THE NEUROSCIENCE OF CROSS-MODAL NARRATIVE COMMUNICATION

PICTURES, PANTOMIMES, AND A THOUSAND WORDS: THE NEUROSCIENCE OF CROSS-MODAL NARRATIVE COMMUNICATION IN HUMANS

BY

YE YUAN, B.SC.

A THESIS

SUBMITTED TO THE SCHOOL OF GRADUATE STUDIES

IN PARTIAL FULFILMENT OF THE REQUIREMENTS

FOR THE DEGREE

DOCTOR OF PHILOSOPHY

MCMASTER UNIVERSITY

McMaster University DOCTOR OF PHILOSOPHY (2017)

Hamilton, Ontario (Psychology, Neuroscience & Behaviour)

TITLE:

Pictures, Pantomimes, and a Thousand Words: The Neuroscience of Cross-Modal Communication in Humans

AUTHOR:

Ye Yuan, B.Sc. (McMaster University)

SUPERVISOR:

Dr. Steven Brown

NUMBER OF PAGES:

xv, 160

LAY ABSTRACT

Communication is the exchange of thoughts and ideas from one person to another, often through the form of narratives. People communicate using speech, gestures, and drawing, or some combination of the three. Although there has been much research on how we understand and produce speech and pantomimes, there is relatively little on drawing, and even less on cross-modal communication. To address these issues, I first documented the brain areas that form a basic drawing network (Chapter 2). Next, I looked at how these brain areas are connected to one another (Chapter 3). Finally, I explored how this basic drawing network responds to a more complex task, and found common brain areas during narrative communication using speech, mime, and drawing (Chapter 4). I discuss how my findings advance our understanding of how the brain communicates narratives cross-modally (Chapter 5).

ABSTRACT

Communication is the exchange of thoughts and ideas from one person to another, often through the form of narratives. People communicate using speech, gesture, and drawing, or some multimodal combination of the three. Although there has been much research on how we understand and produce speech and pantomimes, there is relatively little on drawing, and even less on cross-modal communication. This dissertation presents novel empirical findings that contribute to a better understanding of the brain areas that mediate narrative communication across speech, pantomime, and drawing. Since the neuroscience of drawing was so understudied, I first used functional magnetic resonance imaging (fMRI) to investigate the existence of a basic drawing network in the human brain (Chapter 2). The drawing network was shown to contain three visual-motion areas that process the emanation of the visual image as drawing occurs. Next, to follow up on the poorly-characterized structural connectivity of these areas in the human dorsal visual stream, I used diffusion imaging to explore how these dorsal stream areas are connected (Chapter 3). The tractography results showed structural connectivity for two of the three predicted branches connecting the three visual-motion areas. Finally, I used fMRI to investigate how the basic drawing network is recruited during the more complex task of narrative drawing, and to find common brain areas among narrative speech, pantomime, and drawing (Chapter 4). Results suggest

that people approached narratives in an intrinsically mentalistic fashion in terms of the protagonist, rather than as a mere description of action sequences. Together, these studies advance our understanding of the brain areas that comprise a basic drawing network, how these brain areas are interconnected, and how we communicate stories across three modalities of production. I conclude with a general discussion of my findings (Chapter 5).

ACKNOWLEDGEMENTS

I am especially grateful to my supervisor Dr. Steven Brown for bringing big ideas to the table, for encouraging an open mind to differences in perspectives, and for his endless dedication and effort in cultivating my scholarship. I thank the members of my supervisory committee, Dr. Patrick Bennett and Dr. Judy Major-Girardin, for their guidance and support throughout my graduate degree. To Léa Chauvigné and Matthew Berry, thank you both for the many discussions about life, lab, and leisure.

Outside of academia, I thank my parents for their constant love and support during hard times. To the Hamilton Chinese Christian Fellowship: thank you for being my spiritual family for this past decade. To the Paisley Boys, we have become men of God. It has been an amazing journey, and I look forward to the next time we meet. Finally, I wish to express my deepest gratitude to God, who has never left me in my times of darkness and who has always been the light at the end of the tunnel. *Soli Deo gloria!*

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS AND SYMBOLS

FFA	fusiform face area
fMRI	functional magnetic resonance imaging
IPL	inferior parietal lobule
IPS	intraparietal sulcus
pIPS	posterior intraparietal sulcus
KO	kinetic-occipital
MEP	motor-evoked potential
MT	middle temporal
MFG	middle frontal gyrus
MTG	middle temporal gyrus
LO	lateral occipital
PPA	parahippocampal place area
SPL	superior parietal lobule
SMA	supplementary motor area
Spt	Sylvian parietal temporal region
STG	superior temporal gyrus
STS	superior temporal sulcus
pSTS	posterior superior temporal sulcus
TE	anterior inferotemporal cortex (in monkeys)
TEO	posterior inferotemporal cortex (in monkeys)
TMS	Transcranial magnetic stimulation

DECLARATION OF ACADEMIC ACHIEVEMENT

The following thesis contains five chapters. Chapter 1 provides an overview of the neural processes related to the perception and production of speech, pantomime, and drawing, as well as both aspects of multimodal communication. Chapter 2 is an article published in an open-access peer-reviewed journal and is reproduced verbatim, only reformatted to match the specifications of this thesis. Chapters 3 and 4 are manuscripts in preparation for submission. Chapter 5 summarizes the major findings of the intermediate chapters and proposes avenues for future research.

Chapter 1 – General introduction

Author: Ye Yuan

Chapter 2 – The neural basis of mark making: A functional MRI study of drawing

Authors: Ye Yuan and Steven Brown

Chapter 3 – Structural connectivity of human visual-motion areas within the dorsal visual stream

Authors: Ye Yuan, Jennifer S. W. Campbell, Janine D. Mendola, G. Bruce Pike, and Steven Brown

Chapter 4 – Storytelling is intrinsically mentalistic: An fMRI study of cross-modal narrative production

Authors: Ye Yuan, Judy Major-Girardin, and Steven Brown

Chapter 5 – General discussion

Author: Ye Yuan

CHAPTER 1 – GENERAL INTRODUCTION

Communication is the exchange of thoughts and ideas from one person to another. It is an essential part of human interaction and is often used to convey narratives. The three major modalities by which people communicate ideas to one another are speech, gesture, and drawing, or some multimodal combination of them. This dissertation presents novel empirical research regarding the common brain areas recruited during multimodal narrative production. Although there have been numerous studies investigating the unimodal perception and production of speech and pantomime, there is considerably less research on these aspects of drawing, and nothing at all regarding the multimodal production of narratives. This introduction will briefly outline relevant literature for the perception and production of each modality, offer an overview of the current state of multimodal communicative research, and convey the primary research objectives of the dissertation.

1.1 Speech-specific perception and production

It is well known that speech perception and comprehension implicate Wernicke’s area, found in the posterior superior temporal gyrus (STG; Binder et al., 1997; Wernicke, 1874). This area has been subdivided into a posteromedial subregion that preferentially responds to self-generated non-speech and speech sounds, and an anterolateral subregion that preferentially responds to heard speech as well as during recall of word lists (Wise et al., 2001). In more recent years, advancements in magnetic resonance imaging have provided a more detailed understanding of speech and language areas in the brain, implicating left frontal and peri-Sylvian areas, such as the middle frontal gyrus (MFG), insula, posterior superior temporal sulcus (pSTS), and angular gyrus (Benson et al., 2001; Okada & Hickok, 2006; Wilson et al., 2004). Okada and Hickok (2006) conducted an experiment in which participants passively listened to speech in the first two of four sessions and then performed a covert naming task in the next two sessions. Both speech perception and covert object naming resulted in activation in the superior temporal cortex, supramarginal gyrus, lateral premotor cortex, inferior frontal gyrus, and ventral occipito-temporal areas, demonstrating that there is considerable overlap between brain areas involved in perception and production.

In contrast to the posterior temporal brain areas, anterior temporal regions seem to show greater responses to sentences and sentence-like gibberish (Friederici, Meyer, & von Cramon, 2000; Humphries, Willard, Buchsbaum, & Hickok, 2001; Mazoyer, Murayama, & Dehaene, 1993), as well as to phonetic sounds instead of environmental or artificial control sounds (DeWitt & Rauschecker, 2012). Crinion et al. (2003) conducted an experiment in which participants listened to either narratives or audio-reversed versions of those narratives as control stimuli. The researchers found bilateral activation in the anterior STS/middle temporal gyrus (MTG) in their contrast of conditions, suggesting that these areas are involved in the processing of semantic content at the level of discourse (as opposed to sentence- or word-level processing).

Speech perception does not appear to be an isolated process, and has been shown to invoke speech production areas as well. For example, Fadiga et al. (2002) used transcranial magnetic stimulation (TMS) to assess excitability of the somatotopic tongue area of each participant's primary motor cortex. They simultaneously recorded motor-evoked potentials (MEPs) from the tongue muscles. Participants heard words and pseudo-words in which specific phonemes that require the tongue for pronunciation were either present or absent. The results showed increased excitability of the tongue muscles when participants heard phonemes that required the tongue for pronunciation, compared to when they heard phonemes that did not. Thus, Fadiga et al. (2002) were able to demonstrate that the perception of verbal stimuli activates speech production areas.

Broca's area within the left inferior frontal gyrus is perhaps the best-known of brain areas in speech production (Broca, 1861). In addition to this area, other studies on the production of single words or the performance of word-stem completions reveal brain activations in the left inferior prefrontal cortex, anterior insula, and anterior prefrontal cortices, particularly the orbitofrontal and frontopolar regions (Buckner, Raichle, & Petersen, 1995;

McCarthy et al., 1993). Speech production can be dissociated into planning and execution processes, recruiting the medial and ventrolateral premotor cortices, anterior insula, and cerebellum during planning, and sensorimotor areas and the basal ganglia during execution (Riecker et al., 2005). Feedback of the speech sounds produced is processed in the inferior parietal cortex (Bohland & Guenther, 2006; Geranmayeh et al., 2012). Speech processing, both narrative and non-narrative, implicates a distributed network consisting of frontal, temporal, and parietal areas, though many of these areas also respond to pantomimes, as will be discussed next.

1.2 Pantomime-specific perception and production

Gestures typically fall under the main categories of gesticulation (speech-linked gesture), pantomime, emblem, and finally sign language (McNeill, 1992). Of these categories, only gesticulations (also known as co-speech gestures) are truly multimodal in the sense that they require the expression of both speech and gesture. Although they are communicative, co-speech gestures do not occur in a stand-alone fashion the way that pantomimes do.

Pantomimes, on the other hand, are a category of gestures that McNeill (1992) described as “conveying a narrative line, with a story to tell, produced without speech”. Based on this definition, the semantic brain areas that would be recruited during pantomime comprehension should be similar to those recruited during speech comprehension. Wu and Coulson (2005) used electroencephalography to investigate cortical responses to pantomimes that accompanied soundless cartoon segments. Participants were shown pantomime-cartoon pairs that were either related or unrelated followed by a probe word. The authors found that unrelated pantomime-cartoon pairs elicited brain wave negativity patterns that were similar to perceiving incongruencies in speech, and that related pairs elicited much less negativity-signal than did unrelated pairs.

Perceiving pantomimes necessarily recruits brain areas that respond to biological motion, including occipito-temporal regions such as area KO, as well as major biological-motion centers in the superior temporal sulcus (STS; Grossman et al., 2000; Vaina, Solomon, Chowdhury, Sinha, & Belliveau, 2001). Lower-level visual motion processing centers that do not have a preference for biological motion, such as the V5/MT+ complex, are also implicated in the perception of pantomimes (Grossman et al., 2000; Vaina et al., 2001). The subprocesses of goal-oriented movement and recognition of meaningful actions seem to preferentially activate the inferior parietal lobule (Bonda et al., 1996) and the inferior and middle temporal gyri (Decety et al., 1997), respectively.

The investigation of cognitive and neural processes during pantomime production has traditionally been motivated by the field of apraxia in tool use, where the term apraxia refers to a movement disorder caused by brain damage. Highly influential research by Liepmann and colleagues (see Brown 1988) delineated three major brain areas involved in limb apraxia: a first area that represents sensorimotor information for pantomiming tool use, including the pre- and postcentral gyri and premotor cortices; a second area that plays important roles in the planning and formation of the actions of motoric skills, encompassing the anterior superior parietal lobule (SPL), inferior parietal lobule (IPL), and intraparietal sulcus (IPS); and a third area for retrieving conceptual knowledge, located near the intersection of the occipital, temporal, and parietal lobes. More recent functional neuroimaging research has also found involvement of the inferior frontal cortex, with a strong left lateralization (Goldenberg, Hermsdörfer, Glindemann, Rorden, & Karnath, 2007; Hermsdörfer, Terlinden, Mühlau, Goldenberg, & Wohlschläger, 2007). Additionally, in a review focused on the role of the parietal cortex in spatial cognition, Sack (2009) described key areas of the bilateral prefrontal, premotor, superior parietal, and occipito-temporal cortices, although he did not mention the semantic inferior temporal areas. The parietal and occipito-temporal cortices are also more activated during imagined tool use compared to viewing tool use (Wadsworth & Kana, 2011). Taken together, these studies demonstrate that brain areas important for pantomime perception and production are found in the superior and inferior parietal cortices.

1.3 Drawing-specific perception and production

Vision is by far the dominant modality for perception in humans, and this has stimulated an extensive body of research into the neuroscience of visual processing. The perception of images recruits brain areas that are part of the well-known dorsal and ventral streams of visual perception and cognition (Chao & Martin, 2000; Culham et al., 2003; Grill-Spector et al., 1999; Haxby et al., 2001; Landau & Jackendoff, 1993; Mishkin, Ungerleider, & Macko, 1983; Valyear, Culham, Sharif, Westwood, & Goodale, 2006). Spatial orientation and location are typically classified as dorsal stream functions and are processed in the SPL and in areas near the IPS (Chao & Martin, 2000; Valyear et al., 2006; Zachariou, Klatzky, & Behrmann, 2014), with anterior IPS regions playing a role in the perception of spatial actions such as grasping (Shmuelof & Zohary, 2005). The ventral visual stream, on the other hand, is comprised of object feature processing areas, such as the lateral occipital region (LO) (Bona, Herbert, Toneatto, Silvanto, & Cattaneo, 2014; Grill-Spector et al., 1999; Guggenmos et al., 2015), posterior inferior temporal regions such as V4, and human homologues of monkey temporal-lobe areas TE and TEO (Ungerleider & Mishkin, 1982), as well as ventral temporal regions like the fusiform face area (FFA) and parahippocampal place area (PPA; Haxby et al., 2001). Although the roles of the dorsal and ventral visual streams appear to be distinct, there is some research that suggests the involvement of both the dorsal and ventral streams in shape detection, but only the dorsal stream for object location detection (Zachariou, Klatzky, & Behrmann, 2013). Presumably, spatial information would play a crucial part in the identification of visual object features used during object recognition, thereby activating both visual streams.

Drawing is a common form of image production, and is typically a visuo-manual task performed in conjunction with tool use, such as with a pencil, stick, or paintbrush. Image production is a dynamic sensorimotor process that leaves behind a visual product on some surface, a progression of image development that we have called “emanation” (Yuan & Brown, 2014; see Chapter 2). Gowen and Miall (2007) used functional magnetic resonance imaging (fMRI) to compare the brain’s responses to drawing from memory with tracing an externally cued form. Participants were required to trace or draw simple geometric figures (squares, triangles, and circles) using only their eyes, only their hands (without moving their eyes), or both the eyes and hands, a task very close to naturalistic drawing. The control task was a simple fixation cross upon which participants had to focus their attention. The authors found that the pre-supplementary motor area (pre-SMA), frontal eye fields (FEF), and cerebellum were more active during naturalistic drawing compared to tracing, whereas tracing showed greater activity in the IPS and the superior occipital region close to our designation of area V3A (Yuan & Brown, 2014). In another study, Ogawa and Inui (2009) used computerized copying and tracing tasks in order to manipulate memory guidance versus visual guidance during drawing. Participants were given a visual model curve and had to use a computer mouse to either copy or trace over the curve. When compared to tracing, copying resulted in more activity in the FEF, SMA, IPS, V5/MT+, and V3A. Memory guidance activated frontal motor-planning areas and parietal spatial-processing brain areas, whereas visual guidance activated occipital visuospatial areas in the vicinity of V3A and V5/MT+.

Up until very recently, much of the neuroscientific literature investigating drawing production had many drawbacks that limited the visual feedback given to participants in the scanner. This included tasks that involved air-drawing (Makuuchi, Kaminaga, & Sugishita, 2003), finger-drawing (Ino, Asada, Ito, Kimura, & Fukuyama, 2003), or even covert mental imagery without any overt behavioural activity (Harrington, Farias, Davis, & Buonocore, 2007; Harrington, Farias, & Davis, 2009; Suchan et al., 2002). I address the implications of these limitations in Chapter 2. All in all, image perception and production require a combination of both ventral and dorsal visual processing stream areas, including areas in or near the SPL, IPS, and IPL, as well as visuospatial regions such as V5/MT+ and V3A.

1.4 Multimodal processing

Although there is a large body of research focused on unimodal processing for each of the major communicative modalities just mentioned, human communication is rarely carried out in only one modality. For example, speech is most frequently supplemented with actions or visuals to convey alternate forms of information, transforming the communicative process into one of audio-visual or audio-gestural cognition. Hickok and colleagues have described an area called the Spt that shows involvement in processing speech and motor functions (Hickok & Poeppel, 2000; Hickok & Poeppel, 2004; Hickok, Buchsbaum, Humphries, & Muftuler, 2003). Other perisylvian areas, such as the parietal operculum complex (OP1-4; Eickoff et al., 2010) and even the posterior insula (Kurth et al., 2010) also show auditory-motor functions and connectivity with both auditory and motor areas. Progressing posteriorly in the temporal lobe, the pSTS seems to show involvement in both visual-motor and auditory-motor processes (Van Overwalle & Baetens, 2009; Xu, Gannon, Emmorey, Smith, & Braun, 2009). The pSTS also plays a role in processing communicative intent across modalities (Enrici, Adenzato, Cappa, Bara, & Tettamanti, 2011; Redcay, Velnoskey, & Rowe, 2016).

More recently, Marstaller and Burianová (2015) described shared brain areas during the production of speech, co-speech gestures, and pantomime. In their study, participants were cued with the response modality via icons that represented speech or gesture, and then given a cue noun of a common handheld tool. The task was to produce a verb, an action gesture, or a combination of the two in relation to the cued noun. The fMRI results showed that co-speech gestures did not activate additional brain areas that were not already activated in speech production, and that common brain areas shared between pantomime gestures and co-speech gestures were the MTG and temporo-parietal junction (TPJ). In addition to the unimodal perception and production areas, these multimodal areas are important to the understanding of cross-modal communication.

1.5 Research objectives

Although the major modalities of communication are speech, pantomime, and drawing, research interest in these three has not been equivalent. While there are many experiments investigating how people perceive and produce speech and pantomimes, there is relatively little research on how people draw, and not even a handful of experiments on cross-modal communication. In light of the paucity of drawing research that provided visual feedback to participants, the objectives of this dissertation are to characterize a basic drawing network in the human brain using a novel MRI-compatible drawing tablet (Chapter 2) and to follow up by exploring the structural connectivity of this network (Chapter 3). Additionally, the dissertation addresses the question of how the basic drawing network will respond to more-complex task demands, such as those in the production of narrative drawing, using a multimodal paradigm that also provides affordances to investigate the shared brain areas in multimodal narrative production (Chapter 4). I conclude with a general discussion of my findings (Chapter 5).

Chapter 2 - The neural basis of mark making: A functional MRI study of drawing

Ye Yuan & Steven Brown.

PLoS ONE, *2014, 9(10), e108628*

Keywords: drawing, fMRI, visual, image, picture, copy, geometric

2.1 Abstract

Compared to most other forms of visually-guided motor activity, drawing is unique in that it “leaves a trail behind” in the form of the emanating image. We took advantage of an MRI-compatible drawing tablet in order to examine both the motor production and perceptual emanation of images. Subjects participated in a series of mark making tasks in which they were cued to draw geometric patterns on the tablet's surface. The critical comparison was between when visual feedback was displayed (image generation) versus when it was not (no image generation). This contrast revealed an occipito-parietal stream involved in motion-based perception of the emerging image, including areas V5/MT+, LO, V3A, and the posterior part of the intraparietal

sulcus. Interestingly, when subjects passively viewed animations of visual patterns emerging on the projected surface, all of the sensorimotor network involved in drawing was strongly activated, with the exception of the primary motor cortex. These results argue that the origin of the human capacity to draw and write involves not only motor skills for tool use but also motor-sensory links between drawing movements and the visual images that emanate from them in real time.

2.2 Introduction

Images refer to visual patterns created on surfaces, generally on flat surfaces (such as canvases or cave walls) but also on three-dimensional objects (such as human bodies or ceramic vases). This includes the products of both drawing and writing as well as a third category of images that Elkins [1] refers to as “notation”, including musical notation, mathematical formulas, and a host of other images that are categorized as neither pictures (drawing) nor words (writing).

From a motor-control perspective, drawing can be thought of as being similar to most other forms of visuo-manual activity, including the ones that neuroscientists typically study, such as reaching, grasping, object manipulation, pointing, and gesturing [2]. It involves visual guidance of hand movement towards a target though hand/eye coordination. Drawing also shows similarities with forms of tool use (e.g., joystick movement), as it invariably involves use of a drawing tool. Drawing, therefore, is similar to other forms of manual activity in that it is a dynamic sensorimotor process.

However, it differs from all these forms of motor activity in an important respect: it leaves a trail behind. In other words, an image emerges through the process of performing the motor activity. We will use the term "emanation" to refer to this emergence of an image as drawing progresses. Emanation applies to writing as much as drawing, since writing too is a form of image generation. Therefore, while pictures themselves are generally considered to be static objects – making them standard stimuli for studies of "neuroaesthetics" [3] – drawing itself is a dynamic process both in the sense that it requires visually-guided coordination of the eyes, hands and body, and more uniquely that it involves emanation of an image, in other words the intentional laying down of a trail on a surface as the movement occurs.

Many previous neuroimaging studies of drawing have had a strong limitation in that subjects' perception of emanation during drawing was limited by a lack of visual feedback in the scanning situation. The major modalities for drawing in these studies included drawing in the air [4] and drawing on a pad sitting on the body or on a writing board using a drawing tool [5-8, 9] or a finger alone [10]. Several studies have had subjects draw covertly using mental imagery alone [5, 11, 12]. In some studies of overt production, the eyes were kept closed during the drawing task [7, 10]. In certain studies in which the eyes were open, no visual activations were reported [4]. Curiously enough, the very first imaging study of drawing [13] performed a tracing task using positron emission tomography (PET) in which subjects were indeed able to see their tracings via a back-projection system (see also [9, 14]). Studies of air drawing and imagined drawing, beyond having feedback limitations, provide no behavioral data on subject performance (whereas pad studies produce drawings).

More recently, a small number of drawing studies have used MRI-compatible drawing devices that are able to provide visual feedback to subjects during drawing, thus permitting the capacity to perform tracing tasks in the MRI scanner. These include the use of an MRI-compatible mouse [15-16] and drawing tablet [17-20]. The current study took advantage of the MRI-compatible drawing tablet devised by Tam et al. [19]. This tablet not only provides visual feedback to subjects but furnishes a means of recording all the drawing movements of the subject, permitting video reconstruction of drawing trajectories and thus behavioral performance during drawing. It also provides a means of manipulating visual feedback to the subject during drawing, for example the ability to eliminate visual feedback, as was done in the present study (see also Thaler & Goodale [20]).

Drawing can occur in three principal ways: from memory, as copying, or as tracing. Whereas writing is almost always done from memory, drawing is done equally commonly as copying (for example, drawing a sitter's portrait) and as drawing from memory. In like form, most neuroimaging studies of drawing have had subjects generate images either from memory [6-8, 10, 14, 16, 21] or as a copying task [4, 5, 8, 9, 11, 14, 16, 18]. Tracing has been restricted to the few studies that have provided visual feedback to subjects in the scanner [13-16, 18].

The abovementioned neuroimaging literature for drawing has produced a reliable set of findings. This includes not only expected activations in the left primary motor cortex and right cerebellum for right-handed drawing but quite often activity in the posterior parietal cortex, including the cortex of the intraparietal sulcus (IPS). The IPS is involved in creating a transformation between retinotopic coordinates in visual space and egocentric motor coordinates in effector space, thereby supporting visual guidance of hand movement [22]. Such activity must be coordinated with eye movement as well, since eye position defines retinotopic position. Posterior parietal activations tend to either be either ipsilateral to the motor-cortex activations or bilateral. Other common activations have been found in the frontal eye fields (FEF), inferior frontal gyrus (IFG), and precuneus. Studies that have provided visual feedback to subjects have been the most informative since they have observed visual activations as well. For example, Ogawa and Inui [16] had subjects perform both copying and tracing of curved lines using an MRI-compatible mouse with visual feedback.

While tracing gave no residual activations beyond copying, copying gave additional activations in V1, V2, IPS, IFG and pre-supplementary motor area (SMA), most likely reflecting the greater spatial demands of copying compared to tracing in recreating the visual properties of the drawn object.

The principal objective of the current study was to examine the neural basis of image generation and its emanative component through the performance of mark making tasks while taking advantage of the precision and flexibility conferred by using an MRI-compatible drawing tablet, not least the ability of subjects to see what they were drawing and for visual feedback to be manipulated. A critical comparison was between when visual feedback was displayed on the projected screen (image generation) versus when it was not (“blind drawing”, the situation of many previous imaging studies of drawing). This contrast should allow us to isolate brain areas important for emanation in drawing. In a perceptual control condition, we had subjects passively view an animation of an image emerging in time on the projected surface. This motion-perception task should likewise reveal brain areas important for emanation. Finally, we included the additional condition of copying in order to examine a drawing task that has a stronger spatial-processing demand than a task done from memory. We predicted that, unlike most previous studies of drawing, we would observe activations in parts of the brain involved in motion perception, eye movement, and hand/eye coordination, allowing us to establish a basic sensorimotor network for drawing in the brain, one that includes neural areas for emanation as central components but that are missing in all previous studies in which visual feedback was lacking.

2.3 Materials and methods

2.3.1 – Subjects

Fifteen right-handed subjects (9 females, mean age 25 years old, range 18-35 years old) participated in the study after giving their informed consent (McMaster Research Ethics Board, McMaster University). Handedness was tested using the Edinburgh handedness inventory [23]. Subjects had normal or corrected-to-normal vision and no history of neurological disorders, psychiatric illness, alcohol or substance abuse, and were not taking psychotropic medications. No subject required corrective lenses during the MRI experiment. Subjects received monetary compensation for their participation.

2.3.2 – Apparatus

Drawing was performed on an MRI-compatible (i.e., non-ferromagnetic) drawing tablet developed by Tam et al. [19], as connected to a Hewlett Packard Pavilion dv5 laptop computer running E-Prime 2.0 (Psychology Software Tools, Sharpsburg, PA). Figure 1 of Tam et al. [19] demonstrates the set-up of the tablet and its placement above a subject in an MRI scanner.

The tablet consists of a resistive touch-screen connected to an elevated support platform. The tablet was custom-made to fit the specifications of the GE scanner-bed used in this study. A controller box served as an interface between the tablet and the laptop computer that was used for both stimulus presentation and the recording of drawing data. The dimensions of the screen were 12.8 cm width by 9.2 cm height. Drawing was made using a simple plastic stylus roughly the size and weight of a ballpoint pen. When subjects were placed in the scanner, the drawing tablet was fitted close to the body surface so as to permit easy access with the hands. The right hand was used for drawing (all subjects were right handed), and the left hand rested on the left side of the support platform. A series of calibration tasks was performed for each subject in order to ensure that the projected image was visible to them and that their drawings were well contained within the field of view of the LCD projector.

It is important to note that no drawings actually appeared on the tablet's surface. All drawings were seen via a mirror positioned in the visor of the head coil. The LCD projector presented images onto this visor, and the light was reflected by the mirror to the subject's eyes. This gave the veridical impression to subjects that their drawings were occurring on the tablet's surface. However, this occurred indirectly through information from the computer screen projected onto the visor through the LCD projector. In addition, due to the arrangement of the tablet in the scanner bed, subjects were not able to see the drawing tool during drawing. Hence, the only dynamic visual stimulation that they received during drawing came from the emanating image and not from perceived movement of the stylus tip or their own hand.

2.3.3 – Stimuli

Two categories of stimuli were used in the production tasks (see Figure 1): 1) geometric patterns, of which there were three types (spirals, zigzags, and serpentines), and 2) embellished geometric patterns of the same three types (i.e., geometric patterns with added loops, used for the copying task).

2.3.4 – Tasks

Each participant took part in a one-hour training session on a day prior to the MRI scan in order to become proficient at using the drawing tablet while minimizing head movement as well as to practice the tasks to be performed in the scanner. Training was performed in a simulated scanner environment in which subjects were supine and had the tablet positioned across their lap. Subjects were positioned at a comfortable viewing position below a computer monitor that was mounted to a hinged arm on the wall. Pre-recorded scanner noise was played in the background as subjects performed each training task. During the scanning session, subjects performed each trial as an alternation between 20s periods of fixation and

20s periods of task. Each scan lasted 4 min and consisted of 6 trials of the same condition. During the fixation periods, a black fixation-cross was projected onto the center of a grey background. There were four scans altogether, one for each of the following conditions: perception, mark making, blind drawing, and copying. All stimuli were presented using E-Prime 2.0 running on a Hewlett Packard Pavilion dv5 laptop.

Participants performed the following three drawing tasks in random order. 1) Mark making: participants were prompted for 2s with the written name of a geometric pattern (zigzag, spiral, or serpentine) as well as an arrow indicating the direction in which to draw it (i.e., leftward or rightward). This occurred in the center of the screen. The prompt was then removed, and the subject drew the pattern from memory on the right half of the screen for the 18s remaining in the epoch. The subject was instructed to draw for the duration of the 18s epoch. If the subject reached the edge of the defined drawing space while drawing zigzags or serpentines, they were told to double back in the other direction until the task epoch was over. For all conditions, the drawing direction was balanced across stimuli. Subjects were unable to see either the stylus or their own hand. Thus, the only visual feedback available to them during the drawing tasks was the emanating image. 2) Blind Drawing: this was exactly the same as the mark making task except that the line color for drawing was changed to the background color of the display, thereby removing all visual feedback. This created a condition in which the subject could not see their drawing while making it, a situation common to many neuroimaging studies of drawing (see Introduction). 3) Copying: participants were presented on the left half of the screen with a visual model that was to be copied on the right half of the screen. The model remained visible throughout the task epoch. The stimuli were not the simple geometric stimuli used in the other mark making conditions but rather embellished geometric patterns in which loops were added to the geometric patterns (see Figure 1). The reason for this change was that pilot testing showed that using standard geometrics allowed subjects to ignore the features of the model and simply create the patterns from memory, just as they had in the mark making condition. The introduction of embellishments was a necessary step to keep the subject’s attention focused on the visual features of the model. For each stimulus, a starting point for copying was indicated on the model so as to balance drawing direction across stimuli. The full set of six copying stimuli is presented in Document S1. Finally, 4) a Perception task was performed in which subjects passively viewed animations of abstract line drawings unfurling on the projected screen over the course of 20s. Since pilot testing showed surprisingly widespread brain activations for this task, we had subjects perform it first so as to reduce any contamination of actual drawing on perception. The production tasks were then randomized among themselves after the perception task. For all drawing tasks, motor activity – and thus visual feedback – was limited to the right half of the tablet. Subjects were free to move their eyes during all conditions in order to make the drawing tasks naturalistic.

