two consecutive frames of the recorded movies, motion energy was quantified in every pixel as the sum of the squared differences in the red, green and blue channels and then summed over all pixels. This time course was then mean corrected, convolved with the hemodynamic response function and sampled at the acquisition rate of the fMRI signal (TR = 1.33s). The GLM included three predictors: 1) a boxcar having values 1 for the period from onset of the movie in which the gesturer was shown until the time of button press and zero elsewhere 2) a boxcar having values 1 from the time of button press until the end of the response procedure and zero elsewhere, 3) the mean corrected, convolved and sub sampled motion energy time course. Predictors 1 and 2 were then also convolved with the canonical hemodynamic response functions. The mean parameter estimates of predictors 1 and 3 were then averaged for each participant and tested at the second level using a one-sample t-test (see Fig. 4.9). Predictor 1 but not 3 showed above threshold clusters.

## SUPPLEMENTARY INFORMATION 2

### GRANGER CAUSALITY AND ITS APPLICATION TO FMRI

Granger causality or G-causality was proposed by Clive Granger (Granger, 1969, 1980) and partially based upon earlier ideas of Norbert Wiener (Wiener, 1956). The aim was to give an operational definition of what causality or influence could mean for observations, structured in time, of multiple variables of interest. Granger clearly did not mean to equate such a definition to an interventional notion of causality that demands that one can only establish true causality if a process is actively interfered with and the consequences, observed (c.f. Pearl, 2000). Rather, G-causality quantifies a very useful, pragmatic statistical notion of information transfer (see below) between two stochastic processes.

A stochastic process X(t) is a random variable that is observed repeatedly in time in an orderly fashion. The general idea of G-causality is that Y(t) G-causes another process X(t) if the prediction of X's values improves when we use past values of Y, given that all other relevant information is taken into account. Here, all other relevant information is understood to contain at least the past values of X itself. The vehicle most often used to make predictions in this context is the linear autoregressive (AR) model. Two regressions are compared: one including only the past of X itself, and one additionally considering the past of Y:

$$X_t = \sum_{i=1}^{3} \alpha_i X_{t-i} + \epsilon_t \tag{4.1}$$

$$X_t = \sum_{i=1}^{3} \alpha'_i X_{t-i} + \sum_{i=1}^{3} \beta_i Y_{t-i} + \epsilon_t \tag{4.2}$$

The residual error (variance) reduction from $\sigma^2(\epsilon)$ to $\sigma^2(\epsilon')$ quantifies how much Y G-causes X. In the present study, the index i runs from 1 to 3 (the so-called autoregressive order) which determines how far into the past we look for useful information to predict X's values. The autoregressive order needs to be carefully chosen (see Methods and Fig. 4.10). When G-causality is applied to fMRI, the rich spatial structure in the fMRI signal can be used. fMRI gives 10's of thousands of independent time-courses (for each volume element or voxel) of activity throughout the brain. The technique of fMRI Granger Causality Mapping (GCM) explores all regions in the brain that interact with a single selected reference region using G-causality as a measure of directed

influence or information flow. By employing an autoregressive model containing the reference region and, in turn, every other voxel in the brain, the sources and targets of influence for the reference region can be mapped (Roebroeck et al., 2005).

Here, Granger causality mapping was applied to data from different brains in between brain Granger Causality Mapping (bbGCM). Given a seed selected in one brain (the gesturer), maps are created that specify G-causal influence from the seed in the gesturer to all of the guesser's brain, as well as influence in the reverse direction, i.e. from anywhere in the guesser's brain to the seed in the gesturer's brain. In this way we quantify information transfer (in the Granger causal or information theoretic sense, see below) between two brains during gestural communication (see Figure 4.2). Simulations have shown that if two populations of neurons influence each other strictly in a X → Y direction, if BOLD is used to make inferences about the neural populations, the low-pass filtering effect of the hemodynamic response function and the low sampling rate of fMRI lead to a certain cross-talk between the X → Y and Y → X G-causality, such that the Y → X direction in the BOLD is no longer zero. The same simulations have however also shown that the difference between X → Y and Y → X, called the 'differential G-causality', remains positive (Roebroeck et al., 2005). When applying G-causality to BOLD signal, it is therefore advisable to interpret the differential G-causality, and limit inferences to stating that X influences Y more than the other way around or vice versa.

In addition to detecting potential directed causal influences, G-causality has been shown to be a way to quantify information transfer between two variables. In information theory, a branch of mathematical statistics, Mutual information is used as a mathematical quantity that measures how much one random variable X tells us about another Y (Cover and Thomas, 1991; Shannon and Weaver, 1949). Specifically, mutual information measures to what degree the uncertainty about the value of one random variable X is reduced if we know the value of another random variable Y. Here uncertainty is mathematically quantified as entropy: the higher the entropy of X, the more uncertain we are about X's value. Schreiber (2000) used these information-theoretic concepts to define the notion of transfer entropy as a measure of directed (time-asymmetric) information transfer between stochastic processes. Transfer entropy from Y(t) to X(t) is the degree to which Y(t) disambiguates (i.e. reduces the uncertainty about) the future of X(t) beyond the degree to which X(t) already disambiguates its own future. Interestingly, one can show that, if X(t) and Y(t) values have a Gaussian distribution at each time point, transfer entropy is mathematically equivalent to G-causality quantified within an autoregressive model (Barnett et al., 2009).

In the context of social neuroscience, a key question is to study what brain regions are involved in information transferred from one brain to another. Given that G-causality quantifies time-directed information transfer, bbGCM should be seen as a tool to map which regions of one brain show evidence, in the measured signal

and at the time scale used, of receiving information from (or sending information to) the seed selected in the other brain. Finding significant differential G-causality between regions in two brains then suggest that a causal chain of events (in the stricter interventional sense see Pearl, 2000) may connect these two brain regions. This causal chain of event is necessarily indirect (i.e. mediated through the bodies of the agents and the world, here the gesture and their observation). Such G-causality is then an invitation to:

- Test whether these regions showing G-causality are indeed linked by a causal chain in the interventional, strong sense of causality tout-court, for instance by using TMS to manipulate the region of one brain and measure the effect on the other.
- Manipulate the visual/auditory channels connecting the two individuals to identify how one brain influences the other.

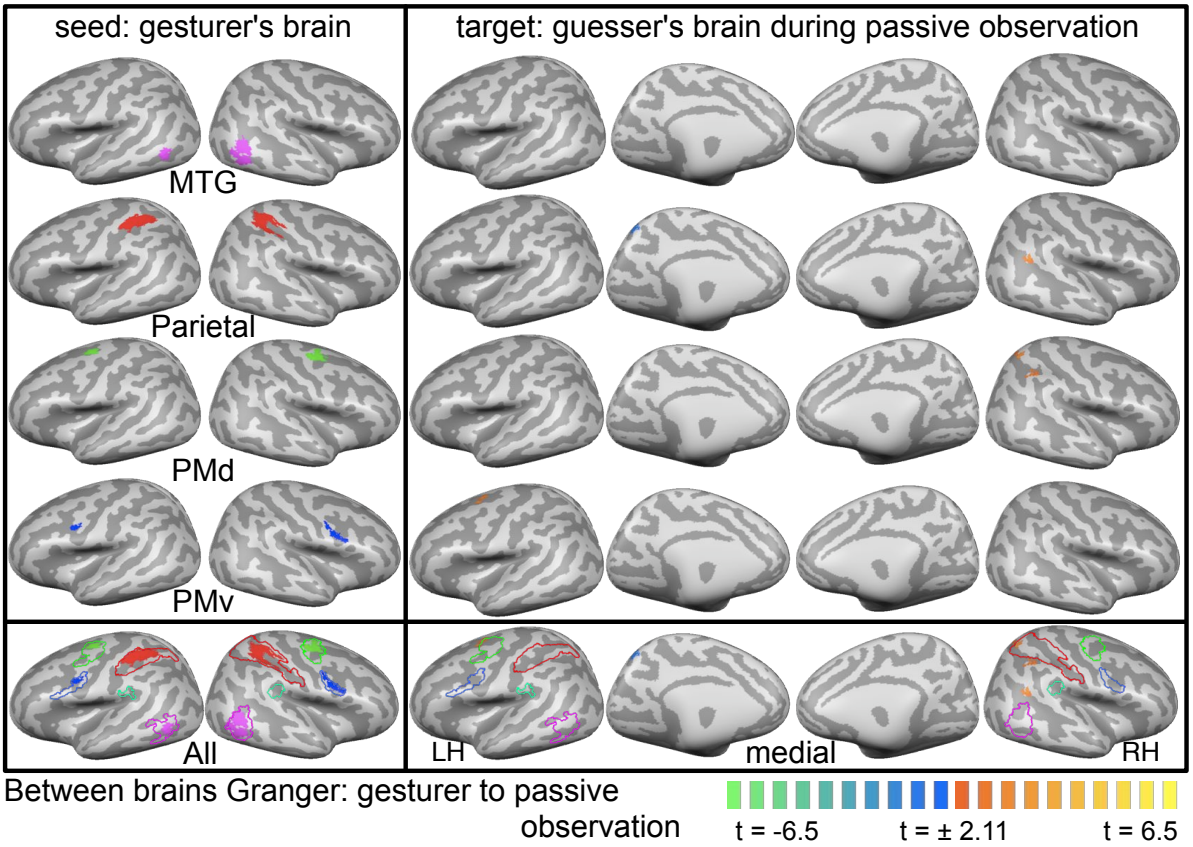



Figure 4.4: Results of second-level bbGCM for gesturer to passive observation (Random Effects, n=18). Conventions as in Fig. 4.2, but using brain activity during passive observation instead of guessing for the guesser.

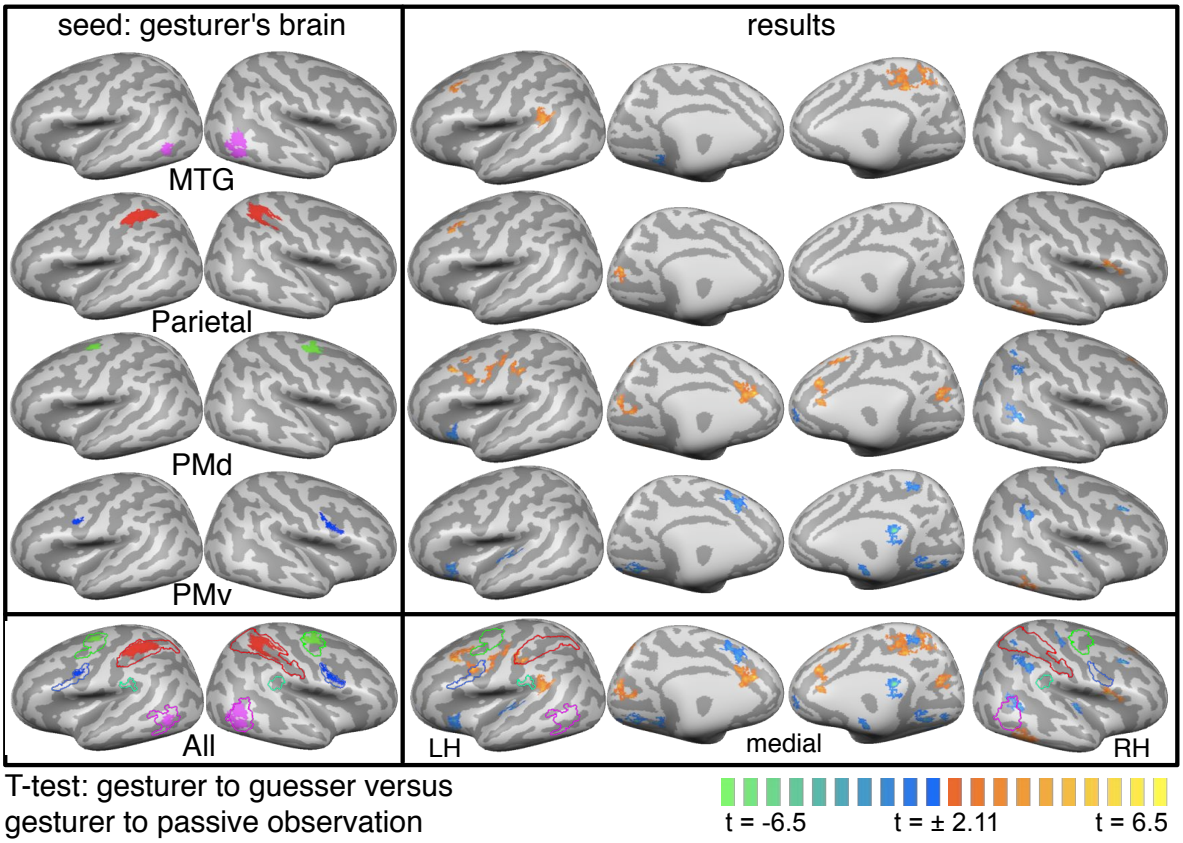



Figure 4.5: Comparison (paired t-test) between gesturer → guesser bbGCM and gesturer → passive observation bbGCM (Random Effects, n=18 participants). Warm colors indicate higher differential Granger values for the original analysis (active guessing). Further conventions as in Fig. 4.2.

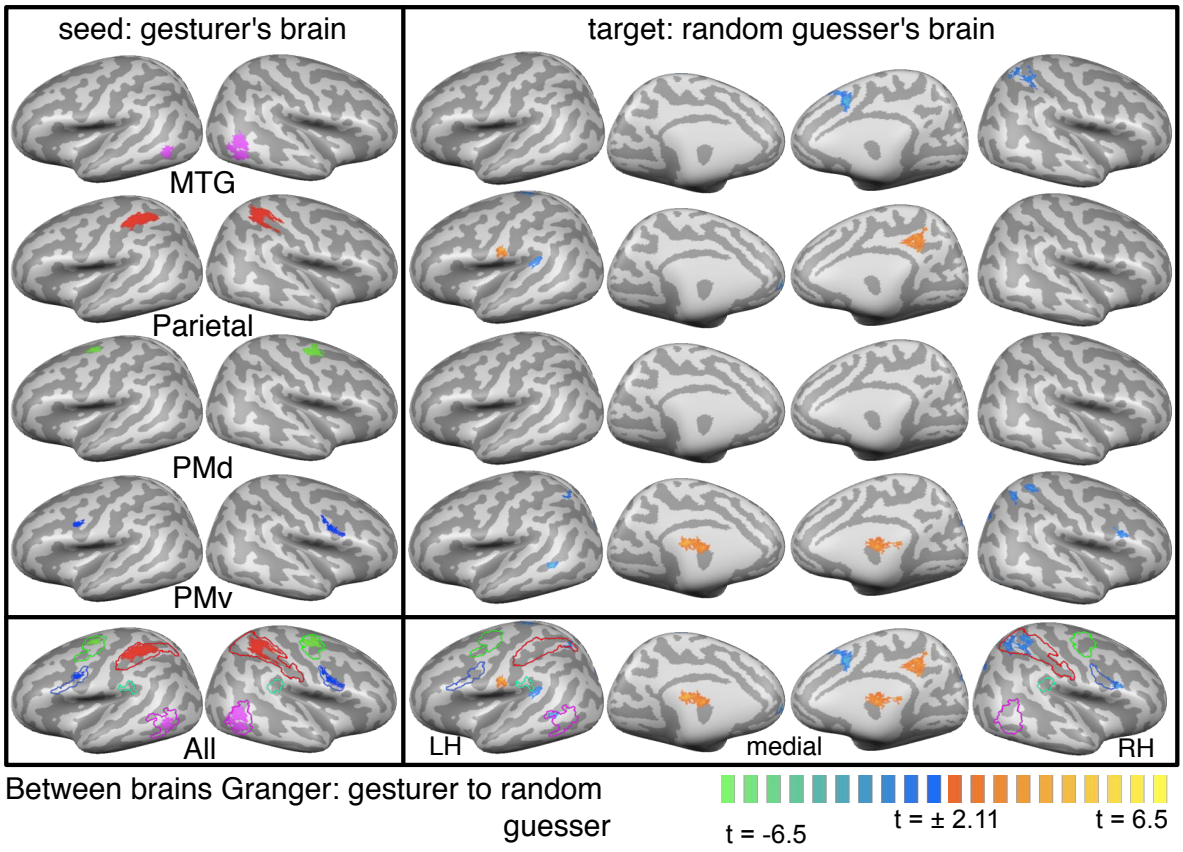



Figure 4.6: bbGCM using brain activity from a randomly selected guesser that had seen the gestures of a different gesturer (Random Effects, n=18). The right side therefore represents the guesser's brain showing t-values of the paired t-test between gesturer → random guesser versus random guesser → gesturer G-causality. Conventions as in Fig. 4.2.

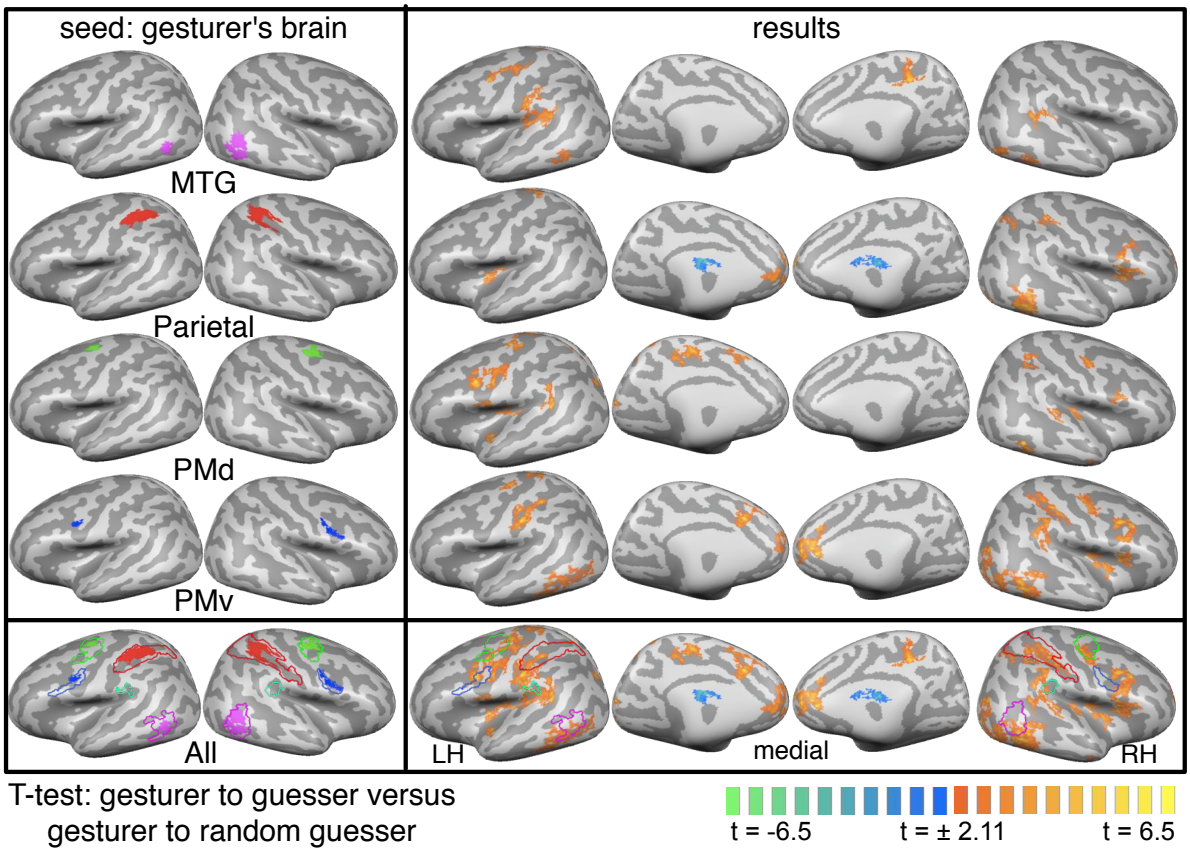



Figure 4.7: Comparison gesturer → guesser bbGCM vs gesturer → random guesser bbGCM (Random Effects, n=18 participants). Warm colors indicate higher differential Granger values for the gesturer → guesser analysis. Conventions as in Fig. 4.2.

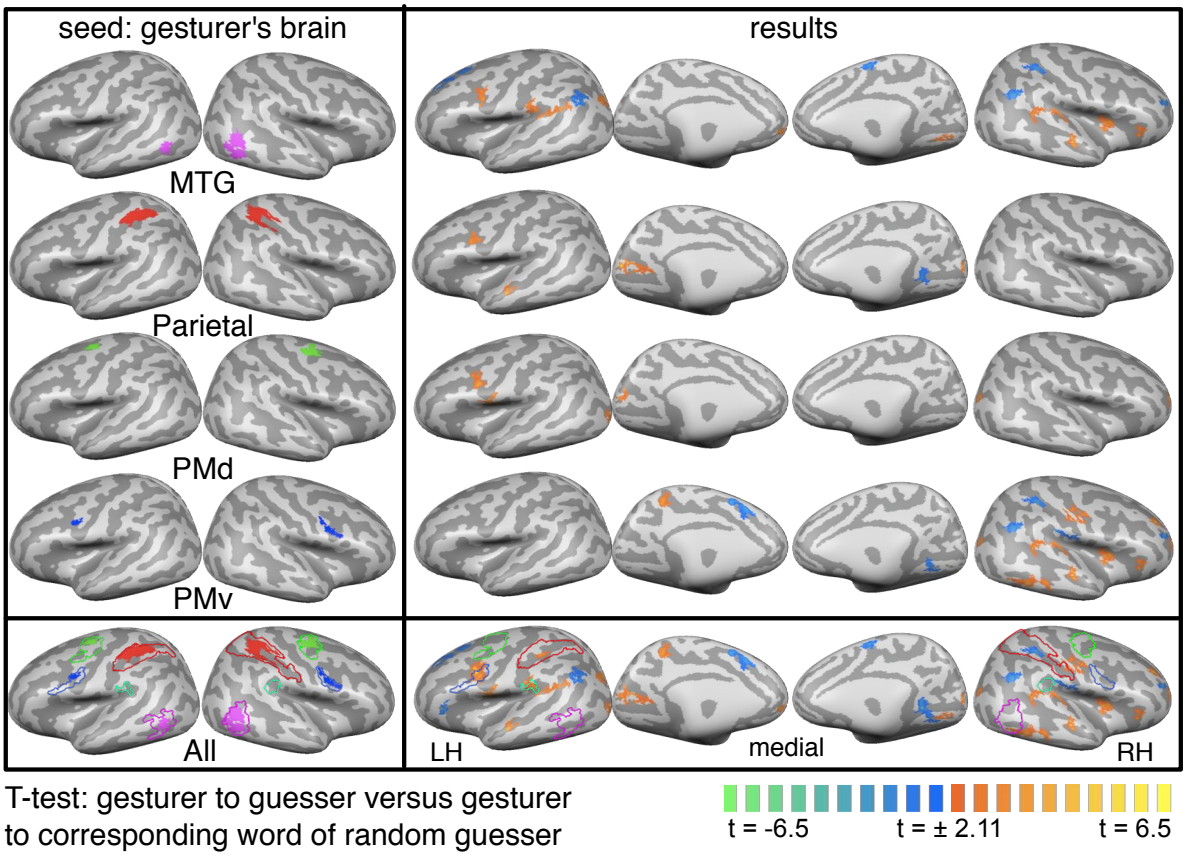



Figure 4.8: Comparison gesturer-to-guesser bbGCM vs one in which the activity of the gesturer was paired with a patchwork of episodes in which a different, randomly selected guesser (that might be different for each word) saw the same word as the original guesser being gestured by a different gesturer (Random Effects, n=18 participants). Conventions as in Fig. 4.2.r

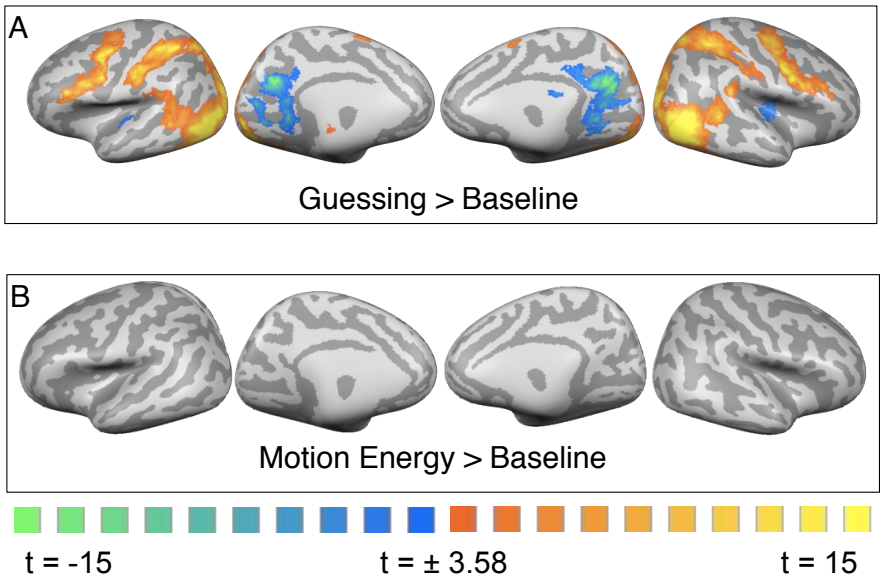



Figure 4.9: Activation maps of all 18 subjects for the GLM with motion energy and the guessing period as regressors. (a) Main effect guessing-baseline. (b) Main effect motion energy-baseline. All images are thresholded at t=3.58 which corresponds to an uncorrected p<0.001.

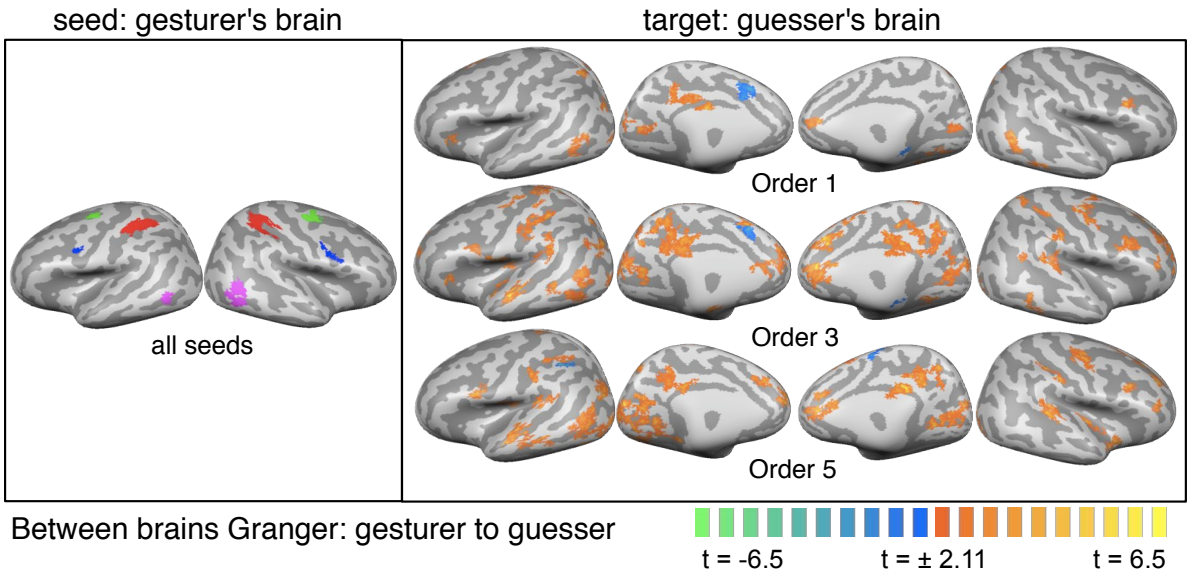



Figure 4.10: bbGCM at temporal windows 1,3 and 5 (Random Effects, n=18). Only the summary representation is shown for each order, all other conventions as in Fig. 4.2.

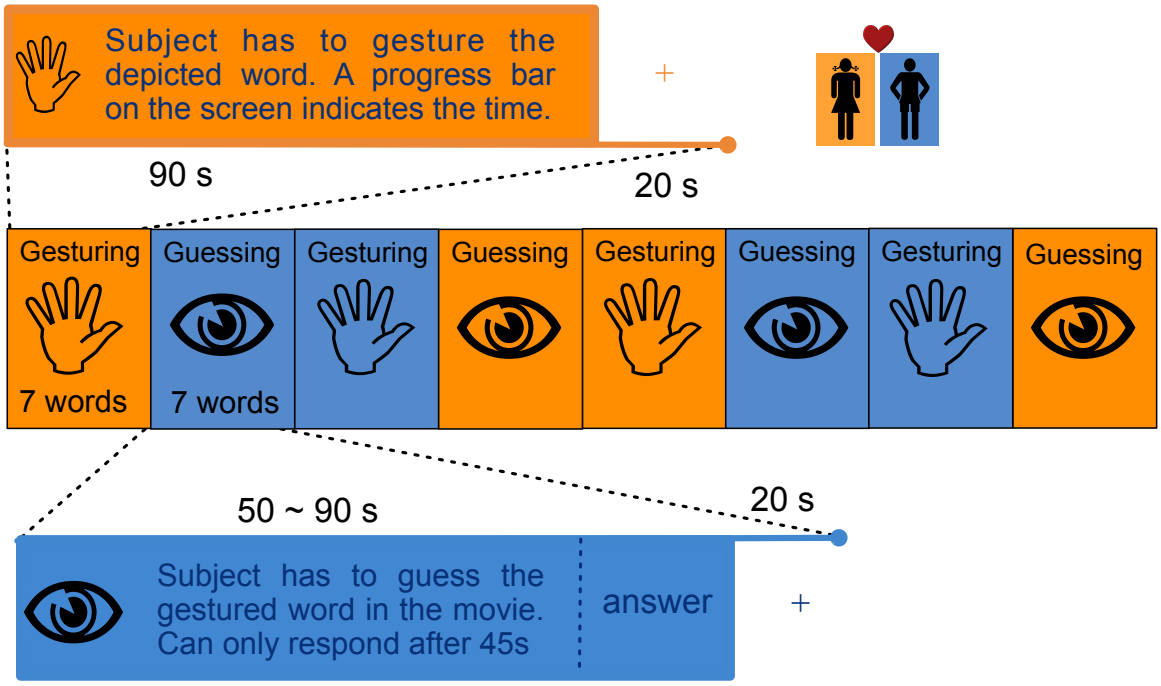



Figure 4.11: Example time line of the paradigm of the Charades experiment. Orange signals times in which the female is in the scanner, blue times in which the male is in the scanner. A hand represents gesturing runs, and eye, guessing runs. Each change of color signifies that one partner went out of the scanner and the other was placed into the scanner. Gesture runs were composed of 7 words, and each word had to be gestured for 90s, as signaled by a progress bar. Guess runs were composed of 7 movies of gestures, and the participant could only provide his/her answer once the progress bar had crossed the 50s mark. Twenty seconds of rest with a fixation cross always separated two words or movies.

# 5

# THE EFFECT OF INTRA- AND INTER-SUBJECT VARIABILITY OF HEMODYNAMIC RESPONSES ON GROUP LEVEL GRANGER CAUSALITY ANALYSES

*Submitted as:* Schippers, M.B., Renken, R., C. Keysers. The effect of intra- and inter-subject variability of hemodynamic responses on group level Granger Causality analyses.

## ABSTRACT

Granger causality analyses aim to reveal the direction of influence between brain areas by analyzing temporal precedence: if a signal change in area A consistently precedes a signal change in area B, then A Granger-causes B. fMRI-based Granger causality inferences are mediated by the hemodynamic response function which can vary across brain regions. This variability might induce a bias in Granger causality analyses. Here we use simulations to investigate the effect of hemodynamic response variability on Granger causality analyses at the level of a group of twenty participants. We used a set of hemodynamic responses measured by Handwerker et al. (2004) and simulated 200 experiments in which time series with known directions of influence are convolved with these hemodynamic responses and submitted to Granger causality analysis. Results show that the average chance to find a significant Granger causality effect when no actual influence is present in the data stays well below the p-level imposed on the second level statistics. Most importantly, when the analyses reveal a significant directed influence, this direction was accurate in the vast majority of the cases. The sensitivity of the analyses however depended on the neuronal delay between the source and target regions and their relative hemodynamic delay. Influences flowing from regions to one with the same or a slower hemodynamic response function were detected in over 80% of the cases when the neuronal delay was at least 100ms. Influences flowing to a region with a faster hemodynamic delay were detected in over 80% of the cases when delays are above 1s.

## 5.1 INTRODUCTION

Granger causality is a measure of directed influence between two time series. Originally conceptualized by Wiener and formalized by Granger (Wiener, 1956; Granger, 1969), it was introduced as a connectivity analysis for fMRI data in 2003 by Goebel

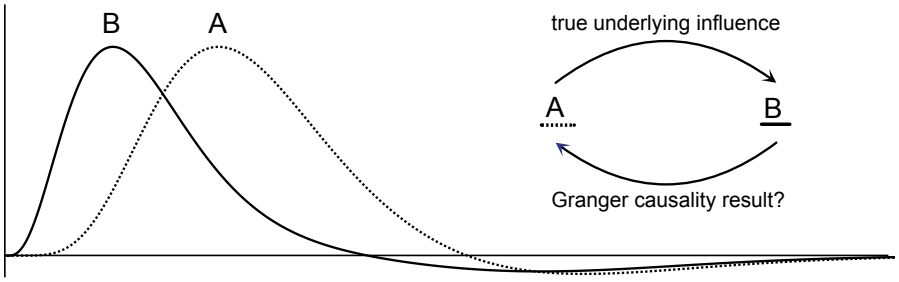



Figure 5.1: A faster hemodynamic response function could lead to an inverted Granger causality result

et al. (2003). Granger formalized causality between two time series using the concept of temporal precedence: if a signal change in A is consistently followed by a signal change in B, A Granger-causes B.

When applied to fMRI, Granger causality indicates the direction of influence between BOLD time series of different brain areas. Results of Granger causality analyses are interpreted as indicating connectivity on a neuronal level. However, fMRI measures BOLD responses rather than neuronal activity directly, therefore this inference is mediated by the hemodynamic response in the brain. The hemodynamic response is not equal across brain regions (Rajapakse et al., 1998; Aguirre et al., 1998; Kruggel and von Cramon, 1999; Handwerker et al., 2004) and this regional variability could cause problems for Granger causality analyses (David et al., 2008; de Marco et al., 2009; Roebroeck et al., 2005; Friston, 2009; Chang et al., 2008). On the one hand, it is feared that spurious Granger causality findings could be reported as a difference in hemodynamic response might introduce temporal relations where there are none. On the other hand a difference in hemodynamic response might invert the reported direction of Granger causality. The intuitive idea behind this last problem is as follows: If region A causes activity changes in region B and region B has a faster hemodynamic response than region A, a Granger causality analysis might indicate a net influence going from B to A rather than the true underlying causality from A to B (see Figure 5.1). Roebroeck et al. (2005) warned against this possible confound and suggested using the modulation of connectivity between different conditions, rather than within one condition.

Deshpande et al. (2009) investigated the effect of differences in hemodynamic response function on the sensitivity of Granger causality analyses in *single subjects*. They found that even when intra-subject differences in hemodynamic response function are present, Granger causality is still sensitive to influences in the order of a hundred milliseconds. This result seems counterintuitive with differences in hemodynamic delay as big as 2.5s (Handwerker et al., 2004). However, as Deshpande et al. (2009)

note, differences between hemodynamic responses are not just due to a temporal shift of the whole response. Rather, parameters that varied most between regions were onset time and time-to-peak (Handwerker et al., 2004). This means that convolution of time series of neuronal activity with a hemodynamic response function is not mathematically equivalent to a shift in time. The difference in hemodynamic response does not remove the characteristic temporal relation between the two time series, which could be an explanation why the effect of regional variability of the hemodynamic response was not as disastrous as initially expected.

