Ventral occipito-temporal cortex function and anatomical connectivity in reading

Tae Twomey

2013

Cognitive, Perceptual and Brain Sciences

University College London

A thesis submitted for the degree of Doctor of Philosophy

DECLARATION

I, Tae Twomey, confirm that the work presented in this thesis is my own. Where information has been derived from other sources, I confirm that this has been indicated in the thesis.

ABSTRACT

Previous functional neuroimaging studies of reading in skilled readers, acquired dyslexia and developmental dyslexia have all shown that the left ventral occipito-temporal cortex (vOT) is involved in visual word recognition. Specifically, a region in the left posterior occipito-temporal sulcus lateral to fusiform gyrus and medial to inferior temporal gyrus has been reported to play an important role. However, the precise functional contribution of this area in reading is yet to be fully explored. In this thesis, I empirically evaluated a claim that vOT responds not only to bottom-up processing demands of the visual stimuli but is also influenced by automatic, top-down non-visual processing demands, as proposed by the Interactive Account of vOT functioning.

The first part of this thesis investigated the functional properties of vOT during reading, using functional magnetic resonance imaging. In the first project, the top-down influences on vOT were investigated, teasing apart visual and non-visual properties of written stimuli. In the second project, using the Japanese orthography I disentangled a word’s lexical frequency from the frequency of its visual form – an important distinction for understanding the neural information processing in regions engaged by reading and further explored the interactive nature of the vOT responses. The second part then investigated the anatomical basis of these functional interactions between vOT and other cortical regions. I used diffusion-weighted magnetic resonance imaging and tractography, the only method currently available to identify and measure white matter fibre pathways non-invasively and in vivo.

My research has demonstrated that vOT integrates bottom-up visual information and top-down predictions from regions encoding non-visual attributes of the stimulus in an interactive fashion. It also illustrated the putative anatomical basis for functional connectivity during reading, which is consistent with the parallel cortical visual pathways seen in other primates. Altogether, the results provide strong support for the Interactive Account.

ACKNOWLEDGEMENTS

I would like to thank all my friends, family and colleagues, who helped me through my PhD. Without you, it would have been a lot harder, if not impossible, for me to get here.

I would also like to thank Joe for everything. He is truly the best PhD supervisor anyone could wish for.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

1. INTRODUCTION

1.1. THE SEARCH FOR A VISUAL WORD CENTRE

Reading is everywhere. We use our reading skills not only for novels and stories but also for gaining an enormous amount of information in everyday life. Yet, the mechanisms that enable this uniquely human skill are still not fully understood. Historically, the search for the answer began with 19th century neurologist seeking to identify a brain region responsible for visual word forms. A century later, the focus of the research had shifted to behavioural accounts of visual word recognition by cognitive psychologists and by behavioural neuropsychologists who focused on the types of processing (routes to reading) rather than a single brain region. More recently, the advances in neuroimaging enabled neuroscientists to look for the visual word form centre once again but this time, in vivo.

One of the earliest investigations into the neural substrates of reading was conducted by Dejerine (Dejerine, 1891). He saw a patient who suddenly became unable to read or write following a stroke in the left hemisphere. The post mortem confirmed a lesion in the inferior part of the left angular gyrus extending to the occipital horn of the lateral ventricle and optic radiations. Dejerine deduced that the left angular gyrus was the visual word centre and damage to it resulted in alexia (inability to read) with agraphia (inability to write) since the patient had no access to the visual word forms to associate with incoming written words or inner speech. The following year, he started observing another patient who had lost the ability to read although his writing was spared (pure alexia, or alexia without agraphia). Shortly before his death, the patient also lost the ability to write (Dejerine, 1892). This enabled Dejerine to dissociate the brain regions responsible for the two conditions, pure alexia versus alexia with agraphia, on the basis of the ages of infarcts at the post mortem. The older lesions were localised in lingual and fusiform gyri, cuneus, the tip of the occipital lobe and corpus callosum while the fresher infarct was found on the left parietal lobe and left angular gyrus. To Dejerine, this confirmed his earlier hypothesis that the angular gyrus was the visual word centre and the damage to this area resulted in alexia with agraphia.

He also concluded that the disruption of the connection in the occipitotemporal region carrying the visual information from the occipital lobe to the angular gyrus was the cause of his patient's pure alexia. His writing was spared because the visual word representations stored in the angular gyrus were intact and could still be fed forward to the later stages of the language system. Dejerine's finding was replicated and his interpretation was supported by several subsequent neuropsychological studies (e.g., Damasio and Damasio, 1983; Geschwind, 1962; Geschwind and Fusillo, 1966; Hinshelwood, 1900). According to the classical neurological model of reading (Geschwind, 1965), visual information proceeds from the left and right visual cortices to the left angular gyrus where visual words are stored. The information is then relayed to the rest of the language system in a feed-forward manner (Figure 1.1).

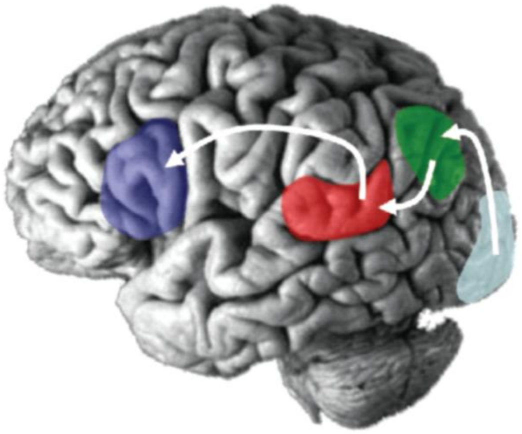

Figure 1.1: The classical neurological model of language. Visual information proceeds from the visual cortex (light blue) to the angular gyrus (green) where visual words are stored. These visual word forms are then relayed to the Wernicke's area (red) to be linked with their auditory word forms. Finally these are linked to motor word forms in Broca's area (dark blue) for articulation. Reprinted from Devlin (2009), with kind permission from Springer Sciences and Business Media.

Later patient studies, however, were not always consistent with Dejerine's account and demonstrated a problem with his model. For instance, cases of pure alexia following left temporo-parietal, instead of occipitotemporal, lesions were reported (Warrington and Shallice, 1980). In addition, more recently it has been shown that lesions in the left occipitotemporal region do not always cause alexia (Hillis et al., 2005). In order to be consistent with these cases, Dejerine's account would require, as a minimum, additional routes to the visual word centre. More problematically however, it has been shown that reading comprehension is not typically impaired following lesions in angular gyrus (Price and Friston, 2002), calling the role of angular gyrus as the visual word centre into question.

Later patient studies on acquired dyslexia led some researchers to suggest multiple processing routes to recognize visual words. Marshall & Newcombe (1973) reported three different types of acquired dyslexia in six patients following a head injury or missile/gunshot wounds in occipital, temporo-parietal or parieto-occipital regions. One patient who had a closed head injury (the location is not known) and another patient with a missile wound in the left occipital region both displayed reading errors that were visually similar to the words presented (e.g., "dug" = "bug"; "met" = "meet"). The authors labelled this type "visual dyslexia" and explained that errors were perceptual rather than linguistic in origin. A further two patients with a missile wounds in their left temporo-parietal regions read written words with grapheme-to-phoneme conversion errors (e.g., /k/ = /s/ as in "insect" = "insist"; /z/ = /s/ as in "phase" = "face"). This type, "surface dyslexia" indicates that the patients are relying on mapping visual input onto its corresponding phonological associations albeit unsuccessfully because they cannot access the semantic knowledge to judge if the output constitute a real word. The final two patients, one with a missile injury in the left temporo-parietal region and one with a gunshot in the left parieto-occipital region, displayed lexico-semantic errors in which the erroneous answers were of the same semantic category as the presented word (e.g., "speak" = "talk"; "diamond" = "necklace"). This type "deep dyslexia" suggests that the patients have access to the semantic content of the visual words without being able to find the phonological associations to them. Based on these observations, Marshall & Newcombe proposed that for successful word perception to occur, the visual input must first be associated with “visual addresses”, a functional component responsible for visual processing, from which it gets associated with both phonological and semantic addresses (components) via separate routes before it is articulated. In normal readers, both routes are available. However, brain injury can cause one of these routes to break down and as a result reading is impaired in a particular fashion. Similar varieties of developmental dyslexia have also been identified which are often explained by the separable semantic (lexical) and phonological (sublexical) routes to reading (Castles and Coltheart, 1993). Therefore, contrary to the classical neurological model of reading, which postulates only single feed-forward route of information processing, it seems that reading requires interactions of multiple routes or processing, each of which can be affected independent of the other, whether it is acquired or developmental.

More recently, neuroimaging studies using positron emission tomography (PET) and functional magnetic resonance imaging (fMRI) have shown that reading typically activates a widely distributed network of regions in occipitotemporal, parietotemporal and inferior frontal regions in addition to the visual cortex (Figure 1.2). The occipitotemporal region in extrastriate cortex encompasses both posterior fusiform gyrus and occipitotemporal sulcus. It is located anterior to V1 and within the ventral, object recognition pathway (Ungerleider and Mishkin, 1982). In the monkey, this visual pathway follows the course of inferior longitudinal fasciculus from striate to inferior temporal regions then reaches ventral frontal cortex via limbic system (Mishkin et al., 1983). The location between V1 and inferior temporal cortex makes this region ideal for visual word recognition and accessing semantics. As such, this region is considered to be the first site of orthographic processing (Cohen et al., 2000). The parietotemporal region, in contrast, is located in the dorsal, object localisation pathway (Ungerleider and Mishkin, 1982), anterior to V1. In the monkey this pathway follows the course of superior longitudinal fasciculus from striate to inferior parietal regions then reaches dorsal frontal cortex via dorsal limbic cortex

(Mishkin et al., 1983). The temporoparietal region is comprised of supramarginal, angular and posterior superior temporal regions. Involved in spatial perception and visuomotor integration, activation here has been seen during writing tasks (Katanoda et al., 2001; Menon and Desmond, 2001; Sugihara et al., 2006) and implicated in mapping orthography to phonology (Shaywitz and Shaywitz, 2008) - a suitable functional role given the close proximity to the auditory cortex. Importantly and interestingly, angular gyrus is not always activated during reading (Price, 2000), inconsistent with this region as a visual word centre (Price et al., 2003) although activation here is correlated with reading skills (Rumsey et al., 1999). Reading activation here is only observed when greater semantic processing is required (Price, 2000; Price and Mechelli, 2005). Supramarginal gyrus has been implicated in computation of phonological representations (Joubert et al., 2004; Moore and Price, 1999; Stoeckel et al., 2009). Posterior superior temporal regions have also been suggested to be involved in phonological processing (Binder and Price, 2001; Démonet et al., 2005; Graves et al., 2008; Hu et al., 2010; Wise et al., 2001). The inferior frontal region is where these two visual pathways converge. It encompasses Broca’s area (pars opercularis and pars triangularis) and the ventral premotor area. Broca’s area is implicated in cognitive processes such as working memory or executive control as well as motor control of speech. Activation here is not only dependent on the cognitive task used (Price, 2000), but also sensitive to extra-linguistic factors such as the novelty of the task (Raichle et al., 1994). Semantic and phonological processing has been attributed to this region too.

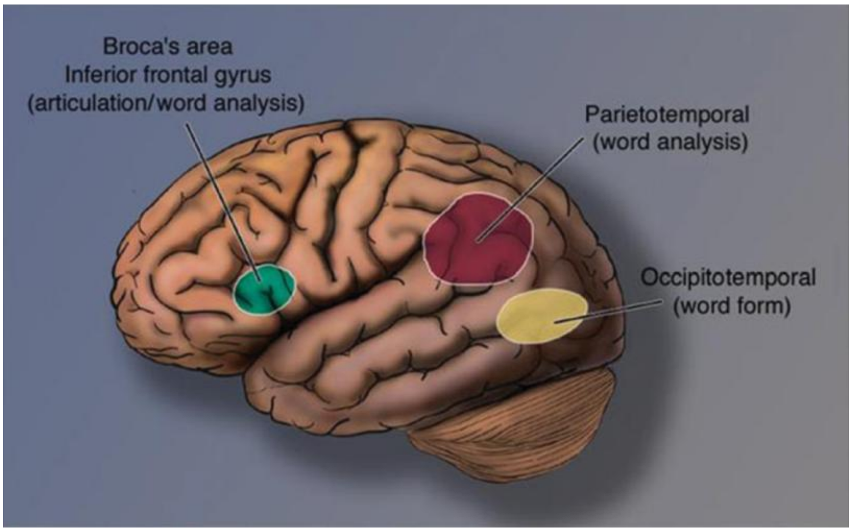

Figure 1.2: Regions activated during reading. Inferior frontal regions are involved in word analysis (semantic and phonological processing) as well as articulation; parietotemporal regions are also involved in word analysis and mapping orthography to phonology; and occipitotemporal regions decodes visual word forms. Reprinted from Shaywitz and Shaywitz (2008), with permission from Cambridge University Press.

Within this network, further functional sub-regions have been identified. For instance, a double dissociation between semantic and phonological processing has been demonstrated within both the inferior frontal regions and inferior parietal lobule. Specifically, greater activation for semantic relative to phonological processing is associated with pars orbitalis and pars triangularis (the anterior and ventral part of the inferior frontal region) and angular gyrus while pars opercularis, premotor cortex (the posterior dorsal part of the inferior frontal region) and supramarginal gyrus are implicated in phonological more than semantic processing (Devlin et al., 2003; McDermott et al., 2003; Mummery et al., 1998; Roskies et al., 2001). In addition, evidence appears to suggest that these semantic and phonological regions correspond to areas identified for lexical (exception words) and prelexical (pseudowords) processing, respectively (Mechelli et al., 2005). Although this seems consistent with the dual routes of processing implicated in the neuropsychological studies mentioned earlier, it is important to emphasise that both of these sets of regions are activated by both types of processing - the significant difference in activation between the two types of processing is relative, not absolute. That is, separate processing does not automatically entail separate anatomical routes. Rather, functional dissociation seems to arise as the result of differential weight put on separate sets of regions within the reading network.

Activations in the reading network are bilateral (Price and Mechelli, 2005) but in general left dominant (Binder et al., 1996; Springer et al., 1999). Compared to alphabetic orthography, a greater involvement of the right hemisphere has been reported for non-alphabetic orthography such as Chinese and Japanese kanji (Bolger et al., 2005; Nakamura et al., 2005c; Tan et al., 2000). In particular, increased activations in bilateral inferior/middle frontal (Tan et al., 2005; Tan et al., 2003) and right occipitotemporal regions (Bolger et al., 2005; Tan et al., 2005; Wu et al., 2012) have been associated with reading logographic orthography. These findings have typically been attributed to the greater demands on visuospatial analyses that are required for visually complex logographic characters (Tan et al., 2001). An alternative interpretation is that these effects reflect the reading strategy that relied more on visual rather than phonological features, which is typical of logographic reading acquisition (Hu et al., 2010), perhaps due to the visual complexity and the lack of transparent grapheme-to-phoneme conversion rules. Interestingly, these “logographic-specific” activation patterns closely match the compensatory system of dyslexic readers of alphabetic orthography, which is activated more relative to the controls (Shaywitz and Shaywitz, 2008), demonstrating that these effects indeed reflect individual differences in reading strategies rather than complex orthographic features specific to logographic characters. In turn, it suggests that the reading network is universal but the learning environment and cognitive abilities could modify reading activation (Hu et al., 2010). It is also important to emphasise that regions in the reading network engage not only in reading but also other tasks both linguistic such as picture naming or speech comprehension and non-linguistic such as object recognition and number processing.

Taken together, early neuropsychological and more recent neuroimaging studies suggest that there is not a single brain region that is exclusively responsible for visual word recognition. Instead, reading seems to arise from a system of multiple cortical regions and the pathways linking them. Therefore, while searching for a cortical region involved in reading such as a visual word centre may have been a good starting point, understanding the reading mechanisms additionally requires investigating information flow within an anatomically distributed system. For instance, can the early neurological model of reading that postulates solely feed-forward flow of information explain all reading behaviour in patients as well as normal readers? What is the anatomical pathway that makes all the necessary communication possible? Answers to these questions are essential if we want to fully explain how reading works.

1.2. VENTRAL OCCIPITOTEMPORAL (vOT) CORTEX IN READING

1.2.1. A NEW "VISUAL WORD CENTRE"?

In recent years, the occipitotemporal region has been attracting attention from researchers investigating visual word recognition. In particular, within the ventral occipitotemporal (vOT) cortex, a small region in the left hemisphere has been consistently found to be activated during tasks that require reading (Herbster et al.; Petersen et al., 1988; Price et al., 1996; Price et al., 1994; Rumsey et al., 1997) but its exact functional role is still under debate. This region is centred on the occipitotemporal sulcus and includes the lateral aspect of the fusiform gyrus and the medial aspect of the inferior temporal gyrus. In group studies, vOT's rostro-caudal location is approximately y= –50 to –60 in MNI (Montreal Neurological Institute) space and is anterior and lateral to V4 (Figure 1.3), although in individuals the peak activation for reading in vOT ranges from y= –45 to –70 (Duncan et al., 2009). Activation in vOT are automatic to visually presented words (Dehaene et al., 2001), reproducible (McCandliss et al., 2003) and found in various tasks that involve visual word recognition such as lexical decision (Fiebach et al., 2002), semantic decision (Chee et al., 2003b), naming (Cohen et al., 2002), silent reading (Kronbichler et al., 2004) and one-back memory task (Hasson et al., 2002). Furthermore, activation here is generally greater than in its right homologue particularly in alphabetic languages (Seghier and Price, 2011) although some studies report that it is bilateral or right-lateralised in Chinese logographs (Bolger et al., 2005; Liu et al., 2007; Tan et al., 2005; Wu et al., 2012) and Japanese Kanji (Kiyosawa et al., 1995; Law et al., 1991; Sakurai et al., 1992). Intracranial recordings (Nobre et al., 1994), magnetoencephalography (MEG) (Cornelissen et al., 2009; Salmelin et al., 1996; Tarkiainen et al., 1999) as well as transcranial magnetic stimulation (TMS) (Duncan et al., 2010) studies have also all provided supporting evidence for reading-related activity in vOT.

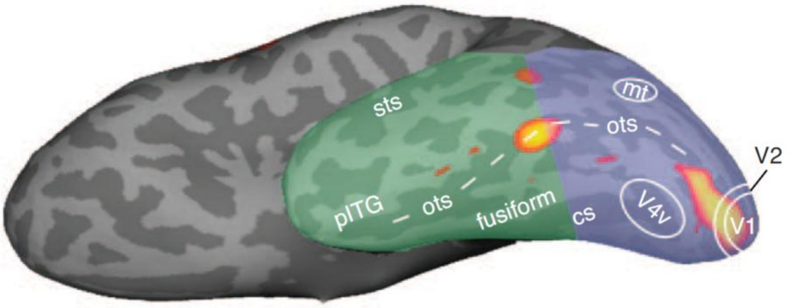

Figure 1.3: Ventral occipitotemporal (vOT) cortex, on an inflated left hemisphere. vOT is centred on the occipitotemporal sulcus (broken white line) and includes the lateral aspect of the fusiform gyrus and the medial aspect of the inferior temporal gyrus. It is located at the transition from the occipital (in blue) to the temporal lobe (in green). Also shown are the regions of activation for visual word recognition (in red-yellow). Abbreviations: cs = collateral sulcus; mt = visual motion area; ots = occipitotemporal sulcus; pITG = posterior inferior temporal gyrus; sts = superior temporal sulcus; V1 = central field of primary visual cortex; V2 = secondary visual cortex; V4v = ventral component of visual area 4. Reprinted from Price and Devlin (2011), with permission from Elsevier, copyright (2011).

More generally, the region encompassing vOT can be functionally segregated into three subdivisions in the posterior-to-anterior direction based on their preferential response to different stimuli or types of processing (Price and Mechelli, 2005). The middle region (y= –50 to –60) responds to words and pseudowords more than false fonts (Herbster et al., 1997) or consonant strings (Cohen et al., 2004; Price et al., 1996). A more posterior region (y= –60 to –70) responds to pseudowords more than words (Mechelli et al., 2003) while a more anterior region (y= –20 to –50) is implicated in semantic processing (Devlin et al., 2006a; Moore and Price, 1999; Mummery et al., 1998; Nakamura et al., 2005b; Spitsyna et al., 2006).

An important property of vOT activation during reading is that it is not specific to a particular language or orthography. In addition to English, activation here has been reported for other alphabetic orthographies such as French (Cohen et al., 2004), German (Kronbichler et al., 2004), Spanish (Carreiras et al., 2007) and Hebrew (Hasson et al., 2002), as well as non-alphabetic orthography such as Chinese (Kuo et al., 2003), Korean (Yoon et al., 2005) and Japanese (Ino et al., 2009; Sakurai et al., 2001). Because the spatial location of vOT corresponds to the lesion site implicated in pure alexia (Beversdorf et al., 1997; Binder and Mohr, 1992; Gaillard et al., 2006; Leff et al., 2001) and activation here is reduced in the developmental dyslexia (Shaywitz and Shaywitz, 2008), vOT is thought to play an important role in orthographic processing (McCandliss et al., 2003; Price and Mechelli, 2005) and features in modern neurological models of reading as the new “visual word form area” (Cohen et al., 2000; Cohen et al., 2002).

1.2.2. READING-SPECIFIC ACCOUNTS OF vOT

Since the importance of vOT in reading was consistently demonstrated, several accounts that assume vOT is dedicated to reading have emerged. Perhaps the most influential account is the Local Combination Detector model by Dehaene and Cohen (Dehaene and Cohen, 2011; Dehaene et al., 2005). It is a model of reading that takes into account the neurophysiology of visual recognition of non-human primates. The ventral visual pathway has a hierarchical organisation and the size of the receptive fields of neurons increases as they ascend the hierarchy, as does the visual complexity of the stimuli that they preferentially respond to. The LCD model proposes a similar hierarchical organisation in which orthographic information is extracted following feed-forward steps that detect progressively more complex visual features. Visual information is encoded through a sequence of stages, from simple feature detectors located in early visual cortex, to letter detectors in V4, to prelexical bigram (letter pairs that comply with the orthotactics of a given language) detectors in vOT and then on to whole word detectors located even more anteriorly in the temporal lobe (Dehaene et al., 2005). Prior to reading acquisition, neurons in vOT are sensitive to visual objects. Through experience and learning to read, these neurons are "recycled" and become attuned to reading (Dehaene and Cohen, 2007).

In order to elucidate the functional role of vOT, Cohen and colleagues (2000) investigated the response to word stimuli presented in either the right (RVF) or left (LVF) visual field in split-brain patients with left hemialexia and healthy controls. Both fMRI and event-related potentials (ERPs) were used to define the spatial and temporal characteristics of visual information processing. For fMRI, the participants silently read real words and for ERPs, these words were contrasted with consonant strings in order to better identify the onset of responses associated with lexicality. In both patients and controls, an early negativity was found around 150 – 160 ms post-stimulus over the posterior electrodes contralateral to the visual field in which the stimuli were presented. Spatially, this activation was identified as the inferior occipitotemporal region corresponding to V4. The authors concluded that letter strings at this stage are still processed in the hemisphere contralateral to the visual hemifield. In contrast, a later negativity of 180 – 200 ms post-stimulus was recorded over the inferior temporal electrodes only in the left hemisphere, irrespective of hemifield presentation. In the patients, the same response was elicited only from the stimulus presented in the RVF. The spatial location of this activation was the left middle portion of fusiform gyrus (vOT) and the latency of this component was consistent with previously found N170 component sensitive to orthographic processing (Nobre et al., 1994; Salmelin et al., 1996; Tarkiainen et al., 1999). Given the latency of this effect, Cohen and colleagues proposed that vOT is the first site where the orthographic information from the two hemifields is combined. Therefore, analogous to Dejerine (1892) designating the angular gyrus to be the visual word centre (i.e., the storage of “optical images of words”), they labelled vOT as the visual word form area (VWFA) (Cohen et al., 2000).

However, as pointed out by Cohen and colleagues (2002), vOT response to orthographic stimuli itself does not necessarily entail that the function of this region is orthographic processing. The response to letter strings may simply be driven by the fact that letters happen to have the type of visual features that the neurons of this region preferentially respond to. In order to determine the functional properties of vOT, Cohen and colleagues (2002) further investigated whether vOT is sensitive to orthographic regularities using words, consonant strings and checkerboards during silent reading/passive viewing and naming tasks. They found greater activation for letter stimuli relative to checkerboards and for words relative to consonant strings. This was interpreted as vOT being sensitive to the letter combinations possible in a given language (orthotactics). Cohen and colleagues thus proposed that vOT is attuned specifically to language. Another difference between words and consonant strings, namely lexicality, was rejected as the explanation since previous findings from their own group as well as others demonstrated equal or greater activation for pseudowords relative to words (Brunswick et al., 1999; Dehaene et al., 2002; Fiez et al., 1999; Xu et al., 2001), which additionally suggested to them that the representation in vOT is essentially prelexical (Dehaene et al., 2002).

Modality specificity of vOT was also investigated. Dehaene and colleagues (2002) found that vOT responded only to written, not spoken, words and pronounceable pseudowords. Moreover, Cohen and colleagues (2004) compared written and spoken words over two tasks, which were designed to focus on phonological (phoneme detection) or orthographic (descender detection) processing, respectively. In addition, in order to test the effect of word (conscious) repetition, the stimulus was preceded by the same stimulus in half of the trials. They found that vOT responded to both novel and repeated written words, regardless of the task. In contrast, for spoken words, novel (unrepeated) stimuli elicited only weak activation and there was no activation to repeated ones. Consequently, they concluded that that response in vOT was mandatory in the visual modality for words and that vOT was a strictly visual (i.e., unimodal) area. According to the authors, responses to auditory stimuli only occurred via top-down processing when the task required manipulation or mental imagery of orthographic representations.

Their next study (Dehaene et al., 2004) led them to identify two sub-regions of vOT, each of which responds to different types of information in visual letter strings. In this priming study investigating location-invariance of vOT response, the target words were primed by unrelated words, the same (repeated) words or circular anagrams of the target words (e.g., range/anger). In addition, each prime – target pair differed in the typographic case so that the pairs were not visually similar (e.g., A/a). This was done to confirm that their previous finding of case-invariance of vOT response (Dehaene et al., 2001) was not due to visual similarities (e.g., O/o) but rather represented true orthographic information. Primes and targets appeared in either the same or different location by one letter position (e.g., #RANGE/anger#). They found repetition suppression in a posterior location (y= –64 to –68) only when the same letter strings appeared in the same location (i.e., the same words in the same location and anagrams in different location (e.g., RANGE#/#anger), suggesting this region may contain location-specific "letter detectors". In contrast, in a more anterior region (y= –56), reduced activation was found for the same words in the same or different locations, suggesting that this region may store a location-invariant representation of visual words. Case-invariance was also replicated for both of these sub-regions. These findings motivated Dehaene and colleagues to propose that neurons in the occipitotemporal region may be hierarchically organised visual feature detectors dedicated to reading that extract progressively more complex, abstract information of written stimuli as one moves anteriorly along fusiform gyrus. Moreover, they made the specific claim that within this posterior-to-anterior gradient, their VWFA (centred on y=-56) contained bigram detectors.

A slightly different reading-specific account of vOT was proposed by Kronbichler and colleagues (2004). They argued that the prelexical explanation of vOT activation by Dehaene and Cohen was inconsistent with neuroimaging results from object recognition studies, which demonstrated a frequency effect in the nearby fusiform regions (e.g., van Turennout et al., 2003) and also visual word recognition studies, which showed lexicality effects (e.g., Mechelli et al., 2003) as well as frequency effects (e.g., Keller et al., 2001; Kuo et al., 2003). These finding were problematic to Dehaene and Cohen's account since lexicality or conceptual frequency should not affect vOT activation if representations here are prelexical unless they differ in terms of bigram frequency. Kronbichler and colleagues therefore investigated vOT responses, parametrically manipulating lexical frequency of German words. Pseudowords were included as the least frequent condition. Various psycholinguistic factors such as letter length, syllable length and crucially, bigram frequency were matched across five word frequency levels to ensure that effects would not be driven by differences in prelexical factors. Using a silent reading task, they found that the vOT response was modulated by, and inversely proportional to, lexical frequency - a property of the whole word. Since bigrams are not words, they should not be sensitive to lexical frequency, especially when bigram frequency was matched across word frequency levels. Therefore, Kronbichler and colleagues suggested that vOT contains orthographic patterns (visual words) that serve as recognition units. Greater vOT response to pseudowords relative to words may be due to the fact that pseudowords partly activate several existing orthographic patterns of real words, with which they share their constituents.