2.3.5 – Image acquisition

Magnetic resonance images were acquired with a GE Medical Systems Signa Excite 3-Tesla MRI at the Imaging Research Centre at St. Joseph's Healthcare Hamilton. The subject's head was firmly secured in the head coil using foam pads placed around the ears. Ear plugs were used to help block out scanner noise.

Functional images sensitive to the blood-oxygen-level-dependent (BOLD) signal were collected with a gradient-echo echo planar imaging (EPI) pulse sequence using standard parameters (TR = 2000 ms, TE = 45ms, flip angle = 90o, 31 slices, 4 mm slice thickness, no gap, matrix size = 64 x 64, field of view = 24cm, voxel size = 3.75 mm x 3.75 mm x 4 mm), effectively covering the whole brain. All functional scans lasted 4 min, resulting in the collection of 120 brain volumes per scan.

High-resolution, T1-weighted structural images were acquired in order to register functional activity onto brain anatomy. The scanning parameters were 3D-FSPGR, IR-prepped, Ti = 450 ms, flip angle = 12 degrees, TR = 7.5 ms, TE = 2.1 ms, field of view = 240 mm x 180 mm, slice thickness = 2 mm, acquisition matrix 320 x 192, 1 average (NEX = 1), receiver bandwidth

= 31.25 kHz, data was interpolated to a 512 x 512 matrix, and the number of slices doubled during reconstruction, resulting in 164 slices.

2.3.6 – Data analysis

Functional images were reconstructed offline, and the scan series was realigned and motion-corrected using BrainVoyager QX 2.4 (Brain Innovation, Maastricht). Motion-correction analysis revealed that subjects displayed very little head movement. Translational and rotational corrections did not exceed an acceptable level of 1.5 mm and 1.5 degrees, respectively, for any subject. During the preprocessing stage, a temporal high-pass filter was applied at a frequency of 0.0078 Hz, or 2 cycles per scan, using the GLM-Fourier algorithm. 3D spatial smoothing was performed using a Gaussian filter with a FWHM kernel size of 4 mm. Following realignment, each functional scan was normalized to the Talairach template [24]. The BOLD response for each task-block was modeled as the convolution of a 20s boxcar with a synthetic hemodynamic response function composed of two gamma functions. Beta weights associated with the modeled hemodynamic responses were computed to fit the observed BOLD-signal time course in each voxel for each subject using the General Linear Model, as implemented in BrainVoyager QX 2.4. The six head-motion parameters were included as nuisance regressors in the analysis. Each subject's data was processed using a fixed-effects analysis, corrected for multiple comparisons using a Bonferroni correction at a threshold of $p<0.05$. Contrast images for each subject were brought forward into a random-effects analysis, where a false discovery rate (FDR) of $p<0.01$ was employed as a correction for multiple comparisons, with a cluster threshold of $k = 25$. Group data were registered onto the inflated brain of one of the subjects within the study (Subject 4), as generated using Brain Voyager. Talairach coordinates were extracted using NeuroElf (neuroelf.net).

2.4 Results

Figure 1 provides examples of the stimuli used for the mark making and copying tasks, as well as representative drawn responses. Document S2 provides representative examples of drawn responses for the blind drawing condition. Figure 2a shows the activation pattern for mark making in contrast with fixation. The Talairach coordinates of the activations are listed in Table 1. Prominent activations related to motor control of the right hand and forearm were found in the left sensorimotor cortex and right posterior cerebellum (lobule V). Additional motor activations were found bilaterally in the frontal eye fields medial to the primary hand activations. Next, while no activation was found in the primary visual cortex, strong activations were found bilaterally in the motion-perception area V5/MT+ (BA 19) and in area LO posterior to it (BA 18). LO is a form-processing area but is thought to be important in processing form from motion [25]. Higher-level motion-related visual activations were seen dorsomedially in area V3A (BA 19) in the right hemisphere. As mentioned in the Methods sections, subjects in the scanner were unable to see the drawing tool, and hence all visual activations reported in this study result from perception of the emanating image, not perception of the moving tool, hence permitting a disambiguation of the two sources. No activity was detected in another well-studied motion-perception area, the posterior superior temporal sulcus (pSTS), which is more associated with the perception of body motion. Parietal activations were found in the inferior parietal lobule (IPL), IPS bilaterally, and in the superior parietal lobule (SPL) in the left hemisphere (BA 7), directly posterior to the sensorimotor cortex. The IPS activations included both the posterior regions referred to by Swisher et al. [26] as IPS1 and IPS2 and the anterior regions referred to as IPS3 and IPS4. Finally, bilateral activations were seen in the middle frontal gyrus (BA 6). In sum, mark making defined the basic motor-sensory components of the drawing network, reflecting the dynamic visual/hand and visual/eye coupling that occurs during the generation of marks and the emergence of images.

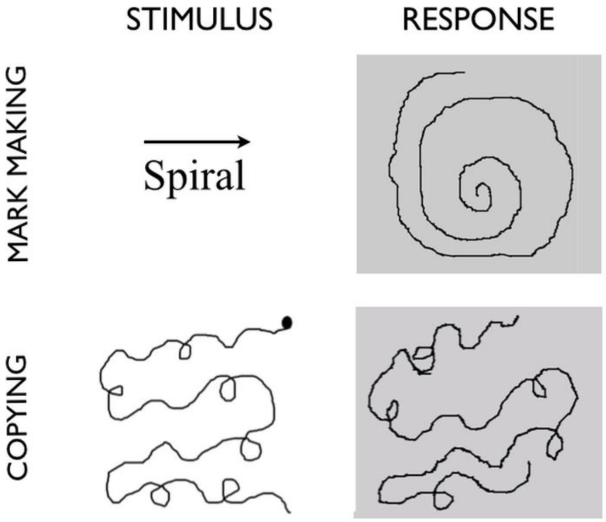


Figure 2.1: Tasks and stimuli for the study.

Representative stimuli and responses for two of the drawing conditions. In mark making, the name of a geometric figure and a direction for drawing it are presented. Blind drawing (not shown in the figure) has exactly the same stimuli, but no response is observable on the drawing tablet. In copying, an object and a starting point for drawing are indicated, and the subject creates a copy of the object in the space to the right, with the object continually in view.

Turning off visual feedback while doing mark making created a condition of blind drawing, which places drawing under purely proprioceptive control and which serves as a motoric control for mark making. Figure 2b shows the contrast of mark making vs. blind drawing (the reverse contrast gave no signal). The Talairach coordinates of these activations and of those for blind drawing vs. fixation are present in Table 1. As expected, all of the motor activity in the left sensorimotor cortex and right cerebellum was eliminated in this subtraction due to the matched motoric nature of the tasks. What remained was the occipito-parietal visual-motion network, including areas V5/MT+, LO, V3A, and IPS1/2, with a strong right-dominant pattern. This group of areas represents the best neural correlate of the phenomenon of emanation occurring during drawing-based image generation.

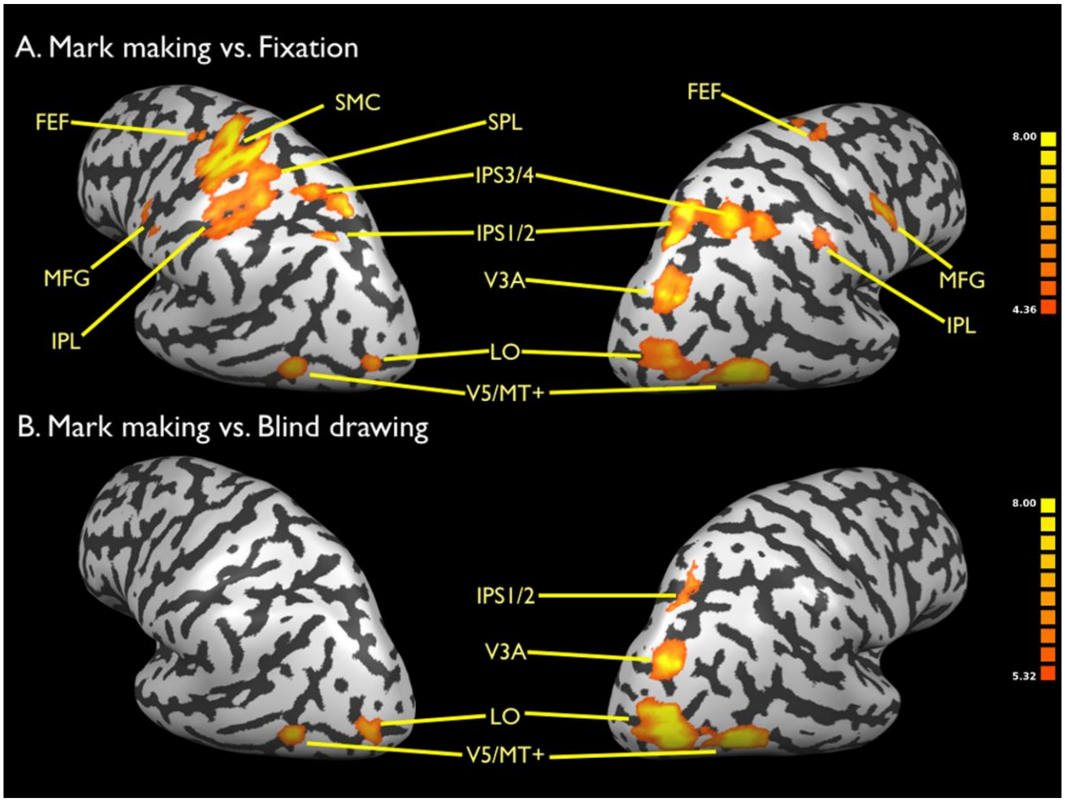


Figure 2.2. Brain activations for mark making

a) Mark making vs. fixation. b) Mark making vs. blind drawing. Data are corrected for multiple comparisons using FDR $p < 0.01$. Activations shown in Figures 2-5 are rendered onto an inflated brain of one of the subjects in the study (Subject 4) as normalized into Talairach space. The color bars in Figures 2-5 reflect the t score of the activated voxels for a given contrast. Abbreviations: FEF: frontal eye fields; IPS1/2: segments 1 and 2 of the intraparietal sulcus; IPS3/4: segments 3 and 4 of the intraparietal sulcus; MFG: middle frontal gyrus; MT: middle temporal; SMC: sensorimotor cortex; SPL: superior parietal lobule.

Table 1	**Mark making**				**Blind drawing**				**Mark making vs. Blind drawing**			
	x	y	z	t	x	y	z	t	x	y	z	t
Right Hemisphere												
Frontal Lobe												
Middle frontal gyrus (BA6)	51	-1	37	12.4	63	2	46	7.3				
Precentral gyrus (BA 4)	27	-16	58	6.9								
Frontal eye fields (BA 6)	18	-13	58	6.7								
Inferior frontal gyrus (BA 6)					51	-1	28	6.6				
Parietal Lobe												
IPS1/2 (BA 7)	21	-67	52	11.8					27	-67	58	10.1
IPS1/2 (BA 7)									24	-58	46	7.4
IPS3/4 (BA 7)	27	-49	49	8.6	30	-43	46	6.3				
Inferior parietal lobule (BA 40)	51	-28	40	8.8								
Precuneus (BA 7)	12	-73	61	8.7								
Inferior parietal lobule (BA 40)	39	-40	43	8.4	51	-31	43	6.6				
Occipital Lobe												
V5/MT+ (BA19)	42	-67	-5	12.7					42	-67	-5	16.7
V3A (BA19)	30	-76	25	11.2					27	-76	25	9.6
LO (BA18)	33	-79	1	9.3					33	-76	-5	11.5
Temporal Lobe												
Middle temporal gyrus (BA 21/37)									45	-49	-5	9.2
Cerebellum												
Lobule V	24	-52	-14	8.2								
Left Hemisphere												
Frontal Lobe												
Sensorimotor cortex (BA 4/3)	-30	-28	67	12.7	-33	-28	67	9.4				
Middle frontal gyrus (BA 6)	-57	-1	37	7.7								
Frontal eye fields (BA 6)	-12	-16	55	7.7	-3	-16	58	7.5				
Parietal Lobe												
Somatosensory cortex (BA3)	-39	-28	49	12	-33	-28	49	8.5				
Postcentral gyrus (BA 2)					-36	-40	67	8.3				
IPS1/2 (BA7)	-21	-67	46	11.3								
Superior parietal lobule (BA 7)	-24	-67	55	8.2								
Paracentral lobule (BA 5)	-9	-31	49	7.7								
IPS3/4 (BA7)	-30	-49	52	7.6								
Inferior parietal lobule (BA 40)	-48	-34	40	7.6	-51	-37	46	5				

(Cont'd on next page)

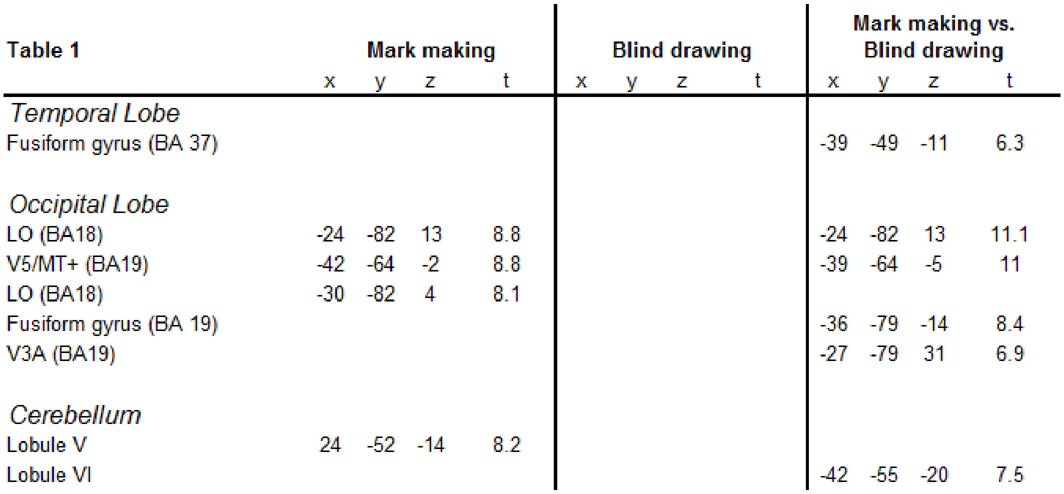


Table 1	**Mark making**				**Blind drawing**				**Mark making vs. Blind drawing**			
	x	y	z	t	x	y	z	t	x	y	z	t
Temporal Lobe												
Fusiform gyrus (BA 37)									-39	-49	-11	6.3
Occipital Lobe												
LO (BA18)	-24	-82	13	8.8					-24	-82	13	11.1
V5/MT+ (BA19)	-42	-64	-2	8.8					-39	-64	-5	11
LO (BA18)	-30	-82	4	8.1								
Fusiform gyrus (BA 19)									-36	-79	-14	8.4
V3A (BA19)									-27	-79	31	6.9
Cerebellum												
Lobule V	24	-52	-14	8.2								
Lobule VI									-42	-55	-20	7.5

Table 2.1 – Talairach coordinates for mark making, blind drawing, and the contrast between mark making and blind drawing

Stereotaxic coordinates are in millimeters along the left-right (x), anterior-posterior (y), and superior-inferior (z) axes. In parentheses after each brain region is the Brodmann area, except for the cerebellum, in which case the anatomical labels of Schmahmann et al. (2000) are used. Abbreviations: IPS: intraparietal sulcus; LO: lateral occipital complex; MT: middle temporal.

Next, activations for copying versus fixation are shown in Figure 3a (lateral view) and Figure 4a (medial view). The Talairach coordinates for these activations are presented in Table 2. While copying showed activations in the same set of regions as mark making, the global level of activation for copying was much more intense than that for mark making, including in areas involved in emanation. In addition, two new systems were present for copying that were not seen in memory-based drawing, and this was highlighted in the contrast of copying vs. mark marking in Figures 3b and 4b. First, activations were seen in the basal ganglia system, including bilateral putamen and ventral thalamus. This system, perhaps in combination with the right IFG, most likely mediates the imitative aspect of copying. Second, strong activity was seen in the primary visual cortex and surrounding areas (BA 17 and 18). Such areas were not detected in mark making. This lower-level visual activity most likely reflects the presence of the static model that the subject had to glance at repeatedly during the drawing process. This result is consistent with the findings of Ferber et al. [18] who, in a similar contrast between copying and drawing from memory, found greater activity in the cuneus and other inferior occipital regions for copying.

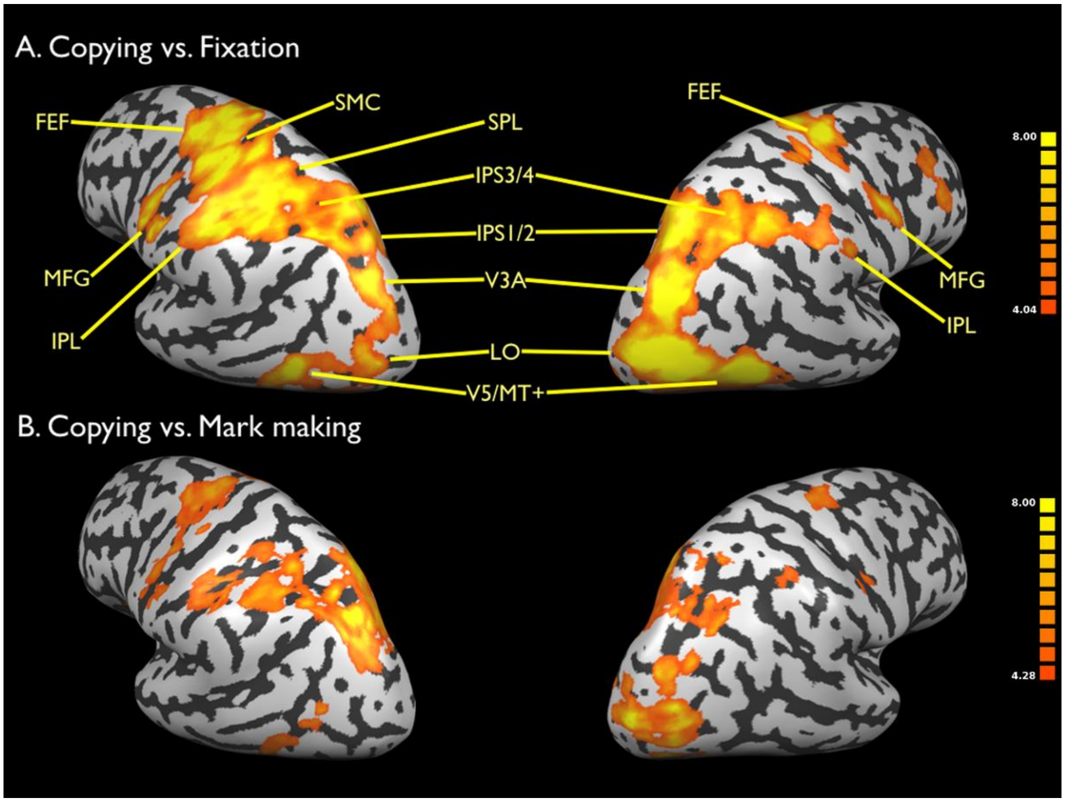


Figure 2.3 – Brain activations for copying: Lateral view

a) Copying vs. fixation. b) Copying vs. mark marking. Data are corrected for multiple comparisons using FDR $p < 0.01$. See legend to Figure 2 for abbreviations.

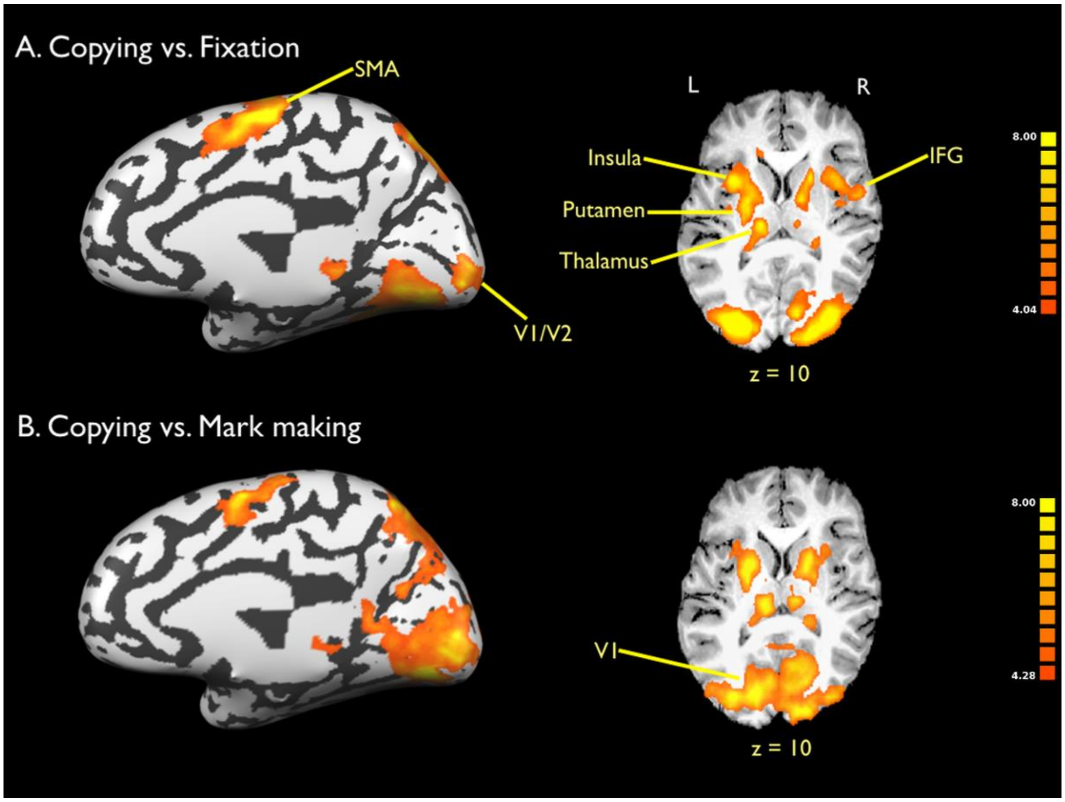


Figure 2.4 Brain activations for copying: Medial view

a) Copying vs. fixation. b) Copying vs. mark marking. Data are corrected for multiple comparisons using FDR $p < 0.01$. The left side of the slices is the left side of the brain, as indicated by the L (left) and R (right) symbols. The Talairach z coordinate is shown below the slices. Abbreviations: IFG: inferior frontal gyrus; SMA: supplementary motor area.

Table 2

	Copying				**Perception**			
	x	y	z	t	x	y	z	t
Right Hemisphere								
Frontal Lobe								
SMA (BA 6)	3	-4	58	12.1				
Superior frontal gyrus (BA 6)					30	-4	67	9.4
Frontal eye fields (BA 6)	24	-13	52	11.7	24	-10	58	8.9
Middle frontal gyrus (BA 6)	54	2	40	9.8	45	-4	40	10.4
	36	-1	34	9.8				
	45	-1	28	8.1				
Parietal Lobe								
IPS1/2 (BA 7)	21	-64	46	11.9	21	-64	49	13.8
	30	-61	55	10.9				
IPS3/4 (BA 7)	33	-49	49	9.1	30	-46	46	11.6
Superior parietal lobule (BA 7)	18	-73	61	9.6	24	-70	61	10
Inferior parietal lobule (BA 40)	39	-37	40	10.4	39	-40	46	11.5
					54	-28	37	9.7
Somatosensory cortex (BA 1/2)					39	-37	73	8.6
Precuneus (BA 7)					6	-70	55	8.1
Occipital Lobe								
LO (BA 19/18)	24	-85	10	15.1	33	-79	7	10.5
	36	-79	1	14.9				
	39	-70	-5	10.6				
V3A (BA 19)	30	-70	22	13.7	30	-76	22	15.5
Lingual gyrus: V2/V1 (BA 18)	27	-82	1	12.7	15	-85	-2	9.4
	21	-73	1	11.8				
V5/MT+ (BA 19/37)	45	-58	-5	11.9	42	-64	-2	13.3
					45	-52	-5	11.5
Fusiform gyrus (BA 19)	33	-61	-11	10.4	24	-76	-11	13.2
Subcortical								
Pulvinar					21	-28	1	9.7
Caudate nucleus	21	14	13	9.6				
Putamen	15	-1	13	9.5				
Caudate nucleus	24	-7	25	8.5				
Globus pallidus	18	-7	4	8				
Cerebellum								
Lobule VI	6	-67	-14	8.2				

(Cont’d on next page)

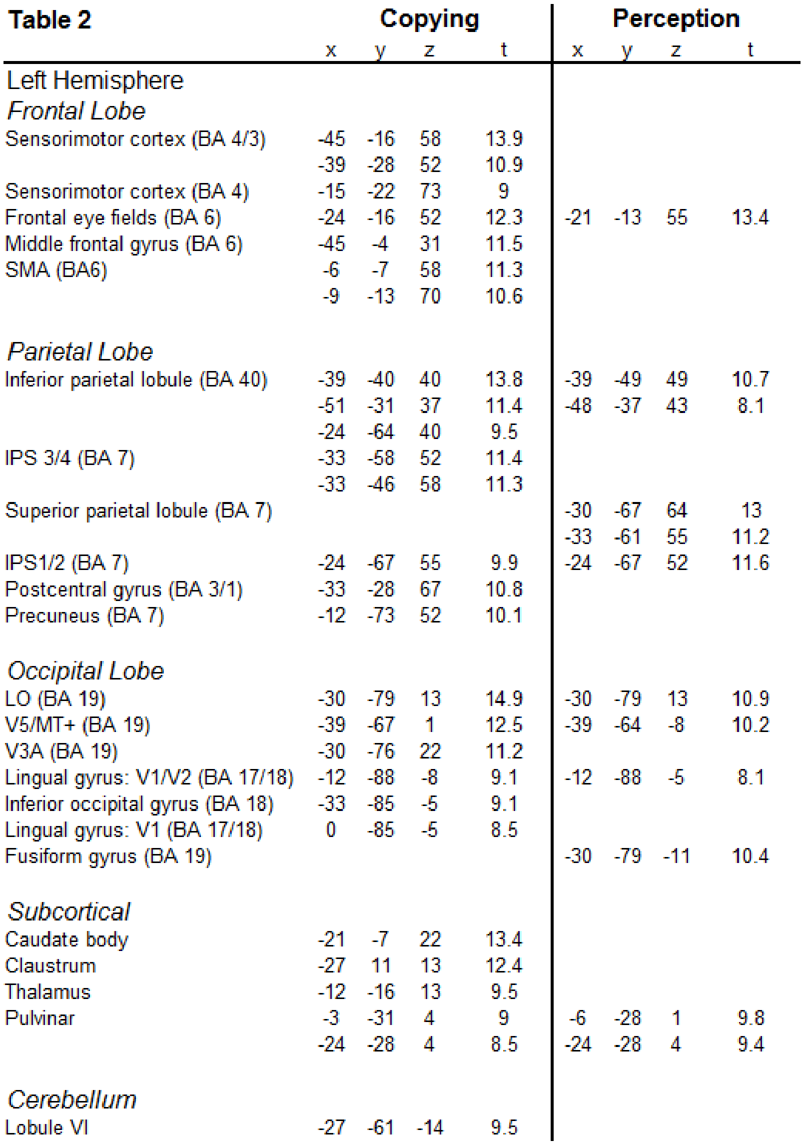


Table 2	**Copying**				**Perception**			
	x	y	z	t	x	y	z	t
Left Hemisphere								
Frontal Lobe								
Sensorimotor cortex (BA 4/3)	-45	-16	58	13.9				
	-39	-28	52	10.9				
Sensorimotor cortex (BA 4)	-15	-22	73	9				
Frontal eye fields (BA 6)	-24	-16	52	12.3	-21	-13	55	13.4
Middle frontal gyrus (BA 6)	-45	-4	31	11.5				
SMA (BA6)	-6	-7	58	11.3				
	-9	-13	70	10.6				
Parietal Lobe								
Inferior parietal lobule (BA 40)	-39	-40	40	13.8	-39	-49	49	10.7
	-51	-31	37	11.4	-48	-37	43	8.1
	-24	-64	40	9.5				
IPS 3/4 (BA 7)	-33	-58	52	11.4				
	-33	-46	58	11.3				
Superior parietal lobule (BA 7)					-30	-67	64	13
					-33	-61	55	11.2
IPS1/2 (BA 7)	-24	-67	55	9.9	-24	-67	52	11.6
Postcentral gyrus (BA 3/1)	-33	-28	67	10.8				
Precuneus (BA 7)	-12	-73	52	10.1				
Occipital Lobe								
LO (BA 19)	-30	-79	13	14.9	-30	-79	13	10.9
V5/MT+ (BA 19)	-39	-67	1	12.5	-39	-64	-8	10.2
V3A (BA 19)	-30	-76	22	11.2				
Lingual gyrus: V1/V2 (BA 17/18)	-12	-88	-8	9.1	-12	-88	-5	8.1
Inferior occipital gyrus (BA 18)	-33	-85	-5	9.1				
Lingual gyrus: V1 (BA 17/18)	0	-85	-5	8.5				
Fusiform gyrus (BA 19)					-30	-79	-11	10.4
Subcortical								
Caudate body	-21	-7	22	13.4				
Claustrum	-27	11	13	12.4				
Thalamus	-12	-16	13	9.5				
Pulvinar	-3	-31	4	9	-6	-28	1	9.8
	-24	-28	4	8.5	-24	-28	4	9.4
Cerebellum								
Lobule VI	-27	-61	-14	9.5				

Table 2.2 – Talairach coordinates for copying and perception

Stereotaxic coordinates are in millimeters along the left-right (x), anterior-posterior (y), and superior-inferior (z) axes. In parentheses after each brain region is the Brodmann area, except for the cerebellum, in which case the anatomical labels of Schmahmann et al. [48] are used. Due to the excessive number of activation foci for copying and perception, we decided to eliminate foci with a t value less than 8.0. The data in Table 1, by contrast, includes activation foci with a t value as low as 6.3. Abbreviations: IPS: intraparietal sulcus; LO: lateral occipital complex; MT: middle temporal; SMA: supplementary motor area.

Finally, we examined drawing perception alone by showing subjects animations of emanating images of abstract patterns (Figure 5a). The Talairach coordinates of these activations are presented in Table 2. Since pilot testing had revealed that this condition gave very strong activity throughout the drawing network, we decided to place this condition first in the scanning session in order to reduce any carryover effects that might come from performing drawing itself. As with the pilot data, the group results showed very strong activity throughout the drawing network, with the exception of the primary motor cortex. Brain areas associated with emanation were strongly activated in this condition, again with a right-hemisphere dominance, just as in production. Figure 5b shows the subtraction of mark making vs. perception. As can be seen, this subtraction eliminated virtually all of the activations for mark marking (compared with Figure 2a), except for the primary motor cortex (Talairach coordinates -36, -25, 55) and contralateral cerebellum (Talairach coordinates 15, -52, -14). Another way of thinking about this subtraction is that it basically resulted in the brain activity produced by blind drawing (see Table 1).