The study of Deshpande et al. (2009) shows how sensitivy of Granger causality is affected by variability in hemodynamic response at the level of the single subject. However, Granger causality at group level is of more interest than at single subject level for most studies. Furthermore, as within subject variability varies across subjects (Aguirre et al., 1998) results from Deshpande et al. (2009) cannot be extrapolated trivially to group results. In this article, we investigate whether differences in hemodynamic response have an effect on *group level* Granger causality results. We use simulations to answer the following three questions. (1) When no actual directed influence is present between two time series, what is the chance to find a significant Granger causality result (i.e. a false positive)? (2) When an actual directed influence is present between two time series, how often will differences in hemodynamic response lead to a significant inverted Granger causality direction and (3) how often will the true direction be detected (sensitivity)?

Hemodynamic response shapes as measured by Handwerker et al. (2004) are used in the current simulations. These hemodynamic response shapes were measured subject by subject for 20 individuals, permitting us to simulate data for an appropriate sized 'group study' that will take realistic intersubject variability into account. Unfortunately, due to the fact that participants were performing a task while their hemodynamic responses were measured, the BOLD latencies conflate neuronal latencies and hemodynamic latencies. This inflates the actual variability of the hemodynamic responses. Ideally, one would use data where local field potentials were measured simultaneously to disentangle these sources of variance. Awaiting such a dataset, the current work will most likely represent a scenario worse than reality.

## 5.2 METHODS

We simulated 200 experiments, each time generating a pair of time series of neuronal activity for sixteen connections for twenty subjects. The time series were generated with a known directed influence which ranged from no influence to a strong influence. Each of these pairs of time series were subsequently convolved with a combination of hemodynamic responses. We used a set of hemodynamic response functions

from four different areas measured by Handwerker et al. (2004). In other words, we simulated a time series of neuronal activity for region A and a time series of neuronal activity for region B which contained a directed influence from region A to B (see next paragraph for more details). We then convolved the first time series with one of the four hemodynamic responses and the second time series with either the same or another hemodynamic response of the same subject. Thereafter, we repeat the same procedure for each and every of the 20 subjects based on the subject-unique set of hemodynamic responses empirically measured by Handwerker et al. (2004). This means that we simulate variability in hemodynamic responses between brain regions and between participants in our data set. We repeated this for every possible combination of hemodynamic responses of that subject. Note that we do not treat the hemodynamic response functions as belonging to a particular brain area, but as a representative set of hemodynamic responses. In this way, we could simulate connections between areas having the same hemodynamic response, but do not interpret this as a connection between an area and itself.

For each connection we then calculated the difference Granger causality result (subtracting the influence of region B on region A from the influence of region A on region B) and assessed this difference statistically with a t-test across subjects. We then calculated the proportion of correct and inverted Granger causality results. A correct result here indicates a significant net influence from region A to region B, while an inverted result indicates a significant net influence from region B to region A. Supplementary Information 5 gives a schematic overview of the simulations.

### 5.2.1 *Generation of fMRI time series*

One of the constraints of Granger causality analyses is that time series should be wide-sense stationary. Hence, in our simulation, only that part of an fMRI time series is generated which does not include on and offset effects. This could represent for example resting-state data, but also the steady-state part of a long stimulus (see e.g. Schippers et al., 2010).

For each simulation, two time series ($x_i$ and $y_i$) are modeled as in Roebroeck et al. (2005). Neuronal interactions (local field potentials) are simulated using a bi-dimensional first-order vector autoregressive process (see Equation 5.1). The autocorrelation of $x_i$ and $y_i$ is set to 0.9, and the influence of $x_i$ onto $y_i$ varies systematically between 0, 0.1, 0.2, 0.5, and 0.9 (from here on indicated as Granger causality strength). Both signals are simulated for 15000 time steps of 10 ms (150 sec). Target time series $y_i$ is then shifted to represent the neuronal transmission delay. We systematically varied this neuronal delay with the following values: 10, 20, 50, 100, 200, 500, 1000, and 2000 ms. Both signals are convolved using subject and area specific hemody-

namic response models and gaussian noise is added to represent physiological noise in the BOLD response. Subsequently, the signal is down sampled to 1 s (resembling an acquisition rate (TR) of an MR-scanner) and gaussian noise is again added to represent acquisition noise. After each step the signals are normalized to zero mean and unit variance. The total amount of noise added was either 20% or 100%. After noise was added at two different stages of the simulation, the standard deviation of the signal had increased to 120% and 200% of the original standard deviation for the respective noise levels. This means the SNR is 5:1 in the 20% noise case and 1:1 in the 100% noise case.

$$\begin{bmatrix} x_i \\ y_i \end{bmatrix} = A \begin{bmatrix} x_{i-1} \\ y_{i-1} \end{bmatrix} + \begin{bmatrix} \varepsilon_{x_i} \\ \varepsilon_{y_i} \end{bmatrix}$$
$$A = \begin{bmatrix} 0.9 & 0 \\ 0.5 & 0.9 \end{bmatrix}, \Sigma = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \tag{5.1}$$

### 5.2.2 *Generation of hemodynamic response function*

Handwerker et al. (2004) obtained models of hemodynamic response functions for primary visual cortex, supplementary eye fields, primary motor cortex and frontal eye fields (in the current article referred to as respectively areas 1 to 4) for twenty participants. Daniel Handwerker kindly provided us with the data. In their study, participants fixated on a cross and responded with a button press to a checkerboard appearing for 200ms on their left or right hemifield. The hemodynamic response shape was computed by averaging over all trials within an anatomically defined region of interest. Subsequently, measured hemodynamic responses were fitted with the sum of two gamma functions. See Supplementary Information 1 and Supplementary Information 2 for details.

### 5.2.3 *Granger Causality*

A time series $x_i$ is said to Granger-cause another time series $y_i$ if including information about the past of $x$ significantly increases the prediction of the current value of $y$ with regard to this prediction based on the past values of $y$ alone (Granger, 1969). Granger Causality was implemented here according to Geweke (1982) using vector autoregressive models. Following this approach, four autoregressive equations are

calculated (see Equations 5.2-5.5).

$$x_i = \sum_{j=1}^{L} a_j x_{i-j} + \epsilon_{1i} \tag{5.2}$$

$$y_i = \sum_{j=1}^{L} a_j y_{i-j} + \eta_{1i} \tag{5.3}$$

Equations 5.2 and 5.3 calculate how much two time series, $x_i$ and $y_i$, can be explained by their own past ($x_{i-j}$ and $y_{i-j}$), resulting in residual error variances $\Sigma_1$ and $\Gamma_1$. The order is represented by $L$ and specifies how many previous time points are taken into account.

$$x_i = \sum_{j=1}^{L} a_j x_{i-j} + \sum_{j=1}^{p} b_j y_{i-j} + \epsilon_{2i} \tag{5.4}$$

$$y_i = \sum_{j=1}^{L} a_j y_{i-j} + \sum_{j=1}^{p} b_j x_{i-j} + \eta_{2i} \tag{5.5}$$

In equations 5.4 and 5.5 the prediction is based on the time series' own past and the past of the other time series. This results in residual error variances $\Sigma_2$ and $\Gamma_2$. The linear influence from $x_i$ to $y_i$ ($F_{x\to y}$, eq. 5.6) and from $y_i$ to $x_i$ ($F_{y\to x}$, eq. 5.7) can now be calculated as the ratio between the variances of the residual error (F-values). A reduction in error variance when including the past of another time series results in a larger F-ratio.

$$F_{x\to y} = \ln\left(\frac{var(\eta_{1i})}{var(\eta_{2i})}\right) \tag{5.6}$$

$$F_{y\to x} = \ln\left(\frac{var(\epsilon_{1i})}{var(\epsilon_{2i})}\right) \tag{5.7}$$

Difference Granger causality ($F_{x\to y} - F_{y\to x}$) was calculated to assess the dominant direction of information flow.

### 5.2.4 *Statistical analysis*

The result of one experiment was obtained by performing a simulation for a single combination of hemodynamic responses for every subject (from here on called a connection). A t-test was then performed to test the null-hypothesis ($F_{x\to y} = F_{y\to x}$) at group level. The procedure is then repeated 200 times (representing 200 experiments),

| | target | | | |
|---|---|---|---|---|
| source | 1 | 2 | 3 | 4 |
| 1 | 0 | 0.34 | 0.57 | 0.68 |
| 2 | -0.34 | 0 | 0.23 | 0.34 |
| 3 | -0.57 | -0.23 | 0 | 0.11 |
| 4 | -0.68 | -0.34 | -0.11 | 0 |

Table 5.1: Overview of difference in the delay parameter ($\delta_1$, see also Supplementary Information 2) between source and target in seconds. The three groups are indicated with dark gray (negative delay group), white (no delay group), and light gray (positive delay group).

leading to a distribution of 200 t-values. We then counted how often t < -T(19, 0.05) and t > T(19, 0.05) and used these proportions of inverted and correct results as dependent variables in our statistical assessments.

The modeled connections can be divided into three groups: (1) *no delay group* which contains connections in which source and target time series are convolved with the same hemodynamic response function, (2) *positive delay group* in which the source time series is convolved with a hemodynamic response function which has a smaller delay parameter than the hemodynamic response function used for the target time series, and, (3) *negative delay group* which contains connections in which the source time series is convolved with a hemodynamic response function which has a larger delay parameter than the hemodynamic response of the target time series. Table 5.1 shows the average differences across subjects in delay parameter between brain areas and the resulting division in three groups.

For each delay group, we separately analyzed how the proportion of correct and inverted results are influenced by noise, neuronal delay, and Granger causality strength (gc-strength). Because the effect of these factors is dependent on the simulations being nested within connections, we modeled the data with a multilevel variance model. We verified this approach by calculating the intraclass correlation coefficients of the proportion of inverted and correct results (Snijders and Bosker, 1999). These indicated that 31% of the variance of inverted results and 17% of the variance of correct results can be explained by the fact that the data has a nested structure. A multilevel model is in these cases an appropriate statistical approach. In all groups, the factors *noise*, *neuronal delay*, and *gc-strength* were modeled as fixed factors, while *connection* was modeled as a random factor.

We also calculated the intraclass correlation coefficient for the results of the simulations in which there was no Granger causality present. This indicated that only 1% of the variance could be explained by the nested structure of the data. In our opinion

this was too few to necessitate a multilevel model. We tested a regression model in which the proportion of false positives is explained by the factors noise and neuronal delay. We tested our models against an alpha level $p \leq 0.05$.

### 5.2.5 *Comparison crosscorrelations in simulation data and experimental data*

To analyze whether our simulations represent a physiologically plausible situation, we compared lagged cross correlations in our generated time series with cross correlations in time series of experimental data. We assessed *lagged* cross correlations because in Granger causality the immediate past ($t-1$ at order 1) of the time series is used. We calculated these crosscorrelations with lag 1 between our simulated pairs of time series for each Granger causality strength (after convolution and downsampling with no neuronal delay added). Next, we calculated crosscorrelations between regions of the putative Mirror Neuron System (consisting of dorsal and ventral premotor cortex, inferior parietal lobule and middle temporal gyrus) using data of a previous experiment (Schippers et al., 2009). We compared these crosscorrelations of 18 subjects in all connections from the experimental data and the simulation data with a two-sampled t-test.

### 5.2.6 *Likelihood of true direction given a detected direction*

It makes sense, from the point of view of reporting the results of our simulations, to examine the frequency with which a certain direction of influence is detected given a certain underlying direction of connectivity. However, neuroimagers will typically be confronted with the reverse question: given that they found a significant differential Granger causality suggesting information flow from region X→Y, how likely is X→Y to be the real direction of influence? In this empirical situation, the experimenter will not know what the underlying strength of connection is (our GC factor) but they can form an educated guess about the length of the neuronal delay between the brain regions involved (see our discussion). The experimenter can also estimate the difference in hemodynamic delay between the regions as described in our method section.

To provide neuroimagers with an estimate of these conditional probabilities (i.e. p(true direction X→Y|detected direction X→Y) and p(true direction Y→X|detected direction X→Y), we will call X the region with the shorter and Y that with the longer hemodynamic delay. We concentrated on the simulations with a 20% noise level that seem closer to the empirical situation (see paragraph 5.3.3) and pooled the results of our simulations at different GC levels. We then identified all simulations in which a particular direction of influence was detected and counted how many of these cases

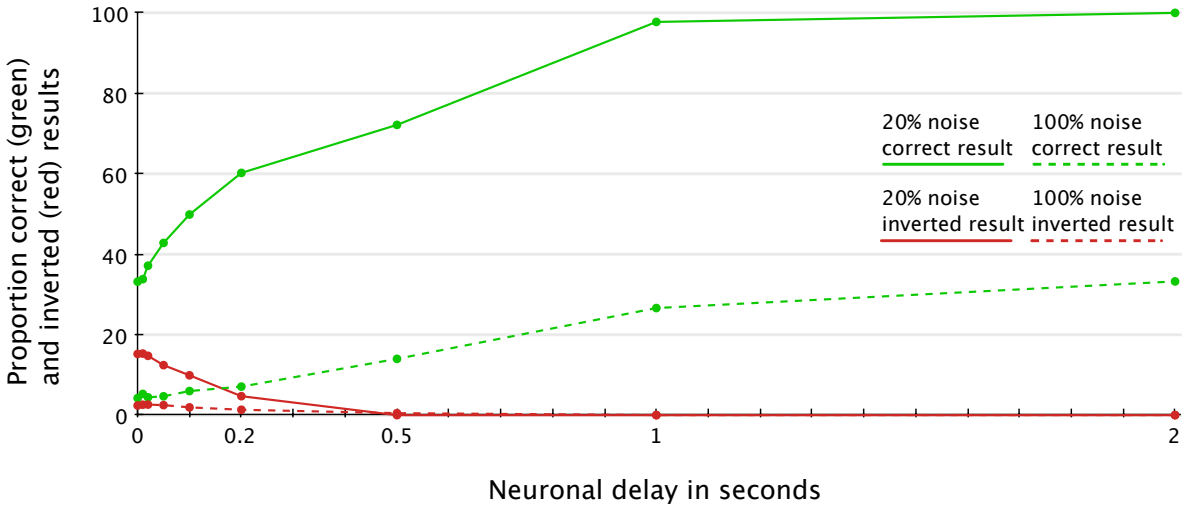



Figure 5.2: Mean proportion of correct and inverted results per neuronal delay.

had originated from simulations in which the underlying direction of influence was in the same or the opposite direction. We did this separately for the various neuronal delays and additionally summarized the situation for short (0.01-0.05s), medium (0.1-0.5s) and long (1-2s) neuronal delays. We limited our analysis to the connections between brain regions with different hemodynamic delays in our simulations to examine if the direction of the detected influence (from the faster region X to the slower region Y or vice-versa) influenced the results.

## 5.3 RESULTS

### 5.3.1 *When no Granger causality is present in the data*

Neither noise-level ($b = 1.5 * 10^{-5}$, $t(284) = 0.61$, $p = 0.54$) nor neuronal delay ($b = 1.9 * 10^{-3}$, $t(284) = 0.82$, $p = 0.4$) has a significant influence on the amount of false positives in the simulations where no Granger-causality was present. The result of a one-tailed t-test shows that on average the amount of false positives stays below 5% ($t(287) = -10.69$, $p < 2.2 * 10^{-16}$). The 95% confidence interval of the proportion of false positives ranged from 0.0397% to 0.0429%. Given that a t-test was thresheld at $p<0.05$, this shows that the distribution of Granger causality is indeed well-behaved in a statistical sense.

### 5.3.2 *When Granger causality is present in the data*

An overview of the results when a Granger causality influence is present is given in Figure 5.2 showing the percentage of correct and inverted results for all simulations per neuronal delay. The number of correct results increases with neuronal delay, while the amount of inverted results decreases. Noise decreases the amount of significant results found, both correct and inverted. Figures 5.3-5.6 show results per connection and per Granger causality strength for a noise level of 20%. See Supplementary Information 3 for the results at a noise level of 100%.

#### *Negative delay group*

In the negative delay group, the amount of inversions is significantly influenced by noise, neuronal delay and their interaction. The influence of neuronal delay on inverted results has the same direction for both noise levels (see Table 5.2b) with a stronger effect for the 20% noise level than for the 100% noise level. The longer the neuronal delay, the fewer inverted results are observed.

Correct results are significantly influenced by noise, neuronal delay and their interaction. At a 20% noise level (see Table 5.2a), neuronal delay has a significant influence, but gc-strength does not. At a 100% noise level, the interaction between neuronal delay and gc-strength is significant. The stronger the gc-strength, the stronger the influence of neuronal delay on the amount of correct results.

Further, it should be noted that in the negative delay group, Granger causality behaves well for neuronal delays $\geq$ 1s. Here inversions are absent and significant Granger causality is always detected with appropriate power (> 80% of cases).

The amount of inverted results would be expected to depend on the average difference in delay between the hemodynamic response functions of the two brain regions, with larger differences leading to more inversions, and on the variance of the timing across participants, with larger variance leading to relatively more false negatives (blue) and fewer inversions (red). The examination of Figure 5.6, in which the data in each column is arranged with the topmost cell having the most favorable and the bottom-most, the least favorable difference in average hemodynamic timing, confirms that inversions roughly increase from top to bottom, but that exceptions exist. For instance comparing the bottom two panels of Figure 5.6, corresponding to 3$\rightarrow$1 and 4$\rightarrow$1, one might be surprised to see that more inversions occur in the 3$\rightarrow$1 than 4$\rightarrow$1 configuration despite the fact that area 3$\rightarrow$1 should have the benefit of a less unfavorable difference in hemodynamic response function on average. We hypothesized that the larger variance in the parameter $\delta_1$ in area 4 compared to 3 (see Supplementary Information 2) would lead area 4 to have comparatively more false

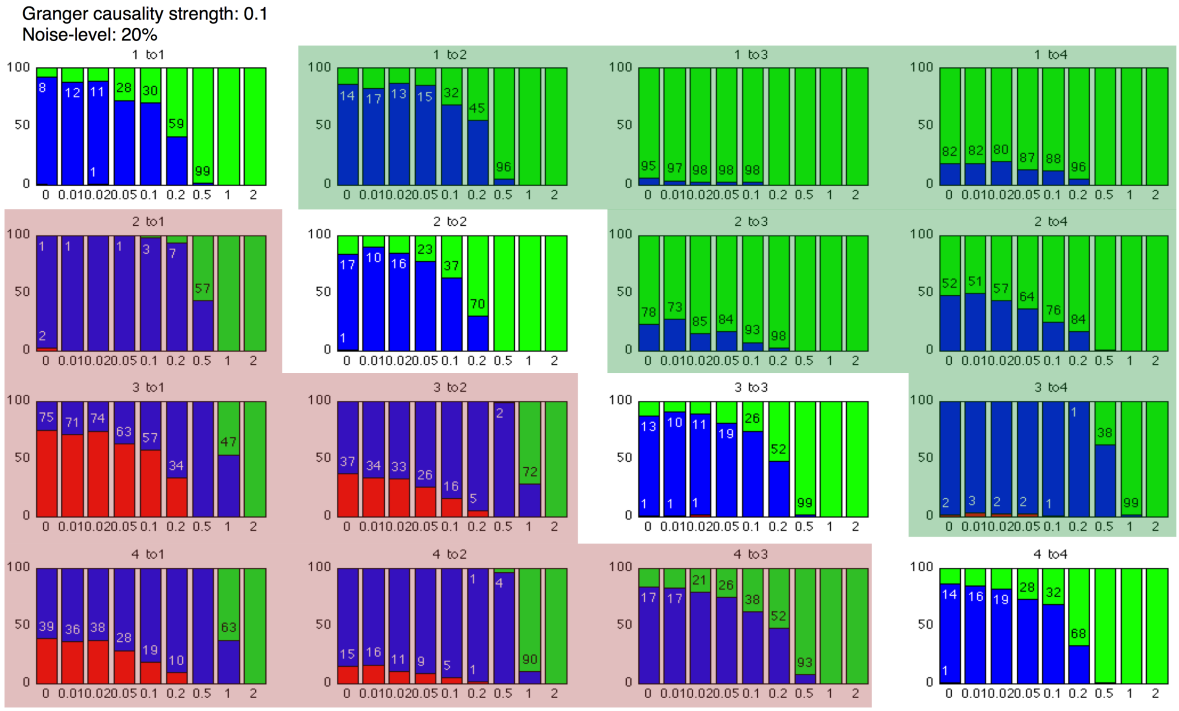



Figure 5.3: Proportion of correct (bright green), inverted (bright red) and non-significant (blue) results per connection and per neuronal delay for a noise-level of 20% and Granger causality strength of 0.1. Source (rows) and target (columns) are sorted on their mean onset parameter. Connections are divided in a positive delay group (green background), a no-delay group (white background) and a negative delay group (red background). A correct result for a given connection and neuronal delay signifies that the group Granger causality analysis indicated the direction of influence that was modeled in the time series (for example for connection '3 to 1', the correct direction of influence goes from area 3 to area 1). An inverted result indicates a significant Granger causality result with a direction opposite of the direction that was modeled in the simulation. A non-significant result indicates a result that did not exceed the threshold.

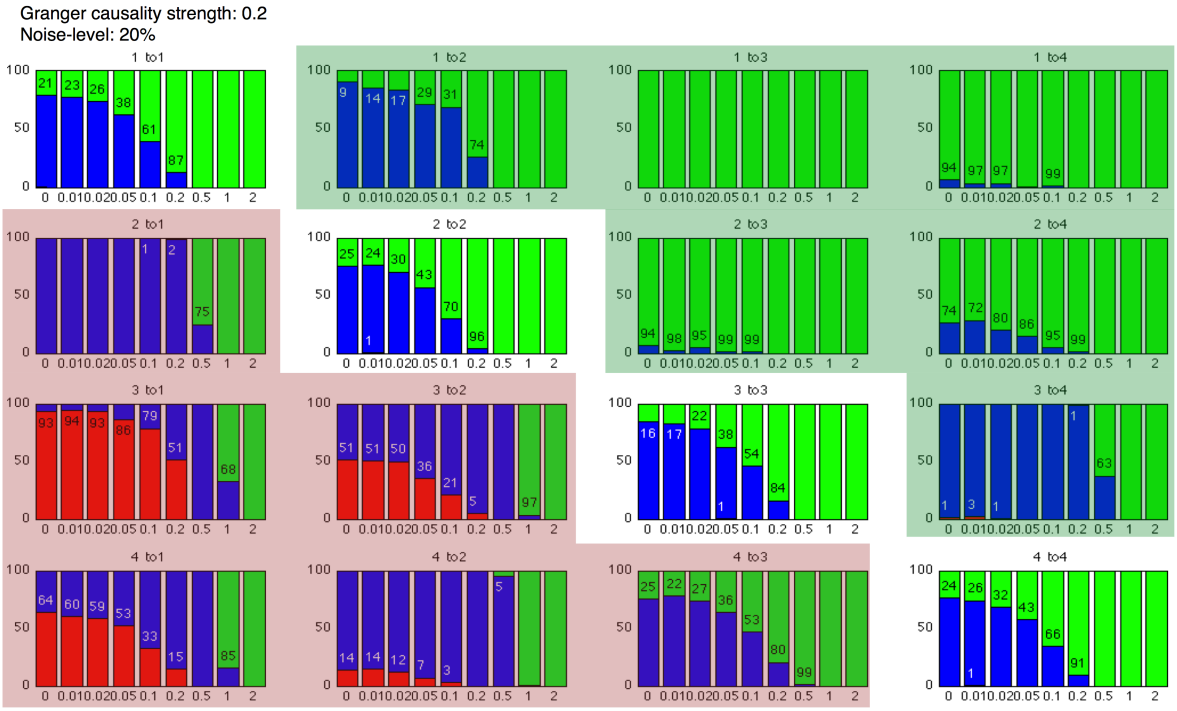



Figure 5.4: Proportion of correct (green), inverted (red) and non-significant (blue) results per connection and per neuronal delay for a noise-level of 20% and Granger causality strength of 0.2. See also caption of Figure 5.3

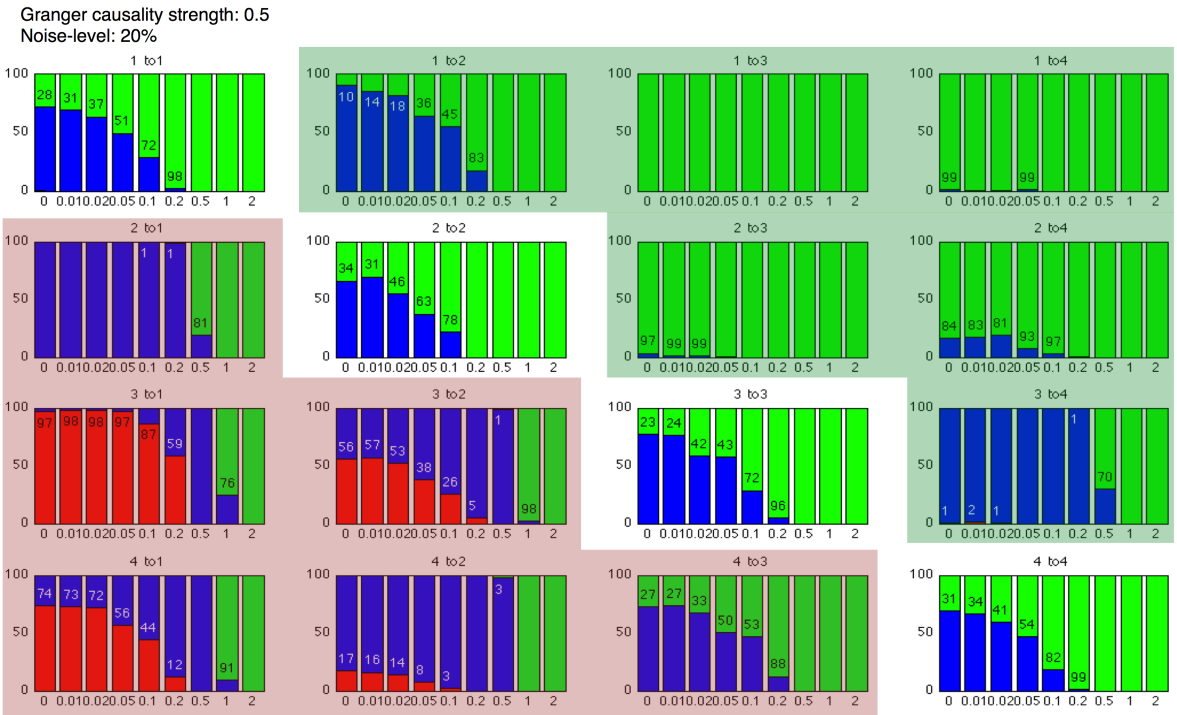



Figure 5.5: Proportion of correct (green), inverted (red) and non-significant (blue) results per connection and per neuronal delay for a noise-level of 20% and Granger causality strength of 0.5. See also caption of Figure 5.3

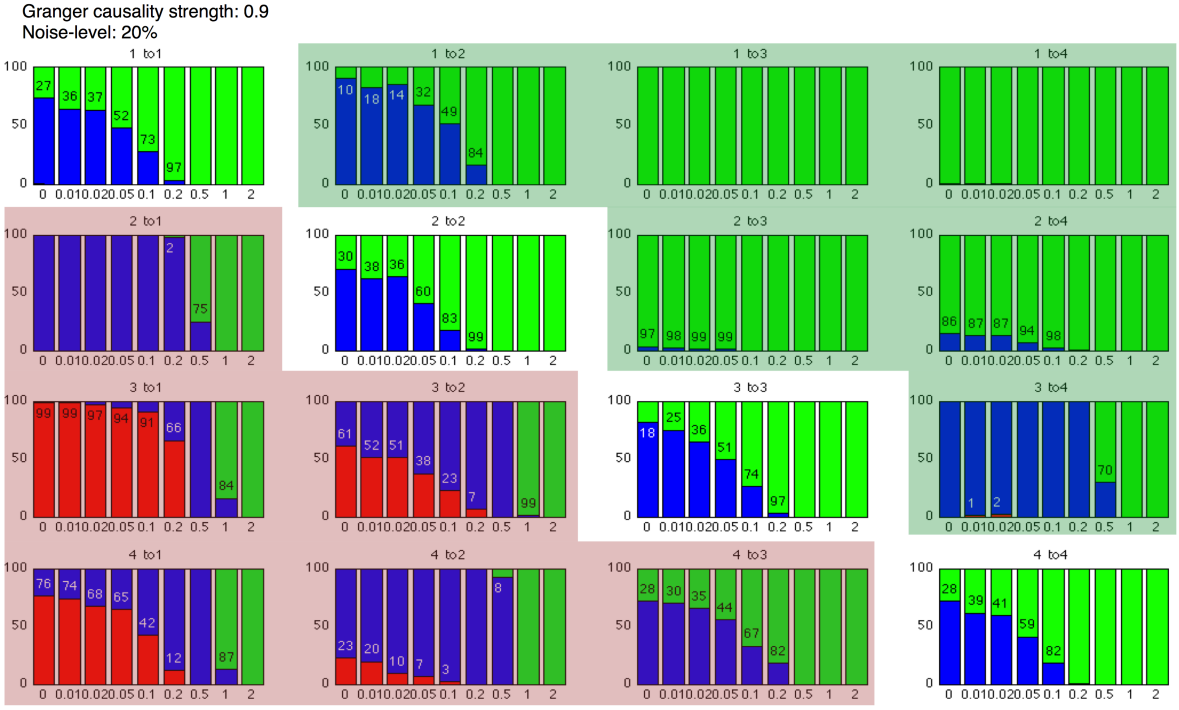



Figure 5.6: Proportion of correct (green), inverted (red) and non-significant (blue) results per connection and per neuronal delay for a noise-level of 20% and Granger causality strength of 0.9. See also caption of Figure 5.3

negatives (because the granger values would be more variable) and less inversions. Accordingly, if area 3 had the same variance in $\delta_1$ as area 4, the 3→1 case would then indeed lead to less inversions than the 4→1 case. To test this hypothesis, we generated artificially a region 3' that has the same variance in parameters as area 4 but the same average $\delta_1$ parameters (i.e. hemodynamic delay) as the original area 3. To do so, we used for area 3' all subject-specific parameters from area 4, but shifted all $\delta_1$ parameters by the difference in average $\delta_1$ between area 4 and 3. Doing so indeed showed that with the same variance in hemodynamic response, 3'→1 demonstrates less inversions than 4→1 (see Supplementary Information 4).

*Positive delay group*

The occurrence of inverted results in the positive delay group was too small (< 0.01%) to perform meaningful statistical analyses. In this group, only the proportion of correct results was assessed statistically.

The amount of correct results is significantly influenced by noise, neuronal delay and gc-strength, with a significant interaction between noise and neuronal delay. At

| **Negative delay group (proportion correct results)** | | | |
|---|---|---|---|
| **A** | 20% noise | | |
| | Estimate | Std. Error | P-value |
| (Intercept) | 0.038920 | 0.068146 | 0.939 |
| GC | 0.039594 | 0.047388 | 0.813 |
| delay | 0.537016 | 0.032521 | <1e-04 |
| GC:delay | 0.007621 | 0.061734 | 1.000 |
| **B** | 100% noise | | |
| | Estimate | Std. Error | P-value |
| GC=0.1 | | | |
| (Intercept) | 0.016212 | 0.004985 | 0.00228 |
| delay | 0.059474 | 0.003585 | <1e-10 |
| GC=0.2 | | | |
| (Intercept) | 0.011177 | 0.010972 | 0.517 |
| delay | 0.100576 | 0.005264 | <1e-10 |
| GC=0.5 | | | |
| (Intercept) | 0.010037 | 0.011411 | 0.608 |
| delay | 0.132431 | 0.006053 | <1e-10 |
| GC=0.9 | | | |
| (Intercept) | 0.007403 | 0.012295 | 0.792 |
| delay | 0.149710 | 0.005765 | <1e-10 |

(a) The effect of gc-strength and neuronal delay on the proportion of correct results in the negative delay group

| **Negative delay group (proportion inverted results)** | | | |
|---|---|---|---|
| | 20% noise | | |
| | Estimate | Std. Error | P-value |
| (Intercept) | 0.22156 | 0.09007 | 0.0472 |
| gc-strength | 0.11232 | 0.05079 | 0.0886 |
| delay | -0.15134 | 0.03486 | <0.001 |
| interaction | -0.07594 | 0.06617 | 0.6065 |
| | 100% noise | | |
| | Estimate | Std. Error | P-value |
| (Intercept) | 0.03357 | 0.007408 | <0.001 |
| gc-strength | 0.014703 | 0.005099 | 0.0134 |
| delay | -0.019791 | 0.003499 | <0.001 |
| interaction | -0.011352 | 0.006642 | 0.2568 |

(b) The effect of gc-strength and neuronal delay on the proportion of inverted results in the negative delay group

Table 5.2: Statistical results for the negative delay group

a noise-level of 20% only neuronal delay has a significant influence on the amount of correct results (see Table 5.3aA). When noise level is increased to 100%, there is a significant interaction between gc-strength and neuronal delay (see Table 5.3aB). The influence of neuronal delay is larger in connections with a stronger Granger causality influence. It should be noted that appropriate statistical power is obtained at neuronal delays around 0.1s or 0.2s depending on Granger causality strength and noise.

*No delay group*

The occurrence of inverted results in the no delay group was too small (< 0.01%) to perform meaningful statistical analyses on. In this group, only the proportion of correct results was assessed statistically.

The amount of correct results is significantly influenced by the interaction between noise and neuronal delay as well as by the three-way interaction between gc-strength, noise and neuronal delay. At a 20% noise level, the influence of gc-strength and neuronal delay on the amount of correct results have the same direction: the stronger the

| **Positive delay group (proportion correct results)** | | | |
|---|---|---|---|
| **A** | 20% noise | | |
| | Estimate | Std. Error | Pr(>\|z\|) |
| (Intercept) | 0.63621 | 0.12235 | <0.001 |
| GC | 0.11163 | 0.05419 | 0.126 |
| delay | 0.23487 | 0.03719 | <0.001 |
| GC:delay | -0.06861 | 0.07060 | 0.729 |
| **B** | 100% noise | | |
| | Estimate | Std. Error | Pr(>\|z\|) |
| GC=0.1 | | | |
| (Intercept) | 0.056683 | 0.007962 | <1e-10 |
| delay | 0.048286 | 0.005357 | <1e-10 |
| GC=0.2 | | | |
| (Intercept) | 0.084099 | 0.017352 | 2.51e-6 |
| delay | 0.079624 | 0.009133 | <1e-10 |
| GC=0.5 | | | |
| (Intercept) | 0.09549 | 0.02060 | 7.1e-6 |
| delay | 0.09549 | 0.01192 | <1e-10 |
| GC=0.9 | | | |
| (Intercept) | 0.09854 | 0.02014 | 1.99e-6 |
| delay | 0.11721 | 0.01268 | <1e-10 |

(a) Results for positive delay group

| **No delay group (proportion correct results)** | | | |
|---|---|---|---|
| **A** | 20% noise | | |
| | Estimate | Std. Error | Pr(>\|z\|) |
| (Intercept) | 0.35599 | 0.04121 | <0.001 |
| GC | 0.27791 | 0.07824 | 0.00137 |
| delay | 0.43760 | 0.05369 | <0.001 |
| GC:delay | -0.17573 | 0.10192 | 0.21702 |
| **B** | 100% noise | | |
| | Estimate | Std. Error | Pr(>\|z\|) |
| GC=0.1 | | | |
| (Intercept) | 0.035270 | 0.006076 | 1.29e-8 |
| delay | 0.092478 | 0.005935 | <1e-10 |
| GC=0.2 | | | |
| (Intercept) | 0.041930 | 0.006730 | 9.32e-10 |
| delay | 0.183667 | 0.008768 | <1e-10 |
| GC=0.5 | | | |
| (Intercept) | 0.03988 | 0.00930 | 3.57e-5 |
| delay | 0.24319 | 0.01212 | <1e-10 |
| GC=0.9 | | | |
| (Intercept) | 0.041050 | 0.008482 | 2.59e-6 |
| delay | 0.246919 | 0.010577 | <1e-10 |

(b) Results for no delay group

Table 5.3: Statistical results for positive and no delay group

gc-strength and the longer the neuronal delay, the more correct results are reported (see Table 5.3bA). At a noise-level of 100% the influence of neuronal delay is larger in connections with a stronger Granger causality influence (see Table 5.3bB).