Critically though, lexical frequency is not just the measure of visual familiarity but also includes frequency of semantics as well as phonology of lexical entries. Similarly, the difference between real words and pseudowords is not limited to the visual-orthographic familiarity. Pseudowords lack semantic content and are phonologically novel. Therefore, it is possible that the word frequency effect found by Kronbichler and colleagues (2004) may in fact reflect these differences in these non-visual properties rather than in visual familiarity correlated with lexical frequency. In order to address these issues, Kronbichler and colleagues (2007) conducted another study in which orthographic familiarity and linguistic properties were manipulated. Stimuli were orthographically and semantically/phonologically familiar real words (e.g., Taxi), orthographically unfamiliar but semantically/phonologically familiar pseudohomophones (e.g., Taksi) and orthographically and semantically/phonologically unfamiliar pseudowords (e.g., Tazi). Participants were asked to decide if the letter strings sounded like an existing word (a phonological lexical decision task). Psycholinguistic factors including bigram frequency were matched across conditions. They found greater activation for pseudowords relative to real words in vOT while there was no significant activation difference between pseudohomophones and pseudowords here or anywhere in the left occipitotemporal region. In other words, it was orthographic familiarity that differentiated activations in vOT, not semantic/phonological familiarity. Kronbichler and colleagues thus concluded that vOT is the site of the “orthographic input lexicon” where orthographic word representations, not prelexical bigrams, are stored that abstract away from details of the visual form. Like logogens (Morton, 1969), entries in this orthographic input lexicon are sensitive to experience, with access to less frequent words requiring greater effort and therefore resulting in greater activation, consistent with familiarity effects found for objects in bilateral occipitotemporal regions and for faces in the right fusiform gyrus.

Their account was supported by a study that replicated their findings with English stimuli using a similar design (Bruno et al., 2008) as well as by an fMRI-rapid adaptation study (Glezer et al., 2009). In order to look for evidence that neurons in vOT are selective for the whole words, rather than prelexical representations such as bigrams, Glezer and colleagues (2009) used an fMRI-rapid adaptation technique, suggested to better probe neuronal tuning (Jiang et al., 2006). With this technique, the level of BOLD response is considered to reflect the similarity of the neuronal populations responding to each of the two successively presented stimuli within a trial. Neural response to the second stimulus is suppressed relative to the first stimuli and the suppression is greater when the first and second stimuli are similar relative to when they are dissimilar. Thus, greater response is expected when the two stimuli activate different populations of neurons while two stimuli activating the identical populations of neurons results in weaker response. Manipulating lexicality (real words, pseudowords) and conditions of the prime – target pairs (same, differ by one letter, different), they hypothesised that if neurons in vOT are tuned to prelexical representations, the vOT response would increase gradually as the difference between the prime and the target increases for both real words and pseudowords. In contrast, if they are selective to whole real words, then this pattern would be seen only for pseudowords but critically, the response for real words should not significantly differ between “differ by one letter” and “different” conditions. This was indeed what they found in individual’s vOT identified in separate localisation scans. They also confirmed that the effects were not driven by the inevitable difference between real words and pseudowords, namely the semantic information, by conducting an additional control experiment of a similar design including semantic relatedness as a factor. Crucially, bigram frequency and trigram frequency were matched across conditions in this study. As a result, Glezer and colleagues concluded that neurons in vOT are selective to the whole words, compatible with the mental orthographic lexicon postulated in some cognitive model of reading (Coltheart, 2004), while still endorsing the kind of hierarchical organisation of this region proposed by Dehaene and Cohen.

Like Dehaene and Cohen (Dehaene et al., 2004; Dehaene et al., 2001), Kronbichler and colleagues (Kronbichler et al., 2009) also tested whether case-deviant forms of real words would affect vOT response. Unlike Dehaene and Cohen however, they used mixed-cases. With a silent reading task comparing real German noun forms (e.g., Taxi), case-deviant forms (e.g., TaXi) and letter-deviant forms (e.g., Taksi), they found that case-deviant forms elicited increased activation in vOT relative to real word forms; and also in right occipitotemporal and left posterior occipitotemporal regions relative to both real word and letter-deviant forms. Kronbichler and colleagues interpreted their findings as further evidence that vOT contains word-specific orthographic representations. It is important to note however, that in contrast, previous PET studies using English words showed no effects on vOT responses for real words in mixed- relative to same-cases (Mayall et al., 2001; Xu et al., 2001). Therefore the increased activation could have been driven instead by the status of capital letters in German. In German, nouns must start with a capital letter. Therefore, the capital letter carries the information that a letter string is a noun when it occurs at the beginning, over and above visual and other linguistic information it encodes in other alphabetic languages such as French or English. In other words, case-deviation means much more than the deviation in visual features in German. If this is correct, then the increased activation for case-deviant forms in German is elicited by non-visual information, either via top-down processing or because this information is stored for each lexical entry in the orthographic input lexicon.

One important matter that has not been explicitly discussed by this theory is what the orthographic input lexicon means at the neuronal level. Given this account, it logically follows that individual neurons are “word detectors” that encode each word, analogous to the bigram detectors of the LCD model. Alternatively, it may be that some distributed pattern of neurons collectively encodes a word. These patterns would have to be unique and orthogonal so that each word is represented by a different pattern. Any overlapping would inevitably entail that each neuron encodes less than a word, which would mean, contrary to their claim, that individual neurons in vOT are tuned to prelexical representations as proposed by the LCD model.

Any complete accounts of reading at the neuronal level should be universally applicable. As mentioned earlier, functional neuroimaging studies have shown that non-alphabetic orthographies such as Chinese (e.g., Kuo et al., 2003; Liu et al., 2007; Tan et al., 2000; Xue et al., 2006) and Japanese (e.g., Ino et al., 2009; Nakamura et al., 2005b; Sakurai et al., 2001) also activate vOT. Reading-specific accounts, however, focus solely on alphabetic orthography and provide little or no explanation for how they can be extended to include non-alphabetic scripts. Bigrams, for instance, do not even make sense in logographic (Chinese), morphographic (Japanese kanji) and syllabographic (Japanese kana) scripts since each character often constitute a word or morpheme carrying semantic information. Indeed, Dehaene and colleagues (2005) point out that local detectors should be selected for “any useful recurrent combination of curves within the cell’s receptive field” through perceptual learning (p.338). Therefore, in Chinese neurons might become tuned to semantic or phonological radicals or even whole characters. However, Chinese characters (or Japanese kanji) are visually much more complex than bigrams and carry semantic information. Thus, according to the LCD model, these must be encoded by larger-size units located more anteriorly than the “VWFA”. In this case, it is unclear what the neurons in “VWFA” encode in these orthographies. In short, the LCD model would require orthography-specific modifications to be viable. While this may be conceivable, another challenge is what happens to multi-script orthography such as Japanese in which the same words can be written in two different scripts (kanji and kana)? Dehaene and colleagues (2005) suggest that the size of the units encoded by neurons in the occipitotemporal region can reflect orthographic transparency and go on to state that Japanese kanji (orthographically opaque and visually complex) requires larger-size unit than Japanese kana (orthographically transparent and visually simple). The LCD model specifies that abstract orthographic codes are recognised through sequential stages detecting progressively more complex visual features. The problem is that the visual features of Japanese kanji and Japanese kana are never combined to produce visual codes. That is, the visual features of the two scripts cannot be pooled together to be the input to the next stage in the hierarchy. Therefore, this explanation would require two separate hierarchical organisations for each script, which then must interact since Japanese kanji and kana do get combined to form a word. Moreover, a single kanji or kana can also be a word. In alphabetic orthography, the LCD model proposes that the whole words are encoded at the anterior end of the hierarchy, possibly in the temporal lobe while a single letter is located more posteriorly. It is unclear how this explanation could be applied to Japanese.

In contrast, the orthographic input lexicon account is applicable to multi-script orthography. Although it is not specified, there could be a single lexical entry for each word, irrespective of visual properties of the word. Alternatively, if the postulated lexicon contains word-specific information such as case identity (Kronbichler et al., 2009), then this explanation can be extended to allow separate entries for the same words written in Japanese kanji and Japanese kana.

The question of generalisability demonstrates a fundamental problem with reading-specific accounts: namely, that they seek an explanation within the linguistic domain yet they largely ignore non-alphabetic languages. It seems that a successful account of vOT must either accommodate orthographic features from both alphabetic and non-alphabetic orthography or depart from linguistic explanations altogether.

Another serious challenge for reading-specific accounts is that although vOT is activated during reading, it also responds to non-orthographic stimuli such as pictures of objects (Bookheimer et al., 1995; Moore and Price, 1999; Murtha et al., 1999), animals (Chao et al., 1999), meaningless non-objects (Phillips et al., 2002; van Turennout et al., 2000) and meaningless 2D shapes during colour naming (Price and Friston, 1997). In order to explain this, the LCD model suggests that there are distinct populations of neurons that are specific to word and to objects within vOT (Dehaene et al., 2002; McCandliss et al., 2003) although this cannot be detected with the currently available spatial resolution of fMRI. More recently it has also been shown that although activation in vOT is greater for words and pseudowords more than false fonts (Herbster et al., 1997), vOT is significantly activated for false fonts too when numbers are used as the baseline (Woodhead et al., 2011a). However, it is unlikely that the LCD model would apply the same explanation and postulate a separate set of neurons specific to false fonts.

Moreover, thorough examinations of patients with pure alexia following left occipitotemporal lesions have shown that their deficits are not specific to reading but include non-orthographic tasks such as picture identification (Behrmann et al., 1998), visual apprehension span and object decisions with fragmented pictures (Starrfelt et al., 2009). This indicates that their problem may be of general visual deficits rather than specific to reading. Therefore, a successful account should be able to explain both reading and general visual deficits. Indeed, it has been demonstrated that activation in this region is greater for high relative to low spatial frequencies (Woodhead et al., 2011b) and patients with left posterior fusiform lesions have reduced sensitivity to high spatial frequency (Roberts et al., 2012), suggesting a possible common cause for deficits in reading and non-orthographic processing.

Furthermore, activation in vOT for words can increase or decrease relative to other orthographic stimuli or pictures depending on the task (Price and Devlin, 2011). For instance, a greater vOT response to words relative to consonant strings was found during silent reading (Cohen et al., 2002) and this effect was interpreted that vOT was sensitive to orthographic rules (i.e., letter combinations). However, during a one-back task, less word-like stimuli also elicited greater vOT response than words (Wang et al., 2011b). Wang and colleagues (2011b) used Chinese characters and parametrically manipulated phonological and semantic cues (radicals) as well as their left-right configurations in order to create a gradient of eight conditions varying in "word-likeness". Contrary to the finding from Cohen and colleagues (2002), activation in vOT was inversely related to the word-likeness of the stimuli. The authors thus concluded that the role of vOT is task-dependent, explaining that the one-back task increased short-term memory demands for the less word-like stimuli due to the absence of phonological or semantic information. Another example found that although vOT activation is generally greater for pictures than words (Price and Devlin, 2011), this too depends on the processing demands of the task. In a study in which stimulus type (words, pictures) was crossed with task (colour decision [white/yellow], categorisation [natural/artefacts]), Starrfelt & Gerlach (2007) found that vOT activation was greater for words relative to pictures during colour decision. Interestingly however, this effect diminished during the categorisation task. According to the authors, the demands for shape processing for pictures vary depending on the task as the levels of identification required differs (i.e., greater for semantic categorisation than colour decision) whereas they are likely to remain constant for words. As a result, they suggested that vOT is not specific to words but the response here reflects shape configuration, which is a common process for both words and pictures. Similarly, vOT response to physically different stimuli with identical semantic and phonological information was shown to depend on the nature of the task. Hellyer and colleagues (2011) used numbers written with words and digits (e.g., "twenty" vs. "20") in numerical (odd or even) and phonological (if a particular phoneme was present) tasks. They found that during the phonological (i.e., linguistic) task, vOT activation did not differ between words and digits despite their physical differences whereas the words elicited greater activation relative to digits during the numerical (i.e., non-linguistic) task. There was no effect of stimulus type in the anatomically defined vOT. Activation thus can be modulated not only by the stimulus type but also by the task demands. These data are difficult for reading-specific accounts to explain since their central claim is that vOT activation is driven by the stimulus, not by the nature of the processing.

These findings attest that trying to understand the functional role of a cortical region putatively responsible for visual word recognition based purely on the stimulus type is not fruitful. Instead, all the evidence suggests that the vOT response is the result of interaction of visual stimuli and their non-visual properties, which can be further modulated by task demands. A complete account of vOT function should be able to address this issue.

1.2.3. AN ALTERNATIVE ACCOUNT OF VOT

An alternative neural account, the Interactive Account (Price and Devlin, 2011), offers a very different theory of vOT function within a predictive coding framework (Friston and Kiebel, 2009; Rao and Ballard, 1999). According to the predictive coding perspective, perceptual inferences involve recurrent interactions between bottom-up sensory input via forward connections and top-down predictions from higher level processing regions via backward connections within a hierarchical organisation. Top-down predictions are automatic, based on prior experience and sent to predict the response at lower level regions. The difference between the prediction and response (prediction error) is then sent to higher regions for further inferences until the error is minimised[1]. Since fMRI

---

[1] To minimise prediction error is to minimise free-energy. This is a necessary biological characteristic under the free-energy principle (Friston, 2010; Friston and Kiebel, 2009).

cannot differentiate between signals originating from forward and backward connections, vOT response reflects bottom-up sensory information, top-down predictions and the mismatch (prediction error) between the two. According to the Interactive Account, vOT represents visual form information and is not specialised for written words. The same vOT neurons that represent spatial configurations important to written words also contribute to other visual stimuli such as objects, scenes and faces. Reciprocal connections with higher order association areas link these visual representations with non-visual properties of the stimulus such as its sound or meaning. Thus, in cognitive terms, vOT continuously and automatically interacts with other regions during reading, acting as an interface between low level visual input and higher level non-visual information such as phonology and semantics (Cai et al., 2010; Devlin et al., 2006a; Hillis et al., 2005; Kherif et al., 2011; Nakamura et al., 2002; Price and Friston, 2005; Xue et al., 2006). There are two crucial differences between this account and reading-specific accounts. First, the Interactive Account is not specific to visual word forms. Since it does not stipulate any stimulus-specific representations, this account is consistent with not only vOT activation in response to non-orthographic stimuli, but also the aforementioned findings that patients with pure alexia due to vOT lesions frequently also demonstrate impairments in non-orthographic tasks (Behrmann et al., 1998; Starrfelt et al., 2009). Similarly, it means that this account does not require additional modification in order to generalise to non-alphabetic orthography. Second, only the Interactive Account proposes that the nature of dynamic processing in vOT is of recurrent and automatic interaction of both feed-forward and top-down feedback, rather than purely feed-forward. Therefore, it can also explain task effects (see below).

In order to investigate the specific functional contribution of vOT and test the predictions from competing accounts of vOT, Devlin and colleagues (2006a) conducted a visual masked priming study with fMRI. Specifically, they manipulated lexicality, orthographic similarity and semantic relatedness between the prime and the target to investigate whether vOT was sensitive to prelexical (the LCD model), lexical (the orthographic lexicon account) or non-visual (the interface/Interactive Account) properties. The first experiment tested whether repetitions across typographic cases would result in neural priming (reduced BOLD signal) over words and pseudowords. This was done to differentiate between the LCD model and the other two accounts. They found that repetitions of words (e.g., cabin – CABIN), but not pseudowords (e.g., solst – SOLST), led to reduced BOLD signal in vOT, suggesting that the vOT response was not purely visual but included processing of non-visual properties (i.e., lexicality). This result is consistent with the orthographic input lexicon account (if there is a single lexical entry for each word, irrespective of the surface form of the word) since only words have entries in the orthographic lexicon that could prime. However it is problematic for the LCD model because lexicality should not concern the bigram detecting neurons and thus the same effect would be expected for words and pseudowords. The second experiment investigated the effects of orthographic similarity and semantic relatedness between the prime and the target in order to differentiate between reading-specific accounts and the alternative interface/interactive explanation. They found neural priming for all orthographically related prime-target pairs (e.g., fasten – FAST, corn – CORNER), which was attenuated when the pairs were also semantically related (e.g., deadly – DEAD). These findings suggest that vOT is involved in visual feature detecting but non-visual properties (semantic relatedness) clearly affect its response, which is inconsistent with both of the reading-specific accounts. Since the orthographic lexicon account suggests that different lexical items would be represented by separate entries, a lexical item cannot prime any other items however visually similar they may be. As for the LCD model, non-visual properties such as semantic relatedness should not matter to the bigram detectors since they are purely visual feature detectors. Based on these findings, the authors suggested that vOT acts as an interface, instantiating visual information sent from the visual cortices via bottom-up projections and associating it with its higher-order non-visual information such as its meaning and pronunciation via top-down connections, without stipulating reading-specific representations.

In order to investigate whether vOT activation was specific to words, Kherif and colleagues (2011) tested repetition suppression and varied stimulus type (words, pictures) of primes and targets and relatedness between them (unrelated, conceptually identical, semantically related, phonologically related) while measuring activation in vOT with fMRI. Even within the same stimulus type, physically different stimuli were used for the conceptually identical prime-target pairs (cross-case for words and different exemplars for pictures) so that the effects would not stem from bottom-up visual differences. They found repetition suppression for conceptually identical but not other prime-target pairs in both anterior and posterior vOT, irrespective of stimulus type between the prime and the target and whether the prime was masked or unmasked. That is, they replicated previous repetition priming effects for cross-case (not mixed-case) words in vOT (Dehaene et al., 2004; Devlin et al., 2006a) and additionally demonstrated the same effect in precisely the same region for pictures as well as for cross-stimulus types. These results are not compatible with either of the reading-specific accounts of vOT since the LCD model would predict repetition suppression only for word-word pairs but not for cross-stimulus or picture-picture pairs. Similarly the orthographic input lexicon account would predict priming effects only for word-word pairs if there is a single lexical entry for each word, irrespective of the surface form of the word. Alternatively, it would predict no priming effects for any prime-target pairs if the lexicon contains word-specific information such as case identity (Kronbichler et al., 2009), since word-word pairs were in different typographic cases. Since the prime-target pairs were never physically identical (even within the same stimulus type), the effect could not have been derived from the bottom-up perceptual differences. Anterior vOT is implicated in semantic processing (Devlin et al., 2006a; Moore and Price, 1999; Mummery et al., 1998; Nakamura et al., 2005b; Spitsyna et al., 2006). Therefore, the effect here could be interpreted as arising from the identical semantic contents between the prime and the target. However, given that posterior vOT is involved in bottom-up visual feature processing (Devlin et al., 2006a; Szwed et al., 2009; Xue et al., 2006), the effect in this region cannot be accounted for in the same way. Instead, it is best explained by the Interactive Account, in terms of minimised prediction error. The fact that this effect was found regardless of whether the prime was masked or unmasked suggests that it was not a result of a conscious strategy but automatic. When the prime and the target are conceptually identical, the prime automatically and efficiently predicts the response elicited by the target thereby minimising prediction error.

Another strength of the Interface Account is that it can also explain the data that supported reading-specific accounts of vOT, which primarily used orthographic stimuli. The primary argument for the orthographic lexicon account rests on the lexical frequency effect. For high frequency words, the association between the visual form and its linguistic properties is stronger than it is for less frequent words. As a result, high frequency words yield less prediction error, which leads to less activation. Similarly, the findings by Glezer and colleagues (2009) above can also be explained in terms of top-down predictions. Words, but not pseudowords, engage strong top-down predictions (from semantic and phonological areas) since they have been encountered previously. For real words, these top-down predictions induced by the prime are “wrong” when the target is a different word, regardless of by how many letters it differs from the prime. The deciding factor is whether the prime successfully predicts the target or not. Therefore, prediction error is equally increased between single letter deviations and multi-letter deviations. Consequently, vOT activation does not differ between them. In contrast, for pseudowords the influential factor is the visual similarity between the prime and the target since pseudowords induce only weak top-down predictions. Therefore, activation could be increased proportional to the degree of physical deviations between the prime and the target. As for the data that supported the LCD model, the same explanation applies to greater activation for words relative to consonant strings (Cohen et al., 2002) and case- or location-invariance of vOT activation (Dehaene et al., 2004). With regard to pseudowords, the two reading-specific accounts reported different results: greater activation relative to words during a phonological lexical decision task (Kronbichler et al., 2004) and no activation difference between pseudowords and words during a repetition detection task (Dehaene et al., 2002). While each account is not consistent with the data that support the other account (see above), the Interactive Account can account for both, in terms of task requirements and attentional demands. Specifically, it predicts greater activation for pseudowords than words during a linguistic task such as phonological lexical decision, due to weaker phonological top-down predictions, which leads to greater prediction error. This difference however, can be attenuated during attentionally demanding tasks such as repetition detection since the demands for visual processing is higher for pseudowords, due to the lack of/weaker semantic and phonological top-down predictions. The gradations of vOT activation to different types of written stimuli during general reading tasks (false fonts < consonant strings < high frequency words < low frequency words < pseudowords) can be explained by the Interactive Account by the interaction of visual word-likeness (greater activation for more word-like stimuli), familiarity (greater activation for less familiar stimuli), top-down predictions (greater for more word-like stimuli) and prediction error (greater for less familiar stimuli). During non-linguistic tasks, this pattern might be altered because these tasks require different bottom-up and top-down processing demands.

1.3. TESTING HYPOTHESES

Having examined the existing literature and evaluated the evidence available to date, it is clear that in order to deepen our understanding of the specific role of this region, it is necessary to test the hypothesis that vOT communicates with other regions within the reading network via both bottom-up and top-down connections as suggested by the Interactive Account, for both alphabetic as well as non-alphabetic orthography. In addition, if there is evidence for these functional interactions, it becomes important to systematically investigate their anatomical basis.

Although electroencephalography (EEG) and MEG both offer temporal resolution at millisecond scale, which may enable separating early bottom-up (feed-forward) from later top-down (feedback) processes, if the vOT response is interactive then these two processes may not be distinct in time and therefore may not be separable based solely on their time courses. Moreover, even if it was possible to separate them, the spatial resolution of EEG and MEG does not allow localisation of the signal to a small cortical area such as vOT. Given that the focus of this thesis is on vOT, it would be inappropriate to lose spatial resolution that is necessary to identify the region. Therefore, in order to evaluate both bottom-up and top-down influences on vOT activation, functional magnetic resonance imaging (fMRI) was used. It is a non-invasive method, which offers sufficient spatial resolution to reliably identify signals from vOT. Given the relatively poor temporal resolution of the technique (i.e., seconds rather than milliseconds), it was critical to design experiments and stimuli to control for bottom-up influences on vOT while manipulating top-down processing demands.

Chapter 3 investigated the two types of top-down effects on vOT, namely task and stimuli. fMRI was used to investigate whether there was evidence that activation in vOT was influenced top-down by the interaction of visual and non-visual properties of the stimuli during visual word recognition tasks. Participants performed two different types of lexical decision tasks that focused on either visual or non-visual properties of the word or word-like stimuli. The design allowed the investigation of how activation in vOT was influenced by a task change to the same stimuli and by a stimulus change during the same task. Both stimulus- and task-driven modulation of vOT activation were found that could be best explained by top-down processing of non-visual aspects of the task and stimuli.

Chapter 4 investigated the differential effects of lexical frequency and visual familiarity on vOT using a non-alphabetic orthography, namely Japanese, with fMRI. In Japanese, the same word can be written in either morphographic Kanji or syllabographic Hiragana and this provided a unique opportunity to disentangle a word's lexical frequency from the frequency of its visual form – an important distinction for understanding the neural information processing in regions engaged by reading. Behaviourally, participants responded more quickly to high than low frequency words and to visually familiar relative to less familiar words, independent of script. Critically, the imaging results showed that visual familiarity, as opposed to lexical frequency, had a strong effect on activation in vOT. Activation here was also greater for Kanji than Hiragana words and this was not due to their inherent differences in visual complexity.

Having investigated the functional properties of vOT, Chapter 5 looked for the anatomical basis of functional interactions between vOT and other regions of the reading network including Broca's area and the inferior parietal lobule. The experiment used diffusion-weighted MRI with probabilistic tractography. Seed masks were manually traced within the cortical grey matter for three rostro-caudal locations along the occipitotemporal sulcus for vOT and the tractography algorithm generated the most probable paths. Additional seed masks in Broca's area, inferior parietal lobule and middle occipital gyrus were also employed in further analyses in order to test more specific hypotheses. The results showed both direct and indirect pathways between vOT and Broca's area. Moreover, vOT was found to be anatomically linked to angular gyrus via the vertical occipital fasciculus of Wernicke. In addition, the same parts of middle occipital gyrus were seen to send projections to both supramarginal gyrus and vOT. The results suggest that vOT is not the first bifurcation point in the reading network and support a model of reading with multiple routes not only from visual to higher-order language areas but also within the language areas.

Chapter 6 discusses the implications of all the data in this thesis and draws general conclusions on the vOT function, its connectivity and its theories. Limitations and future directions are also discussed in this section. In addition, a revised neuroanatomical model of reading is proposed, based on the data in this thesis.

2. GENERAL METHODS

In this thesis, Magnetic Resonance Imaging (MRI) was used as the primary method. The first two experiments used functional MRI while the third used diffusion-weighted MRI. The overview of these methods is briefly described and the relevance of these methods to my research questions is discussed in this section.

2.1. OVERVIEW OF MAGNETIC RESONANCE IMAGING (MRI)

MR signal generation is based on the quantum mechanical property of atomic nuclei called spin. Atomic nuclei are said to possess this property if they have a magnetic moment and angular momentum. While many atomic nuclei possess this property, hydrogen-1 is most commonly used with MRI since they are abundant in the human body. The nuclei of hydrogen-1 spin around their own axis due to thermal energy and as protons are positively charged, this spinning induces magnetic field due to the laws of electromagnetic induction. As a result, these nuclei act like small bar magnets. In a normal environment, their magnetic moments are randomly aligned and therefore there is no net magnetisation. In a strong external magnetic field, however, their magnetic moments align either parallel (low-energy state) or antiparallel (high-energy state) to the direction of the magnetic field. There are usually more nuclei in low- than high-energy state producing a net magnetisation parallel to the direction of the magnetic field. This is called longitudinal magnetisation. In addition, the nuclei initiate a secondary spin around the direction of the magnetic field (i.e., precession) but are out of phase with each other. The frequency of the precession (called the “Larmor frequency”) is proportional to the strength of the external magnetic field and depends on the isotope of the atom. When a radiofrequency (RF) pulse is applied at the same frequency as the precession frequency of the nuclei at 90 degrees to the magnetic field, the RF energy is absorbed by some of the nuclei in low-energy state turning them into high-energy state (i.e., resonance). The correct amount of

RF energy can equate the number of nuclei in low- and high-energy state producing an equilibrium and as a result, the net magnetisation vector is flipped 90 degrees to the magnetic field, resulting in transverse magnetisation that precesses in this plane. The voltage induced here is measured as the MR signal. While the RF pulse is on, the spins are in phase. Once the RF pulse is switched off, the MR signal begins to decay as the spins starts to lose their phase coherence due to the spin-spin interactions (i.e., decay of transverse magnetisation or T2 decay) and also due to the inhomogeneities of the external magnetic field (i.e., T2* decay). The net magnetisation returns to the longitudinal plane as the nuclei releases the energy absorbed from the RF pulse and return to the low-energy state (i.e., recovery of longitudinal magnetisation or T1 recovery). These exponential changes of net magnetisation over time (i.e., relaxation) are inherent to the different tissue types and also depend on the environment external to the spinning nuclei, namely the field strength. The T1 relaxation time depends on the rate of energy exchange between the nuclei and their environment. In turn, this rate is determined by the size and the density of the molecular structure that affect both the efficacy of energy absorption and the speed of molecular motion. Water has a long T1 relaxation time (i.e., relatively inefficient in energy exchange) as its molecular motion is fast compared to the Larmor frequency. In addition, the water molecules are small and spaced out, which makes it inefficient to absorb and dispose of energy. The T2 relaxation time is determined by the density of the molecules, which affects the efficacy of the spin-spin interaction. Thus water has long T2 relaxation time as the molecules are spaced out. The T2* relaxation time is dependent on the geometric and compositional variations in the applied magnetic field due to varying magnetic susceptibility differences such as the level of deoxyhaemoglobin in the blood.

MR images are dependent on the density of hydrogen-1 protons in the tissues and the contrast made by their relaxation time differences. Therefore the differences in the concentration of water in the tissues contribute to the image. However, these differences must be measured optimally in order to generate sufficient contrast. An appropriate pulse sequence maximises the difference between the contrast parameter of interest (e.g., T1 relaxation time) and the other effect (e.g., T2 relaxation time). This is done by setting the extrinsic contrast parameters such as repetition time (TR) and echo time (TE) accordingly. T1-weighted images provide excellent anatomical detail by differentiating between different tissue types. In order to emphasise T1-weighted signal and attenuate T2-weighted signal, both the TR and the TE must be short. In contrast, for T2-weighted images, both the TR and the TE are long. T2-weighted scans can also provide excellent anatomical detail. In T2*-weighted images, the contrast is based on the differences in the inhomogeneity of magnetic field. This contrast is generated by setting the TR long and the TE to a medium value. T2*-weighted images can reveal subtle magnetic changes due to levels of oxygenation in the blood and thus are the primary scan used for functional magnetic resonance imaging.

In this thesis, two sequences were used. A Fast Low Angle Shot **(**FLASH) sequence was used for acquiring T1 weighted images for detailed anatomical scans and an echo planar imaging (EPI) sequence was used for both functional and diffusion weighted imaging.

2.2. FUNCTIONAL MAGNETIC RESONANCE IMAGING (FMRI)

In chapters 3 and 4, functional magnetic resonance imaging (fMRI) with blood-oxygen-level-dependent (BOLD) contrast was used. fMRI is a non-invasive neuroimaging method with a reasonable spatial and temporal resolution although the fMRI signal is not a direct measure of neuronal activities.