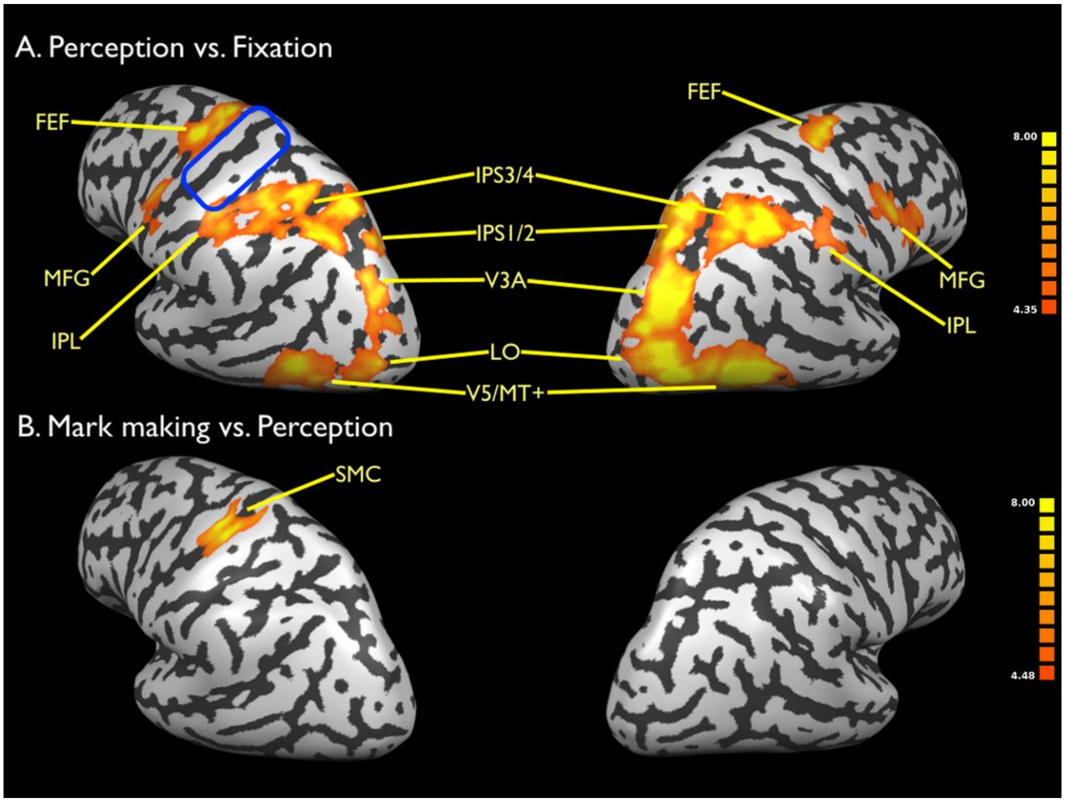


Figure 2.5 – Brain activations for motion perception

a) Perception vs. fixation. b) Mark making vs. perception. Data are corrected for multiple comparisons using FDR $p < 0.01$. The blue oval in panel a indicates the region of the sensorimotor cortex for copying not activated in motion perception. See legend to Figure 2 for abbreviations.

2.5 Discussion

We used functional MRI to explore the neural basis of the generation of images through mark making, including its defining property of emanation. Our use of an MRI-compatible drawing tablet allowed us to manipulate visual feedback, in contrast to many previous studies of drawing, where subjects obtained no feedback of their drawing activity. The blind drawing task essentially mimicked the situation of all overt drawing studies in which subjects did not have access to visual feedback (i.e., through drawing on a pad or drawing in the air) and hence did not perceive image formation. Providing visual feedback to subjects using the projected display created the more naturalistic situation of subjects perceiving images as they generated them in real time, hence allowing us to identify components of the motion-perception system of the brain's dorsal visual stream that mediate the perception of the emanating image.

The process of mark making from memory, involving the production of uninterrupted geometric patterns, defined the basic components of the drawing network, with brain areas involved in hand movement (M1, SMA, cerebellum), eye movement (FEF), visual motion perception (V5/MT+, V3A, LO), and sensorimotor coupling (IPS, IPL, SPL). Activity in this system as a whole was modulated quantitatively by the allocentric requirements of the drawing task, being lowest for blind drawing and highest for copying and perception. Emanation was associated with an occipito-parietal stream along the posterior aspect of the brain, extending dorsomedially from V5/MT+ to the posterior IPS, and encompassing the motion-related area V3A. To our surprise, passive perception of emanation was an extremely strong stimulus for the drawing network, eliciting activity in the areas involved in motor planning, even though subjects were explicitly instructed to passively view the emanation presented to them. This finding might suggest that some motor planning processes are automatic and are not under conscious control. In contrast, blind drawing gave only the motor components of the system plus activity in the left SPL.

2.5.1 – Emanation: Pictures result from trailing

It is interesting to note that the neural system for drawing is strikingly similar to that for gesture production [27]. We can classify drawing movements as a form of instrumental (transitive) gesture. From a cognitive standpoint, drawing might simply be gesturing that leaves a trail behind. Studies of drawing in which subjects moved their finger in the air ([4]; see also [28-30] for writing) are, in reality, studies of pantomime production. Ekman and Friesen [31] referred to gestures of this type as “pictographs”, making an allusion to drawing. Given the longstanding interest in the gestural origins of language through processes like pantomime [32-33], it might be the case that figurative drawing emerged from iconic gesturing processes like pantomime through the realization that such movements could leave a trail behind, perhaps first occurring using fingers or sticks in media like the earth or sand or even on the human body. In this regard, a key area that mediates emanation, namely V3A, appears to have undergone evolutionary modification in humans compared to monkeys [34]. This neural change might have relevance to the evolution of species-specific capacity for drawing in humans.

An important visual component of drawing compared to most other visuomotor tasks that people engage in is that visual information accumulates through trailing as the drawing progresses. The comparison between mark making and blind drawing revealed an occipito-parietal stream extending up the posterior aspect of the brain from V5/MT+ through V3A to the posterior IPS, with right hemisphere dominance during both production and perception. This stream was also active during copying and the passive perception of emanating images. Since subjects could not see the drawing tool in our experimental set-up, visual emanation could only come from the image alone and not from perception of the hand or drawing tip, thereby disambiguating these various sources of visual stimulation. Activity in this occipito-parietal stream is thus a neural marker of emanation, as shown in other studies of drawing in which visual feedback was present during image generation due to the use of MRI-compatible devices. In particular, our results are concordant with the contrast between visual feedback and no visual feedback in Thaler and Goodale's [20] analysis of line drawing.

Another brain area important for high-level motion perception, namely the pSTS [35-36], was not active in any of the conditions in this study. This is in distinction to many studies of gesture perception, where the pSTS is commonly seen [37-38]. This supports the association of the pSTS with the perception of biological motion, namely the motion of articulated bodies that move in the manner typical of animals [39]. An ALE meta-analysis of action observation and imitation [27] reported activity in V5/MT+, pSTS and IPS, but not in V3A. These findings suggest that the emanation system is engaged more strongly by trailing than by the perception of hand or body movement alone. The preliminary conclusion from this is that the pSTS, but not V3A, is activated by the perception of others' gestures and actions, and that V3A, but not the pSTS, is activated by the perception of emanation during drawing and writing.

Why might V3A be a critical area for drawing emanation when it seems not to respond to the perception of biological motion? V3A is well known to be directionally-selective and to be responsive to coherent motion compared to random motion [40]. It might therefore be involved in extracting form from motion [41], a function that is of importance in drawing, since form unfolds over time through a motion-based process of trailing. For example, Ellamil et al. [17], in a study of creative drawing using the same drawing tablet employed in this study, found right V3A to be active during the generative phase of drawing compared to an evaluative phase that followed it. In addition, V3A has also been shown to be responsive to ego-motion, in other words self-motion through space [40-41]. Thus, in contrast to the pSTS's responsiveness to the motion of others, V3A, along with areas like V6 and the IPS [42], might be more responsive to the motion of oneself. The optic flow that is perceived during emanation in our experiment is paradoxical in that it is based neither on an object moving through space nor on the subject moving through space relative to a fixed spatial reference frame. Instead, it represents the outcome of self-generated movement and is thus a proxy for self-motion. Although our MRI set-up dissociated emanation from hand and tool movement, there is a strong correlation between the motions of the hand, drawing tool and the emerging image during naturalistic drawing. So, V3A activation in our experiment might represent a response to neither object motion nor self-motion per se but instead to self-generated motion. It is expected that the emanation system would be even more engaged if the hand and drawing tool were perceivable during drawing. Although subjects were not able to see their hand or the drawing tool in our set-up, it would be quite interesting to compare the effects of viewing hand or drawing-tool movement without emanation vs. the emanation without perception of hand/tool movement that occurred in our set-up. Finally, the V3A activations in our study were adjacent to a region called SPOC (superior parieto-occipital cortex [43]) that is implicated into visuomotor functions related to reaching, pointing, and grasping. This area seems to be involved in encoding motor affordances, such as the reachability of an object by the hand. This might have relevance not only to our tool-based drawing tasks but to our motion-perception condition as well, especially if this task was processed by subjects as a type of virtual drawing.

2.5.2 – Limitations

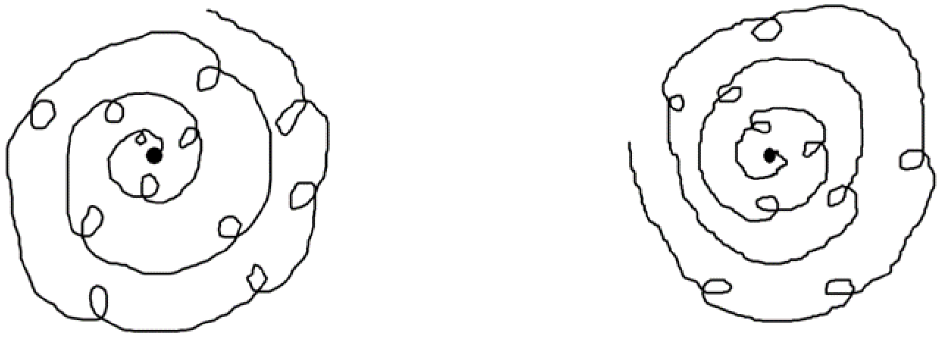
Two additional issues that would be important to explore in order to develop a more comprehensive understanding of the drawing system of the brain are figurativity and flattening. The current study used geometric figures as stimuli, but it is important to look at more-quotidian items as drawing stimuli as well, such as cars or houses. Harrington, Farias and Davis [11] performed a comparison between the copying of figurative vs. abstract models, but did so using visual imagery in the absence of actual drawing. Their results demonstrated overall similarity between these two categories of stimuli but significant differences in the fusiform gyrus, basal ganglia, and inferior frontal gyrus. The fusiform gyrus is part of the visual ventral stream, and so its preferential activation for figurative compared with abstract images might be indicative of the “object” status of figurative items in the object-recognition pathway of the inferior temporal lobe.

Finally, copy-based drawing of natural scenes requires a dimensional reduction (i.e., a flattening) from the three dimensions of visual perception to the two dimensions of the drawing space. Artists work extensively with monocular depth cues in order to create a sense of three-dimensional perspective in drawn images [44]. How they achieve this is not understood at the neural level. What is better understood is the perceptual system involved in depth processing [45]. Georgieva et al. [34] carried out a study of depth processing based on binocular-disparity cues and found an activation profile in motion-perception areas very similar to our occipito-parietal stream, including areas V5/MT+, V3A, V7/VIPS, and the posterior IPS. This implies an overlap in this system between disparity-based depth perception and the perception of trailing. This system might therefore be engaged in artists not only when they perceive depth in a model to be drawn but also when they transform its three-dimensional features into a two-dimensional form during the process of image generation.

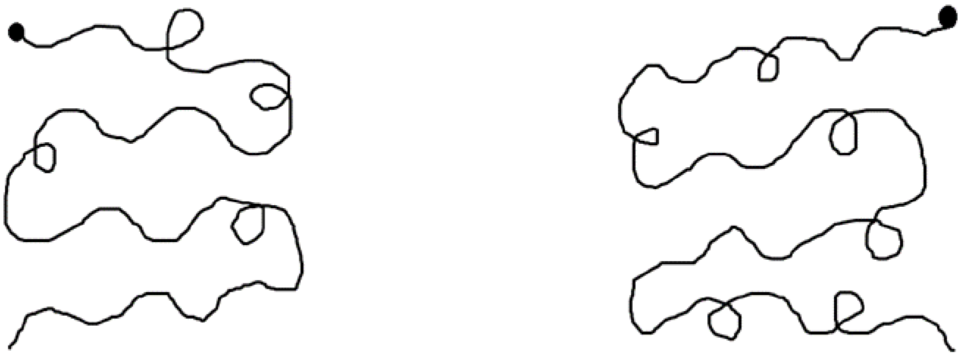
2.6 Acknowledgements

We thank Ellen Dissanayake for critical reading of the manuscript.

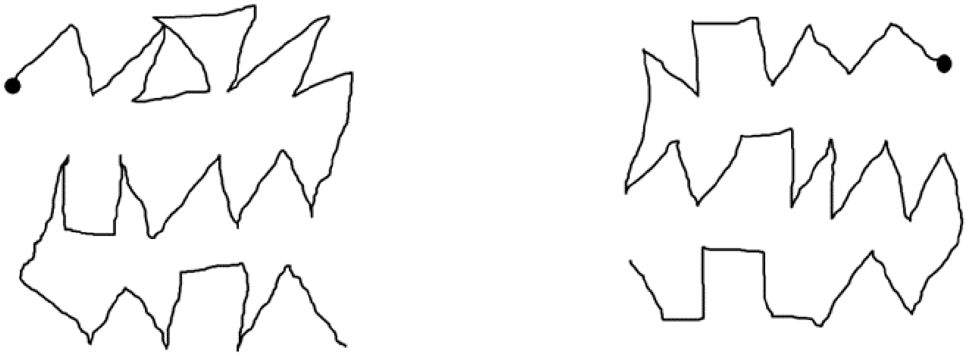
2.7 Supporting Information



Counterclockwise Spiral Clockwise Spiral



Left-to-right serpentine Right-to-left serpentine



Left-to-right zigzag Right-to-left zigzag

SI Figure 2.1 – Copy stimuli

Counterclockwise Spiral

Clockwise Spiral

Left-to-right serpentine

Right-to-left serpentine

Left-to-right zigzag

Right-to-left zigzag

SI Figure 2.2 – Blind drawing response

2.8 References

1. Elkins J (1999) The domain of images. New York: Cornell University Press.

2. Culham JC, Valyear KF (2006) Human parietal cortex in action. Curr Opin Neurobiol 16: 205-212.

3. Kawabata H, Zeki S (2004) Neural correlates of beauty. J Neurophysiol 91: 1699-1705.

4. Makuuchi M, Kaminaga T, Sugishita M (2003) Both parietal lobes are involved in drawing: A functional MRI study and implications for constructional apraxia. Cog Brain Res 16: 338-347.

5. Harrington GS, Farias D, Davis CH, Buonocore MH (2007) Comparison of the neural basis for imagined writing and drawing. Hum Brain Mapp 28: 450-459.

6. Kawashima R, Tajima N, Yoshida H, Okita K, Sasaki T, et al. (2000) The effect of verbal feedback on motor learning: A PET study. Neuroimage 12: 698-706.

7. Lerner A, Shill H, Hanakawa T, Bushara K, Goldfine A, et al. (2004) Regional cerebral blood flow correlates of the severity of writer’s cramp symptoms. Neuroimage 21: 904-913.

8. Miall RC, Gowen E, Tchalenko J (2009) Drawing cartoon faces: A functional study of the cognitive neuroscience of drawing. Cortex 45: 394-406.

9. Schaer K, Jahn G, Lotze M (2012) fMRI-activation during drawing a naturalistic or sketchy portrait. Beh Brain Res 233: 209-216.

10. Ino T, Asada T, Ito J, Kimura T, Fukuyama H (2003) Parieto-frontal networks for clock drawing revealed with fMRI. Neurosci Res 45: 71-77.

11. Harrington GS, Farias D, Davis CH (2009) The neural basis for simulated drawing and the semantic implications. Cortex 45: 386-393.

12. Suchan B, Yágüez L, Wunderlich G, Canavan AGM, Herzog H, et al. (2002). Neural correlates of visuospatial imagery. Beh Brain Res 131: 163-168.

13. Frutiger SA, Strother SC, Anderson JR, Sidtis JJ, Arnold JB, et al. (2000) Multivariate predictive relationship between kinematic and functional activation patterns in a PET study of visuomotor learning. Neuroimage 12: 515-527.

14. Gowen E, Miall RC (2007) Differentiation between external and internal cuing: An fMRI study comparing tracing with drawing. Neuroimage 36: 396-410.

15. Ogawa K, Inui T, Sugio T (2007) Neural correlates of state estimation in visually guided movements: An event-related fMRI study. Cortex 43: 289-300.

16. Ogawa K, Inui T (2009) The role of the posterior parietal cortex in drawing by copying. Neuropsychologia 47: 1013-1022.

17. Ellamil M, Dobson C, Beeman M, Christoff K (2012) Evaluative and generative modes of thought during the creative process. Neuroimage 59: 1783-1794.

18. Ferber S, Mraz R, Baker R, Graham SJ (2007) Shared and differential neural substrates of copying versus drawing: A functional magnetic resonance imaging study. Neuroreport 18: 1089-1093.

19. Tam F, Churchill NW, Strother SC, Graham SJ (2010) A new tablet for writing and drawing during functional MRI. Hum Brain Mapp 32: 240-248.

20. Thaler L, Goodale MA (2011) Neural substrates of visual spatial coding and visual feedback control for hand movements in allocentric and target-directed tasks. Front Hum Neurosci 5: 92.

21. Diciotti S, Cecchi P, Ginestroni A, Mazzoni LN, Pesaresi I, et al. (2010) MR-compatible device for monitoring hand tracing and writing tasks in fMRI with an application to healthy subjects. Concepts Magn Reson Part A 36A: 139-152.

22. Grefkes C, Fink GR (2005) The functional organization of the intraparietal sulcus in humans and monkeys. J Anat 207: 3-17.

23. Oldfield RC (1971) The assessment and analysis of handedness: The Edinburgh inventory. Neuropsychologia 9: 97-113.

24. Talairach J, Tournoux P (1988) Co-planar stereotaxic atlas of the human brain. New York: Thieme Medical Publishers.

25. Niemeier M, Goltz HC, Kuchinad A, Tweed DB, Vilis T (2005) A contralateral preference in the lateral occipital area: Sensory and attentional mechanisms. Cereb Cortex 15: 325-331.

26. Swisher JD, Halko MA, Merabet LB, McMains SA, Somers DC (2007) Visual topography of human intraparietal sulcus. J Neurosci 27: 5326-5337.

27. Caspers S, Zilles K, Laird AR, Eickhoff SB (2010) ALE meta-analysis of action observation and imitation in the human brain. Neuroimage 50: 1148-1167.

28. Katanoda K, Yoshikawa K, Sugishita M (2001) A functional MRI study on the neural substrates of writing. Hum Brain Mapp 13: 34-42.

29. Omura K, Tsukamoto T, Kotani Y, Ohgami Y, Yoshiwaka K (2004) Neural correlates of phoneme-to-grapheme conversion. Neuroreport 15: 949-953.

30. Sugihara G, Kaminaga T, Sugishita M (2006) Interindividual uniformity and variety of the “writing center”: A functional MRI study. Neuroimage 32: 1837-1849.

31. Ekman P, Friesen WV (1969) The repertoire of nonverbal behavior: Categories, origins, usage, and codings. Semiotica 1: 49-98.

32. Armstrong DF, Stokoe WC, Wilcox SE (1995) Gesture and the nature of language. New York: Cambridge University Press.

33. Coraballis MC (2003) From hand to mouth: The origins of language. Princeton, NJ: Princeton University Press.

34. Georgieva S, Peeters R, Kolster H, Todd JT, Orban GA (2007) The processing of three-dimensional shape from disparity in the human brain. J Neurosci 29: 727-742.

35. Grossman E, Donnelly M, Price R, Pickens D, Morgan V, Neighbor G, Blake R (2000) Brain areas involved in perception of biological motion. J Cog Neuro 12: 711-720.

36. Hein, G, Knight RT (2008) Superior temporal sulcus: It's my area, or is it? J Cog Neuro 20: 2125-2136.

37. Gallagher HL, Frith C (2004) Dissociable neural pathways for the perception and recognition of expressive and instrumental gestures. Neuropsychologia 42: 1725-1736.

38. Dick AS, Goldin-Meadow S, Solodkin A, Small SL (2012) Gesture in the developing brain. Dev Sci 15: 165-180.

39. Beauchamp MS, Lee KE, Haxby JV, Martin, A (2002) Parallel visual motion processing streams for manipulable objects and human movements. Neuron 28: 149-159.

40. Cardin V, Sherrington R, Hemsworth L, Smith AT (2012) Human V6: Functional characterisation and localisation. PLoS One 7: e47685.

41. Galletti C, Fattori P (2003) Neuronal mechanisms for detection of motion in the field of view. Neuropsychologia 41: 1717-1727.

42. Arnoldussen DM, Goossens J, van den Berg AV (2011) Adjacent visual representations of self-motion in different reference frames. Proc Nat Acad Sci 108: 11668-11673.

43. Fischer E, Bülthoff HH, Logothetis NK, Bartels A (2012) Human areas V3A and V6 compensate for self-induced planar visual motion. Neuron 73: 1228-1240.

44. Livingstone M (2002) Vision and art: The biology of seeing. New York: Harry N. Abrams

45. Backus BT, Fleet D, Parker AJ, Heeger DJ (2001) Human cortical activity correlates with stereoscopic depth perception. J Neurophysiol 86: 2054-2068.

46. Schmahmann JD, Doyon J, Toga AW, Petrides M, Evans AC (2000) MRI atlas of the human cerebellum. San Diego: Academic Press.

Chapter 3 – Structural connectivity of human visual-motion areas within the dorsal visual stream

Ye Yuan, Jennifer S. W. Campbell, Janine D. Mendola, G. Bruce Pike, & Steven Brown.

Unpublished manuscript.

Keywords: dorsal visual stream, diffusion, human brain, connectivity

3.1 Abstract

The visual system of primates is divided into a ventral "what" pathway for object recognition, and a dorsal "how" pathway for visual guidance of motor activity. Our recent functional MRI study (Yuan & Brown, 2014) identified a series of visual-motion areas as being critical for visual guidance of hand movement during drawing, including V5/MT+, V3A, and the posterior part of the intraparietal sulcus (pIPS). Despite the established importance of these areas for motion perception and visual guidance of action, their structural connectivity is poorly understood. Using diffusion imaging and probabilistic tractography, we identified tracts connecting V5/MT+, V3A/B, and pIPS.

These results identify novel structural connections underlying visual-motion areas within the dorsal stream of the human brain, and suggest a model of information flow during visuomotor tasks like drawing.

3.2 Introduction

Humans are the most dexterous of primate species. Object manipulation and tool use require ongoing cross-talk between the motor system and object-related information from the visual system. The dorsal stream of the central visual system is specialized not only for the perception of object location and motion but also for visual guidance of hand movement, especially object-related movements (e.g., Culham & Valyear 2006; Goodale & Milner, 1992; Mishkin, Ungerleider, & Macko, 1983; Ungerleider & Haxby, 1994). A functional magnetic resonance imaging (fMRI) study from our lab (Yuan & Brown, 2014) investigated the neural basis of drawing using an MRI-compatible drawing tablet. By manipulating the visual feedback that participants experienced during drawing, we were able to identify several brain areas important for visually-guided hand movement during the drawing of geometric patterns: V5/MT+, V3A, and the posterior part of the intraparietal sulcus (pIPS). The aim of the present study is to examine the structural connectivity among these functional areas.

V5/MT+ is one of the principal motion-perception areas of the brain (Zeki et al., 1991; Huk et al., 2002; Watson et al., 1993). It responds to both biological and non-biological visual motion (Peuskens et al., 2005) and to implicit motion in still images (Kourtzi and Kanwischer, 2000; Senior et al., 2000). V3A, too, responds to visual motion (Bartels, Zeki, and Logothetis, 2008; Fischer et al. 2012; Tootell et al., 1997). Both V3A and V5/MT+ also respond to binocular depth cues (Buckthought et al., 2011; DeAngelis & Newsome, 1999), thus contributing to the perception of depth. The posterior intraparietal sulcus (pIPS) has a well-established role in visually-guided grasping (Culham et al., 2003) and spatial attention (Dinstein et al., 2008), and shares organizational similarities with the macaque IPS (Culham & Kanwisher, 2001; Swisher et al., 2007; Sakata et al., 1997).

There has been extensive research on the connectivity of these areas in monkeys. Tracer-injection studies in the macaque have demonstrated reciprocal connections between V3A and MT (e.g., Ungerleider & Desimone, 1986; Maunsell & van Essen, 1983; Lewis & van Essen, 2000), the latter of which is the homologue of human V5/MT+. Furthermore, V3A is also reciprocally connected with regions of the IPS (Felleman & van Essen, 1991; Nakamura et al., 2001). Finally, MT is reciprocally connected with intraparietal areas (Maunsell & van Essen, 1983; Ungerleider & Desimone, 1986; Felleman & van Essen, 1991).

In contrast to the extensive monkey connectivity literature, there is virtually no research describing these connections in humans. In fact, there is little reliable information about tracts in the human brain posterior to the well-characterized arcuate fasciculus (e.g., Catani & Thiebaut de Schotten, 2008). For example, most diffusion-based human brain atlases show virtually no tracts in the occipital lobe (Catani & Thiebaut de Schotten, 2008; Oishi et al., 2011). While there has been some discussion regarding fasciculi interconnecting the human occipital, temporal, and parietal lobes (Bartsch et al., 2013; Martino et al., 2013; Martino & García-Porrero, 2013), such work has thus far failed to describe specific connections among functionally-defined brain areas. The recent “rediscovery” of the vertical occipital fasciculus of Wernicke (VOF; Yeatman et al., 2014) has stimulated discussion about tracts connecting the ventral occipital/temporal lobes with the superior occipital and inferior parietal regions (Takemura et al., 2015). The VOF is thus an important, though little-explored, candidate for a white matter bundle containing tracts of relevance for the dorsal visual stream in humans.

The principal aim of the present study was to use diffusion imaging and probabilistic tractography to explore the connectivity between V5/MT+, V3A, and pIPS in the human brain *in vivo* in order to better understand the underlying neuroanatomical structure of dorsal visual stream areas. To our knowledge, only one study has reported connectivity between V5/MT+ and V3A in humans (Kim et al., 2006), but only in a single individual. Hence, a major objective of our study was to quantify connectivity at the group level.

3.3 Materials and methods

3.3.1 – Participants

Sixteen right-handed individuals (9 females, mean age 25 years old, range 18–35 years old) participated in the study after giving their informed consent (McMaster Research Ethics Board, McMaster University). Four participants were excluded due to motion artifacts and outlier brain activity, leaving 12 participants in the final analyses. Participants had normal or corrected-to-normal vision (using corrective lenses) and no history of neurological disorders, psychiatric illness, alcohol or substance abuse, and were not taking psychotropic medications. They received monetary compensation for their participation.

3.3.2 – fMRI tasks

Participants performed a cortical retinotopic mapping, a functional localizer for V5/MT+, and a functional mapping of the network involved in drawing, as performed in Yuan and Brown (2014). The cortical representation of retinotopic space was determined using a phase-encoded design in which the cardinal axes of visual space (eccentricity and polar angle) were mapped separately (Engel et al. 1997). The stimuli consisted of two different high-contrast, multi-coloured expanding checkerboard patterns. The “rotating wedge” stimulus (10 deg angle) swept through the polar angle dimension in a counterclockwise manner, while the “expanding ring” stimulus mapped eccentricity by starting from the center of the visual field and expanding outward. Eccentricity stimuli traversed space with a logarithmically changing rate, as has been used previously (Sereno et al. 1995; Conner et al., 2007). These phase-encoded single-direction stimuli used a 64s cycle, completing 8 cycles per scan, for a total of 512s.

The functional localizer task for V5/MT+ consisted of low-contrast oscillating rings (expanding-contracting) that were presented for 8 cycles (Tootell et al., 1995). It has also been shown that area V3A is often identifiable with this localizer (Tootell et al., 1997). Each cycle comprised 16s of moving rings and 16s of static rings, for a total task duration of 256s. Participants were instructed to fixate on a central crosshair for the duration of the scan. In order to develop regions-of-interest (ROI) for the tractography analyses, we had participants perform a drawing task with an MRI-compatible drawing tablet exactly as described for the “mark making” condition in Yuan and Brown (2014), in which participants drew geometric patterns (spirals, zigzags, and serpentines) with concurrent visual feedback. The drawing task comprised 6 epochs, alternating between 20s of drawing and 20s of fixation, for a total task duration of 240s. The results replicated the pattern found in Yuan and Brown (2014) as well as a meta-analysis of a sizeable literature using similar drawing and writing tasks (Yuan & Brown, 2015).

3.3.3 – Image acquisition

Diffusion images were acquired with a GE Medical Systems Signa Excite 3-Tesla MRI at the Imaging Research Centre at St. Joseph's Healthcare Hamilton. Diffusion images were collected using an echo-planar imaging spin-echo pulse sequence (TR = 9000 ms; TE = 83.4 ms; slice thickness = 2.0 mm; slice gap = 0 mm; matrix size = 128x128, FOV = 25.6 cm; 70 slices per volume, 2 mm isovoxel). Six T_2-weighted volumes without diffusion gradients and 64 volumes with unique diffusion-encoding gradients at a strength of b=1200 s/mm^2 were collected over a total time of 10m39s.

For the fMRI tasks, functional images sensitive to the blood-oxygen-level-dependent (BOLD) signal were collected with a gradient-echo echo planar imaging (EPI) pulse sequence using standard parameters (TR = 2000 ms, TE = 45 ms, flip angle = 90 degrees, 31 slices per volume, 4 mm slice thickness, no slice gap, matrix size = 64×64, field of view = 24 cm, voxel size = 3.75 mm×3.75 mm×4 mm), effectively covering the whole brain.

Anatomical T_1 images were collected for each participant (3D-FSPGR, IR-prepped, TI=900ms; TE=3.22 ms; flip angle = 9 degrees; receiver bandwidth = 31.25 kHz; NEX = 1; slice thickness = 1mm; slice gap = 0mm; FOV = 24cm; slices = 164; matrix size = 512 x 512).