As for the positive delay group, inversions are also extremely rare in the no delay group (<0.01%) and appropriate sensitivity is obtained for neuronal delays of 0.1s or 0.2s, depending on Granger causality strength and noise.

### 5.3.3 *Comparison crosscorrelations in simulation data and experimental data*

To investigate whether our simulations represent a physiologically plausible situation, we compared lagged cross correlations in our generated time series with cross correlations in time series of experimental data. At a noise level of 20% the averaged lagged crosscorrelations in the time series of our simulations (over all subjects and all connections) are 0.39, 0.5, 0.56, and 0.56 for Granger causality strengths 0.1, 0.2, 0.5 and 0.9 respectively. At a noise-level of 100% these crosscorrelation drop to respectively 0.09, 0.11, 0.13, and 0.13. The average lagged crosscorrelation between pMNS

| Short neuronal delays (0.01-0.05s) | | | | |
|---|---|---|---|---|
| Neuronal delay | Detected direction | Underlying direction X→Y | Y→X | Proportion correct |
| 0.01 | X→Y | 2993 | 488 | 0.86 |
| | Y→X | 16 | 60 | 0.79 |
| 0.02 | X→Y | 3036 | 450 | 0.87 |
| | Y→X | 10 | 69 | 0.87 |
| 0.05 | X→Y | 3222 | 406 | 0.89 |
| | Y→X | 4 | 88 | 0.96 |
| 0.01-0.05 | X→Y | 9251 | 1344 | 0.87 |
| | Y→X | 30 | 217 | 0.88 |

| Medium neuronal delays (0.1-0.5s) | | | | |
|---|---|---|---|---|
| Neuronal delay | Detected direction | Underlying direction X→Y | Y→X | Proportion correct |
| 0.1 | X→Y | 3396 | 317 | 0.91 |
| | Y→X | 1 | 134 | 0.99 |
| 0.2 | X→Y | 3723 | 169 | 0.96 |
| | Y→X | 0 | 167 | 1.00 |
| 0.5 | X→Y | 4471 | 0 | 1.00 |
| | Y→X | 0 | 365 | 1.00 |
| 0.1-0.5 | X→Y | 11590 | 486 | 0.96 |
| | Y→X | 1 | 666 | 1.00 |

| Long neuronal delays (1-2s) | | | | |
|---|---|---|---|---|
| Neuronal delay | Detected direction | Underlying direction X→Y | Y→X | Proportion correct |
| 1 | X→Y | 4798 | 0 | 1.00 |
| | Y→X | 0 | 1140 | 1.00 |
| 2 | X→Y | 4800 | 0 | 1.00 |
| | Y→X | 0 | 1200 | 1.00 |
| 1-2 | X→Y | 9599 | 666 | 0.9 |
| | Y→X | 0 | 2340 | 1.00 |

Table 5.4: Likelihood of true direction given a detected direction

areas in the experimental data is 0.82. Crosscorrelations in the real experimental data are significantly higher than in the simulation data for all Granger causality strengths and all noise-levels (in all cases df=287, $t \geq 30.78$, $p<0.001$), but were closer to the real experimental data in the 20% than in the 100% noise condition.

### 5.3.4 *Likelihood of true direction given a detected direction*

Our results of the likelihood of detecting the true direction are summarized in Table 5.4. While reading this table, it should be noted that 200 simulations were performed for each of the 4 GC strengths, and that 6 connections had underlying connectivity from the fast to the slower hemodynamic response region (X→Y), and 6 in the opposite direction (Y→X). This means that at each delay, 4800 simulations were performed in each direction. Accordingly, each of the columns at one delay could sum to up to 4800 simulations at most, if and only if each simulation leads to a significantly non-zero differential Granger causality. Summaries for each range of latencies were simply the sum of the simulations from the included latencies and therefore included 14400 simulations in each direction for the short and medium latencies and 9600 for the long delays.

Inspecting this table reveals three important findings. First, in all cases, a detected

direction of influence was always at least three times as likely to correspond to the true underlying direction of influence than to the opposite underlying direction of influence. At short delays (0.01-0.05s), the detected direction was accurate in over 85%, at medium delays (0.1-0.5s) in over 95%, and at long delays (1-2s) in 100% of the cases. Second, the proportion of detected influences that were accurate was generally similar in the X→Y and Y→X direction. Third, comparing how many simulations lead to a detection in the X→Y and Y→X direction (independently of the true underlying direction of influence) reavealed that the former was over 40 (for short), 18 (for medium) and 4 times (for long delays) more likely to be detected than the latter.

## 5.4 DISCUSSION

In this article, we focused on the question whether regional differences in hemodynamic response function can have an impact on group-level Granger causality analyses. More specifically, we simulated 200 experiments that measured 16 pairs of time series (connections) with a known directed influence between them. Each time series in a pair was convolved with a different combination of hemodynamic responses. We then calculated Granger causality and tested how often a group level result (across 20 subjects) would be correctly detected (hit), falsely rejected (miss), or detected but in the inverted direction (inversion).

The results of the simulations indicate several things. First, when no underlying directed influence is simulated between two time series, the average chance to find a significant group level result, with this set of hemodynamic responses, stays well below the set threshold of 5% for the t-test assessing whether average Granger causality $\neq 0$.

Second, when the hemodynamic response of the source has a smaller or equal delay parameter than the hemodynamic response of the target (the positive delay and no delay groups), the chance to find an inverted effect is near zero (<0.01%). However, when the difference in hemodynamic delay between source and target goes against the direction of influence (negative delay group), the chance to find an inverted result becomes substantial, in particular when the added neuronal delay between the two time series is small (see Figures 5.3 - 5.6). Overall, only long neuronal delays of 1-2s ensure that the chance of inversions is insignificant.

Third, when the analysis of the simulated data revealed a significant differential Granger causality value, the direction of the detected causality correctly reflects the simulated direction of influence in over 79% of the cases.

Finally our simulations show that the sensitivity of Granger causality is adequate (>80% hit ratio) for neuronal delays from 0.1 or 0.2s and longer when the source region had the same or a faster hemodynamic delay than the target. When the source had a slower hemodynamic delay than the target, sensitivity of Granger causality is

adequate at longer neuronal delays of 1-2s.

*Worst case scenario*

It should be emphasized that our simulations represent a worst-case scenario. We assessed the physiological plausibility of our simulations by comparing lagged cross correlations between time series from our simulation with these from data of a previous experiment (Schippers et al., 2009). This reveals that the dependencies in the time series of our simulations are lower than those in real experimental data. Dependencies between time series are therefor underestimated in our simulations, thereby representing a scenario worse than reality.

Furthermore, we used models of hemodynamic response functions that were measured with fMRI while participants were presented with a simple task (Handwerker et al., 2004). Latency differences (see Table 5.1) are thus influenced both by the 'intrinsic' hemodynamic response of that brain area and the delay of neuronal transfer between different brain areas due to the task. Our simulations thus overestimate hemodynamic latencies again representing a scenario worse than reality.

*Neural latencies*

Our simulations show that Granger causality results are strongly dependent on the neuronal latency between two areas. Single cell recording in the macaque monkey (Schmolesky et al., 1998) reveal that in the monkey, median latencies increase by approximately 20ms from one brain region in the visual hierarchy to the next. Simulated neuronal delays of 100ms therefore correspond to about 5 processing stages in the monkey, which could correspond to detecting an interaction between V1 and the FEF (Essen et al., 1992). In humans, these delays are less well studied, but due to increases in brain size, neuronal delays are probably longer. Nishitani and Hari (2002) for instance recorded MEG signals while participants viewed lip movements. They measured differences in peak responses of 100ms between responses in the occipital cortex and the inferior frontal gyrus, and 200ms between the occipital cortex and the primary motor cortex. Accordingly, the neuronal latencies in our simulations could be split in three ranges. Short connections (e.g. V1 → V2 or V1 → V3) include short latencies of 0.01, 0.02 and 0.05s. Long connections (e.g. V1 → IFG) include medium latencies of 0.1, 0.2 and 0.5. The longest latencies we used (1 or 2s) are probably more representative of communication between brains (e.g. Schippers et al., 2010).

*How useful is Granger causality mapping?*

The main conclusions of our simulations regard the sensitivity of Granger causality analysis and their validity. In terms of sensitivity, we show that Granger causality is not very sensitive for short range connections within a brain. At such short latencies (0.01-0.05s) sensitivity is low throughout our simulations. Long-range connections can be adequately studied using Granger causality as long as differences in hemodynamic response do not go against the direction of information transfer. In these cases sensitivity is adequate (> 80%), inversions are rare, and false positives are below 5%. Finally, Granger causality mapping is ideal for measuring situations with long delays such as the ones encountered in the analysis of information transfer across brains (Schippers et al., 2010) where sensitivity is high, inversions inexistent, and false positives well controlled, independently of realistic variations in hemodynamic response. In terms of validity (whether a detected Granger causality correctly reflected the true underlying direction of influence), we found that at a realistic noise level the majority of detected causalities accurately reflected the underlying direction of influence. Coupled with the fact that less than 5% of cases in which no connection was simulated lead to a false positive Granger causality result, this indicates that Granger causality analysis remain extremely valid even in the context of differences in hemodynamic responses.

Several researchers have suggested that differences in hemodynamic response functions between brain areas could lead to spurious Granger causality results (David et al., 2008; de Marco et al., 2009; Roebroeck et al., 2005; Friston, 2009; Chang et al., 2008). Roebroeck et al. (2005) suggested that contrasting the differential Granger causality values between different conditions may be the only way to avoid these spurious effects. Because none of these researchers directly simulated realistic differences in hemodynamic responses across individuals, these considerations were based on intuition. Here, using realistic variations of hemodynamic response function at a group level, we show that spurious findings are actually rare. We therefore suggest that differential Granger causality is a valid method to determine the dominant direction of information flow even within a single experimental condition at the group level. This greatly increases the applicability of Granger causality by allowing its application to situations in which control conditions not involving similar connections may be difficult to perform.

*Estimation of hemodynamic response*

The current simulations show that the difference in hemodynamic delay between two brain areas contains important information about the sensitivity (but not the validity)

of short and long range within brain Granger causality analyses. If the source area of a connection has a hemodynamic delay that is as short or shorter than that of the target region, Granger causality analyses have good sensitivity when a $p<0.05$ criteria is used. At this threshold the majority of simulated influences were detected. When the difference is negative however, i.e. information flows from a region with a slower to a region with a faster hemodynamic response, the risk of not detecting existing connections is large. Accordingly, estimating the difference in hemodynamic delay between two areas can be used to shed light on negative findings in Granger causality analysis. If experimenters fail to detect a hypothesized information flow from a brain region with a faster to one with a slower response at a permissive threshold ($p<0.05$), this negative finding is unlikely to be due to a lack of sensitivity and challenges the hypothesis. If experimenters had hypothesized a direction of information flow from the slower to the faster region, and fail to detect this influence, they need to keep in mind that sensitivity is lower, and a negative finding is less challenging to their hypothesis. Importantly, pooling over the range of relationships between the delays of the two regions we explored based on empirically measured HRF variance, our data suggests that on average, if experimenters do find a significant differential Granger causality result, they can be relatively confident that the result is in the correct direction.

## 5.5 CONCLUSIONS

Based on the outcome of our simulations, examining the differential Granger causality across a group of participants provides a valid measure of underlying effective connectivity. If a significant differential Granger causality is detected, this finding is (a) unlikely to have arisen without a true underlying information flow between these regions and (b) the detected direction of influence is likely to reflect the true direction of influence. The sensitivity of the method however greatly varies based on the neuronal delay separating the source and the target region and the difference in hemodynamic delay between these regions. If information flows from a region to one with the same or a slower hemodynamic delay, Granger causality has adequate sensitivity for longer within brain connections corresponding to neuronal delays of at least 100ms. If the source region has a slower hemodynamic delay than the target region, only between brain connections with delays of at least 1s can be detected with appropriate sensitivity. Accordingly, our simulations suggest that when Granger causality reveals significant results, these results can be trusted even in the context of systematic differences in hemodynamic delays. If the analysis however does not lead to significant results, the relative lack of sensitivity of the method should be kept in mind in particular when used to analyze connections within a brain.

ACKNOWLEDGEMENTS

The research was supported by a VIDI grant of the Dutch science foundation (N.W.O.) and a Marie Curie Excellence Grant of the European Commission to CK. Data were collected by Handwerker et al. (2004) with the support of NIH grants MH63901 and NS40813.

# BIBLIOGRAPHY


Aguirre, G. K., Zarahn, E., and D'esposito, M. (1998). The variability of human, bold hemodynamic responses. *Neuroimage*, 8(4):360–9.

Chang, C., Thomason, M. E., and Glover, G. H. (2008). Mapping and correction of vascular hemodynamic latency in the bold signal. *Neuroimage*, 43(1):90–102.

David, O., Guillemain, I., Saillet, S., Reyt, S., Deransart, C., Segebarth, C., and Depaulis, A. (2008). Identifying neural drivers with functional mri: an electrophysiological validation. *PLoS Biology*, 6(12):2683–97.

de Marco, G., Devauchelle, B., and Berquin, P. (2009). Brain functional modeling, what do we measure with fmri data? *Neuroscience Research*, 64(1):12–9.

Deshpande, G., Sathian, K., and Hu, X. (2009). Effect of hemodynamic variability on granger causality analysis of fmri. *Neuroimage*, 52(3):884–896.

Essen, D. C. V., Anderson, C. H., and Felleman, D. J. (1992). Information processing in the primate visual system: an integrated systems perspective. *Science*, 255(5043):419–23.

Friston, K. (2009). Causal modelling and brain connectivity in functional magnetic resonance imaging. *PLoS Biology*, 7(2):e33.

Geweke, J. (1982). Measurement of linear dependence and feedback between multiple time series. *Journal of the American Statistical Association*, 77(378):304–313.

Goebel, R., Roebroeck, A., Kim, D.-S., and Formisano, E. (2003). Investigating directed cortical interactions in time-resolved fmri data using vector autoregressive modeling and granger causality mapping. *Magnetic resonance imaging*, 21:1251–1261.

Granger, C. (1969). Investigating causal relations by econometric models and cross-spectral methods. *Econometrica*, 37(3):424–438.

Handwerker, D. A., Ollinger, J. M., and D'Esposito, M. (2004). Variation of bold hemodynamic responses across subjects and brain regions and their effects on statistical analyses. *Neuroimage*, 21(4):1639–51.

Kruggel, F. and von Cramon, D. Y. (1999). Temporal properties of the hemodynamic response in functional mri. *Human brain mapping*, 8(4):259–71.

Nishitani, N. and Hari, R. (2002). Viewing lip forms: cortical dynamics. *Neuron*, 36(6):1211–20.

Rajapakse, J. C., Kruggel, F., Maisog, J. M., and von Cramon, D. Y. (1998). Modeling hemodynamic response for analysis of functional mri time-series. *Human brain mapping*, 6(4):283–300.

Roebroeck, A., Formisano, E., and Goebel, R. (2005). Mapping directed influence over the brain using granger causality. *Neuroimage*, 25:230–242.

Schippers, M. B., Gazzola, V., Goebel, R., and Keysers, C. (2009). Playing charades in the fmri: are mirror and/or mentalizing areas involved in gestural communication? *PLoS ONE*, 4(8):e6801.

Schippers, M. B., Roebroeck, A., Renken, R., Nanetti, L., and Keysers, C. (2010). Mapping the information flow from one brain to another during gestural communication. *Proceedings of the National Academy of Sciences of the United States of America*, 107(20):9388–93.

Schmolesky, M. T., Wang, Y., Hanes, D. P., Thompson, K. G., Leutgeb, S., Schall, J. D., and Leventhal, A. G. (1998). Signal Timing Across the Macaque Visual System. *Journal of Neurophysiology*, 79(6):3272–3278.

Snijders, T. and Bosker, R. (1999). *Multilevel analysis: an introduction to basic and advanced multilevel modeling*. London: Sage.

Wiener, N. (1956). Theory of prediction. *Modern Mathematics for Engineers, Series 1*.

# SUPPLEMENTARY INFORMATION 1

## OVERVIEW SET OF HEMODYNAMIC RESPONSES

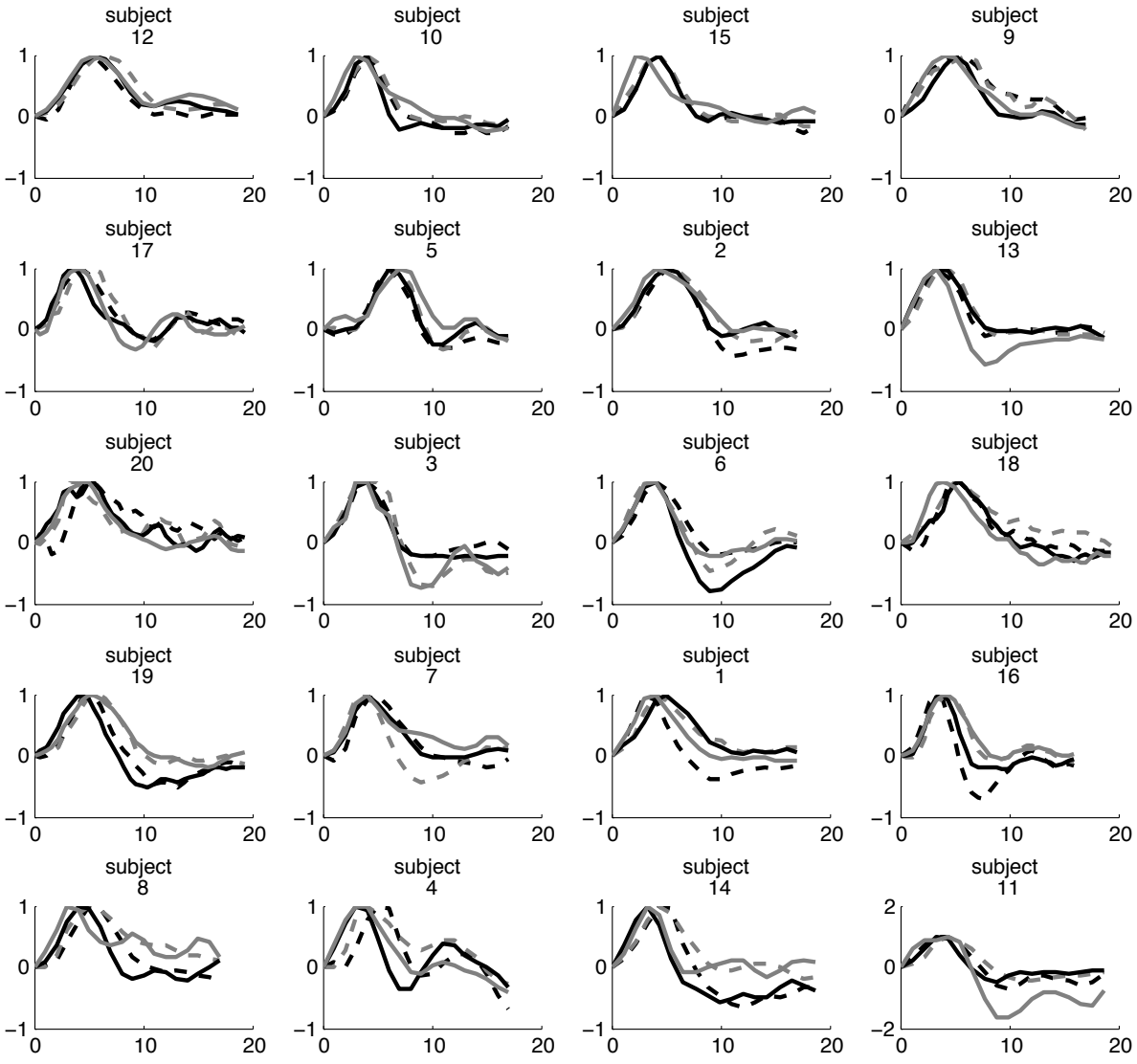



Figure 5.7: Hemodynamic response shapes for the four measured areas in 20 different subjects, sorted on amount of variability.

# SUPPLEMENTARY INFORMATION 2

## SPREAD PARAMETERS OF HEMODYNAMIC RESPONSE MODELS

The hemodynamic responses were obtained and published by Handwerker et al. (2004). These HRFs were modeled by the sum of two gamma functions:

$$y(t) = \underbrace{A_1\left(\frac{x(t)-\delta_1}{\tau_1}\right)^2 e^{-\left(\frac{x(t)-\delta_1}{\tau_1}\right)}}_{peak} + \underbrace{A_2\left(\frac{x(t)-\delta_2}{\tau_2}\right)^2 e^{-\left(\frac{x(t)-\delta_2}{\tau_2}\right)}}_{undershoot} + C$$

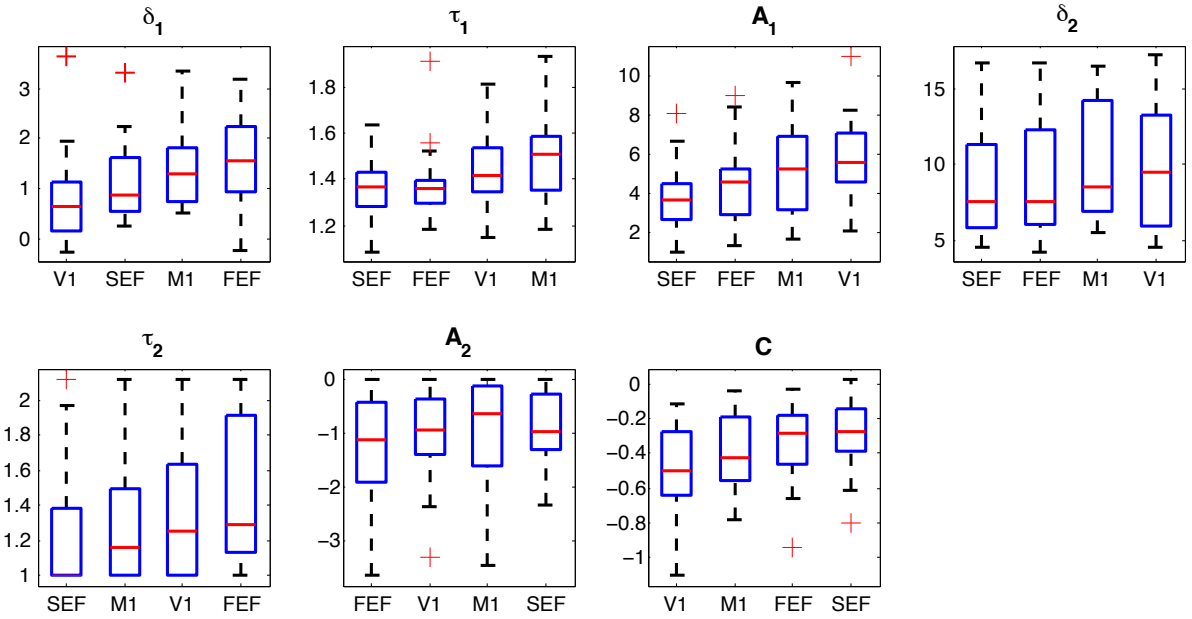



Figure 5.8: Spread of the parameters that were used to fit two gamma-functions on the measured hemdynamic responses. See Figure 1 of Handwerker et al. (2004) for a graphical explanation for the two-gamma functions.

# SUPPLEMENTARY INFORMATION 3

## RESULTS SIMULATIONS WITH 100% NOISE LEVEL

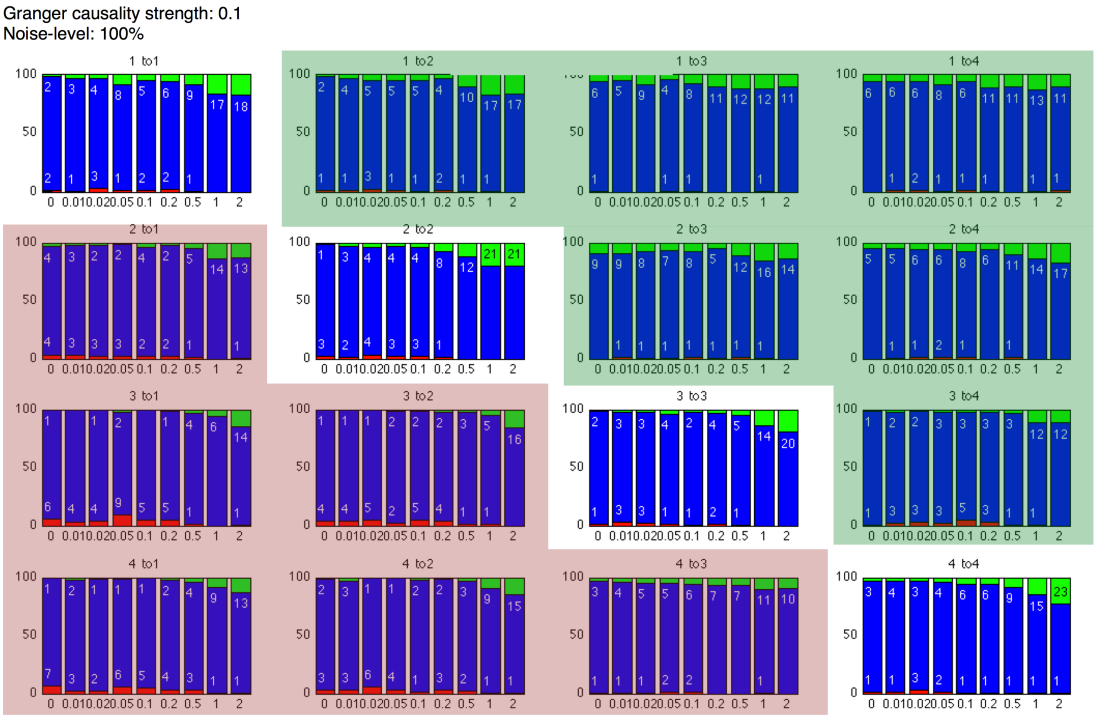



Figure 5.9: Proportion of correct (bright green), inverted (bright red) and non-significant (blue) results per connection and per neuronal delay for a noise-level of 100% and Granger causality strength of 0.1. Source (rows) and target (columns) are sorted on their mean onset parameter. Connections are divided in a positive delay group (green background), a no-delay group (white background) and a negative delay group (red background). A correct result for a given connection and neuronal delay signifies that the group Granger causality analysis indicated the direction of influence that was modeled in the time series (for example for connection '3 to 1', the correct direction of influence goes from area 3 to area 1). An inverted result indicates a significant Granger causality result with a direction opposite of the direction that was modeled in the simulation. A non-significant result indicates a result that did not exceed the threshold.

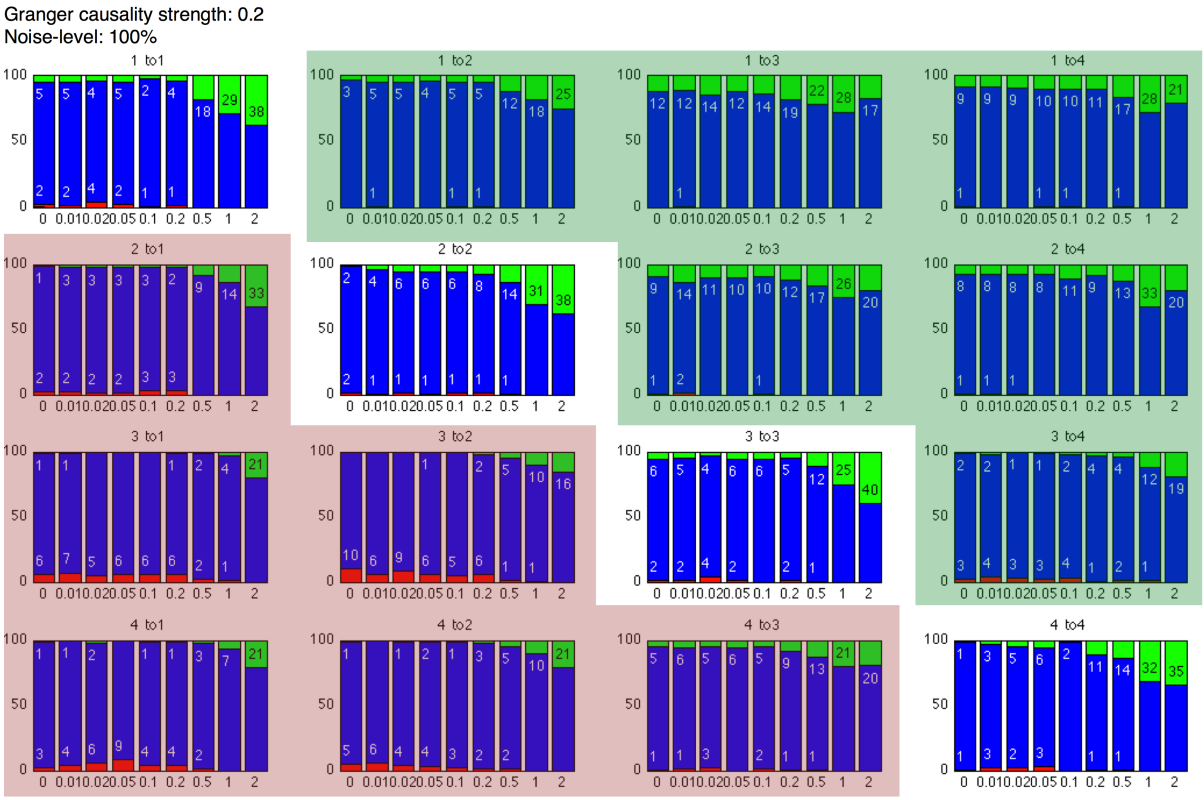



Figure 5.10: Proportion of correct (green), inverted (red) and non-significant (blue) results per connection and per delay for a noise-level of 100% and Granger causality strength of 0.2. See also caption of Figure 5.9

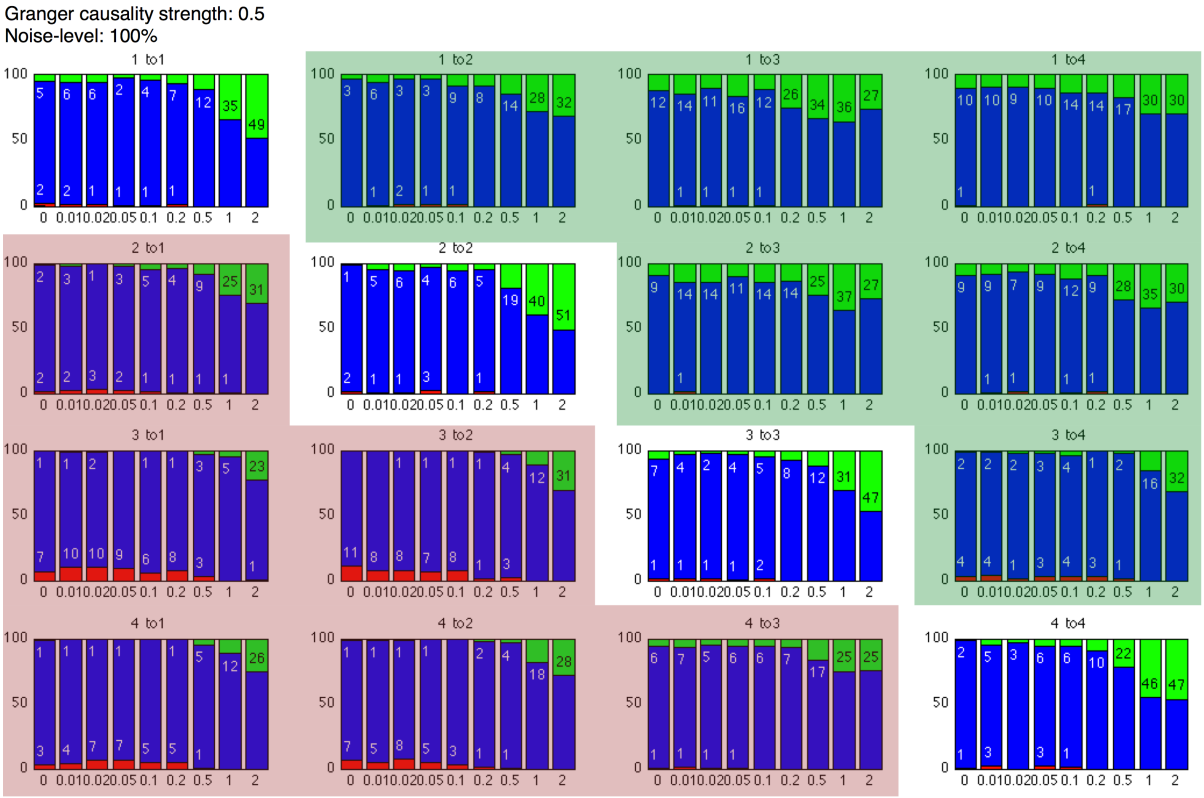



Figure 5.11: Proportion of correct (green), inverted (red) and non-significant (blue) results per connection and per delay for a noise-level of 100% and Granger causality strength of 0.5. See also caption of Figure 5.9

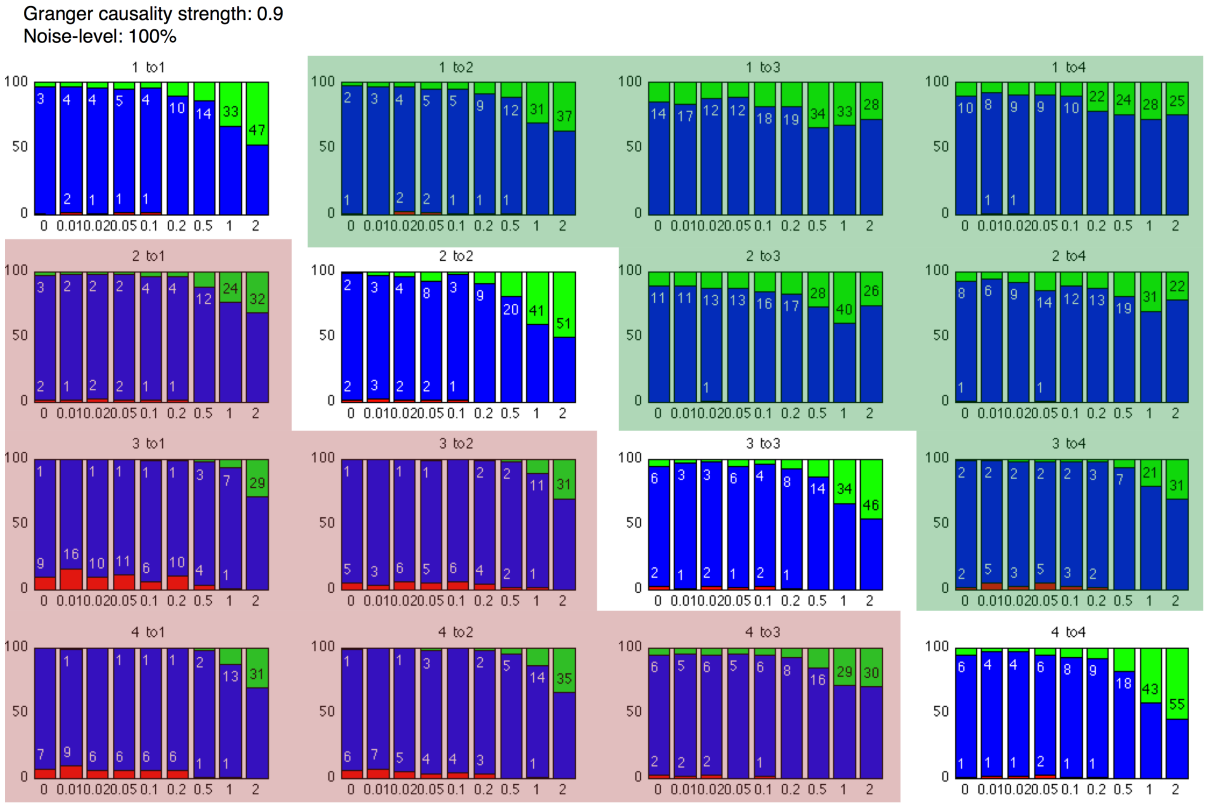



Figure 5.12: Proportion of correct (green), inverted (red) and non-significant (blue) results per connection and per delay for a noise-level of 100% and Granger causality strength of 0.9. See also caption of Figure 5.9

## CONNECTION 3'

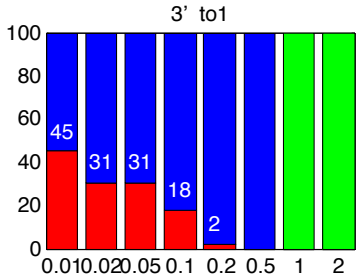



Figure 5.13: Results of Granger causality simulations in a connection with artificially created hemodynamic responses. The hemodynamic response of the source is created to have the same variance in parameters over subjects as area 4, but on average a smaller delay parameter (hemodynamic delay, $\delta_1$). All subject-specific parameters from area 4 are used to create area 3', but the $\delta_1$ parameters are shifted by the difference in average $\delta_1$ between area 4 and 3.