The blood oxygen level dependent (BOLD) contrast is based on the magnetic susceptibility difference of oxyhaemoglobin and deoxyhaemoglobin. While oxyhaemoglobin is diamagnetic and has no net magnetic moment, deoxyhaemoglobin is paramagnetic thus distorts the external magnetic field and adds to the field inhomogeneity producing a T2*-weighted effect (Pauling and Coryell, 1936). In other words, the relaxation time of blood is dependent on oxygenation. Neuronal activity causes physiological changes such as a regional increase in cerebral blood flow, cerebral blood volume and oxygen consumption rate (Logothetis, 2008). As a result, deoxyhaemoglobin decreases, increasing the regional signal intensity in T2*-weighted images and manifests as the BOLD signal (Ogawa et al., 1992). These physiological changes take place over much longer time (> 10 sec) than the actual neuronal responses (< a second). Therefore BOLD signal is an indirect measure of neuronal activities.

The BOLD signal is a measure of physiological changes that are correlated with the neuronal activities. While the direct relationship between the BOLD signal and the underlying neuronal activities is yet to be fully understood, currently available evidence seems to suggest that the BOLD response is approximately linear to the underlying neuronal activities and reflects local field potentials (LFPs) in perisynaptic activity, rather than the rate of spiking (Berens et al., 2013; Ekstrom, 2010; Logothetis, 2002, 2003, 2008; Logothetis et al., 2001; Nair, 2005), although studies in which the BOLD signal, LFPs and spiking activity dissociate have also been reported (see Ekstrom, 2010 for a review). LFPs arise from both excitatory and inhibitory postsynaptic potentials (Berens et al., 2013; Mitzdorf, 1985, 1987). As a result, the BOLD signal is a sum of the two and cannot distinguish between excitatory and inhibitory processes at the cellular level.

Spatial resolution

The fundamental point regarding spatial resolution of fMRI is that what is achievable is far greater in scale than the site of the neuronal activity to be localised. The standard spatial resolution of fMRI is around 8-27 cubic millimetres and although it has been demonstrated that BOLD signal can be localised within 1.1mm at 1.5T (Engel et al., 1997), this still contains millions of cells. In addition, the same vascular system feeds large areas of cortex and thus signal from neighbouring voxels may be correlated based solely on blood flow. Typically, a voxel of 3x3x3mm contains millions of neurons, axons and glia, all of which contribute to metabolic demands. In addition, there may be large or small blood vessels as well as capillaries included in a voxel. The BOLD signal is therefore an aggregate of signals from all of these different types of tissue.

Choosing a spatial resolution of fMRI largely depends on the signal-to-noise ratio (SNR), which is the ratio of the signal relative to the undesired noise in the background. The SNR is a function of the volume of a voxel. Therefore, a 3x3x3mm voxel has more than three times as much SNR as a 2x2x2mm voxel. The appropriate voxel size depends on one’s research question. For instance, larger voxel size (3x3x3mm) may be sufficient for investigating the whole brain whereas smaller voxel size is naturally more appropriate for investigating a single region. However, although smaller voxel size provides more detailed images, reducing the voxel size will lead to a loss of sensitivity. Relatively small BOLD signal changes in response to higher-order processing such as cognitive tasks may not be reliably detectible. In addition to a loss of sensitivity, smaller voxels require longer signal acquisition time, which can cause distortions in the images due to T2* decay that occurs over a long acquisition period. In this thesis, the voxel size of 3x3x3mm was used for fMRI because both fMRI experiments investigated vOT response to different written stimuli during cognitive tasks (lexical decision). Furthermore, since vOT is much larger than 3x3x3mm, it was not necessary to compromise SNR to improve spatial resolution further.

2.3. DIFFUSION-WEIGHTED MAGNETIC RESONANCE IMAGING (DW-MRI) & TRACTOGRAPHY

In chapter 5, anatomical pathways of vOT were investigated using DW-MRI and tractography.

2.3.1. DW-MRI

Diffusion-weighted MRI (DW-MRI) generates images sensitive to the random displacement of molecules due to thermal energy that does not require any external force. The speed of diffusion is intrinsic to the medium and determined by factors such as temperature, molecular size and the microstructure of the environment. Without structural restrictions, diffusion is isotropic (i.e., the same in all directions) and this is the case in the brain grey matter, for instance. However, in a tubular environment such as within the brain white matter fibres, diffusion is restricted across the direction of the fibres while it is still relatively free along the fibres, resulting in anisotropic (i.e., not isotropic) diffusion. Diffusion can be sensitised in MRI to distinctively label tissue microstructure.

DW-MRI utilises strong bipolar gradients that cause phase shift in spins that have displaced, resulting in a signal drop due to increased field inhomogeneity. Spins that are restricted due to the environmental structure do not displace. Consequently they remain in phase and produce higher signal. Diffusion weighting is controlled by an extrinsic contrast parameter, the b-value measured in s/mm^2, which controls the gradient strength, gradient duration and the interval between the gradients. Typical b-values lie between 500 and 1000 s/mm^2. Signal attenuation is dependent on the apparent diffusion coefficient (a measure of the magnitude of diffusion) and b-value. The image contrast is based on the signal attenuation due to diffusion, relative to the signal acquired without diffusion weighting (i.e., when b=0). That is, DW-MRI measures the displacement, but not the directionality, of water molecules. Signal is acquired in many directions so that multiple (possible) directions within a voxel can be measured. Scans are typically repeated and then averaged to increase SNR.

2.3.2. TRACTOGRAPHY

Tractography is an analytic method that allows visualisation of the white matter pathways by integrating voxel-wise diffusion orientations from the DW-MRI data, using a fibre-tracing algorithm. It is important to note that tractography is therefore an indirect method to investigate the white matter fibre structures since it makes assumptions based on the diffusion profile, not the actual fibre orientation. Moreover, it does not detect synapses and cannot infer directionality. However, the results have been demonstrated to be largely in agreement with invasive tracer studies in non-human primates (e.g., Behrens et al., 2003a) and also post-mortem dissections of human fibre paths (e.g., Holl et al., 2011; Sarubbo et al., 2011) as well as histological findings (Hagmann et al., 2003), although there are some inconsistencies concerning terminations. Nevertheless, validation of one's tractography results against surgical dissection, animal studies as well as prior anatomical knowledge is prudent where possible.

Estimating the local diffusion profile

After quantifying diffusion with DW-MRI, the first step is to estimate the local diffusion profile at each voxel. Diffusion tensor imaging (DTI), for instance, models the local diffusion shape as a 3D tensor and estimates the principal diffusion direction (PDD), the direction of least hindrance to diffusion. A tensor is a 3D mathematical model that characterises the diffusion with three eigenvectors and their eigenvalues. The orientation of a tensor is associated with the principal eigenvector, which has the largest eigenvalue corresponding to the PDD. As such, DTI is suitable for identifying major white matter bundles. However, it fails in voxels where there is more than one population of fibres such as crossing or "kissing" fibres (Figure 2.1).

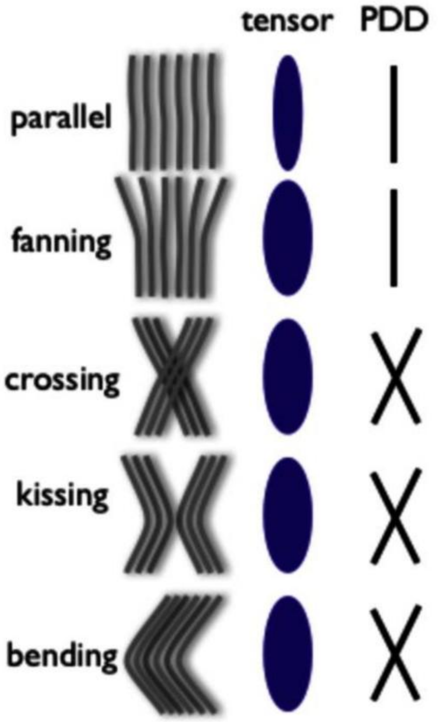

Figure 2.1: Illustration of possible fibre structures, their corresponding tensor orientations and the PDDs. Note that the PDDs for parallel and fanning structures are the same. Similarly, crossing, kissing and bending structures cannot be differentiated by their PDDs. Adapted from Jbabdi and Johansen-Berg (2011), with permission from Mary Ann Liebert, Inc.

DTI cannot correctly model these configurations because the model does not assume multiple directions. As a result, the PDDs may be very different from the underlying fibre orientations. More recently DTI with multiple tensors have been used to tackle multiple fibre orientations (e.g., Parker and Alexander, 2005). However, multi-tensor models produces superfluous directions when there is only single fibre orientation to recover since the model assumes that there are as many fibre orientations as the tensors. Given that the spatial resolution of DW-MRI is relatively large (2-3mm isotropic voxels), it is probable that a single voxel contains more than one fibre direction. However, the information on the number of distinct fibre populations present in a given voxel is rarely available. Therefore, for multi-tensor models to be effective, further techniques that deal with this model selection problem are required. In addition, the currently available techniques cannot cope with more than two distinct fibre directions (Seunarine and Alexander, 2009). Furthermore, the local diffusion profile is in fact complex. As diffusion occurs both along and across the fibre direction, the diffusion profile and fibre orientation do not exactly match. There are also noise and artefacts in the data. These can make the estimate of a single PDD less reliable but crucially, with this method, there is no means of measuring how reliable the estimated PDDs are. This is important since a local error can accumulate when estimating a continuous path from the local PDDs (the next step; see below), creating a path that may not resemble the trajectory of the true fibre structure.

An alternative method, which can tackle these problems, is to estimate the local diffusion profile in the form of a *distribution* of diffusion directions and the number of such directions, rather than as a single PDD. This approach deals with uncertainty in DW-MRI by calculating the uncertainty in the measurement of the local fibre structure. In this thesis, the “ball and stick” model was used to characterise the local diffusion profile and Bayesian methods were used to calculate uncertainty (Behrens et al., 2007; Behrens et al., 2003b). This two-component model assumes that diffusion profile belongs to either an anisotropic population in and around white matter fibres (the fibre direction) or an isotropic one in free water. The noise is modelled separately. This model is fitted to the DW-MRI data at each voxel and the uncertainty given the assumptions made by the model is calculated as posterior density functions (pdfs) using Bayes’ theorem: the posterior density [of the parameters, given the data and the model] is proportional to the likelihood [of seeing the data given the parameters and the model], multiplied by the prior belief [of the parameters given the model]. In this fashion, multiple paths can be estimated within a voxel, each of which is associated with different certainty. For this reason, this method is particularly robust with multiple-fibre orientations.

Model-free methods that do not make any assumptions on the diffusion forms are also available, such as diffusion spectrum imaging (DSI) (Wedeen et al., 2005), Q-ball imaging (Tuch, 2004), spherical deconvolution (SD) (Tournier et al., 2004) or persistent angular structure (PAS) MRI (Jansons and Alexander, 2003). At present the accuracy of the subsequent tractography with these methods is considered to be better (especially with DSI and PASMRI) than model-based approaches. However, the data acquisition requirements are also higher (e.g., a greater number of measurements, longer acquisition time, higher b-values) (Seunarine and Alexander, 2009), making these difficult or impractical methods to apply, given the resources available.

Tractography algorithms

Once the voxel-wise diffusion profiles are estimated, the final step of tractography is to estimate continuous paths using an algorithm using either deterministic or probabilistic approaches. Deterministic tractography begins with a "seed", a starting point (a single voxel) or a mask (a region) and simply tracks sequentially through voxels following their PDDs step by step until the stopping criteria are met. Typical stopping criteria are a minimum FA (fractional anisotropy) threshold and a maximum degree of curvature. Consequently, deterministic tractography cannot trace through or seed from areas with low diffusion anisotropy such as in grey matter. In other words, it is not possible to determine whether the white matter paths identified actually reach the cortical grey matter. Thus, it is not suitable for investigating if two cortical regions communicate with each other. In addition, as mentioned earlier, this method suffers from local modelling errors that can accumulate over a distance when joining the PDDs because there is no means to measure how accurate the recovered paths are.

In contrast, probabilistic tractography propagates uncertainty through the data. Each voxel is represented with a distribution of diffusion directions. In theory, all possible directions need to be integrated into paths. However, this is not practical as the number of possible paths could be astronomical. This problem is overcome by sampling approaches, which repeatedly sample a direction from the probability distribution. Directions with higher probability are more likely to be sampled than those with low probability, effectively integrating all the possible paths with correct weight. In this thesis, Metropolis Hastings Markov Chain Monte Carlo sampling method was used to estimate paths. Moreover, a Bayesian model selection technique, automatic relevance determination (ARD), was applied to model parameters so that it assessed the most appropriate number of distinct fibre orientations at each voxel and ensured that the samples were drawn from each of the distinct fibre populations if and only if these parameters were supported by the data. In the absence of such data, these parameters were set to zero and did not contribute to the probability. In other words, with ARD, tracing propagates through distinct fibre populations, only if the data suggests that multiple fibre orientations exist, which results in better estimation of fibre orientations and their certainty.

As with deterministic tractography, probabilistic tractography begins with a seed. An important difference is that this can be within the cortical grey matter. At the seed, it draws a sample from the probability distribution, moves a fixed distance along this direction and samples again. This procedure is repeated until it reaches the stopping criteria, which are much less strict than those of deterministic tractography – typically the edge of the brain or a lenient curvature threshold to prevent looping back to the current location. In other words, there is no minimum FA value for a voxel, allowing probabilistic tractography to trace paths into the grey matter. The number of samples is specified by the user. Larger numbers obviously represent the population more accurately. In practice, many studies have used 5000 or fewer (Behrens et al., 2007; Eickhoff et al., 2010; Tomassini et al., 2007) and 5000 samples per seed voxel is found to be sufficient to reach convergence (http://fsl.fmrib.ox.ac.uk/fsl/fdt/fdt_probtrackx.html; Carreiras et al., 2009). Consequently, the two-ROI experiments reported in Chapter 5 used 5000 samples while the first, single-ROI analyses used 25000 samples in order to optimise sensitivity in the absence of further constrains to limit the paths.

It is also important to emphasise that the continuous paths produced by probabilistic tractography are the spatial probability distribution of paths from the seed, not the distribution of the paths from the seed (Behrens and Jbabdi, 2009). The results (the value) at each voxel is the discretised spatial probability distribution, calculated by the number of paths going through the voxel divided by the total number of samples (the number of samples multiplied by the number of voxels in the seed). It indicates how probable that a streamline from the seed passes through a given voxel. This means that the probability value is in part dependent on the size of the seed masks, which can vary across subjects and the voxel size. It is possible to normalise the inter-subject differences in mask sizes by calculating the total number of samples sent out first and then threshold this number equally across subjects (e.g., Rilling et al., 2008). However, this may or may not be necessary depending on one's research question. For example, normalisation is particularly important when there might be quantitative and qualitative differences such as across different species (e.g., macaques vs. humans), groups (e.g., patients vs. healthy controls) or analyses (e.g., single mask vs. multiple masks). However, normalisation is not always used especially when anatomical connectivity between particular regions is of interest. In this case, thresholding can be applied to the number of samples per voxel, rather than the total number of samples sent out from a seed mask (e.g., Ciccarelli et al., 2006; Ford et al., 2010).

Thresholding is difficult and there is no "gold-standard" at present. Choosing the appropriate percentage requires checking whether it allows anatomically plausible paths while it correctly rejects extraneous paths. Interpreting the results is thus largely dependent on one's anatomical knowledge and as mentioned earlier, validating the results against the findings from other methods is valuable. In addition, when additional masks such as target masks are used in order to limit the paths, thresholding may not be reported (e.g., Anwander et al., 2007) since the target masks eliminate both extraneous and irrelevant paths. In the experiments presented in Chapter 5, the threshold of 0.1% of the number of samples per voxel was used without normalisation, after examining the effects of various thresholds on the data. Similar thresholds (0.1 – 0.2% of the number of samples per voxel) were used in previous studies (e.g., Devlin et al., 2006b; Ford et al., 2010; Heiervang et al., 2006).

The great advantage of probabilistic over the deterministic tractography is the ability to trace through areas of low anisotropy such as grey matter or where the data are noisy (i.e., areas with high uncertainty). Therefore, probabilistic tractography is more suitable for investigating anatomical connections between cortical regions conducted in this thesis. As the aim of Chapter 5 was to investigate the anatomical pathways of vOT, it was vital to be specific about the cortical site of the start and/or termination points where the FA is low. Probabilistic tractography allows seeding grey matter such as vOT. Thus it is possible to determine whether the pathways recovered actually link vOT and other cortical regions thereby confirming that these regions have the means of communication. Seeding white matter near vOT would not produce the same results since it cannot differentiate paths linking vOT and other regions from those that are physically near but in fact merely “passing-by”.

3. INTERACTION OF TOP-DOWN INFLUENCE AND BOTTOM-UP INFORMATION

A version of this chapter was published as “Top-down modulation of ventral occipito-temporal responses during visual word recognition” by Tae Twomey, Keith J. Kawabata Duncan, Cathy J. Price & Joseph T. Devlin in *Neuroimage* (2011) 55: 1242–1251.

3.1. INTRODUCTION

Although cognitive models of reading emphasise the importance of interactive processing during visual word recognition, most neuroanatomical models of reading have focused on the feed-forward flow of information. In the classic neurological model of reading, as described in Chapter 1, visual input arrives at the occipital pole and projects to the angular gyrus where visual word forms are stored (Dejerine, 1891, 1892). These then link to auditory word forms in the posterior superior temporal lobe (i.e., Wernicke's area) and from there to articulatory motor patterns in the inferior frontal gyrus (i.e., Broca's area). In this linear fashion, a written word is recognised, converted into a sound then motor pattern, and read aloud. More recent studies elaborate additional anatomical territories (Bitan et al., 2009; Dehaene et al., 2005; Frost et al., 2008; Price and Mechelli, 2005), allow for multiple parallel pathways (Devlin, 2009; Mechelli et al., 2005), and characterise the functional contributions of the component regions differently (Shaywitz and Shaywitz, 2008). Even so, most neural models of reading continue to involve an essentially feed-forward, staged processing dynamic (Dehaene et al., 2005; Kronbichler et al., 2004).

At a behavioural level it is generally agreed that reading requires interaction between visual and non-visual properties of the written stimulus. A classic example is the "word superiority effect" where there is a perceptual advantage for identifying letters in words relative to visually matched letter strings that do not form words (Cattell, 1886). Reicher (1969), for instance, used a two-alternative single letter choice paradigm with masked tachistoscopic displays of visual letters, words or anagrams. Participants were instructed to identify one of the two letters that appeared in four-letter words (e.g., WORD), its anagrams composing nonwords (e.g., ORWD) or single letters (e.g., D). The performance was found to be more accurate for the four-letter words than the other two conditions. The results were replicated by Wheeler (1970). This phenomenon suggests, *inter alia*, that word recognition is not merely recognition of a set of individual letters but instead implies an interaction amongst letters that form a word (Wheeler, 1970). The fact that letter detection is affected by whether or not the stimulus is a word – namely, by information not present in the visual display – illustrates that this information is automatically retrieved and fed back to affect visual processing (McClelland and Rumelhart, 1981). Another clear example is the finding that when participants make lexical decisions (i.e., decide whether a letter string forms a real word), they are slower to reject an item that sounds like a word (i.e., pseudohomophones, e.g., “furst”) than one that that does not (e.g., “gurst”, McCann et al., 1988). McCann and colleagues (1988) conducted a visual lexical decision experiment with words (“yes” response) and both monosyllabic pseudohomophones and nonwords (“no” response). Within the “no” responses, the reaction times for the pseudohomophones were slower and the accuracies were lower than the nonwords. This pseudohomophone effect illustrates that automatic retrieval of phonological and/or semantic information that is not essential for task performance can nonetheless affect behaviour. In addition, it has been shown that visual word recognition is easier when a word is preceded by a semantically related word (e.g., “butter” preceded by “bread”). This suggests that the context provided by the first lexical item, which is not encoded visually, supports identification of subsequent words by way of feedback connections particularly when the task elicits semantic access as demonstrated with word naming (Reimer et al., 2008) or lexical decision (Smith and Besner, 2001). Furthermore, this facilitation is not limited to real words but includes nonwords that bear visual resemblance to existing real words. Rosson (1983) compared naming latency of words preceded by a word or nonword, which can be either semantically related or unrelated (e.g., “SHEEP” preceded by “LAMB” - semantically related real words; or by “FAMB” - semantically “related” nonword [semantically related nonwords were such that they were visually similar to semantically related real words]). He found that in addition to semantically related words, “semantically related” nonwords decreased naming latency of the subsequent words. The possibility that this effect was caused simply by visual similarity rather than access to lexical information was rejected following another experiment in which pronunciation of nonwords was successfully biased by preceding words that had semantic associations with a related word (e.g., nonword “LOUCH” preceded by either “FEEL” to facilitate the pronunciation similar to “TOUCH”, or “SOFA” to facilitate the pronunciation similar to “COUCH”). Altogether these results were interpreted as evidence for lexical information influencing the pronunciation of nonwords. These and other similar observations (Frost, 1998) demonstrate the need for feedback connections linking non-visual to visual information processing, thus creating an interactive (rather than feed-forward) system for visual word recognition (Coltheart et al., 2001; Harm and Seidenberg, 2004; Jacobs et al., 2003; McClelland and Rumelhart, 1981; Perry et al., 2007; Plaut et al., 1996; Rumelhart and McClelland, 1982).

Indeed, interaction is an integral part of cognitive and computational models. The interactive activation model (McClelland and Rumelhart, 1981; Rumelhart and McClelland, 1982) for instance, assumes perception as an interactive process, in which many different levels of processing occur in parallel, both spatially within a level and across levels. For visual recognition, there are visual feature, letter and word levels, together with an (unspecified) higher level that provides top-down input to these lower levels. According to this model, our perception is governed by the interaction of simultaneous top-down (conceptually driven) and bottom-up (data driven) processing (Figure 3.1). Its successors such as the triangle model (Seidenberg and McClelland, 1989) and the dual-route cascading (DRC) model (Coltheart et al., 2001) each postulate different routes and mechanisms. However, they all share this fundamental assumption of interactivity and maintain that this is necessary for their performance to match the observed behaviour.

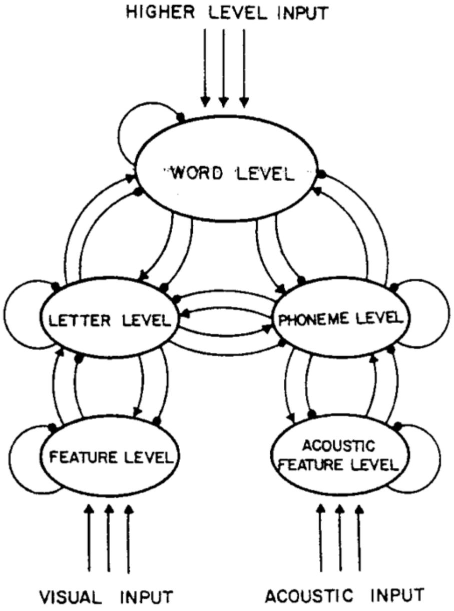

Figure 3.1: The general conception of the interactive activation model. Lines with arrow heads indicate excitatory connections and the ones with circular ends indicate inhibitatory connections. Interaction occurs both within and across levels. Copyright © 1981 by the American Psychological Association. Reproduced with permission from McClelland and Rumelhart (1981).

It is also important, however, to note that purely feed-forward cognitive accounts reject any involvement of top-down feedback and suggests instead that the word superiority effect and the pseudohomophone advantage can be explained without any interactions (Johnston, 1981; Norris et al., 2000; Paap et al., 1982). According to Norris and colleagues (2000), letter information that influences word recognition is available from the lexicon itself and is activated by the visual input. The word superiority effect, for instance, occurs because information (e.g., whether a letter string contains "m") is fed forward to the decision making stage, rather than information regarding the lexicality of a letter string being fed back to influence the perception of the stimuli. Since only real words are in the lexicon, this information is available only for words, resulting in the latency advantage for words. At any rate, feed-forward accounts maintain that reading does not require any interactions of bottom-up and top-down information but rather follows steps of feed-forward processing. Nevertheless, if this is the case, this effect should be reflected in the neural responses in the regions implicated in the decision making, not in the early perceptual stage of processing.

This discrepancy between cognitive interactivity, on the one hand, and feed-forward (neuroanatomical or cognitive) models, on the other, is particularly relevant to theories of ventral occipito-temporal (vOT) cortex functioning during reading and can be used to separate competing accounts of vOT. As described in the previous chapter, this region of extrastriate visual cortex is consistently engaged during visual word recognition and damage to the area can result in severe reading deficits (Behrmann et al., 1998; Cohen et al., 2000; Leff et al., 2001; Philipose et al., 2007; Starrfelt et al., 2009) and is thought to play an important role in orthographic processing (McCandliss et al., 2003; Price and Mechelli, 2005). A detailed account can be found in the preceding chapter, but briefly, the LCD model (Dehaene and Cohen, 2011; Dehaene et al., 2005) suggests that visual information is encoded through a sequence of stages, from simple feature detectors located in early visual cortex, to letter detectors in V4, to bigram detectors in vOT, and then on to whole word detectors located even more anteriorly in the temporal lobe (Dehaene et al., 2005). In other words, orthographic information is progressively extracted following hierarchical, feed-forward steps that detect progressively more complex visual features. Although by this account vOT receives primarily bottom-up visual information, the authors note that certain attentional manipulations can also provide a top-down signal such as when participants are asked to visualise written words (Cohen et al., 2004; Cohen et al., 2002). For example, although auditory words do not typically engage vOT (Dehaene et al., 2002; Spitsyna et al., 2006), a recent study found that when participants selectively attended to auditory words it produced activation within the region (Yoncheva et al., 2010) and the authors suggested that participants accessed the orthographic codes associated with the auditory stimuli during a challenging auditory task. This type of strategic attentional control to recruit orthographic representations in the absence of visual input, however, is fundamentally different from the automatic interactions between visual and non-visual (e.g., phonological or semantic) properties of a visual stimulus such as a word. A clear distinction between these two types of top-down influences is that strategic influences are purposely exerted to help perform the task whereas automatic top-down influences are involuntary. Consequently, these influences could interfere with task performance, as demonstrated by the pseudohomophone effect discussed above. Another example of automatic interaction was demonstrated in the previously mentioned study by Kherif and colleagues (2011). In this study, vOT activation for reading object names was suppressed when primed with a masked picture of the same object relative to a masked picture of a different object, suggesting that non-visual processing that is common to words and pictures (e.g., semantics and phonology) was influencing vOT activation. Crucially, these could not be expectation-driven attentional effects because the visual masked priming paradigm precluded subjects from having conscious awareness of the primes. Instead, these priming effects provide strong evidence of automatic interactions between the different types of visual and non-visual information important for reading words. These interactions are the type of top-down processing, carried in the feedback connections, that are crucial to cognitive and computational models of reading (Coltheart et al., 2001; Harm and Seidenberg, 2004; Jacobs et al., 2003; Perry et al., 2007; Plaut et al., 1996) but missing from most neuroanatomic models (e.g., Cohen et al., 2002; Dehaene et al., 2005; Kronbichler et al., 2004). An alternative neural model the Interactive Account (Price and Devlin, 2011), as previously described, suggests that vOT continuously and automatically interacts with other regions during reading, acting as an interface associating bottom-up visual form information critical for orthographic processing with top-down higher order linguistic properties of the stimuli (Cai et al., 2010; Devlin et al., 2006a; Hillis et al., 2005; Kherif et al., 2011; Nakamura et al., 2002; Price and Friston, 2005; Xue et al., 2006). These non-visual properties are sent into vOT as automatic top-down predictions and interact with the bottom-up visual information so that the response to the stimuli can be predicted (Price and Devlin, 2011). Thus, vOT activation is a combination of top-down prediction, bottom-up sensory information and the mismatch between the two (prediction error). Unlike the feed-forward LCD model, interaction is an integral part of the Interactive Account.

Ideally, evidence for the direction of information flow in the reading network requires effective connectivity analyses that measure how activity in one region is influenced by activity in other regions. Such inferences are possible with dynamic causal modelling (DCM) of fMRI data (Friston et al., 2003). However, current implementations of this technique can only test the interactions among a limited number of regions specified in the model. DCM therefore relies on knowing, *a priori*, where top down inputs to vOT are coming from. Without specifying such regions, evidence for top-down signals into vOT would be hard to obtain with DCM. Effective connectivity, however, is not the only way to assess top-down effects. Standard fMRI experiments can also be informative when they are designed so that the contribution of top-down influence could be teased apart from bottom-up responses.

The aim of this study was to investigate whether activation in vOT during visual word recognition is influenced by top-down non-visual information. Participants performed two different types of lexical decision tasks that focused attention on either visual (i.e., orthographic) or non-visual (i.e., phonological or semantic) properties of the stimulus. In one, participants were asked to decide whether the letter string was a real English word or not. Half of the stimuli were words (e.g., “brain”) and the other half were pseudohomophones – that is, pronounceable nonwords that sound like real words such as “brane.” When performing this task, participants had to focus on the visual properties of the stimuli to make the correct response since phonological and semantic properties of the stimuli would not differentiate a real word from a pseudohomophone. In the other task, participants were asked to decide whether the letter string on the screen sounded like a real word or not. Half of the stimuli were pseudohomophones (e.g., “beest”) and the other half were pseudowords (e.g., “beal”). In this task, participants had to focus on the phonological (and possibly semantic) properties of the stimuli to make the correct response since the visual properties of the stimuli were insufficient to perform the task as neither type of stimuli was visually a word.