3.3.4 – Data analysis

Functional image analyses were conducted using BrainVoyager QX (version 2.8.0, Brain Innovation). Images were reconstructed offline, and the scan series was realigned and motion-corrected. We excluded any participants that required translational and rotational corrections that exceeded an acceptable level of 1.5 mm and 1.5 degrees, respectively. During the preprocessing stage, a temporal high-pass filter was applied at a frequency of 0.0078 Hz, or 2 cycles per scan, using the GLM-Fourier algorithm. 3D spatial smoothing was performed using a Gaussian filter with a FWHM kernel size of 4 mm. Following realignment, each functional scan was normalized to the Talairach template (Talairach & Tournoux, 1988). The BOLD response for the drawing task was modeled as the convolution of a 20s boxcar with a synthetic hemodynamic response function composed of two gamma functions, whereas that for the motion localizer was modeled as the convolution of a 16s boxcar with a similar hemodynamic response function. The six head-motion parameters were included as nuisance regressors in the analysis. Each participant's data was processed using a fixed-effects analysis, corrected for multiple comparisons using a Bonferroni correction at a threshold of $p<0.05$. Contrast images for each subject were brought forward into a random-effects analysis, where a false discovery rate (FDR) of $p<0.01$ was employed as a correction for multiple comparisons, with a cluster threshold of $k=25$. Group data were registered onto a template inflated brain, generated using Brain Voyager. Talairach coordinates were extracted using NeuroElf (neuroelf.net). Retinotopic volumes were analyzed in BrainVoyager using a cross-correlation algorithm that identifies the regions of maximal response by subdividing each period into 8 segments of 4 volumes (for a total of 32 lags) and then convolving the reference function with the hemodynamic response function. Eccentricity and polar maps were created separately for each hemisphere of all participants.

Diffusion weighted images were processed using FMRIB Software Library (FSL, Jenkinson et al., 2012). Data were corrected for participant head motion and eddy-current-induced distortions, and subsequently underwent brain extraction to remove non-brain tissue from the analysis. A crossing-fiber model was used as the basis for performing probabilistic tractography (see Behrens et al., 2007 for details). We specified a maximum of two fiber orientations per voxel, with a burn in of 2000 and a weighting of 1.

Regions-of-interest (ROIs) for the arcuate fasciculus were created based on the descriptions of Makris et al. (2005). ROIs for V5/MT+, V3A, and pIPS were created using an in-house semi-automated script. For each participant, the script identified the locations of the peak voxels of sub-clusters from the participant’s own “mark making” condition in Talairach space. We also took into consideration the anatomical landmarks around each of the reported peak-voxel locations so as to choose the peaks closest to each participant’s own anatomical areas (see below). We used the chosen locations to create spherical ROIs with a radius of 5 mm. The spherical ROIs were then translated into each participant’s diffusion space and manually edited as needed to ensure that they did not cross sulcal boundaries.

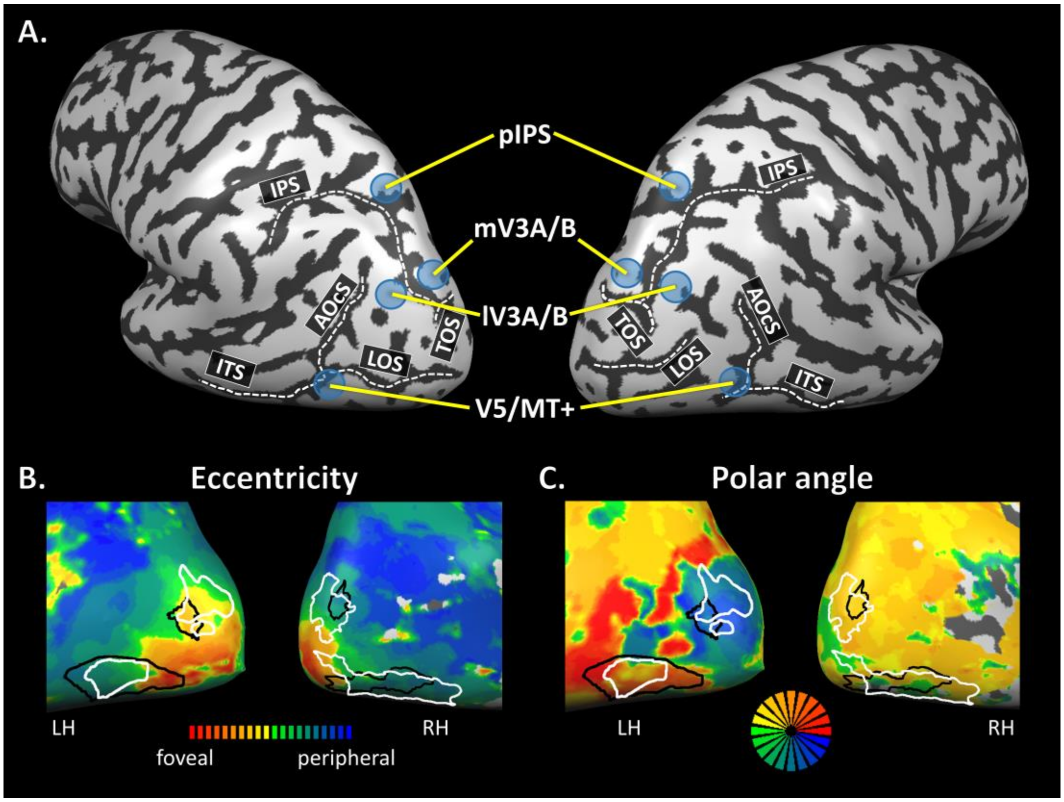


Figure 3.1: Regions of interest (ROIs) for probabilistic tractography.

A) Schematic of ROIs used in the analyses overlaid onto a 3D inflated rendering of a template brain, with relevant anatomical landmarks labeled. B) Overlay of motion localizer (black outline) and drawing (white outline) results on a retinotopic eccentricity map for one representative participant on an inflated Talairach-normalized template brain. C) Overlay of motion localizer (black outline) and drawing (white outline) results on a retinotopic polar-angle map for one representative participant on an inflated Talairach-normalized template brain. Abbreviations: anterior occipital sulcus (AOcS), intraparietal sulcus (IPS), inferior temporal sulcus (ITS), lateral occipital sulcus (LOS), transverse occipital sulcus (TOS).

For example, Figure 1A shows a representation of the ROIs on a sample inflated brain, as well as a combined overlay of the drawing activation (white outline), motion localizer (black outline), and retinotopy results of a representative participant. To create each V5/MT+ ROI, we chose peak activation coordinates from the drawing activations that matched the anatomical landmarks, which we identified as the cortex on the posterior bank of the anterior occipital sulcus (AOcS; see Petrides 2012), also known as the ascending limb of the inferior temporal sulcus (ALITS; e.g., Watson et al., 1993) and the pre-occipital notch (e.g., Malikovic et al., 2007). The V3A landmark was the region of cortex that is found near the junction of the transverse occipital sulcus (TOS) and the paroccipital IPS (IPS-PO). Since it was unclear whether the activations in our fMRI analysis were medial or lateral to the TOS/IPS-PO junction (i.e., they tended to be inside the sulcus), we decided to create separate ROIs for the medial and lateral parts of V3A, thus labeling these ROIs as mV3A and lV3A, respectively (see Results for additional labeling considerations). The pIPS ROIs were created based on peak activations in the cortical region just anterior to the sulcus of Brissaud (sB; see Petrides 2012; Zlatkina & Petrides, 2014) and medial to the IPS. In order to optimize tractography precision, we only included a participant's specific ROIs in our analysis if the peak-voxel location matched the anatomical landmarks. In total, there were 11 participants that contributed to the left hemisphere analyses, and 12 participants that contributed to the right hemisphere analyses. The functional and anatomical definitions of the ROIs were validated using retinotopy and the motion localizer.

Probabilistic tractography was performed individually (5000 samples per voxel, max steps = 2000, step length = 0.5mm, curvature threshold = 80 degrees, FA threshold=0.2) using each participant's ROIs as both the seed and termination masks (e.g., running tractography using mV3A as seed ROI and pIPS as termination, then running it again but reversing the seed and termination ROIs, and finally combining both results using binary masks to eliminate voxels that extended beyond the ROIs). Group average streamline count maps for each tract were calculated after warping individual results from participants' diffusion space into Talairach space using spatial normalization that applied warp coefficients and affine transformation matrices generated during linear and non-linear registration (Greve & Fischl, 2009). These results were then thresholded such that voxels with less than an intensity value (representing the number of streamlines) of 25 were zeroed in order to eliminate spurious connections. The value of 25 is close to thresholds customarily used in such experiments (Miller et al., 2012). The final results were then viewed using the Multi-image Analysis GUI (Mango, http://ric.uthscsa.edu/mango/) to generate orthogonal and oblique views. Boxplots and statistical analyses were carried out using a combination of FSL tools and the R statistical analysis packages (http://www.r-project.org).

3.4 Results

3.4.1 – Validation of ROIs

The V5/MT+ ROI was created using individual peak-voxel coordinates from the drawing task. These coordinates were found to be almost exactly matching the peak activations from the participant's own motion localizer analysis. As displayed in Figure 1B, retinotopic analysis showed the typical pattern of V5/MT+, responding to both the foveal and peripheral regions in the contralateral visual field (see Georgieva et al., 2009; Kolster et al., 2010; Abdollahi et al., 2014). The V3A ROI was likewise created using the drawing activations. These drawing activations were found to mostly include the motion localizer activations and areas that displayed characteristic retinotopic behavior of V3A, namely responding to both the foveal and peripheral regions in the contralateral visual field (see Sereno et al., 1995; Tootell et al., 1997; Wandell et al., 2007). It is worth nothing that, in our previous fMRI study, we obtained activations in V5 and V3A using a condition of motion perception that had no drawing component to it (Yuan & Brown, 2014), and that the coordinates were nearly identical to those obtained for drawing both in that experiment and in the current experiment.

As mentioned in the Methods section, two ROIs were generated for V3A, one medial to the TOS/IPS-PO junction and one lateral to it. While the medial ROI overlaps well with literature descriptions of V3A, the lateral ROI could potentially map onto V3B. While earlier functional activation studies placed V3B ventral (and lateral) to V3A (Van Oostende et al., 1997; Smith et al., 1998; Vaina et al., 2001), more recent retinotopic analyses place it more directly lateral to V3A (Press et al., 2001; Larsson & Heeger, 2006; Swisher et al., 2007; Wandell et al. 2007; Georgieva et al., 2009; Abdollahi et al., 2014; Wang et al., 2015). Because of uncertainties in the location of V3B in the literature (Yamamoto et al., 2012), we have opted to refer to the medial ROI as mV3A/B and the lateral ROI as lV3A/B.

We qualitatively evaluated all tractography results on their well-formedness, taking note of cases where the results did not look visually plausible to represent true tracts (i.e., if a result most strongly connected across a sulcus, due to resolution limitations, rather than through the white matter, or resembles unlikely shapes such as loops, spirals, and other anomalies). In general, our results showed that the lV3A/B ROI – i.e., the V3A/B ROI that occurred on the lateral bank of the TOS – generated a greater number of visually well-formed streamlines with both the V5/MT+ and pIPS ROIs than did the mV3A/B ROI, and also had higher streamline counts, although some connectivity was also seen for mV3A/B.

3.4.2 – Connectivity between V5/MT+ and lV3A/B

We found probabilistic connectivity (i.e., nonzero streamline counts) for all participants in both hemispheres between V5/MT+ and lV3A/B (see Figure 2 for 3D and sagittal and axial visualizations of these connections). Note that the majority of these tracts cannot be seen in the axial plane, as most of their trajectories are oriented in a sagittal manner. These tracts progressed from the lV3A/B ROIs ventrally and laterally toward the V5/MT+ ROI, although the left and right hemisphere tracts differed somewhat in their shape. These tracts were located in a position clearly posterior to the arcuate fasciculus in each hemisphere (see Figure 4B). Figure 2C demonstrates that the streamline count was significantly higher in the right hemisphere (Wilcoxon signed rank test, $p<0.05$). Additionally (not shown in figures), the streamline count was significantly higher between V5/MT+ and lV3A/B compared to the medial counterpart in the left hemisphere (Wilcoxon signed rank test, $p<0.05$), whereas these differences were present, but did not reach significance, in the right hemisphere (Wilcoxon signed rank test, $p>0.09$).

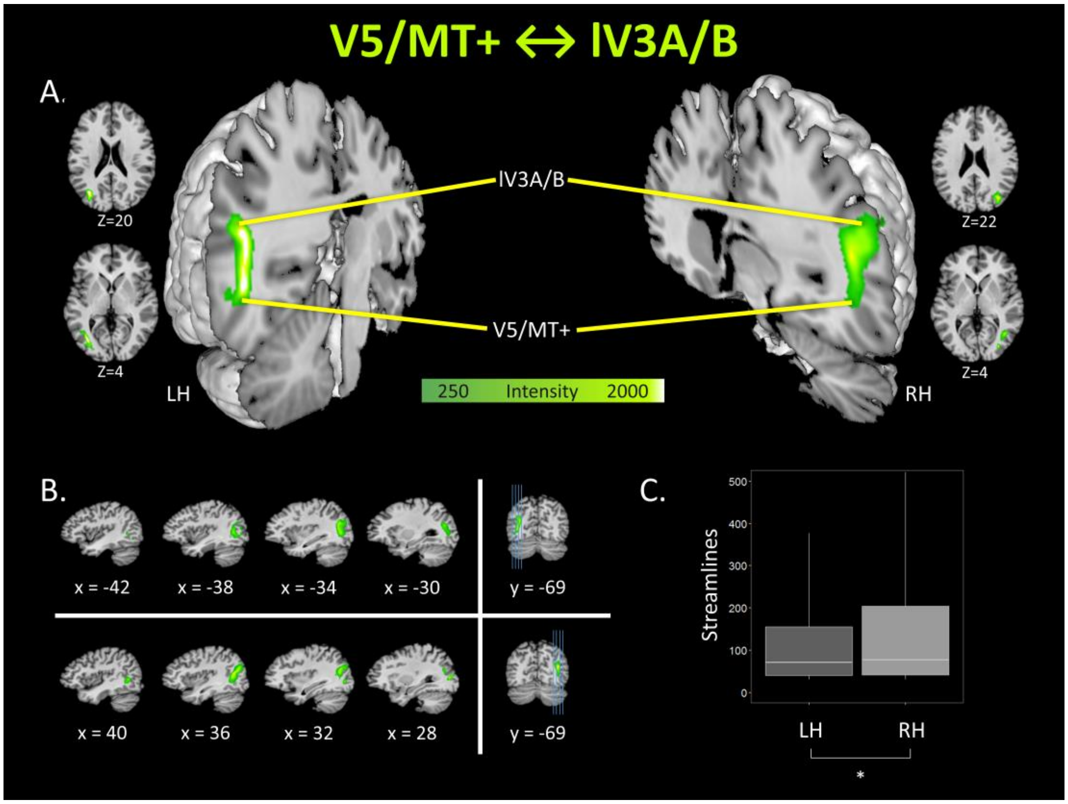


Figure 3.2: The V5/MT+ ↔ lV3A/B tract.

Group-average tractography for the V5/MT+ ↔ lV3A/B tract in the left and right hemispheres rendered on A) oblique and axial slice views of a cut-out brain, and B) a series of sagittal slices. C) A boxplot of connectivity values for the left and right V5/MT+↔lV3A/B tracts.

3.4.3 – Connectivity between lV3A/B and pIPS

We found probabilistic connectivity for all participants between lV3A/B and pIPS in both hemispheres (Figure 3). As with the previous results, the majority of these tracts cannot be seen in the axial plane, as most of their trajectories are oriented in a sagittal manner. These tracts progressed anteriorly and superiorly. We found voxel intensities of the connection between pIPS and lV3A/B to be significantly higher compared to the medial counterpart in the right hemisphere (Wilcoxon signed rank test, $p < 0.05$). Such differences were not seen statistically in the left hemisphere (Wilcoxon signed rank test, $p > 0.56$). The tract from the lV3A/B ROI progressed anteriorly and medially, crossing beneath the IPS, before bending dorsally to reach the pIPS ROI (Figure 3A). As with the V5/MT+↔lV3A/B tracts described previously, the pIPS↔lV3A/B tract is reliably posterior to the arcuate fasciculus and does not seem to overlap it (Figure 4B). The right pIPS↔lV3A/B tract appears to have a similar shape as the left-hemisphere tract (as shown in Figure 3A), and shows a significantly higher streamline count (Wilcoxon signed rank test, $p < 0.05$). Figure 3B shows sagittal sliceviews for additional details and more conventional visualization.

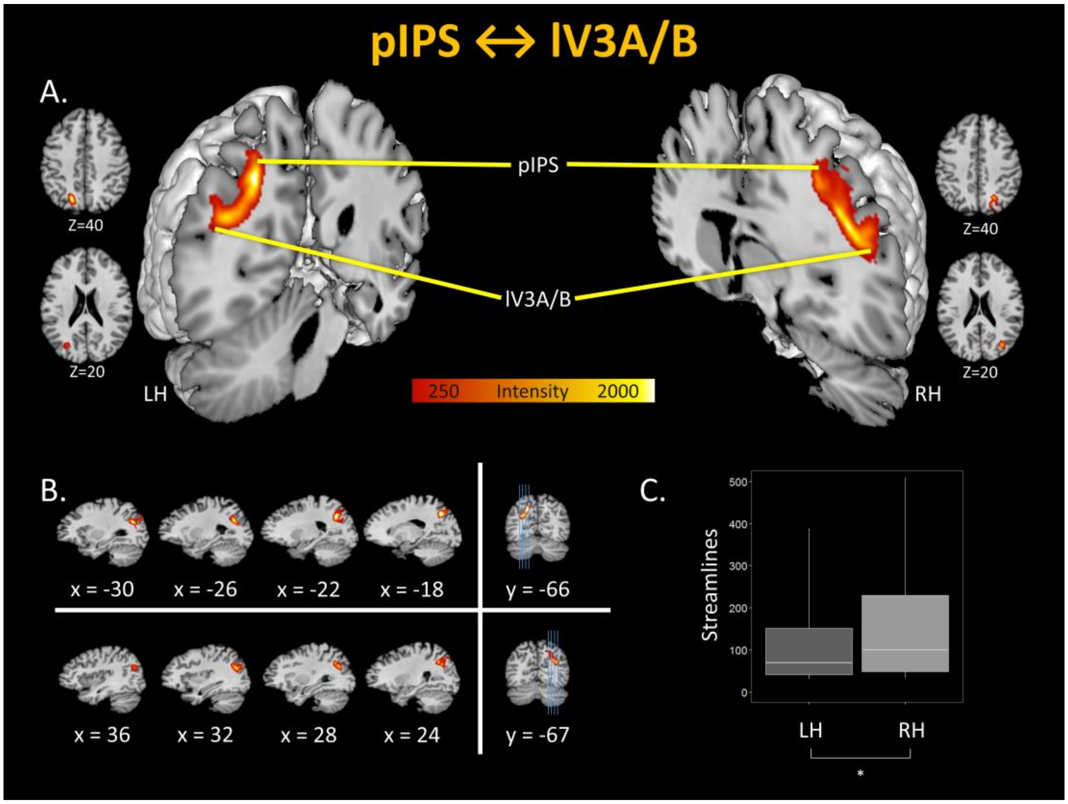


Figure 3.3: The pIPS ↔ lV3A/B tract.

Group-average tractography for the pIPS ↔ lV3A/B tract in the left and right hemispheres rendered on A) oblique and axial slice views of a cut-out brain, and B) a series of sagittal slices. C) A boxplot of connectivity values for the left and right pIPS↔lV3A tracts.

3.4.4 – Connectivity between V5/MT+ and pIPS

In contrast with our other ROI pairs, we did not find strong evidence of a continuous tract between V5/MT+ and the pIPS in either hemisphere in any of our participants. The probabilistic tractography results did not remain continuous at a low streamline threshold of 25.

3.5 Discussion

Our results demonstrate a strong likelihood of white matter pathways between V5/MT+, lV3A/B, and pIPS regions, which are pivotal components of the dorsal visual stream that process visual-motion information during visual guidance of hand movement, as occurs during drawing and a multitude of other motor skills. Specifically, our results showed connectivity between V5/MT+ and both the medial and lateral counterparts of V3A/B, as well as between lateral V3A/B and the pIPS. These tracts were found in both hemispheres, posterior to and not overlapping with the well-characterized arcuate fasciculus. Figure 4 presents a summary of the tractography results for both hemispheres. While an evaluation of connectivity from these ROIs to adjacent cortex would provide further insight to the specificity of these tracts, such an analysis is beyond the scope of this paper. In addition, while we understand that there is an inherent gyral bias in our tractography methods (see Reveley et al., 2015), a resolution of this bias is beyond the scope of this paper.

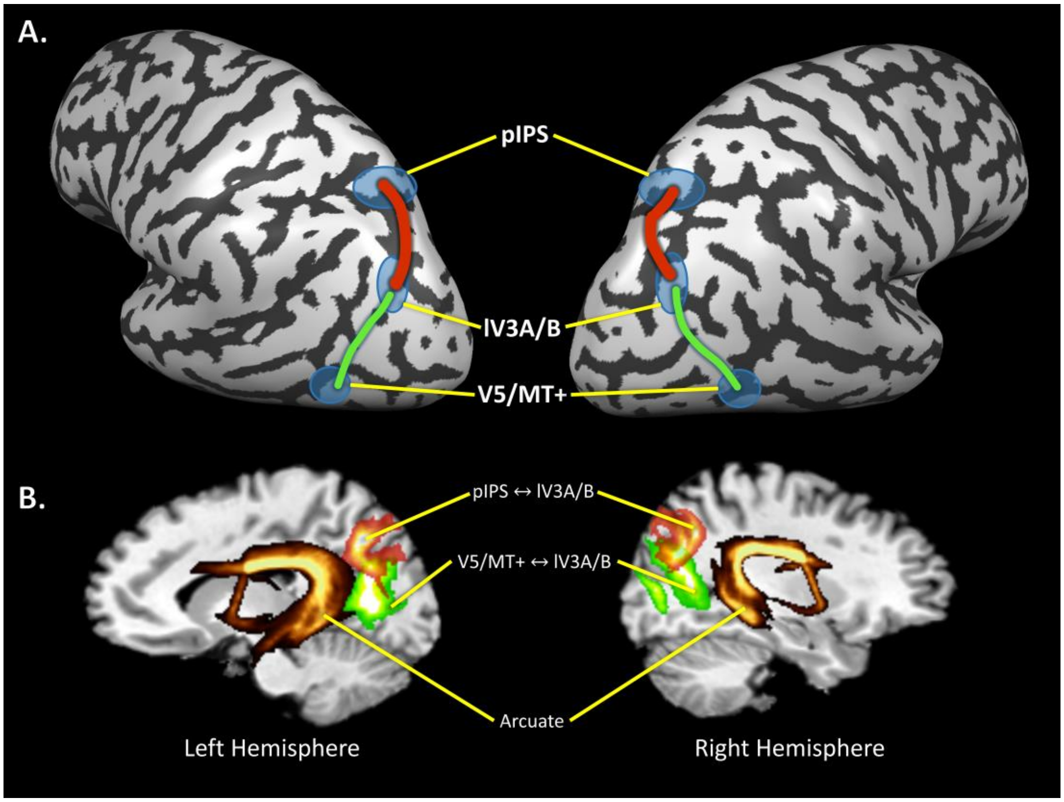


Figure 3.4: Summary of results.

A) Schematic illustration of the V5/MT+↔lV3A/B (green) and pIPS↔lV3A/B tracts (orange) superimposed on an inflated 3D rendering of a template brain from the group. B) Maximum intensity projection showing the V5/MT+↔lV3A/B and pIPS↔lV3A/B tracts in relation to the arcuate fasciculus and to one another.

3.5.1 – The V5/MT+ ↔ lV3A/B tract

The V5/MT+↔lV3A/B tract was found in what we believe to be the VOF (Wakana et al., 2004; Yeatman et al., 2013, 2014), also referred to as the perpendicular fasciculus in the early literature (Meynert, 1872; Gray, 1918, p.843), although the definition of this fasciculus is still debated (see below). Wakana et al. (2004) demonstrated ascending short-range association fibers (U-fibers) connecting the inferior temporal region with the superior occipital region, and grouped these fibers with the VOF. The proposed terminations of the VOF encompass important regions of the dorsal visual-processing stream, including V3A/B and V5/MT+, although connections from other functional areas in the vicinity, such as the visual word form area (VWFA; Yeatman et al., 2013) and fusiform face area (Kim et al., 2006), may send and/or receive fibers via this fasciculus. Our current analysis is not able to rule out the presence of these extraneous connections. Given that connectivity between V3A/B and V5 is bidirectional in the monkey (e.g., Ungerleider & Desimone, 1986), additional research is needed to elucidate the direction of information flow between lV3A/B and V5/MT+ during visually-guided motor activity in humans.

The VOF itself has been a source of confusion regarding its potential integration into the vertical (fourth) branch of the superior longitudinal fasciculus (SLF IV; Bartsch et al., 2013; Martino & García-Porrero, 2013), although Martino & García-Porrero (2013) disambiguated SLF IV from the VOF based on location: SLF IV connects inferior temporal to inferior parietal regions, whereas the VOF connects inferior temporal to superior occipital regions. More recently, Yeatman et al. (2014) described the VOF as connecting the ventrolateral and dorsolateral visual regions, which would include our V5/MT+ and lV3A/B ROIs. In addition, Takemura et al. (2015) demonstrated that the VOF contains fibers passing between the human ventral-stream area V4 (hV4-VO1) and the dorsal-stream region V3A/V3B, which suggests that the VOF may integrate information between the two visual streams. As with our results, these authors found their fibers to be distinctly posterior to the arcuate fasciculus.

Figure 4B shows a maximum intensity projection in the sagittal orientation in order to clearly visualize the posterior position of the V5/MT+↔lV3A/B tracts in both hemispheres in relation to the arcuate fasciculus. Our results are close, if not within, the known trajectory of the VOF, suggesting that this fasciculus is the most likely candidate to contain these obliquely-oriented tracts.

3.5.2 – The pIPS ↔ lV3A/B tract

We found significantly greater streamline counts between lateral V3A/B and the pIPS in both hemispheres compared to their medial counterparts. Although the connections between the lV3A/B and pIPS are well documented in the monkey literature (e.g., Felleman & van Essen, 1991; Nakamura et al., 2001), diffusion-based research on these connections in the human brain is still unavailable. Our results suggest that these regions are connected by cortical association fibers (also known as U-fibers, see Wakana et al., 2004), which connect nearby regions of the cortex. The arc of this pathway is clearly displayed in Figure 3A. Given the evidence of connectivity between these areas in the monkey brain and the similarity of neural organization between monkeys and humans, it is reasonable to find this connectivity in the human brain.

3.5.3 – The V5/MT+ ↔ pIPS tract

Although evidence exists for white matter bundles connecting the inferior temporal and inferior parietal regions of the human brain based on postmortem dissection analysis (e.g., Bartsch et al., 2013; Kim et al., 2006; Martino & García-Porrero, 2013), we did not find sufficient evidence for a direct pathway between V5/MT+ and the pIPS using probabilistic tractography. One factor accounting for this could be the placement of ROIs. Since our drawing-based ROIs are in different locations in the brain compared to the documented terminations of the temporo-parietal connections (e.g., Caspers et al., 2011), our results cannot be taken to mean that these connections do not exist in the human brain. Given the evidence of these connections in the macaque brain (e.g., Distler et al., 1993; Maunsell & van Essen, 1983), we would expect to find these connections in the human brain as well, albeit with a different placement of ROIs. Additionally, factors of geometry, such as length and curvature of the tract as well as influences from neighboring tracts, could weaken connectivity measurements.

In addition to length, curvature, and partial volume averaging, tractography suffers from other sources of error and bias. For example, it is known that tractography has an inherent gyral bias. Addressing such sources of bias is beyond the scope of this work. However, we note that the distinct connections from the lateral and medial V3A/B ROIs could in fact be more continuous than they appear. It is also clear that diffusion MRI tractography suffers from significant numbers of false positive results (Maier-Hein et al, 2016, Campbell et al., 2014). Importantly, tractography also suffers from false negative results, and a direct pathway between V5/MT+ and the pIPS may indeed exist, but not be visible with our current methods. Improvement in attainable spatial resolution may increase the detectability of this pathway. The streamline counts reported here reflect our confidence that the data support a connection between the designated ROIs. They will also be biased by ROI size and shape relative to the underlying tract geometry. With these confounds in mind, this work shows the most likely course of these pathways should they exist.

3.5.4 – A model of visual motion information transfer

Our results demonstrated both that the V5/MT+ ROI was significantly more strongly connected with the lateral V3A/B ROI than the medial counterpart and that the pIPS ROI formed visually plausible tracts to the lateral V3A/B ROI, but not to the medial counterpart. This distinction in connectivity is consistent with the hypothesis that there might be functional differences between the types of information transferred through the lateral and medial

V3A/B ROIs, which may correspond to the distinction between V3A and V3B proper (Press et al., 2001; Larsson & Heeger, 2006; Swisher et al., 2007; Wandell et al. 2007; Georgieva et al., 2009; Abdollahi et al., 2014; Wang et al., 2015).

Importantly, the ROIs for the present study were derived from an fMRI analysis of a visuomotor task, implicating these brain areas in the visual guidance of hand movement, which would suggest roles in both perceptual and motor-planning aspects of drawing. Our results suggest a lateral pathway by which information is shared between the V5/MT+, lV3A/B, and pIPS regions. This pathway looks to be significantly stronger in the right hemisphere compared to the left.

3.6 Conclusion

Using probabilistic tractography, we were able to demonstrate connectivity among V5/MT+, lV3A/B, and pIPS. These areas have been shown to play crucial roles in visual-motion processing and visual guidance of hand movement. To our knowledge, our study is the first to demonstrate connectivity among these human visual-motion processing areas in a group-level analysis, as well as to suggest a model of visual-motion information flow that distinguishes between lateral and medial V3A/B ROIs based not only on sulcal location but diffusion-based connectivity.

3.7 Acknowledgements

We are grateful to Michael Petrides for his comments, anatomical expertise, and critical insight. We thank Matthew Berry for assistance in the creation of the figures. We thank Michel Belyk, Léa Chauvigné, and Veronika Zlatkina for their comments. We also thank the anonymous reviews for their insight, comments, and suggestions. This work was funded by a grant from the Natural Sciences and Engineering Research Council (NSERC) of Canada to S.B. (Grant number: 371336). The authors declare no competing financial interests.

3.8 References

Abdollahi RO, Kolster H, Glasser MF, Robinson EC, Coalson TS, Dierker D, Jenkinson M, Van Essen DC, Orban GA (2014) Correspondences between retinotopic areas and myelin maps in human visual cortex. NeuroImage 99:509-524.

Bartels A, Zeki S, Logothetis NK (2008) Natural vision reveals regional specialization to local motion and to contrast-invariant, global flow in the human brain. Cereb Cortex 18(3):705-717.

Bartsch AJ, Geletneky K, Jbabdi S (2013) The temporoparietal fiber intersection area and Wernicke perpendicular fasciculus. Neurosurgery 73(2):E381-E382.