# SUPPLEMENTARY INFORMATION 5

## SIMULATIONS OVERVIEW

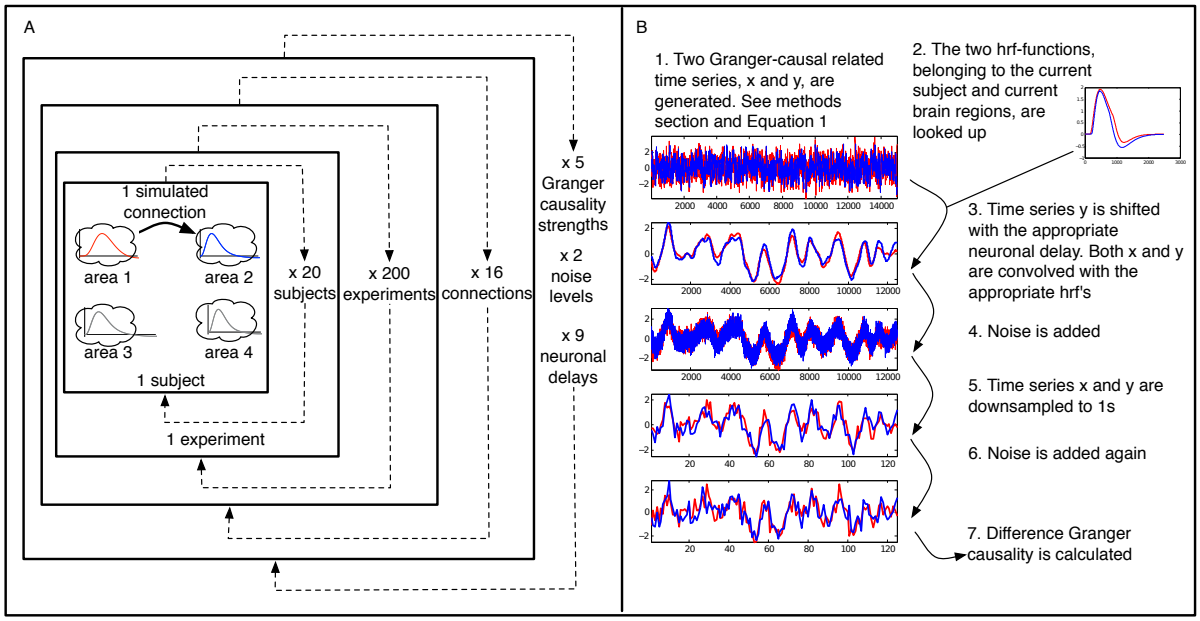



Figure 5.14: Overview of the procedure used to generate the simulated data for the analysis presented in chapter 5.

Figure 5.14 shows the procedure used to generate the simulated data for the analysis presented in this paper. (a) For each of the 20 subjects, Handwerker et al. (2004) measured the hemodynamic response in four brain areas. We picked two of these hemodynamic responses of one subject, and used these to generate simulated data (see panel (b)). This procedure is then repeated for all 20 participants, using subject-specific hemodynamic responses in each case. This represents the experimental unit of our simulation, reflecting what would be the dataset of a 20 subject experiment looking at one connection between two brain regions. We then repeat this 200 times to generate a distribution of results. The same is then done by picking another two brain regions (e.g. area 1 → area 3), and so on for all 16 possible connections (we also simulated area 1→ area 1 to look at connections between regions with identical

hemodynamic responses). This was done using 5 different Granger causality strengths, 2 noise levels and 9 neuronal delays. (b) for each of the simulation, we generated two time series using a bi-dimensional first-order vector autoregressive process (see Equation 5.1 in main article). Both signals are simulated for 15000 time steps of 10 ms (150 sec). Target time series $y_i$ is then shifted to represent the neuronal transmission delay. Both signals are convolved using subject and area specific hemodynamic response models and gaussian noise is added to represent physiological noise in the BOLD response. Subsequently, the signal is down sampled to 1 s (resembling an acquisition rate (TR) of an MR-scanner) and gaussian noise is again added to represent acquisition noise. After each step the signals are normalized to zero mean and unit variance. The total amount of noise added was either 20% or 100%. After noise was added at two different stages of the simulation, the standard deviation of the signal had increased to 120% and 200% of the original standard deviation for the respective noise levels.

# 6

# MAPPING THE FLOW OF INFORMATION WITHIN THE PUTATIVE MIRROR NEURON SYSTEM DURING GESTURE OBSERVATION

*Submitted as:* Schippers, M.B., Keysers, C. Mapping the flow of information within the putative mirror neuron system during gesture observation.

ABSTRACT

The putative Mirror Neuron System may either function as a strict feed-forward system or as a dynamic control system. A strict feed-forward system would predict that action observation leads to a predominantly temporal → parietal → premotor flow of information in which a visual representation is transformed into motor-programs which contribute to action understanding. Instead, a dynamic feedback control system would predict that the reverse direction of information flow predominates because of a combination of inhibitory forward and excitatory inverse models. Here we test which of these conflicting predictions best matches the information flow within the putative Mirror Neuron System (pMNS) and between the pMNS and the rest of the brain during the observation of comparatively long naturalistic stretches of communicative gestures. We used Granger causality to test the dominant direction of influence. Our results fit the predictions of the dynamic feedback control system: we found predominantly an information flow within the pMNS from premotor to parietal and middle temporal cortices. This is more pronounced during an active guessing task than while passively reviewing the same gestures. In particular, the ventral premotor cortex sends significantly more information to other pMNS areas than it receives during active guessing than during passive observation.

## 6.1 INTRODUCTION

Prediction of future events is a key feature of many cognitive abilities, such as language and action and has even been suggested as "one of the main unitary principles of cognition" (Bubic et al., 2010). Applied to the social domain, correct prediction of other's behavior helps in understanding other's intentions (Blakemore and Frith, 2005; Falck-Ytter et al., 2006), allows us to coordinate our behavior with others (Kokal et al., 2009; Sebanz et al., 2006) and is important for effective communication (Garrod and

Pickering, 2009).

Prediction of other's behavior can be conceptualized on several levels and different timescales. For example, on a high cognitive level, humans can deliberately mentalize about beliefs and desires of others and how these will potentially influence their future behavior (Frith and Frith, 1999). On a lower cognitive level and a smaller time scale however, prediction of other's behavior already occurs while the other person is still performing the action (Umiltà et al., 2001; Urgesi et al., 2010). The putative Mirror Neuron System (pMNS), a set of brain regions involved both in observing the actions of others and in programming similar actions (Gazzola and Keysers, 2008), has been suggested as a network that might be involved in predicting other people's behavior on the basis of perceptual-motor mapping (Blakemore and Frith, 2005; Urgesi et al., 2010; Keysers and Perrett, 2004; Hesse et al., 2009; Kilner et al., 2007b,a; Lamm et al., 2007; Schubotz, 2007; Sebanz and Knoblich, 2009). The pMNS is a network in the human brain consisting of ventral and dorsal premotor cortex, anterior inferior parietal lobule and adjacent somatosensory area BA2 and middle temporal gyrus (Gazzola and Keysers, 2008; Keysers and Gazzola, 2009; Keysers, 2009). These areas of the human brain, normally associated with planning, preparation, execution and proprioception of our own actions, were found to be also involved in the hearing or observation of actions (Gazzola and Keysers, 2008; Chong et al., 2008; Dinstein et al., 2007; Filimon et al., 2007; Gazzola et al., 2006). These areas show a tight coupling between perception and action, which has led to the idea that we understand the actions of others in part by transforming them into the motor vocabulary of our own actions.

Internal models (forward and inverse models) could play a key role in the transformation from observing an action to simulating this action internally and predicting its sensory consequences (Wolpert et al., 2003). Forward and inverse models were originally conceptualized in the context of action control (Wolpert et al., 2003). During action execution the forward model sends an efference copy of the action to sensory areas to predict sensory and proprioceptive consequences of that action (Greenwald, 1970; Voss et al., 2006). These predicted consequences can then be compared with actual sensory input. If a mismatch is detected between actual and predicted sensory input, the action can be adjusted. An inverse model follows the reverse path of a forward model: it calculates what motor commands should be executed to achieve a certain goal or end-state (Wolpert et al., 2003). It has been suggested that the pMNS hosts such forward and inverse models. One of the consequences of these models in the pMNS is that they would automatically work together to achieve a prediction of other's intentions and behaviors based on simulation (Gazzola and Keysers, 2008; Keysers and Perrett, 2004; Kilner et al., 2007b,a; Lamm et al., 2007; Blakemore and Decety, 2001; Csibra and Gergely, 2007; Miall, 2003). In particular during action observation the inverse model could be used to infer the motor program that could have

caused the sensory input (Gazzola and Keysers, 2008; Kilner et al., 2007b; Blakemore and Decety, 2001). Hebbian learning has been suggested to be a candidate process by which these models are instantiated (Gazzola and Keysers, 2008; Keysers and Perrett, 2004). A conceptually similar model for the acquisition of these internal models is the associative sequence learning theory (Heyes et al., 2000).

Within the pMNS the forward model for visual information would follow the path of premotor cortices → posterior parietal lobule → middle temporal gyrus (MTG), such that predicted visual consequences of the executed action are represented in the MTG (Gazzola and Keysers, 2008). The inverse model would then go in the opposite direction of this path, MTG → posterior parietal lobule → premotor cortices. During action observation, forward and inverse models can work together to achieve an interpretation of the observed action. The observed action is first internally simulated in the premotor cortex, which is reached through the inverse model. Then the forward model can be used to predict the sensory consequences of the internally simulated action and be compared to the actual consequences to assess the accuracy of that prediction (Gazzola and Keysers, 2008; Wolpert et al., 2003).

Single cell recordings in the monkey give us further information about the working of these forward and inverse models. When a monkey sees the actions of others, about 10% of the neurons responsible for executing an action increase their firing rate (Gallese et al., 1996; Keysers et al., 2003) while very few reduce their firing rate (Kraskov et al., 2009). Accordingly, inverse models seem to be primarily excitatory. In contrast, about 50% of neurons responding to the sight of a particular movement in the temporal lobe decrease their firing rate if the monkey is causing this movement (Hietanen and Perrett, 1993, 1996), while none seems to augment its firing rate during motor execution. Accordingly, forward models appear to be primarily inhibitory, and seem to subtract expected visual input from the actual input. A similar phenomenon can be observed in the somatosensory domain: we are unable to tickle ourselves, because a copy of the motor program inhibits the somatosensory processing of self induced tickling (Blakemore et al., 1998). This architecture predicts that the pMNS acts like a negative feedback loop: when we see a predictable chain of events, the beginning is fully represented in the visual cortex and triggers motor programs through the inverse model. These motor programs are then forwarded to predict future visual stimuli. If these stimuli conform to the predictions, they will be inhibited. The visual → premotor stream of information is reduced and the premotor representations triggered will not be substantially updated. If the visual information violates the predictions, it will not be inhibited, leading to a renewed visual → premotor stream of information and an update of the motor representations.

Many have taken what is known about the anatomy and function of the mirror neuron system to mean that during action observation the flow of information goes from temporal lobe → posterior parietal lobe → premotor cortex (thus visual information

represented in the temporal lobe triggers activity in mirror neurons of the posterior parietal lobe which in turn triggers activity in mirror neurons in the premotor cortex). This in turn is thought to contribute to an understanding of the actions of others (see Kilner et al., 2007b, for an excellent discussion of this issue). We will call this account a strict feed-forward account of action observation. In contrast, seeing the pMNS as a combination of excitatory inverse and inhibitory forward models, as described above, makes the pMNS a dynamic control system rather than a strict feed-forward recognition system. Importantly, these two visions make different predictions with regard to the dominant direction of information flow during the observation of longer sequences of actions.

A strict feed-forward account would predict that even while looking at longer sequences of actions, the dominant direction of information flow is anterior, from posterior visual to more anterior premotor regions. Instead, the dynamic control system account makes the opposite prediction. When a sequence of actions starts, the anterior flow of information briefly dominates. As soon as the first motor programs are activated, however, forward models would start predicting the future visual input in a posterior flow of information from premotor to visual areas. Because this posterior flow is inhibitory (Keysers and Perrett, 2004; Hietanen and Perrett, 1993, 1996), the more predictable the observed action sequence, the more visual input is inhibited, and the more the anterior flow of information is therefore reduced. Accordingly, because action sequences are to a great extend predictable, the dynamic control system model would predict that except for the initial phase of any action observation, the posterior flow of information would be dominant.

To test which of these conflicting predictions best fits empirical data, we performed an experiment in which participants were engaged in the game charades. They were required to deduce what word their partner was gesturing while viewing a video of these gestures of at least 50s. This data has previously been analyzed to examine the relationship between the brain activity of the sender and receiver of the gestures (Schippers et al., 2010, 2009), and results suggested that the pMNS indeed processes the moment-to-moment state of the brain of the sender during gesture interpretation. Here we will use the data of the observers' brain to investigate the dominant direction of information flow within the pMNS during gesture observation using differential Granger causality mapping (Schippers et al., 2010; Roebroeck et al., 2005; Goebel et al., 2003; Jabbi and Keysers, 2008). This method contrasts the degree to which the BOLD signal of one brain region predicts or is predicted by that in another brain region (Roebroeck et al., 2005; Goebel et al., 2003).

Before applying Granger causality to this data, we examined if Granger causality can indeed deduce the dominant flow of neural information even when applied to fMRI data, and in particular, whether differences in the hemodynamic response between different brain regions and between participants (Handwerker et al., 2004)

would make Granger causality analysis unreliable. Roebroeck et al. (2005) had alerted potential Granger causality users to this risk: if neural activity in region A consistently predicts neural activity in region B more than the other way around, but the hemodynamic delay of region A is longer than that of region B, the BOLD signal of region A might turn out to be predicted by that of B more than the other way around. The effect of this hemodynamic variability has been assessed on a single-subject level by Deshpande et al. (2009). However, noone had, to our knowledge, assessed whether for the kind of group study we wanted to conduct, Granger causality analyses were indeed rendered unreliable because of intra- and inter-subject variance in the hemodynamic response functions. To directly assess this risk, we obtained the hemodynamic response functions of 4 brain regions measured in 20 subjects by Handwerker et al. (2004), and used them to convolve time series of simulated neural data of known directional influence. This ensured that we could apply the same group analysis used in this manuscript on data with realistic differences in hemodynamic delay, and check how often a detected direction of influence corresponded to the actual direction of influence. The data, submitted to this journal (Schippers et al., subm), show that although Granger causality analysis often failed to detect a significant directed influence when there actually was one. Whenever Granger causality analysis did detect a significant directed influence between two brain regions, the detected direction of dominant influence was actually accurate in over 80% of the cases. We take these results to mean that we might fail to detect any directed information flow in our experiment using Granger causality, but that if we find such directed information flow, the detected dominant direction of information flow would inform us about which model best describes the working of the pMNS: if the anterior flow in the visual → premotor direction dominates, a feed-forward account of the pMNS is best supported by the data. If on the other hand the posterior flow in the premotor → visual direction dominates, a dynamic control system model would better fit the data.

Roebroeck et al. (2005) however suggested that the safest way to apply Granger causality is to perform calculations in two different conditions. If the difference in information flow (as measured by Granger causality) between regions A and B changes from one condition to another, it is unlikely that this change could be an artifact of hemodynamic delay differences, which should be condition independent (Roebroeck et al., 2005). To enable this conservative approach, we showed our participants the same gesture movies again on a later day, with an instruction to try not to interpret the gestures. We reasoned that without an aim to decode the gestures, the pMNS of the observer should generate fewer predictions. Accordingly, contrasting the results of Granger causality during active guessing with those during passive observation should evidence more posterior flow of information in the active guessing condition.

## 6.2 MATERIAL AND METHODS

*The data used for this analysis is the same as the one used to study information flow between brains in Schippers et al. (2010). Accordingly, the methods of data acquisition have been described elsewhere (Schippers et al., 2010). Here we will therefore focus on the issues regarding the specific data analysis applied here to measure the dominant direction of information flow within the pMNS. A detailed methods description is also attached as Supplementary Methods.*

### *Granger causality analyses*

Granger causality (G-causality) utilizes the concept of temporal precedence to formalize causality between two time series: if a signal change in A is consistently followed by a signal change in B, A Granger-causes B. Mathematically, this is calculated by comparing two regression equations: one in which the current value of a time series $y_i$ is explained by its own past ($y_{i-j}$) with one in which the same time series $y_i$ is explained both by its own past and the past of another time series ($x_{i-j}$). The ratio between the error variances of the former and the latter quantifies the influence $x$ exerts on $y$. Detailed descriptions of Granger causality can be found in Roebroeck et al. (2005) and Schippers et al. (2010).

Granger causality analyses were performed as described in Roebroeck et al. (2005). In short, given two time-series (for a seed and another point on the cortical surface), autoregressive models are estimated that quantify G-causality. Given a seed, maps are created that specify G-causal influence from the seed to the rest the brain, as well as influence in the reverse direction, i.e. from anywhere in the brain to the seed. These two directions of G-causality are then subtracted from each other to generate differential G-causality maps, such that positive values indicate more G-causality from the seed to the target than from the target to the seed. This differential G-causality measure was used for two reasons. First, it generates values that are approximately normally distributed, with a mean of zero under the null hypothesis of an absence of (indirect) causal relationship at the neural level between seed and target, and are thus suitable for parametric testing at the second level (see below and Roebroeck et al. (2005). Second, it has been shown that when the BOLD signal is used to estimate G-causality, the differential G-causality (i.e. X→Y - Y→X) is more robust than testing the individual components due to the filtering properties of the hemodynamic response and the relatively low sampling rate of fMRI (Roebroeck et al., 2005). A separate differential G-map was calculated for each of the 8 seed regions (see below) for each participant. These differential G-causality maps were then taken, separately for each seed region map, to the second level (see below) and thresholded

for multiple comparisons at p<0.05 using a cluster threshold determined by a Monte Carlo simulation method (Forman et al., 1995; Hagler et al., 2006). The order of the estimated autoregressive models was 1, i.e. taking one preceding time point into account to predict the current activity, corresponding to 1.33 seconds (1TR). This interval covers neuronal transfers within a brain.

*Seed ROIs*

The ROIs that were used as seeds in the Granger causality analysis were defined as those 'putative mirror' areas that were active both during gesturing and guessing using a traditional General Linear Model (GLM) analysis on the same data. The GLM for gesturing was estimated using the entire period in which the gesture was executed as the only regressor. The GLM for the guess runs included two regressors: 1) the period from onset of the movie in which the gesturer was shown until the time of button press (with which participants indicated they thought they knew what was being gestured) and 2) from button press until the participant had given a response (i.e. had chosen one of the answers from a multiple choice menu). All regressors were convolved with canonical hemodynamic response functions. The mean parameter estimates of the contrasts gesturing versus baseline (20s fixation cross) and guessing versus baseline were tested at the second level using a one-sample t-test. Both results of the second-level random effects analysis of gesturing and guessing versus baseline were thresholded at p<0.0001 (uncorr.). All these vertices also survive an FDR-correction of p< 0.05. We used this stringent criterion to reduce the size of our seeds. The resulting maps were binarized (i.e. contained value 1 at above-threshold vertices[1] and 0 at below-threshold vertices). The binary maps were multiplied to implement a logical 'and'. From the resulting clusters, we used the typical mirror regions: the middle temporal gyrus (LH: 69, RH: 256 vertices), the ventral (LH: 42, RH: 123 vertices) and dorsal premotor cortex (LH: 69, RH: 147 vertices) and a larger cluster in the parietal lobe. We excluded one cluster at the posterior end of the Sylvian fissure. To reduce the size of our parietal lobe seed, we used only that part of the ROI that overlapped with area BA2 (LH: 529, RH: 552 vertices), because BA2 was found to be the most consistent location of mirror voxels in a previous analysis of the execution and observation of goal directed behavior (Gazzola and Keysers, 2008) and because there is increasing interest in the role of the somatosensory cortex in action simulation (Caspers et al., 2010; Keysers et al., 2010). The location of BA2 was defined by projecting the maximum probability map of BA2 from the anatomy toolbox (Eickhoff et al., 2005) onto a cortical surface segmentation of the Colin brain (Holmes et al., 1998). For informative purposes, we also included a seed region that

1 A vertex is a node on the cortical surface

contains all shared vertices of the parietal lobe minus those vertices of BA2. Results of this analysis are presented in Supplementary Figure 6.5 and are extremely similar to those obtained from BA2. These particular ROIs were chosen because our primary aim was to examine the information flow within the pMNS. For visual orientation, we show the seeds in the figures (solid colors) together with an outline of the pMNS areas when results of guessing and gesturing are thresheld at $p<0.001$ (uncorr.).

*Time series*

The input of the Granger causality analysis consisted of the average time course of the seed during guessing or passive observation as well as all the corresponding time courses of the vertices of the rest of the brain during guessing or passive observation. The time courses were truncated to contain only those parts that reflected the steady-state part of either the guessing or the passive observation. We excluded 1) 15TR from the beginning and 5TR from the end of the observation, thereby removing the on- and offset transients, 2) the period from the button press of the guesser until the onset of the next gesture, and 3) the baseline fixation cross period between two trials. Additionally, information about the beginning and end of each separate gesture-part was taken into account in the Granger causality analysis, such that autoregressive model estimation was pooled over calculations on separate blocks rather than calculated over a single time-course with all blocks concatenated. On average, participants watched the gesture movies for 58 seconds, which corresponds to ±43 TR. This means that on average 43 - 15 - 5 = 23 TR per trial were included in the Granger analyses, cumulating to around 23 x 14 = 322 TR per participant.

*Second level Granger analysis*

Random effects' testing was performed by t-tests at the second level. Two-tailed one-sample t-tests were computed for a single differential G-causality maps, with a null-hypothesis of a zero value (i.e. there is as much G-causality from seed → target than from target → seed). Two-tailed paired t-tests were computed when differential G-causality maps were compared between conditions, with a null-hypothesis of a zero value for the difference of the maps. Random effects t-maps were then statistically thresholded at $p<0.05$ and corrected for multiple comparisons by using a Monte Carlo simulation-based cluster-size threshold adjustment (Hagler et al., 2006).

## 6.3 RESULTS

### 6.3.1 *Behavioral Results*

During guessing the participants were asked to consider each movie for at least 50 seconds. After the 50s they could push the button when they thought they knew what was being gestured to enter the multiple-choice menu. The average latency to response was 58 seconds. Participants were equally accurate on both categories: 82.5% of the object words were guessed correctly against 86.5% of the action words ($t(17)=-1$, $p>.33$). We did not find a significant difference between the two types of gestures, neither in terms of latency to respond (58.7s ± 6.5s for action and 60.8s ± 6.8s for object words, $t(17)=1.16$, $p>.26$) nor in terms of accuracy (6.06 ± 0.73sd correct out of 7 action and 5.78 ± 1.11sd correct our of 7 object words, $t(17)=-1$, $p>.33$). Words that were guessed incorrectly were watched significantly longer than words that were guessed correctly: 58s ± 5s for the 289 correct guesses versus 68s ± 12s for the 47 incorrect guesses ($t(16)=-4.41$, $p<.0005$).

### 6.3.2 *G-causality during guessing*

Results of Granger causality from pMNS areas to the rest of the brain during guessing are visualized in Figure 6.1. The ventral premotor cortex (PMv) influences large parts of the brain more than the other way around, including all other pMNS areas. This is evidenced in two ways. First, taking the PMv as a seed region (Fig. 6.1, fourth row) shows large areas being more influenced by the PMv than the other way around. Second, the PMv seems to send more information than it receives also when the seed is placed in the middle temporal gyrus (MTG, Fig. 6.1, first row, blue blob) or in the somatosensory cortex of the pMNS (Fig. 6.1, second row, blue blob). Furthermore, all nodes of the pMNS influence the middle temporal gyrus significantly more than the MTG influences them. The dorsal premotor cortex also influences the somatosensory cortex in the parietal node of the pMNS network more than it receives from that area.

### 6.3.3 *G-causality during passive observation*

G-causality originating from pMNS seeds during passive observation shows influences from both the dorsal and ventral premotor cortex to all other pMNS areas (see figure 6.2). The middle temporal gyrus receives significantly more influence from all other pMNS areas than it sends. The parietal node mainly receives from other pMNS areas and sends to part of the MTG.

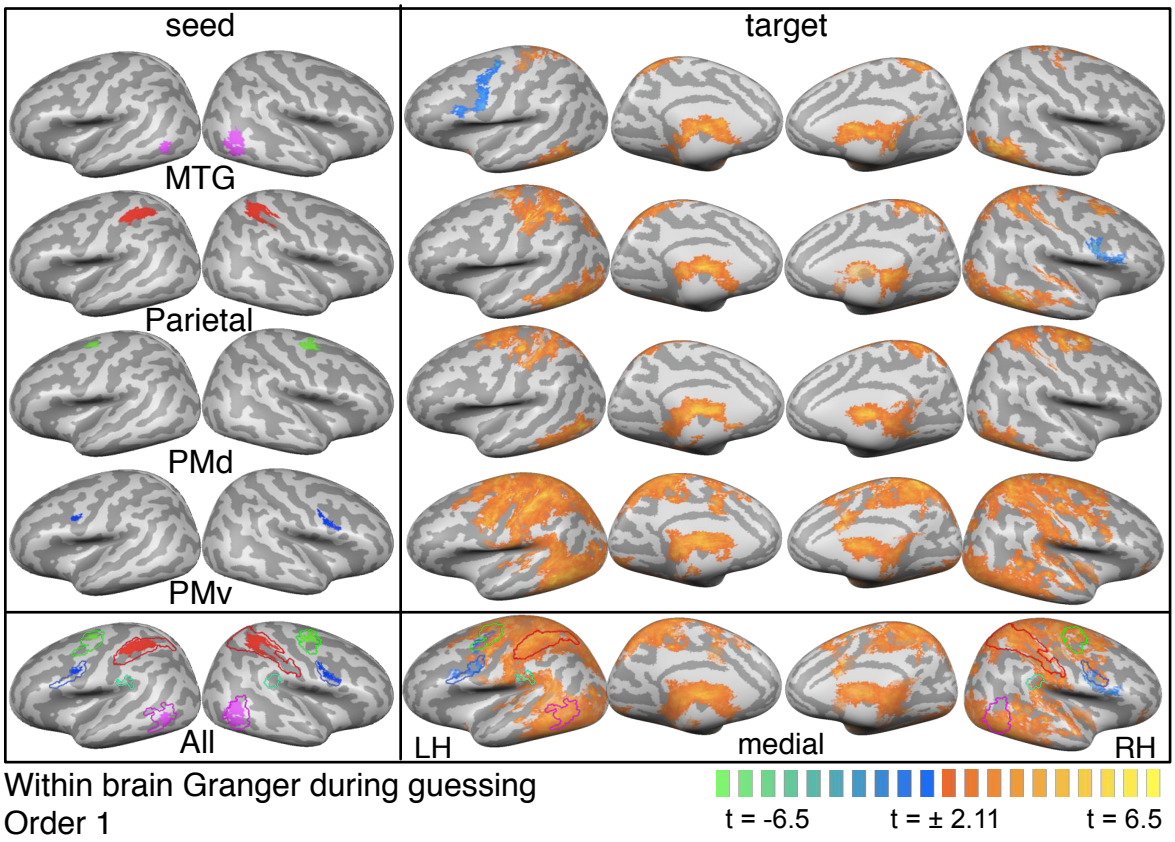



Figure 6.1: Results of second-level G-causality mapping during active guessing. Granger analyses executed separately for the left and right seed are shown together. The left two columns represent the seed regions from which the analyses have been performed. The right side represents the guesser's brain showing t values of the paired t test between seed→target and target→seed G-causality (random effects, n = 18). Warm colors suggest that the seed region (shown on the left) sent more information to these vertices than the other way around; cold colors indicate that the seed region received more information from these vertices than the other way around. Upper four rows: differential G-causality originating from the seeds on the left. Bottom row: on the left a summary of all seeds (solid colors) and on the right G-causality maps with an outline of the pMNS at a less stringent threshold. G-causality maps are statistically thresheld at P<0.05 corrected for multiple comparisons by using a Monte Carlo simulation-based cluster-size threshold adjustment. MTG, middle temporal gyrus; PMd, dorsal premotor cortex; PMv, ventral premotor cortex; LH, left hemisphere; RH, right hemisphere.

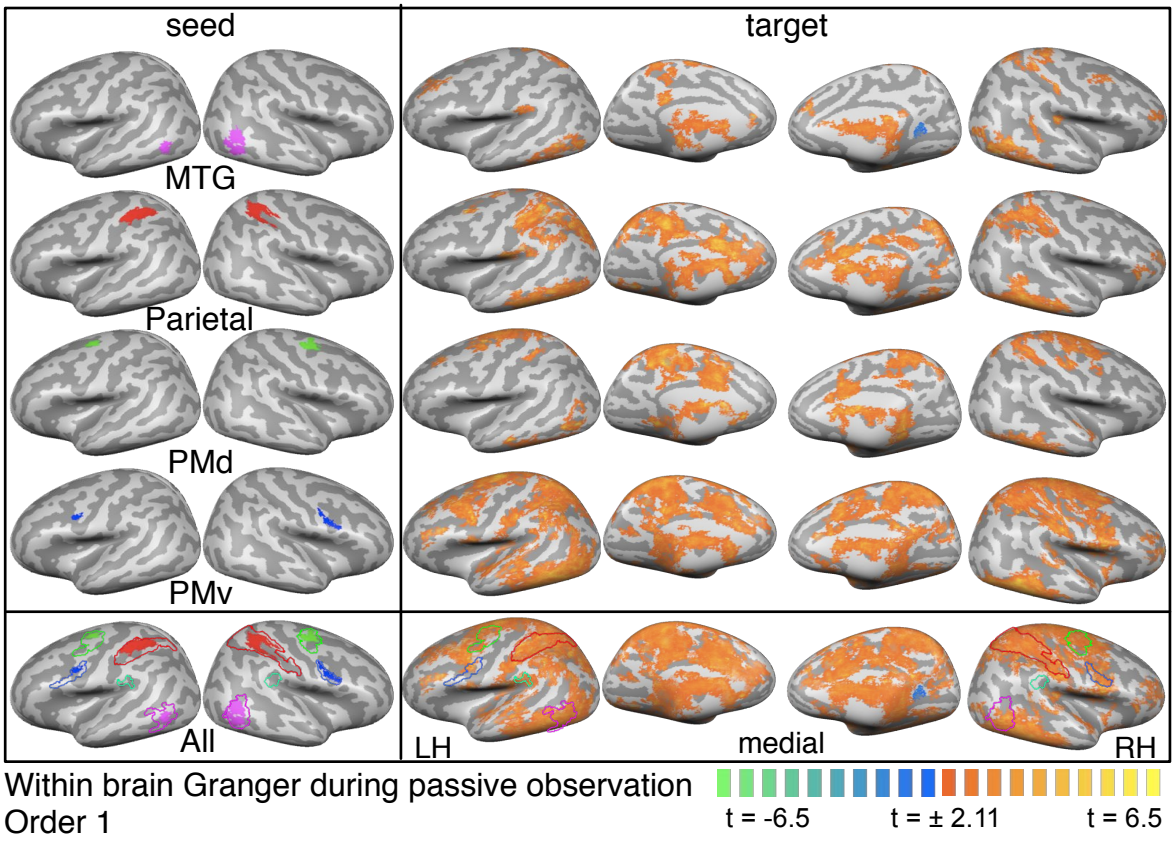



Figure 6.2: Results of second-level G-causality during passive observation. Conventions as in Fig. 6.1

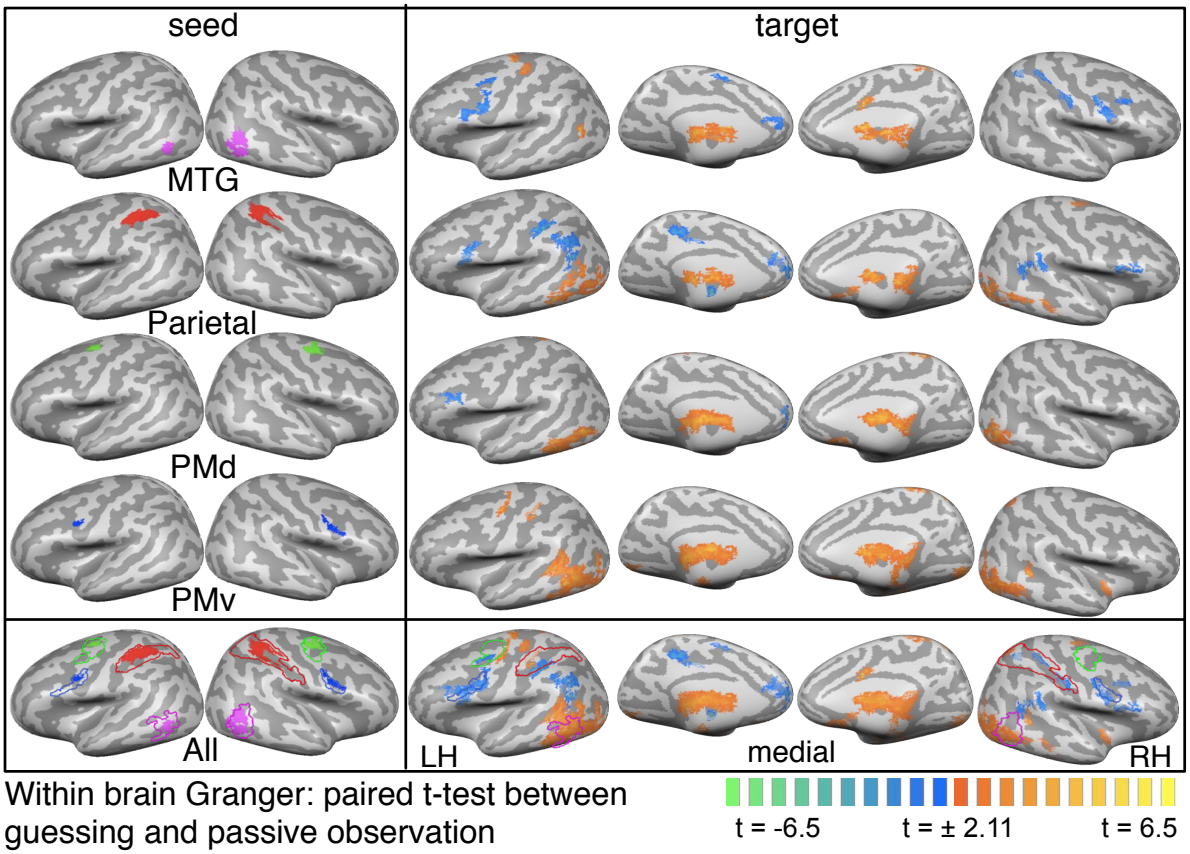



Figure 6.3: Comparison (paired t test) between G-causality during guessing and G-causality during passive observation (random effects, n = 18). Warm colors indicate significantly higher differential Granger values during guessing, thus more net seed → target influence during guessing than during passive observation. Cold colors indicate significantly lower differential Granger values during guessing, thus more target → seed influence during guessing than during passive observation. Further conventions as in Fig. 6.1

### 6.3.4 *Comparing G-causality during guessing with G-causality during passive observation*

Results of the paired t-test between G-causality during guessing and passive observation shows that the posterior flow of information is more dominant during guessing (see Figure 6.3). In particular the ventral premotor cortex sends significantly more to other pMNS areas during guessing than during passive observation. The middle temporal gyrus receives significantly more during guessing than during passive observation. The somatosensory cortex in the parietal pMNS node seems to play a relay role. During active guessing compared to passive viewing, it receives more from ventral premotor cortex than it sends, while it sends more to visual cortices than it receives.