Unlike previous studies that only used a single task ("Does the item sound like a word?" Bruno et al., 2008; Kronbichler et al., 2007; van der Mark et al., 2009), this design enabled me to examine two different types of top-down processing, namely stimulus-driven and task-driven effects. Stimulus effects were evaluated within task by carefully matching the stimuli on a range of visual properties (see below) such that if processing was primarily feed-forward, vOT activation would be expected to be comparable across conditions. If, on the other hand, the region also receives feedback from higher order areas, then non-visual properties would be expected to significantly modulate vOT activation levels. Task effects were evaluated by holding the stimulus constant and comparing the activations to pseudohomophones across tasks. Feed-forward accounts predict that pseudohomophone activations in vOT would either be comparable across tasks (as the stimuli were carefully matched) or possibly increased for orthographic relative to phonological lexical decisions. In the case of a purely feed-forward account, increased activation in vOT during the orthographic relative to phonological task could be based solely on increased local processing demands without requiring any feedback interactions. In contrast, increased activation in vOT during the phonological relative to orthographic task would indicate greater interactions between regions involved in phonological and orthographic processing, consistent with feedback connections linking these areas. Here these predictions were tested using functional magnetic resonance imaging.

3.2. MATERIAL AND METHODS

Participants

20 monolingual native English speakers (11M, 9F) participated in this study. All were from the British Home Counties (i.e., southern England) with the same regional accent, which was important for consistent pronunciation of nonwords. The data from four participants were excluded in total: one subject was excluded due to excessive motion inside the scanner (>3mm); one subject was excluded due to task performance that was not significantly above chance (i.e., <65% accuracy); and two subjects were excluded because unexpected structural abnormalities were present in their T1 images. The ages of the remaining 16 (9M, 7F) participants ranged from 19 to 43 (M=30). All were right-handed and none reported any history of neurological problems or reading difficulties. The experiment was approved by the NHS Berkshire Research Ethics Committee.

Tasks & Stimuli

There were two lexical decision tasks that forced participants to attend to different aspects of the stimuli. The first task emphasised visual over non-visual properties of the stimuli whereas the second emphasised non-visual over visual information. Consequently, these will be referred to as the ‘orthographic’ and ‘phonological’ lexical decision tasks, respectively. For both tasks, participants viewed a string of letters presented sequentially. For the orthographic lexical decision task, participants were instructed to decide whether the string formed an existing English word or not. For the phonological lexical decision task, participants were asked to decide whether the string sounded like an existing English word or not (Figure 3.2a).

A behavioural pre-test was conducted with an independent set of 52 (28M, 24F) participants to pilot the stimuli and establish baseline performance in a reasonably large sample. The stimuli were provided by Sarah White and the data were collected by Caroline Ellis, Odette Mergrin, Stephanie Burnett and Sue Ramsden at the Wellcome Trust Centre for Neuroimaging, UCL. All participants were monolingual native English speakers aged 17 to 69 (M=27). For the orthographic lexical decision task, there was no significant difference in accuracy between words and pseudohomophones (93.7% vs. 93.1%, *t*(51)=.50, *p*=.622) but responses to words were significantly faster (779 vs. 1052msec, *t*(51)=11.37, *p*<.001). For the phonological lexical decision task, responses to pseudohomophones were less accurate than to pseudowords (85.1% vs. 88.9%, *t*(51)=2.02, *p*=.049) but were significantly faster (1061 vs. 1478, *t*(51)=10.78, *p*<.001), possibly indicating a speed-accuracy trade-off. Anecdotally it became clear that because the participants in this behavioural pilot study came from geographically diverse areas of the UK, different regional accents contributed additional variability to the phonological lexical decision task due to different pronunciations of nonwords. Even so, a fairly large sample size ensured an adequate estimate of baseline performance. Given the smaller sample used in the fMRI study, I chose to recruit from a more uniform population of accents to minimise this variability.

Following the behavioural pre-test, stimuli were revised to exclude ambiguous items and the final stimulus set used for the fMRI tasks was comprised of 48 stimuli in each condition (192 stimuli in total). Stimuli were all monosyllabic and balanced for the number of letters (M=4.5, *F*(3,188)=1.07, *p*=.364), frequency of single letters (M=281379, *F*(3,188)=.196, *p*=.899), bigram frequency (M=1553, *F*(3,188)=1.52, *p*=.211), trigram frequency (M=258, *F*(3,188)=1.85, *p*=.141) and orthographic neighbourhood (M=6.1, *F*(3,188)=.13, *p*=.943) based on N-Watch (Davis, 2005). For the word condition, the mean frequency per million words of British English was 76 as derived from the Celex database (Baayen and Pipenbrook, 1995) and the mean familiarity rating was 430 and was calculated from the MRC Psycholinguistic Database (Coltheart, 1981). For each task, the full set of 96 stimuli was divided evenly into two runs of 48 trials. For the orthographic lexical decision task, it was ensured that no pairs of a real word and its pseudohomophone (e.g., “brain” and “brane”) occurred in the same run in order to avoid any priming effects. A different set of pseudohomophones were used in the phonological lexical decision task to ensure that no stimulus was repeated across tasks in order to avoid any priming effects and to avoid switching response type from “no” (in orthographic task) to “yes” (in phonological task) for the identical stimuli. These two sets of pseudohomophones will be referred to as PH_1 (orthographic task) and PH_2 (phonological task) to emphasise the fact that the stimulus sets were independent. The base words of PH_1 and PH_2 were balanced for frequency (M=59, *t*(58)=1.10, *p*=.275) and familiarity (M=457, *t*(85)=1.40, *p*=.165) to ensure that if differences are observed between pseudohomophones across tasks, these are the result of task-differences rather than potential psycholinguistic confounds. The order of both tasks and stimulus sets within a task were fully counter-balanced across participants.

A mixed block and event-related design was used. Participants performed a 33s block of trials which included both “yes” and “no” responses in a pseudorandomised order. These were separated by 16s blocks of fixation which served as an implicit baseline. Each trial began with a fixation cross presented for 500msec. A stimulus was then presented for 200msec, followed by a jittered inter-stimulus interval of 1800 – 4800msec (M=3300msec). Therefore, the average trial length was four seconds. Stimuli were presented in a block of 8 trials. Over a run, there were six blocks of task performance and five blocks of rest. Therefore, each run lasted 4.85mins and there were a total of four runs (two per task). Responses were made with a button press, using either the index or middle finger of their right hand to indicate “yes” and “no”. The response fingers were fully counter-balanced across participants. The stimuli were projected onto a screen and viewed via mirrors attached to the head coil. Participants practiced each task inside the scanner before the main runs began. No items that were used in the practice runs occurred during the main experiment.

MRI acquisition

Whole-brain imaging was performed on a Siemens Avanto 1.5 T MR scanner at the Birkbeck-UCL Neuroimaging (BUCNI) Centre in London. The functional data were acquired with a gradient-echo EPI sequence (TR=3000msec; TE=50msec; FOV=192×192; matrix=64×64) giving a notional resolution of 3×3×3 mm. Each run consisted of 97 volumes and as a result, the four runs together took 19.4 min. In addition, a high-resolution anatomical scan was acquired (T1-weighted FLASH, TR=12msec; TE=5.6msec; 1 mm^3 resolution).

Analyses

Items whose accuracy was below 65% were excluded from all analyses (n=10). Reaction times (RTs) were recorded from the onset of the stimulus. To minimise the effect of outliers, median RTs for correct responses per condition per subject were used in the statistical analyses (Ulrich and Miller, 1994). Because the two tasks used different types of stimuli (words and pseudohomophones vs. pseudohomophones and pseudowords), the experimental design was not factorial. Consequently, the data were analysed using a repeated measures 1×4 analysis of variance (ANOVA) with Condition as the independent variable. For the behavioural data, accuracy and RTs were the dependent measures. Where Mauchly's test indicated significant non-sphericity in the data, a Greenhouse-Geisser correction was applied. When there was a main effect of Condition, planned comparisons used paired t-tests to evaluate differences between the two conditions per task to evaluate stimulus effects and between the two pseudohomophone conditions to evaluate task effects.

The imaging data were processed using FSL 4.0 (www.fmrib.ox.ac.uk/fsl). The first two volumes were discarded in order to allow for T1 equilibrium. The data were then realigned to remove small head movements (Jenkinson et al., 2002), smoothed with a 6 mm full width at half maximum Gaussian kernel, and pre-whitened to remove temporal autocorrelation (Woolrich et al., 2001). The preprocessed data from each subject were then entered into a first level statistical analysis and modelled as events using a general linear model. The two main regressors corresponded to the correct trials from the two task conditions (per task) and these were convolved with a double gamma canonical hemodynamic response function (Glover, 1999). Eight additional regressors-of-no-interest were added: i) errors trials (Murphy and Garavan, 2004), ii) six estimated motion parameters, and iii) RTs. It is important to note that the inclusion of RTs in the model only accounts for first-order (i.e., linear) effects and therefore higher-order (i.e., polynomial) relations between effort (as indexed by RTs) and BOLD signal may remain. Nonetheless, simple correlations between effort and BOLD signal were treated as a covariate-of-no-interest in order to model systematic differences in effort between conditions seen in the behavioural pilot. To remove low frequency confounds, the data were high-pass filtered with a cut-off point of 100 seconds. The contrasts of interest at the first level were the two experimental conditions relative to fixation per task. First level results were registered to the Montreal Neurological Institute (MNI)-152 template using a 12 degree of freedom affine transformation (Jenkinson and Smith, 2001) and all subsequent analyses were conducted in the MNI standard space. A second level fixed-effects model combined the two first level runs into a single, subject-specific analysis (per task) which was then entered into a third level, mixed effects analysis to draw inferences at the population level (Beckmann et al., 2003; Woolrich et al., 2004).

The first analysis identified areas of activation that were common to all four conditions using a linear contrast to compute their mean activity (i.e., [1 1 1 1]) and inclusively masked it with each condition relative to fixation at Z>3.1 (i.e., masking with [1 0 0 0], [0 1 0 0], [0 0 1 0], and [0 0 0 1]). A second analysis used a 1 × 4 ANOVA to identify areas showing significant differences across conditions (i.e., a main effect of Condition identified using an F-contrast). These were characterised by plotting the mean effect sizes per condition in a sphere (5mm radius) centred on the peak coordinate.

Since the primary aim of this study was to investigate the top-down modulation on left vOT, an *a priori* anatomical mask for this region was defined. The main anatomical areas of interest are the occipito-temporal sulcus and adjacent regions on the crests of the fusiform and inferior temporal gyri: areas consistently activated by visual word recognition tasks (Bitan et al., 2007; Cai et al., 2010; Cohen et al., 2000; Devlin et al., 2006a; Duncan et al., 2009; Fiez and Petersen, 1998; Frost et al., 2005; Herbster et al., 1997; Kronbichler et al., 2007; Price et al., 1996; Rumsey et al., 1997; Shaywitz et al., 2004; van der Mark et al., 2009). Because the precise coordinates vary along a rostro-caudal axis, standard space coordinates ranging from X= −30 to −54 and Y= −45 to −70 were used to delineate this region. In addition, the depth of the sulcus coupled with the fact the temporal lobe is angled downwards required a range of Z-coordinates as well (Z= −30 to −4). Together these coordinates describe a rectangular prism that conservatively encompass the anatomical regions-of-interest but also include parts of the cerebellum that were not of interest. Consequently these were manually removed from the mask. A small volume correction determined that a voxel threshold of $Z>3.2$ corresponded to $p<.05$ after correcting for the number of independent comparisons within the region (Worsley et al., 1996) and this was used for all vOT analyses. With an unconstrained, whole brain search, a corrected voxel-wise p-value of .05 corresponded to $Z>4.6$. To minimise Type II errors, activations present at $Z>4.0$ are also reported as trends.

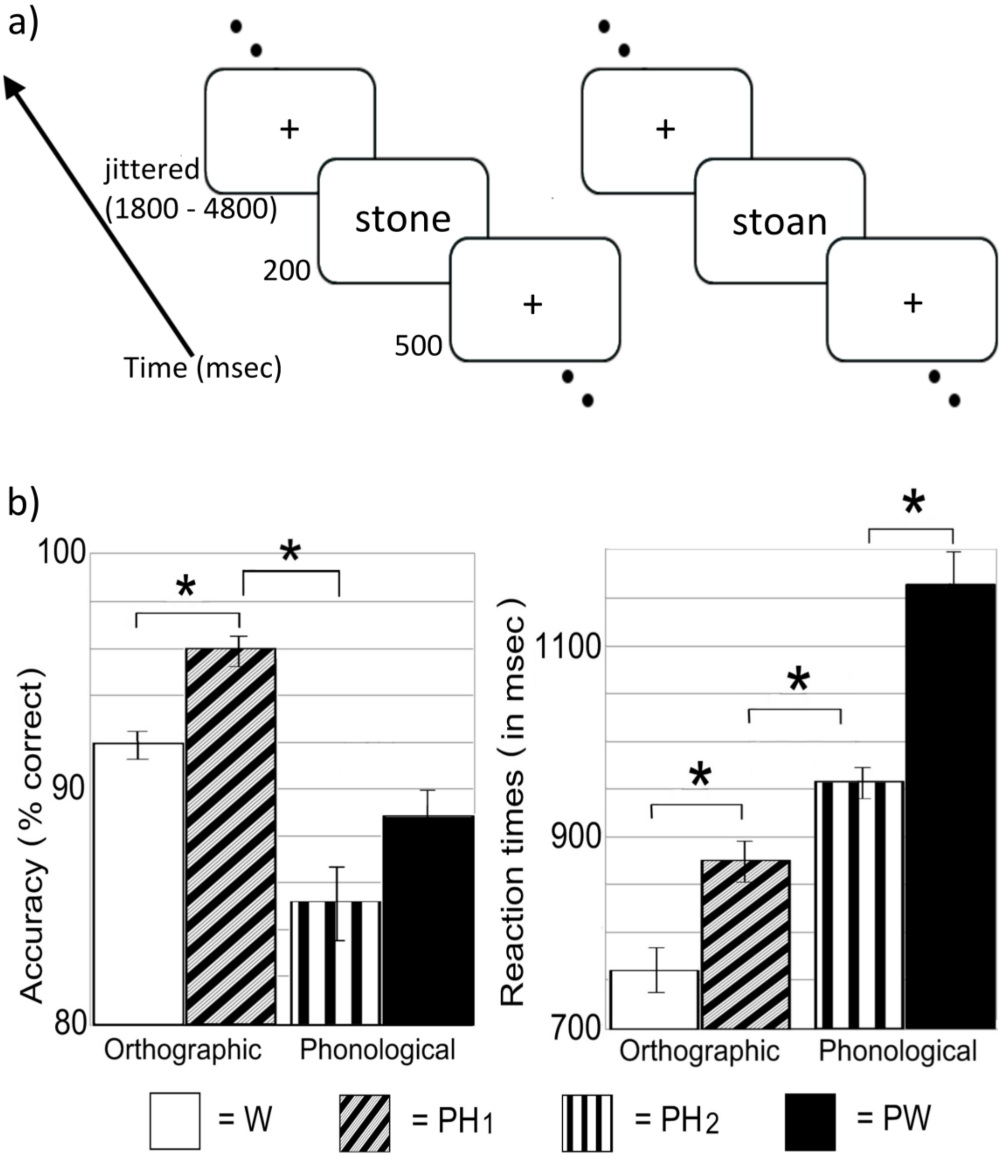

Figure 3.2: a) Schematised task. Each trial began with a fixation cross presented for 500msec. A stimulus was then presented for 200msec, followed by a jittered inter-stimulus interval of 1800 – 4800msec (M = 3300msec). b) Mean accuracy and reaction times for all four conditions. An asterisk (*) indicates *p*<.05. Abbreviations: W = Words (orthographic task), PH1 = Pseudohomophones (orthographic task), PH2 = Pseudohomophones (phonological task) and PW = Pseudowords (phonological task).

3.3. RESULTS

Behavioural results

The behavioural data (Figure 3.2b) demonstrated significant differences across conditions for both accuracy (*F*(3,45)=11.98, *p*<.001) and RTs (*F*(1, 22)=31.90, *p*<.001, with Greenhouse-Geisser correction). Moreover, Figure 3.2b clearly shows evidence of both stimulus- and task-related differences. In the orthographic task, responses to words were less accurate (92% vs. 96%, *t*(15)= 2.98, *p*=.009) but faster (761 vs. 874msec, *t*(15)= 6.76, *p*< .001) than responses to pseudohomophones. A similar pattern was present in the phonological task. Here, responses to pseudohomophones were numerically less accurate (85% vs. 89%, *t*(15)=1.74, *p*=.102) but significantly faster (956 vs. 1162msec, *t*(15)=5.30, *p*< .001) than responses to pseudowords. In other words, like the behavioural pre-test, these results suggest that participants may have adopted a speed-accuracy trade-off within each task. Therefore, when analysing the imaging data, I considered only correct trials and explicitly modelled RTs on a trial-by-trial basis to account for these first order, systematic differences between conditions. In addition to these stimulus effects, there was also a significant task effect when comparing the pseudohomophone conditions. Responses were more accurate (96% vs. 85%, *t*(15)= 4.69, *p*< .001) and faster (874 vs. 956msec, *t*(15)= 2.32, *p*= .035) when participants made orthographic relative to phonological lexical decisions. In summary, the behavioural results demonstrate both stimulus- and task-effects on behaviour, consistent with top-down influences in visual word recognition (McCann et al., 1988).

Imaging results: Common system

I began by identifying the common system of regions activated by all four conditions (Figure 3.3). As expected, there was strong bilateral activation in vOT centred on the posterior occipito-temporal sulcus that extended inferiorly into lobule VI of the cerebellum. In addition, there was bilateral activation in the early visual cortices of the calcarine sulcus, in the intraparietal sulcus, the deep frontal operculum and at the junction of the inferior frontal and precentral sulci. There was also left hemisphere activation in the pre-SMA, the anterior supramarginal gyrus and within sensori-motor cortices that included the omega-knob marker for the hand area (Yousry et al., 1997). In other words, these results correspond closely to previous lexical decision studies, validating the success of the task (Carreiras et al., 2007; Devlin et al., 2006a; Fiebach et al., 2007; Gold et al., 2006; Kiehl et al., 1999; Mummery et al., 1999; Rumsey et al., 1997). Table 3.1 provides the full details of these activations and illustrates that for each region, there is activation in each of the four conditions. Presumably these reflect common aspects of the two tasks including not only visual word recognition, but also sustaining attention, maintaining a cognitive set and making manual responses.

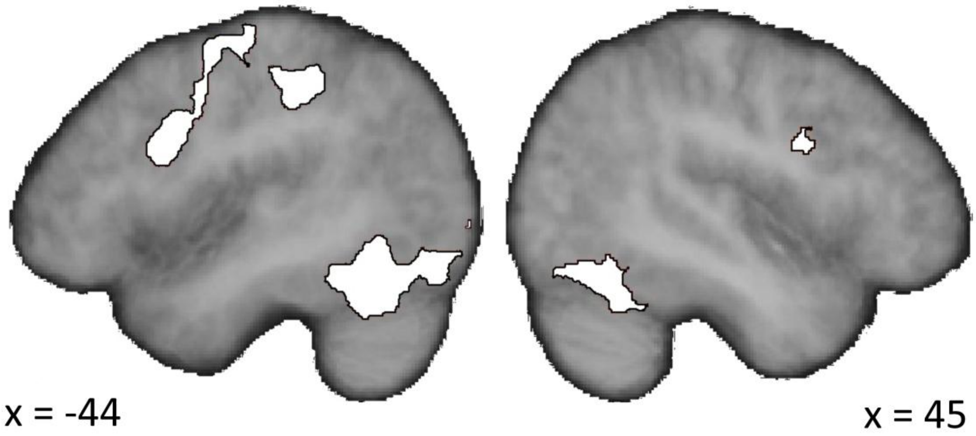

Figure 3.3: The brain areas commonly activated for all four conditions relative to fixation. Activations are thresholded at *Z*> 3.1 and shown as white areas (outlined in black) on two parasagittal slices through the mean structural image of the group in standard (i.e., MNI152) space.

Table 3.1: Common activations across the four conditions relative to fixation. For each peak in the mean activation contrast, its anatomical location, *Z*-score and standard space (i.e., MNI152) coordinate is displayed. In addition, the *Z*-score at that peak is shown for each of the four individual conditions relative to fixation to illustrate that activation was present for all four conditions.

Region		*Z*-score	Mean peak coordinate			*Z*-score relative to rest			
			x	y	z	Orthographic		Phonological	
						W	PH$_1$	PH$_2$	PW
Occipital									
L	vOT	11.6	–44	–56	–15	4.5	4.6	5.1	5.1
R	vOT	8.7	45	–63	–13	3.5	3.4	3.3	3.9
L	Calcarine sulcus	9.5	–7	–76	8	4.0	4.4	4.5	4.0
R	Calcarine sulcus	9.2	9	–74	12	4.2	4.3	4.3	3.8
Parietal									
L	Intra-parietal sulcus	10.2	–27	–52	46	3.8	4.1	5.0	4.7
R	Intra-parietal sulcus	9.0	27	–56	47	4.0	4.4	4.8	4.0
L	Supramarginal gyrus	10.4	–48	–33	46	3.6	4.0	4.4	3.7
L	Parietal operculum	8.7	–54	–17	18	5.0	3.7	3.4	4.2
L	Postcentral gyrus	10.3	–40	–21	50	3.7	4.0	3.1	3.1
Frontal									
L	Frontal operculum	8.5	–31	24	2	3.7	4.3	4.0	4.2
R	Frontal operculum	9.5	33	25	–3	4.9	4.6	4.5	4.4
L	IFS/PCS junction	11.1	–42	7	26	4.1	4.4	5.2	4.5
R	IFS/PCS junction	9.3	44	5	28	4.2	3.6	3.3	3.6
L	Pre-SMA	10.6	–3	15	45	4.8	5.1	5.6	5.3
L	Precentral gyrus	9.1	–44	–1	40	3.6	4.5	4.1	4.8
Subcortical									
L	Cerebellum (VI)	8.1	–6	–73	–20	3.8	4.5	4.0	4.5
R	Cerebellum (VI)	10.3	21	–52	–22	4.7	4.5	4.6	4.3
R	Cerebellum (VI)	10.1	35	–49	–23	5.3	4.9	4.3	4.7
R	Cerebellum (VI)	8.8	11	–25	–22	4.3	3.9	3.6	4.1
L	Putamen	7.3	–26	–1	0	3.6	3.9	3.8	3.9
L	Thalamus (MD)	8.1	–12	–18	5	3.8	3.9	3.7	4.8

Abbrev: W=Words (orthographic task), PH$_1$=Pseudohomophones (orthographic task), PH$_2$=Pseudohomophones (phonological task) and PW=Pseudowords (phonological task); vOT=ventral occipito-temporal cortex, IFS=inferior frontal sulcus, PCS=precentral sulcus, SMA=supplementary motor area, MD=mediodorsal nucleus.

The critical analysis, however, looked for activation differences across the four conditions reflecting the different top-down processing demands. Areas that were significantly affected by Condition were identified from the F-map of the one-way ANOVA and fell into two classes. The first set included ventral occipito-temporal cortex and pars opercularis (POp) – where activation was increased during all conditions relative to fixation. The second set included the angular gyrus, medial prefrontal cortex, and precuneus – areas showing significant deactivations. Although I report the second set of effects for completion, I focus on the top-down processing effects in the region of interest (vOT) and in POp which showed the same pattern of effects as vOT.

Activations

The most significant effect in the F-map was located in posterior occipito-temporal sulcus at [–44, –54, –12; *Z*=3.5], precisely in the region of the so-called "visual word form area" (Cohen et al., 2000; Cohen et al., 2002; but see Price and Devlin, 2003). Figure 3.4a shows the region and illustrates how its BOLD signal response profile differed across the four conditions. Planned comparisons of vOT responses revealed that, within both tasks, there were significant stimulus effects. In the orthographic lexical decision task, there was greater activation for pseudohomophones than for words (*t*(15)=2.23, *p*=.041) mirroring the RT pattern. In contrast, for phonological lexical decisions the effect sizes went in the opposite direction to the behavioural results, with significantly greater activation for pseudohomophones than pseudowords, (*t*(15)=4.42, *p*< .001). Finally, the direct comparison of the two pseudohomophone conditions revealed significant task-related differences with greater activation in the phonological than the orthographic task (*t*(15)= 2.70, *p*= .017), once again mirroring the RT pattern.

This same pattern of activation was also observed in a region of left POp [–51, +10, +16], although it was only a trend (*Z*=4.3). As in vOT, there was a significantly greater activation for pseudohomophones relative to words in orthographic lexical decisions (*t*(15)=2.92, *p*=.010), significantly more activation for pseudohomophones relative to pseudowords in phonological lexical decisions (*t*(15)=3.01, *p*=.009) and a significantly more activation for pseudohomophones in the phonological task relative to those in the orthographic task (*t*(15)=4.56, *p*<.001). In sum, both vOT and POp showed a similar pattern of activation, consistent with top-down modulation.

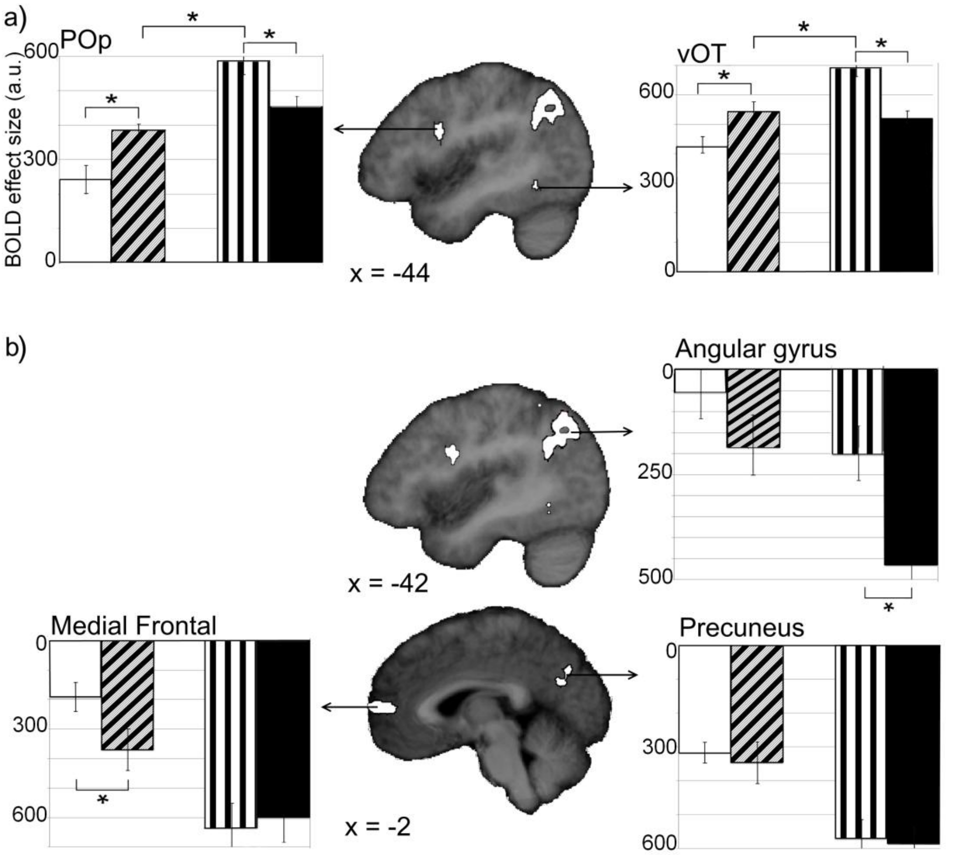

Figure 3.4: Regions whose activations differed across the four conditions. Also shown are bar plots of the BOLD signal per condition relative to fixation in each region: a) The top panel illustrates stimulus- and task-dependent modulation of activation in left ventral occipitotemporal (vOT) cortex and left pars opercularis (POp). The BOLD response profile in these two regions was essentially identical and did not follow the RT profile (Figure 3.2b) and thus could not be explained solely in terms of effort. Note that the opercular activation was not part of common activation seen at the junction of the inferior frontal and precentral sulci because words, unlike the other three conditions, were not significantly activated relative to fixation (*Z*=1.6); b) The bottom panel illustrates significant differences across conditions due to deactivations and are consistent with stimulus- and task-independent responses seen in the default network. Statistical threshold = *p*<.05 (* = significant). Activations are thresholded at *Z*>3.09 and only clusters with significant, or nearly significant, activations are shown (i.e., *Z*>3.2 in the vOT region-of-interest or *Z*>4.0 across the whole brain).