Behrens TEJ, Berg HJ, Jbabdi S, Rushworth MFS, Woolrich MW (2007) Probabilistic diffusion tractography with multiple fibre orientations: What can we gain? NeuroImage 34(1):144-155.

Buckthought A, Jessula S, Mendola JD (2011) Bistable percepts in the brain: FMRI contrasts monocular pattern rivalry and binocular rivalry. PLoS One 6:e20367.

Campbell JSW, Pike GB (2014) Potential and limitations of diffusion MRI tractography for the study of language. Brain Lang 131:65-73.

Caspers S, Eickhoff SB, Rick T, von Kapri A, Kuhlen T, Huang R, Shah NJ, Zilles K (2011) Probabilistic fibre tract analysis of cytoarchitectonically defined human inferior parietal lobule areas reveals similarities to macaques. NeuroImage 58(2):362-380.

Catani M, Thiebaut de Schotten M (2008) A diffusion tensor imaging tractography atlas for virtual in vivo dissections. Cortex 44(8):1105-1132.

Conner IP, Odom JV, Schwartz T, Mendola JD (2007) Retinotopic maps and foveal suppression in visual cortex of amblyopic adults. J Physiol 583(1):159-173.

Culham JC, Danckert SL, DeSouza JFX, Gati JS, Menon RS, Goodale MA (2003) Visually guided grasping produces fMRI activation in dorsal but not ventral stream brain areas. Exp Brain Res 153(2):180-189.

Culham JC, Kanwisher NG (2001) Neuroimaging of cognitive functions in human parietal cortex. Curr Opin Neurobiol 11(2):157-163.

Culham JC, Valyear KF (2006) Human parietal cortex in action. Curr Opin Neurobiol 16(2):205-212.

DeAngelis GC, Newsome WT (1999) Organization of disparity-selective neurons in macaque area MT. J Neurosci 19(4):1398-1415.

Dinstein I, Gardner JL, Jazayeri M, Heeger DJ (2008) Executed and observed movements have different distributed representations in human aIPS. J Neurosci 28(44):11231-11239.

Distler C, Boussaoud D, Desimone R, Ungerleider LG (1993) Cortical connections of inferior temporal area TEO in macaque monkeys. J Comp Neurol 334(1):125-150.

Engel SA, Glover GH, Wandell BA (1997) Retinotopic organization in human visual cortex and the spatial precision of functional MRI. Cereb Cortex 7(2):181-192.

Felleman DJ, van Essen DC (1991) Distributed hierarchical processing in the primate cerebral cortex. Cereb Cortex 1(1):1-47.

Fischer E, Bülthoff HH, Logothetis NK, Bartels A (2012) Human areas V3A and V6 compensate for self-induced planar visual motion. Neuron 73(6):1228-1240.

Georgieva S, Peeters R, Kolster H, Todd JT, Orban GA (2009) The processing of three-dimensional shape from disparity in the human brain. J Neurosci 29(3):727-742.

Goodale MA, Milner AD (1992) Separate visual pathways for perception and action. Trends Neurosci 15(1):20-25.

Gray H (1918) Anatomy of the human body (Lewis WH ed). Philadelphia, PA: Lea & Febiger.

Greve DN, Fischl B (2009) Accurate and robust brain image alignment using boundary-based registration. NeuroImage 48(1):63-72.

Huk AC, Dougherty RF, Heeger DJ (2002) Retinotopy and functional subdivision of human areas MT and MST. J Neurosci 22(16):7195-7205.

Jenkinson M, Beckmann CF, Behrens TEJ, Woorich MW, Smith SM (2012) FSL. NeuroImage 62(2):782-790.

Kim M, Ducros M, Carlson T, Ronen I, He S, Ugurbil K, Kim D-S (2006) Anatomical correlates of the functional organization in the human occipitotemporal cortex. Mag Reson Imaging 24(5):583-590.

Kourtzi Z, Kanwisher N (2000) Activation in human MT/MST by static images with implied motion. J Cognitive Neurosci 12:48-55.

Larsson J, Heeger, DJ (2006) Two retinotopic visual areas in human lateral occipital cortex. J Neurosci 26(51):13128-13142.

Lewis JW, van Essen DC (2000) Corticocortical connections of visual, sensorimotor, and multimodal processing areas in the parietal lobe of the macaque monkey. J Comp Neurol 428(1):112-137.

Maier-Hein K, Neher P, Houde J-C, Cote M-A, Garyfallidis E, Zhong J, Chamberland M, …, Descoteaux M (2016) Tractography-based connectomes are dominated by false-positive connections. doi: https://doi.org/10.1101/084137.

Makris N, Kennedy DN, McInerney S, Sorensen AG, Wang R, Caviness VS, Pandya DN (2005) Segmentation of subcomponents within the superior longitudinal fascicle in humans: a quantitative, in vivo, DT-MRI study. Cereb Cortex 15(6):854-69.

Malikovic A, Amunts K, Schleicher A, Mohlberg H, Eickhoff SB, Wilms M, Palomero-Gallagher N, Armstrong E, Zilles K (2007) Cytoarchitectonic analysis of the human extrastriate cortex in the region of V5/MT+: A probabilistic, stereotaxic map of area hOc5. Cereb Cortex 17(3):562-574.

Martino J, da Silva-Freitas R, Caballero H, de Lucas EM, García-Porrero JA, Vázquez-Barquero A (2013) Fiber dissection and diffusion tensor imaging tractography study of the temporoparietal fiber intersection area. Neurosurgery 72(1 Suppl Operative):87-97; discussion 97-98.

Martino J, García-Porrero JA (2013) In Reply: Wernicke perpendicular fasciculus and vertical portion of the superior longitudinal fasciculus. Neurosurgery 73(2):E382-E383.

Maunsell JH, van Essen DC (1983) The connections of the middle temporal visual area (MT) and their relationship to a cortical hierarchy in the macaque monkey. J Neurosci 3(12):2563-2586.

Meynert T (1872) Handbuch der Lehre von den Geweben des Menschen und der Thiere (Stricker S ed), pp. 694-808. London, UK: The New Sydenham Society.

Miller KL, McNab JA, Jbabdi S, Douaud G (2012) Diffusion tractography of post-mortem human brains: Optimization and comparison of spin echo and steady-state free precession techniques. NeuroImage 59(3):2284-2297.

Mishkin M, Ungerleider LG, Macko KA (1983) Object vision and spatial vision: Two cortical pathways. Trends Neurosci 6:414-417.

Nakamura H, Kuroda T, Wakita M, Kusunoki M, Kato A, Mikami A, Sakata H, Itoh K (2001) From three-dimensional space vision to prehensile hand movements: The lateral intraparietal area links the area V3A and the anterior intraparietal area in macaques. J Neurosci 21(20):8174-8187.

Oishi K, Faria A, van Zijl PCM, Mori S (2011) MRI atlas of human white matter (2nd ed). New York, NY: Elsevier Academic Press.

Peuskens H, Vanrie J, Verfaillie K, Orban GA (2005) Specificity of regions processing biological motion. Eur J Neurosci 21:2864-2875.

Petrides M (2012) The human cerebral cortex: An MRI atlas of the sulci and gyri in MNI stereotaxic space. Waltham, MA: Elsevier Academic Press.

Press WA, Brewer AA, Dougherty RF, Wade AR, Wandell BA (2001) Visual areas and spatial summation in human visual cortex. Vision Res 41:1321-1332.

Reveley C, Seth AK, Pierpaoli C, Silva AC, Yu D, Saunders RC, Leopold DA, Ye FQ (2015) Superficial white matter fiber systems impeded detection of long-range cortical connections in diffusion MR tractography. Proc Natl Acad Sci USA 112:E2820-E2828.

Sakata H, Taira M, Kusunoki M, Murata A, Tanaka Y (1997) The parietal association cortex in depth perception and visual control of hand action. Trends Neurosci 20(8):-350–357.

Senior C, Barnes J, Giampietroc V, Simmons A, Bullmore ET, Brammer M, David AS (2000) The functional neuroanatomy of implicit-motion perception or ‘representational momentum’. Curr Biol 10(1):16-22.

Sereno MI, Dale AM, Reppas JB, Kwong KK, Belliveau JW, Brady TJ, Rosen BR, Tootell RBH (1995). Borders of multiple visual areas in humans revealed by functional magnetic resonance imaging. Science 268(5212):889-893.

Smith AT, Greenlee MW, Singh KD, Kraemer FM, Hennig J (1998) The processing of first- and second-order motion in human visual cortex assessed by functional magnetic resonance imaging (fMRI). J Neurosci 18(10):3816-3830.

Swisher JD, Halko MA, Merabet LB, McMains SA, Somers DC (2007) Visual topography of human intraparietal sulcus. J Neurosci 27(20):5326-5337.

Takemura H, Rokem A, Winawer J, Yeatman JD, Wandell BA, Pestilli F (2015) A major human white matter pathway between dorsal and ventral visual cortex. Cereb Cortex 26(5):2205-2214.

Talairach J, Tournoux P (1988) Co-planar stereotaxic atlas of the human brain. New York: Thieme Medical Publishers.

Tootell RBH, Reppas JB, Kwong KK, Malach R, Born RT, Brady TJ, Rosen BR, Belliveau JW (1995) Functional analysis of human MT and related visual cortical areas using magnetic resonance imaging. J Neurosci 15:3215-3230.

Tootell RBH, Mendola JD, Hadjikhani NK, Ledden PJ, Liu AK, Reppas JB, Sereno MI, Dale AM (1997) Functional analysis of V3A and related areas in human visual cortex. J Neurosci 17(18):7060-7078.

Ungerleider LG, Desimone R (1986) Cortical connections of visual area MT in the macaque. J Comp Neurol 248(2):190-222.

Ungerleider LG, Haxby JV (1994) “What” and “where” in the human brain. Curr Opin Neurobiol 4:157-165.

Vaina LM, Solomon J, Chowdhury S, Sinha P, Belliveau JW (2001) Functional neuroanatomy of biological motion perception in humans. Proc Natl Acad Sci USA 98(20):11656-11661.

Van Oostende S, Sunaert S, Van Hecke P, Marchal G, Orban GA (1997) The kinetic occipital (KO) region in man: An fMRI study. Cereb Cortex 7(7):690-701.

Wakana S, Jiang H, van Zijl PCM (2004) Fiber tract-based atlas of human white matter anatomy. Radiology 230:77-87.

Wandell BA, Dumoulin SO, Brewer AA (2007) Visual field maps in human cortex. Neuron 56:366-383.

Wang L, Mruczek REB, Arcaro MJ, Kastner S (2015) Probabilistic maps of visual topography in human cortex. Cereb Cortex 25(10):3911-3931.

Watson JD, Myers R, Frackowiak RS, Hajnal JV, Woods RP, Mazziotta JC, Shipp S, Zeki S (1993) Area V5 of the human brain: Evidence from a combined study using positron emission tomography and magnetic resonance imaging. Cereb Cortex 3(2):79-94.

Yamamoto H, Fukunaga M, Takahashi S, Mano H, Tanaka C, Umeda M, Ejima Y (2012) Inconsistency and uncertainty of the human visual area loci following surface-based registration: Probability and entropy maps. Hum Brain Mapp 33(1):121-129.

Yeatman JD, Rauschecker AM, Wandell BA (2013) Anatomy of the visual word form area: adjacent cortical circuits and long-range white matter connections. Brain Lang 125(2):146-155.

Yeatman JD, Weiner KS, Pestilli F, Rokem A, Mezer A, Wandell BA (2014) The vertical occipital fasciculus: A century of controversy resolved by in vivo measurements. P Natl Acad Sci USA 111(48):E5214-E5223.

Yuan Y, Brown S (2014) The neural basis of mark making: a functional MRI study of drawing. PloS One 9(10):e108628.

Yuan Y, Brown S (2015) Drawing and writing: An ALE meta-analysis of sensorimotor activations. Brain Cognition 98:15-26.

Zeki S, Watson JD, Lueck CJ, Friston KJ, Kennard C, Frackowiak RS (1991) A direct demonstration of functional specialization in human visual cortex. J Neurosci 11:641-649.

Zlatkina V, Petrides M (2014) Morphological patterns of the intraparietal sulcus and the anterior intermediate parietal sulcus of Jensen in the human brain. P R Soc B 281(1797):20141493.

Chapter 4 – Storytelling is intrinsically mentalistic: An fMRI study of cross-modal narrative production

Ye Yuan, Judy Major-Girardin, Steven Brown

Unpublished manuscript.

Keywords: narrative, brain, speech, pantomime, drawing, fMRI, communication

4.1 Abstract

People utilize multiple expressive modalities for communicating narrative ideas about past events. The three major ones are speech, pantomime, and drawing. The current study used functional magnetic resonance imaging in order to identify common brain areas that mediate narrative communication across these three sensorimotor mechanisms. In the scanner, participants were presented with short narrative prompts, akin to newspaper headlines (e.g., “Surgeon finds scissors inside of patient”). The task was to generate a representation of the event, either by describing it verbally through speech, by pantomiming it gesturally, or by drawing it on a tablet. In a control

condition designed to remove sensorimotor activations, participants described the spatial properties of individual objects (e.g., “binoculars”). Each of the three modality-specific subtractions produced the same results, with activations in key components of the mentalizing network, including the temporo-parietal junction, posterior superior temporal sulcus, and precuneus. Conjunction analysis revealed that these areas constitute a cross-modal “narrative hub” that transcends the three modalities of communication. The involvement of these areas in narrative production suggests that people adopt an intrinsically mentalistic and character-oriented perspective when engaging in storytelling, whether using speech, pantomime, or drawing.

4.2 Introduction

Theories of language origin can be dichotomized into “vocal” and “gestural” models (Armstrong & Wilcox 2007; Arbib, 2012; Corballis, 2002; MacNeilage & Davis, 2005; McGinn, 2015). Gestural models posit that manually-produced symbols evolved earlier than those produced vocally, and that speech was a replacement for a pre-established symbolic system that was mediated by gestures alone. Importantly, the kind of gesturing that gestural models allude to is pantomime, or iconic gesturing. Iconic gesturing through pantomime is thought to have predated symbolic gesturing, passing through an intermediate stage that Arbib (2012) refers to as “proto-symbol”.

From a neuroscientific perspective, these theories of language origin establish a fundamental contrast between two different sensorimotor routes for the conveyance of language, namely the audio-vocal route for speech and the visuo-manual route for pantomime. Language is an inherently multimodal phenomenon, not least through the gesturing that occurs during speech (Beattie, 2016; Kendon, 2015; McNeill, 2005). Humans have yet a third means of conveying semantic ideas, and that is through the generation of images, such as via drawing and writing (Elkins, 1999). We have argued elsewhere that the capacity for drawing is an evolutionary offshoot of the system for producing iconic gestures such as pantomimes (Yuan & Brown, 2014). Drawing is essentially a tool-use gesture that “leaves a trail behind” in the form of resulting image. Overall, speech, pantomime, and image generation comprise a “narrative triad”, representing the three major modalities by which humans have evolved to referentially communicate their ideas to one another.

Perhaps the most important function of language is the communication of narrative, conveying the actions of agents, or “who did what do whom”. Agency is one of the primary elements that is communicated through syntactic structure (Tallerman, 2015). While word order varies across languages, 96% of languages place the subject (the agent) before the thing that the subject acts upon (Tomlin, 1986). Hence, an “agent first” organization of sentences seems to be an ancestral feature of language grammar (Jackendoff, 1999), and gestural models of language origin highlight this type of phrase organization as well (Armstrong & Wilcox, 2007). While language is well-designed to communicate agency through syntax, it typically does so in a multimodal manner, combining speech and gesture. A basic question for the evolutionary neuroscience of human communication is whether the conveyance of narrative is linked to specific sensorimotor modalities (vocal vs. manual) or whether there are cross-modal narrative areas in the brain that transcend these modalities. This question led us to design an experiment in which we would explore for the first time whether cross-modal brain areas mediate the communication of narrative ideas using speech, pantomime, and drawing as the triad of production modalities.

Most previous neuroimaging studies of cross-modal communication are perceptual, and we are not aware of production studies that have compared any pair of functions among speech, pantomime, and drawing in healthy adults. Before considering the relevant perceptual studies, we will first examine a handful of studies that have explored the basic network for narrative production, focusing on speech as the modality. AbdulSabur et al. (2014), in a combined fMRI and positron emission tomography (PET) study, had participants learn a series of 12 simple stories, based on a standardized set of three-picture stimuli, and then recount the stories aloud in the scanner when seeing the story's title alone. The control condition was the recitation of standard nursery rhymes. Because this condition involved the production and perception of speech, most of the sensorimotor activations for speech were washed out in the subtraction of storytelling minus nursery rhymes. In theory, what should be left over are areas involved in the narrative content of the stories. This subtraction revealed areas involved in both language processing and mentalizing (i.e., character processing), including the left inferior frontal gyrus (IFG), dorsomedial prefrontal cortex, precuneus, superior parietal lobule, posterior superior temporal sulcus (pSTS), cerebellum, and basal ganglia. In addition, there was a prominent activation in the cortex of the temporo-parietal junction (TPJ), another area that is implicated in both language processing (especially semantics) and mentalizing (Carter & Huettel, 2013) and which we will argue is a strong candidate for being a cross-modal hub area for narrative processing.

In Hassabis et al. (2014), participants became familiarized prior to their scan with four characters having contrastive personality traits. In the scanner, participants were required to mentally play out vignettes involving prescribed events happening to the characters in prescribed locations (e.g., the character's drink being spilled while in a bar), and focus on the actions, thoughts, and feelings of the character. Hence, the participant had to mentally simulate a narrative involving the protagonist. Hassabis et al. (2014) observed activations across most of the areas described in the analysis of AbdulSabur et al. (2014), including the TPJ. Interestingly, nearly identical results were obtained when the participant imagined themselves (rather than a character) in the prescribed scenarios, consistent with the results of Awad et al. (2007), in which participants generated self-referential propositional speech (“tell me what you did last weekend”), compared with a baseline condition of counting. The results of these studies suggest that narrative production is not just about recounting a sequence of events but of conveying embodied episodes in which the perspective of a protagonist is automatically assumed as a default process, as shown by strong activations in the mentalizing areas like the TPJ, precuneus, and medial prefrontal cortex. In other words, they suggest a character-driven mechanism of narrative processing in the brain.

Looking now to the multimodal perceptual studies, no neuroimaging study has compared images with gestures, to the best of our knowledge (although see Wu and Coulson, 2011, for an electroencephalography study). However, several studies have compared speech with gestures. Xu et al. (2009) had participants view video clips of an actor performing gestures (pantomimes or emblematic gestures) or listen to an actor speaking words having the same meaning as the observed gestures. A major point of overlap was found not in the TPJ but more ventrally in the posterior superior temporal sulcus (pSTS) bilaterally. Other studies that have compared speech with gesture have found either similar effects to Xu et al. (2009) in the pSTS alone (Kircher et al., 2009), effects in both the TPJ and pSTS

(Andric et al., 2013; Redcay et al., 2016), or effects in neither the TPJ nor the pSTS (Straube et al., 2012). Cross-modal comparisons between language and images have highlighted similar areas. In an early PET study, Vanderberghe et al. (1996) compared the processing of pictures with the processing of single words (visually presented) having the same content as the pictures. They found overlapping activation in the left TPJ (among other areas). Jouen et al. (2015) explored convergent activations related to semantic processing across modalities using an fMRI study in which participants viewed pictures of everyday events or read sentences describing these same types of events. Like Vanderberghe et al. (1996), they found converging activations in the region of the left TPJ bordering on the pSTS. Overall, while the role of the TPJ in language-based narrative is compelling, its importance for narrative-based gestures and images is still unclear, with more evidence of convergence being found in the pSTS than more dorsally in the TPJ.

The principal objective of this fMRI study was to carry out the first tri-modal production study of narrative processing with the aim of identifying a “narrative hub” in the brain. In order to do this, we had participants read simple headlines (for example, “Surgeon finds scissors inside of patient”) and then generate the narrative described in the headline using either speech (as in a news brief), pantomime (as in the game of Charades), or drawing (as in the game of Pictionary), where the latter was done using an

MRI-compatible drawing table that allowed participants to see their drawings (Yuan & Brown, 2014). All headlines described transitive actions carried out by protagonists, in keeping with a view of narrative based on agency. As a way of controlling for sensorimotor differences among the modalities and in order to hone in on the narrative content per se of the task, we had participants perform a control task in which they were presented with the names of objects (e.g., “binoculars”) and were asked to describe the spatial properties of each object (again either through speech, pantomime, or drawing), while avoiding any mention of the object’s uses or human interactions with it. This permitted a cognitive contrast between narration (a recounting of the actions of a protagonist) and description (an enumeration of an object’s properties, separate from a person’s interaction with it). We performed the “narration versus description” contrast for each of the three modalities individually. This subtraction permitted us to eliminate the sensorimotor components of the tasks (i.e., audiovocal activations for speech, and visuomotor activations for pantomime and drawing), and thereby isolate components specifically associated with narrative processing of the protagonist’s actions. We then ran a conjunction of the three narration-versus-description subtractions in order to see if there were any brain areas that were commonly activated across the three narrative modalities of communication, while controlling for sensorimotor differences. Based on the literature mentioned above, we predicted that the

TPJ and/or the adjacent pSTS would serve such a function. The TPJ in particular is an attractive candidate for this role in cross-modal narrative since it is involved in the processing of language, theory-of-mind, and agency, hence combining the linguistic and character-related aspects of narrative.

4.3 Methods

4.3.1 – Stimulus validation

A set of 60 headline stimuli was devised by the authors. All of them were subject-verb-object declarative statements in the present tense describing narrative events as transitive actions carried out by a protagonist (typically gender-neutral) on some object or person. Examples include “Surgeon finds scissors inside of patient” and “Fisherman rescues boy from freezing lake”. Headlines ranged in length from 5 to 8 words. Word-frequency analysis was performed using the Corpus of Contemporary American English (COCA; http://corpus.byu.edu/coca) to ensure that the headlines did not contain words with outlier frequency ratings (i.e., in excess of 1.5 times the interquartile range of the group word frequencies). From a narrative standpoint, the headlines were designed to convey “newsworthy” events that one might find in a newspaper. Half were designed to convey a positive-valenced outcome and half a negative-valenced outcome.

A stimulus-validation experiment was carried out using 32 undergraduate students (mean age 19.8 years, 24 females) in order to equate the headlines to be used in the fMRI experiment for level of difficulty across the three modalities of production. Since 60 stimuli could not be accommodated into a one-hour experiment, the stimuli were randomly divided into four groups of 15 headlines. While in a sound booth, participants were asked to produce a representation of each of the 15 headlines using all three modalities (speaking, pantomiming, and drawing) in a randomized order. Headlines were presented to participants using a laptop computer (HP Pavilion dv5-2050ca) running E-Prime 2.0 Standard (build 2.0.10.356). Audiovisual behavioral data were recorded using a tripod-mounted video camera (Canon FS200). For each trial, participants were given 4s to read the headline and the associated modality of production, followed by a 12s planning phase. After a 2s “Ready” signal, the word “Begin” indicated the start of a 30s production phase. An auditory tone signaled the end of the production phase, and a 5s fixation cross was shown before the next trial began. Each of the 45 trials (15 headlines x 3 modalities) lasted 51s, for a total session time of just over 38m.

After the experiment, the participants were asked to rate the headlines that they saw during the experiment for emotional valence (positive, neutral, or negative) and difficulty of production (1-easy to 3-difficult) for each headline per modality. Video recordings were used to assess the time required for production. The goal was to assemble a collection of headlines that were not significantly different across modalities in terms of difficulty and that had a completion time of longer than 18s in order to prevent participants from finishing early during the fMRI study (which contained task epochs of 18s).

Mean modality difficulty and headline difficulty scores were calculated by collapsing across all modalities and by collapsing across all stimuli, respectively. Emotional valence ratings were tallied to determine how each headline was perceived by participants. A one-way analysis of variance (ANOVA) was used to test for differences in difficulty scores across speech, pantomime, and drawing for each of the 60 headlines. Post hoc t-tests were used to examine pairwise differences (i.e., speech vs. mime, mime vs. drawing, and speech vs. drawing). Video recordings of the production times were analyzed in a similar fashion as the questionnaire data using a one-way ANOVA to determine whether there were differences in production times among modalities, with post hoc t-tests used to examine pairwise differences.

The aim of the validation study was to identify a 24-headline subset of the original 60-headline stimulus set to be used in the fMRI experiment that showed comparable difficulty levels for the three modalities of narrative production. However, not surprisingly, speaking a headline was generally rated to be easier by participants than miming or drawing it. We dealt with this difference in difficulty in two manners. As a first step, we ignored the speech condition, and attempted to create a 24-headline set that was equated for perceived difficulty between miming and drawing ($p > 0.05$ in the pairwise comparison). The second step involved introducing a modification of the vocalization style for the speech task. In the validation experiment, participants spoke in a standard manner. However, for the fMRI experiment, we wanted participants to speak in a slower and more controlled manner, as has been done in previous vocal studies from our lab (Brown et al., 2008). Hence, participants were instructed during a training session on a day prior to the scanning session to speak with their teeth together so as to reduce jaw movement, and to vocalize at a slow pace, roughly 2 words per second. Compared with the production style of the validation experiment, this should have increased the difficulty level of the speech task. We tested this manipulation on a separate group of undergraduate participants ($n=20$), and demonstrated that, while pre-manipulation difficulty ratings were comparable between this group of participants and the original validation group, post-manipulation difficulty ratings increased to be similar to the pantomime and drawing difficulty in the validation experiment ($p>0.05$).

We would also point out that, while pantomime and drawing share the same effector system of the hands, wrists and arms, speech utilizes a completely different sensorimotor system comprised of the vocal tract and auditory system. In this regard, its presence in the 3-way conjunction analysis should bias the results toward erring on the side of false negatives, rather than false positives. Hence, even if the speech task were overall less difficult than the two manual tasks, this would actually minimize a confounding effect of the speech task on the 3-way conjunction since it would tend to minimize speech-specific activations compared to the (difficulty-matched) manual tasks.

The control task for the fMRI study consisted of an object description task. Control stimuli were created as single objects and as natural or associative pairs relating to the concepts present in the headline stimuli. All control objects were inanimate. In order to minimize the effects of different categorical associations, the control stimuli were chosen to be in the same semantic category as the headline scenarios. Examples included "helicopter" for the headline "Pilot lands plane safely during storm", and "football helmet" for "Quarterback throws long pass to win game". Note that the control stimuli were words (just as with the headlines), not visual images of objects. Control stimuli were validated by two expert artists with over 10 years of drawing experience. Both artists gave a rating of how well they felt that they could draw the control objects within a time limit of 20s using a 5-point Likert-like scale (1= "I did a bad job"; 5= "I'm happy with what I drew"). All control objects were rated at least 2+ by both artists.

4.3.2 – Participants

Twenty-four right-handed individuals who did not participate in the stimulus-validation experiment participated in the fMRI experiment after giving their informed consent (McMaster Research Ethics Board, McMaster University). In order to ensure that the drawing task could be performed in a competent manner, we recruited participants who had a minimum of 2 years of fine arts training. Most of the participants were undergraduate majors in a studio arts program. Two participants were excluded due to head motion and one participant was excluded for responding to one headline with the incorrect modality, resulting in 21 participants in the analyses (17 females, mean age 20.4). The mean fine arts training of the participants was 5.5 years.

Participants had normal or corrected-to-normal vision (using corrective lenses) and no history of neurological disorders, psychiatric illness, alcohol or substance abuse, and were not taking psychotropic medications. They received monetary compensation for their participation. Participants attended a one-hour training session on a day prior to the fMRI experiment in order to become familiarized with the task timing and to learn how to perform all of the tasks in a highly controlled manner so as to minimize head, jaw, and body movement.

4.3.3 – Stimuli

In the MRI scanner, stimuli were presented to participants using a laptop computer (HP Pavilion dv5-2050ca) running E-Prime 2.0 Standard (build 2.0.10.356). Each headline was paired uniquely to a visual “modality icon” (a voice icon for speech, a hand icon for mime, and a pencil icon for drawing), resulting in 24 headline-modality pairs that did not contain duplicate headlines. No participant produced a given headline with more than one modality (i.e., there was no within-subject repetition), and the full set of headline-modality pairings was achieved in a between-subject manner by creating three stimulus sets across the pool of participants, as produced using a Latin squares approach. Control stimuli (names of objects) were paired to the three modalities using the same approach, but completely independently of the headlines. In other words, even though the control objects were generated based on a pairing to the semantic content of the headlines, the control objects were separated from the headlines as follows. The 24 headlines were initially randomly distributed across the 4 fMRI scans, but were then assigned to those scans. Within a scan, they were presented in random order, but the 4 scans were presented in a fixed order. Once this assignment of headlines to scans had been made, the control stimuli for the set of headlines for one scan were assigned to another scan, such that the control stimuli derived from a set of headlines were never in the same scan as those headlines themselves. In addition, the 3 stimulus sets were set up such that a control stimulus was never performed in the same modality as its associated headline. For example, the headline might be mimed in one scan and the control object drawn in another scan. As mentioned above, the performance of all headlines in all three modalities was achieved in a between-subject manner such that no participant ever performed a given headline or object in more than one modality, and that no participant ever performed a headline and its associated control object in the same modality or in the same scan.

4.3.4 – fMRI tasks

During a task epoch, each stimulus (a headline or object name) and the associated modality icon were displayed for 8s, during which time the participants were instructed to plan what they were going to do but to not physically respond. After the planning phase, the screen was replaced by a grey canvas, and participants were given 18s to present the stimulus item using the assigned modality. There was then a 4s “Stop” signal that indicated the end of the production phase, followed by a “Ready” screen for 2s as a transition between stimuli. Each task epoch thus lasted 32s and was analyzed as a single unit. During the fixation trials, a crosshair was displayed for 16s, followed by “Ready” for 2s, lasting a total of 18s and was also analyzed as a single unit. Each of the 4 scans had a duration of 7m (420s).

Participants performed 4 functional MRI scans comprised of both narrative production and object description in all three modalities (speech, mime, and drawing). Each scan contained 6 narrative production trials (2 of each modality), 6 object description trials (2 of each modality), and 2 fixation trials. For the speech modality, participants were instructed to verbally tell a story befitting the given headline or to verbally describe the spatial properties of a given object. For the narrative condition, the instruction during the training session was "Your task is to tell a story centered on the protagonist of the headline". As mentioned above, participants were required to use a very slow rate of speech in order to minimize the difference between the amount of elements that could be verbally produced compared to those being mimed or drawn. For the mime modality, participants produced bimanual pantomimes in order to depict a given headline or object. For the drawing modality, participants drew on an MRI-compatible drawing tablet (Tam et al., 2010; Yuan & Brown, 2014) using their dominant hand. They had full visual feedback of their drawings during the task.

In order to minimize the narrative content of the object description task, we explicitly instructed participants to focus on the structural properties of the objects and to avoid describing the objects' uses, since the latter would introduce a human actor into the description. Hence, the contrast between the narrative task and the object task was designed to emphasize the unique presence of a protagonist in the narrative condition.