## 6.4 DISCUSSION

In this article we investigated the dominant direction of information flow within the putative Mirror Neuron System (pMNS) and between the pMNS and the rest of the brain during the observation of comparatively long naturalistic stretches of communicative gestures. Letting participants observe these gestures for at least 50s allowed using Granger causality to test the dominant direction of influence. We reasoned that if the pMNS is a strict feed-forward system, the dominant direction of influence should be from visual → premotor cortices. If instead it functions as a dynamic feedback control system, with a combination of inhibitory forward and excitatory inverse models, guessing the meaning of gestures of others should be accompanied by a stronger information flow from premotor regions to more posterior pMNS regions than the other way around. Our results fit the latter. We found a net information flow within the pMNS from premotor to parietal and middle temporal cortices. This is more pronounced during an active guessing task than while passively reviewing the same gestures. In particular, the ventral premotor cortex sends significantly more information to other pMNS areas during active guessing than during passive observation.

This posterior flow of information in the pMNS (e.g. premotor → parietal → MTG) supports the idea that the main function of the pMNS could be a predictive one (Keysers and Perrett, 2004; Kilner et al., 2007b; Miall, 2003). Posterior influences between pMNS areas would thus correspond to a forward model that inhibits expected sensory input on the basis of the motor programs triggered by the inverse model during the observation of the beginning of the action (Blakemore and Frith, 2005; Urgesi et al., 2010; Gazzola and Keysers, 2008; Keysers and Perrett, 2004; Kilner et al., 2007b; Csibra and Gergely, 2007; Miall, 2003).

If the dynamic control theory is right, then the relationship between visual cortices (MTG) and premotor cortices should change during the period of guessing. At first, before the gestures can be predicted, the MTG should send to the premotor cortices more or as much as it receives. As the gestures become more predictable, information flow from premotor cortex to MTG should start to prevail. To examine this possibility, we also performed Granger causality from the MTG seed during the initial 15 volumes corresponding to the relatively earlier phase of guessing. As expected, comparing the analyses during the early and later phase revealed a reversal of dominant information flow. In the early period, a forward information flow dominates, with occipital cortices sending more information to the MTG than it receives, and the MTG sending more information to frontal areas than it receives. In contrast, in the later period, the same MTG region now receives more information from the premotor cortex than it sends to that regions, and information from the occipital lobe becomes less prominent (see Figure 6.4)

In particular when the modulation of connectivity between passive observation and guessing is considered, it can be seen that the inverse model is more dominant during guessing than during passive observation. This would fit with the task requirements, as during the guessing condition participants where actively trying to process what was being gestured while during passive observation this tendency was being inhibited. This data therefore suggests that the posterior information flow from premotor to visual cortices is augmented more than the anterior information flow when participants deliberately decode the actions of others. These connectivity results complement the finding that activity in the pMNS can be influenced in a top-down fashion by instructions given to participants (Engel et al., 2008).

The task during the passive observation condition, to try not to interpret the gestures, is somewhat odd, akin to the proverbial: don't think about a pink elephant. Our participants will probably not have been able to fully refrain from interpreting the gestures they were looking at. However, upon debriefing, participants reported more interpretation during guessing than during passive observation. Our results show a significant modulation of connectivity between these two tasks, which suggests that the instructions indeed changed the way participants processed the movies.

Our results further show that the medial prefrontal cortex (mPFC) changes its connectivity with the MTG, parietal lobule and PMd (see Figure 6.3, row 1 to 3) when the two experimental conditions are compared. Examining results in the two conditions separately reveals that this change is due to the fact that differential G-causality between the pMNS and the mPFC does not differ significantly from zero during active guessing, but does so during passive viewing, where the mPFC receives more information than it sends to various nodes of the pMNS. Given that differential G-causality can have a value of zero either because there is no information flow between two regions or because the information flow is symmetrical, this result should not be in-

terpreted as suggesting that the mPFC and the pMNS are not exchanging information during gesture guessing. Instead they may exchange information on a symmetrical basis. During passive viewing however, this region seems to receive more information than it sends to the pMNS. Given that the mPFC is an area associated with social cognition, in particular reasoning about others (Amodio and Frith, 2006; Brunet et al., 2000; Castelli et al., 2002; Frith and Frith, 2006; Gallagher et al., 2000) and it has been suggested that it might receive information from the mirror neuron system during action observation (Agnew et al., 2007; Blakemore et al., 2004; Decety and Chaminade, 2003; Keysers and Gazzola, 2007; Ohnishi et al., 2004; Uddin et al., 2007) and gesture observation (Schippers et al., 2010) this could add more mentalist interpretations to the observed actions or gestures (Keysers and Gazzola, 2007). If people actively engage in mentalizing, such cognitions might in turn influence the pMNS (Keysers and Gazzola, 2007). One interpretation of our pattern of connectivity during active guessing and passive viewing could therefore be that during passive viewing, information flow dominates from the pMNS → mPFC, because the processes occurring in the pMNS are more automatic than the more deliberate mentalizing triggering mPFC activity. If instructed to guess what word the other person is trying to gesture, however, deliberate mentalizing is encouraged, providing more opportunity for the mPFC → pMNS information flow. This would even out with the more automatic pMNS → mPFC information flow and lead to a zero difference between the directions of information flow. Evidence for the fact that the mPFC indeed processes the gestures more during active guessing comes from our previous study where we show that the mPFC of the guesser does reflect the brain activity of the gesturer more during active guessing than passive viewing (Schippers et al., 2010).

In this experiment, we try to deduce the dominant direction of neural information flow by applying Granger causality to fMRI measured BOLD signal. We therefore base most of our conclusions on the difference in differential G-causality between guessing and passive viewing, according to the recommendation of Roebroeck et al. (2005). One might challenge these conclusions, however, on the basis that seeing the same movies again might have lead to a reduction in visual processing, and therefore to less overall flow of information in the brain. Accordingly, the differences in differential Granger causality we found across conditions could be due to a reduction of information flow in the passive viewing condition, rendering our results more difficult to interpret. It is important therefore, that simulations (Schippers et al., subm) show that if Granger causality detects a directed influence (as it did in this experiment), this direction correctly corresponds to the underlying dominant neural direction of influence in over 80% of the cases even within a single experimental condition. This means that even if we only consider the active guessing condition, the fact that we find the premotor regions to send more information to the visual regions than the other way around based on BOLD, we have an 80% chance, that the actual direction of dominant

neural information flow indeed is in the direction predicted by the dynamic feedback control model of pMNS functioning.

ACKNOWLEDGEMENTS

The research was supported by a VIDI grant of the Dutch science foundation (N.W.O.) and a Marie Curie Excellence Grant of the European Commission to CK.

# BIBLIOGRAPHY


Agnew, Z. K., Bhakoo, K. K., and Puri, B. K. (2007). The human mirror system: a motor resonance theory of mind-reading. *Brain research reviews*, 54(2):286–93.

Amodio, D. M. and Frith, C. D. (2006). Meeting of minds: the medial frontal cortex and social cognition. *Nature Reviews of Neuroscience*, 7(4):268–77.

Blakemore, S. J. and Decety, J. (2001). From the perception of action to the understanding of intention. *Nature Reviews of Neuroscience*, 2(8):561–7.

Blakemore, S. J. and Frith, C. (2005). The role of motor contagion in the prediction of action. *Neuropsychology*, 43(2):260–7.

Blakemore, S. J., Winston, J., and Frith, U. (2004). Social cognitive neuroscience: where are we heading? *Trends in Cognitive Sciences*, 8(5):216–22.

Blakemore, S. J., Wolpert, D. M., and Frith, C. D. (1998). Central cancellation of self-produced tickle sensation. *Nature Neuroscience*, 1(7):635–40.

Brunet, E., Sarfati, Y., Hardy-Baylé, M., and Decety, J. (2000). A pet investigation of the attribution of intentions with a nonverbal task. *Neuroimage*, 11(2):157–166.

Bubic, A., von Cramon, D. Y., and Schubotz, R. I. (2010). Prediction, cognition and the brain. *Frontiers in Human Neuroscience*, 25:1–15.

Caspers, S., Zilles, K., Laird, A. R., and Eickhoff, S. B. (2010). Ale meta-analysis of action observation and imitation in the human brain. *Neuroimage*, 50(3):1148–67.

Castelli, F., Frith, C. D., Happè, F., and Frith, U. (2002). Autism, asperger syndrome and brain mechanisms for the attribution of mental states to animated shapes. *Brain*, 125:1839–1849.

Chong, T. T.-J., Cunnington, R., Williams, M. A., Kanwisher, N., and Mattingley, J. B. (2008). fmri adaptation reveals mirror neurons in human inferior parietal cortex. *Current biology*, 18(20):1576–80.

Csibra, G. and Gergely, G. (2007). 'obsessed with goals': functions and mechanisms of teleological interpretation of actions in humans. *Acta Psychologica*, 124(1):60–78.

Decety, J. and Chaminade, T. (2003). When the self represents the other: A new cognitive neuroscience view on psychological identification. *Conscious and Cognition*, 12:577–596.

Deshpande, G., Sathian, K., and Hu, X. (2009). Effect of hemodynamic variability on granger causality analysis of fmri. *Neuroimage*, 52(3):884–896.

Dinstein, I., Hasson, U., Rubin, N., and Heeger, D. J. (2007). Brain areas selective for both observed and executed movements. *Journal of Neurophysiology*, 98(3):1415–27.

Eickhoff, S., Stephan, K. E., Mohlberg, H., Grefkes, C., Fink, G. R., Amunts, K., and Zilles, K. (2005). A new spm toolbox for combining probabilistic cytoarchitectonic maps and functional imaging data. *Neuroimage*, 25(4):1325–1335.

Engel, A., Burke, M., Fiehler, K., Bien, S., and Rösler, F. (2008). What activates the human mirror neuron system during observation of artificial movements: bottom-up visual features or top-down intentions? *Neuropsychology*, 46(7):2033–42.

Falck-Ytter, T., Gredebäck, G., and von Hofsten, C. (2006). Infants predict other people's action goals. *Nature Neuroscience*, 9(7):878–9.

Filimon, F., Nelson, J. D., Hagler, D. J., and Sereno, M. I. (2007). Human cortical representations for reaching: mirror neurons for execution, observation, and imagery. *Neuroimage*, 37(4):1315–28.

Fischl, B., Sereno, M. I., and Dale, A. M. (1999). Cortical surface-based analysis. ii: Inflation, flattening, and a surface-based coordinate system. *Neuroimage*, 9(2):195–207.

Forman, S. D., Cohen, J. D., Fitzgerald, M., Eddy, W. F., Mintun, M. A., and Noll, D. C. (1995). Improved assessment of significant activation in functional magnetic resonance imaging (fmri): use of a cluster-size threshold. *Magnetic resonance in medicine*, 33(5):636–47.

Frith, C. D. and Frith, U. (1999). Interacting minds–a biological basis. *Science*, 286(5445):1692–1695.

Frith, C. D. and Frith, U. (2006). The neural basis of mentalizing. *Neuron*, 50(4):531–534.

Gallagher, H. L., Happè, F., Brunswick, N., Fletcher, P. C., Frith, U., and Frith, C. D. (2000). Reading the mind in cartoons and stories: An fmri study of 'theory of mind' in verbal and nonverbal tasks. *Neuropsychology*, 38:11–21.

Gallese, V., Fadiga, L., Fogassi, L., and Rizzolatti, G. (1996). Action recognition in the premotor cortex. *Brain*, 119(2):593–609.

Garrod, S. and Pickering, M. (2009). Joint action, interactive alignment, and dialog. *Topics in Cognitive Science*, 1(2):292–304.

Gazzola, V., Aziz-Zadeh, L., and Keysers, C. (2006). Empathy and the somatotopic auditory mirror system in humans. *Current biology*, 16(18):1824–9.

Gazzola, V. and Keysers, C. (2008). The observation and execution of actions share motor and somatosensory voxels in all tested subjects: Single-subject analyses of unsmoothed fmri data. *Cerebral Cortex*, 19(6):1239–1255.

Goebel, R., Roebroeck, A., Kim, D.-S., and Formisano, E. (2003). Investigating directed cortical interactions in time-resolved fmri data using vector autoregressive modeling and granger causality mapping. *Magnetic resonance imaging*, 21:1251–1261.

Greenwald, A. G. (1970). Sensory feedback mechanisms in performance control: with special reference to the ideo-motor mechanism. *Psychological Review*, 77(2):73–99.

Hagler, D. J., Saygin, A. P., and Sereno, M. I. (2006). Smoothing and cluster thresholding for cortical surface-based group analysis of fmri data. *Neuroimage*, 33(4):1093–103.

Handwerker, D. A., Ollinger, J. M., and D'Esposito, M. (2004). Variation of bold hemodynamic responses across subjects and brain regions and their effects on statistical analyses. *Neuroimage*, 21(4):1639–51.

Hesse, M. D., Sparing, R., and Fink, G. R. (2009). End or means–the "what" and "how" of observed intentional actions. *Journal of Cognitive Neuroscience*, 21(4):776–90.

Heyes, C., Ray, E., Peter, J., Slater, J., and Timothy, J. (2000). What is the significance of imitation in animals? *Advances in the Study of Behavior*, Volume 29:215–245.

Hietanen, J. K. and Perrett, D. I. (1993). Motion sensitive cells in the macaque superior temporal polysensory area. i. lack of response to the sight of the animal's own limb movement. *Experimental brain research*, 93(1):117–28.

Hietanen, J. K. and Perrett, D. I. (1996). Motion sensitive cells in the macaque superior temporal polysensory area: response discrimination between self-generated and externally generated pattern motion. *Behavioural brain research*, 76(1-2):155–67.

Holmes, C. J., Hoge, R., Collins, L., Woods, R., Toga, A. W., and Evans, A. C. (1998). Enhancement of mr images using registration for signal averaging. *Journal of computer assisted tomography*, 22(2):324–33.

Jabbi, M. and Keysers, C. (2008). Inferior frontal gyrus activity triggers anterior insula response to emotional facial expressions. *Emotion*, 8(6):775–80.

Keysers, C. (2009). Mirror neurons. *Current Biology*, 19(21):R971–3.

Keysers, C. and Gazzola, V. (2007). Integrating simulation and theory of mind: from self to social cognition. *Trends in Cognitive Sciences*, 11(5):194–6.

Keysers, C. and Gazzola, V. (2009). Expanding the mirror: vicarious activity for actions, emotions, and sensations. *Current Opinion in Neurobiology*, 19(6):666–71.

Keysers, C., Kaas, J. H., and Gazzola, V. (2010). Somatosensation in social perception. *Nature Reviews of Neuroscience*, 11(6):417–28.

Keysers, C., Kohler, E., Umiltà, M., Nanetti, L., Fogassi, L., and Gallese, V. (2003). Audiovisual mirror neurons and action recognition. *Experimental Brain Research*, 153(4):628–636.

Keysers, C. and Perrett, D. (2004). Demystifying social cognition: A hebbian perspective. *Trends in Cognitive Sciences*, 8(11):501–507.

Kilner, J. M., Friston, K. J., and Frith, C. D. (2007a). The mirror-neuron system: a bayesian perspective. *Neuroreport*, 18(6):619–23.

Kilner, J. M., Friston, K. J., and Frith, C. D. (2007b). Predictive coding: an account of the mirror neuron system. *Cognitive Processes*, 8(3):159–166.

Kokal, I., Gazzola, V., and Keysers, C. (2009). Acting together in and beyond the mirror neuron system. *Neuroimage*, 47(4):2046–56.

Kraskov, A., Dancause, N., Quallo, M. M., Shepherd, S., and Lemon, R. N. (2009). Corticospinal neurons in macaque ventral premotor cortex with mirror properties: a potential mechanism for action suppression? *Neuron*, 64(6):922–30.

Lamm, C., Fischer, M. H., and Decety, J. (2007). Predicting the actions of others taps into one's own somatosensory representations–a functional mri study. *Neuropsychology*, 45(11):2480–91.

Miall, R. C. (2003). Connecting mirror neurons and forward models. *Neuroreport*, 14(17):2135–7.

Ohnishi, T., Moriguchi, Y., Matsuda, H., Mori, T., Hirakata, M., Imabayashi, E., Hirao, K., Nemoto, K., Kaga, M., Inagaki, M., Yamada, M., and Uno, A. (2004). The neural network for the mirror system and mentalizing in normally developed children: an fmri study. *Neuroreport*, 15(9):1483–7.

Oldfield, R. (1971). The assessment and analysis of handedness: the edinburgh inventory. *Neuropsychology*, 9(1):97–113.

Roebroeck, A., Formisano, E., and Goebel, R. (2005). Mapping directed influence over the brain using granger causality. *Neuroimage*, 25:230–242.

Schippers, M., Renken, R., and Keysers, C. (subm). The effect of intra- and inter-subject variability of hemodynamic responses on group level granger causality analyses. *Neuroimage*.

Schippers, M. B., Gazzola, V., Goebel, R., and Keysers, C. (2009). Playing charades in the fmri: are mirror and/or mentalizing areas involved in gestural communication? *PLoS ONE*, 4(8):e6801.

Schippers, M. B., Roebroeck, A., Renken, R., Nanetti, L., and Keysers, C. (2010). Mapping the information flow from one brain to another during gestural communication. *Proceedings of the National Academy of Sciences of the United States of America*, 107(20):9388–93.

Schubotz, R. I. (2007). Prediction of external events with our motor system: towards a new framework. *Trends in Cognitive Sciences*, 11(5):211–8.

Sebanz, N., Bekkering, H., and Knoblich, G. (2006). Joint action: bodies and minds moving together. *Trends in Cognitive Sciences*, 10(2):70–6.

Sebanz, N. and Knoblich, G. (2009). Prediction in joint action: What, when, and where. *Topics in Cognitive Science*, 1(2):353–367.

Uddin, L. Q., Iacoboni, M., Lange, C., and Keenan, J. P. (2007). The self and social cognition: the role of cortical midline structures and mirror neurons. *Trends in Cognitive Sciences*, 11(4):153–7.

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

Urgesi, C., Maieron, M., Avenanti, A., Tidoni, E., Fabbro, F., and Aglioti, S. M. (2010). Simulating the future of actions in the human corticospinal system. *Cerebral Cortex Advance Access*. DOI 10.1093/cercor/bhp292.

Voss, M., Ingram, J. N., Haggard, P., and Wolpert, D. M. (2006). Sensorimotor attenuation by central motor command signals in the absence of movement. *Nature Neuroscience*, 9(1):26–7.

Wolpert, D. M., Doya, K., and Kawato, M. (2003). A unifying computational framework for motor control and social interaction. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences*, 358(1431):593–602.

## SUPPLEMENTARY METHODS

### PARTICIPANTS

Twelve couples (total: 24 participants) were scanned while playing the game charades. We excluded participants who had moved more than the voxel size during the gesturing phase. All of the analyses in this paper were performed on the remaining 18 participants. The mean age of the participants was 27.5 ± 3.8 years. Each couple consisted of a man and a woman involved in a romantic relationship for at least 6 months. Our aim was not to study specifically romantic processes, but simply to let participants feel as comfortable as possible during the game. Participants were prescreened to exclude those with a history of neurological or psychiatric illness. Participants were also asked not to drink coffee before scanning commenced. The participants freely consented to participating in the study by signing an informed consent form and were scaled for their right-handedness on the Edinburgh Righthandedness scale (Oldfield, 1971). This entire study was approved by the Medical Ethics Committee of the University Medical Center Groningen (2007/080).

### TASK / EXPERIMENTAL DESIGN

The experiment consisted of two separate sessions on different days. In the first session, the couple was required to play the game of charades. In the second, detailed anatomical scans and a passive observation control condition were acquired. For the game of charades, participants took turns going into the scanner, alternating gesturing and guessing of words. Words were either objects (for example nutcracker, watch, pencil sharpener) or actions (for example painting, knitting, shaving). Each participant performed two gesture and two guess runs in which they gestured 14 words and guessed 14 words in total (7 per run). The set of words used was the same for each couple, but word order was randomized between participants. After the last gesture-session, a T1-weighted anatomical image was acquired.

*Gesture run*: during a gesture run, the participant was presented with a word on the screen and was instructed to communicate this word to his or her partner by means of gestures. Every word had to be gestured for 90 seconds. Prior to scanning participants were trained not to repeat the same gesture over and over again, but to keep generating new gestures to provide their partner with multiple sources of information. The participant could see how much time he/she needed to keep gesturing by a progress

bar on the screen. A fixation cross was presented for 20 s after each word, which served as our baseline. The gestures were recorded from the control room of the MR-scanner with a video camera (Sony DSR-PDX10P). After the participant had gestured seven words, he/she was taken out of the scanner and went into the waiting room, while his/her partner went into the scanner to guess what he/she had gestured. During this changeover, the experimenter cut the recording of the gestures into movies of 90s in which the participant gestured a word. To ensure that the movies were cut at exactly the moment the word was presented to the gesturing participant, the stimulus computer's sound card emitted a sound at the beginning of word presentation. The output of the sound card was connected to the audio input of the video camera, thus allowing the auditory signal to serve as a marker for cutting. To minimize the amount of head motion in the participants, the upper arms of the participant were fixed to the bed by means of a Velcro strap band. This left the participant free to gesture with his lower arms and fingers, which still allowed 86% percent correct gesture recognition.

*Guess run*: during a guess run, the participant was shown the movies that were recorded in the gesture run of their partner. The task they had to perform was to guess what their partner was trying to gesture to them. Participants were asked to consider the gestures for at least 50 seconds before committing to a specific interpretation of the gestures. This was done to ensure at least 50 seconds of data in each trial to examine the time course of activity using Granger causality. This was done by showing a progress bar below the movie, changing from red to green after 50 seconds, indicating the beginning of the period (50-90s post stimulus onset) during which participants could decide on their interpretation of the gestures, whenever they felt confident, by pressing a button on their 4-button button-box, triggering the appearance of a multiple choice screen. In the multiple-choice menu they had to choose the correct word from five alternatives. One of the alternatives was always 'none of the above' and the correct answer was always present in the multiple-choice menu. The correct answer was never the option 'none of the above'. This marked the end of a trial. Two consecutive trials were separated by 20s of a white fixation cross against a black background, which served as our baseline.

*Passive observation run*: As a control condition for the guess run, the participants watched the movies they had seen during the guessing condition again. This time, they were instructed not to guess what was gestured, but only to passively view them. We are aware that such instructions cannot ensure that participants entirely stopped to interpret the gestures, but at debriefing, participants reported having interpreted the gestures at least less consistently than during the guess run. To keep the run exactly the same as the original guess run, the movie stopped at the moment the participant during the original run had pushed the button. The same multiple-choice menu then appeared and the participant had to answer again. This time, however, they had to select the word written in green letters. The green word was the correct

answer. A fixation cross was presented between two consecutive trials for 20 seconds and served as our baseline.

### DATA ACQUISITION

Functional imaging data was recorded with a Philips 3.0T MR scanner, using gradient echo planar imaging (EPI) and an 8-channel head coil using SENSE technology. T2* weighted images revealed changes in blood oxygen level. Volume repetition time (TR) was 1.33 seconds. The whole brain was scanned in 28 (axial) slices with a thickness of 4.5mm. Further imaging parameters include echo time (TE) 28 ms, field of view 224 x 224 mm, 64 x 62 matrix, SENSE acceleration factor 2.4, ensuing 3D voxel size 3.5 x 3.5 x 4.5mm. This set of imaging parameters were chosen to cover the entire neo-cortex while at the same time providing a TR short enough to expect sufficient power in a Granger Causality analysis (Roebroeck et al., 2005). In the first session, a fast structural image ("fast anatomy") was acquired of the participant's brain, while in the second session an additional structural image of higher resolution was acquired. Both were structural, T1-weighted images acquired with a T1TFE sequence (echo time 3.5 ms, repetition time 7.6 ms, 224 x 160 x 256 matrix, 1 x 1 x 1 mm3 voxels).

### DATA PRE-PROCESSING

All analyses and preprocessing was performed in BrainVoyager QX 1.10 along with custom written C++ code for the Granger causality analyses. The pre-processing steps included slice scan time correction, 3D motion correction and temporal filtering (consisting of linear trend removal and a high pass filter with a cut-off at 0.004Hz). The images were not smoothed spatially. Functional images were co registered with the structural images and morphed into Talairach space. The structural images were corrected for inhomogeneity (to improve segmentation results), normalized into Talairach space after which the cortical grey matter/white matter boundary was segmented into a topologically correct surface representation. After segmentation, the cortical surface representations of all subjects were aligned using a cortical curvature based alignment procedure. This procedure aligns the sulci and the gyri of the different brains using their cortical curvature-maps (Fischl et al., 1999). Included in the cortex-based alignment was also the Colin (27) brain (Holmes et al., 1998). The fMRI time courses were resampled on the curvature aligned cortical surface representations.

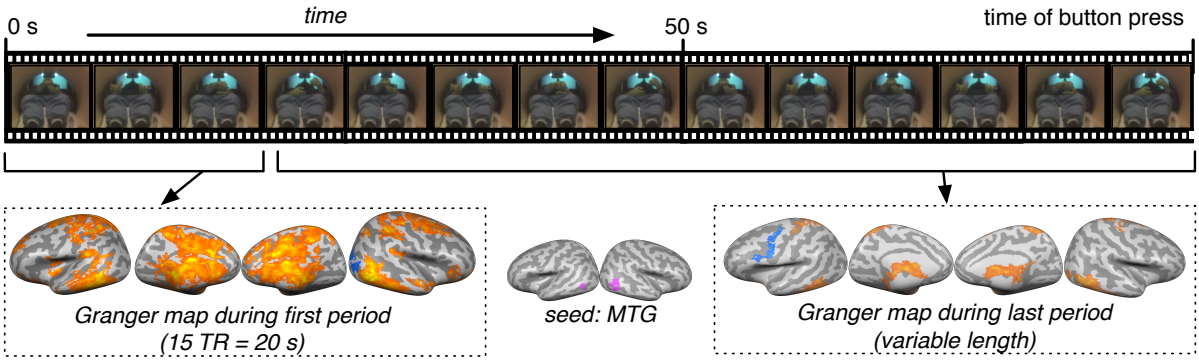



Figure 6.4: Comparison between Granger causality results between the first period of guessing (left in the figure, starting at the first volume until the 15th volume) and the last period of guessing (right in the figure, starting at the 16th volume until the volume at the time of button press). The seed region is the middle temporal gyrus.

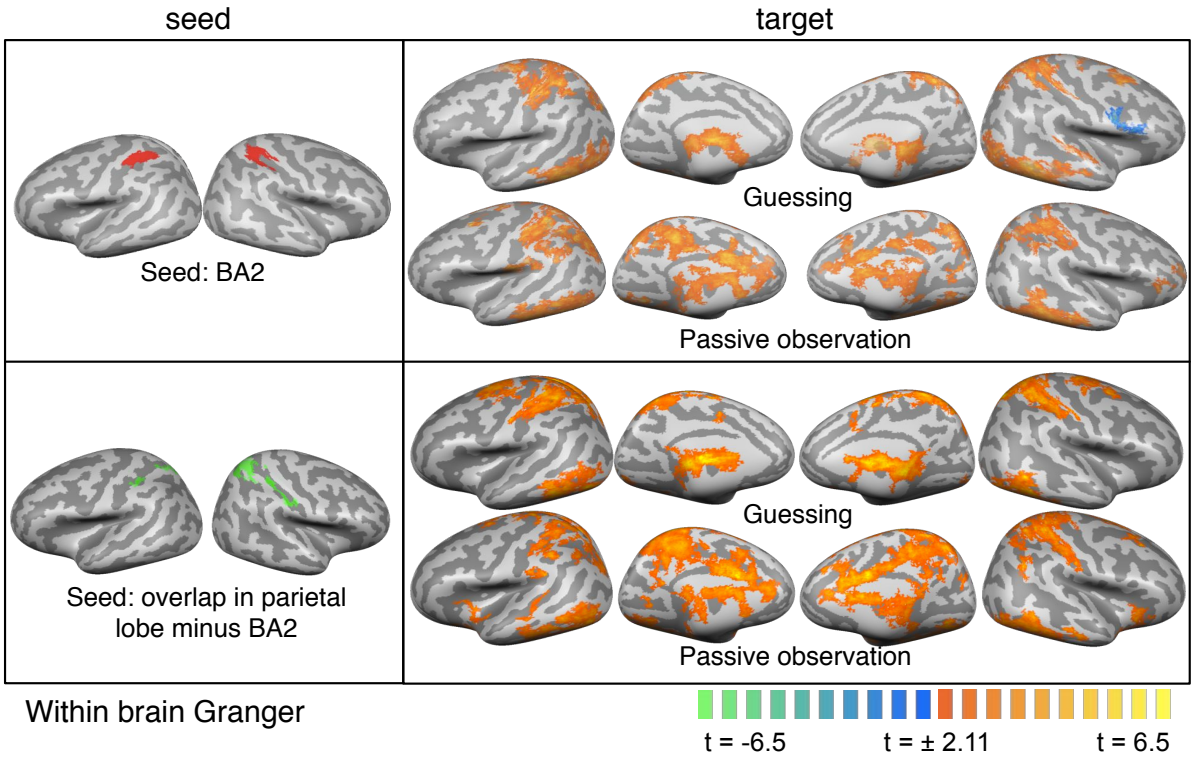



Figure 6.5: Comparison between Granger causality results of two seed regions in the parietal lobe during guessing and passive observation. The upper row shows results for the BA2 seed region (which is used in the main article). The bottom row shows results for a seed region that is composed of the overlap between gesturing and guessing in the random effects analyses minus the vertices comprising BA2.

# 7

# CONCLUDING REMARKS

Social neuroscience is essentially interested in how the human brain works in situations in which at least two brains are present. Even so, most studies have measured brain activity of only one person in a social situation. The current thesis describes results of the charades experiment in which brain activity of two persons was measured pseudo-simultaneously. Participants engaged in an actual game of charades in which they took turns as a sender (gesturing) or receiver (guessing) of gestures. The experiment was set up in a way that participants very naturally generated and interpreted communicative gestures. This concluding chapter shortly summarizes the results and describes implications of this research.

## 7.1 IMPLICATIONS FOR THE MIRROR SYSTEM

The putative Mirror Neuron System (pMNS) is a network of areas in the human brain that have the property of being active both when an action is executed and when it is observed (Buccino et al., 2001; Grafton et al., 1996; Grèzes et al., 1998; Grèzes and Decety, 2001; Grèzes et al., 2003; Nishitani and Hari, 2000, 2002; Perani et al., 2001; Gazzola et al., 2007b,a; Gazzola and Keysers, 2008). In Chapter 3, we studied the involvement of pMNS areas in gestural communication using a whole brain analysis in which we masked activity during guessing with the activity during gesturing. This showed that brain regions associated with the pMNS for goal-directed, transitive actions were recruited during gestural communication. Furthermore, we investigated the involvement of pMNS areas through a region-of-interest analysis, localizing pMNS areas with a mask of a different study (Gazzola et al., 2007a). Combining the study of Gazzola et al. (2007a) with the results of the current study show that the same set of voxels in the brain is involved in (a) mapping the object-directed hand actions of others onto the neural substrates involved in executing similar object-directed hand actions and (b) mapping the gestures of others onto the neural substrates involved in executing similar gestures.

In Chapter 4 we quantified for the first time connectivity between brains using fMRI and explicitly tested the resonance property for pMNS areas put forward by simulation theory (Gallese and Goldman, 1998; Gallese et al., 2004; Rizzolatti et al., 2001). Results of our between brains Granger causality mapping indicated that activity of pMNS regions of the guesser indeed show a significant Granger causal relation with pMNS

regions of the gesturer. This means that pMNS regions in the guesser on a moment-by-moment basis track the activity of the pMNS regions of the gesturer and in this sense resonate with the pMNS activity in the gesturer. Importantly, we performed these analyses not on the full time series of the gesturer and guesser, but excluded the beginning and end of each interaction. This ensures that the Granger causal relations between the pMNS areas of the gesturer and the guesser are not dominated by the relatively trivial fact that pMNS regions in the observer will 'turn on' after these regions were turned on in the gesturer, but that even during an ongoing interaction, the observer's pMNS continues to track the ups and downs of the gesturer's pMNS (and therefore, motor system).

In Chapter 6, we gained further insights in the function of the pMNS by investigating the flow of information from and to pMNS areas during gesture observation. We compared hypotheses generated by two different models of the mirror system. Our results were most compatible with the conceptualization of the pMNS as a dynamic control system. During the relatively long gesture observation blocks (at least 50s), the dominant direction of information flow was posterior (premotor → parietal → temporal). In our opinion this indicates a predictive function of pMNS areas. When we start observing a sequence of actions, the anterior flow of information briefly dominates. As soon as the first motor programs are activated, however, forward models would start predicting the future visual input in a posterior flow of information from premotor to visual areas. Because this posterior flow is inhibitory (Keysers and Perrett, 2004; Hietanen and Perrett, 1993, 1996) and the visual sequence of gestures is relatively predictable, visual input is increasingly inhibited, and the anterior flow of information is decreased.

## 7.2 IMPLICATIONS FOR MENTALIZING

The ventral medial prefrontal cortex (vmPFC) is known to play a role in social situations, in particular in situations in which reasoning about other people's mental states is required (Amodio and Frith, 2006; Frith and Frith, 2006). For a long time, the field of social neuroscience had been divided in those that believed that the pMNS was the most important mean towards understanding others and those that thought that mentalizing regions like the vmPFC plaid the lead role. It was recently proposed, however, that rather than debating which system, pMNS or vmPFC, is more important, we need to start thinking of how they interact to jointly enable human social understanding (Keysers and Gazzola, 2007, 2009).

In Chapter 3, we used a classical GLM analysis and found that pMNS areas are consistently activated above baseline during gesture observation. The vmPFC, in contrast, does not show an involvement in gesture observation, when such an involvement is

defined as 'above baseline activity', neither when analyzed with a whole-brain analysis nor when analyzed with a region-of-interest analysis. The between brains Granger causality analysis in Chapter 4, however, revealed directed influences between pMNS areas of the gesturer and the vmPFC of the guesser. These influences were present from gesturer to guesser, but not from gesturer to passive observer. Finally, using a within brain connectivity analysis in Chapter 6, we demonstrated that the pMNS sends information to the vmPFC during passive observation, while no differential Granger causality influence was found between pMNS and vmPFC during active guessing. A zero finding of differential Granger causality can indicate either no information flow at all, or a symmetrical information flow. Since our between brains analysis indicated a relationship between pMNS areas of the gesturer and the vmPFC of the guesser, it seems unlikely that no information flow between the pMNS areas of the guesser and the vmPFC of the guesser is present. Therefore, we interpret this finding as a symmetrical flow of information between pMNS areas and the vmPFC.