Deactivations

A very different pattern of significant differences across conditions was observed within the left angular gyrus [–42, –65, +47; *Z*=4.9]. Here, all four conditions showed deactivation relative to fixation, and moreover, the magnitude of the deactivation corresponded to the amount of effort required, with the largest effects in conditions showing the longest RTs (Figure 3.4b). The fact that the magnitude of the deactivations was greater in conditions with the longest RTs despite including RTs as a covariate-of-no-interest in the statistical model indicates a non-linear (e.g., higher order) relation between effort and BOLD signal reductions. Two additional areas showing a trend for significant differences across conditions also demonstrated deactivations relative to fixation, namely the medial prefrontal cortex [–2, +63, +8; *Z*=4.3] and the precuneus [–4, –65, +29; *Z*=4.1]. Together these three regions are often considered core components of the “default mode network” (Binder et al., 1999; Greicius et al., 2003; Mazoyer et al., 2001; Raichle et al., 2001; Raichle and Snyder, 2007; Shulman et al., 1997), which is consistent with the deactivations relative to fixation observed here. Indeed, greater deactivation within the default mode network has even been shown to correlate with increasing effort (Lin et al., 2010).

3.4. DISCUSSION

The aim of this study was to investigate whether activation in vOT during commonly used word recognition tasks was influenced by top-down processing of non-visual properties of the visual stimuli. I used words, pseudohomophones and pseudowords in two separate lexical decision tasks in order to manipulate the processing demands on visual and non-visual aspects of the written stimuli. The findings demonstrated that activation in the left vOT (at x= –44, y= –54, z= –12; the precise location of the so-called “visual word form area”) was significantly different across the four conditions and the pattern of activation here could not be predicted by differences in response times. In order to characterise the observed effect, I begin by discussing the stimulus effects within each task and then turn to the task effects seen for pseudohomophones.

Accurate performance on the orthographic task required participants to ignore non-visual properties of the stimulus and focus instead on its specific visual form since all stimuli could be associated with phonological (and semantic) information. Here I found greater activation for pseudohomophones relative to words in vOT (Figure 3.4), replicating previous studies (Bruno et al., 2008; Kronbichler et al., 2007; van der Mark et al., 2009). This finding is difficult to reconcile with a feed-forward account of progressively larger orthographic detectors (Dehaene et al., 2005) because words and pseudohomophones were carefully matched for pre-lexical visual properties such as letter, bigram and trigram frequencies. As previously described, Kronbichler and colleagues (2004) suggested an alternative feed-forward hypothesis in which the visual forms of whole words are stored in vOT, presumably as word detectors analogous to the bigram detectors proposed by Dehaene and colleagues (2005). By this account, pseudohomophones partially activate multiple word detectors yielding greater activation than a single, fully-active word detector (Kronbichler et al., 2004). Although consistent with findings from the orthographic task, this explanation runs into difficulties explaining the results from the phonological task.

The phonological lexical decision task required that unfamiliar visual forms were ignored and instead focused on the phonological (and perhaps semantic) properties of the letter strings. Here I found a significantly greater activation for pseudohomophones relative to pseudowords. Moreover, this activation difference went in the opposite direction to the behavioural difference, effectively ruling out effort as a possible explanation and suggesting the difference had to relate to processing the stimuli themselves. According to Kronbichler and colleagues (2004), both types of stimuli would be expected to partially activate word detectors to similar extents, yielding comparable activation levels for pseudowords and pseudohomophones. Clearly, this was not the case. Instead, pseudohomophones produced significantly greater activation than pseudowords in vOT despite being matched on their orthographic properties. As a result, this finding suggests that the difference in activation was most likely driven by non-visual properties that differentiate the two conditions. Although both are pronounceable and therefore have an associated phonological pattern, these phonological patterns are only familiar for pseudohomophones where they correspond to existing words. Greater vOT activation may reflect the differential cost of integrating these non-visual phonological and semantic properties with their visual forms via feedback projections to vOT. In other words, the finding that non-visual properties modulated activation in vOT demonstrates that this region does more than relay visual information forward to the language system; it interactively integrates bottom-up visual signals with top-down higher order information that is not present in the visual stimuli.

Given the theoretical importance of the finding, it is worth noting that two recent studies have also found greater vOT activation for pseudohomophones relative to pseudowords in a similar task (Bruno et al., 2008; van der Mark et al., 2009). Both studies used a similar phonological lexical decision task (“Does the item sound like a word?”), although their stimuli included real words (“taxi”) in addition to pseudohomophones (“taksi”) and pseudowords (“tazi”). In this design, real words benefit from a familiar orthographic pattern that facilitates “yes” responses relative to pseudohomophones and thus reduces vOT activation, consistent with the claim that lexical visual word forms are stored in the area (Kronbichler et al., 2007). Like the current study, van der Mark and colleagues (2009) reported significantly enhanced vOT activation for pseudohomophones relative to pseudowords which was also present numerically, but not reliably, in the study by Bruno et al. (2008). This effect, however, is difficult to reconcile within a lexical visual word form account (Kronbichler et al., 2007; Kronbichler et al., 2004) without positing some form of feedback from non-visual properties of the stimuli that modulates vOT activation levels.

Finally, in addition to these stimuli-effects, I observed a significant effect of task on vOT activation when the stimuli were held constant, namely greater activation for pseudohomophones during phonological relative to orthographic lexical decisions. This novel finding is at odds with feed-forward accounts that predict either: i) no modulation in activations for pseudohomophones across tasks because the stimuli are the same in both cases or ii) greater activation for orthographic task due to increased orthographic processing demands. Because the stimuli were held constant (i.e., the two tasks used a carefully matched set of pseudohomophones), the change in vOT activation cannot be driven by the stimuli themselves but must instead be a consequence of the different non-visual processing demands required by the two tasks. For instance, this task effect may reflect the additional phonological demands on decoding or assembly which is essential for the phonological task but not for the orthographic task (cf. Dietz et al., 2005; Hellyer et al., 2011). In other words, the increase seen during the phonological lexical decision task is an index of top-down modulation that is consistent with the Interactive Account.

If correct, this hypothesis offers a single, principled explanation for all the current findings and is consistent with previous studies whose results are difficult to explain without an interactive framework (Cai et al., 2010; Devlin et al., 2006a; Kherif et al., 2011). In both the orthographic and phonological tasks, activation for pseudohomophones was greater than for words or pseudowords, respectively, indicating increased processing demands. Presumably, these increased demands are caused by the conflicting visual and non-visual properties of pseudohomophones (Harm and Seidenberg, 2004). Pseudohomophones initially activate semantic information consistent with their phonological form, although this is rapidly suppressed (Harm and Seidenberg, 2004; Lukatela and Turvey, 1994). If vOT plays a role integrating this information, then the top-down semantic signal will conflict with the bottom-up visual information, requiring additional processing to suppress the inappropriate semantic pattern, thus increasing activation for pseudohomophones relative to words or pseudowords where there is no such conflict. In other words, it is precisely the integration of visual and non-visual information that drives the activation observed in vOT. In terms of the Interactive Account, pseudohomophones activate strong top-down predictions from phonological and semantic regions because they are associated with the non-visual properties of their root words (i.e., those of “brain” for “brane”). However, these automatic top-down predictions are incorrect since they do not accurately predict the response to the visual input, resulting in a mismatch thus greater vOT activation. This prediction error is then sent forward until the error is minimised by suppressing the inappropriate top-down prediction and the response is correctly predicted. In contrast, for words, the top-down predictions match the response to the visual input while pseudowords do not elicit any mismatching semantic predictions. Therefore unlike pseudohomophones, neither words nor pseudowords suffer from a conflict between top-down and bottom-up information, which explains the stimulus-driven effects. Furthermore, as I found, such conflict will have a greater effect on pseudohomophones during the phonological task relative to the orthographic task because the phonological task requires greater top-down phonological processing while the orthographic task could be performed on the basis of bottom-up orthographic processing. This interactivity between bottom-up visual information and top-down linguistic codes easily explains why vOT lateralisation follows hemispheric language dominance in individuals (Cai et al., 2010) and can also account for non-visual priming effects observed in vOT (Devlin et al., 2006a; Kherif et al., 2011).

Could the current findings be explained by cognitive feed-forward accounts such as that of Norris et al. (2000)? According to this hypothesis, apparent top-down effects such as word superiority or pseudohomophone effects occur not at the level of processing the stimulus, but rather during the decision making process. Both functional neuroimaging and lesion-deficit studies with neurological patients have consistently associated decision making processes with prefrontal regions (Fleming et al.; Walton et al., 2004; Weller et al., 2007), consistent with the stimulus- and task-driven modulation I observed in POp. This explanation runs into difficulty, however, accounting for the similar pattern of activation observed in vOT, a unimodal sensory area, unless of course it is due to feedback projections from prefrontal regions. In other words, the fact that effects I observed were present in the early perceptual stages of processing is incompatible with a strictly feed-forward explanation based on decision making (Norris et al., 2000).

A clear prediction of the Interactive Account is that for integration to occur in vOT, it should be functionally connected with other components of the cortical language system during reading. Indeed, previous studies have shown intrinsic functional connections linking vOT with Broca's area (Bitan et al., 2005; Mechelli et al., 2005). Furthermore, recent studies investigating resting-state functional connectivity suggest that a strong intrinsic connectivity exists between Broca's area and ventral occipito-temporal regions even during rest (Koyama et al., 2010; Smith et al., 2009). Thus it was of considerable interest that the activation pattern in POp, a core region of Broca's area, matched that in vOT. This is also consistent with Hellyer et al. (2011) in which increased activation in vOT during phonological relative to numerical task with orthographical and digital forms of numbers was accompanied by an increase in activation in inferior frontal gyrus. These observations suggest a possible functional linkage between these regions that may contribute to top-down influence on vOT. Confirmation will require evidence of effective connectivity that demonstrates top-down modulation of vOT activity by Broca's area.

Taken together, the current findings demonstrate that activation in vOT during reading is influenced by non-visual properties of written stimuli and emphasise that interactivity is as important for neural accounts as it is for cognitive and computational models (Coltheart et al., 2001; Harm and Seidenberg, 2004; Jacobs et al., 2003; McClelland and Rumelhart, 1981; Perry et al., 2007; Plaut et al., 1996; Rumelhart and McClelland, 1982). It is worth noting that this conclusion is not specific to reading but rather is in line with a growing literature demonstrating that visual object recognition or face processing cannot be a hierarchical, feed-forward process either (Bar et al., 2006; Cohen Kadosh et al., 2010; Gazzaley et al., 2007; Gilaie-Dotan et al., 2009; Kveraga et al., 2007; Schrader et al., 2009). These studies challenge the traditional view of serial, bottom-up visual object or face recognition and instead support non-hierarchical mechanisms which integrate top-down feedback to influence recognition process (see also Bar, 2003; Bullier, 2001; Bullier and Nowak, 1995). Together these studies highlight a need to focus not only on the nature of neuronal representations, but also on the dynamics of this information processing. Critically, this involves elucidating both the functional and anatomical connectivity, which will hopefully help to close the gap between cognitive and neuroanatomical models of reading.

4. LEXICAL FREQUENCY AND VISUAL FAMILIARITY

A version of this chapter was published as “Dissociating visual form from lexical frequency using Japanese” by Tae Twomey, Keith J. Kawabata Duncan, John S. Hogan, Kenji Morita, Kazumasa Umeda, Katsuyuki Sakai & Joseph T. Devlin in *Brain & Language* 125: 184-193 (2013).

The data collected for this experiment were used in a DCM study and published as "Inter-and Intrahemispheric Connectivity Differences When Reading Japanese Kanji and Hiragana" by Keith J. Kawabata Duncan, Tae Twomey, ‘Ōiwi Parker Jones, Mohamed L. Seghier, Tomoki Haji, Katsuyuki Sakai, Cathy J. Price, and Joseph T. Devlin in *Cerebral Cortex* (2013).

4.1. INTRODUCTION

Of all the world's languages, Japanese uniquely relies on multiple written scripts for its everyday use. Books, magazines, and advertisements all mix morphographic[2] Kanji with syllabographic Hiragana such that no adult text consists solely of one script (although some children's books are written only in Hiragana). As a result, Japanese adults are equally familiar with both scripts and many individual words can be written multiple ways. For instance, a word such as "apple" is as common in Hiragana (りんご) as Kanji (林檎). In many cases, however, one form will be more common than the other. For instance, "mischief" is usually written in Hiragana (わんぱく) but sometimes occurs as Kanji (腕白). The fact that the same word can be written in different scripts means that Japanese offers a unique opportunity to disentangle the frequency of a written word (i.e., its lexical frequency) from the frequency of its visual form (i.e., its visual familiarity). Lexical frequency is a measure of how frequently a lexical item is encountered. Thus, it is the frequency of the concept of a word and a property of the whole word, regardless of its form. In contrast, visual familiarity is specific to the visual form of a word and the familiarity of its sublexical constituents (e.g., individual letters or orthographic neighbours) may also contribute to this factor. In alphabetic languages, however, written lexical frequency and visual familiarity are essentially the same thing – both measure how frequently a word appears in print. Consequently, much of the literature focuses on lexical frequency as a key factor in understanding the nature of the neural information processing in brain regions related to reading (Fiebach et al., 2002; Hauk et al., 2008) and/or these two terms are used interchangeably (e.g., Kronbichler et al., 2007), despite the

---

[2] Although Japanese Kanji and Chinese Hanzi characters are visually identical, I use the term "morphographic" when referring to Kanji characters because most characters function as grammatical morphemes. I use the term "logographic" for Chinese characters since this is the term widely used in the literature, even though the characters too are largely morphographic (Rogers, 2005). In contrast, Hiragana characters refer to individual mora, a phonological timing unit that largely corresponds to a syllable in Japanese.

possibility that visual familiarity and lexical frequency may have differential effects.

This is particularly relevant to theories of the left ventral occipito-temporal cortex (vOT). As previously noted, vOT activation is greater for low than high frequency words (Chee et al., 2003b; Hauk et al., 2008; Joubert et al., 2004; Kronbichler et al., 2004) and this effect is not limited to alphabetic languages but is also seen in non-alphabetic Chinese (Kuo et al., 2003; Lee et al., 2004). These lexical frequency effects are central to the orthographic input lexicon account (Bruno et al., 2008; Glezer et al., 2009; Kronbichler et al., 2007; Kronbichler et al., 2004; van der Mark et al., 2009). As described in the previous chapter, this account claims that this area stores orthographic word representations since lexical frequency inversely modulates vOT response. Like logogens (Morton, 1969), entries in this orthographic input lexicon are sensitive to experience, with access to less frequent words requiring greater effort and therefore resulting in greater activation. An alternate explanation for frequency effects in vOT offered by the Interactive Account is that they arise from the interaction of bottom-up and top-down constraints. Reciprocal connections with higher order association areas link visual representations such as words, objects, scenes and faces with non-visual properties of the stimulus such as its sound (phonology) or meaning (semantics). Non-visual information is sent to vOT as top-down predictions in order to predict the response to the visual stimuli. High frequency written words are more familiar visual patterns and thus have more accurate top-down predictions into vOT reducing prediction error (the mismatch between the response and the predictions) and therefore activation. In contrast, low frequency words result in greater prediction error, increasing the processing demands on vOT and thereby increasing the activation (Price and Devlin, 2011).

The aim of the current study was to test this Interactive Account using Japanese to differentiate between the frequency of a word and its visual familiarity. This distinction is not possible in most languages as individual words can only be correctly written one way. For instance, in alphabetic languages like English, it is possible to write a word differently (e.g., “brane” for “brain”) but literate readers immediately recognise these as incorrectly spelled which is very different from seeing two different forms of a correctly spelled word. Another possibility would be to test bilinguals with the same word in different languages (e.g., “米” in Japanese and “rice” in English) but again, this is not optimal since the difference in script is confounded by a difference in language. In this example, the semantic properties of “米” and “rice” are not identical because unlike the English word, the Japanese word means “uncooked rice” and there is a separate word for “cooked rice”. In other words, the “same” words in two different languages often are subtly different, confounding differences in their visual form. Therefore, only those languages in which a word can be correctly written in multiple forms suffice for dissociating visual familiarity from lexical frequency. Although Chinese (Hanzi vs. Pinyin) and Korean (Hanja vs. Hangul) offer this possibility, in these languages only one form is in daily use, making one script much more familiar than the other. Japanese, on the other hand, is unique in its reliance on multiple forms for everyday use and therefore provides a fertile ground for testing theories of vOT function in reading.

According to the Interactive Account, visual familiarity is expected to strongly modulate activation in vOT, consistent with the hypothesis that the region plays a more domain-general role in representing visual patterns, of which written words are only one example. Lexical frequency, on the other hand, is predicted to interact with visual familiarity such that only low frequency words are affected by visual familiarity. For highly frequent words, vOT will receive sufficiently accurate top-down predictions to quickly and accurately match the bottom-up visual information regardless of whether the visual form is more or less familiar. In contrast, less frequent words will send less accurate top-down predictions to vOT, resulting in greater prediction error and increased processing demands.

Consequently, activation in vOT for low frequency words will benefit from greater visual familiarity.

Another interesting aspect of Japanese visual word recognition that can be tested is the script differences between Kanji and Hiragana. Previous studies found greater vOT activation for Kanji relative to Hiragana during various tasks such as reading aloud (Sakurai et al., 2000), size judgement (Ha Duy Thuy et al., 2004), lexical decision (Ha Duy Thuy et al., 2004), semantic decision (Nakamura et al., 2005a) and real word recognition (judging which one of the two stimuli is the real word, Ino et al., 2009). These effects are in general interpreted as the consequence of visual or semantic differences between the two scripts. Kanji and Hiragana characters are visually quite different as most Kanji characters are composed of many strokes, making them visually more complex than Hiragana words. Thus, one of the possibilities of greater vOT activation for Kanji relative to Hiragana is the greater demands on visuospatial analysis required for Kanji (Dehaene, personal communication, Ino et al., 2009). This visual feature explanation is plausible given that the vOT is located in the extrastriate visual cortex and consistent with recent studies on spatial frequency effects in occipitotemporal cortex (Mercure et al., 2008; Woodhead et al., 2011b). For instance, Mercure and colleagues (2008) investigated the lateralisation of the N170 event-related potential (ERP) component, an index of complex visual processing, in bilateral occipitotemporal regions using fill-spectrum and spatially (high- or low-pass) filtered pictures of faces and words (dominant spatial frequency: words > faces). They found that the amplitude of N170 for words greater in the left than the right hemisphere; and within the left hemisphere, N170 was larger for words than faces. In addition, N170 lateralisation was attenuated with low-pass filtered stimuli demonstrating that the left lateralisation of words was influenced by spatial frequency. Similarly, Woodhead and colleagues (2011b) compared fMRI activation for words, faces and scrambled words and faces in order to test the effect of spatial frequency in bilateral occipitotemporal regions. The scrambled stimuli were created with sine-wave gratings so that no higher-order information was left in the stimuli that might bias activations. Their analyses showed that the left occipitotemporal region was preferentially activated by high more than low spatial frequencies, even without any structure or semantic information left in the stimuli (Woodhead et al., 2011b). Both of these studies are compatible with the visual complexity explanation of the stronger vOT activation for Kanji relative to Hiragana. At the same time however, it has also been reported that visual complexity did not modulate vOT activation during a linguistic task. Hellyer and colleagues (2011) showed that vOT activation across orthographical and numerical forms of the same numbers (e.g., “twenty” vs. “20) were equal during a phoneme detection but not numerical task (to decide whether a number is odd or even), despite the fact that the former was visually more complex than the latter. While the authors acknowledge that this effect may be driven by the difference in the task difficulty, they also offer another plausible explanation that the top-down modulation from the phonological system is exerted only during the phoneme detection task. This interpretation is indeed supported by their finding that left posterior frontal cortex, a region implicated in phonological processing, was also strongly activated preferentially during the phoneme detection task. These findings question whether the stronger vOT activation for Kanji relative to Hiragana during linguistic tasks is purely driven by the visual complexity difference. Instead, all these studies suggest that, although there may be a perceptual preference for Kanji in vOT given its higher spatial frequency than Hiragana, this particular difference may be attenuated during linguistic tasks due to the top-down influence from the higher-order language systems. Thus, it is possible that the observed script difference in vOT activation during linguistic tasks may not reflect the spatial frequency differences, but in fact the differences in their higher-order properties (see below), or the combination of the two (cf. Mercure et al., 2008).

Indeed, Kanji and Hiragana strongly differ in terms of their higher-order properties. Single Kanji characters carry some semantic information and some even have multiple meanings whereas Hiragana has no semantic values whatsoever. In many cases the meaning of a Kanji word can be inferred from the meaning of the individual Kanji characters in the word. The word for "advertisement" (広告) for instance, contains a character that means "wide" (広) and "inform" (告). However, at times the meaning of the individual Kanji characters in a Kanji word does not correspond to the meaning of the whole word. For instance, in the word for "mischief" (腕白), the first character signifies "arm" and the second, "white" when they occur in isolation. Therefore, in such Kanji words, the correct meaning cannot be computed but must be learnt via association, ignoring the meaning of each character. In other words, semantic information of each Kanji character in Kanji words can either aid or hinder inferring or learning the correct meaning when one encounters a novel or less frequent Kanji word. In highly familiar Kanji words, the meaning of the individual characters is not explicitly required. However, reading Kanji words may prompt automatic access to all the available semantic information within the words. Therefore many researchers suggest that the greater vOT activation for Kanji relative to Hiragana is due to the increased reliance on, the greater activation induced by or the interference from the lexico-semantic system via the ventral route, in which vOT is located (Ha Duy Thuy et al., 2004; Ino et al., 2009; Nakamura et al., 2005a; Sakurai et al., 2000).

Yet another possible explanation for greater vOT activation for Kanji relative to Hiragana is the differences in phonology between the two scripts. Mapping between Kanji characters and their phonology is inconsistent and opaque. Each Kanji character is composed of strokes and although some of them can be segmented into separate radicals that may provide phonological or semantic cues, the pronunciations of Kanji characters are arbitrary and cannot be computed from its constituents. Therefore they must be memorised. In addition, most Kanji characters have multiple pronunciations and the correct pronunciation is dependent on the context in which they appear. In contrast, the pronunciations of Hiragana characters are almost completely regular, transparent and independent of the context. Therefore, these orthographic transparency differences could lead to greater vOT activation for Kanji (see below).

Needless to say, the semantic and phonology explanations are not mutually exclusive. Within the Interactive Account framework, both of these semantic and phonological effects would manifest as differences in the top-down signals of higher-order, non-visual properties. Therefore, consistent with previous studies, the Interactive Account makes one further prediction that vOT activation will be greater for Kanji relative to Hiragana. Although the current design cannot separate the two effects, greater prediction error for Kanji is expected due to: i) the inconsistent mapping between Kanji characters and their phonology as opposed to Hiragana words where there is a consistent mapping between characters and syllables, ii) the greater associations of Kanji characters and their semantics than Hiragana or iii) the combination of the two. Consequently, the magnitude of the BOLD signal is expected to be greater for Kanji than for Hiragana.

The current experimental design aimed to evaluate these hypotheses. Participants performed a lexical decision task where they decided whether visual stimuli represented real Japanese words. Stimuli were written in either Kanji or Hiragana and fully crossed with visual familiarity. In other words, each word appeared twice in the course of the experiment, once in Kanji and once in Hiragana. Half of the words were more commonly written in Kanji while the other half were more common in Hiragana. This design allowed me to look for main effects of visual familiarity (high vs. low) and script (Kanji vs. Hiragana) as well as their potential interaction. A second analysis recoded all the words into four sets divided according to lexical frequency (i.e., independent of script) and looked for frequency effects and their interaction with visual familiarity. In this fashion, I could independently evaluate the effects of visual familiarity and lexical frequency on vOT as well as look for potential script differences.

4.2. MATERIAL AND METHODS

Participants

40 native Japanese speakers participated in this study although the data from six were excluded due to either excessive motion inside the scanner (i.e., motion greater than the dimensions of a voxel; n=3 participants) or due to poor performance (i.e., accuracy less than 60% in one or more conditions; n=3 participants). Consequently, only data from 34 participants (13M, 21F, aged 21-62) were included in the final analyses. Since it was hypothesised that the amount of exposure to written Japanese may affect the activation, two groups of native Japanese readers were tested: university students in Tokyo with daily exposure to written Japanese (n=15, 10M, 5F, aged 21-31) and Japanese ex-patriots who had lived outside Japan for a minimum of three years and thus had reduced exposure to written Japanese in their daily lives (n=19, 3M, 16F, 29-62).[3] The imaging analyses, however, revealed no significant interactions between Group (Tokyo vs. London) and any other factor. As a result, the results presented here collapse over Group despite including it as a factor in the analyses to better model structured variance in the data.

All participants were native Japanese speakers born and educated in Japan through to at least secondary school. Consequently, all were literate adult readers in Japanese familiar with both Kanji and Hiragana. In addition, all were right-handed except for one who was confirmed to be ambidextrous according to the Edinburgh Handedness Inventory (Oldfield, 1971). None reported a history of reading difficulties or neurological problems. Each of the London participants had lived outside of Japan, China and Korea (where morphographic or logographic scripts are used) for at least for three years (range: 3-34 years, mean=11). Testing

---

[3] Subject recruitment and data collection in Tokyo were carried out by Keith Kawabata Duncan.

in Tokyo was approved by the ethics committees of the Graduate School of Medicine, the University of Tokyo (#2968), and the ethics committee of the Brain Science Institute, Tamagawa University (C21-4). In London, ethical approved was granted by the NHS Berkshire Research Ethics Committee (06/Q1602/20).

Experimental procedures

The participants' task was to view strings of characters and decide whether the string formed an existing Japanese word or not. The task involved 60 words, each of which was presented twice – once in Kanji and once in Hiragana. Half of the words are most commonly written in Kanji and the other half are most commonly written in Hiragana. An equal number of nonwords, divided evenly between Kanji and Hiragana, were included to ensure adequate task performance.

A trial began with a fixation cross presented for 500msec. A stimulus (written horizontally using the MS Gothic font) was then presented for 500msec, followed by a jittered inter-stimulus interval of 1–4 sec (mean = 2.5 sec). Therefore, the average trial length was 3.5 sec. Stimuli were presented in blocks of 15 trials (lasting 57 sec) in a pseudorandomised order which included both "yes" and "no" responses. These were separated by 12 sec blocks of fixation that served as an implicit baseline. Over a run, there were eight blocks of task and eight blocks of rest. Therefore, each run lasted 9 minutes and 21 seconds. There were two runs. Responses were made with a button press, using either the index or middle finger of their right hand to indicate "yes" and "no." The response fingers were fully counter-balanced across participants. The stimuli were projected onto a screen and viewed via mirrors attached to the head coil. Participants practised each task inside the scanner before the main runs began. No items that were used in the practice runs occurred during the main experiment.

Stimuli

The word stimuli were obtained from the NTT Japanese Psycholinguistic Database (Amano and Kondo, 2003a, b) by identifying 30 words that had a higher visual familiarity score when written in Kanji than any other script (including Hiragana, Katakana and mixed scripts). Another set of 30 words was found that had a higher visual familiarity score when written in Hiragana than any other script. The Kanji words were then transliterated into Hiragana (“展望”→” てんぼう”) and the Hiragana words transliterated into Kanji (“とんち”→” 頓知”) producing 120 words (60 written in Kanji, 60 written in Hiragana). It is important to note that while this transcription changes the visual familiarity of the word, the lexical frequency of the word remains constant. All words were 2 or 3 characters in length when written in Kanji, and between 2 and 6 when written in Hiragana. The resulting word set had four different conditions, each with 30 items (high visual familiarity Kanji words, low visual familiarity Kanji words, high visual familiarity Hiragana words, and low visual familiarity Hiragana words) corresponding to a 2 × 2 factorial design with Script (Kanji, Hiragana) and Visual Familiarity (High, Low) as factors.

The stimuli were carefully matched along several different dimensions summarised in Table 4.1. These values, from the NTT database (Amano and Kondo, 2003a, b), were analysed with a 2 × 2 ANOVA. Across conditions words were matched for mora length (a measure of phonological complexity). In addition, the analysis confirmed the main effect of Visual Familiarity and demonstrated that this did not interact with Script. Conceptual familiarity (derived by summing familiarity ratings for the visual word and for its auditory form) was matched across Script but naturally it was not possible to match across Visual Familiarity. Finally, Hiragana words had significantly more characters but fewer strokes than Kanji words, which is an inevitable difference between the scripts. It is worth noting, however, that effects of visual complexity and word length are expected to manifest in early visual cortices (Hsu et al., 2011; Mechelli et al., 2000; Tarkiainen et al., 1999) rather than in higher order visual regions like vOT.

Table 4.1: Mean psycholinguistic properties per condition for the visual familiarity analyses

	Kanji		Hiragana		Effect of		
	High	Low	High	Low	Script	VisFam	Interaction
Mora length*	2.77	3.03	3.03	2.77	*t=1.46 (p=.145)*		
Visual familiarity	5.31	3.66	5.21	4.09	0.9 (n.s)	69.5 (p<.001)	2.3 (n.s.)
Combined familiarity	5.36	3.91	5.25	4.06	0.0 (n.s.)	64.5 (p<.001)	0.5 (n.s.)
Number of characters	2.00	2.07	3.07	2.90	77.7 (p<.001)	0.2 (n.s.)	1.2 (n.s.)
Number of strokes	19.17	19.70	7.60	6.63	217.0 (p<.001)	0.8 (n.s.)	0.8 (n.s.)