4.3.5 – Image acquisition

Functional images sensitive to the blood-oxygen-level-dependent (BOLD) signal were collected with a gradient-echo echo planar imaging (EPI) pulse sequence using standard parameters (TR = 2000 ms, TE = 45 ms, flip angle = 90 degrees, 31 slices per volume, 4 mm slice thickness, no slice gap, matrix size = 64×64, field of view = 24 cm, voxel size = 3.75 mm×3.75 mm×4 mm), effectively covering the whole brain. Over each scan, 210 volumes of data were collected. Anatomical T1 images were collected for each participant (3D-FSPGR, IR-prepped, TI=900ms; TE=3.22 ms; flip angle = 9 degrees; receiver bandwidth = 31.25 kHz; NEX = 1; slice thickness = 1mm; slice gap = 0mm; FOV = 24cm; slices = 164; matrix size = 512 x 512).

4.3.6 – Data analysis

Functional image analyses were conducted using BrainVoyager QX (version 2.8.0, Brain Innovation). Images were reconstructed offline, and the scan series was realigned and motion-corrected. During the preprocessing stage, a temporal high-pass filter was applied at a frequency of 0.0078 Hz, or 2 cycles per scan, using the GLM-Fourier algorithm. 3D spatial smoothing was performed using a Gaussian filter with a FWHM kernel size of 4 mm. Following realignment, each functional scan was normalized to the Talairach template (Talairach & Tournoux, 1988). The BOLD response for the drawing task was modeled as the convolution of a 32s (task) or 18s (fixation) boxcar with a synthetic hemodynamic response function composed of two gamma functions. The six head-motion parameters were included as nuisance regressors in the analysis. As mentioned previously, two participants were excluded due to head motion in excess of 1.5 mm of translation or 1.5 degrees of rotation in one or more dimensions, and one participant was excluded for responding to a headline in an incorrect modality. Each participant's data was processed using a fixed-effects analysis, corrected for multiple comparisons using a Bonferroni correction at a threshold of $p<0.05$ for low-level (task vs. fixation) subtractions and FDR $p < 0.05$ for high-level (narrative vs. object) comparisons.

Each individual participant's results was forwarded into a random effects analysis to produce the group results ($n=21$). For the low-level subtraction analyses, we contrasted each task x modality (narrative/object x speak/mime/draw) combination to the fixation baseline condition. This resulted in six subtractions: narrative speech > fixation; object speech > fixation; narrative mime > fixation; object mime > fixation; narrative draw > fixation; and object draw > fixation. We then performed 3 high-level intra-modal subtractions between narrative and the object control: 1) narrative speech > object speech; 2) narrative mime > object mime; and 3) narrative draw > object draw. Finally, we performed 3 pairwise conjunctions between the high-level subtractions for each modality pair, as well as the 3-way conjunction of interest ("narrative speech > object speech" ∩ "narrative mime > object mime” ∩ “narrative draw > object draw”). All low-level analyses were thresholded at FDR $p < 0.005$, the higher-order subtractions at uncorrected $p < 0.0005$, and the conjunction analyses at uncorrected $p < 0.005$. The uncorrected thresholds were determined a posteriori to improve exploratory discussion.

It is important to note that, while the drawing tablet provides information about behavioral performance in the scanner, we were not able collect behavioral data on speech or mime production in the scanner. Had our focus been on a subtraction analysis, this might have been problematic. However, since our principal goal was to carry out a conjunction analysis, this would emphasize what the three modalities share, and would thus offset behavioral differences among modalities with respect to task difficulty.

4.4 Results

Figure 1 shows the brain activations for the Narrative > Fixation contrast for each modality (FDR corrected, $p < 0.005$; Talairach coordinates in Supplementary Table 1). Prominent activations were found in the primary sensorimotor cortex (SMC) related to the particular effector system employed by each modality: for speech, activity was found ventrally in the orofacial precentral gyrus, whereas for both mime and drawing, activity was found more dorsally in the somatotopic representations for the hand, wrist, and arm. The supplementary motor area (SMA) was found to be commonly activated across the three modalities. Pantomime and drawing, but not speech, showed strong activations throughout the posterior parietal cortex associated with visuomotor processing. This included intraparietal sulcus (IPS) and superior parietal lobule (SPL). In addition, both of the visuomotor tasks showed activation in the frontal eye fields (FEF) and in visual-motion area V5/MT+, neither of which were seen in the speech task. Modality-specific activations included the auditory association cortex of the posterior superior temporal gyrus (pSTG) for speech, the supramarginal gyrus (SMG) for miming, and area V3A for drawing. Lower-level visual areas were seen in all tasks, in part driven by the presentation of the prompt during the planning phase (either a headline or object-word), although the visual motion that occurred during task production for pantomime and drawing worked to make these activations much stronger than that for speech.

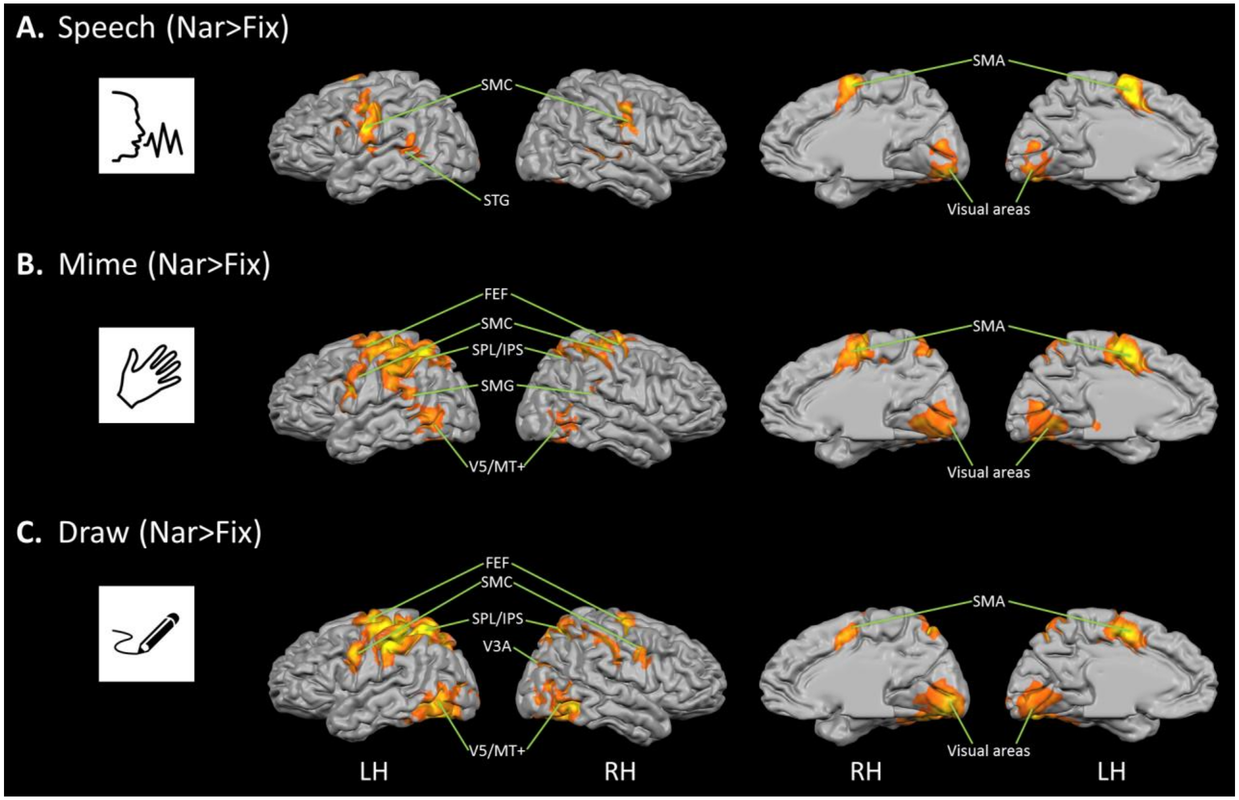


Figure 4.1: Narrative > Fixation subtractions

Brain activations for the Narrative (Nar) > Fixation (Fix) contrast for A) speech, B) mime, and C) drawing are presented. All results are displayed at FDR $p < 0.005$ on a template brain showing the left hemisphere (LH) and right hemisphere (RH). The two brains at left are lateral views, and the two brains at right are medial views. The icons in each row are the same ones that participants saw during stimulus presentation. Abbreviations (alphabetically): FEF, frontal eye fields; IPS, intraparietal sulcus; MTG, middle temporal gyrus; SMA, supplementary motor area; SMC, sensorimotor cortex; SMG, supramarginal gyrus; SPL, superior parietal lobule; STS, superior temporal sulcus; V5/MT+, middle temporal complex. Low-level visual-processing areas corresponding to V1 and V2 are labeled as “visual areas”.

In order to isolate the brain areas specific to narrative generation, we performed the high-level Narrative > Object subtraction for each modality (Figure 2, uncorrected $p < 0.0005$; Talairach coordinates in Table 1). Virtually all of the sensorimotor areas seen in Figure 1 were eliminated in this subtraction for each modality, suggesting that the control condition was well-matched to the narrative condition for these features. What was seen instead were areas associated with mentalizing, social cognition, semantics, and discourse processing. Similar patterns were seen in each of the three modalities. The most common areas across the three were the pSTS and TPJ bilaterally. The precuneus was seen in speech and drawing, although at slightly different locations. Activity in this area was present for mime as well, but it did not survive at the current threshold. Next, the middle temporal gyrus (MTG), a semantic-processing area, was present bilaterally for speech and drawing, but not for mime. Finally, the anterior superior temporal sulcus (aSTS) that is associated with discourse processing was present for speech and pantomime, but not drawing. The effect was stronger in the right hemisphere for speech and the left hemisphere for mime.

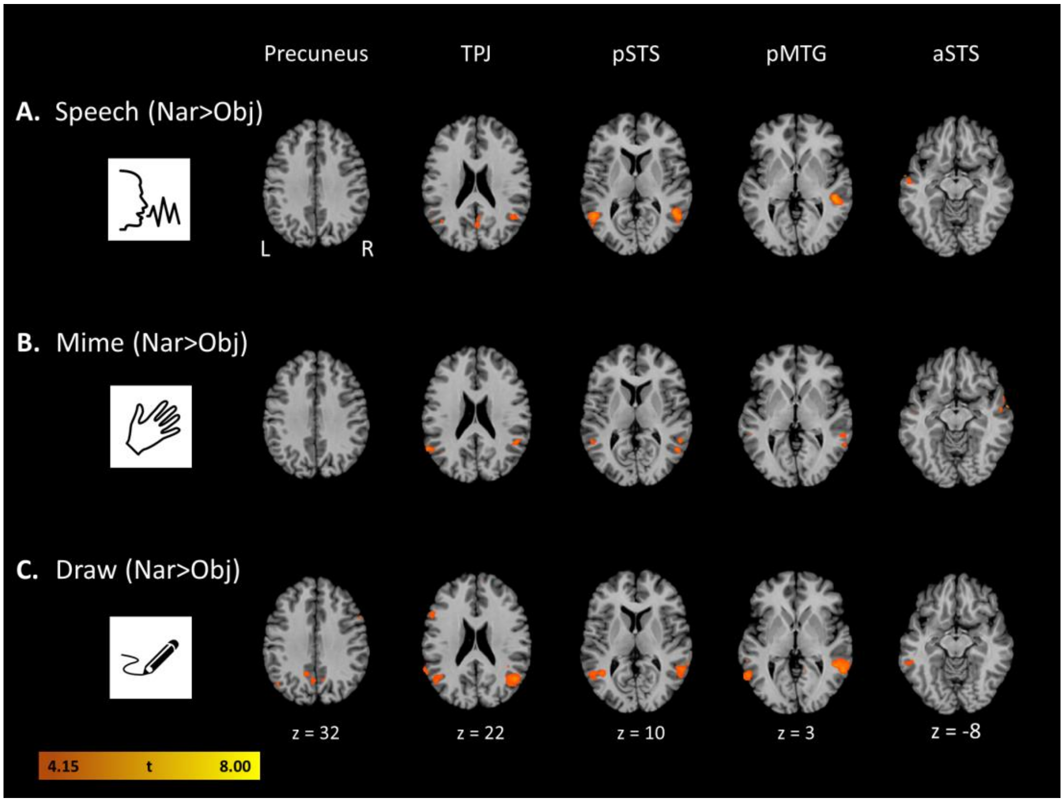


Figure 4.2: Narrative > Object subtractions

Axial views of the Narrative (Nar) > Object (Obj) high-level contrast for A) speech, B) mime, and C) drawing are displayed at uncorrected $p < 0.0005$. The icons in each row are the same ones that participants saw during stimulus presentation. The figure is set up to show activations in the precuneus, temporoparietal junction (TPJ), posterior superior temporal sulcus (pSTS), posterior middle temporal gyrus (pMTG), and anterior superior temporal sulcus (aSTS). The Talairach z coordinate is shown below each column of slices. The left side of the slice is the left side of the brain.

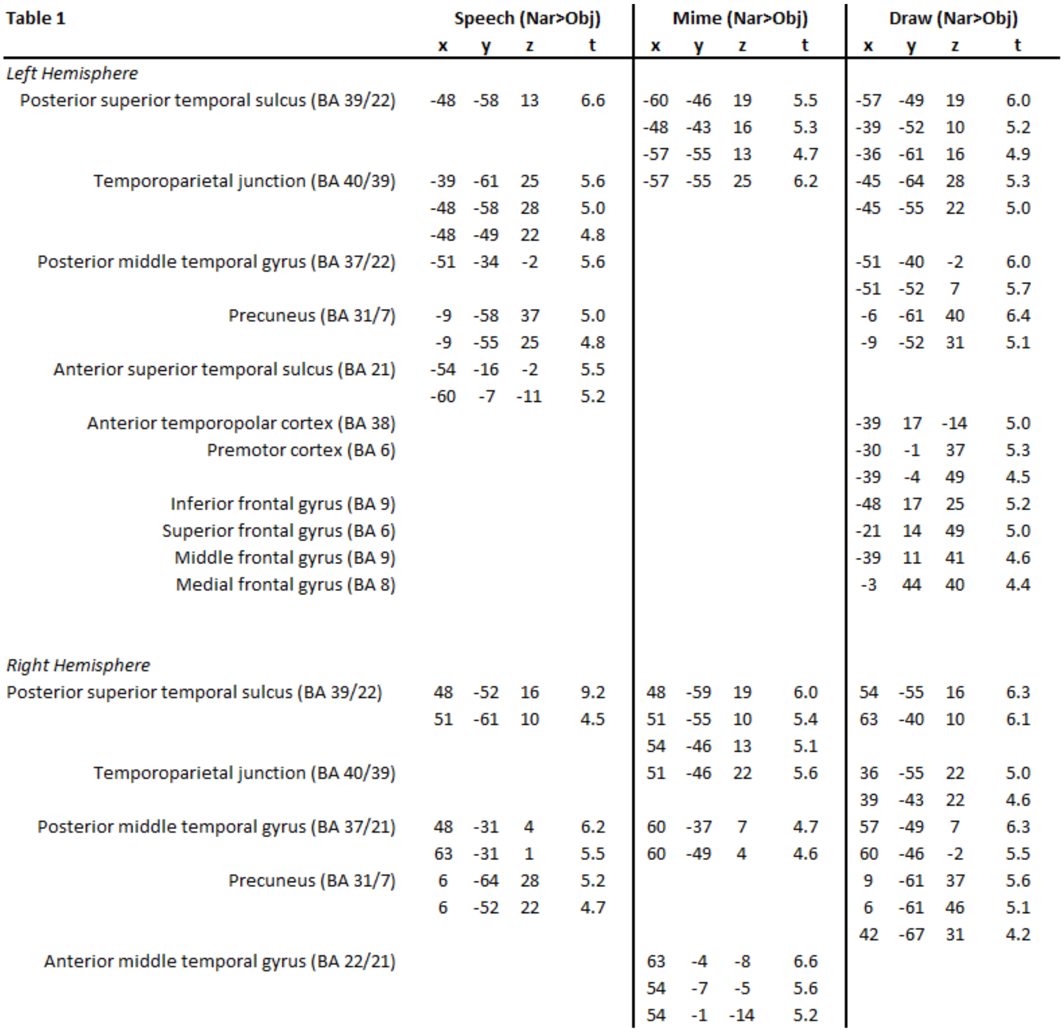


Table 1	Speech (Nar>Obj)				Mime (Nar>Obj)				Draw (Nar>Obj)			
	x	y	z	t	x	y	z	t	x	y	z	t
Left Hemisphere												
Posterior superior temporal sulcus (BA 39/22)	-48	-58	13	6.6	-60	-46	19	5.5	-57	-49	19	6.0
					-48	-43	16	5.3	-39	-52	10	5.2
					-57	-55	13	4.7	-36	-61	16	4.9
Temporoparietal junction (BA 40/39)	-39	-61	25	5.6	-57	-55	25	6.2	-45	-64	28	5.3
	-48	-58	28	5.0					-45	-55	22	5.0
	-48	-49	22	4.8								
Posterior middle temporal gyrus (BA 37/22)	-51	-34	-2	5.6					-51	-40	-2	6.0
									-51	-52	7	5.7
Precuneus (BA 31/7)	-9	-58	37	5.0					-6	-61	40	6.4
	-9	-55	25	4.8					-9	-52	31	5.1
Anterior superior temporal sulcus (BA 21)	-54	-16	-2	5.5								
	-60	-7	-11	5.2								
Anterior temporopolar cortex (BA 38)									-39	17	-14	5.0
Premotor cortex (BA 6)									-30	-1	37	5.3
									-39	-4	49	4.5
Inferior frontal gyrus (BA 9)									-48	17	25	5.2
Superior frontal gyrus (BA 6)									-21	14	49	5.0
Middle frontal gyrus (BA 9)									-39	11	41	4.6
Medial frontal gyrus (BA 8)									-3	44	40	4.4
Right Hemisphere												
Posterior superior temporal sulcus (BA 39/22)	48	-52	16	9.2	48	-59	19	6.0	54	-55	16	6.3
	51	-61	10	4.5	51	-55	10	5.4	63	-40	10	6.1
					54	-46	13	5.1				
Temporoparietal junction (BA 40/39)					51	-46	22	5.6	36	-55	22	5.0
									39	-43	22	4.6
Posterior middle temporal gyrus (BA 37/21)	48	-31	4	6.2	60	-37	7	4.7	57	-49	7	6.3
	63	-31	1	5.5	60	-49	4	4.6	60	-46	-2	5.5
Precuneus (BA 31/7)	6	-64	28	5.2					9	-61	37	5.6
	6	-52	22	4.7					6	-61	46	5.1
									42	-67	31	4.2
Anterior middle temporal gyrus (BA 22/21)					63	-4	-8	6.6				
					54	-7	-5	5.6				
					54	-1	-14	5.2				

Table 4.1: Talairach coordinates for the Narrative > Object subtractions across modalities

Talairach coordinates for the peak activations for the Narrative > Object subtraction for speech, mime, and drawing. Brodmann areas (BA) are indicated in parentheses.

In order to look for cross-modal narrative areas in the brain, we ran conjunction analyses for the Narrative > Object subtractions just described. This included the three pairwise conjunctions and the single three-way conjunction. The results are shown in Figure 3 at a slightly reduced threshold than shown in Figure 2 (uncorrected $p < 0.005$), with Talairach coordinates reported in Tables 2 (pairwise conjunctions) and 3 (3-way conjunction). All pairwise conjunctions showed bilateral activity in TPJ and pSTS, as well as in the precuneus. These areas were also present in the three-way conjunction. Bilateral aSTS activity was only shared between speech and mime (panel A), but not between speech and drawing (panel B), mime and drawing (C), or the three-way conjunction.

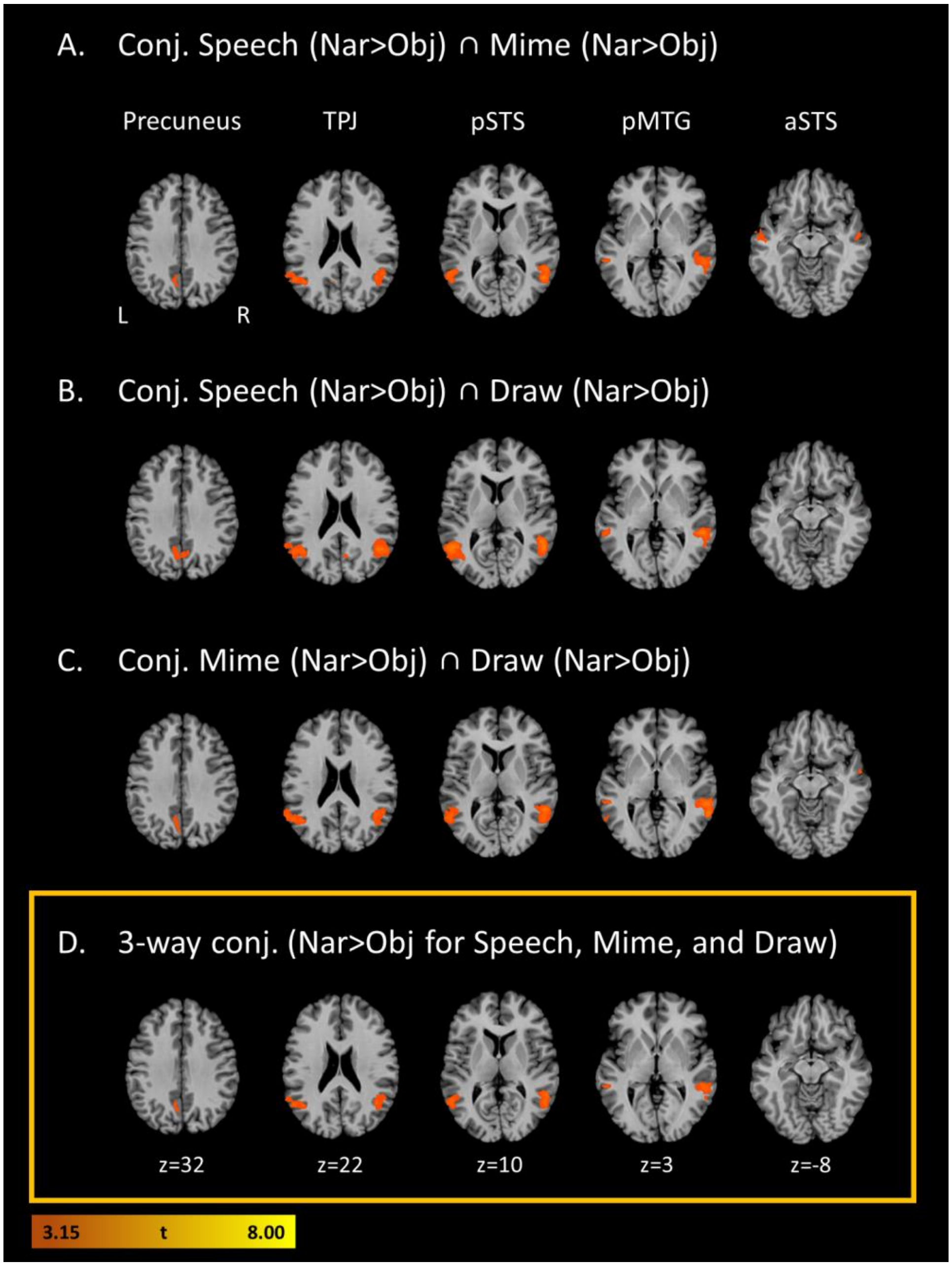


Figure 4.3: Conjunction analyses

Axial views of the conjunction results from the Narrative (Nar) > Object (Obj) contrasts are presented. The first three panels show pairwise conjunctions for A) speech and mime, B) speech and drawing, C) mime and drawing, respectively. The last panel shows the 3-way conjunction of speech, mime and draw. The figure is set up to show activations in the precuneus, temporoparietal junction (TPJ), posterior superior temporal sulcus (pSTS), posterior middle temporal gyrus (pMTG), and anterior superior temporal sulcus (aSTS). Results are displayed at $p < 0.005$, uncorrected. The Talairach z coordinate is shown below each column of slices. The left side of the slice is the left side of the brain.

Table 2	**Speech (Nar>Obj) ∩ Mime (Nar>Obj)**				**Speech (Nar>Obj) ∩ Draw (Nar>Obj)**				**Mime (Nar>Obj) ∩ Draw (Nar>Obj)**			
	x	**y**	**z**	**t**	**x**	**y**	**z**	**t**	**x**	**y**	**z**	**t**
Left Hemisphere												
Posterior superior temporal sulcus (BA 39/22)	-48	-49	13	4.9	-48	-52	10	5.3	-57	-49	19	4.7
					-39	-58	16	4.6	-51	-58	10	4.1
									-48	-46	10	3.9
Temporoparietal junction (BA 40/39)	-51	-55	22	4.6	-51	-55	19	4.8	-42	-58	25	4.6
	-39	-58	22	4.3	-39	-58	25	4.2	-51	-58	22	4.4
					-48	-58	28	4.2				
					-60	-52	22	3.9				
Posterior middle temporal gyrus (BA 37/22)	-54	-37	1	4.4	-51	-34	-2	4.8	-54	-37	1	4.2
									-51	-52	1	3.5
Precuneus (BA 31)	-9	-58	25	4.3	-6	-58	37	4.5	-6	-52	31	3.8
	-3	-67	28	4.0	-9	-52	28	4.4	-3	-67	31	3.7
					-9	-49	37	4.0				
					-6	-58	46	3.9				
Anterior superior temporal sulcus (BA 21)	-48	-13	-8	4.3								
	-54	-4	-8	4.0								
	-57	-13	-8	3.9								
Right Hemisphere												
Posterior superior temporal sulcus (BA 39/22)	48	-55	19	5.9	54	-55	16	6.3	54	-46	13	5.1
	54	-46	10	4.7	51	-46	10	5.0	48	-58	19	4.9
Temporoparietal junction (BA 40/39)	48	-43	25	3.9	45	-55	22	5.0	51	-49	22	4.6
Posterior middle temporal gyrus (BA 37/21)	48	-28	1	4.4	57	-37	-2	5.1	60	-49	4	4.6
	54	-37	4	4.2					60	-37	7	4.6
									45	-43	7	3.8
									45	-28	1	3.6
Precuneus (BA 31/7)					3	-58	37	4.1				
					3	-67	28	4.1				
Anterior middle temporal gyrus (BA 22/21)	57	-7	-8	3.8	45	-28	-2	4.1	57	-4	-8	4.5
	51	-1	-14	3.7								
Superior frontal gyrus (BA 9)					3	53	28	3.3				

Table 4.2: Talairach coordinates for the pairwise conjunctions of the Narrative > Object subtractions

Talairach coordinates for the peak activations for the pairwise conjunctions of the Narrative > Object subtractions for speech and mime; speech and drawing; and mime and drawing. Brodmann areas (BA) are indicated in parentheses.

Table 3	**3-way conjunction**			
	x	**y**	**z**	**t**
Left Hemisphere				
Posterior superior temporal sulcus (BA 39/22)	-51	-52	13	4.5
Temporoparietal junction (BA 40/39)	-51	-58	22	4.4
	-39	-58	25	4.2
Posterior middle temporal gyrus (BA 37/22)	-54	-37	1	4.2
Precuneus (BA 31/7)	0	-64	28	3.6
	-6	-55	31	3.3
Right Hemisphere				
Posterior superior temporal sulcus (BA 39/22)	51	-55	13	4.9
	54	-46	10	4.7
Temporoparietal junction (BA 40/39)	42	-55	22	4.3
	51	-49	22	4.3
Posterior middle temporal gyrus (BA 37/22)	54	-37	4	4.2

Table 4.3: Talairach coordinates for the 3-way conjunction

Talairach coordinates for the 3-way conjunction of the Narrative > Object subtractions for speech, mime, and drawing. Brodmann areas (BA) are indicated in parentheses.

4.5 Discussion

The principal objective of this fMRI study was to carry out the first tri-modal production study of narrative processing with the aim of identifying a narrative hub in the brain. We examined this by looking at the contrast between a narration condition and an object-description condition, and then conjoining this contrast across the three major modalities of narrative expression: speech (language), pantomime, and drawing. In keeping with our predictions based on the narrative literature for speech, we found an effect in the TPJ as well as in the pSTS and precuneus. These areas are strongly associated with "character" processing, as related to both mentalizing and emotional expression (including a person's facial expression, vocal prosody, and expressive body motion). We will argue below that a potentially unifying explanation for our activation profile might relate to the concept of agency.

During both everyday conversations and performances of theatrical works, people tell stories about themselves and others through a process of narration. In most stories, there is a central protagonist (be it oneself or a character) who drives the actions of the narrative and who serves as a focal point by which perceivers (listeners or readers) understand the goal structure of the story's actions (Abbott, 2008; Mandler, 1984; Rumelhart, 1975; Stein & Glen, 1979). As a result, people see the central conflict of the story from the protagonist's perspective. The narration condition of our experiment was designed to tap into protagonist processing. Since objects were part of both the control condition and the narrative condition (the headlines described transitive actions on objects), what was unique to the narrative condition was a protagonist and his/her actions. Hence, the activation results most likely reflect this.

In our initial pilot testing of the control condition, we allowed participants to describe the uses of the objects they were describing. We noticed that this decreased the subtraction signal quite a bit (data not shown), most likely because a conception of the uses of an object requires the presence of a virtual agent who serves as the vehicle for the action on the object. In response to these pilot results, we placed restrictions on the control task such that participants were required to focus exclusively on the structural properties of an object, and not its uses. This served to accentuate the contrast between the narrative condition and the description condition. It is interesting to point out that many studies of mentalizing about people's psychological states use a control condition that is based on "trait judgments" about the same people, for example their gender or physical features (Denny et al., 2012). Hence, while our narrative condition was in no way explicitly mentalistic, our experimental design showed similarities to many studies of theory-of-mind processing that use trait judgments as the control condition for their subtractions.

Two major neural systems have been invoked in understanding human action, including through narrative. The mirror neuron system is a sensorimotor system that is thought to mediate an understanding of the actions of agents (Arbib, 2012). From our standpoint, we can think of this system as being a "gestural" system, since it deals with the visually-perceivable motor gestures of an agent. This is in contrast to the mentalizing system, which deals with inferring the unobservable mental states of these agents (Frith & Frith, 2003; Nichols & Stich, 2003). Although our narrative task was, on the surface, quite gestural – the headlines were statements of transitive actions with no mentalistic content – a key question for the interpretation of the imaging results is whether people intrinsically tend to adopt a mentalistic stance when they recount stories about protagonists, even if the task does not require them to infer the mental states of the agents. If so, it would suggest that people approach storytelling less as a recounting of the event-sequences that make up the plot (as in studies of “event perception”, Radvansky & Zacks, 2011) than as a connection with the mental states of the characters. In other words, people may carry out storytelling in a character-based, rather than a plot-based, manner as their default mode of operation. In fact, even for studies of event perception, the activations we observed are those that are associated with characters, rather than objects, space, or causation in their narrative stimuli (Zacks et al., 2010).