Our between brains Granger causality method could reveal tight relationships between pMNS areas and the vmPFC in our data, while a classic GLM analysis of the same data, as presented in Chapter 3, did not show simple on/off activation of the vmPFC of the guesser. This apparent incongruency between results obtained using the same data but different methods might be partly attributed to the fact that we used a passive baseline in the GLM analyses. A passive baseline (e.g. a fixation cross) is not as passive as it might seem. A network of areas (the 'default' network) has been found to augment their activity during baseline when compared to many different types of tasks (Raichle et al., 2001; Gusnard et al., 2001). In effect, activity in these areas is suspended during task performance, leading to a decrease of activity. As the vmPFC is considered part of the default network (Gusnard et al., 2001) this might have masked our results.However, even though the vmPFC might not seem to be involved in guessing communicative gestures when analyzed with a classical GLM, activity in the vmPFC of the guesser does contain information about the time course of activity in the regions involved in planning and executing gestures in the gesturer (see Supplementary Information 2 of Chapter 4).

The vmPFC is typically involved in inferential processes (Keysers and Gazzola, 2007; van der Meer et al., 2010; Overwalle and Baetens, 2009), while the pMNS is engaged by more direct perceptual processes. This difference in types of processes in the pMNS and the vmPFC could also provide an interpretation of the pattern of connectivity within a brain during active guessing and passive viewing (see Chapter 6). During passive viewing, information flow could dominate from the pMNS → vmPFC, because the processes occurring in the pMNS are more automatic than the more inferential vmPFC activity. If instructed to guess what word the other person is trying to gesture, however, deliberate mentalizing is encouraged, providing more opportunity for the vmPFC to process the gestures and hence to send information in the vmPFC → pMNS

information flow. This would even out with the more automatic pMNS → vmPFC information flow that dominates during passive viewing and lead to a zero difference between the directions of information flow.

## 7.3 IMPLICATIONS FOR GRANGER CAUSALITY

Using Granger causality analyses to investigate directed influences within and between brains might make the results vulnerable to the alleged confound caused by variability in hemodynamic delay (David et al., 2008; de Marco et al., 2009; Roebroeck et al., 2005; Friston, 2009; Chang et al., 2008). This has been investigated on a single subject level by Deshpande et al. (2009), but the effect on group level has been in great need of investigation. We took this opportunity and performed simulations in which we systematically varied (1) the directed influence between two time series, (2) the neuronal delay between them, (3) the hemodynamic responses with which the time series were convolved and (4) the level of noise in the signal. The results, described in Chapter 5, indicate that when no influence is present between two brain areas, there is only a marginal chance (< 5%) to report significant Granger causality results (false positives). Furthermore, when an influence is present and the neuronal delay between two areas is longer than 1 second, no inversions are observed, even when the difference in hemodynamic delays goes against the direction of influence (thus with a longer hemodynamic delay for the source than for the target). Additionally, the sensitivity of Granger causality group analyses at this length of neuronal delay is adequate (>80% for a realistic noise level). Granger causality is therefore an adequate method to investigate influences in a *between* brains context, in which neuronal delays are in the order of seconds. Our simulations further indicate that the sensitivity of Granger causality in a within brain situation is rather limited with short neuronal delays. However, they also show that if a significant Granger causal relation is found between two brain regions, the direction of this reported influence is correct in ~80% of the cases. Granger causality on a group level is thus less vulnerable for hemodynamic variability than has often been assumed. Our results therefore argue that Granger causality is a valid method to investigate connectivity within a brain.

## 7.4 GENERAL IMPLICATIONS

The research described in this thesis underlines three increasingly important aspects of social cognitive neuroscience: naturalistic experimental situations, the use of data-driven analysis methods, and quantifying relations *between* brains.

The importance of creating as naturalistic situations as possible in the MR-scanner is increasingly appreciated (Redcay et al., 2010; Hari and Kujala, 2009; Mathiak and

Weber, 2006; Hasson et al., 2004). For example, stimuli nowadays often consist of video recordings of social situations, rather than static images (see e.g. van der Gaag et al., 2007; Gazzola et al., 2007a; LaBar et al., 2003). Furthermore, the more constraints are imposed on a social situation, the less naturalistic it might feel for the participants. In a seminal study, Hasson et al. (2004) did not constrain his participants at all, but let them freely watch a Hollywood movie, similar to how they would watch a movie at home. The current research extends this type of free-form experiments by letting participants engage in the actual social game charades. We did not impose a strict temporal structure on the timing of the experiment. Gestures could be made freely with hardly any constraints: no restrictions on the amount of hand movements or eye gaze, no fixed number of repetitions for one word and no separation in planning and execution phase.

The use of such an unconstrained social interaction as experimental design, however, has implications for the data analysis. When a strict temporal structure is imposed on the experimental setting, a classical GLM analysis can be used. This method hinges on the knowledge of the timing of cognitive processes. We used this method, for example, to investigate the general involvement of pMNS and mentalizing areas in gestural communication (see Chapter 3). To quantify influences between brains during naturalistic interactions, a more exact timing of cognitive processes would be necessary, other than the on- and offset of the video recordings. This exact timing, however, is unknown in our experiment in which social interaction was deliberately left unconstrained. A quantification of the stream of gestures would be necessary, but this is far from trivial: it is unknown what aspects of the highly multidimensional videorecording are actually relevant for the brain activity of the guesser, making the generation of predictors for a GLM arbitrary. Instead, we used between brains Granger causality, which has the elegant property of circumventing this problem altogether by using brain activity of the gesturer to quantify brain activity of the guesser (see Chapter 4). Between brains Granger causality is thereby capable of directly testing those theories, like the pMNS 'resonance' theory (Gallese and Goldman, 1998; Rizzolatti and Arbib, 1998), which are formulated not as a link between a stimulus and a neural state, but between the neural states of two individuals.

Finally, quantification of influences between brains are of increasing interest in social cognitive neuroscience. This started with the development of techniques to measure two brains simultaneously. The research group of Montague developed, for example, the 'hyperscanning' technique with which two scanners could be coupled during an experiment (Montague et al., 2002). Babiloni et al. (2006) instead measured EEG simultaneously from different participants and showed that causal relations exist between prefrontal cortices during a cooperative card-game. Here, we have introduced between brains Granger causality mapping as a data analysis approach for fMRI data to quantify influences and information flow between brains. This thesis is,

to our knowledge, one of the very, if not the first to describe this type of connectivity between brains. We hope our work helped paving the way to answer exciting research questions such as how the connection between brains is modulated in people with a social deficit, such as autism spectrum disorder.

# BIBLIOGRAPHY


Amodio, D. M. and Frith, C. D. (2006). Meeting of minds: the medial frontal cortex and social cognition. *Nature Reviews of Neuroscience*, 7(4):268–77.

Babiloni, F., Cincotti, F., Mattia, D., Mattiocco, M., Fallani, F. D. V., Tocci, A., Bianchi, L., Marciani, M. G., and Astolfi, L. (2006). Hypermethods for eeg hyperscanning. *Conference Proceedings: Annual International Conference of the IEEE Engineering in Medicine and Biology Society IEEE Engineering in Medicine and Biology Society Conference*, 1:3666–9.

Buccino, G., Binkofski, F., Fink, G. R., Fadiga, L., Fogassi, L., Gallese, V., Seitz, R. J., Zilles, K., Rizzolatti, G., and Freund, H.-J. (2001). Action observation activates premotor and parietal areas in a somatosopic manner: An fmri study. *The European Journal of Neuroscience*, 13:400–404.

Chang, C., Thomason, M. E., and Glover, G. H. (2008). Mapping and correction of vascular hemodynamic latency in the bold signal. *Neuroimage*, 43(1):90–102.

David, O., Guillemain, I., Saillet, S., Reyt, S., Deransart, C., Segebarth, C., and Depaulis, A. (2008). Identifying neural drivers with functional mri: an electrophysiological validation. *PLoS Biology*, 6(12):2683–97.

de Marco, G., Devauchelle, B., and Berquin, P. (2009). Brain functional modeling, what do we measure with fmri data? *Neuroscience Research*, 64(1):12–9.

Deshpande, G., Sathian, K., and Hu, X. (2009). Effect of hemodynamic variability on granger causality analysis of fmri. *Neuroimage*, 52(3):884–896.

Friston, K. (2009). Causal modelling and brain connectivity in functional magnetic resonance imaging. *PLoS Biology*, 7(2):e33.

Frith, C. D. and Frith, U. (2006). The neural basis of mentalizing. *Neuron*, 50(4):531–534.

Gallese, V. and Goldman, A. (1998). Mirror neurons and the simulation theory of mind-reading. *Trends in Cognitive Sciences*, 12:493–501.

Gallese, V., Keysers, C., and Rizzolatti, G. (2004). A unifying view of the basis of social cognition. *Trends in Cognitive Sciences*, 8(9):396–403.

Gazzola, V. and Keysers, C. (2008). The observation and execution of actions share motor and somatosensory voxels in all tested subjects: Single-subject analyses of unsmoothed fmri data. *Cerebral Cortex*, 19(6):1239–1255.

Gazzola, V., Rizzolatti, G., Wicker, B., and Keysers, C. (2007a). The anthropomorphic brain: The mirror neuron system responds to human and robotic actions. *Neuroimage*, 35:1674–1684.

Gazzola, V., van der Worp, H., Mulder, T., Wicker, B., Rizzolatti, G., and Keysers, C. (2007b). Aplasics born without hands mirror the goal of hand actions with their feet. *Current biology*, 17(14):1235–40.

Grafton, S. T., Arbib, M. A., Fadiga, L., and Rizzolatti, G. (1996). Localization of grasp representations in humans by positron emission tomography. 2. observation compared with imagination. *Experimental Brain Research*, 112(1):103–111.

Grèzes, J., Armony, J., Rowe, J., and Passingham, R. E. (2003). Activations related to mirror and canonical neurones in the human brain: An fmri study. *Neuroimage*, 18:928–937.

Grèzes, J., Costes, N., and Decety, J. (1998). Top-down effect of strategy on the perception of human biological motion: A pet investigation. *Cognitive Neuropsychology*, 15:553–582.

Grèzes, J. and Decety, J. (2001). Functional anatomy of execution, mental simulation, observation, and verb generation of actions: a meta-analysis. *Human Brain Mapping*, 12(1):1–19.

Gusnard, D. A., Akbudak, E., Shulman, G. L., and Raichle, M. E. (2001). Medial prefrontal cortex and self-referential mental activity: relation to a default mode of brain function. *Proceedings of the National Academy of Sciences of the United States of America*, 98(7):4259–64.

Hari, R. and Kujala, M. V. (2009). Brain basis of human social interaction: from concepts to brain imaging. *Physiological Reviews*, 89(2):453–79.

Hasson, U., Nir, Y., Levy, I., Fuhrmann, G., and Malach, R. (2004). Intersubject synchronization of cortical activity during natural vision. *Science*, 303(5664):1634–40.

Hietanen, J. K. and Perrett, D. I. (1993). Motion sensitive cells in the macaque superior temporal polysensory area. i. lack of response to the sight of the animal's own limb movement. *Experimental brain research*, 93(1):117–28.

Hietanen, J. K. and Perrett, D. I. (1996). Motion sensitive cells in the macaque superior temporal polysensory area: response discrimination between self-generated and externally generated pattern motion. *Behavioural brain research*, 76(1-2):155–67.

Keysers, C. and Gazzola, V. (2007). Integrating simulation and theory of mind: from self to social cognition. *Trends in Cognitive Sciences*, 11(5):194–6.

Keysers, C. and Gazzola, V. (2009). Expanding the mirror: vicarious activity for actions, emotions, and sensations. *Current Opinion in Neurobiology*, 19(6):666–71.

Keysers, C. and Perrett, D. (2004). Demystifying social cognition: A hebbian perspective. *Trends in Cognitive Sciences*, 8(11):501–507.

LaBar, K. S., Crupain, M. J., Voyvodic, J. T., and McCarthy, G. (2003). Dynamic perception of facial affect and identity in the human brain. *Cerebral Cortex*, 13(10):1023–33.

Mathiak, K. and Weber, R. (2006). Toward brain correlates of natural behavior: fmri during violent video games. *Human brain mapping*, 27(12):948–56.

Montague, P. R., Berns, G., Cohen, J., McClure, S., Pagnoni, G., Dhamala, M., Wiest, M., Karpov, I., King, R., Apple, N., and Fisher, R. (2002). Hyperscanning: simultaneous fmri during linked social interactions. *Neuroimage*, 16(4):1159–1164.

Nishitani, N. and Hari, R. (2000). Temporal dynamics of cortical representation for action. *Proceedings of the National Academy of Sciences of the United States of America*, 97(2):913–8.

Nishitani, N. and Hari, R. (2002). Viewing lip forms: cortical dynamics. *Neuron*, 36(6):1211–20.

Overwalle, F. V. and Baetens, K. (2009). Understanding others' actions and goals by mirror and mentalizing systems: a meta-analysis. *Neuroimage*, 48(3):564–84.

Perani, D., Fazio, F., Borghese, N., Tettamanti, M., Ferrari, S., Decety, J., and Gilardi, M. (2001). Different brain correlates for watching real and virtual hand actions. *Neuroimage*, 14(3):749–758.

Raichle, M. E., MacLeod, A. M., Snyder, A. Z., Powers, W. J., Gusnard, D. A., and Shulman, G. L. (2001). A default mode of brain function. *Proceedings of the National Academy of Sciences of the United States of America*, 98(2):676–82.

Redcay, E., Dodell-Feder, D., Pearrow, M. J., Mavros, P. L., Kleiner, M., Gabrieli, J. D. E., and Saxe, R. (2010). Live face-to-face interaction during fmri: a new tool for social cognitive neuroscience. *Neuroimage*, 50(4):1639–47.

Rizzolatti, G. and Arbib, M. A. (1998). Language within our grasp. *Trends in Neurosciences*, 21(5):188–194.

Rizzolatti, G., Fogassi, L., and Gallese, V. (2001). Neurophysiological mechanisms underlying the understanding and imitation of action. *Nature Reviews of Neuroscience*, 2(9):661–70.

Roebroeck, A., Formisano, E., and Goebel, R. (2005). Mapping directed influence over the brain using granger causality. *Neuroimage*, 25:230–242.

van der Gaag, C., Minderaa, R. B., and Keysers, C. (2007). Facial expressions: what the mirror neuron system can and cannot tell us. *Social Neuroscience*, 2(3-4):179–222.

van der Meer, L., Costafreda, S., Aleman, A., and David, A. S. (2010). Self-reflection and the brain: a theoretical review and meta-analysis of neuroimaging studies with implications for schizophrenia. *Neuroscience and Biobehavioral Reviews*, 34(6):935–46.

# ABBREVIATIONS

BA - Brodmann Area

BBGCM - Between Brains Granger Causality Mapping

BOLD - Blood Oxygen Level Dependent

EPI - Echo Planar Imaging

FDR - False Discovery Rate

FIR - Finite Impuls Response

FMRI - Functional Magnetic Resonance Imaging

G-CAUSALITY - Granger Causality

GCM - Granger Causality Mapping

GLM - General Linear Model

HRF - Hemodynamic Response Function

MPFC - Medial Prefrontal Cortex

MRI - Magnetic Resonance Imaging

OP - Operculum

PMNS - Putative Mirror Neuron System

PSTS - Posterior Superior Temporal Sulcus

PTOM - Putative Theory of Mind areas

ROI - Region of Interest

RTPJ - Right Temporal Parietal Junction

SPM - Statistical Parametric Mapping Software

STS - Superior Temporal Sulcus

TMS - Transcranial Magnetic Stimulation

TOM - Theory of Mind

TPJ - Temporal Parietal Junction

TR - Repetition Time

VMPFC Ventral Medial Prefrontal Cortex

# ENGLISH SUMMARY

People involved in a social interaction can be thought of as being temporarily connected. This connection starts, for example, in the motor system of person A, which leads to observable behavior that is perceived and interpreted by the primary sensory and higher cortices of person B. This in turn can lead to activity in the motor system of person B, which leads to observable behavior that can be perceived and interpreted by person A, and so on. This connection is dynamic and dependent on events in the interaction, such as changes in direction of eye gaze, the words being said and the gestures being made. In this thesis, we investigate such a communicative connection on a neural level.

This research is inspired both by theories explaining how humans understand each other and by the methodological advancements in connectivity analyses. Over the last years, two important ideas have been developed about how people interact and communicate: the discovery of the mirror system and the idea of mentalizing. These two concepts form the background of this research and are described in the following two sections. Furthermore, the development of Granger causality mapping, an analysis method to investigate directed influences between brain regions, allowed us to perform the kind of connectivity analyses we needed.

## THE MIRROR SYSTEM

The idea of mirroring is that our brain 'mirrors' or 'echoes' the perceived actions of other people. This idea was inspired by the discovery of mirror neurons in the nineties (Gallese et al., 1996; Pellegrino et al., 1992). Mirror neurons were discovered in a lab in Parma while researchers were measuring neurons in area F5 (ventral premotor cortex) of the Macaque monkey. The researcher noticed that these neurons not only fire at the moment the monkey grasped a peanut, but also when the monkey was observing the researcher grasping a peanut. It was known that these neurons are involved in execution of goal-directed actions performed with the hand or with the mouth (Kurata and Tanji, 1986; Rizzolatti et al., 1988), but now it became clear that these neurons also exhibited sensory characteristics. These neurons thus represent both the execution of an action as well as the observation of an action. This discovery constitutes an important step in neuroscience, because it confirmed the idea that perception and action are inextricably linked on a neural level. This idea was already present in psychological theories. James Gibson, for example, claimed that perception

in essence consists of perceiving 'action affordances' (Gibson, 1986).

Shortly after the first discovery of mirror neurons in the Macaque monkey, the question rose whether a similar mirroring mechanism might also be present in the human brain. Because measuring individual neurons is hardly possible without damaging a brain, research initially focused on the question whether brain *areas* exist, which show activity both during execution and observation of an action (Buccino et al., 2001; Grafton et al., 1996; Grèzes et al., 1998; Grèzes and Decety, 2001; Grèzes et al., 2003; Nishitani and Hari, 2000, 2002; Perani et al., 2001; Gazzola et al., 2007b,a; Gazzola and Keysers, 2008). Brain areas with mirroring properties were indeed identified and were collectively dubbed the human mirror neuron system (Keysers and Gazzola, 2009). These areas comprise the ventral and dorsal premotor cortex, the inferior parietal lobe, and the middle temporal lobe (see Figure 1). Furthermore, several other brain areas show an overlap between experience and observation of for example emotions (Bastiaansen et al., 2009; Wicker et al., 2003), and sensations (Singer et al., 2004; Keysers and Perrett, 2004). Innovative experiments that make use of, for example, 'cross-modal repetition suppression', have shown that there is scientific evidence for the existence of individual mirror neurons in the human brain (Kilner et al., 2009; Lingnau et al., 2009; Chong et al., 2008; Mukamel et al., 2010).

Because mirror neurons form a link between actions we perform ourselves and actions we see others do, it is assumed that mirror neurons have a function in understanding other people (see Rizzolatti and Sinigaglia, 2010, for a recent overview of literature). When observing someone else performing an action, our own motor representations of that action become active, 'as if ' we are doing this action ourselves. This is the core idea of simulation theory: we understand other people's actions, because we internally simulate these actions in our own brain (Goldman, 1992; Gibson, 1986; Gallese, 2003).

Important for the research in the current thesis is that simulation theory makes a prediction about mirror neurons. Simulation theory claims that mirror neurons in the brain of the observer resonate with the mirror neurons in the brain of the actor. The term 'resonance' is used loosely here and refers to the ups and downs in the brain activity of one person that cause similar ups and downs in the brain activity of the other person. (Gallese and Goldman, 1998; Gallese et al., 2004; Rizzolatti et al., 2001). This resonance property is investigated in Chapter 4 of the current thesis.

## MENTALIZING

Besides simulation as a mechanism to understand others, people have the capacity to think about and understand others on a more reflective level. This can be illustrated with a typical scene from a soap opera, such as the Bold and the Beautiful: Taylor

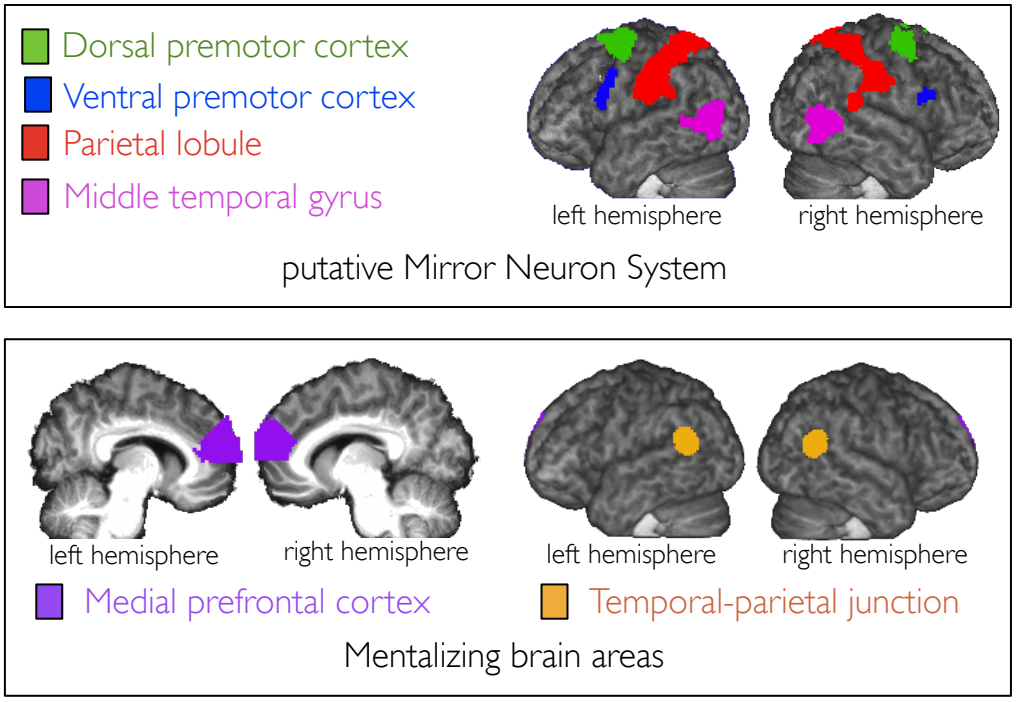



Figure 1: Areas constituting the putative Mirror Neuron System and the mentalizing system in the human brain.

and Ridge are about to get married. Without Taylor knowing this, Brooke is about to confess that she is pregnant with Ridge's baby. She hopes that by revealing this she can prevent this marriage from happening. To understand and appreciate such a situation, we have to be able to track what all characters involved know, what they do not know and predict what they will think and feel when they will find out. This ability to attribute mental states, beliefs and desires to others is called having a Theory of Mind (Premack and Woodruff, 1978; Wimmer and Perner, 1983). The process of reasoning about other people's mental states is often referred to as mentalizing (Frith and Frith, 1999). Mentalizing involves different brain areas than mirroring (Amodio and Frith, 2006; Fletcher et al., 1995; Frith and Frith, 2006, 2003; Gallagher et al., 2000; Gusnard et al., 2001), the most important being the ventral medial prefrontal cortex and the temporal-parietal junction (see Figure 1).

## CHARADES

There has been a lot of debate about which of the two theories (theory of mind or simulation theory) can explain most of human interpersonal understanding (Gallagher, 2007; Hickok, 2009; Saxe and Wexler, 2005). Many now believe that mentalizing is a

separate mechanism from the more basic, low-level motor simulation (Uddin et al., 2007; de Lange et al., 2008; Brass et al., 2007; Overwalle and Baetens, 2009). The debate has therefore shifted to the more fruitful question of how these two systems work together to achieve a full understanding of other people (Keysers and Gazzola, 2007). The charades experiment that is central to this thesis was set up in part to investigate this issue.

In the charades experiment couples of participants played the game of charades by taking turns gesturing and guessing concepts in the MR-scanner. Participants were presented with a word on the screen (either an action or an object, for example nutcracker, knitting, shaving) and were instructed to convey the meaning of this word by gestures. Their gestures were recorded on video and presented to their partners, who went into the MR-scanner to guess what their partner had gestured. The partners had to push a button when they thought they knew what word was being portrayed. After a while, the two partners switched roles, with the former gesturer becoming the guesser and vice-versa. Both partners guessed and gestured 14 words in total. On a different day, they returned for the control condition in which they observed exactly the same movies of their gesturing partner, but now with the instruction to try not to interpret the gestures. In this way, we recorded the brain activity of both partners in the social interaction, as if they were scanned simultaneously. The set-up of the experiment allowed us to investigate several questions: (1) Are mirroring and mentalizing areas involved during the production and interpretation of gestures? (2) Does brain activity of the guesser resonate with brain activity of the gesturer? (3) How do areas of the mirror system in the guesser work together to achieve this interpretation? The following three sections discuss these issues.

## INVOLVEMENT OF MIRRORING AND MENTALIZING AREAS

Our first research question is whether mirroring and mentalizing areas are involved during the production and interpretation of communicative gestures. Our hypothesis was that both the mirror system and the mentalizing system would be active during the charades game. The reasoning behind this was that on the one hand the mirror system would be involved in transforming observed gestures into the motor programs that could produce such a gesture, while on the other hand the mentalizing system would be involved in interpreting the communicative intentions of the other player. The results, described in Chapter 3 of this thesis, support in part our hypothesis. Gesture production and gesture interpretation recruit many of the same brain areas, showing a large overlap. The areas that are active are in part indeed areas of the mirror system. When the time course of activity during the production and interpretation of gestures is considered, it can be seen that all mirror areas elevate their activity

substantially above baseline for the whole period of the gesture. The mirror system is thus indeed involved in gestural communication.

The ventral medial prefrontal cortex (vmPFC), which is considered the most important of the mentalizing system, is not found to be active during any stage of the game. This was in disagreement with our hypothesis. Another study did find activity in this region during the interpretation of gestures, but used gestures that refer to inner mental states (Gallagher and Frith, 2004). The gestures in our experiment, however, were either actions or objects (e.g. knitting, nutcracker, board game). Speculative, it could be this difference in gestures that explains the absence of activity in the vmPFC.

A different explanation could be that the vmPFC is part of the so-called 'default network' of the brain. The default network is a conglomerate of areas that show a distinct activity during the abscence of a clear cognitive task (during 'rest'). This network was discovered in a meta-study, in which activity during rest was compared to activity during all kinds of tasks. No matter which task was being performed, during the rest period similar areas were consistently active. One of them is the ventral medial prefrontal cortex (Gusnard et al., 2001; Raichle et al., 2001). When attention is not focused on the external world, these areas seem to process information that has to do with the self, such as autobiographical memories, future plans and thinking about the perspective of others (Buckner et al., 2008). An explanation for the absence of vmPFC activity during gestural communication could be that the vmPFC was active both during the game as well as during the rest period (looking at a fixation cross). This could have made it hard to measure a difference in activity.

### RESONANCE AND CONNECTIVITY BETWEEN BRAINS

The results described above were calculated by using a classic fMRI analysis. This means that we have tested which brain regions increase their level of activity when a task starts and decrease their level of activity when a task is finished. What we furthermore were interested in was whether areas of the mirror system would show a resonance between gesturer and guesser. The charades experiment was set up in such a way that brain activity of both gesturer and guesser were measured, as if they they were scanned simultaneously. This enabled us to determine whether brain activity of the guesser is influenced by the brain activity of the gesturer. In other words, we used Granger causality (further described below) to test where in the brain of the guesser, brain activity could be predicted by brain activity of the mirror system of the gesturer. The results, described in Chapter 4 show that activity in mirror areas of the guesser do indeed track the activity of mirror areas of the gesturer accurately. This means that a consistent relationship exists between mirror areas of the gesturer and the guesser.

Surprisingly, we found that the ventral medial prefrontal cortex of the guesser has

a similar relation with mirror areas of the gesturer. We did not expect this, since a classical analysis of the same data did not show an involvement of this area. These results indicate that the vmPFC does play a role in the observation and interpretation of communicative gestures.

Besides the original analysis between the gesturer and the guesser, we have also calculated several control analyses. Participants had performed a control condition in which they were shown the gesture movies of their partners again, but were asked to try not interpret them ('passive observation'). De temporal relations between the mirror systems and the vmPFC that we found between gesturer and guesser was strongly decreased in an analysis in which we paired the gesturer to the passive observer. This shows that the temporal relation between the two brains is dependent on the task that is performed, even with exactly the same stimuli. Of course, it is very difficult to observe, but not interpret. Participants reported, however, to have interpreted less during the passive observation condition than during the guessing condition.

When the analysis was performed between a gesturer and a random guesser (thus using brain activity of someone who had observed gestures of a different gesturer), a similar effect of decreased temporal relations with regard to the orginal analysis was found. Even when we paired the gesturer to a random guesser, but paired the words such that they corresponded between gesturer and random guesser, the temporal relations between brains decreased strongly when compared to the temporal relations between gesturer and guesser.

### CONNECTIVITY WITHIN THE MIRROR SYSTEM

Our last research question that we tested with the charades experiment concerned the influences from and to mirror areas within the brain of the guesser. We compared hypotheses generated by different models of the mirror system. A classic account of the mirror system could be described as strictly 'feed-forward' (Kilner et al., 2007a). This means that a visual representation of an observed action is transformed into the corresponding motor programs, which leads to an understanding of the action. On a neuronal level, this would mean that the information flow would predominantly go from temporal cortex → parietal cortex → premotor cortex. An alternative account sees the mirror system as a dynamic feedback control system. This model is inspired by both 'forward' and 'inverse' models from motor control theory (Wolpert et al., 2003; Voss et al., 2006; Greenwald, 1970). During action execution, a copy of motor commands is sent to sensory areas to predict sensory and proprioceptive consequences of that action. When, for example, you reach for a tea cup, you expect your own arm and hand to start stretching out in front of you. This is called the forward model.

An inverse model, in contrast to a forward model, does exactly the opposite: when you see somebody reaching for a tea cup, the inverse model calculates what motor commands should be executed to achieve this reaching behavior. The dynamic feedback control theory assumes that forward models are inhibitory, which means that they can decrease the amount of activity in brain areas that they are connected with (Hietanen and Perrett, 1993, 1996). An illustrative example is formed by the fact that it is difficult to tickle yourself: the accurate expectation of your own tickling hand decreases the sensitivity of the tickle (Blakemore et al., 1998). In contrast, inverse models work excitatory and thus can increase the activity of areas with which they are connected (Gallese et al., 1996; Keysers et al., 2003; Kraskov et al., 2009). The dynamic feedback control theory thus combines inhibitory forward and excitatory inverse models. On a neural level, this predicts an information flow from premotor cortex → parietal cortex → temporal cortex.

The results, described in Chapter 6 show that the information flow during gesture interpretation is predominantly posterior: from premotor cortex → parietal cortex → temporal cortex. These results fit better with the mirror system functioning as a dynamic feedback control system than as a strict feed-forward system. The combination of inhibitory forward and excitatory inverse models leads to the following chain of events in the brain of the guesser: when we see a predictable chain of events, the beginning is fully represented in the visual cortex and triggers motor programs through the inverse model. These motor programs are then 'forwarded' to predict future visual (and somatosensory) stimuli. If these stimuli conform to the predictions, they will be inhibited. The visual → premotor stream of information is reduced and the premotor representations triggered will not be substantially updated. If the visual information violates the predictions, it will not be inhibited, leading to a renewed visual → premotor stream of information and an update of the motor representations. This account of the mirror system is in line with the idea that its main function is a predictive one (Keysers and Perrett, 2004; Kilner et al., 2007b; Miall, 2003).

### GRANGER CAUSALITY

To investigate directed influences between brain regions within and between brains, we used Granger causality mapping. Granger causality is a measure of directed influence between two time series. Originally conceptualized in the econometric field by Wiener and formalized by Granger (Wiener, 1956; Granger, 1969), it was introduced as a connectivity analysis for fMRI data in 2003 by Goebel et al. (2003) and Roebroeck et al. (2005). Clive Granger formalized causality between two time series using the intuitively appealing concept of temporal precedence: if a signal change in A is consistently followed by a signal change in B, A Granger-causes B. Mathematically,

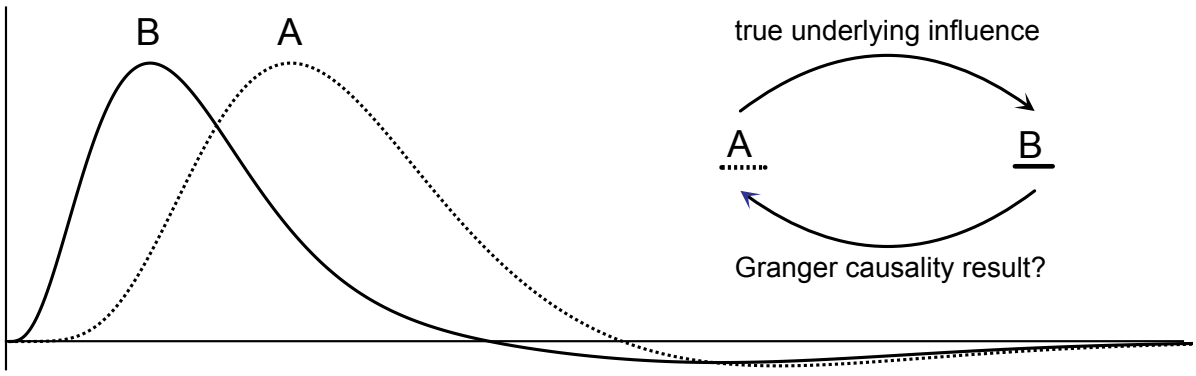



Figure 2: Could a faster hemodynamic response lead to an inverted Granger causality result?

this is calculated by comparing two regression equations: one in which the current value of a time series $y_i$ is explained by its own past ($y_{i-j}$) with one in which the same time series $y_i$ is explained both by its own past and the past of another time series ($x_{i-j}$). This results in error variances, whose F-ratio quantifies the influence $x$ exerts on $y$. The converse influence of $y$ on $x$ is also calculated and the difference between the two reveals the dominant direction of influence between the two time series. In this way, Granger causality provides a statistical measure of directed influences between brain regions and has the advantage that it does not require an underlying anatomical model. It maps influences from a certain seed region to the rest of that brain and vice versa.

Results of Granger causality analyses of fMRI data are interpreted as indications of information flow on a neuronal level. This is an indirect inference, however, as fMRI measures BOLD responses rather than neuronal activity directly. The BOLD response (Blood Oxygen Level Dependent response) is essentially a measure of changes in deoxyhemoglobin level triggered by changes in neural activity. It is assumed to originate from neural activity, predominantly synaptic (Logothetis et al., 2001; Logothetis and Wandell, 2004), but unfolds later in time (~ 4 - 6s). The temporal characteristics of the hemodynamic response that links neural activity to changes in the BOLD signal is not equal across brain regions and participants (Rajapakse et al., 1998; Aguirre et al., 1998; Kruggel and von Cramon, 1999; Handwerker et al., 2004) and this variability is assumed to cause problems for Granger causality analyses (David et al., 2008; de Marco et al., 2009; Roebroeck et al., 2005; Friston, 2009; Chang et al., 2008). One fear is that a systematic difference in hemodynamic response between two regions might introduce temporal precedence where there was none, leading to the report of spurious Granger causality findings. Another fear is that a difference in hemodynamic response might invert the reported direction of Granger causality. The intuitive

idea behind this is as follows: If region A causes neural activity changes in region B and region B has a faster hemodynamic response than region A, a Granger causality analysis of the BOLD signal might indicate a net influence going from B to A rather than the true underlying neural causality that goes from A to B (see Figure 2).