***Because the same words were used in the Kanji High and Hiragana Low conditions as well as in the Kanji Low and Hiragana High conditions, script-independent measures were the same for these conditions. Consequently, differences were assessed with a t-test rather than an ANOVA and are marked in italics.**

The same stimuli were then regrouped according to their lexical frequency – a script-independent measure of how often the word occurs in print regardless of its visual form (i.e., Kanji or Hiragana). Lexical frequency values were calculated by summing the frequency of the Kanji and Hiragana word forms, taken from the NTT database (Amano and Kondo, 2003a, b). For example, the Japanese word pronounced /tembo:/ has a lexical frequency value of 6984, since its written Kanji form ("展望") has a frequency of 6979 and its Hiragana form ("てんぼう") has a frequency value of 5. The frequencies for the 60 lexical items were then divided into quartiles so that those within the lower quartile (i.e., low frequency words) could be compared to those within the upper quartiles (i.e., high frequency words) in order to maximise the distinction between them. Because the stimuli were originally chosen according to their visual familiarity scores across scripts, the distribution of the (log of the) lexical frequencies was nearly uniform over the 60 words. Note that it proved impossible to fully balance visual familiarity, script and lexical frequency into a factorial design, forcing the data to be interrogated in two separate analyses. In order to separate visual familiarity into high and low, each lexical item needed to be presented in both scripts. In contrast, lexical frequency was independent of script. As a result, a full factorial design that included Visual Familiarity, Lexical Frequency and Script was impossible to generate. There were no significant differences between high and low frequency items in terms of number of mora, number of characters, or total stroke count (Table 4.2)

Table 4.2: Mean psycholinguistic properties for the lexical frequency analysis

	Upper	Lower	t
Mora length	2.8	3.1	1.8 (n.s.)
Lexical frequency	12100	25	4.5 (p<.001)
Visual familiarity	4.9	4.1	2.6 (p=.011)
Number of characters	2.5	2.6	0.7 (n.s.)
Number of strokes	14.0	13.1	0.4 (n.s.)

Finally, Kanji nonwords were created by combining random Kanji characters that together did not form a word. These were matched 1:1 with the real Kanji words for number of strokes and characters. Hiragana nonwords were created by combining random Hiragana characters that together did not form a word. These were matched 1:1 with the real Hiragana words for number of strokes and characters.

MRI acquisition

For the subjects scanned in Tokyo, whole-brain imaging was performed on a Siemens 3T MRI scanner at the Brain Science Research Center at Tamagawa University. The functional data were acquired with a gradient-echo EPI sequence (TR = 3000msec; TE = 25msec; FOV = 192mm; matrix = 64 × 64) giving a notional resolution of 3 × 3 × 3. For participants in London, whole brain imaging was performed on a Siemens 1.5T MRI scanner at the Birkbeck-UCL Neuroimaging (BUCNI) centre. The functional data were acquired with a gradient-echo EPI sequence (TR = 3000msec; TE = 50msec; FOV = 192mm; matrix = 64 × 64) giving a notional resolution of 3 × 3 × 3. In both cases, a run consisted of 187 volumes and as a result the two runs together took 18 minutes 42 seconds. In addition, a high-resolution ($1mm^3$) T1-weighted anatomical scan was acquired for localizing the functional data on the individual's brain anatomy.

Analyses

In the both the behavioural and imaging data, items whose accuracy was at chance (≤ 50%) were excluded from all analyses (n = 9) and only correct trials were analysed. Reaction times (RTs) were recorded from the onset of the stimulus and anticipatory responses (i.e., RTs < 300ms) were trimmed (0.05% of trials). To minimise the effect of outliers, median RTs per condition per subject were used in the statistical analyses (Ulrich and Miller, 1994). The behavioural data were analysed using a mixed 2 × 2 × 2 analysis of variance (ANOVA) with Script (Kanji, Hiragana), Visual Familiarity (High, Low) as within-subject factors and Group (Tokyo, London) as a between-subject factor. Accuracy and RTs were the dependent measures. In addition, the behavioural data were then re-grouped into quartiles according to lexical frequency of the stimuli and analysed using a repeated-measures 4 × 2 × 2 ANOVA with Lexical Frequency (Upper, Upper Middle, Lower Middle, Lower) and Visual Familiarity (High, Low) as within-subject factors and Group (Tokyo, London) as a between-subject factor.

The imaging data were processed using SPM8 (Wellcome Trust Centre for Neuroimaging, London UK, http://www.fil.ion.ucl.ac.uk/spm/). The first four volumes in the Tokyo (i.e., 3T) data and two volumes in London (i.e., 1.5T) data were discarded in order to allow for T1 equilibrium. All functional volumes were spatially realigned and unwarped to adjust for minor distortions in the B0 field due to head movement (Andersson et al., 2001). They were then normalised to the MNI-152 EPI template, maintaining the original 3 × 3 × 3mm resolution. Finally, images were smoothed with an isotropic 8 mm full-width half-maximum Gaussian kernel. Time-series from each voxel were high-pass filtered (1/128 Hz cut-off) to remove low-frequency noise and signal drift. The preprocessed functional volumes were then analysed in two separate GLMs. One investigated the effects of visual familiarity and script while the other investigated the effect of lexical frequency and visual familiarity. In both cases, a first-level, fixed-effects analysis combined the two runs from each participant and the estimated effect sizes were entered into a second-level, random-effects analysis to estimate the population effect. At the first level, the onsets of stimuli were modelled as delta functions convolved with a canonical haemodynamic response function (Glover, 1999), which provided regressors for the general linear model. The appropriate contrast images, averaged over sessions, were then generated in all subjects for each condition versus fixation.

The first analysis included four word conditions (Kanji high visual familiarity, Kanji low visual familiarity, Hiragana high visual familiarity and Hiragana low visual familiarity), two nonwords conditions (Kanji, Hiragana) and a condition for incorrect and excluded trials (Murphy and Garavan, 2004). Fixation was not modelled and served as an implicit baseline. The four word-relative-to-rest contrasts were computed and entered into a second-level, 2 × 2 × 2 ANOVA with Script (Kanji, Hiragana), Visual Familiarity (High, Low) as within-subject factors and Group (Tokyo, London) as a between-subject factor. I first identified areas of common activations for all word conditions using a linear contrast to compute their mean activity and inclusively masking it with each condition relative to fixation at $p = 0.001$. From this analysis I computed statistical contrasts of the two conditions within a factor, inclusively masking them with common activations of these conditions at $p = 0.05$.

The second analysis investigated the effect of lexical frequency independent of script. Here, words were divided into quartiles based on their frequency and entered into a second-level, 4 × 2 × 2 ANOVA with Lexical Frequency (Upper, Upper middle, Lower middle, Lower) and Visual Familiarity (High, Low) as within-subject factors and Group (Tokyo, London) as a between-subject factor.

Since the primary aim of the current study was to investigate effects in vOT, an *a priori* anatomical mask for this region-of-interest (ROI) was defined in the same fashion as in Chapter 3. The main anatomical areas of interest were, as before, the occipitotemporal sulcus and adjacent regions on the crests of the fusiform and inferior temporal gyri: areas consistently activated by visual word recognition tasks (Bitan et al., 2007; Cai et al., 2010; Cohen et al., 2000; Devlin et al., 2006a; Duncan et al., 2009; Fiez and Petersen, 1998; Frost et al., 2005; Herbster et al., 1997; Kronbichler et al., 2007; Price et al., 1996; Rumsey et al., 1997; Shaywitz et al., 2004). However, the precise coordinates along a rostro-caudal axis differed slightly from the anatomical masks defined in Chapter 3 due to the anatomical differences between the Caucasian and Japanese head shapes. The standard space coordinates used to delineate this region ranged from X= −36 to −54, Y= −45 to −66 and Z= −30 to −6. As in the previous chapter, these coordinates together described a rectangular prism that conservatively encompassed the region of vOT sensitive to visual word recognition as well as the anatomically adjacent lobule VI of the cerebellum. Because the cerebellum was both anatomically and functionally distinct, it was manually removed from the ROI mask.

For all imaging analyses, activations were considered significant based on voxel-level inference of $p < 0.05$, corrected for multiple comparisons either within the ROI ($Z > 3.30$) or across the entire brain ($Z > 4.60$). In order to visualise the pattern of activation within a region, I plotted the mean effect size per condition within a 5mm-radius sphere centred on the peak coordinate. No inferential statistics were based on these effect size plots.

4.3. RESULTS

Behavioural results

The behavioural results are illustrated in Figure 4.1, where the left panel displays accuracy scores and the right reaction times. Within the accuracy data, there was a significant main effect of Visual Familiarity ($F(1,33)=32.8$, $p < 0.001$), confirming that less visually familiar word forms were more difficult. This was qualified by a significant Visual Familiarity × Script interaction ($F(1,33)=26.6$, $p < 0.001$), indicating that the visual familiarity advantage was significant for Kanji ($t(33)=6.6$, $p < 0.001$) but not for Hiragana ($t(33) = 0.9$, $p = 0.365$, n.s.). In addition, there was a significant main effect of Script ($F(1,33)=14.5$, $p = 0.001$) indicating responses to Hiragana were more accurate than to Kanji. The analysis of the reaction time data revealed a similar pattern of results. There was a main effect of Visual Familiarity ($F(1,33)=37.5$, $p < 0.001$) with responses to less visually familiar forms taking longer than those to highly familiar forms (854 vs. 775ms). This was qualified by a significant interaction ($F(1,33)=10.1$, $p=0.003$) indicating that familiarity effect was larger for Kanji (118ms) than Hiragana (39ms). The main effect of Script was not significant ($F(1,33)=2.7$, $p = 0.106$).

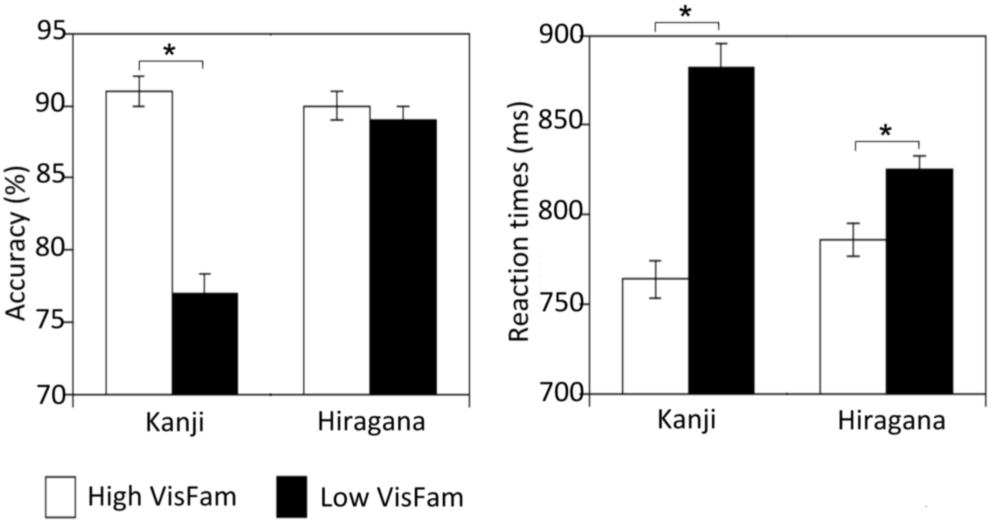

Figure 4.1: Mean accuracy (left panel) and reaction times (right panel) for four conditions. An asterisk indicates significant at $p = 0.05$ (corrected for multiple comparisons).

The second behavioural analysis focused on lexical frequency and visual familiarity and the data are shown in Figure 4.2. There was a main effect of Lexical Frequency for both accuracy ($F(3,99)=35.9$, $p < 0.001$) and RTs ($F(3,99)=24.4$, $p < 0.001$). From the figure it is clear that the lower the lexical frequency, the more difficult the word was with both lower accuracy and longer response times. There was also a main effect of Visual Familiarity for both accuracy ($F(1,33)=42.5$, $p < 0.001$) and RTs ($F(1,33)=88.9$, $p < 0.001$). In addition, there was a significant interaction for RTs ($F(3,99)=5.5$, $p = 0.002$), indicating that visual familiarity had a greater effect on low relative to high frequency items although the interaction for accuracy was not significant ($F(3,99)=1.4$, $p = 0.257$). This pattern of results remained exactly the same when only the Upper and Lower frequency quartiles were included in the ANOVA. In other words, both lexical frequency and visual familiarity strongly affected overall performance of the task. The question then becomes: to what extent do these two factors affect activation in vOT?

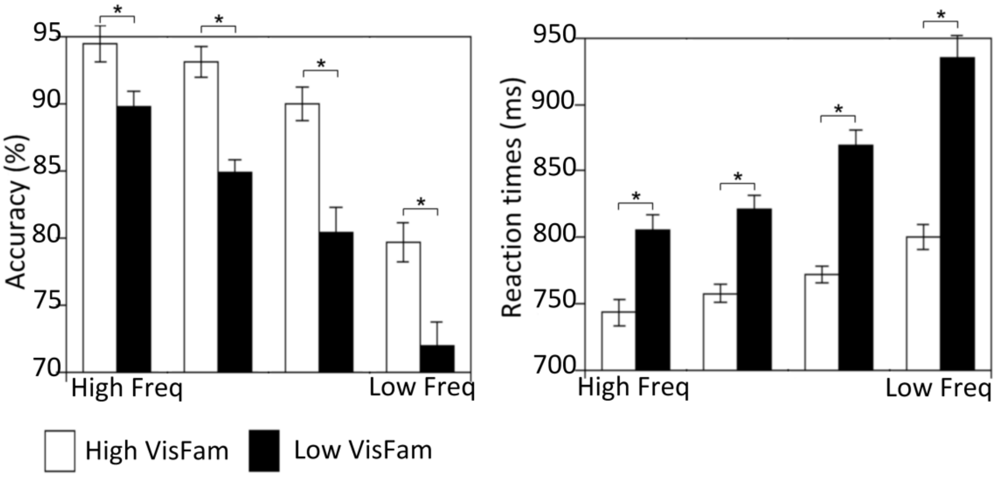

Figure 4.2: Mean accuracy (left panel) and reaction times (right panel) for the Lexical Frequency × Visual Familiarity analysis. An asterisk indicates significant at $p = 0.05$ (corrected for multiple comparisons).

Imaging results

The first imaging analysis identified brain regions commonly activated by all four word conditions, in order to determine whether vOT (among other regions) was engaged by both Kanji and Hiragana words, independent of their visual familiarity. As expected, vOT was strongly activated bilaterally, centred on the posterior occipitotemporal sulcus and extending laterally into inferior temporal gyrus, medially onto the crest of the fusiform gyrus and inferiorly into lobule VI of the cerebellum. Other bilateral activations included pars opercularis, the pre-supplemental motor area (pre-SMA), the intraparietal sulcus, a mid-cingulate region, and parts of the basal ganglia. In addition, there were several activations only seen in the left hemisphere including Broca's complex (i.e., pars triangularis, pars orbitalis), the deep frontal operculum, the supramarginal gyrus, and a small cluster in the anterior fusiform gyrus (see Table 4.3 for full details). These results are consistent with previous visual word recognition studies conducted in alphabetic (Carreiras et al., 2007; Devlin et al., 2006a; Fiebach et al., 2007; Hauk et al., 2008) and logographic (Booth et al., 2006; Chen et al., 2002; Fu et al., 2002; Hu et al., 2010; Kuo et al., 2003; Tan et al., 2001) languages, indicating a common system engaged by visual word processing, independent of script.

Table 4.3: Common activation across the four conditions relative to fixation. For each region, the standard space (MNI) coordinates of the peak voxel and the *Z*-score for the main effect of words relative to rest at that voxel are shown. In addition, the final four columns display the *Z*-score for each condition relative to rest at the same coordinate.

Region		**Mean peak coordinate**				***Z* relative to rest**			
		x	y	z	*Z*-score	Kanji Hi	Kanji Lo	Hira Hi	Hira Lo
Frontal									
L	pars opercularis	–42	8	31	8.19	6.55	7.99	7.80	7.96
R	pars opercularis	45	5	31	6.84	4.22	6.49	5.94	5.18
L	pars orbitalis	–39	35	1	6.92	3.19	5.87	6.23	6.88
L	pars triangularis	–48	29	19	7.91	5.85	7.41	7.52	7.86
L	frontal operculum	–33	20	1	7.76	4.29	6.74	6.38	7.25
L	pre-SMA	–6	14	52	7.86	6.41	7.35	7.35	7.69
R	pre-SMA	6	14	52	6.39	3.59	5.27	5.11	6.57
R	cingulate gyrus	6	–7	28	5.41	3.73	4.41	3.96	5.14
L	cingulate gyrus	–6	–7	28	5.24	3.39	4.58	3.96	4.76
Parietal									
R	IPS	33	–58	43	6.73	4.61	5.70	5.57	5.83
L	IPS	–30	–58	43	7.84	6.31	7.43	7.20	7.63
L	SMG	–47	–37	49	7.63	5.81	6.38	7.19	7.67
Occipital/Temporal									
L	vOT	–45	–64	–14	8.46	8.12	7.99	7.88	8.17
R	vOT	44	–55	–20	7.42	7.09	6.51	5.65	6.76
L	ITG	–60	–43	–14	5.56	4.04	4.47	5.84	5.00
R	ITG	63	–40	–14	4.82	3.41	3.41	4.03	4.61
L	anterior fusiform gyrus	–39	–28	–26	5.62	5.17	5.21	3.44	3.78
Subcortical									
R	cerebellum (VI)	33	–55	–26	8.17	7.81	7.33	7.58	7.98
L	cerebellum (VI)	–39	–52	–26	8.01	7.67	7.46	7.27	7.48
R	globus pallidus	12	2	–2	5.40	4.16	4.32	4.37	4.41
L	Putamen	–21	5	4	7.69	5.49	6.02	6.41	7.01
R	Putamen	21	5	–2	5.06	3.37	3.55	4.31	5.05
L	Thalamus	–9	–13	7	6.76	3.89	5.79	5.77	6.35
R	caudate nucleus (body)	21	–10	22	5.16	3.36	5.10	3.19	4.63

L	caudate nucleus (head)	–15	8	13	5.50	3.53	4.46	4.02	5.51
R	caudate nucleus (head)	15	8	13	4.91	3.77	3.62	3.29	4.99

Abbreviations: IPS = intra-parietal sulcus; ITG = inferior temporal gyrus; SMA = supplemental motor area; SMG = supramarginal gyrus; vOT = ventral occipito-temporal cortex.

Next, I asked whether visual familiarity modulated vOT activation. The comparison of low relative to high visual familiarity items revealed significant left vOT activation in the ROI [–45 –58 –11, *Z*=3.34, *p*=0.042] and this did not interact with Script (Figure 4.3a & b). Outside of the ROI, the whole brain search revealed activation in left pars triangularis, left pars opercularis, and in the frontal operculum bilaterally (Table 4.4) – three regions previously associated with low relative to high frequency effects in alphabetic languages (Carreiras et al., 2006; Fiebach et al., 2002; Kronbichler et al., 2007; Kronbichler et al., 2004). In each of these regions, there was a (non-significant) interaction with Script such that low visual familiarity items increased activation for Kanji more than for Hiragana relative to high visual familiarity items. No regions showed significant activation for the contrast of high relative to low visual familiarity, even when the statistical threshold was lowered to $p < 0.001$, uncorrected for multiple comparisons.

Table 4.4: Activation for low relative to high visual familiarity.

Region		**Mean peak coordinate**			**Z-score**			
		x	y	z	Lo > Hi	Kanji Lo > Hi	Hiragana Lo > Hi	VisFam ×Script
L	vOT	–45	–58	–11	3.34	2.80	1.96	0.95
L	pars opercularis	–42	5	31	5.22	5.54	1.61	3.48
L	pars triangularis	–42	29	7	4.92	5.00	1.82	2.87
L	frontal operculum	–30	29	1	5.06	4.72	2.45	2.24
R	frontal operculum	30	20	–5	5.38	5.56	1.89	3.34

Next I turned to the effects of lexical frequency. I contrasted the lowest frequency items to the highest in order to maximise the difference in frequency. Activation associated with lexical frequency was found in several left frontal regions, including pars opercularis, pars triangularis, a region of anterior paracingulate sulcus and the right deep frontal operculum (Table 4.5), consistent with previous studies (Carreiras et al., 2006; Fiebach et al., 2002; Hauk et al., 2008; Kronbichler et al., 2004). In addition, the ROI analysis identified a peak in lateral inferior temporal gyrus [–54 –55 –14, *Z*=4.15, *p*=0.003] adjacent to, but not overlapping, the activation seen in vOT for visual familiarity (Figure 4.3c). In fact it was approximately 1cm lateral to the visual familiarity peak and located in the inferior temporal gyrus, rather than the occipitotemporal sulcus.

Table 4.5: Activation for Lower > Upper frequency

Region		Mean peak coordinate			*Z*-score
		x	y	z	
L	pars triangularis	–45	26	19	4.65
L	pars opercularis	–45	11	28	4.74
L	inferior frontal sulcus	–51	17	34	4.74
L	paracingulate sulcus	–3	26	46	6.05
R	frontal operculum	30	23	–5	4.67

Within vOT, there was no main effect of Lexical Frequency (Figure 4.3d). There was, however, a small peak for the interaction of Lexical Frequency and Visual Familiarity at [–45 –64 –8], although this did not reach statistical reliability (*Z*=2.81, *p*=0.177). Nonetheless, the pattern of activation across conditions suggests that visual familiarity modulated low frequency words but not high frequency words (see Figure 4.3e). Words in the middle frequency quartiles showed intermediate sized visual familiarity effects.

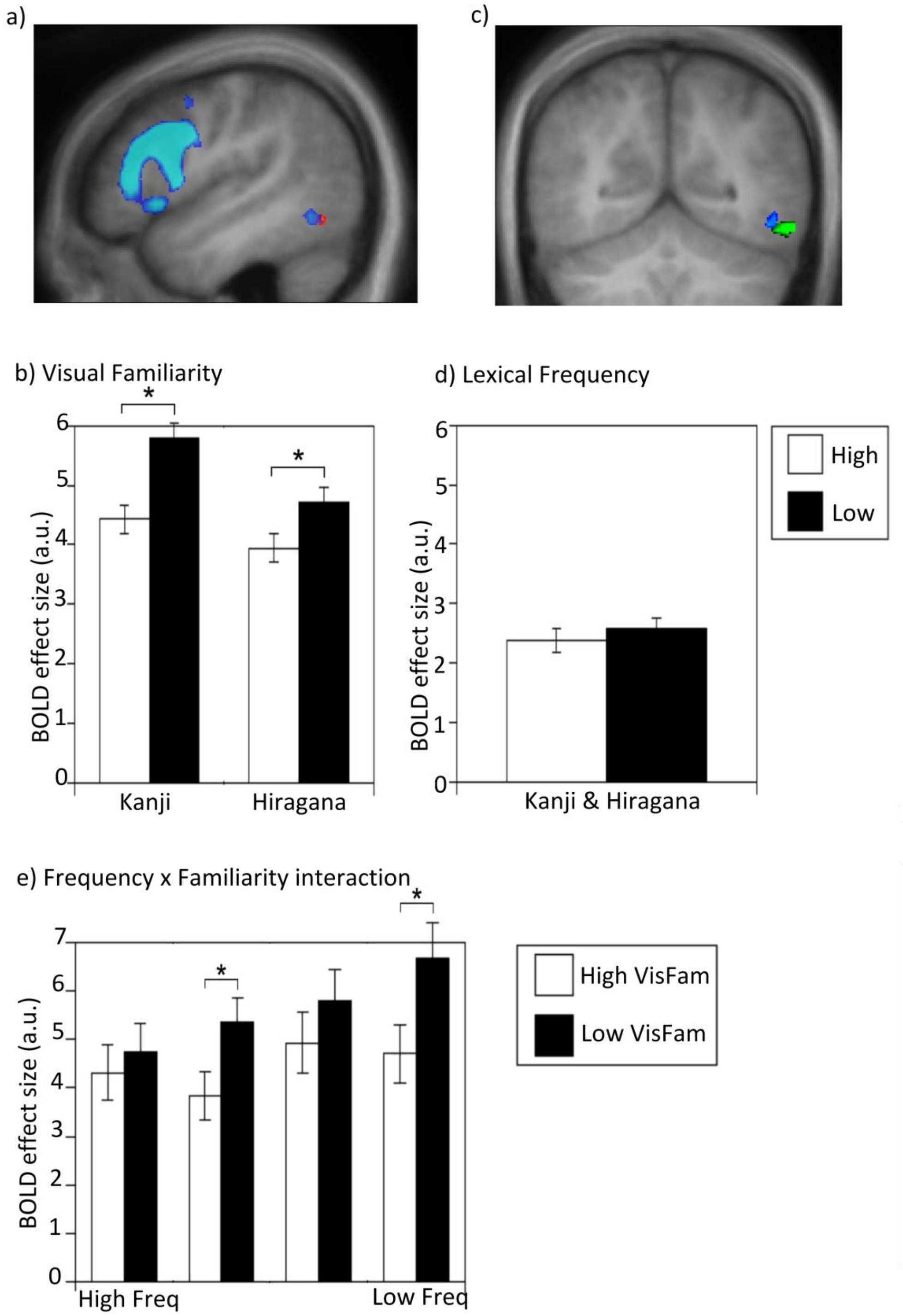

Figure 4.3: a) The overlap of Visual Familiarity (low>high) in blue and Script (Kanji > Hiragana) in red at x=–48. b) Bar plot for the BOLD signal per condition relative to fixation at x=–45, y=–58, z=–11 (in arbitrary unit), showing Visual Familiarity effect. c) Visual Familiarity effect (low > high) in blue and Lexical Frequency effect (Lower > Upper) in green at y=–58. Also shown are bar plots for the BOLD signal per condition relative to fixation at x=–45, y=–58, z=–11 for d) Lexical Frequency and e) Frequency x Familiarity interaction.

Finally, I turned to the question of whether the different scripts, Kanji and Hiragana, influenced vOT activation. Relative to Hiragana, Kanji produced significantly greater activation within vOT (Figure 4.3a & b). The peak was slightly posterior to the visual familiarity peak, although the clusters of activation were largely overlapping (Figure 4.3a). Outside the vOT region-of-interest no significant activation was found in the whole brain search. The opposite comparison of Hiragana relative to Kanji revealed no significant activation.

To investigate whether the increased activation in vOT for Kanji relative to Hiragana effect was driven by the inevitable difference in the visual complexity across scripts, a 2×2×2 ANOVA with Script (Kanji, Hiragana), Visual Complexity (high, low), and Group (Tokyo, London) was run. Since the two scripts differed in both the number of strokes and of characters, the total number of strokes per trial was used as the measure of visual complexity. There was no main effect of Visual Complexity nor interaction between Visual Complexity and the other two factors within vOT, even at a lenient statistical threshold of $p<0.001$ uncorrected for multiple comparisons. In other words, there was no evidence that the vOT activation observed for Kanji relative to Hiragana was a by-product of the greater visual complexity of Kanji words. Outside of the region-of-interest, high relative to low visual complexity was associated with activation in the left calcarine sulcus [–15 –94 –5, $Z=3.68$] and the right lingual gyrus [18 –91 –20, $Z=3.90$] at a threshold of $p=0.001$ (uncorrected).

4.4. DISCUSSION

The aim of the current study was to test whether visual familiarity and lexical frequency have separable effects on activation levels in vOT, as predicted by the Interactive Account. The results confirmed that visual familiarity, as opposed to lexical frequency, had a strong effect on vOT activation that was qualified by a small (but non-significant) interaction. Visual familiarity had essentially no effect on the most frequent words but a greater effect on the least frequent. In contrast, lexical frequency modulated activation in a region of the inferior temporal gyrus lateral to the visual familiarity effect in vOT. Finally, vOT also showed higher activation for Kanji than Hiragana words, although this was not due to their inherent differences in visual complexity. These findings place important constraints on understanding the nature of neural information processing in the region.

Given that vOT is a region of extrastriate visual cortex, it is perhaps not surprising that the region is sensitive to the familiarity of visual patterns. Indeed, a visual familiarity effect for faces in vOT has been previously reported (Eger et al., 2005). Although written words are a special form of familiar visual patterns, they too appear to be sensitive to this basic property of the visual system (Nazir et al., 2004; Xue et al., 2006; Xue et al., 2008). Within a predictive coding account, this is implemented in terms of more accurate top-down predictions for highly familiar visual patterns. This reduces the prediction error between the top-down and the bottom-up signals, reducing the regional BOLD signal. It is worth noting, however, that this visual familiarity effect is likely to be task-dependent and only present in tasks that place strong demands on integrating bottom-up visual form information with top-down non-visual properties of the stimulus. In the current experiment, linking visual forms to their sound and meaning is important for either recognizing them as a word or correctly rejecting them as a nonword. Tasks with similar demands on vOT processing such as reading aloud or reading for meaning would also be expected to demonstrate greater vOT activation for less visually familiar words. In contrast, one-back tasks or purely perceptual decisions may not show a visual familiarity effect for written words because neither places significant demands on integrating bottom-up visual and top-down non-visual information (Hellyer et al., 2011; Price and Devlin, 2011; Wang et al., 2011b).