4.5.1 – The unimodal analyses: Low-level and high-level subtractions

The low-level subtractions against fixation (Figure 1) revealed mainly sensorimotor areas involved in task performance. On the production side, this included the sensorimotor cortex and SMA. For the sensorimotor cortex, we observed the expected somatotopic distinction between two general effector systems used for communication: the orofacial motor cortex for the speech task, and both the hand and arm motor areas for the two manual tasks. The pantomime and drawing conditions also showed prominent activations in the posterior parietal cortex not seen in the speech condition, including the SPL, IPS, and IPL. Such activations are thought to mediate visuomanual translation during visually-guided hand-movement tasks, and are common areas of activation across studies of both pantomime (Vingerhoets & Clauwaert, 2015) and drawing (Yuan & Brown, 2014, 2015). Additionally, perceptual areas that were stimulated by the outcome of the production process were activated, including auditory areas for speech and visual areas for miming and drawing. The latter included area V3A that we have described previously as being important for perceiving the image that dynamically accumulates as a result of the process of drawing (Yuan & Brown, 2014; see also Thaler & Goodale, 2011). This activation was considerably weaker in the pantomime condition, in which this type of visual accumulation does not occur. Activations in low-level visual areas were also due to the presence of the stimulus prompts, and were therefore present in the speech conditions as well (Figure 1).

Next, the high-level subtractions were designed to eliminate the abovementioned sensorimotor activations and thereby isolate components specifically associated with narrative processing of the protagonist's actions. This permitted a cognitive contrast between narration (a recounting of the actions of a protagonist) and description (an enumeration of an object's properties, separate from a person's interaction with it). The efficacy of the control condition was shown by the fact that virtually all of the sensorimotor areas were eliminated in the narration-versus-description subtraction for each modality (Figure 2), suggesting that sensorimotor activations were well accounted for by the control condition.

What was left over in each case were areas associated with the mentalizing and social-cognition networks, which will be discussed in the next section on the conjunction analysis. For now, we focus on two additional areas that came up in the high-level subtractions, namely the anterior superior temporal sulcus (aSTS, BA 38) and the posterior middle temporal gyrus (pMTG, BA 21). The aSTS came up in the high-level subtractions for both speech (left hemisphere) and pantomime (right hemisphere), but not drawing. This is a part of the brain that, in auditory perception studies, is thought to show a preference for stimuli that are complex and coherent, compared to either elemental stimuli or complex stimuli that are meaningless or incoherent (Scott et al., 2000, 2006). For example, it shows a preference for the perception of sentences, compared to words or phonemes (DeWitt & Rauschecker, 2012). This supports a role of the anterior temporal region in discourse processing, text integration, and the generation of meaning beyond the single-sentence level (Mason & Just, 2006), something that would be important for narrative processing (Mar, 2011). In a study by Brown et al. (2006) on the vocal generation of completions for both sentence fragments and musical-melody fragments, the authors found the aSTS to be activated in both the speech and music conditions. It is interesting that we saw activation of this area for pantomime in the present study, since pantomime tends to share with speech its linear, sequential nature, whereas individual drawings are not typically created in a linear and episodic fashion. An exception is the processing of comic strips, whose perception does show activation in the anterior temporal region (Osaka et al., 2014). This linear stringing together of events is a critical part of narrative processing. The remaining activations in the unimodal high-level contrasts relate to the mentalizing network and will be discussed in the next section about cross-modal processing.

The pMTG's connection with narrative relates to its central role in semantic processing. Not only is the pMTG a key semantic area of the brain (Binder et al., 2009), but its function is strongly cross-modal, as shown by its activation in semantic tasks using visual words, spoken words, pictures, and gestures (Krieger-Redwood et al., 2015; Visser et al., 2012; Xu et al., 2009). Hence, the pMTG might function as an amodal semantic area, although some studies link it most strongly to visual processing (Devereux et al, 2013). In Huth et al.'s (2016) fine-grained analysis of semantic processing throughout the brain, the pMTG was found to be associated with "numeric, tactile and visual concepts", while the TPJ was found to be associated with "social, emotional and mental concepts". This suggests that, while the TPJ might deal with semantic concepts related to the characters of a story, the pMTG might deal with more general semantic aspects of the story's setting, such as visual features of the scene. Given that the pMTG projects to the TPJ region via the vertical limb of the arcuate fasciculus (Catani et al., 2005), this might suggest a sequential model of narrative processing in which setting information in the pMTG is transmitted to and combined with character information in the TPJ and pSTS.

4.5.2 – The cross-modal analyses: 2-way and 3-way conjunctions

The three unimodal subtractions showed strikingly similar patterns of activation among themselves, focusing mainly on areas of the mentalizing and social-cognition networks involved in character processing. There were few additional brain areas in each subtraction beyond those that were common across the three (Figure 2). As a result, the cross-modal conjunctions, whether 2-way or 3-way, showed these same areas. The pMTG was present bilaterally in all conjunctions, while the aSTS was present only in the pairwise conjunction between speech and pantomime for the reasons suggested above. Aside from this, the remaining areas were common across the analyses. These included the TPJ, pSTS, and precuneus. These areas were not seen for the most part in the narrative condition against fixation, but mainly came up in the high-level contrasts. The only one with a clear connection with sensory processing is the pSTS, although it too has a connection with "character" processing to the extent it is about multimodally perceiving the speech prosody, facial expression and body movement of others as part of the process of social cognition (Biau et al., 2016; Campanella & Belin, 2007; Deen et al., 2015; Kreifelts et al., 2009). As mentioned in the Introduction, the pSTS is a prominent area of overlapping activity in cross-modal studies of language and narrative processing for speech, gesture, and visual images. While the pSTS is typically seen in studies of perception, we demonstrate its role in production as well, as seen in previous studies of narrative production (AbdulSabur et al., 2014).

4.5.3 – A cross-modal narrative hub?

Humans possess a “narrative triad” of three major modalities for conveying narrative, each of which is highlighted in one of the general branches of the arts: speech/language in theatre and literature; pantomime in mime theatre and narrative forms of dance; and image generation in drawing, painting, sculpting, and related forms of graphic art. Our experiment attempted to factor out the modality-specific components of narrative in order to identify an area or set of areas that would be common to the multiple forms of narrative expression and would thus serve as an amodal conceptual area for narrative generation. One hypothesis that we had in mind for this was that a cross-modal production area should also be a region of production/perception overlap, since such an overlap would be one indicator of the pleiotropy of the brain area. While we did not examine narrative perception per se in our experiment (perception was an integral component of all of the task conditions), both Silbert et al. (2014) and AbdulSabur et al. (2014) compared narrative production with perception, and found overlap in a large number of areas. Among them were all of the areas seen in our conjunction analyses, including the TPJ, pSTS, precuneus, pMTG, and aSTS. Hence, this satisfies our criterion that these areas be multifunctional, not only cross-modally, but also sensorimotorically.

The activations obtained in both the high-level subtractions and conjunctions suggest that narrative production is more associated with character than plot, despite the field of literary studies prioritizing plot over character since the time of Aristotle. The activations showed an orientation toward encoding features of the protagonist. The TPJ and precuneus are components of the classic mentalizing network involved in making inferences about the beliefs, desires and emotions of other people as well as oneself (Denny et al., 2012; Frith & Frith, 2003). This network is typically probed in neuroimaging studies with tasks that have participants explicitly think about the beliefs and emotions of others, including false-belief paradigms. Hence, it is virtually always seen in perceptual tasks, rather than production tasks.

As mentioned above, there was nothing overtly mentalistic about our narrative condition. Participants were asked to recount short narratives about transitive actions conveyed in the headlines. Nothing about the task required them to mentalize about the protagonist. Given that the mirror system emphasizes observable gestures, and that the mentalizing system emphasizes unobservable mental states, our tasks would seem to better fit the profile of the mirror system, since they were oriented towards describing the observable behaviors (gestures) of agents, as well as event sequencing. A gestural emphasis is seen in studies of pantomime production, where people have to act out simple actions or represent objects (Schippers et al., 2009). One could argue that by removing the object through our control condition, what would be left over would be the action itself, hence perhaps reflected in the activation of premotor areas or related areas involved in motor sequencing, such as inferior frontal gyrus and/or basal ganglia (Shmuelof & Krakauer, 2011). However, we did not see that pattern. It appears that participants, instead of focusing on the actions per se, focused on the protagonist and his/her underlying mental states. Perhaps the activation most indicative of a gestural interpretation was the pSTS. However, even that area reflects the social aspects of communicative expression, rather than motoric parameters related to, say, grasping an object (as the surgeon might grasp the scissors found inside of the patient in one of our headlines), which is a classic stimulus for the mirror system.

One possible interpretation of the results is that people assume a mentalistic stance as their default mode of processing stories, oriented towards the characters in the stories. Because we did not manipulate any aspect of mentalizing in our narrative condition, we had no control over whether participants did or did not process the headlines mentalistically when performing them. While we can easily imagine that participants could have engaged in this mode of processing for speech, we find it unlikely that they would have done so for drawing, where participants simply drew a scenic description of the headline, for example a patient lying on an operating table, and a surgeon next to him finding a pair of scissors inside of him. Regarding pantomime, there are two distinct modes of performing a pantomime (Boyatzis and Watson, 1993; O'Reilly, 1995; Suddenhoff, Flecher-flinn, and Johnston, 1999): egocentrically as an open-handed gesture (as in miming a tennis serve with an imaginary racquet in one hand and an imaginary tennis ball in the other) or allocentrically as an act of object substitution (as such when a person mimes "call me" by using their hand to embody a telephone receiver). While these two modes of pantomime involve strikingly different relationships between the mimer's body and the mimed object, it does not seem that one format is inherently more mentalistic than the other, although one could argue that the egocentric mode is more self-oriented and that the allocentric mode more other-oriented.

In future studies, we could directly manipulate mentalistic processing by the participants. We could compare a recounting task (i.e., simply describing what happened without reference to motivation or causation) with a mentalistic task (i.e., describing what happened with reference to the protagonist's goals and the consequences therein). In Schippers et al. (2009), participants either produced or perceived pantomimic gestures with the aim of gauging the relative involvement of the mirror and mentalizing systems. While components of both systems were involved in perceiving pantomime (including the TPJ), only mirror-neuron areas were involved in production. Half of the stimuli in that study were objects, which would have fit into our control condition. The other half were actions, e.g., "peel fruit", but without explicit protagonists. As mentioned above, when we allowed pilot participants to include the uses of objects in our control condition, this had the effect of reducing activity in mentalizing areas, most likely because of the virtual agent that underlies egocentric pantomiming.

Overall, our results suggest that when people generate narratives, they assume a mentalistic stance that is driven by their psychological conception of the protagonist, rather than a purely gestural approach to the observable event-sequence of actions carried out by the protagonist. In doing so, they cognitively prioritize character over plot. While the mentalizing system is invoked in perceptual studies of narrative (Mar, 2011), what is lacking is a comparable production system for narrative, not least one from a character-centered perspective. The mirror system is the classic system for aligning perception and production. However, it is principally a gestural system. A theoretical understanding of narrative requires a comparable sensorimotor system for mentalistic production and perception.

In fact, we would argue that *narration is the production-counterpart to the perceptual process of theory-of-mind*. If theory-of-mind is the private process of inferring a character's motivations, beliefs and emotions, narration is the public process of externalizing such motivations, beliefs and emotions through acts of communication. Comparisons between story production and perception (AbdulSabur et al., 2014; Silbert et al., 2014) have demonstrated that all of the areas seen in our high-level subtractions and conjunctions were comparably present in production and perception. Other areas included the inferior frontal gyrus, anterior temporopolar cortex, and dorsolateral prefrontal cortex.

One concept that has the potential to unify many key aspects of narrative processing is agency, which refers to the intentionality to act and the sense of voluntary control over one's action and one's ability to achieve desired outcomes (Haggard, 2017). The vast majority of the psychology literature on agency deals with "self-agency", in other words perceptions about one's own capacity to act. What is lacking is a concept of "other-agency" that seems to underlie third-person storytelling. Much narrative, whether during conversation or in literature, is about recounting the agency of protagonists. According to both literary theory and psychological models of story grammars, stories are generally about the goal-directed behaviors of protagonists, their efficacy at overcoming obstacles, and their capacity to achieve their goals through intentional actions, often to solve problems

(Abbott, 2008; Mandler, 1984; Rumelhart, 1975; Stein & Glen, 1979). Stories are very much outcome-driven, based on an arc-like sequence of events (the “dramatic arc”) that resembles the psychological progression of problem-solving episodes. Narrative models contain all of the ingredients of models of self-agency, but apply them to “others”. Hence, storytelling is typically a third-person recounting of the agency of protagonists.

Interestingly, neuroimaging analyses of instrumental agency show an effect in the TPJ. However, the connection is much stronger with “external” agency than self-agency. External agency, in the context of the operant actions looked at in this literature, refers not to social agency in human interactions (as per stories) but rather the ability to control the outcomes of instrumental actions, such as pushing a button to cause a tone to be emitted. To the best of our knowledge, there is minimal neuroimaging literature exploring social agency between interaction partners, not least the agency of people other than oneself. One function of the TPJ activations seen in our subtractions and conjunctions could be related to the attribution of agency to protagonists, again consistent with a general orientation towards characters, rather than episodic sequences, in stories. It is interesting to point out that a meta-analytic comparison of other-judgments versus self-judgments in mentalizing paradigms revealed that many brain areas were commonly activated between other- and self-processing, while the TPJ bilaterally and the precuneus were among a small number of areas that showed a preference for other-processing over self-processing in the direct contrast (Denny et al., 2012). This is again compatible with the notion that the TPJ might be specialized for processing “other” agency more than self-agency, as would be important in narrative production and perception. Overall, components of the mentalizing network (TPJ, precuneus) in combination with the pSTS seem to constitute a set of hub areas for narrative production, with a special focus on the protagonist, perhaps related to their agency.

4.5.4 – Evolutionary implications

Evolutionary models of language origin typically contrast speech and pantomime accounts, as mentioned in the Introduction. In our experiment, the comparison between speech and pantomime, when examined against a common fixation condition, revealed clear sensorimotor differences (Figure 1), both somatotopic differences motorically and sensory differences perceptually (in contrast to some studies that have shown far more overlap between them; Andric et al., 2013). It was only when their respective control conditions were subtracted out that a deep similarity between them appeared, as related to mentalizing, social cognition, and semantics. Interestingly, there was minimal evidence of inferior frontal gyrus (IFG) activation in either the low-level or high-level subtractions, despite the established involvement of this area in sequential motor tasks (Shmuelof &

Krakauer, 2011). Instead, the aSTS was seen commonly for speech and pantomime (but not drawing), in keeping with its role in discourse processing cross-modally. Likewise for the pMTG and its role in cross-modal semantic processing.

Both vocal and gestural models of language attempt to account for the origins of syntax. As mentioned in the Introduction, language grammar seems to have an intrinsically narrative structure to it, being efficient at describing who did what to whom, in other words agency. Standard subject-verb-object models of syntactic structure (Tallerman, 2015) essentially encapsulate the kinds of transitive actions that we examined in our headlines. A large majority of languages operate on an Agent First basis, putting the actor before either the action or the target of the action. To the extent that agency is one of the most fundamental things that is conveyed in grammars (and that is lacking in so-called proto-languages; Bickerton, 1995), then our results have application to evolutionary models of language. In particular, the imaging results that were obtained in the most purely linguistic condition (speech) were replicated almost identically in the non-linguistic conditions of pantomime and drawing. This cross-modal similarity suggests that the capacity of syntax to represent agency can be achieved through non-linguistic means employing essentially the same brain network.

A number of biological theories of language propose that syntax emerged from basic processes of motor sequencing (Arbib, 2012; Fitch, 2011; Jackendoff, 2011). While this might account for grammar's connection with object-directed actions – in other words, the gestural level of representation – it may not do justice to the sense of agency that is well-contained in syntactic structure. Hence, we suggest that another important evolutionary ingredient in the emergence of syntax – beyond the "plot" elements contained in motor sequencing – would be the incorporation of circuits that mediate the sense of agency, not least "other" agency. To be clear, we are not arguing that the TPJ and pSTS are syntax areas. We are simply suggesting that, while circuits in the IFG more typically associated with syntax (Zaccarella & Friederici, in press) might mediate the gestural level of language, the TPJ might have a stronger connection with agents in the overall scheme of language, discourse, and narrative. Agency can be conveyed linguistically through speech and sign, but it can also be conveyed non-linguistically through pantomime (iconic gesturing) and drawing.

Given that high-level functions like agency, theory-of-mind and narrative seem to be human-specific functions, it is possible that novel areas in the human brain mediate them? Mars et al. (2013) examined the resting state connectivity of the TPJ region in monkeys and humans, and observed that the human TPJ was most similar in its connectivity fingerprint to the monkey mid-STS, rather than to a monkey homologue of the TPJ. The authors proposed that "one potential evolutionary route for the human TPJ region to have appeared is by an expansion and subsequent division and specialization of mid-STS regions” in the common ancestor of humans and monkeys (p. 10809). In other words, while the TPJ might be evolutionarily derived from an ancestral mid-STS structure, it might in fact be a novel specialization of human brain. According to Patel et al. (2015), an anatomico-functional region resembling the TPJ seems to be absent in monkeys, and therefore the TPJ might be a human-specific brain area. While the fMRI tasks that examined these cross-species differences looked much more at visual attention than linguistic or mentalizing processes, they do hold hope for the TPJ region as being a novel area for cross-modal narrative processing.

4.6 Conclusions

In this first tri-modal fMRI study of narrative production, we observed results that suggest that people generate stories in an intrinsically mentalistic fashion focused on the protagonist, rather than in a purely gestural manner related to the observable action sequence. The same set of mentalizing and social-cognition areas came up with each of the three modalities of production that make up the narrative triad, pointing to a common set of cognitive operations across modalities. These operations are most likely rooted in character processing, as related to a character’s intentions, motivations, beliefs, emotions, and actions. Hence, narratives – whether spoken, pantomimed, or drawn – seem to be rooted in the communication of “other-agency”. Given that theories of the origins of language are polarized with respect to whether language began as speech or pantomime, a brain network that can cross-modally mediate the communication of agency, both self and other, might lay the conceptual foundations for language’s syntactic system as well as for the multimodal expression of narratives in art forms spanning from theatre to dance to painting.

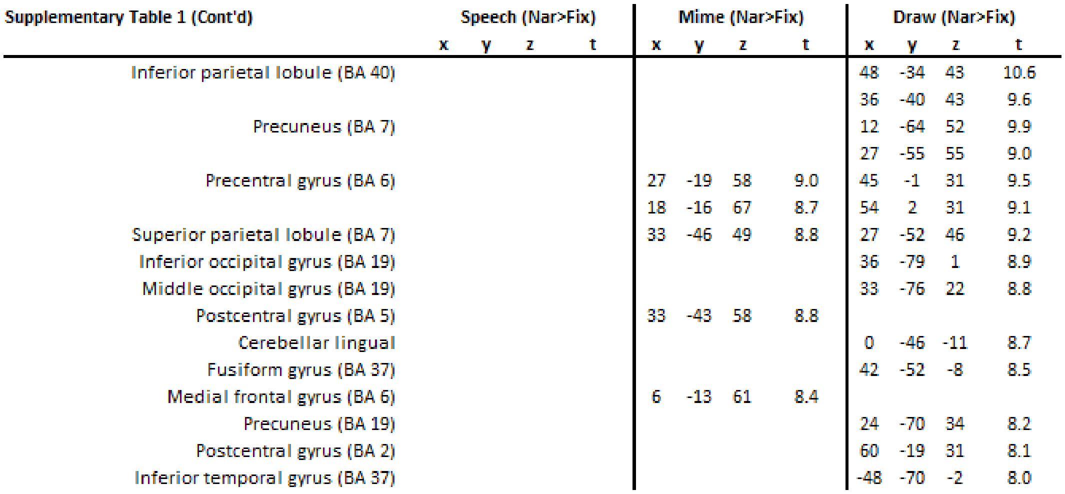
4.7 Acknowledgements

This work was supported by a grant to S.B. from NSERC.

4.8 Supplementary Information

Supplementary Table 1	Speech (Nar>Fix)				Mime (Nar>Fix)				Draw (Nar>Fix)			
	x	y	z	t	x	y	z	t	x	y	z	t
Left Hemisphere												
Medial frontal gyrus (BA 6)	-6	-1	55	15.2	-3	-7	55	10.4	-3	-4	52	9.1
Postcentral gyrus (BA 3/1/2)	-60	-10	25	11.1	-30	-34	52	10.4	-33	-34	49	14.3
					-42	-25	52	9.7	-51	-28	43	11.6
					-27	-28	64	8.1				
Precentral gyrus (BA 6/4)	-48	-16	40	9.3	-36	-16	52	11.2	-24	-16	55	12.0
					-27	-16	55	9.8	-54	-1	31	9.9
									-36	-16	55	9.6
									-36	-25	64	8.2
Inferior parietal lobule (BA 40)					-42	-43	46	11.9	-42	-43	49	14.0
					-36	-43	55	9.1	-42	-34	43	12.2
									-36	-43	37	9.5
Superior parietal lobule (BA 7)					-33	-58	58	9.4	-33	-58	52	12.2
									-24	-61	58	10.0
									-24	-67	43	10.0
									-15	-64	55	9.6
									-27	-52	43	8.9
Inferior frontal gyrus (BA 9)					-51	5	25	11.6				
Lingual gyrus (BA 17)									-18	-85	-2	11.2
									-21	-76	4	10.2
Supramarginal gyrus (BA 40)					-39	-43	37	10.7				
Fusiform gyrus (BA 18)									-18	-85	-14	10.1
Postcentral gyrus (BA 40)					-39	-34	55	9.9				
					-42	-28	43	8.2				
Culmen	-24	-58	-23	9.1								
Cingulate gyrus (BA 24)					-9	2	43	8.9				
Precuneus (BA 7)					-15	-61	49	8.8				
Middle frontal gyrus (BA 6)					-27	-7	55	8.6				
					-18	-7	61	8.4				
Right Hemisphere												
Precentral gyrus (BA 4)	48	-13	37	10.1								
Postcentral gyrus (BA 3)	57	-10	25	9.8								
Declive	15	-61	-17	9.4	30	-58	-20	10.6	3	-67	-11	12.3
	3	-70	-11	9.1	12	-55	-14	9.5	21	-67	-17	8.3
	6	-79	-11	8.3	3	-67	-11	9.1				
Superior temporal gyrus (BA 41)	42	-28	7	8.6								
Lingual gyrus (BA 17/18)	21	-79	-8	8.4	15	-82	-5	9.3	15	-82	-5	11.8
									12	-91	-2	8.1
Postcentral gyrus (BA 40)					36	-34	58	10.0				
Culmen					6	-55	-5	9.1	6	-55	-8	14.9
									30	-49	-23	9.4
									3	-61	-20	8.9
Middle temporal gyrus (BA 37)									54	-52	-8	11.1
Middle frontal gyrus (BA 6)									27	-13	61	10.7
									54	2	40	8.4

(Cont’d on next page)



Supplementary Table 1 (Cont'd)	Speech (Nar>Fix)				Mime (Nar>Fix)				Draw (Nar>Fix)			
	x	y	z	t	x	y	z	t	x	y	z	t
Inferior parietal lobule (BA 40)									48	-34	43	10.6
									36	-40	43	9.6
Precuneus (BA 7)									12	-64	52	9.9
									27	-55	55	9.0
Precentral gyrus (BA 6)					27	-19	58	9.0	45	-1	31	9.5
					18	-16	67	8.7	54	2	31	9.1
Superior parietal lobule (BA 7)					33	-46	49	8.8	27	-52	46	9.2
Inferior occipital gyrus (BA 19)									36	-79	1	8.9
Middle occipital gyrus (BA 19)									33	-76	22	8.8
Postcentral gyrus (BA 5)					33	-43	58	8.8				
Cerebellar lingual									0	-46	-11	8.7
Fusiform gyrus (BA 37)									42	-52	-8	8.5
Medial frontal gyrus (BA 6)					6	-13	61	8.4				
Precuneus (BA 19)									24	-70	34	8.2
Postcentral gyrus (BA 2)									60	-19	31	8.1
Inferior temporal gyrus (BA 37)									-48	-70	-2	8.0

SI Table 4.1

Talairach coordinates for the Narrative > Fixation subtraction for each modality, corrected for multiple comparisons using FDR $p < 0.005$.

4.9 References

Abbott, H. P. (2008). *The Cambridge introduction to narrative*, 2nd ed. Cambridge, UK: Cambridge University Press.

AbdulSabur, N. Y., Xu, Y., Liu, S., Chow, H. M., Baxter, M., Carson, J., & Braun A. R. (2014). Neural correlates and network connectivity underlying narrative production and comprehension: A combined fMRI and PET study. *Cortex*, 57, 107-127.

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

Arbib, M. (2012). *How the brain got language: The mirror system hypothesis.* Oxford: Oxford University Press.

Armstrong, D. F., & Wilcox, S. E. (2007). *The gestural origin of language.* Oxford: Oxford University Press.

Awad, M., Warren, J. E., Scott, S. K., Turkheimer, F. E., & Wise R. J. (2007). A common system for the comprehension and production of narrative speech. *Journal of Neuroscience, 27*, 11455-11464.

Bickerton, D. (1995). *Language and human behavior.* Seattle: University of Washington Press.

Beattie, G. (2016). *Rethinking body language: How hand movements reveal hidden thoughts.* London: Routledge.

Biau, E., Morís Fernández, L., Holle, H., Avila, C., & Soto-Faraco, S. (2016). Hand gestures as visual prosody: BOLD responses to audio-visual alignment are modulated by the communicative nature of the stimuli. *Neuroimage, 132*, 129-137.

Binder, J. R., Desai, R. H., Graves, W. W., & Conant, L. L. (2009). Where is the semantic system? A critical review and meta-analysis of 120 functional neuroimaging studies. *Cerebral Cortex, 19,* 2767-2796.

Boyatzis, C. J., & Watson, M. W. (1993). Preschool children's symbolic representation of objects through gestures. *Child Development, 64,* 729-735.

Brown, S., Martinez, M. J., & Parsons, L. M. (2006). Music and language side by side in the brain. *European Journal of Neuroscience, 23,* 2791-2803.

Campanella, S., & Belin, P., 2007. Integrating face and voice in person perception. *Trends in Cognitive Sciences, 11,* 535-543.

Carter, R. M., & Huettel, S. A. (2013). A nexus model of the temporal-parietal junction. *Trends in Cognitive Sciences, 17,* 328-336.

Catani, M., Jones, D. K., & ffytche, D. H. (2005). Perisylvian language networks of the human brain. *Annals of Neurology, 57,* 8-16.

Corballis, M. C. (2002). *From hand to mouth: The origins of language.* Princeton: Princeton University Press.

Deen, B., Koldewyn, K., Kanwisher, N., & Saxe R. (2015). Functional organization of social perception and cognition in the superior temporal sulcus. *Cerebral Cortex, 25,* 4596–4609.

Denny, B. T., Kober, H., Wager, T. D., & Ochsner, K. N. (2012). A meta-analysis of functional neuroimaging studies of self- and other judgments reveals a spatial gradient for mentalizing in medial prefrontal cortex. *Journal of Cognitive Neuroscience, 24,* 1742-1752.

Devereux, B. J., Clarke, A., Marouchos, A., & Tyler, L. K. (2013). Representational similarity analysis reveals commonalities and differences in the semantic processing of words and objects. *Journal of Neuroscience, 33,* 18906-18916.

DeWitt, I., & Rauschecker, J. P. (2012). Phoneme and word recognition in the auditory ventral stream. *Proceedings of the National Academy of Sciences, 109,* E505-514.

Fitch, W. T. (2011). The evolution of syntax: An exaptationist perspective. *Frontiers in Evolutionary Neuroscience, 3,* 9.

Frith, U., & Frith, C. D. (2006). Development and neurophysiology of mentalizing. *Philosophical Transactions of the Royal Society of London B, 358,* 459-473.

Haggard, P. (2017). Sense of agency in the human brain. *Nature Reviews Neuroscience, 18,* 197-208.

Hassabis, D., Spreng, R. N., Rusu, A. A., Robbins, C. A., Mar, R. A., & Schacter, D. L. (2014). Imagine all the people: How the brain creates and uses personality models to predict behavior. *Cerebral Cortex, 24,* 1979-1987.

Huth, A. G., de Heer, W. A., Griffiths, T. L., Theunissen, F. E., & Gallant, J. L. (2016). Natural speech reveals the semantic maps that tile human cerebral cortex. *Nature, 532,* 453-458.

Kendon, A. (2015). Semiotic diversity in utterance production and the concept of 'language'. *Philosophical Transactions of the Royal Society of London B, 369,* 20130293.

Kreifelts, B., Ethofer, T., Shiozawa, T., Grodd, W., & Wildgruber, D. (2009). Cerebral representation of non-verbal emotional perception: fMRI reveals audiovisual integration area between voice- and face-sensitive regions in the superior temporal sulcus. *Neuropsychologia, 47,* 3059-3066.

Jackendoff, R. (1999). Possible stages in the evolution of the language capacity. *Trends in Cognitive Sciences, 3,* 272-279.

Jouen, A. L., Ellmore, T. M., Madden, C. J., Pallier, C., Dominey, P. F., & Ventre-Dominey, J. (2015). Beyond the word and image: Characteristics of a common meaning system for language and vision revealed by functional and structural imaging. *Neuroimage, 106,* 72-85.

Kircher, T., Straube, B., Leube, D., Weis, S., Sachs, O., Willmes, K., Konrad, K., & Green, A. (2009). Neural interaction of speech and gesture: Differential activations of metaphoric co-verbal gestures. *Neuropsychologia, 47,* 169-79.

Krieger-Redwood, K., Teige, C., Davey, J., Hymers, M., & Jefferies, E. (2015). Conceptual control across modalities: Graded specialisation for pictures and words in inferior frontal and posterior temporal cortex. *Neuropsychologia, 76,* 92-107.

MacNeilage, P. F., & Davis, B. L. (2005). The frame/content theory of evolution of speech: A comparison with a gestural-origins alternative. *Interaction Studies, 6,* 173-199.

Mandler, J. M. (1984). *Stories, scripts, and scenes: Aspects of schema theory.* Hillsdale, NJ: L. Erlbaum Associates.

Mar, R. A. (2011). The neural bases of social cognition and story comprehension. *Annual Review of Psychology, 62,* 103-34.

Mars, R. B., Sallet, J., Neubert, F. X., & Rushworth, M. F. (2013). Connectivity profiles reveal the relationship between brain areas for social cognition in human and monkey temporoparietal cortex. *Proceedings of the National Academy of Sciences, 110,* 10806-10811.

Mason, R. A., & Just, M. A. (2006). Neuroimaging contributions to the understanding of discourse processes. In M. Traxler & M. A. Gernsbacher (eds.), *Handbook of psycholinguistics* (pp. 765-799). Amsterdam: Elsevier

McGinn, C. (2015). *Prehension: The hand and the emergence of humanity.* Cambridge, MA: MIT Press.

McNeill, D. (2005). *Gesture and thought*. Chicago: University of Chicago Press.

Nichols, S., & Stich, S. (2003). *Mindreading: An integrated account of pretense, self- awareness and understanding other minds.* Oxford: Oxford University Press.