In a between-brains analysis, these hemodynamic differences between brain regions and individuals pose less of a problem because the time lags between neural activity in the two brains, in the order of seconds, are larger than differences in hemodynamic delay, in the order of tenth of seconds (Handwerker et al., 2004). In a within-brain analysis, however, this issue is more pressing. Therefore, we examined whether Granger causality can indeed deduce the dominant flow of neural information even when differences in hemodynamic response are present between different brain regions and between participants. We used computer simulations to systematically investigate the effect of the following factors: (1) the directed influence between two time series, (2) the neuronal delay between them, (3) the hemodynamic responses with which the time series were convolved and (4) the level of noise in the signal. The results, described in Chapter 5, indicate that when no influence is present between two brain areas, there is only a marginal chance (< 5%) to report significant Granger causality results (false positives). Furthermore, when an influence is present and the neuronal delay between two areas is longer than 1 second, no inversions are observed, even when the difference in hemodynamic delays goes against the direction of influence (thus with a longer hemodynamic delay for the source than for the target). Additionally, the sensitivity of Granger causality group analyses at this length of neuronal delay is adequate (>80% for a realistic noise level). Granger causality is therefore an adequate method to investigate influences in a *between* brains context, in which neuronal delays are in the order of seconds.

With regard to Granger causality in a within brain situation, our simulations indicate that sensitivity is rather limited with short neuronal delays. However, we also found that if a significant Granger causal relation is found between two brain regions, the direction of this reported influence is correct in ~80% of the cases. Granger causality on a group level is thus less vulnerable for hemodynamic variability than has often been assumed. Our results therefore argue that Granger causality is a valid method to investigate connectivity within a brain.

## CONCLUSIONS

Concluding, we can say that both the mirror system and the mentalizing system are involved in playing the game charades. Even though using a classical GLM analysis the vmPFC was not found to substantially increase its activity above baseline, Granger causality between brains revealed that activity in this area of the guesser accurately

follows the activity in mirror areas of the gesturer. We have introduced and validated Granger causality as an analysis method to investigate influences *between* brains. Applying this method on brain activity of mirror areas during social interaction shows a resonance between these areas between brains. This indicates that it is important to investigate both brains involved in the interaction, because it can reveal relationships that would otherwise have stayed hidden.

# BIBLIOGRAPHY


Aguirre, G. K., Zarahn, E., and D'esposito, M. (1998). The variability of human, bold hemodynamic responses. *Neuroimage*, 8(4):360–9.

Amodio, D. M. and Frith, C. D. (2006). Meeting of minds: the medial frontal cortex and social cognition. *Nature Reviews of Neuroscience*, 7(4):268–77.

Bastiaansen, J. A. C. J., Thioux, M., and Keysers, C. (2009). Evidence for mirror systems in emotions. *Philosophical transactions of the Royal Society of London Series B, Biological sciences*, 364(1528):2391–404.

Blakemore, S. J., Wolpert, D. M., and Frith, C. D. (1998). Central cancellation of self-produced tickle sensation. *Nature Neuroscience*, 1(7):635–40.

Brass, M., Schmitt, R. M., Spengler, S., and Gergely, G. (2007). Investigating action understanding: inferential processes versus action simulation. *Current biology*, 17(24):2117–21.

Buccino, G., Binkofski, F., Fink, G. R., Fadiga, L., Fogassi, L., Gallese, V., Seitz, R. J., Zilles, K., Rizzolatti, G., and Freund, H.-J. (2001). Action observation activates premotor and parietal areas in a somatosopic manner: An fmri study. *The European Journal of Neuroscience*, 13:400–404.

Buckner, R. L., Andrews-Hanna, J. R., and Schacter, D. L. (2008). The brain's default network: anatomy, function, and relevance to disease. *Annals of the New York Academy of Sciences*, 1124:1–38.

Chang, C., Thomason, M. E., and Glover, G. H. (2008). Mapping and correction of vascular hemodynamic latency in the bold signal. *Neuroimage*, 43(1):90–102.

Chong, T. T.-J., Cunnington, R., Williams, M. A., Kanwisher, N., and Mattingley, J. B. (2008). fmri adaptation reveals mirror neurons in human inferior parietal cortex. *Current biology*, 18(20):1576–80.

David, O., Guillemain, I., Saillet, S., Reyt, S., Deransart, C., Segebarth, C., and Depaulis, A. (2008). Identifying neural drivers with functional mri: an electrophysiological validation. *PLoS Biology*, 6(12):2683–97.

de Lange, F. P., Spronk, M., Willems, R. M., Toni, I., and Bekkering, H. (2008). Complementary systems for understanding action intentions. *Current biology*, 18(6):454–7.

de Marco, G., Devauchelle, B., and Berquin, P. (2009). Brain functional modeling, what do we measure with fmri data? *Neuroscience Research*, 64(1):12–9.

Fletcher, P. C., Happè, F., Frith, U., and Baker, S. (1995). Other minds in the brain: A functional neuroimaging study of 'theory of mind' in story comprehension. *Cognition*, 57(2):109–128.

Friston, K. (2009). Causal modelling and brain connectivity in functional magnetic resonance imaging. *PLoS Biology*, 7(2):e33.

Frith, C. D. and Frith, U. (1999). Interacting minds–a biological basis. *Science*, 286(5445):1692–1695.

Frith, C. D. and Frith, U. (2006). The neural basis of mentalizing. *Neuron*, 50(4):531–534.

Frith, U. and Frith, C. D. (2003). Development and neurophysiology of mentalizing. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences*, 358(1431):459–473.

Gallagher, H. L. and Frith, C. D. (2004). Dissociable neural pathways for the perception and recognition of expressive and instrumental gestures. *Neuropsychology*, 42(13):1725–36.

Gallagher, H. L., Happè, F., Brunswick, N., Fletcher, P. C., Frith, U., and Frith, C. D. (2000). Reading the mind in cartoons and stories: An fmri study of 'theory of mind' in verbal and nonverbal tasks. *Neuropsychology*, 38:11–21.

Gallagher, S. (2007). Simulation trouble. *Social Neuroscience*, 2(3-4):353–65.

Gallese, V. (2003). The manifold nature of interpersonal relations: The quest for a common mechanism. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences*, 358:517–528.

Gallese, V., Fadiga, L., Fogassi, L., and Rizzolatti, G. (1996). Action recognition in the premotor cortex. *Brain*, 119(2):593–609.

Gallese, V. and Goldman, A. (1998). Mirror neurons and the simulation theory of mind-reading. *Trends in Cognitive Sciences*, 12:493–501.

Gallese, V., Keysers, C., and Rizzolatti, G. (2004). A unifying view of the basis of social cognition. *Trends in Cognitive Sciences*, 8(9):396–403.

Gazzola, V. and Keysers, C. (2008). The observation and execution of actions share motor and somatosensory voxels in all tested subjects: Single-subject analyses of unsmoothed fmri data. *Cerebral Cortex*, 19(6):1239–1255.

Gazzola, V., Rizzolatti, G., Wicker, B., and Keysers, C. (2007a). The anthropomorphic brain: The mirror neuron system responds to human and robotic actions. *Neuroimage*, 35:1674–1684.

Gazzola, V., van der Worp, H., Mulder, T., Wicker, B., Rizzolatti, G., and Keysers, C. (2007b). Aplasics born without hands mirror the goal of hand actions with their feet. *Current biology*, 17(14):1235–40.

Gibson, J. (1986). The ecological approach to visual perception. *Boston, Houghton Mifflin*.

Goebel, R., Roebroeck, A., Kim, D.-S., and Formisano, E. (2003). Investigating directed cortical interactions in time-resolved fmri data using vector autoregressive modeling and granger causality mapping. *Magnetic resonance imaging*, 21:1251–1261.

Goldman, A. (1992). In defence of the simulation theory. *Mind and Language*, 7:104–119.

Grafton, S. T., Arbib, M. A., Fadiga, L., and Rizzolatti, G. (1996). Localization of grasp representations in humans by positron emission tomography. 2. observation compared with imagination. *Experimental Brain Research*, 112(1):103–111.

Granger, C. (1969). Investigating causal relations by econometric models and cross-spectral methods. *Econometrica*, 37(3):424–438.

Greenwald, A. G. (1970). Sensory feedback mechanisms in performance control: with special reference to the ideo-motor mechanism. *Psychological Review*, 77(2):73–99.

Grèzes, J., Armony, J., Rowe, J., and Passingham, R. E. (2003). Activations related to mirror and canonical neurones in the human brain: An fmri study. *Neuroimage*, 18:928–937.

Grèzes, J., Costes, N., and Decety, J. (1998). Top-down effect of strategy on the perception of human biological motion: A pet investigation. *Cognitive Neuropsychology*, 15:553–582.

Grèzes, J. and Decety, J. (2001). Functional anatomy of execution, mental simulation, observation, and verb generation of actions: a meta-analysis. *Human Brain Mapping*, 12(1):1–19.

Gusnard, D. A., Akbudak, E., Shulman, G. L., and Raichle, M. E. (2001). Medial prefrontal cortex and self-referential mental activity: relation to a default mode of brain function. *Proceedings of the National Academy of Sciences of the United States of America*, 98(7):4259–64.

Handwerker, D. A., Ollinger, J. M., and D'Esposito, M. (2004). Variation of bold hemodynamic responses across subjects and brain regions and their effects on statistical analyses. *Neuroimage*, 21(4):1639–51.

Hickok, G. (2009). Eight problems for the mirror neuron theory of action understanding in monkeys and humans. *Journal of Cognitive Neuroscience*, 21(7):1229–43.

Hietanen, J. K. and Perrett, D. I. (1993). Motion sensitive cells in the macaque superior temporal polysensory area. i. lack of response to the sight of the animal's own limb movement. *Experimental brain research*, 93(1):117–28.

Hietanen, J. K. and Perrett, D. I. (1996). Motion sensitive cells in the macaque superior temporal polysensory area: response discrimination between self-generated and externally generated pattern motion. *Behavioural brain research*, 76(1-2):155–67.

Keysers, C. and Gazzola, V. (2007). Integrating simulation and theory of mind: from self to social cognition. *Trends in Cognitive Sciences*, 11(5):194–6.

Keysers, C. and Gazzola, V. (2009). Expanding the mirror: vicarious activity for actions, emotions, and sensations. *Current Opinion in Neurobiology*, 19(6):666–71.

Keysers, C., Kohler, E., Umiltà, M., Nanetti, L., Fogassi, L., and Gallese, V. (2003). Audiovisual mirror neurons and action recognition. *Experimental Brain Research*, 153(4):628–636.

Keysers, C. and Perrett, D. (2004). Demystifying social cognition: A hebbian perspective. *Trends in Cognitive Sciences*, 8(11):501–507.

Kilner, J. M., Friston, K. J., and Frith, C. D. (2007a). The mirror-neuron system: a bayesian perspective. *Neuroreport*, 18(6):619–23.

Kilner, J. M., Friston, K. J., and Frith, C. D. (2007b). Predictive coding: an account of the mirror neuron system. *Cognitive Processes*, 8(3):159–166.

Kilner, J. M., Neal, A., Weiskopf, N., Friston, K. J., and Frith, C. D. (2009). Evidence of mirror neurons in human inferior frontal gyrus. *The Journal of Neuroscience*, 29(32):10153–9.

Kraskov, A., Dancause, N., Quallo, M. M., Shepherd, S., and Lemon, R. N. (2009). Corticospinal neurons in macaque ventral premotor cortex with mirror properties: a potential mechanism for action suppression? *Neuron*, 64(6):922–30.

Kruggel, F. and von Cramon, D. Y. (1999). Temporal properties of the hemodynamic response in functional mri. *Human brain mapping*, 8(4):259–71.

Kurata, K. and Tanji, J. (1986). Premotor cortex neurons in macaques: Activity before distal and proximal forelimb movements. *The Journal of Neuroscience*, 6(2):403–411.

Lingnau, A., Gesierich, B., and Caramazza, A. (2009). Asymmetric fmri adaptation reveals no evidence for mirror neurons in humans. *Proceedings of the National Academy of Sciences of the United States of America*, 106(24):9925–30.

Logothetis, N. K., Pauls, J., Augath, M., Trinath, T., and Oeltermann, A. (2001). Neurophysiological investigation of the basis of the fmri signal. *Nature*, 412(6843):150–7.

Logothetis, N. K. and Wandell, B. A. (2004). Interpreting the bold signal. *Annual Review of Physiology*, 66:735–69.

Miall, R. C. (2003). Connecting mirror neurons and forward models. *Neuroreport*, 14(17):2135–7.

Mukamel, R., Ekstrom, A. D., Kaplan, J., Iacoboni, M., and Fried, I. (2010). Single-neuron responses in humans during execution and observation of actions. *Current biology*, 20(8):750–756.

Nishitani, N. and Hari, R. (2000). Temporal dynamics of cortical representation for action. *Proceedings of the National Academy of Sciences of the United States of America*, 97(2):913–8.

Nishitani, N. and Hari, R. (2002). Viewing lip forms: cortical dynamics. *Neuron*, 36(6):1211–20.

Overwalle, F. V. and Baetens, K. (2009). Understanding others' actions and goals by mirror and mentalizing systems: a meta-analysis. *Neuroimage*, 48(3):564–84.

Pellegrino, G., Fadiga, L., Fogassi, L., Gallese, V., and Rizzolatti, G. (1992). Understanding motor events: A neurophysiological study. *Experimental Brain Research*, 91:176–180.

Perani, D., Fazio, F., Borghese, N., Tettamanti, M., Ferrari, S., Decety, J., and Gilardi, M. (2001). Different brain correlates for watching real and virtual hand actions. *Neuroimage*, 14(3):749–758.

Premack, D. and Woodruff, G. (1978). Does the chimpanzee have a theory of mind? *Behavioral and Brain Sciences*, 1:515–526.

Raichle, M. E., MacLeod, A. M., Snyder, A. Z., Powers, W. J., Gusnard, D. A., and Shulman, G. L. (2001). A default mode of brain function. *Proceedings of the National Academy of Sciences of the United States of America*, 98(2):676–82.

Rajapakse, J. C., Kruggel, F., Maisog, J. M., and von Cramon, D. Y. (1998). Modeling hemodynamic response for analysis of functional mri time-series. *Human brain mapping*, 6(4):283–300.

Rizzolatti, G., Camarda, R., Fogassi, L., Gentilucci, M., Luppino, G., and Matelli, M. (1988). Functional organization of inferior area 6 in the macaque monkey. ii. area f5 and the control of distal movements. *Experimental Brain Research*, 71:491–507.

Rizzolatti, G., Fogassi, L., and Gallese, V. (2001). Neurophysiological mechanisms underlying the understanding and imitation of action. *Nature Reviews of Neuroscience*, 2(9):661–70.

Rizzolatti, G. and Sinigaglia, C. (2010). The functional role of the parieto-frontal mirror circuit: interpretations and misinterpretations. *Nature Reviews of Neuroscience*, 11(4):264–74.

Roebroeck, A., Formisano, E., and Goebel, R. (2005). Mapping directed influence over the brain using granger causality. *Neuroimage*, 25:230–242.

Saxe, R. and Wexler, A. (2005). Making sense of another mind: the role of the right temporo-parietal junction. *Neuropsychology*, 43(10):1391–9.

Singer, T., Seymour, B., O'Doherty, J., Kaube, H., Dolan, R., and Frith, C. D. (2004). Empathy for pain involves the affective but not sensory components of pain. *Science*, 303(5661):1157–1162.

Uddin, L. Q., Iacoboni, M., Lange, C., and Keenan, J. P. (2007). The self and social cognition: the role of cortical midline structures and mirror neurons. *Trends in Cognitive Sciences*, 11(4):153–7.

Voss, M., Ingram, J. N., Haggard, P., and Wolpert, D. M. (2006). Sensorimotor attenuation by central motor command signals in the absence of movement. *Nature Neuroscience*, 9(1):26–7.

Wicker, B., Keysers, C., Plailly, J., Royet, J., Gallese, V., and Rizzolatti, G. (2003). Both of us disgusted in my insula: The common neural basis of seeing and feeling disgust. *Neuron*, 40:655–664.

Wiener, N. (1956). Theory of prediction. *Modern Mathematics for Engineers, Series 1*.

Wimmer, H. and Perner, J. (1983). Beliefs about beliefs: Representation and constraining function of wrong beliefs in young children's understanding of deception. *Cognition*, 13(1):103–128.

Wolpert, D. M., Doya, K., and Kawato, M. (2003). A unifying computational framework for motor control and social interaction. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences*, 358(1431):593–602.

# NEDERLANDSE SAMENVATTING (SUMMARY IN DUTCH)

*In gewijzigde vorm gepubliceerd als:* Schippers, M. B. (2010). Hersenen in interactie. *Tijdschrift voor Neuropsychologie*, 5(3).

Wanneer twee mensen met elkaar communiceren, dan ontstaat er een soort tijdelijke verbinding tussen hen. Deze verbinding bestaat uit een keten van gebeurtenissen en begint bijvoorbeeld bij de hersenactiviteit in de motorische cortex van de ene persoon. Deze activiteit leidt tot gedrag, bijvoorbeeld het maken van een gebaar, dat wordt gezien door de andere persoon. Deze bekijkt en interpreteert dit gebaar wat leidt tot activiteit in de visuele, sensorische en associatieve cortices. Dit kan dan weer leiden tot hersenactiviteit in de motorische cortex, tot een ander gebaar en zo verder. Dit proefschrift beschrijft een hersenonderzoek naar zo'n dergelijke indirecte verbinding tussen twee mensen.

Voordat ik in ga op het onderzoek wat we hebben uitgevoerd, is het belangrijk om iets te weten over de achtergrond en inspiratie waarop dit onderzoek is gebaseerd. De afgelopen jaren zijn er twee belangrijke ideeën ontwikkelt over hoe mensen elkaar begrijpen en met elkaar kunnen communiceren: het idee van een spiegelsysteem en een 'redeneersysteem'. Deze twee ideeën vormen de basis van dit onderzoek en worden beschreven in de volgende paragraven. Verder heeft de ontwikkeling van 'Granger causaliteit', een analysemethode om verbindingen tussen hersengebieden vast te stellen ook een belangrijke rol gespeeld, deze wordt hierna beschreven.

## HET SPIEGELSYSTEEM

Het idee van spiegelen is dat ons brein de handelingen van andere mensen 'nabootst'. Aan de basis van dit idee staat de ontdekking van spiegelneuronen ('mirror neurons') in de jaren negentig (Gallese et al., 1996; Pellegrino et al., 1992). Deze spiegelneuronen zijn min of meer per toeval ontdekt in een lab in Parma tijdens het meten van neuronen in het gebied F5 (ventrale premotorische cortex) van de Makaak aap. Een onderzoeker merkte op dat deze neuronen niet alleen reageerden op het moment dat de aap zelf een pinda oppakte, maar ook op het moment dat de aap naar de onderzoeker keek terwijl deze een pinda oppakte. Het was bekend dat deze neuronen betrokken zijn bij het uitvoeren van doelgerichte handelingen met de handen en met de mond. Maar nu werd opeens duidelijk dat deze gebieden ook sensorische eigenschappen bezitten (Kurata and Tanji, 1986; Rizzolatti et al., 1988). Deze neuronen

representeren hiermee zowel het uitvoeren van een handeling als de waarneming van die handeling. De ontdekking van spiegelneuronen had een grote impact, omdat hiermee het vermoeden werd bevestigd dat waarnemen en handelen sterk aan elkaar gekoppeld zijn. Dit idee speelde al langer een rol in psychologische theorieën. James Gibson beweerde bijvoorbeeld dat perceptie bestaat uit het direct waarnemen van handelingsmogelijkheden (Gibson, 1986).

Kort na de eerste ontdekking van spiegelneuronen wilde men weten of de menselijke hersenen ook zo'n dergelijk mechanisme bezitten. Omdat het meten van een individuele neuron vrijwel niet mogelijk is zonder een brein te beschadigen, richtten onderzoeken zich op de vraag of er wellicht hersen*gebieden* bestaan die activiteit laten zien tijdens zowel het uitvoeren als het waarnemen van een handeling (Buccino et al., 2001; Grafton et al., 1996; Grèzes et al., 1998; Grèzes and Decety, 2001; Grèzes et al., 2003; Nishitani and Hari, 2000, 2002; Perani et al., 2001; Gazzola et al., 2007b,a; Gazzola and Keysers, 2008). Dat blijkt inderdaad zo te zijn en de gebieden met deze eigenschap vormen samen het menselijke spiegelsysteem (Keysers and Gazzola, 2009). Deze gebieden zijn de ventrale en dorsale premotorische cortex, de inferieure parietale cortex en de middelste superieure temporele gyrus (zie Figuur 3). Er bestaan overigens niet alleen spiegelgebieden die een overlap in activiteit laten zien voor het uitvoeren en waarnemen van handelingen, maar ook voor het ervaren en het waarnemen van emoties en sensaties, zoals walging, aanraking en pijn (Wicker et al., 2003; Keysers and Perrett, 2004; Singer et al., 2004; Bastiaansen et al., 2009). Innovatieve experimenten, die bijvoorbeeld gebruik maken van 'cross-modal repetition suppression', hebben inmiddels wetenschappelijk bewijs geleverd voor het bestaan van individuele spiegelneuronen in de menselijke hersenen (Kilner et al., 2009; Lingnau et al., 2009; Chong et al., 2008; Mukamel et al., 2010).

Doordat spiegelneuronen een directe link leggen tussen de handelingen die we zelf uitvoeren en de handelingen die we anderen zien doen, wordt aangenomen dat spiegelneuronen een functie hebben in het begrijpen van wat de ander aan het doen is (zie Rizzolatti and Sinigaglia, 2010, voor een recent overzicht van de literatuur). Bij het zien van een handeling van iemand anders wordt de motorische representatie van deze handeling in de eigen hersenen actief, alsof deze handeling zelf wordt uitgevoerd. Dit idee vormt de kern van de *simulatietheorie*: we begrijpen wat een ander doet doordat we deze handeling als het ware simuleren in onze eigen hersenen (Goldman, 1992; Gibson, 1986; Gallese, 2003).

Belangrijk voor het onderzoek in dit proefschrift is dat de simulatietheorie een voorspelling maakt over spiegelneuronen. Deze theorie beweert namelijk dat spiegelneuronen in het brein van degene die een handeling waarneemt resoneren met de spiegelneuronen van degene die de handeling uitvoert. De term 'resonantie' wordt hier losjes gebruikt en er wordt mee bedoeld dat de pieken en dalen in de hersenactiviteit van het motorsysteem van de ene persoon overeenkomstige pieken en dalen

veroorzaakt in de hersenactiviteit van het motorsysteem in de andere persoon (Gallese and Goldman, 1998; Gallese et al., 2004; Rizzolatti et al., 2001). In Hoofdstuk 4 van dit proefschrift wordt deze bewering over resonantie onderzocht.

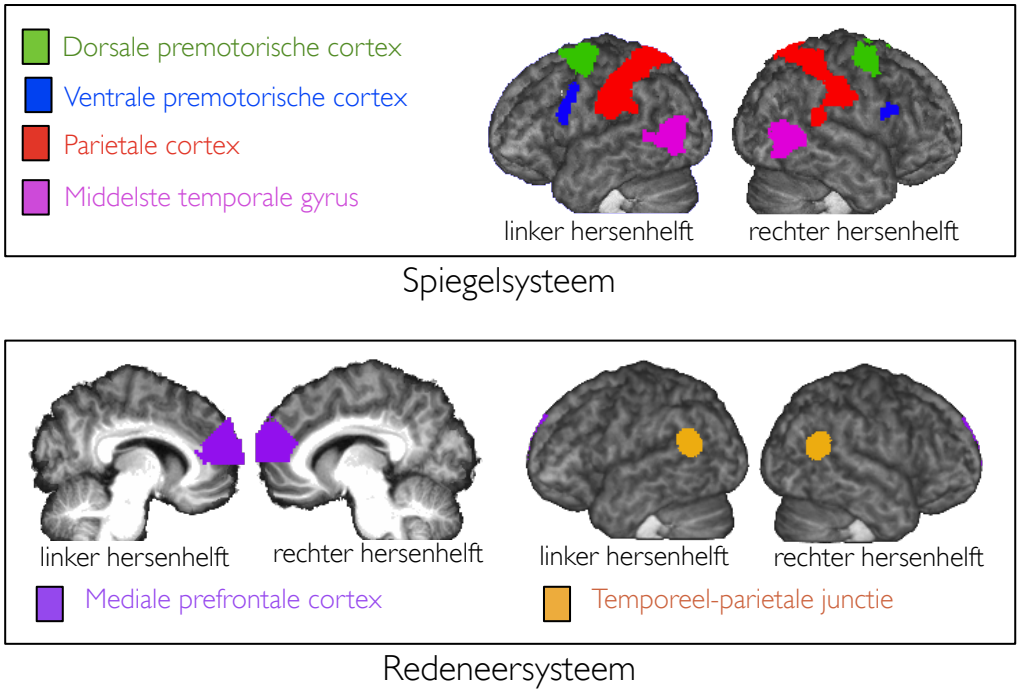
### HET REDENEERSYSTEEM

Naast dit spiegelmechanisme waarmee we anderen begrijpen, bezitten we ook een meer reflectief vermogen om na te denken over wat er in anderen omgaat. Denk bijvoorbeeld aan een typische scene uit een soap, zoals The Bold and the Beautiful: Taylor and Ridge staan op het punt om met elkaar in het huwelijk te treden. Zonder dat Taylor dit weet, staat Brooke op het punt om te vertellen dat ze zwanger is van Ridge, hopende dat ze hiermee de bruiloft kan verhinderen. Om zo'n situatie te kunnen begrijpen en te kunnen waarderen, moeten we in staat zijn om bij te houden wat de verschillende personen wel en niet weten en wat ze zullen denken op het moment dat ze het te horen zullen krijgen. Dit soort bewuste denkprocessen wordt in de literatuur wel 'Theory of Mind' (ToM) genoemd (Premack and Woodruff, 1978; Wimmer and Perner, 1983) en vindt plaats in andere gebieden dan de spiegelgebieden (Frith and Frith, 1999, 2006). Het zijn de 'redeneergebieden' (zie Figuur 3), die actief zijn tijdens bijvoorbeeld het interpreteren van (strip)verhalen en het nadenken over jezelf en anderen (Amodio and Frith, 2006; Fletcher et al., 1995; Frith and Frith, 2006, 2003; Gallagher et al., 2000; Gusnard et al., 2001). De belangrijkste twee gebieden van dit redeneersysteem zijn de ventrale mediale prefrontale cortex en de temporeel-parietale junctie.

### HINTS

In de literatuur is er veel gediscusseerd over de vraag welk van deze twee mechanismes het best kan verklaren van hoe mensen elkaar begrijpen (Gallagher, 2007; Hickok, 2009; Saxe and Wexler, 2005). Inmiddels wordt er door vele aangenomen dat twee verschillende systemen die andere functies vervullen (Uddin et al., 2007; de Lange et al., 2008; Brass et al., 2007; Overwalle and Baetens, 2009). De vraag die nu centraal staat is hoe deze twee mechanismen van spiegelen en redeneren met elkaar samenwerken (Keysers and Gazzola, 2007). Het 'Hints experiment', waarin we gebruik maken van het communiceren door middel van gebaren, is opgezet om deze vraag te kunnen onderzoeken.

In het Hints experiment speelden twaalf romantische koppels het spel Hints terwijl hun hersenactiviteit gemeten werd. De vrouw ging bijvoorbeeld als eerste in de MRI scanner, terwijl haar man in de wachtkamer zat te wachten. Zij kreeg dan woorden te zien op het scherm (bv. 'wenteltrap') en moest deze communiceren naar haar partner




Figuur 3: Hersengebieden die deel uitmaken van het spiegelsysteem (boven) en het redeneersysteem (onder).

door middel van gebaren. Deze gebaren werden opgenomen op video. Nadat ze zeven woorden had uitgebeeld, ging ze uit de scanner en was het de beurt aan haar man. Deze moest in de scanner raden wat zijn vrouw aan hem probeerde duidelijk te maken. Hij kreeg de videobeelden te zien en moest op een knop drukken als hij dacht het geraden te hebben en kon dan antwoord geven via een menu. Nadat hij alle woorden gezien had, mocht hij op zijn beurt woorden gaan uitbeelden. Zijn vrouw ging daarna weer in de scanner zodat man en vrouw beide veertien woorden hadden uitgebeeld en geraden. Op deze manier konden we de hersenactiviteit van de man en de vrouw tijdens het uitbeelden en het raden van dezelfde gebaren naast elkaar leggen, alsof ze tegelijkertijd gescand waren. De opzet van dit experiment stelde ons in staat om meerdere vragen te onderzoeken: (1) Wat is de rol van het spiegel- en het redeneersysteem tijdens het maken en interpreteren van gebaren? (2) Vertoont de hersenactiviteit in de spiegelgebieden van degene die uitbeeldt en degene die aan het raden is een resonantie effect? (3) Hoe werken de spiegelgebieden van degene die aan het raden is met elkaar samen? In de volgende drie paragrafen wordt nader ingegaan op deze drie vragen.

## DE ROL VAN HET SPIEGELSYSTEEM EN HET REDENEERSYSTEEM

Onze eerste onderzoeksvraag is wat de rol is van het spiegelsysteem en het redeneersysteem tijdens het maken en interpreteren van gebaren. Onze voorspellingen waren dat zowel het spiegelsysteem als het redeneersysteem actief zouden zijn tijdens het spelen van het spel hints. De redenatie hierachter was dat aan de ene kant het spiegelsysteem ervoor zorgt dat geobserveerde gebaren omgezet worden in motorische programma's voor soortgelijke handelingen en dat aan de andere kant de redeneergebieden voor een interpretatie van de intenties van de ander zorgen. De resultaten beschreven in Hoofdstuk 3 van dit proefschrift ondersteunen gedeeltelijk onze voorspellingen. Tijdens het maken van de gebaren en het raden van de gebaren zijn er grote overeenkomsten in hersenactiviteit. De gebieden die actief zijn, komen overeen met de spiegelgebieden. Als we daarnaast kijken naar hoe de activiteit verloopt tijdens het maken en het raden van een gebaar, dan is te zien dat de spiegelgebieden tijdens deze hele periode een verhoogde activiteit vertonen. Het spiegelsysteem is dus inderdaad betrokken bij het communiceren door middel van gebaren.

De ventrale mediale prefrontale cortex (vmPFC), het gebied dat als het belangrijkste van de redeneergebieden wordt beschouwd, is daarentegen helemaal niet actief tijdens het gehele spel. Dit hadden wij niet verwacht. Een ander onderzoek vond namelijk wel activiteit in de vmPFC tijdens de interpretatie van gebaren, maar maakte gebruik van gebaren die verwijzen naar innerlijke toestanden (Gallagher and Frith, 2004). De gebaren in ons onderzoek bestonden uit handelingen of objecten (bv. fruit uitpersen, breien, bordspel en notenkraker). We kunnen speculerend dat het verschil in gebaren de activiteit in de vmPFC bepaalt. Een andere verklaring voor de afwezigheid van activiteit in de vmPFC kan worden gegeven door het feit dat de vmPFC ook een onderdeel vormt van het zogenaamde 'default netwerk' in het brein. Dit netwerk is actief als het brein 'in rust' is. Het werd ontdekt in een metastudie, waarbij van een groot aantal onderzoeken de hersenactiviteit tijdens rust werd vergeleken met die tijdens het uitvoeren van een taak. Ongeacht welke taak werd uitgevoerd waren tijdens de rustperiodes steeds dezelfde gebieden actief, waaronder de ventrale mediale prefrontale cortex (Gusnard et al., 2001; Raichle et al., 2001). Wanneer de aandacht niet op de externe wereld gericht is, verwerken deze gebieden informatie die te maken heeft met het zelf, zoals autobiografische herinneringen, toekomstplannen en nadenken over perspectief van anderen (Buckner et al., 2008). De verklaring waarom wij in ons onderzoek geen activiteit vinden in de vmPFC kan zijn dat de vmPFC zowel actief was tijdens de taak als tijdens de rustperiode (het kijken naar een kruis op het scherm tussen de gebaren door). Hierdoor is het moeilijk een verschil te meten.

## RESONANTIE EN VERBINDINGEN TUSSEN TWEE BREINEN

De hierboven beschreven resultaten zijn verkregen door de fMRI data te analyseren op een conventionele manier. Dat wil zeggen dat we getoetst hebben welke gebieden hun activiteit verhogen als een taak begint en deze verlagen als de taak is afgelopen. Waar we hiernaast in geïnteresseerd waren was de vraag of de spiegelgebieden van degene die aan het raden was, zou resoneren met de hersenactiviteit van degene die aan het uitbeelden was. Het Hints experiment is zodanig opgezet dat de hersenactiviteit van beide personen tijdens het spel gemeten was. Hierdoor konden we een link leggen tussen de hersenactiviteit van beide personen. We hebben de analysemethode Granger causaliteit gebruikt (zie paragraaf 'Granger causaliteit' voor een uitgebreide uitleg van deze methode). Hierdoor konden we bepalen of activiteit in hersengebieden van de persoon die raadt wordt beinvloed door activiteit in hersengebieden van degene die uitbeeldt. Met andere woorden, we hebben getest in welke gebieden in het brein van degene die raadt, de hersenactiviteit zich laten voorspellen door de hersenactiviteit in de spiegelgebieden van degene die uitbeeldt. De resultaten beschreven in Hoofdstuk 4 laten zien dat de spiegelgebieden van degene die aan het raden is inderdaad nauwgezet de spiegelgebieden van degene die aan het uitbeelden is volgen. Dat wil zeggen, als een spiegelgebied van de uitbeelder het ene moment iets actiever wordt en daarna iets minder actief, dan gebeurde dit vervolgens ook in de spiegelgebieden van degene die aan het raden was.

Verrassend genoeg bleek dat de ventrale mediale prefrontale cortex van degene die raadt ook een nauwgezette relatie heeft met de spiegelgebieden van degene die aan het uitbeelden is. Dit was niet geheel verwacht omdat we met een klassieke analysemethode hadden gevonden dat de activiteit in dit gebied tijdens het raden van gebaren niet wezenlijk verschilt van diens activiteit tijdens rust. Dit betekent dat dit gebied wel degelijk een rol speelt in het observeren en interpreteren van de gebaren van iemand anders.

Naast de oorspronkelijke analyse tussen degene die uitbeeldt en degene die raadt, hebben we ook een aantal controleanalyses uitgevoerd. Zo hadden de deelnemers van het experiment de opnames van de gebaren van hun partner later (d.w.z na het spelen van Hints) nogmaals gezien, maar nu met de instructie om te proberen geen interpretatie te vormen ('passieve observatie'). De nauwgezette temporele relaties tussen de spiegelgebieden en de ventrale mediale prefrontale cortex die gevonden waren tussen degene die raadt en degene die uitbeeldt, vermindert sterk als degene die raadt wordt vervangen door iemand die slechts passief observeert. Dit laat zien dat de relatie tussen hersengebieden afhankelijk is van de taak die de deelnemers uitvoeren, zelfs bij exact dezelfde stimuli. Uiteraard is het erg moeilijk om te kijken en tegelijkertijd niet te interpreteren, maar de deelnemers aan het onderzoek vertelden wel dat ze in ieder geval minder interpreteerden dan tijdens het echte

spel. Hetzelfde effect treedt op als de analyse wordt uitgevoerd tussen degene die uitbeeldt en een willekeurige andere deelnemer die andere gebaren had gezien. Ook een controleanalyse waarbij we, woord voor woord, de hersenactiviteit van degene die uitbeeldt koppelen aan de hersenactiviteit van een willekeurige rader die een overeenkomstig woord had gezien (maar van een ander persoon), laat zien dat dit resonantie effect sterk vermindert.

### VERBINDINGEN TUSSEN GEBIEDEN VAN HET SPIEGELSYSTEEM

De laatste onderzoeksvraag die we hebben getest met het Hints experiment is de vraag hoe de spiegelgebieden van degene die aan het raden is met elkaar samenwerken. We hebben hiervoor twee modellen van het spiegelsysteem met elkaar vergeleken die hierover tegenstrijdige voorspellingen doen.