Like visual familiarity, lexical frequency would also be expected to modulate the accuracy of top-down predictions into vOT. Indeed, in alphabetic languages, where there is essentially a single visual pattern per word, lexical frequency does affect vOT activation (Chee et al., 2003b; Hauk et al., 2008; Joubert et al., 2004; Kronbichler et al., 2004). In Japanese, however, where a word can be written in different scripts, there was no significant main effect of lexical frequency on vOT activation and only weak evidence of its interaction with visual familiarity. For highly frequent words, visual familiarity had no effect on vOT activation whereas activation for low frequency words was modulated by visual familiarity. In contrast, lexical frequency was found to significantly modulate activation in a region of the left inferior temporal gyrus lateral to the area in vOT showing a visual familiarity effect. Previous studies have argued that this is a functionally separate region engaged by multi-modal semantic processing rather than by visual word forms (Cohen et al., 2004; Moore and Price, 1999) and the current results are consistent with this.

Finally, although both Kanji and Hiragana strongly engaged vOT, there was significantly greater activation for Kanji relative to Hiragana. Because Kanji words generally have more strokes and fewer characters, they tend to be more visually complex than words written in Hiragana. Supplemental analyses, however, showed that the effects of visual complexity manifested in early visual cortices rather than in vOT, consistent with most previous studies (Hsu et al., 2011; Hellyer et al., 2011; Tarkiainen et al., 1999 but see Szwed et al., 2011 who failed to find any significant effects of visual complexity). Thus, the activation difference in vOT is assumed to reflect the different links between the surface form of the word and its non-visual properties (semantics and phonology). Specifically, a Kanji word composed of more than one character carries the meaning of the whole word as well as its individual characters, which may or may not automatically interfere with the meaning of the whole word. In contrast, a Hiragana word is associated with only one meaning (unless, of course, it is homophonic). Moreover, the relation between a Kanji word and its phonological form is largely arbitrary and depends critically on the combination of characters present in the word. Hiragana characters, on the other hand, are nearly 100% consistent in their pronunciation with each one representing a one-to-one mapping to a mora. Consequently, the contribution to prediction error from semantic and phonological regions to vOT is much less for Hiragana than for Kanji, resulting in lower vOT activation.

According to the Interactive Account (Price and Devlin, 2011), all words generate top-down predictions that arrive at vOT and support visual forms consistent with the word (Devlin et al., 2006a; Kherif et al., 2011). For instance, a word such as “yen” sends top-down predictions to vOT that support its Kanji (円), Hiragana (えん), and symbolic (¥) form. These predictions are highly accurate precisely because the word is so common, resulting in essentially equal activation across scripts. In contrast, a less common word such as ”wit” sends less accurate top-down predictions to vOT supporting both its Hiragana (とんち) and Kanji (頓知) forms. In this case, the fact that Hiragana is the more visually familiar form results in less prediction error (and therefore less activation) than the Kanji form. By this account, then, a visual familiarity benefit is principally expected for low, but not highly, frequent words – a pattern demonstrated in Figure 4.3e, but only weakly. Further studies will be required to establish whether this prediction of an interaction between visual familiarity and lexical frequency in vOT is reliable.

Can the current results also be understood in terms of orthographic input lexicon accounts that posit specialised representations of whole word orthographic patterns? In its strongest form, entries in the lexicon are truly “lexical” and abstract away from visual properties of the word such as capitalisation, font, size and even script. Consequently there is a single lexical entry for a word regardless of its script. Obviously, this version is incompatible with the current findings because it cannot explain the activation differences in vOT due to visual familiarity nor script. Some authors have argued, however, that lexical entries in an orthographic input lexicon are specific not only to words but also letter or case identities (Kronbichler et al., 2009). In the case of Japanese orthography, this would entail separate entries for the Kanji and Hiragana forms. By this account, the ease of accessing the form would be modulated by one's experience with the pattern (i.e., its visual familiarity), and is consistent with the main effect of visual familiarity observed here. The observed differences in activation for Kanji and Hiragana, however, are problematic for accounts that claim vOT is an orthographic input lexicon (Glezer et al., 2009; Kronbichler et al., 2004). If every word in an orthographic lexicon has separate Kanji and Hiragana entries and the accessibility of each entry depends on its visual familiarity, then Hiragana and Kanji should produce equivalent activation unless they differ in terms of visual familiarity. Indeed, some previous experiments that reported greater vOT activation for Kanji relative to Hiragana may have confounded script with visual familiarity (Ha Duy Thuy et al., 2004; Nakamura et al., 2005a; Sakurai et al., 2000), but in the current experiment, visual familiarity was carefully balanced across scripts so that Kanji was no more familiar than Hiragana (see also Ino et al., 2009). Nonetheless, I still observed a significant increase in activation for Kanji relative to Hiragana that may prove a challenge for the orthographic input lexicon account.

The current findings also raise an important methodological point about the use of alphabetic scripts in reading research. For many valid reasons, the majority of reading research has been conducted with alphabetic languages and has produced considerable advances in the cognitive and neural mechanisms underlying reading (Price and Mechelli, 2005; Shaywitz et al., 2004; Ziegler and Goswami, 2005). In addition, it has repeatedly been shown that both alphabetic and non-alphabetic scripts such as logographs engage essentially the same neuroanatomical system during reading (Booth et al., 2006; Chee et al., 2003a; Chee et al., 2000; Chen et al., 2002; Fu et al., 2002; Hu et al., 2010; Kuo et al., 2003; Tan et al., 2001). This finding, however, means that inferences drawn regarding the nature of regional information processing need to be consistent with a range of writing systems (i.e., alphabetic, syllabographic, logographic, etc) in order to explain the common information processing across scripts. For instance, Chinese readers engage a region of left middle frontal gyrus (MFG) more strongly than English readers (Tan et al., 2005; Tan et al., 2003) and it is theoretically possible that this is due to language-specific neuronal responses. A more parsimonious explanation, however, is that Chinese Hanzi characters increase visuospatial working memory demands relative to letters and this explanation has the advantage of being consistent with reports of left MFG activation in non-linguistic studies of visual working memory (Bledowski et al., 2010; Kravitz et al., 2011). Moreover, it offers a principled account for why this differential activation disappears in Chinese and English dyslexics (Hu et al., 2010). In short, a unified, cross-cultural account of the neural information processing underlying reading requires a systematic investigation of a range of different languages and scripts.

The current study took advantage of a unique property of the Japanese writing system in order to better characterise neural information processing in vOT during visual word recognition. Here I demonstrate a dissociation between lexical frequency and visual familiarity not possible in alphabetic languages and use the findings to evaluate competing theories of vOT function.

5. ANATOMICAL CONNECTIVITY OF vOT

5.1. INTRODUCTION

Previous chapters established that vOT responses are the combination of bottom-up visual information from the primary visual cortex and top-down non-visual information from other regions within the language system and that interaction is an important aspect of vOT function. For this interaction to occur, vOT and these other regions of the language system should be functionally connected. That is, the neurophysiological responses in these distant regions should be temporally correlated, an indication that these regions are functionally interacting. Indeed, Chapter 3 demonstrated that the activation pattern across the experimental conditions in POp, a region of Broca's area implicated in language, mirrored that in vOT, suggesting a possible functional link between these regions.

More specifically however, inter-regional interactions can be investigated by directly measuring functional connectivity (Friston, 1994). Previous functional connectivity studies have indeed suggested several links between vOT and specific regions within the language system. For instance, Horwitz and colleagues (1998) investigated inter-regional correlations between angular gyrus (ANG), a region in the classical neurological model of reading, and the visual association areas. They tested both dyslexic and non-dyslexic readers of English using PET with pseudowords and low frequency exception words (i.e., real words whose spelling does not comply with the orthotactics of English such as "yacht") in single word reading (aloud) tasks. They found that in the left hemisphere, the regional cerebral blood flow in ANG was correlated with that in vOT during both pseudoword and exception word reading in the normal, but not in the dyslexic, readers, suggesting that this functional link is indeed important for reading.

Broca's area has also been reported to be functionally connected to vOT. For instance, Bitan and colleagues (2005) investigated functional connectivity during spelling and rhyming tasks on visual stimuli with dynamic causal modelling (DCM, Friston et al., 2003). They found endogenous connections[4] (i.e., common to all stimulus-types, irrespective of tasks) between vOT and inferior frontal gyrus (IFG), partially overlapping Broca's area. Endogenous connections between vOT and the opercular part of Broca's area have also been demonstrated with non-alphabetic orthography, Japanese kanji and hiragana, in a lexical decision task (Kawabata Duncan et al., 2013). Furthermore, differential functional connections have been demonstrated for sub-regions of IFG encompassing Broca's area and vOT. Bokde and colleagues (2001) divided IFG into two sub-regions: Broca's area, implicated in phonological processing and the orbitofrontal part, implicated in semantic processing. They used words, pseudowords, consonant letter strings and false font strings in order to separate semantic processing, which would be required for words only, from phonological processing, which would be involved in all stimuli except for the false fonts. Functional connectivity was measured in terms of the inter-regional correlation of the BOLD response during a one-back matching task. As they predicted, only word stimuli resulted in functional connectivity between both of the sub-regions and occipital, occipitotemporal and temporal regions including vOT. In contrast, pseudowords and consonant strings showed correlation only between these regions and Broca's area. Thus, this study suggests that vOT is functionally connected to different parts of IFG depending on the cognitive processes induced by the stimuli - Broca's area during phonological processing and the orbitofrontal part during semantic processing.

Similarly, a functional connectivity study using DCM has also found endogenous connections between sub-regions of the prefrontal cortex including Broca's area and those of vOT. Mechelli and colleagues (2005) compared neuronal responses for regular words, pseudowords and exception words and tested the hypothesis that the differential responses elicited by these different word-types were associated with distinct regional couplings within the reading system. In this

---

[4] These connections are also referred to as "intrinsic" connections, particularly in older papers.

study, in addition to the prefrontal cortex, vOT was also divided into three sub-regions: anterior, middle and posterior, based on their differential functional specialisations reported in previous studies. The details were described in Chapter 1 but briefly, the posterior part (y = –60 to –70 in MNI space) responded to pseudowords more than familiar words while the reverse was true for the anterior part (y = –40 to –50). In contrast, the middle part (y = –50 to –60) was constantly activated during visual word recognition and by various written stimuli but the pattern of response is complex and also modulated by the task (Price and Devlin, 2011). They found an endogenous connection between the anterior vOT and the triangular part of Broca's area, which was also stronger for (i.e., modulated by) exception words relative to pseudo and regular words. They also found a functional connection between posterior vOT and dorsal premotor cortex, which was stronger for pseudowords relative to exception and regular words. Finally, there was a connection between the middle vOT and the opercular part of Broca's area, which was modulated by exception words relative to pseudo and regular words. The finding that the sub-regions of vOT are functionally connected to the sub-regions of Broca's area differently, which is further modulated during processing of particular word-type, also suggests that such functional couplings might be subserved by different anatomical connections.

Another region that has been shown to be functionally connected to vOT is the supramarginal gyrus (SMG). Van der Mark and colleagues (2011) examined functional connectivity of vOT in five, posterior-to-anterior sub-regions spanning from y = –80 to –30 (in MNI coordinates). The "VWFA" was centred on y = –54. A phonological lexical decision task ("does XXX sound like a word?") with German words, pseudohomophones, pseudowords and false font strings was used and both dyslexic and non-dyslexic children were tested. Using seed-voxel correlation mapping (Biswal et al., 1995), a method to measure functional connectivity as the correlation of time series between the seed voxel(s) and all other voxels, they found that only the "VWFA" was functionally connected to typical language areas including SMG in the controls. Similarly, Kawabata Duncan and colleagues (2013) used Japanese kanji and hiragana words and nonwords in a lexical decision task and investigated functional connectivity amongst four regions including vOT. They found endogenous connections between vOT and SMG (in addition to pars opercularis mentioned above).

These functional connectivity studies suggest that vOT is functionally connected to Broca's area and to the inferior parietal lobule (ANG & SMG). Regardless of whether the connections are feed-forward, feedback or interactive, in order for distant cortical regions to be functionally connected, there must exist underlying anatomical connections that enable such communications between them. These can be i) direct anatomical connections between two regions, ii) indirect via another region(s) or iii) indirect via a common input (Friston, 1994). The knowledge of the anatomical basis of functional connections is invaluable to our understanding of vOT function during reading and carries important implications for models of reading. However, currently very little is known about the anatomical connectivity of vOT.

In humans, anatomical connectivity studies are traditionally done post-mortem. The dissection of human brains for instance, has been widely used to investigate both long and short fibres (e.g., Martino et al., 2010a; Martino et al., 2010b; Sarubbo et al., 2011; Türe et al., 2000). Although this method is very informative, it is also time consuming and extremely complex since the fibres are often related to one another structurally and delineating one fibre system could lead to destruction of another (Türe et al., 2000). In addition, since the procedure advances based on one's assumption on the directions of the fibres, the results could also be biased (Axer et al., 2012) and it could lead to erroneous and artificial pathways, particularly when crossing fibres are encountered. Neuronal tracing (Oztas, 2003), such as that of Wallerian (anterograde) degeneration, has also been used to study connectivity in humans. An injury to the neuronal cell body can cause the distant part of the axon to degenerate towards the terminations. The extent of these anterogradely degenerating axons can be traced with staining methods such as the Nauta method (e.g., Clarke, 1994; Miklossy et al., 1991) on histological sections, which can reveal the degenerating axon terminations. Although this method can demonstrate regions (terminations) that are anatomically linked, it cannot always successfully delineate the pathway between them (Schmahmann and Pandya, 2009). Moreover, since this procedure requires a lesion, in human brains its use is limited to the area with naturally occurring lesions. Therefore, it does not allow systematic experiments to be conducted. In addition, the limited availability of the brains means that in general these post-mortem methods are not easily accessible.

Alternatively, non-human primates can be used to study anatomical connectivity. One advantage of using monkeys is that it is possible to select any site to test one's specific question. In order to detect the distribution of degenerating fibres for instance, a lesion can be intentionally inflicted in monkeys, which is needless to say not permitted in humans. Another procedure, which is limited to non-human primates, is neuronal tracing using autoradiography (Schmahmann and Pandya, 2009). Radiolabelled tracer is injected into a selected area of the brain of living monkeys and the tracer is transported anterogradely by the axoplasmic flow. After a certain period of time, depending on the type of the tracer, the animal is then sacrificed and the tissue is analysed using autoradiography. This method can reliably characterise the origin and the terminations as well as the neuronal pathways.

Although anatomical connectivity has been studied routinely in non-human primates using these invasive methods, the results cannot always be extended to humans without caution for a number of reasons. First of all, there are gross anatomical differences between monkey and human brains. Of particular relevance to vOT is that while humans have two gyri on the ventral surface of the temporal lobe, monkeys have only one. Therefore, the single gyrus medial to the occipitotemporal sulcus could be the homologue of fusiform or parahippocampal gyrus. Indeed, it is sometimes labelled as fusiform gyrus (Petr et al., 1949) while others prefer to assign the name parahippocampal gyrus to it (Felleman and Van Essen, 1991; Suzuki and Amaral, 1994; Van Hoesen, 1982; Whitlock and Nauta, 1956). Moreover, this discrepancy cannot be settled by the cytoarchitectonics either. According to the classic work of Brodmann (1909), the human vOT region falls within BA37. Von Economo (1929) defines the same region as PH[5] and more specifically, fusiform gyrus as area TF. Unfortunately, Brodmann's description for BA37 was anatomical and he did not provide sufficient information on distinctive cytoarchitectonics to allow a comparative study. The descriptions of von Economo in human brains on the other hand, were well detailed and there are some quite striking similarities between the descriptions of human TF and a monkey area TF, defined by von Bonin and Bailey (1947), which is located medially to the occipitotemporal sulcus. Specifically, these two areas share the following common features: layer II is not very dense in cells; layer III is relatively thick with medium-sized pyramidal cells; layer IV has granule cells (although the size differ across species); layer V displays a darker middle band with a light band on the either side; and finally layer VI has a sharp boundary with the white matter. Furthermore, both regions were more recently subdivided into two cytoarchitectonic areas in a similar fashion. In humans, two novel areas were identified within the posterior fusiform gyrus: FG1, which is medially located extending on the lateral bank of collateral sulcus; and FG2, which is laterally located encompassing the occipitotemporal sulcus (Caspers et al., 2013). The monkey area TF was similarly divided into two sub-regions: the medial portion TFm and the lateral portion TFl (Suzuki and Amaral, 2003a; Suzuki and Amaral, 2003b). Once gagain, there are some notable similarities in these sub-regions between the two species. For instance, in the medial portions FG1 and TFm, layer III is homogeneous through the layer and not very dense in cells, and there is no

---

[5] According to Von Economo (1929) cytoarchitectonics of the posterior segment of the temporal lobe resembles that of the parietal lobe. Therefore, this area is labelled as "PH" with "P" denoting to the parietal lobe although it is anatomically located in the temporal lobe.

clear boarder between layers V and VI. The lateral portions FG2 and TFl both have a highly dense layer II with a prominent layer IV and a cell dense layer VI. These similarities certainly raise a possibility that the monkey TF might be homologous to the human vOT. However, some reservations remain since they are anatomically not the closest match as the monkey TF occupies the medial temporal lobe.

Anatomically, the closest match to the human vOT is the monkey area TEO[6] (Bailey et al., 1950; Iwai and Mishkin, 1969; von Bonin and Bailey, 1947) and more precisely, its ventral part TEOv that encompasses the posterior occipitotemporal sulcus (Kravitz et al., 2013). However, this classification was based on behavioural and anatomical evidence such as inter-hemispheric connections that separated TEO from both more anteriorly located TE and more posteriorly located OA (Iwai and Mishkin, 1969) but crucially, the cytoarchitectural identity of TEO remained unclear (Bailey et al., 1950; Iwai and Mishkin, 1969). Therefore, it is not known whether area TEO bears any similarities to the human vOT cytoarchitectonically. All in all, the problem is that no studies have directly compared the cytoarchitectonics of human and monkey vOT[7], unlike the ventrolateral prefrontal cortex (Petrides and Pandya, 2002). Therefore, all things considered, it is not straightforward to determine the monkey homologue of human vOT on the ground of anatomy and cytoarchitectonics at present.

---

[6] In the book “The Neocortex of Macaca Mulatta”, von Bonin and Bailey (1947) state: “the fusiform gyrus is covered with a thinner cortex which has distinct differences from TE” (p.42). As pointed out by Boussaoud and colleagues (1991), in Bonin and Bailey’s section drawings, the posterior TE corresponding to posterior fusiform gyrus is indeed labelled with a question mark. They describe this area as a transitional zone between areas OA and TE since it bears general similarity to both of these areas. Only in the photomicrograph section, this area is labelled “TEO”.

[7] The closest comparison made is between the human PH and the monkey TE. According to von Bonin and Bailey (1947), no match was found between these regions (von Bonin and Bailey, 1947, p.49)

There are also functional differences across species. Crucially, non-human primates cannot read. Therefore, it is impossible to directly compare this function across species. On the other hand, like humans their social interactions rely heavily on face recognition which, in humans, is known to engage roughly the same part of the fusiform gyrus as reading, but more strongly in the right hemisphere (Haxby et al., 2002; Kanwisher et al., 1997). In monkeys, however, face processing primarily activates superior temporal sulcus although more recently evidence for face selectivity in the ventral temporal lobe has also emerged (Pinsk et al., 2009; Tsao et al., 2008; Ku et al., 2011). Since function of a cortical region is constrained by its anatomical connectivity (Passingham et al., 2002), functional differences in fusiform gyrus across species might indicate different anatomical connectivity. The anatomical connections that mediate human reading ability may be different or even absent in monkeys. Indeed, the human prefrontal and temporal association cortices are disproportionately large relative to the non-human primates (Rilling, 2006). The increase in the temporal lobe has been found to be mostly in the white matter (Rilling and Seligman, 2002) and it is possible that this is the result of the evolution of human language, subserving the link between the temporal and frontal lobe (Rilling, 2006). Given these points, the anatomical connectivity that mediates reading, particularly in vOT, should be studied in human subjects, in vivo, non-invasively.

The aim of the current study, therefore, was to investigate the anatomical connectivity of vOT directly in humans using diffusion-weighted MRI with probabilistic tractography – a tool that has become available more recently to investigate human anatomical connectivity non-invasively, in vivo. Specifically, I examined whether a direct, cortico-cortical, intra-hemispheric anatomical link exists between vOT and the regions for which functional connections are implicated: i) Broca’s area, ii) angular gyrus and iii) supramarginal gyrus. In so doing, vOT was divided into three rostro-caudal sub-regions given their differential response patterns to word-types. Seed masks were placed within the cortical grey matter in these rostro-caudal positions along the occipitotemporal sulcus and the probabilistic tractography algorithm generated the most probable paths from the masks to the rest of the brain. Both left and right hemispheres were investigated and hemispheric asymmetries were also examined. The analyses took three stages. The first set of analyses looked for cortico-cortical pathways from vOT with no further constraints. The second set investigated the more specific hypotheses that brain regions functionally connected to vOT are linked to vOT via direct cortico-cortical connections. Additional seed masks encompassing such regions were created and utilised in a two region-of-interest (ROI) approach. The final set of analyses investigated the hypothesis that middle occipital gyrus may be the common source of input for SMG and vOT, again using the two ROI approach.

5.2. MATERIAL AND METHODS

Participants

10 right-handed, healthy volunteers participated in the experiment (6M, 4F). The ages of the participants ranged from 18 to 37 (M=24). No history of neuropsychiatric conditions or substance abuse was reported. All participants provided informed consent. The experiment was approved by the Berkshire Research Ethics Committee (Reference: 06/Q1602/20).

DW-MRI acquisition

DW-MRI data were acquired on a 1.5 T Siemens Sonata MRI scanner with 40 mT/m gradients at the Oxford Centre for Clinical Magnetic Resonance Research by Dr. Joseph Devlin. An echo-planar imaging sequence was used (60 contiguous axial slices, 2.5 mm thickness, in-plane resolution of 2.5 mm × 2.5 mm) and three sets of data were acquired for each participant. Each set comprised 3 non-diffusion-weighted and 60 diffusion-weighted images acquired with a b-value of

1000 s/mm^2 uniformly distributed across 60 gradient directions. The diffusion gradients were uniformly distributed through space using a scheme optimised for white matter tractography (Jones et al., 1999). In addition, a high-resolution structural image (3D Turbo FLASH; TR=12 ms; TE= 5.65 ms; flip angle=19°; 1 × 1 × 1 mm) was also acquired for anatomic localisation.

Tractography

The data were corrected for eddy currents and small head movements by registering them to an initial reference volume (Jenkinson et al., 2002). The three data sets were then averaged to improve the signal-to-noise ratio and the results were used to compute 3D probability distributions of fibre direction at each voxel (Behrens et al., 2007; Behrens et al., 2003b). Probabilistic tractography (Behrens et al., 2003b) was run from several different seed masks using the FMRIB's Diffusion Toolbox (FDT 2.0, http://www.fmrib.ox.ac.uk/fsl/fdt). With this method, a pathway starts in a seed voxel. The spatial probability density function at that voxel is randomly sampled and the path moves one step (0.5mm) in the sampled direction. This process is repeated until the path reaches a pre-determined end point such as a target region or the edge of the brain where it stops. The individual steps are then integrated into a continuous path. The same process is repeated thousands of times from each voxel in the seed mask to generate an empirically-defined set of paths from the seed region.

For all analyses, the tracing continued until i) it reached the edge of the brain; ii) it intersected a termination mask (see below) used to prevent paths from “jumping” over intervening sulci; or iii) it intersected an exclusion mask (see below) so that any inter-hemispheric paths via the posterior corpus callosum and paths that travel towards the spine via brain stem, both of which were not of interest, were excluded from the analyses. Moreover, since the tractography algorithm could allow tracing of paths that loop back on themselves, a loop check was employed to ensure that such paths would be terminated. Similarly, the angle at which a path is allowed to turn (i.e., a curvature threshold) was ±80 degrees, in order to avoid anatomically implausible trajectories. Paths were disregarded if their probability of connection did not survive the threshold (see below).

Three different sets of analyses were conducted. The first set of tractography analyses seeded the vOT region to identify paths linking vOT to other brain regions. This set of analyses was blind to any *a priori* hypotheses about specific regions where the paths might terminate. However, the search was limited to the ipsilateral neocortex. The first tractography analyses used 25,000 samples per voxel in the seed mask to represent the population of pathways as accurately as possible. This is a very conservative approach as previous studies typically found that 5000 or fewer samples were sufficient to reach convergence (Behrens et al., 2007; Carreiras et al., 2009; Eickhoff et al., 2010; Tomassini et al., 2007). Indeed, my own analysis with the current data confirmed that although paths traced at 5000 samples had lower probability and did not extend as far as the paths traced at 25,000 samples, when no thresholding was applied they largely overlapped and this difference diminished when the same threshold of 0.1% of the number of samples from a single voxel was applied (Figure 5.1). In other words, the results were essentially identical with either 5,000 or 25,000 samples, but the latter was chosen to optimise sensitivity.

A second set of analyses investigated the more specific hypotheses that brain regions functionally connected to vOT such as the supramarginal gyrus (SMG), angular gyrus (ANG) and Broca's area (pars opercularis [POp] and pars triangularis [PTr]) are linked to vOT via direct cortico-cortical connections. These regions were chosen based on the literature reporting distinct *functional* connections between vOT and these regions (Booth et al., 2002; Horwitz et al., 1998; Mechelli et al., 2005). A two region-of-interest (ROI) approach was used in which each vOT mask was coupled with each non-vOT mask (e.g., anterior vOT & SMG, anterior vOT & ANG), resulting in 24 separate analyses (3 vOT regions × 4 non-vOT regions

× 2 hemispheres). This tractography method traces paths from each of the masks and only retains paths that reached the other mask increasing sensitivity by considering paths in both directions and minimising the effect of accumulated errors. In addition, as these analyses delineate the paths between the two ROIs, the trajectory of the path between them can be more easily identified than a single ROI analyses. Moreover, when both seed masks are placed in the cortical grey matter with two ROI approaches, it can be ensured that the traced paths reach, rather than pass by, the grey matter. Since this is a more sensitive method than the one used in the first set of analyses, the number of samples traced from each voxel was 5000 rather than 25000 used in the first analysis.

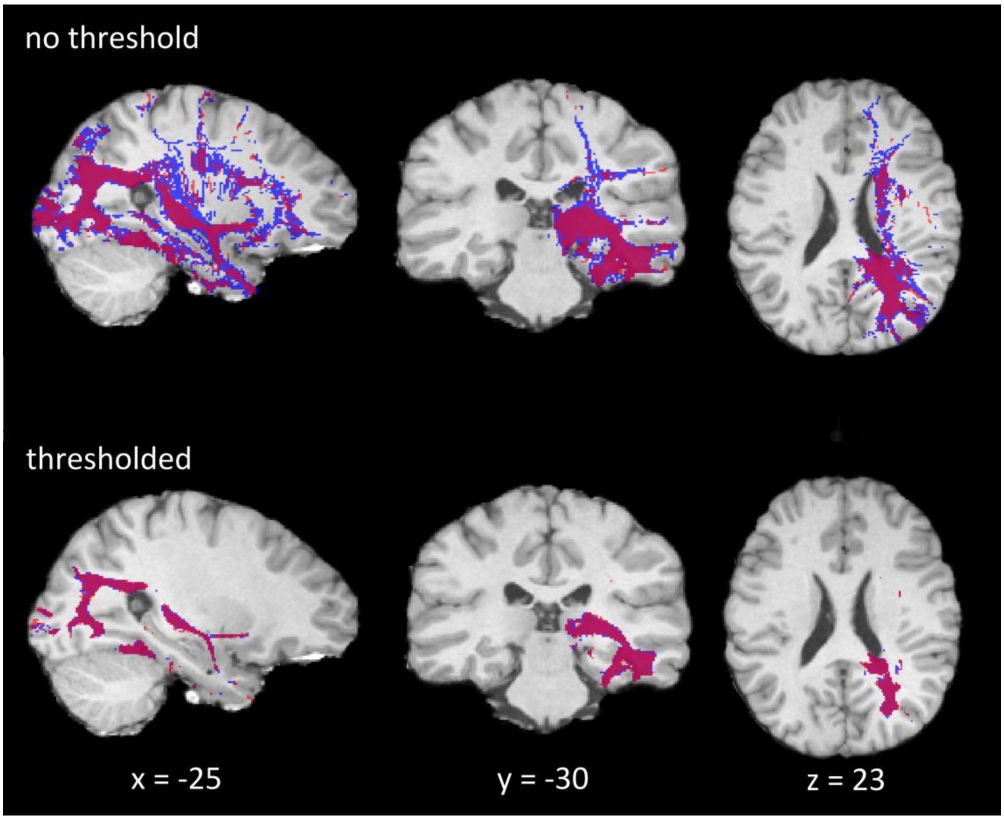

Figure 5.1: Comparison of analyses that used 5000 (red) and 25000 (blue) samples seeded from anterior vOT in the left hemisphere in one subject (standard space). Top panel shows the results with no thresholds and bottom panel shows the same results with the threshold of 0.1% of the number of samples from a single voxel (i.e., 5 for 5000 samples and 25 for 25000 samples). The difference between the two analyses diminishes greatly when thresholding is applied indicating no significant differences in the identified pathways.