O'Reilly, A. W. (1995). Using representations: Comprehension and production of actions with imagined objects. *Child Development, 44,* 309-314.

Osaka, M., Yaoi, K., Minamoto, T., & Osaka, N. (2014). Serial changes of humor comprehension for four-frame comic Manga: *An fMRI study. Scientific Reports, 4,* 5828.

Patel, G. H., Yang, D., Jamerson, E. C., Snyder, L. H., Corbetta, M., & Ferrera, V. P. (2015). Functional evolution of new and expanded attention networks in humans. *Proceedings of the National Academy of Sciences, 112,* 9454-9459.

Radvansky, G. A., & Zacks, J. M. (2011). Event perception. *Wiley Interdisciplinary Reviews: Cognitive Science, 2,* 608-620.

Redcay, E., Velnoskey, K. R., & Rowe, M. L. (2016). Perceived communicative intent in gesture and language modulates the superior temporal sulcus. *Human Brain Mapping, 37,* 3444-3461.

Rumelhart, D. E. (1975). Notes on a schema for stories. In D. G. Brown & A. Collins (Eds.), Representation and understanding: *Studies in cognitive science* (pp. 211-236). New York: Academic Press.

Schippers, M. B., 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,* e6801.

Scott, S. K., Blank, S. C., Rosen, S., & Wise, R. J. S. (2000). Identification of a pathway for intelligible speech in the left temporal lobe. *Brain, 123,* 2400-2406.

Scott, S. K., Rosen, S., Lang, H., & Wise, R. J. (2006). Neural correlates of intelligibility in speech investigated with noise vocoded speech: A positron emission tomography study. *Journal of the Acoustical Society of America, 120,* 1075-1083.

Shmuelof, L., and Krakauer, J. W. (2011). Are we ready for a natural history of motor learning? *Neuron, 72,* 469-476.

Silbert, L. J., Honey, C. J., Simony, E., Poeppel, D., & Hasson, U. (2014). Coupled neural systems underlie the production and comprehension of naturalistic narrative speech. *Proceedings of the National Academy of Sciences, 111,* E4687-4696.

Stein, N. L., & Glen, C. G. (1979). An analysis of story comprehension in elementary school children. In R. O. Freedle (ed.) New directions in discourse processing: Vol. 2. *Advances in discourse processes* (pp. 53-120). Hillsdale, NJ: Erlbaum.

Straube, B., Green, A., Weis, S., & Kircher, T. (2012). A supramodal neural network for speech and gesture semantics: An fMRI study. *PLoS ONE, 7,* e51207.

Suddendorf, T., Fletcher-Flinn, C., & Johnston, L. (1999). Pantomime and theory of mind. *Journal of Genetic Psychology, 160,* 31-45.

Talairach, J, & Tournoux, P. (1988). *Co-planar stereotaxic atlas of the human brain.* New York: Thieme Medical Publishers.

Tallerman, M. (2015). *Understanding syntax*, 4th ed. London: Routledge.

Tam, F., Churchill, N. W., Strother, S. C., & Graham, S. J. (2010). A new tablet for writing and drawing during functional MRI. *Human Brain Mapping, 32,* 240-248.

Thaler, L., & Goodale, M. A. (2011). Neural substrates of visual spatial coding and visual feedback control for hand movements in allocentric and target-directed tasks. *Frontiers in Human Neuroscience, 5,* 92.

Tomlin, R. S. (1986). *Basic word order: Functional principles.* London: Croom Helm.

Vandenberghe, R., Price, C., Wise, R., Josephs, O., & Frackowiak, R. S. (1996). Functional anatomy of a common semantic system for words and pictures. *Nature, 383,* 254-256.

Vingerhoets, G., & Clauwaert, A. (2015). Functional connectivity associated with hand shape generation: Imitating novel hand postures and pantomiming tool grips challenge different nodes of a shared neural network. *Human Brain Mapping, 36,* 3426-3440.

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

Wu, Y. C., & Coulson, S. (2005). Meaningful gestures: Electrophysiological indices of iconic gesture comprehension. *Psychophysiology, 42,* 654-667.

Xu, J., Gannon, P. J., Emmorey, K., Smith, J. F., & Braun, A. R. (2009). Symbolic gestures and spoken language are processed by a common neural system. *Proceedings of the National Academy of Sciences, 106,* 20664-20669.

Yuan, Y., & Brown S. (2014). The neural basis of mark making: A functional MRI study of drawing. *PLoS ONE, 9,* e108628.

Yuan, Y., & Brown S. (2015). Drawing and writing: An ALE meta-analysis of sensorimotor activations. *Brain and Cognition, 98,* 15-26.

Zaccarella, E., & Friederici, A. D. (In press). The neurobiological nature of syntactic hierarchies. *Neuroscience and Biobehavioral Reviews.*

Zacks, J. M., Speer, N. K., Swallow, K. M., & Maley, C. J. (2010). The brain's cutting-room floor: Segmentation of narrative cinema. *Frontiers in Human Neuroscience, 4,* 168.

Chapter 5 – General Discussion

We conducted a series of human brain imaging experiments in order to fill a void in the existing neuroscience literature. First, Chapter 2 outlined the basic drawing network in the human brain and highlighted several areas that process the dynamic accumulation of visual motion during drawing, or "emanation". Next, Chapter 3 described two white matter tracts that connected three of these emanation-related areas. One of these tracts recently resurfaced in the neuroscience literature (the vertical occipital fasciculus, VOF), but the other had no name or existing documentation. Finally, Chapter 4 presented evidence to suggest that the process of narrative production seems to be intrinsically mentalistic and character-based, and that this mentalistic orientation may occur implicitly, regardless of expressive modality.

Most of the neuroscience studies of drawing thus far have used tasks that are very removed from natural drawing, including drawing in the air or on the body (Ino, Asada, Ito, Kimura, & Fukuyama, 2003; Makuuchi, Kaminaga, & Sugishita, 2003), and even covert drawing (Suchan et al., 2002; Harrington, Farias, Davis, & Buonocore, 2007; Harrington, Farias, & Davis, 2009). Since the field has lacked ecological validity, we used a specially-designed drawing tablet in an fMRI experiment that examined many variants of drawing tasks, such as copying, tracing, and blind drawing. The basic drawing network of the human brain includes not only areas that process static visual images and motoric commands for the eyes, arms, and wrists, but also visual-motion areas in the temporo-occipital (V5/MT+), superior occipital (V3A), and posterior parietal (posterior IPS) cortices (Smith et al., 1998) that process the dynamic emergence, or “emanation”, of the visual product that results from the process of drawing. The recruitment of brain areas that process this kind of visual motion and accumulation suggests that, although the product of drawing is static, the process of drawing itself is highly dynamic. The comparison between drawing with visual feedback and blind drawing resulted in activity exclusively in the emanation network. This supports the idea that drawing can be conceptualized as gestural movements that leave a trail behind in the form of the emerging image. It also complements contemporary interest in the gestural origins of language through processes such as pantomime (Armstrong, Stokoe, & Wilcox, 1995; Corballis, 2003), as further supported by an overlapping of brain activations, pointing to a shared effector system for pantomime and drawing.

Having characterized a brain network involved in emanation during drawing, we were curious about whether these areas were anatomically connected. However, we quickly learned that there was minimal information about tracts in the occipital lobe of the human brain. Hence, we decided to use diffusion imaging in order to explore the structural connectivity between the three emanation areas that we characterized in Chapter 2. We only found two out of the three possible tracts: one connecting V5/MT+ and V3A, and another connecting V3A and posterior IPS; we did not find connectivity between V5/MT+ and IPS. The former tract seems to be part of the VOF, which only recently regained interest after fading from the research spotlight for more than a century (Meynert, 1872; Gray, 1918, p.843; Takemura et al., 2015; Yeatman et al., 2013, 2014). The latter tract connects two important areas in the dorsal visual stream, but has been remarkably ignored in the literature. Although the directionality of information transfer can be studied directly in non-human primates using tracer injection (e.g., Ungerleider & Desimone, 1986), this method cannot be employed to study human brains. Since diffusion imaging does not provide data on directionality, our experiment was limited in its ability to offer information on the flow of visual information along these pathways.

While our first study (Chapter 2) only looked at basic geometric drawing, we expanded our work to examine complex narrative drawing (Chapter 4). In our second fMRI experiment, participants were presented with narrative headlines (e.g., “Surgeons finds scissors inside of patient”) and had to draw them. The drawing network observed in this task was remarkably similar to the one found in the first experiment, suggesting that the same basic drawing network is recruited in both simple and complex drawing tasks with almost no differences.

The new fMRI study had two components not present in the first study. First, it was a tri-modal narrative production study, aiming to compare for the first time production across drawing, speech, and pantomime. Second, we employed a control task (an object description task) that would eliminate the sensorimotor components of the tasks so as to reveal brain areas associated with the protagonist of the narrative, rather than sensory or motor components of production.

To our surprise, we found the identical set of areas in all three Narrative vs. Object subtractions, implicating these as multimodal areas involved in protagonist processing during narrative production (and perception). In particular, we found areas involved in mentalizing (TPJ, precuneus), social cognition (pSTS), and semantic processing (pMTG). The engagement of such areas suggests that narrative processing may be intrinsically mentalistic and character-based, rather than simply a description of a sequence of events. Additionally, mentalizing seemed to happen implicitly, without our participants being specifically instructed to do so. Finally, these mentalistic brain areas (precuneus, TPJ, and pSTS) were recruited regardless of expressive modality. Semantic processing in the mid-STS/pMTG was also shared among the three modalities of speech, pantomime, and drawing. Furthermore, we saw the anterior temporal cortices activated bilaterally for both speech and pantomime, but not in drawing. These areas are typically involved in higher-order discourse processing (Binder et al., 2011; Xu et al., 2005). The absence of these areas in drawing suggests that the drawing modality may be the least discourse-like in nature.

It is important to keep in mind that the narrative production tasks contained nothing explicitly mentalistic in their instructions. Since all of the narratives were of object-directed actions, one might more closely associate them with the mirror-neuron system and its function in action observation, especially for object-directed actions (Rizzolatti & Craighero, 2004; Iacaboni et al., 2007). However, our results can be interpreted to mean that, rather than focusing on the actions themselves, participants are more concerned with the underlying goals and emotions of the protagonist. Therefore, one interpretation of these findings is that narrative processing is implicitly mentalistic and perhaps closely related to the concept of agency, in other words a person’s intentionality and sense of control over the ability to achieve desired outcomes (Haggard, 2017), as is important in story processing.

In terms of the overall processing of narratives, the results in this dissertation have shown that narrative generation transcends modality. Cross-modal brain areas are recruited during the generation of an implicitly mentalistic, character-based narrative. Subsequently, information is passed on to motor areas for planning and execution, recruiting modality-specific production areas that were seen in our low-level subtractions. Finally, sensory-feedback processing recruits modality-specific perceptual areas, such as the posterior superior temporal gyrus (pSTG) during speech, and V5/MT+ and V3A for visual motion during both pantomime and drawing.

5.1 Future directions

In Chapter 4, the emanation network was shown to be activated by both pantomime and drawing, both of which contain elements of visual motion. However, is the *accumulation* of visual motion (as in drawing) necessary to stimulate this network, or does this network respond to any visual motion (as in pantomiming), including but not limited to emanation? In order to help answer these questions, future research would require an experimental configuration where participants could see their gestures as well as their drawings, as is now possible in certain MRI scanners that tilt the scanner bed slightly upward to permit participants to perceive their own gestures. A possible prediction is that this network might respond in a piecewise manner, where visual motion itself may activate only a portion of the network (as seen in Figure 4.1B) and the accumulation of visual motion may activate another portion (as seen in Figure 4.1C). Alternatively, the emanation network activity may be modulated parametrically, initially activated by gestural motion alone but then intensified by emanation (i.e., accumulation of images). Future research on the emanation network can provide a better understanding of the functional characteristics of these brain areas.

Next, the interaction between mentalizing areas and semantic-processing areas is currently not well understood. The results in Chapter 4 showed recruitment of the semantic network comprised of the temporo-parietal junction (TPJ), posterior middle temporal gyrus (pMTG), and anterior STS (aSTS). But how does this network share information with the precuneus and other mentalizing areas? One possibility is that the TPJ may serve as some intermediary connection between these two functional networks, since this area is implicated in both cognitive functions and both networks (Moseley & Pulvermüller, 2014; Saxe & Kanwisher, 2003). There has also been some functional connectivity work on the parcellation of the TPJ into a posterior subregion that is more closely associated with classic mentalizing areas, and an anterior subregion that shows greater functional connectivity with traditional semantic areas in the anterior temporal lobe (Mars et al., 2011). Our understanding of the interaction between these networks would greatly benefit from future studies that explore structural connectivity between classic mentalizing areas and the semantic network using techniques such as high angular resolution diffusion imaging (HARDI).

In our narrative study, we did not explicitly control the mentalistic aspect of narratives. However, future studies could introduce this manipulation in order to see what pattern of activations would result from an explicitly mentalistic narrative expression compared to a purely event-based one. One prediction is that the expression of event-based narratives may only elicit activity in the areas that process event or gesture sequences, such as the mirror-neuron system. On the other hand, there may still be some aspect of mentalizing that persists even during event-based narratives and that may be enhanced when events are less mundane or predictable. For example, consider two event sequences: “She grabbed her keys and walked to the car” and “She grabbed her keys and scratched the car”. The first narrative would naturally give rise to questions regarding the person’s destination, whereas the second one arouses concerns of her motivation, conflict, and emotions. These impulses may manifest quite strongly and could be difficult to suppress. Accordingly, one might predict that the second narrative should elicit activity in both the semantic and mentalizing areas, whereas the first narrative might not.

5.2 Clinical implications

Since narrative communication is so integral to our daily lives, dysfunction in any part of the networks that subserve it could have severe repercussions. For example, how would someone’s visual-motion processing be affected by damage or disruption to areas within the emanation network? From earlier TMS studies, it is known that disruption of V5/MT+ function results in a range of deficits, from a perceptual slowing of visual motion (McKeefry et al., 2008) to complete visual motion blindness (Beckers & Homberg, 1992). A recent double-dissociation study (Cai et al., 2014) found that impaired

V5/MT+ activity affected the perception of global but not local coherent visual motion, whereas impaired V3A activity impacted local but not global visual-motion perception. An impaired IPS affected perception of both types of motion. Clinical research on patients with akinetopsia (motion blindness) could examine how their brains would process our narrative communication task, given the importance of visual motion in the pantomime and drawing modalities.

The work in this dissertation has also highlighted the importance of mentalistic processes during the communication of narratives. It is well known that autistic individuals lack the ability to infer the beliefs and thoughts of others (Baron-Cohen, Leslie, & Frith, 1985). In other words, autistic individuals do not seem to mentalize. If narratives are intrinsically mentalistic, how might the brain activity of high-functioning autistic individuals reflect this cognitive deficiency? Moreover, would narratives produced by autistic individuals still show this inherent mentalistic feature seen in the healthy individuals in our own study? Perceptual work has demonstrated the recruitment of the TPJ and medial prefrontal cortex during the reading of cartoons and narratives by healthy individuals (Gallagher et al., 2000), but also that the TPJ activity of autistic individuals does not distinguish between intentional, emotional, or physical inference by remaining hyperactive during all three conditions (Mason et al., 2008). Autistic children show a deficit in processing intentional causality, or agency, and have decreased functional connectivity between TPJ and motor areas for agency, as well as reduced fractional anisotropy in the temporal lobe (Kana et al., 2014). Taken together, we might predict that the TPJ would be hyperactive in autistic individuals compared to healthy controls and that there would be either a smaller amount of coherent connections between the mentalizing and semantic networks, or less coherence altogether in the supporting fiber bundles. Additional clinical research would allow us to pursue the practical implications of these questions.

5.3 Concluding remarks

Communication is one of the most important aspects of human socialization, and usually takes the form of narratives. People communicate these narratives using speech, pantomime, drawing, or some combination of the three. Although speech and pantomime have received much attention over the last several decades, the body of drawing research pales in comparison, and there is virtually nothing on the multimodal production of narrative communication. This dissertation has provided a series of functional and structural brain imaging experiments to improve our understanding of the human brain in terms of drawing and cross-modal narrative communication. We were able to describe the basic drawing network in the human brain as well as isolate brain areas that processed the emanation of drawing. We found that these emanation areas were connected structurally by fasciculi such as the VOF and a previously unknown fasciculus connecting the superior occipital and posterior parietal cortices. Narrative drawing activated the same basic drawing network as that involved in drawing geometric forms. Finally, we compared narrative drawing to narrative production using speech and pantomime as well. While clear sensorimotor differences were seen, the use of an object-description condition allowed us to isolate the narrative aspect of the task from the sensorimotor components and thereby reveal a set of shared narrative areas that transcended the modality of production. These were areas involved in mentalizing, social cognition, and semantics. These results provided evidence that the communication of narratives across the three modalities of expression is inherently mentalistic and focused more on the character than on a sequence of events.

References

Armstrong, D. F., Stokoe, W. C., & Wilcox, S. E. (1995). *Gesture and the nature of language*. New York: Cambridge University Press.

Baron-Cohen, S., Leslie, A. M., & Frith, U. (1985). Does the autistic child have a “theory of mind”?. *Cognition*, *21*(1), 37-46.

Beckers, G., & Homberg, V. (1992). Cerebral visual motion blindness: transitory akinetopsia induced by transcranial magnetic stimulation of human area V5. *Proceedings of the Royal Society of London B: Biological Sciences*, *249*(1325), 173-178.

Benson, R. R., Whalen, D. H., Richardson, M., Swainson, B., Clark, V. P., Lai, S., & Liberman, A. M. (2001). Parametrically dissociating speech and nonspeech perception in the brain using fMRI. *Brain and Language*, *78*(3), 364-396.

Binder, J. R., Frost, J. A., Hammeke, T. A., Cox, R. W., Rao, S. M., & Prieto, T. (1997). Human brain language areas identified by functional magnetic resonance imaging. *Journal of Neuroscience, 17*(1), 353-362.

Binder, J. R., Gross, W. L., Allendorfer, J. B., Bonilha, L., Chapin, J., Edwards, J. C., ... & Koenig, K. (2011). Mapping anterior temporal lobe language areas with fMRI: A multicenter normative study. *Neuroimage*, *54*(2), 1465-1475.

Bohland, J. W., & Guenther, F. H. (2006). An fMRI investigation of syllable sequence production. *Neuroimage*, *32*(2), 821-841.

Bona, S., Herbert, A., Toneatto, C., Silvanto, J., & Cattaneo, Z. (2014). The causal role of the lateral occipital complex in visual mirror symmetry detection and grouping: an fMRI-guided TMS study. *Cortex*, *51*, 46-55.

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

Broca, P. (1861). Remarks on the seat of the faculty of articulated language, following an observation of aphemia (loss of speech). *Bulletin de la Société Anatomique*, *6*, 330-357.

Brown, J. W. (1988). *Agnosia and apraxia: Selected papers of Liepmann.* Hillsdale, NJ: Lange, and Potzl, Lawrence Erlbaum Associates.

Buckner, R. L., Raichle, M. E., & Petersen, S. E. (1995). Dissociation of human prefrontal cortical areas across different speech production tasks and gender groups. *Journal of Neurophysiology*, *74*(5), 2163-2173.

Cai, P., Chen, N., Zhou, T., Thompson, B., & Fang, F. (2014). Global versus local: double dissociation between MT+ and V3A in motion processing revealed using continuous theta burst transcranial magnetic stimulation. *Experimental Brain Research*, *232*(12), 4035-4041.

Chao, L. L., & Martin, A. (2000). Representation of manipulable man-made objects in the dorsal stream. *Neuroimage*, *12*(4), 478-484.

Corballis, M. C. (2003). *From hand to mount: The origins of language*. Princeton, NJ: Princeton University Press.

Crinion, J. T., Lambon-Ralph, M. A., Warburton, E. A., Howard, D., & Wise, R. J. (2003). Temporal lobe regions engaged during normal speech comprehension. *Brain*, *126*(5), 1193-1201.

Culham, J. C., Danckert, S. L., De Souza, J. F., Gati, J. S., Menon, R. S., & Goodale, M. A. (2003). Visually guided grasping produces fMRI activation in dorsal but not ventral stream brain areas. *Experimental Brain Research*, *153*(2), 180-189.

Decety, J., Grezes, J., Costes, N., Perani, D., Jeannerod, M., Procyk, E., ... & Fazio, F. (1997). Brain activity during observation of actions. Influence of action content and subject's strategy. *Brain*, *120*(10), 1763-1777.

DeWitt, I., & Rauschecker, J. P. (2012). Phoneme and word recognition in the auditory ventral stream. *Proceedings of the National Academy of Sciences*, *109*(8), E505-E514.

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

Friederici, A. D., Meyer, M., & von Cramon, D. Y. (2000). Auditory language comprehension: An event-related fMRI study on the processing of syntactic and lexical information. *Brain and Language, 74*(2), 289-300.

Gallagher, H. L., Happé, F., Brunswick, N., Fletcher, P. C., Frith, U., & Frith, C. D. (2000). Reading the mind in cartoons and stories: An fMRI study of 'theory of mind' in verbal and nonverbal tasks. *Neuropsychologia, 38*(1), 11-21.

Geranmayeh, F., Brownsett, S. L., Leech, R., Beckmann, C. F., Woodhead, Z., & Wise, R. J. (2012). The contribution of the inferior parietal cortex to spoken language production. *Brain and Language*, *121*(1), 47-57.

Goldenberg, G., Hermsdörfer, J., Glindemann, R., Rorden, C., & Karnath, H. O. (2007). Pantomime of tool use depends on integrity of left inferior frontal cortex. *Cerebral Cortex*, *17*(12), 2769-2776.

Gowen, E., & Miall, R. C. (2007). Differentiation between external and internal cuing: An fMRI study comparing tracing with drawing. *Neuroimage*, *36*(2), 396-410.

Gray, H. (1918). *Anatomy of the human body* (W. H. Lewis, ed.). Philadelphia, PA: Lea & Febiger.

Grill-Spector, K., Kushnir, T., Edelman, S., Avidan, G., Itzchak, Y., & Malach, R. (1999). Differential processing of objects under various viewing conditions in the human lateral occipital complex. *Neuron*, *24*(1), 187-203.

Grossman, E., Donnelly, M., Price, R., Pickens, D., Morgan, V., Neighbor, G., & Blake, R. (2000). Brain areas involved in perception of biological motion. *Journal of Cognitive Neuroscience*, *12*(5), 711-720.

Guggenmos, M., Thoma, V., Haynes, J. D., Richardson-Klavehn, A., Cichy, R. M., & Sterzer, P. (2015). Spatial attention enhances object coding in local and distributed representations of the lateral occipital complex. *Neuroimage*, *116*, 149-157.

Harrington, G. S., Farias, D., & Davis, C. H. (2009). The neural basis for simulated drawing and the semantic implications. *Cortex*, *45*(3), 386-393.

Harrington, G. S., Farias, D., Davis, C. H., & Buonocore, M. H. (2007). Comparison of the neural basis for imagined writing and drawing. *Human Brain Mapping*, *28*(5), 450-459.

Haxby, J. V., Gobbini, M. I., Furey, M. L., Ishai, A., Schouten, J. L., & Pietrini, P. (2001). Distributed and overlapping representations of faces and objects in ventral temporal cortex. *Science*, *293*(5539), 2425-2430.

Hermsdörfer, J., Terlinden, G., Mühlau, M., Goldenberg, G., & Wohlschläger, A. M. (2007). Neural representations of pantomimed and actual tool use: Evidence from an event-related fMRI study. *Neuroimage*, *36*, T109-T118.

Humphries, C., Willard, K., Buchsbaum, B., & Hickok, G. (2001). Role of anterior temporal cortex in auditory sentence comprehension: An fMRI study. *Neuroreport, 12*(8), 1749-1752.

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

Ino, T., Asada, T., Ito, J., Kimura, T., & Fukuyama, H. (2003). Parieto-frontal networks for clock drawing revealed with fMRI. *Neuroscience Research*, *45*(1), 71-77.

Kana, R. K., Libero, L. E., Hu, C. P., Deshpande, H. D., & Colburn, J. S. (2012). Functional brain networks and white matter underlying theory-of-mind in autism. *Social Cognitive and Affective Neuroscience*, *9*(1), 98-105.

Landau, B., & Jackendoff, R. (1993). Whence and whither in spatial language and spatial cognition?. *Behavioral and Brain Sciences*, *16*(02), 255-265.

Makuuchi, M., Kaminaga, T., & Sugishita, M. (2003). Both parietal lobes are involved in drawing: A functional MRI study and implications for constructional apraxia. *Cognitive Brain Research*, *16*(3), 338-347.

Mars, R. B., Sallet, J., Schüffelgen, U., Jbabdi, S., Toni, I., & Rushworth, M. F. (2011). Connectivity-based subdivisions of the human right “temporoparietal junction area”: Evidence for different areas participating in different cortical networks. *Cerebral Cortex*, *22*(8), 1894-1903.

Mason, R. A., Williams, D. L., Kana, R. K., Minshew, N., & Just, M. A. (2008). Theory of mind disruption and recruitment of the right hemisphere during narrative comprehension in autism. *Neuropsychologia*, *46*(1), 269-280.

Mazoyer, B. M., Tzourio, N., Frak, V., Syrota, A., Murayama, N., Levrier, O., ... & Mehler, J. (1993). The cortical representation of speech. *Journal of Cognitive Neuroscience*, *5*(4), 467-479.

McCarthy, G., Blamire, A. M., Rothman, D. L., Gruetter, R., & Shulman, R. G. (1993). Echo-planar magnetic resonance imaging studies of frontal cortex activation during word generation in humans. *Proceedings of the National Academy of Sciences*, *90*(11), 4952-4956.

McKeefry, D. J., Burton, M. P., Vakrou, C., Barrett, B. T., & Morland, A. B. (2008). Induced deficits in speed perception by transcranial magnetic stimulation of human cortical areas V5/MT+ and V3A. *Journal of Neuroscience*, *28*(27), 6848-6857.

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

Meynert, T. (1872). Vom Gehirne der Säugethiere. *Handbuch der Lehre von den Geweben des Menschen und der Thiere*, *2*, 694-808.

Mishkin, M., Ungerleider, L. G., & Macko, K. A. (1983). Object vision and spatial vision: Two cortical pathways. *Trends in Neurosciences*, *6*, 414-417.

Moseley, R. L., & Pulvermüller, F. (2014). Nouns, verbs, objects, actions, and abstractions: local fMRI activity indexes semantics, not lexical categories. *Brain and Language*, *132*, 28-42.

Ogawa, K., & Inui, T. (2009). The role of the posterior parietal cortex in drawing by copying. *Neuropsychologia*, *47*(4), 1013-1022.

Okada, K., & Hickok, G. (2006). Left posterior auditory-related cortices participate both in speech perception and speech production: Neural overlap revealed by fMRI. *Brain and Language*, *98*(1), 112-117.

Riecker, A., Mathiak, K., Wildgruber, D., Erb, M., Hertrich, I., Grodd, W., & Ackermann, H. (2005). fMRI reveals two distinct cerebral networks subserving speech motor control. *Neurology*, *64*(4), 700-706.

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

Sack, A. T. (2009). Parietal cortex and spatial cognition. *Behavioural Brain Research*, *202*(2), 153-161.

Saxe, R., & Kanwisher, N. (2003). People thinking about thinking people: the role of the temporo-parietal junction in “theory of mind”. *Neuroimage*, *19*(4), 1835-1842.

Scott, S. K., & Johnsrude, I. S. (2003). The neuroanatomical and functional organization of speech perception. *Trends in Neurosciences, 26*(2).

Shmuelof, L., & Zohary, E. (2005). Dissociation between ventral and dorsal fMRI activation during object and action recognition. *Neuron*, *47*(3), 457-470.

Smith, A. T., Greenlee, M. W., Singh, K. D., Kraemer, F. M., & Hennig, J. (1998). The processing of first-and second-order motion in human visual cortex assessed by functional magnetic resonance imaging (fMRI). *Journal of Neuroscience*, *18*(10), 3816-3830.

Suchan, B., Yágüez, L., Wunderlich, G., Canavan, A. G., Herzog, H., Tellmann, L., ... & Seitz, R. J. (2002). Neural correlates of visuospatial imagery. *Behavioural Brain Research*, *131*(1), 163-168.

Takemura, H., Rokem, A., Winawer, J., Yeatman, J. D., Wandell, B. A., & Pestilli, F. (2016). A major human white matter pathway between dorsal and ventral visual cortex. *Cerebral Cortex*, *26*(5), 2205-2214.

Ungerleider, L. G., & Desimone, R. (1986). Cortical connections of visual area MT in the macaque. *Journal of Comparative Neurology*, *248*(2), 190-222.

Vaina, L. M., Solomon, J., Chowdhury, S., Sinha, P., & Belliveau, J. W. (2001). Functional neuroanatomy of biological motion perception in humans. *Proceedings of the National Academy of Sciences*, *98*(20), 11656-11661.

Valyear, K. F., Culham, J. C., Sharif, N., Westwood, D., & Goodale, M. A. (2006). A double dissociation between sensitivity to changes in object identity and object orientation in the ventral and dorsal visual streams: a human fMRI study. *Neuropsychologia*, *44*(2), 218-228.

Wadsworth, H. M., & Kana, R. K. (2011). Brain mechanisms of perceiving tools and imagining tool use acts: A functional MRI study. *Neuropsychologia*, *49*(7), 1863-1869.

Wernicke, C. (1874). *Der aphasische Symptomencomplex: eine psychologische Studie auf anatomischer Basis.* Breslau: Max Cohn & Weigert.

Wilson, S. M., Saygin, A. P., Sereno, M. I., & Iacoboni, M. (2004). Listening to speech activates motor areas involved in speech production. *Nature Neuroscience*, *7*(7), 701-702.

Wise, R. J., Scott, S. K., Blank, S. C., Mummery, C. J., Murphy, K., & Warburton, E. A. (2001). Separate neural subsystems within Wernicke's area. *Brain*, *124*(1), 83-95.

Wu, Y. C., & Coulson, S. (2005). Meaningful gestures: Electrophysiological indices of iconic gesture comprehension. *Psychophysiology*, *42*(6), 654-667.

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

Xu, J., Kemeny, S., Park, G., Frattali, C., & Braun, A. (2005). Language in context: Emergent features of word, sentence, and narrative comprehension. *Neuroimage*, *25*(3), 1002-1015.

Yeatman, J. D., Rauschecker, A. M., & Wandell, B. A. (2013). Anatomy of the visual word form area: Adjacent cortical circuits and long-range white matter connections. *Brain and Language*, *125*(2), 146-155.

Yeatman, J. D., Weiner, K. S., Pestilli, F., Rokem, A., Mezer, A., & Wandell, B. A. (2014). The vertical occipital fasciculus: A century of controversy resolved by in vivo measurements. *Proceedings of the National Academy of Sciences*, *111*(48), E5214-E5223.

Zachariou, V., Klatzky, R., & Behrmann, M. (2014). Ventral and dorsal visual stream contributions to the perception of object shape and object location. *Journal of Cognitive Neuroscience*, *26*(1), 189-209.