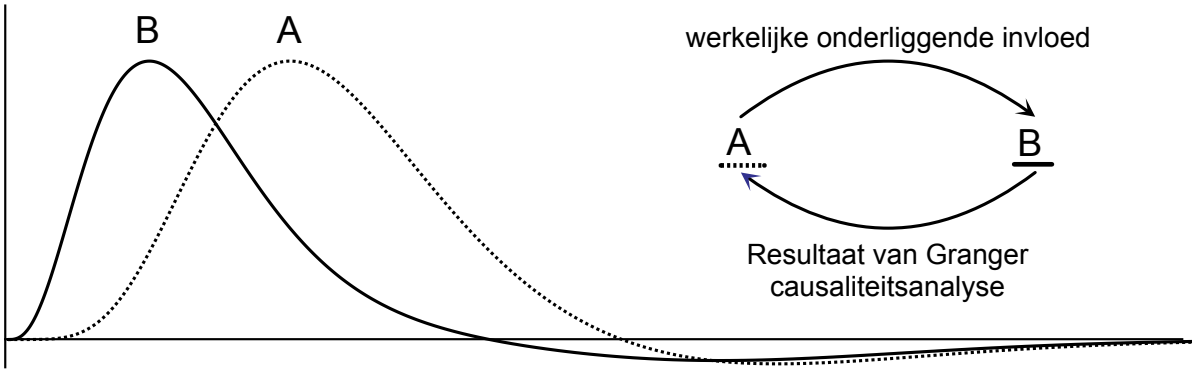
Een gangbaar model van het spiegelsysteem beschrijft het spiegelsysteem als puur 'feed-forward' (Kilner et al., 2007a). Dit houdt in dat de visuele representatie van een waargenomen handeling wordt omgezet in het bijbehorende motorische programma, wat vervolgens leidt tot het begrijpen van die handeling. Op neuraal niveau betekent dit dat de informatiestroom binnen het spiegelsysteem de volgende weg aflegt: temporele cortex → parietale cortex → premotorische cortex. Een alternatief voor dit model is om het spiegelsysteem te beschouwen als een dynamisch feedback controle systeem. Dit systeem is gebaseerd op het idee van 'forward' en 'inverse' modellen uit motor controle theorieën (Wolpert et al., 2003; Voss et al., 2006; Greenwald, 1970). Tijdens het uitvoeren van een handeling wordt een kopie van het motorische programma doorgestuurd naar sensorische cortices, zodat een voorspelling kan worden gemaakt over wat voor sensorische input verwacht kan worden. Als je bijvoorbeeld een kopje oppakt, dan verwacht je je eigen arm en hand voor je te zien die naar het kopje reiken. Dit wordt het forward model genoemd. Het inverse model, daarentegen, doet precies het omgekeerde: als je iemand een kopje op ziet pakken, dan zorgt het inverse model dat de visuele representatie omgezet wordt in de bijbehorende motorische representaties die nodig zouden zijn geweest om deze sensorische input te veroorzaken. De dynamische feedback controle theorie neemt aan dat de forward modellen een inhiberende werking hebben. Ze kunnen dus de activiteit van het hersengebied waarmee ze in verbinding staan verminderen (Hietanen and Perrett, 1993, 1996). Een mooi voorbeeld hiervan is bijvoorbeeld dat het moeilijk is om jezelf te kietelen: de precieze verwachting van je eigen kriebelende hand vermindert de gevoeligheid van de aanraking (Blakemore et al., 1998). Daarentegen hebben inverse modellen voornamelijk een stimulerende werking en versterken daarmee de activiteit van het gebied waarmee ze in verbinding staan (Gallese et al., 1996; Keysers et al., 2003; Kraskov et al., 2009). De dynamische feedback controle theorie is dus een com-

binatie van inhiberende forward- en stimulerende inverse modellen en op neuraal niveau betekent dit dat de informatiestroom gedurende het waarnemen van gebaren de volgende weg af zou moeten leggen: premotorische cortex → parietale cortex → temporele cortex.

De resultaten van onze connectiviteitsanalyse beschreven in Hoofdstuk 6 laten zien dat de informatiestroom tijdens het interpreteren van gebaren voornamelijk in posterieure richting gaat: premotorische cortex → parietale cortex → temporele cortex. Dit is overeenstemming met de dynamische feedback controle theorie. De combinatie van inhiberende forward- en stimulerende inverse modellen die deze theorie voorstelt, leidt tot de volgende keten van gebeurtenissen in het brein: als we een serie voorspelbare handelingen zien, dan wordt het inverse model aangesproken en de visuele representaties worden omgezet in de benodigde motorische programma's. Deze motorische programma's worden vervolgens door het forward model doorgestuurd naar de sensorische cortex om voorspellingen te maken over mogelijke toekomstige input. Als deze voorspelling accuraat blijkt te zijn, dan wordt de input geïnhibeerd. Komen de voorspelling en de werkelijke input echter niet met elkaar overeen, dan wordt de nieuwe input niet geïnhibeerd en dat leidt tot de activatie van andere motorische programma's. Dit benadrukt dat de functie van het spiegelsysteem te maken heeft met het voorspellen wat anderen gaan doen (Keysers and Perrett, 2004; Kilner et al., 2007b; Miall, 2003).

## GRANGER CAUSALITEIT

Om de verbindingen tussen twee breinen en tussen de gebieden van het spiegelsysteem te kunnen onderzoeken, hebben we gebruik gemaakt van de analysemethode Granger causaliteit. Een Granger causaliteitsanalyse berekent of twee hersengebieden elkaar beïnvloeden en in welke richting deze invloed voornamelijk gaat. Deze methode is oorspronkelijk ontwikkeld in het vakgebied van econometrie om de stijging en daling van aandelen te kunnen voorspellen (Wiener, 1956; Granger, 1969). Goebel et al. (2003) en Roebroeck et al. (2005) hebben Granger causaliteit vervolgens geïntroduceerd als een methode om connectiviteit in fMRI data te analyseren. Granger causaliteit is gebaseerd op het intuïtieve idee dat een oorzaak altijd eerder plaatsvindt dan het gevolg. In termen van hersenactiviteit betekent dit dat als een verandering in activiteit in gebied A consistent gevolgd wordt door een activiteitsverandering in gebied B, dan zeggen we dat gebied A een Granger-causale invloed heeft op gebied B. Wiskundig wordt dit uitgerekend door verschillende regressiemodellen met elkaar te vergelijken: één waarin de huidige waarde van een signaal, $y_i$, verklaard wordt door zijn eigen verleden, $y_{i-j}$, en één waarin de huidige waarde van een signaal, $y_i$, wordt verklaard door zowel zijn eigen verleden, $y_{i-j}$, als het verleden van een




Figuur 4: Kan een snellere hemodynamische respons leiden tot een schijnbare omgekeerde invloed bij een Granger causaliteitsanalyse?

ander signaal, $x_{i-j}$. De F-verhouding van de variantie van de residuen van deze twee regressiemodellen kwantificeert de mate van invloed die $x$ uitoefent op $y$. Ook wordt de omgekeerde invloed berekend die $y$ uitoefent op $x$. Door deze twee invloeden van elkaar af te trekken wordt de dominante richting van invloed berekend. Granger causaliteit is dus een statistische maat van de gerichte invloed tussen twee hersengebieden. Het heeft als voordeel dat er geen onderliggend anatomisch model hoeft te worden gespecificeerd, het brengt simpelweg de gebieden in kaart die worden beïnvloed door een bepaald gebied en andersom.

Het feit dat we Granger causaliteit gebruiken als analysemethode brengt potentiële gevaren met zich mee. De resultaten van een Granger causaliteitsanalyse worden namelijk geïnterpreteerd als zijnde informatiestromen op een neuraal niveau. Maar wat we echter meten met fMRI is de hemodynamische respons, in plaats van neurale activiteit. De hemodynamische respons is een gevolg van neurale activiteit, maar ontwikkelt zich met een vertraging van ongeveer 4 tot 6 seconden ten opzichte van deze neurale activiteit (Logothetis et al., 2001; Logothetis and Wandell, 2004). De eigenschappen van de hemodynamische respons zijn niet overal in het brein of bij alle mensen hetzelfde (Rajapakse et al., 1998; Aguirre et al., 1998; Kruggel and von Cramon, 1999; Handwerker et al., 2004). Zo kan een hemodynamische respons in het ene hersengebied een vertraging hebben van 3 seconden, terwijl de vertraging in een ander gebied bijvoorbeeld wel 5 seconden kan zijn. In de literatuur wordt vaak aangenomen dat deze verschillen problemen kunnen veroorzaken voor Granger causaliteitsanalyses (David et al., 2008; de Marco et al., 2009; Roebroeck et al., 2005; Friston, 2009; Chang et al., 2008). Het zou bijvoorbeeld kunnen zijn dat een systematisch verschil in hemodynamische respons leidt tot een schijnbare oorzaak-gevolg relatie, terwijl er op neuraal niveau geen werkelijk invloed bestaat. Daarnaast zou een verschil in hemodynamische respons tussen twee hersengebieden de resultaten van

een Granger causaliteitsanalyse kunnen omdraaien. Stel je bijvoorbeeld voor dat een hersengebied A veranderingen in neurale activiteit in hersengebied B veroorzaakt. Gebied B heeft echter een veel snellere hemodynamische respons dan gebied A (zie Figuur 4). Het lijkt dan net of activiteit in gebied B altijd voorloopt op activiteit in gebied A, terwijl dit alleen een effect is van de hemodynamische respons. Een Granger causaliteitsanalyse is gebaseerd op de metingen van deze hemodynamische responsen en kan dan als resultaat een netto invloed van gebied B naar gebied A geven in plaats van de daadwerkelijk onderliggende neurale invloed die van gebied A naar gebied B gaat.

Granger causaliteitsanalyses tussen de hersenactiviteit van twee verschillende breinen zijn niet erg vatbaar voor deze fouten, omdat de verschillen in hemodynamische respons kleiner zijn dan de tijd die het duurt om een gebaar te zien en te verwerken (Handwerker et al., 2004). Voor onze analyses van het spiegelsysteem binnen één brein, zou het echter wel rol kunnen spelen. Voordat we deze analyses hebben uitgevoerd, hebben we daarom eerst onderzocht of Granger causaliteit wel een goede methode hiervoor is. Door middel van computersimulaties hebben we onderzocht wat het effect is van verschillen in hemodynamische respons op Granger causaliteitsanalyses als deze worden uitgevoerd op groepsniveau. De volgende factoren zijn systematisch onderzocht: (1) de mate van invloed tussen twee signalen, (2) de neuronale afstand tussen de twee signalen (d.w.z. de tijd die het kost om informatie over te dragen), (3) de hemodynamische responsen van de signalen en (4) de hoeveelheid ruis in het signaal.

De resultaten beschreven in Hoofdstuk 5 laten zien de kans klein is (<5%) dat er een significant resultaat wordt gerapporteerd bij een Granger causaliteitsanalyse als er geen onderliggende neurale beïnvloeding bestaat tussen twee signalen. De kans op valse positieven is dus marginaal. Wanneer er wel een beïnvloeding bestaat tussen de twee signalen en de neuronale afstand tussen de signalen groter is dan 1 seconde (bijvoorbeeld wanneer Granger causaliteit wordt toegepast tussen twee breinen), dan worden er geen omgekeerde resultaten gevonden. Zelfs niet als het verschil in hemodynamische responsen tegen de richting van de beïnvloeding gaat (dus als de bron een langzamere hemodynamische respons heeft dan het doelwit). Ook is de sensitiviteit bij deze neuronale afstand boven de 80% (als een realistische hoeveelheid ruis is toegevoegd aan het signaal), wat voldoende is voor een dergelijke analyse. Hiermee hebben we laten zien dat Granger causaliteit een goede methode is om de informatiestroom tussen twee breinen in kaart te brengen.

Wat betreft een Granger causaliteitsanalyse binnen één brein laat de sensitiviteit nog wat te wensen over, vooral bij korte neurale afstanden. Daar staat echter tegenover dat als er een significant Granger causaliteitsresultaat is gevonden, deze in meer dan 80% van de gevallen de juiste richting aangeeft. De resultaten van Granger causaliteit op groepsniveau zijn dus minder gevoelig voor verschillen in hemodynamische

respons dan vaak wordt aangenomen. Granger causaliteit is daarmee een valide methode om informatiestromen binnen één brein te onderzoeken. Hierbij moet wel worden aangetekend dat veel invloeden niet gevonden zullen worden, maar dat dit uiteraard geen bewijs is voor het niet bestaan van deze invloeden.

## CONCLUSIES

In menselijke interactie zoals bij het spelen van het spel Hints, zijn zowel het spiegelsysteem als de redeneergebieden in de hersenen betrokken. We hebben dat kunnen vaststellen door aan te tonen dat tijdens het interpreteren van gebaren, de activiteit van de ventrale mediale prefrontale cortex van de een sterk de spiegelgebieden volgt van de ander, terwijl dat met de klassieke dataanalysemethoden niet kan worden aangetoond. Als methode hebben wij Granger causaliteit geïntroduceerd en gevalideerd, waarmee de verbinding tussen twee breinen, de interactie, geanalyseerd kan worden. Het toepassen van deze analysemethode op spiegelgebieden laat zien dat er inderdaad sprake is van resonantie tussen twee breinen. Het is dus van belang om bij het onderzoeken van het sociale brein beide kanten van een sociale interactie te analyseren. Het analyseren van relaties tussen breinen in plaats van binnen één brein levert nieuwe inzichten op die anders niet aan het licht komen.

# BIBLIOGRAFIE


Aguirre, G. K., Zarahn, E., and D'esposito, M. (1998). The variability of human, bold hemodynamic responses. *Neuroimage*, 8(4):360–9.

Amodio, D. M. and Frith, C. D. (2006). Meeting of minds: the medial frontal cortex and social cognition. *Nature Reviews of Neuroscience*, 7(4):268–77.

Bastiaansen, J. A. C. J., Thioux, M., and Keysers, C. (2009). Evidence for mirror systems in emotions. *Philosophical transactions of the Royal Society of London Series B, Biological sciences*, 364(1528):2391–404.

Blakemore, S. J., Wolpert, D. M., and Frith, C. D. (1998). Central cancellation of self-produced tickle sensation. *Nature Neuroscience*, 1(7):635–40.

Brass, M., Schmitt, R. M., Spengler, S., and Gergely, G. (2007). Investigating action understanding: inferential processes versus action simulation. *Current biology*, 17(24):2117–21.

Buccino, G., Binkofski, F., Fink, G. R., Fadiga, L., Fogassi, L., Gallese, V., Seitz, R. J., Zilles, K., Rizzolatti, G., and Freund, H.-J. (2001). Action observation activates premotor and parietal areas in a somatosopic manner: An fmri study. *The European Journal of Neuroscience*, 13:400–404.

Buckner, R. L., Andrews-Hanna, J. R., and Schacter, D. L. (2008). The brain's default network: anatomy, function, and relevance to disease. *Annals of the New York Academy of Sciences*, 1124:1–38.

Chang, C., Thomason, M. E., and Glover, G. H. (2008). Mapping and correction of vascular hemodynamic latency in the bold signal. *Neuroimage*, 43(1):90–102.

Chong, T. T.-J., Cunnington, R., Williams, M. A., Kanwisher, N., and Mattingley, J. B. (2008). fmri adaptation reveals mirror neurons in human inferior parietal cortex. *Current biology*, 18(20):1576–80.

David, O., Guillemain, I., Saillet, S., Reyt, S., Deransart, C., Segebarth, C., and Depaulis, A. (2008). Identifying neural drivers with functional mri: an electrophysiological validation. *PLoS Biology*, 6(12):2683–97.

de Lange, F. P., Spronk, M., Willems, R. M., Toni, I., and Bekkering, H. (2008). Complementary systems for understanding action intentions. *Current biology*, 18(6):454–7.

de Marco, G., Devauchelle, B., and Berquin, P. (2009). Brain functional modeling, what do we measure with fmri data? *Neuroscience Research*, 64(1):12–9.

Fletcher, P. C., Happè, F., Frith, U., and Baker, S. (1995). Other minds in the brain: A functional neuroimaging study of 'theory of mind' in story comprehension. *Cognition*, 57(2):109–128.

Friston, K. (2009). Causal modelling and brain connectivity in functional magnetic resonance imaging. *PLoS Biology*, 7(2):e33.

Frith, C. D. and Frith, U. (1999). Interacting minds–a biological basis. *Science*, 286(5445):1692–1695.

Frith, C. D. and Frith, U. (2006). The neural basis of mentalizing. *Neuron*, 50(4):531–534.

Frith, U. and Frith, C. D. (2003). Development and neurophysiology of mentalizing. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences*, 358(1431):459–473.

Gallagher, H. L. and Frith, C. D. (2004). Dissociable neural pathways for the perception and recognition of expressive and instrumental gestures. *Neuropsychology*, 42(13):1725–36.

Gallagher, H. L., Happè, F., Brunswick, N., Fletcher, P. C., Frith, U., and Frith, C. D. (2000). Reading the mind in cartoons and stories: An fmri study of 'theory of mind' in verbal and nonverbal tasks. *Neuropsychology*, 38:11–21.

Gallagher, S. (2007). Simulation trouble. *Social Neuroscience*, 2(3-4):353–65.

Gallese, V. (2003). The manifold nature of interpersonal relations: The quest for a common mechanism. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences*, 358:517–528.

Gallese, V., Fadiga, L., Fogassi, L., and Rizzolatti, G. (1996). Action recognition in the premotor cortex. *Brain*, 119(2):593–609.

Gallese, V. and Goldman, A. (1998). Mirror neurons and the simulation theory of mind-reading. *Trends in Cognitive Sciences*, 12:493–501.

Gallese, V., Keysers, C., and Rizzolatti, G. (2004). A unifying view of the basis of social cognition. *Trends in Cognitive Sciences*, 8(9):396–403.

Gazzola, V. and Keysers, C. (2008). The observation and execution of actions share motor and somatosensory voxels in all tested subjects: Single-subject analyses of unsmoothed fmri data. *Cerebral Cortex*, 19(6):1239–1255.

Gazzola, V., Rizzolatti, G., Wicker, B., and Keysers, C. (2007a). The anthropomorphic brain: The mirror neuron system responds to human and robotic actions. *Neuroimage*, 35:1674–1684.

Gazzola, V., van der Worp, H., Mulder, T., Wicker, B., Rizzolatti, G., and Keysers, C. (2007b). Aplasics born without hands mirror the goal of hand actions with their feet. *Current biology*, 17(14):1235–40.

Gibson, J. (1986). The ecological approach to visual perception. *Boston, Houghton Mifflin*.

Goebel, R., Roebroeck, A., Kim, D.-S., and Formisano, E. (2003). Investigating directed cortical interactions in time-resolved fmri data using vector autoregressive modeling and granger causality mapping. *Magnetic resonance imaging*, 21:1251–1261.

Goldman, A. (1992). In defence of the simulation theory. *Mind and Language*, 7:104–119.

Grafton, S. T., Arbib, M. A., Fadiga, L., and Rizzolatti, G. (1996). Localization of grasp representations in humans by positron emission tomography. 2. observation compared with imagination. *Experimental Brain Research*, 112(1):103–111.

Granger, C. (1969). Investigating causal relations by econometric models and cross-spectral methods. *Econometrica*, 37(3):424–438.

Greenwald, A. G. (1970). Sensory feedback mechanisms in performance control: with special reference to the ideo-motor mechanism. *Psychological Review*, 77(2):73–99.

Grèzes, J., Armony, J., Rowe, J., and Passingham, R. E. (2003). Activations related to mirror and canonical neurones in the human brain: An fmri study. *Neuroimage*, 18:928–937.

Grèzes, J., Costes, N., and Decety, J. (1998). Top-down effect of strategy on the perception of human biological motion: A pet investigation. *Cognitive Neuropsychology*, 15:553–582.

Grèzes, J. and Decety, J. (2001). Functional anatomy of execution, mental simulation, observation, and verb generation of actions: a meta-analysis. *Human Brain Mapping*, 12(1):1–19.

Gusnard, D. A., Akbudak, E., Shulman, G. L., and Raichle, M. E. (2001). Medial prefrontal cortex and self-referential mental activity: relation to a default mode of brain function. *Proceedings of the National Academy of Sciences of the United States of America*, 98(7):4259–64.

Handwerker, D. A., Ollinger, J. M., and D'Esposito, M. (2004). Variation of bold hemodynamic responses across subjects and brain regions and their effects on statistical analyses. *Neuroimage*, 21(4):1639–51.

Hickok, G. (2009). Eight problems for the mirror neuron theory of action understanding in monkeys and humans. *Journal of Cognitive Neuroscience*, 21(7):1229–43.

Hietanen, J. K. and Perrett, D. I. (1993). Motion sensitive cells in the macaque superior temporal polysensory area. i. lack of response to the sight of the animal's own limb movement. *Experimental brain research*, 93(1):117–28.

Hietanen, J. K. and Perrett, D. I. (1996). Motion sensitive cells in the macaque superior temporal polysensory area: response discrimination between self-generated and externally generated pattern motion. *Behavioural brain research*, 76(1-2):155–67.

Keysers, C. and Gazzola, V. (2007). Integrating simulation and theory of mind: from self to social cognition. *Trends in Cognitive Sciences*, 11(5):194–6.

Keysers, C. and Gazzola, V. (2009). Expanding the mirror: vicarious activity for actions, emotions, and sensations. *Current Opinion in Neurobiology*, 19(6):666–71.

Keysers, C., Kohler, E., Umiltà, M., Nanetti, L., Fogassi, L., and Gallese, V. (2003). Audiovisual mirror neurons and action recognition. *Experimental Brain Research*, 153(4):628–636.

Keysers, C. and Perrett, D. (2004). Demystifying social cognition: A hebbian perspective. *Trends in Cognitive Sciences*, 8(11):501–507.

Kilner, J. M., Friston, K. J., and Frith, C. D. (2007a). The mirror-neuron system: a bayesian perspective. *Neuroreport*, 18(6):619–23.

Kilner, J. M., Friston, K. J., and Frith, C. D. (2007b). Predictive coding: an account of the mirror neuron system. *Cognitive Processes*, 8(3):159–166.

Kilner, J. M., Neal, A., Weiskopf, N., Friston, K. J., and Frith, C. D. (2009). Evidence of mirror neurons in human inferior frontal gyrus. *The Journal of Neuroscience*, 29(32):10153–9.

Kraskov, A., Dancause, N., Quallo, M. M., Shepherd, S., and Lemon, R. N. (2009). Corticospinal neurons in macaque ventral premotor cortex with mirror properties: a potential mechanism for action suppression? *Neuron*, 64(6):922–30.

Kruggel, F. and von Cramon, D. Y. (1999). Temporal properties of the hemodynamic response in functional mri. *Human brain mapping*, 8(4):259–71.

Kurata, K. and Tanji, J. (1986). Premotor cortex neurons in macaques: Activity before distal and proximal forelimb movements. *The Journal of Neuroscience*, 6(2):403–411.

Lingnau, A., Gesierich, B., and Caramazza, A. (2009). Asymmetric fmri adaptation reveals no evidence for mirror neurons in humans. *Proceedings of the National Academy of Sciences of the United States of America*, 106(24):9925–30.

Logothetis, N. K., Pauls, J., Augath, M., Trinath, T., and Oeltermann, A. (2001). Neurophysiological investigation of the basis of the fmri signal. *Nature*, 412(6843):150–7.

Logothetis, N. K. and Wandell, B. A. (2004). Interpreting the bold signal. *Annual Review of Physiology*, 66:735–69.

Miall, R. C. (2003). Connecting mirror neurons and forward models. *Neuroreport*, 14(17):2135–7.

Mukamel, R., Ekstrom, A. D., Kaplan, J., Iacoboni, M., and Fried, I. (2010). Single-neuron responses in humans during execution and observation of actions. *Current biology*, 20(8):750–756.

Nishitani, N. and Hari, R. (2000). Temporal dynamics of cortical representation for action. *Proceedings of the National Academy of Sciences of the United States of America*, 97(2):913–8.

Nishitani, N. and Hari, R. (2002). Viewing lip forms: cortical dynamics. *Neuron*, 36(6):1211–20.

Overwalle, F. V. and Baetens, K. (2009). Understanding others' actions and goals by mirror and mentalizing systems: a meta-analysis. *Neuroimage*, 48(3):564–84.

Pellegrino, G., Fadiga, L., Fogassi, L., Gallese, V., and Rizzolatti, G. (1992). Understanding motor events: A neurophysiological study. *Experimental Brain Research*, 91:176–180.

Perani, D., Fazio, F., Borghese, N., Tettamanti, M., Ferrari, S., Decety, J., and Gilardi, M. (2001). Different brain correlates for watching real and virtual hand actions. *Neuroimage*, 14(3):749–758.

Premack, D. and Woodruff, G. (1978). Does the chimpanzee have a theory of mind? *Behavioral and Brain Sciences*, 1:515–526.

Raichle, M. E., MacLeod, A. M., Snyder, A. Z., Powers, W. J., Gusnard, D. A., and Shulman, G. L. (2001). A default mode of brain function. *Proceedings of the National Academy of Sciences of the United States of America*, 98(2):676–82.

Rajapakse, J. C., Kruggel, F., Maisog, J. M., and von Cramon, D. Y. (1998). Modeling hemodynamic response for analysis of functional mri time-series. *Human brain mapping*, 6(4):283–300.

Rizzolatti, G., Camarda, R., Fogassi, L., Gentilucci, M., Luppino, G., and Matelli, M. (1988). Functional organization of inferior area 6 in the macaque monkey. ii. area f5 and the control of distal movements. *Experimental Brain Research*, 71:491–507.

Rizzolatti, G., Fogassi, L., and Gallese, V. (2001). Neurophysiological mechanisms underlying the understanding and imitation of action. *Nature Reviews of Neuroscience*, 2(9):661–70.

Rizzolatti, G. and Sinigaglia, C. (2010). The functional role of the parieto-frontal mirror circuit: interpretations and misinterpretations. *Nature Reviews of Neuroscience*, 11(4):264–74.

Roebroeck, A., Formisano, E., and Goebel, R. (2005). Mapping directed influence over the brain using granger causality. *Neuroimage*, 25:230–242.

Saxe, R. and Wexler, A. (2005). Making sense of another mind: the role of the right temporo-parietal junction. *Neuropsychology*, 43(10):1391–9.

Singer, T., Seymour, B., O'Doherty, J., Kaube, H., Dolan, R., and Frith, C. D. (2004). Empathy for pain involves the affective but not sensory components of pain. *Science*, 303(5661):1157–1162.

Uddin, L. Q., Iacoboni, M., Lange, C., and Keenan, J. P. (2007). The self and social cognition: the role of cortical midline structures and mirror neurons. *Trends in Cognitive Sciences*, 11(4):153–7.

Voss, M., Ingram, J. N., Haggard, P., and Wolpert, D. M. (2006). Sensorimotor attenuation by central motor command signals in the absence of movement. *Nature Neuroscience*, 9(1):26–7.

Wicker, B., Keysers, C., Plailly, J., Royet, J., Gallese, V., and Rizzolatti, G. (2003). Both of us disgusted in my insula: The common neural basis of seeing and feeling disgust. *Neuron*, 40:655–664.

Wiener, N. (1956). Theory of prediction. *Modern Mathematics for Engineers, Series 1*.

Wimmer, H. and Perner, J. (1983). Beliefs about beliefs: Representation and constraining function of wrong beliefs in young children's understanding of deception. *Cognition*, 13(1):103–128.

Wolpert, D. M., Doya, K., and Kawato, M. (2003). A unifying computational framework for motor control and social interaction. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences*, 358(1431):593–602.

# ACKNOWLEDGEMENTS

During my PhD period I have heard the most scary stories (supervisors not interested in the work of their students, students slowly drifting a part because of bad communication or no chance of collaboration). Fortunately, for me this was not the case. Many people helped me to finish this PhD project. I would like to thank all the people without whom this thesis could not have been written. I would also like to use this opportunity to thank all those persons that made life so interesting during the last four years.

✠ ✠ ✠

**Christian**, it has been inspirational to work with someone who is so quick of mind and scientifically brilliant as you. I admire the way you can oversee and comprehend the full complexity of obtained results and the relevant steps that consequently should be taken to bring it to a coherent story and article. You have added a level of insight into the articles published in this thesis of which I can only hope to ever achieve it.

**Remco**, wat was het prettig om met jou te werken. Altijd enthousiast, altijd bereid tijd vrij te maken en altijd klaar om 'nerd points' te scoren. Zonder jou zou ik dit project niet af hebben kunnen maken, dank!

**Professor Goebel**, thank you for your help in designing the Charades experiment. Even though we have not met frequently, the notes that I made during our meetings have accompanied me for the full project.

**Alard**, dank voor de prettige samenwerking. De paar keren in Maastricht was het altijd gezellig en heb je me steeds een grote stap verder gebracht. Een goede samenwerking resulterend in een goed resultaat!

**Valeria**, thank you for all your help in this project. You have showed me what science is all about!

**Eleonora**, it was such a pleasure working with you. We wrote the bookchapter so efficiently together... and then it took another two years until it was finally published! I hope you will find your way back to Groningen at a certain point in your life, would be gezellig!

**Luca**, I never met someone who is so dedicated to his work. You were always there and always willing to help not only me but many other people as well, thank you!

**Jo**, thank you for your help in analysing those > 69 million Granger simulations. Without you, I would have been so lost! **Lawrie**, thank you for continuing with my data and all the complicated scripts. I wish it brings you as much luck as it brought me. **Peter Hinrich**, dank voor je hulp bij het schrijven van het voorstel voor de 'Enlighten Your Research' contest. **Stephen Peuchen**, dank voor je hulp bij het indienen van het subsisidievoorstel bij het National Institute of Health. **Michael Spezio**, **Paolo Toffanin** and **Dan Arnstein**, thank you for your valuable comments on the PNAS-article. **Daniel Handwerker**, thank you for selflessly giving us the hemodynamic response data.

✠ ✠ ✠

**Harma** en **Jojanneke**, lieve paranimfen! Tijdens deze vier jaar hebben we met zijn drieën, maar ook alledrie onderling een hechte en diepgaande vriendschap ontwikkeld die zich uit in sauna-bezoeken, frisbeëen, breiclubjes, bitterballen eten, aquajoggen en andere zaken. Hoe anders had deze periode eruit gezien als jullie er niet waren geweest. Mijn innige dank dat ik jullie beiden als vriendinnen mag beschouwen!

Colleagues of the Social Brain Lab, thank you for having shared these four years with me! **Idil**, it took a bit for us to get used to each other, but when we did, we had such a great time! I have such good memories of Naples and Amsterdam and of course in the lab. I hope you will have a great time in Nijmegen! **Jo**, how many times haven't we walked up and down the corridor of the NiC to get some tea?! I miss these moments, but I'm happy you have found a great job close to your family. **Leonardo**, I have come to know you as someone with a huge body of knowledge, skills and intelligence, but also with the biggest inclination to underestimate these. I hope one day you can appreciate your own abilities! Many thanks also to **Annerose, Christiaan, Marc, Mbemba, Nicola, Piray, Sanny**, and **Sui**.

Veel dank aan alle collega's in de gang van het NiC! Vooral de 'dolma dames': **Anne-Marthe, Ans, Leonie, Lisette, Marte**. Maar natuurlijk ook **Ana, Branislava, Cris, Dave** (bedankt voor de inspirerende methodology meetings), **Edith, Esther, Hiske, Jan, Jan-Bernard, Katrien, Marjolijn, Marten, Martijn, Peter, Ramona, Richard, Ruud** en **Shipooo**.

**Bram** en **Brigitte**, ik heb veel van jullie geleerd nog voordat ik aan deze periode was begonnen. Het heeft me zeker beter voorbereid op alle stress. En het was altijd erg

gezellig als we met zijn drieën bij elkaar waren, bedankt!

Zonder goede ondersteuning, kom je nergens, bedankt **Gerry, Hedwig, Anita** (bedankt voor het leren scannen!), **Tinie, Evelien, Diana, Janine**.

⁂ ⁂ ⁂

Many thanks to **all the people of Gronical Dizziness**! Playing Ultimate Frisbee has definitely been an effective way to let out any frustrations... and what a great team-effort we made during the European Championships in 2009!!

**Simon**, I enjoy very much every time you come over for dinner. You are now officially in my circle of trust! **Siebren, Jose, Ray, Tita, Carmen, Paco, Tal**. I have met you all through Paolo, but consider each of you a friend of mine :-) Thank you all for your friendship!

**Harma, Kirsten, Zhen Chih, Lisa, Nanna, Rasa**, and **Pien**, every month I look forward to our cosy gathering of the Strikkenclub: knitting, chatting, drinking wine, can life get any better?!

**Mimi, Edwin, Anna, Freek**, jullie gastvrijheid, openhartigheid, intelligentie en reflectievermogen maakt keer op keer indruk op mij. Dank voor jullie bijzondere vriendschap.

**Vivika** en **Wouter**, ondanks dat de aanleiding in-en in triest was, was het bijzonder om jullie het afgelopen jaar zo intensief mee te hebben gemaakt.

**Anneloes, Augustijn, Ernest, Gaby, Jaap, Jeroen, Laura, Marko, Marlies, Michiel, Mijke, Tim, Wieke**. Dank voor jullie vriendschap lieve vrienden! Dat we nog heel lang gezellig samen mogen 'wibbel-wabbelen'.

**Peter, Corrie, Mariken, Harry, Elles, Stefanie**, we hebben elkaar de afgelopen 4 jaar niet heel vaak gezien, maar de 5 jaar daarvoor des te intensiever. Ik koester de goede herinneringen en hoop jullie zeker niet uit het oog te verliezen.

**Suzanne**, we hebben veel meegemaakt samen en het is moeilijk te onderschatten wat jij voor mij betekent. Dank voor alles!

**Anna, Giorgio, Matteo, Silvia, Angela**. Non é facile avere un figlio/fratello che vive nell'estero e neanche avere 'una straniera' nella famiglia. Mi avete accolto senza

pregiudizzi e mi son subito sentita parte della famiglia, vi sono molto grata. Spero di vedervi spesso. Grasie!

**Tilly, Gerard, Martijn, Michiel, Bahareh**. Waar zou ik zijn zonder goede thuisbasis waar ik altijd op terug kan vallen wat er ook gebeurt?! Dank voor het vanzelfsprekende vertrouwen, de altijd aanwezige steun en natuurlijk de gezelligheid!

**Paolo**, you are the one who has had to deal most with all the ups and downs in this period. Now that we both have obtained our PhD's, it's time to go on some new adventures together. I can't wait!

Groningen, januari 2011

# PUBLICATION LIST

## PEER-REVIEWED

- Schippers, M. B., Gazzola, V., Goebel, R., and Keysers, C. (2009). Playing charades in the fmri: are mirror and/or mentalizing areas involved in gestural communication? *PLoS ONE*, 4(8):e6801.
- Schippers, M. B., Roebroeck, A., Renken, R., Nanetti, L., and Keysers, C. (2010). Mapping the information flow from one brain to another during gestural communication. *Proceedings of the National Academy of Sciences of the United States of America*, 107(20):9388-93.

## ACCEPTED FOR PUBLICATION IN NEUROIMAGE

- Schippers, M. B. and Keysers, C. (in press). Mapping the flow of information within the putative mirror neuron system during gesture observation. *Neuroimage*.
- Schippers, M. B., Renken, R., and Keysers, C. (in press). The effect of intra- and inter-subject variability of hemodynamic responses on group level granger causality analyses. *Neuroimage*.

## BOOKCHAPTER

- Rossi, E.⋆, Schippers, M. B.⋆, and Keysers, C. (2011). Broca's area: linking perception and production in language and actions. In Han, S., editor, *Culture and Neural Frames of Cognition and Communication*, On Thinking. Vol. 3. Springer Berlin Heidelberg
  ⋆ Authors contributed equally to this bookchapter.

## DUTCH

- Schippers, M. B. (2010). Hersenen in interactie. *Tijdschrift voor Neuropsychologie*, 5(3).