A third set of analyses investigated the hypothesis that middle occipital gyrus may be the common input source for SMG and vOT. A two region-of-interest (ROI) approach was used with this set of analyses in which the MOG mask was coupled with each of the vOT masks and with the SMG mask, resulting in 8 separate analyses ([SMG & 3 vOT regions] × 2 hemispheres). As with the second set of analyses, the number of samples traced from each voxel for the third set of analyses was 5000.

Anatomical regions-of-interest

Anatomical regions of interest were defined on each participant's high-resolution T1 image after translating and rotating this into the standard reference space defined by the Montreal Neurological Institute MNI-152 template using a rigid body transformation. These masks were used in the tractography analyses as seed regions, as target regions or as constraints such as termination and exclusion masks (see below) for tracing.

Ventral occipito-temporal (vOT) masks

Three rostro-caudal positions along occipitotemporal sulcus were chosen based on posterior-to-anterior changes in both functional specialisation (Ben-Shachar et al., 2007a; Ben-Shachar et al., 2007b; Mechelli et al., 2005; Mechelli et al., 2004; Moore and Price, 1999; Price et al., 2003) and functional connectivity (Mechelli et al., 2005; Seghier et al., 2008). Each mask was manually traced over contiguous coronal slices. The medial edge of the mask was the crest of fusiform gyrus and it extended to the crest of inferior temporal gyrus. This included the lateral part of the fusiform gyrus, the entire occipito-temporal sulcus (medial wall, fundus and lateral wall), and the medial inferior temporal gyrus (Figure 5.2a). Where a split fusiform gyrus was present (4 out of 20 hemispheres), only the lateral fusiform gyrus was used for masking. The most anterior vOT mask encompassed y= –42 to –46, the middle vOT mask from y= –52 to –56 and the posterior vOT mask from y= –62 to –66 in standard space coordinates. Masks were defined in each hemisphere in each participant. From the total of 60 masks (anterior, middle, posterior vOT in LH and RH), 3 were lost due to unusual gyrification in the form of medial-to-lateral sulci that made tracing vOT masks problematic (Figure 5.2b). In addition, in one hemisphere the left middle vOT mask was excluded due to a medial-to-lateral sulcus cutting through the location of the seed mask (Figure 5.2c). As a result, 29 left vOT masks (10 x anterior, 9 x middle, 10 x posterior) and 27 right vOT masks (9 x anterior, 9 x middle, 9 x posterior) were included in the analyses for a total of 56 seed masks.

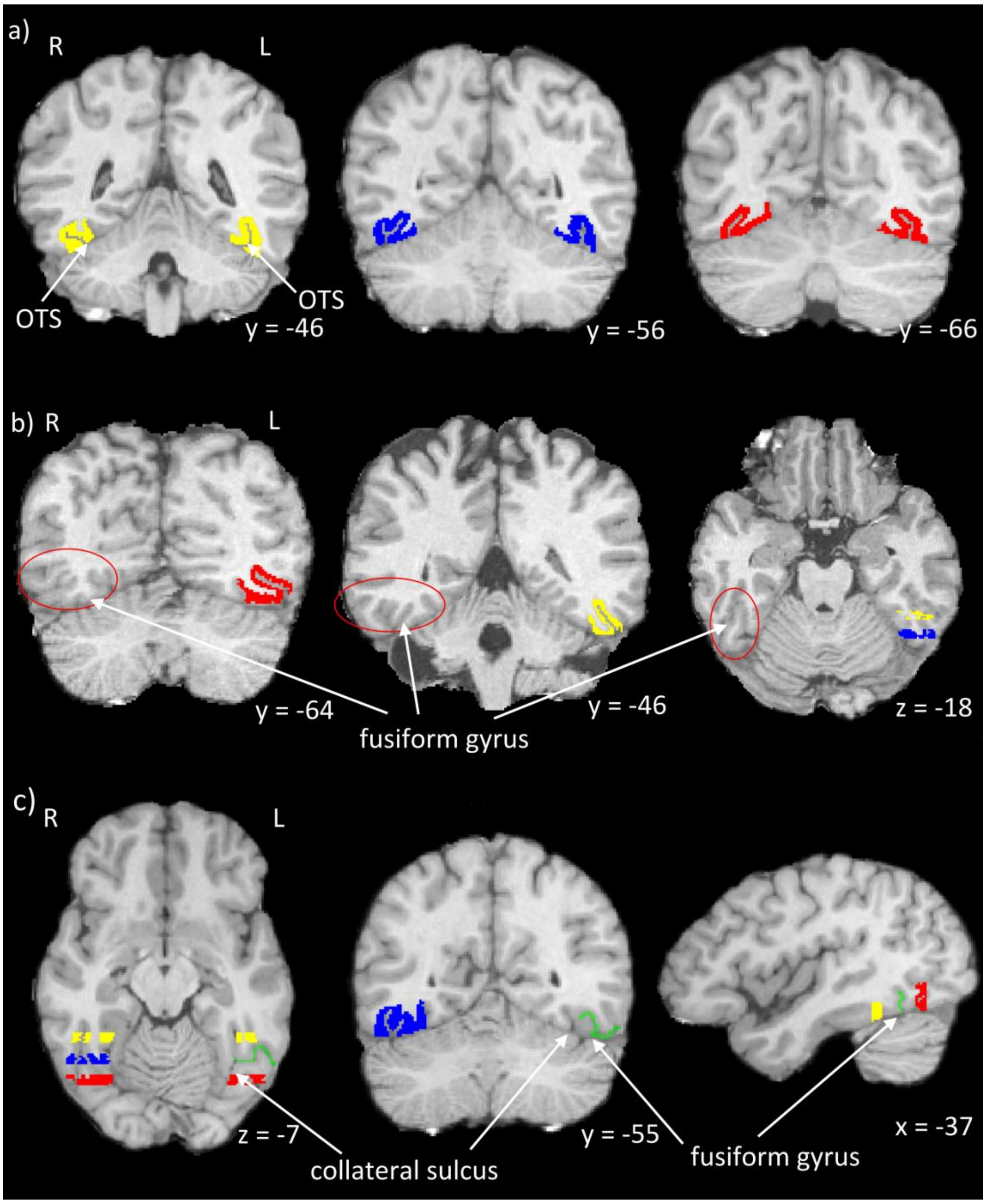

Figure 5.2: vOT masks. a) From left to right: anterior, middle and posterior vOT are masked in yellow, blue and red respectively. vOT masks included the lateral crest of fusiform gyrus and the medial crest of inferior temporal gyrus. b) Unusual gyrification (indicated with red ovals) made it impossible to identify the lateral crest of fusiform gyrus necessary for masking. c) Interfering sulci. The presence of a sulcus cutting through fusiform gyrus at around y = –55 (indicated with green lines) meant masking left middle vOT was not possible for this subject.

Non-vOT masks

An additional five cortical territories were anatomically identified: supramarginal gyrus, angular gyrus, pars opercularis, pars triangularis and middle occipital gyrus. The first four are areas reported to be functionally connected with vOT (Bitan et al., 2005; Bokde et al., 2001; Horwitz et al., 1998; Kawabata Duncan et al., 2013; Mechelli et al., 2005; van der Mark et al., 2011). The latter was included in order to test the hypothesis that functional connectivity between vOT and SMG resulted from the regions receiving a common source of input, namely MOG. All five regions were automatically delineated using Freesurfer's cortical parcellation tool (Desikan et al., 2006; Fischl et al., 2004). An automated process was employed in order to avoid any systematic bias in manually identifying these regions. Both regions of the inferior parietal lobule (SMG and ANG) as well as the subdivisions of Broca's area (POp and PTr) have greater anatomical variability than vOT, which is relatively uniform with clear anatomical landmarks. Specifically, in the inferior parietal lobule (IPL), while the sulcus intermedius primus (Jensen sulcus) separates SMG and ANG, there is high variability in the presence of additional gyri anterior or, more importantly, posterior to SMG (Kiriyama et al., 2009; Naidich et al., 1995). These gyral patterns make the boundary between SMG and ANG uncertain since it is not always straightforward to see whether such an extra gyrus belongs to SMG or ANG. Similarly in the inferior frontal gyrus, there is high variability in the configuration of the horizontal and ascending rami of the Sylvian fissure (Juch et al., 2005), which together define the subdivision of POp and PTr (e.g., V-shape, Y-shape). Moreover, a diagonal sulcus, whose presence is also highly variable (Keller et al., 2007; Naidich et al., 1995), can be mistaken for the ascending ramus of the Sylvian fissure. Consequently, an automated, unbiased method was used to delineate these regions. Specifically, for each subject, first the pial surface and the border of the grey and white matter were delineated for each non-vOT ROI, based on the anatomical region labels generated by Freesurfer. These were then transformed into volumes and converted into FSL space (Figure 5.3). Next, the space between the pial surface and the grey/white border was manually filled in. Visual inspection found that in some cases the results included brain regions that did not belong to the cortical territories to be delineated or did not include the sections that were required. In such cases, manual corrections were made.

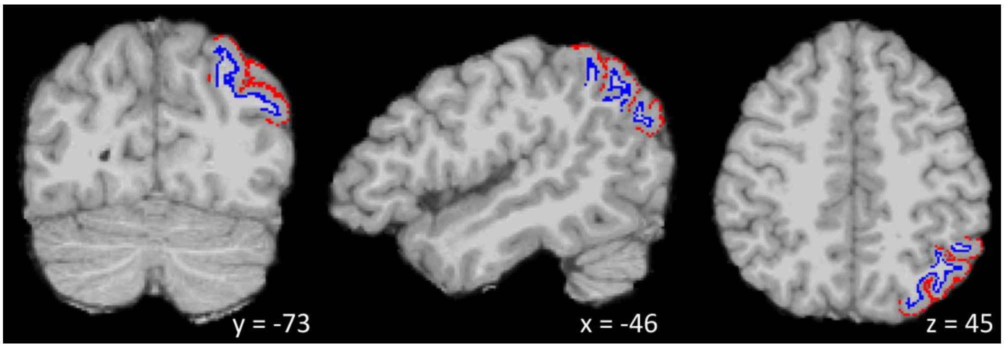

Figure 5.3: Delineated pial surface (in red lines) and the border of grey and white matter (in blue lines) for angular gyrus in LH. The space between the red and blue lines was then manually filled in.

To ensure that both the manually and automatically defined anatomical masks only include grey matter, each participant's T1-weighted structural scan was segmented into grey matter, white matter and cerebro-spinal fluid (CSF) using a probabilistic tissue segmentation algorithm (Zhang et al., 2001). This provides a probability that each voxel is grey matter, white matter or CSF with the sum totalling 100%. A whole-brain grey matter mask was generated by selecting all voxels that had at least a 20% change of being gray matter. This was then applied to the ROI masks such that only voxels that had a 20% or greater chance of being grey matter were included. In other words, white matter and CSF were systematically excluded from all ROI masks.

Finally, in order to minimise erroneous and irrelevant paths in the tractography analyses, two additional masks were defined. First, the tractography algorithm often ignores sulci given the relatively large voxel size used in DW-MR images. Therefore, when unconstrained, it may identify anatomically impossible tracts such as those going from one gyrus to another by crossing a sulcus (“jumping sulci”). Similarly, paths may run from the ventral surface of the cerebrum into cerebellum crossing the interstice. In order to prevent such paths, sulci in occipital and posterior temporal lobes as well as the gap between cerebrum and cerebellum were manually traced in each hemisphere. It was then combined with the CSF mask for the whole brain and used as a “termination mask” (Figure 5.4a). This ensured that the paths generated by the tracing algorithm were terminated when they reach any voxels within this mask and did not continue further. Further additions were made in the parietal and frontal lobes when tracts were found to cross sulci after the tractography was run. In other words, the inclusion of a termination mask helped to avoid false positive due to the coarse spatial resolution of the DWI data (2.5 × 2.5 × 2.5mm) relative to the anatomical image (1 × 1 × 1mm). Importantly, paths that encounter termination masks contribute to the calculation of the connectivity distribution. That is, the paths that have eventually been terminated are not deleted and the number of paths that go through the voxels is still taken into account in the calculation of the discretised spatial probability distribution, which is the number of paths going through the voxel divided by the total number of samples (the number of samples multiplied by the number of voxels in the seed). Second, since the aim of the study was to investigate the anatomical basis of functional connectivity of vOT with other cortical regions within the same hemisphere, inter-hemispheric paths and those directed subcortically were irrelevant. Therefore, in order to restrict the investigation to cortico-cortical intra-hemispheric connections, masks comprised of the midline (i.e., x= 0) and the brain stem at z= –10 were defined and used as an “exclusion mask” (Figure 5.4a). Tracts that reach voxels within this mask such as inter-hemispheric paths via corpus callosum and those that travel towards the spine via brain stem were completely eliminated and thus did not contribute to the connectivity distribution.

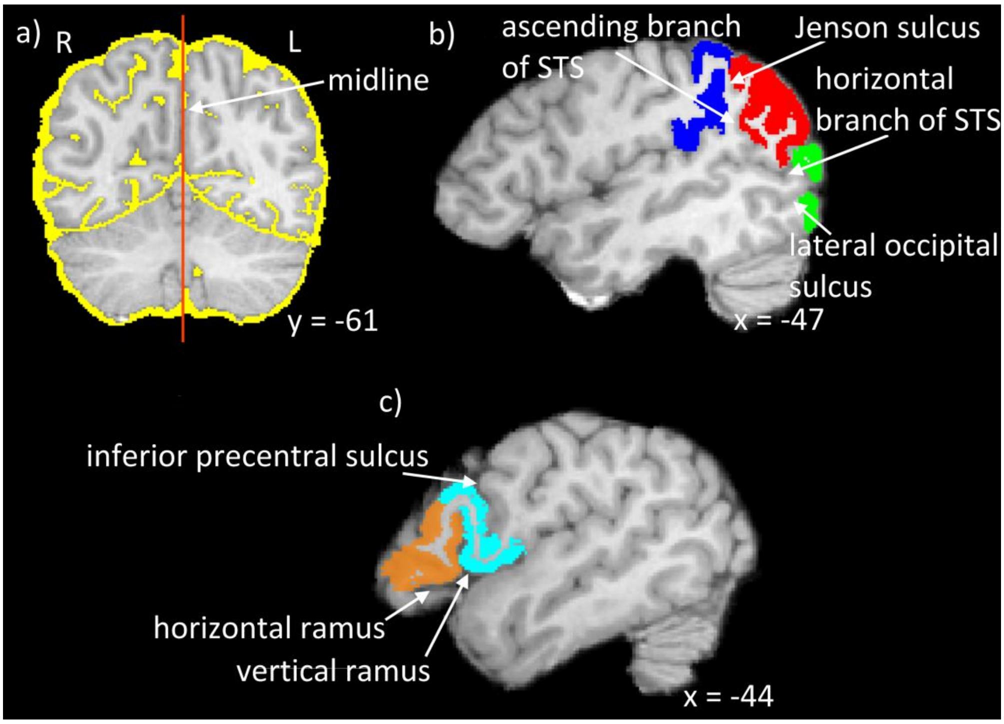

Figure 5.4: Anatomical masks. a) Termination mask comprised of manually traced sulci and CSF mask at y = –61. Also shown is the midline included in the exclusion mask in dark orange. The masked brain stem in the exclusion mask is not shown. b) Masks for supramarginal gyrus in dark blue, angular gyrus in red and middle occipital gyrus in light green at x = –47. STS = superior temporal sulcus. c) Masks for pars opercularis in light blue and pars triangularis in orange at x = –44.

Thresholding

Analyses of the tractography results require thresholding to eliminate noise while retaining meaningful paths. Thresholds here determine the minimum number of pathways going through a single voxel for that voxel to be included in the results.

In the absence of a "gold standard" for thresholding tractography results (details in Chapter 2), the vOT single mask analyses were first compared across various thresholds in order to examine the effects of thresholding at different levels. Figure 5.5 shows the results of a single seed tractography analysis originating in anterior vOT in the left hemisphere thresholded at 0, 1, 5, 10, 25 and 50 tracts for all subjects. As can be seen by these comparisons, the difference is minor across thresholds of above 5 whereas even thresholds of 1 or 5 eliminate noise (speckle-like paths) relative to no threshold. Therefore, for the vOT single seed analyses, the results were thresholded at 25 to reduce noise and maintain a conservative estimate of likely pathways.

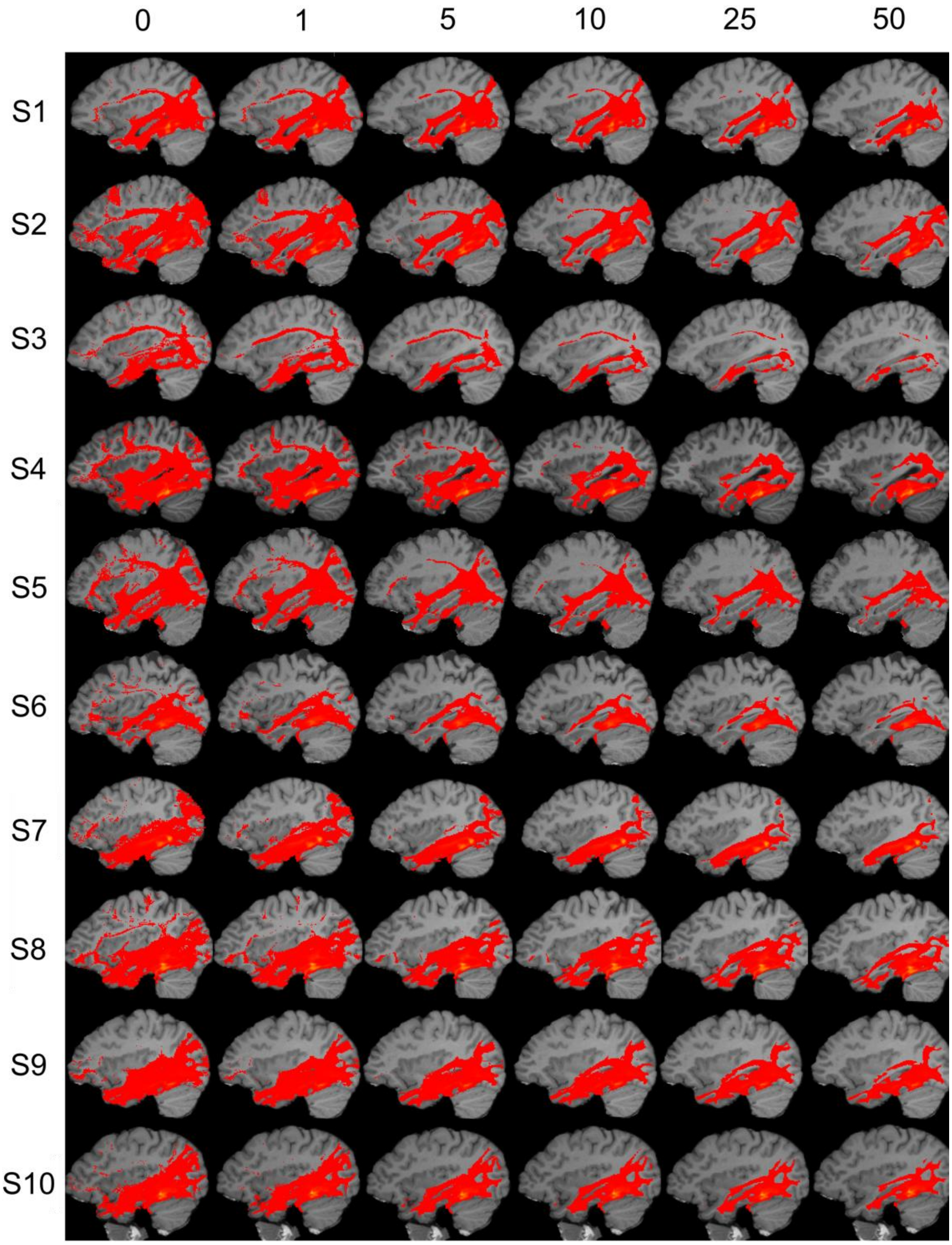

Figure 5.5: Tractography results seeded from anterior vOT in the left hemisphere with 25000 samples shown on a single parasagittal plane (X= –32) for all subjects (S1-S10) in standard space at six different levels of thresholding (0, 1, 5, 10, 25, 50). The basic pattern of results remains largely the same over the range of thresholds. However noise (speckle-like paths) can clearly be seen in the lower thresholds (0, 1) but is mostly absent for thresholds of 5 or larger. The differences in the results become marginal when the threshold is greater than 5.

5.3. RESULTS

vOT single seed analyses

The aim of this analysis was to conduct a search for pathways linking vOT to other ipsilateral cortical territories. Tractography from all vOT seed masks generated widespread paths terminating in temporal pole areas, middle/inferior occipital regions and prefrontal cortex. The trajectories of the inferior longitudinal fasciculus (ILF) and inferior fronto-occipital fasciculus (IFOF) were clearly delineated. In addition, a set of local U-fibres within vOT were identifiable. These tracts were seen regardless of the rostro-caudal position or hemisphere of the seed.

The paths from each vOT seed mask spread superiorly into the deep white matter and laterally into inferior temporal gyrus as well as to the adjacent vOT seed masks before travelling to further regions. This can be illustrated by changing the thresholds from high to low since the local connections have higher probability than further long-distance connections (Figure 5.6).

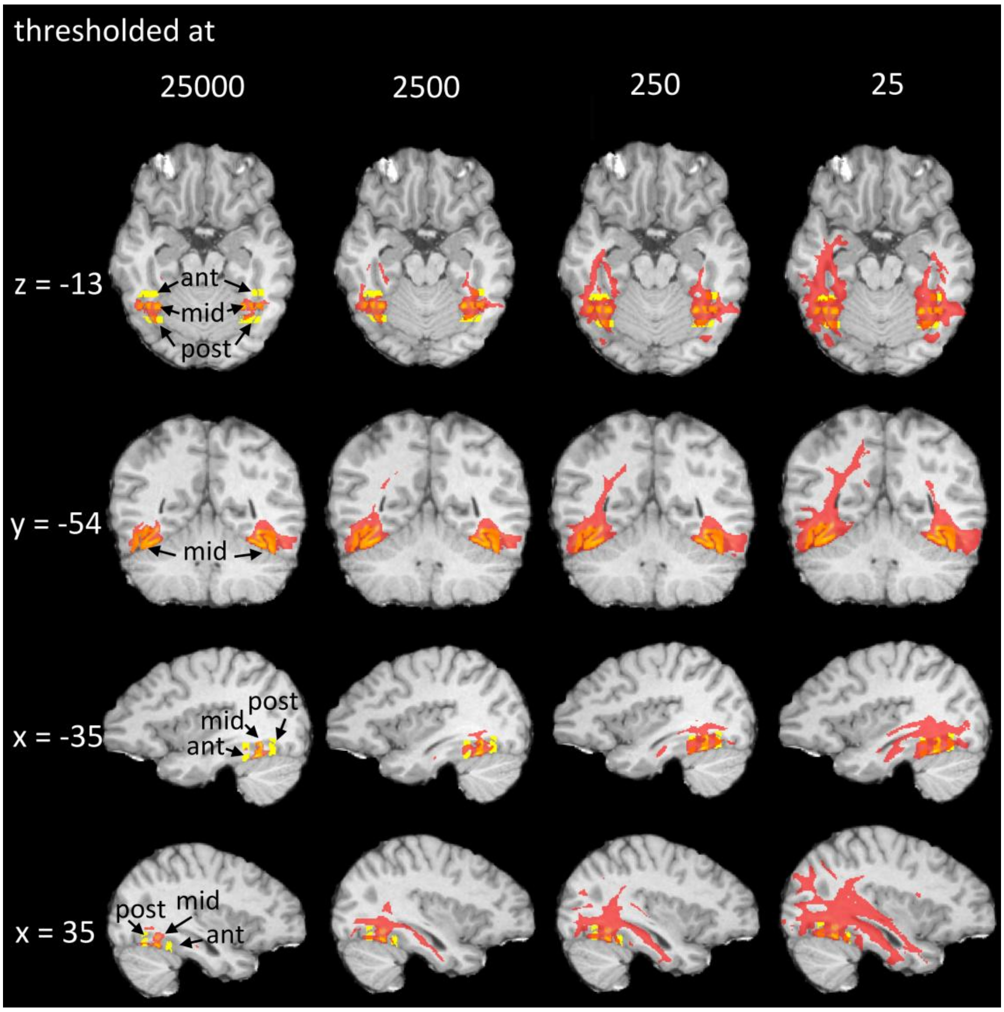

Figure 5.6: Tractography results seeded from left and right middle vOT seed masks in one subject (standard space). The paths (in red) are shown at different thresholds of 25000, 2500, 250 and 25 in order to illustrate the course of the tracts. Also shown are the vOT seed masks in yellow (ant: anterior vOT; mid: middle vOT and post: posterior vOT). Top row: axial views at z = –13; second row: coronal views at y = –54; third row: parasagittal views at x = –35 and bottom row: parasagittal views at x = 35.

The paths traced from vOT could be seen to take three separate courses. First, a path can be seen running across temporal and occipital lobes, following a trajectory inferior to the lateral ventricles (green arrow heads in Figure 5.7a). At the posterior end it reached inferior/middle occipital regions in most cases, posterior to the temporoparietal fibre intersection area (TPFIA) described by

Martino and colleagues (2013a) where seven white matter tracts intersect (Figure 5.8). At the anterior end it reached the temporal pole extensively in most cases in its dorsal and ventral, as well as lateral and medial, parts. This bundle of associative fibres was identified in 96% of the cases in RH and 93% in LH. The posterior and anterior termination points of this path and the number of cases (percentage) in which these were found are listed in Table 5.1. The trajectory and these termination points are consistent with previous anatomical (Crosby et al., 1962), dissection (Türe et al., 2000) as well as DTI (Catani et al., 2002; Catani et al., 2003; Wakana et al., 2004; Yeatman et al., 2013) reports of the inferior longitudinal fasciculus (ILF) (Catani et al., 2003; Crosby et al., 1962).

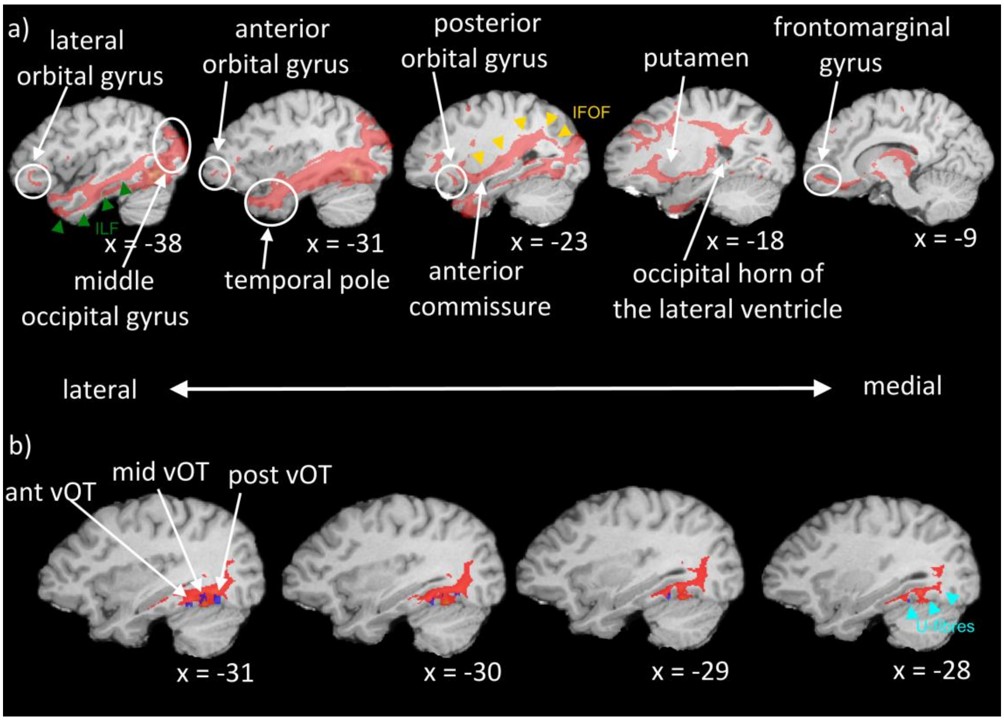

Figure 5.7: Paths traced from left posterior vOT in a single subject (standard space) is shown in red. There were no systematic differences across the rostro-caudal position or hemisphere of the seed masks. The top panel a): IFOF (labelled with yellow arrow heads) and ILF (labelled with green arrow heads) can be identified. ILF reaches the temporal pole in the anterior and middle occipital gyrus at the posterior end. The characteristic curve of IFOF wrapping around the bulb of the occipital horn of the lateral ventricle and the putamen can be seen at x = –23. At the posterior end this paths branch out superiorly and inferiorly (x = –31) and at the anterior end it reaches orbitofrontal cortex. The posterior terminations of IFOF and ILF merge and are difficult to visually separate. The paths are thresholded at 25 and the termination points are indicated with white circles. The bottom panel b): Local U-fibres (labelled with light blue arrow heads) linking three rostro-caudal vOT sub-regions. The locations of anterior, middle and posterior vOT seed masks in blue can be easily seen in more lateral slices while the U-shape of these paths becomes clearer as it moves more medially (at x = –28). The paths traced from middle vOT are shown at a higher thresholded at 2500 in order to demonstrate the U-fibres.

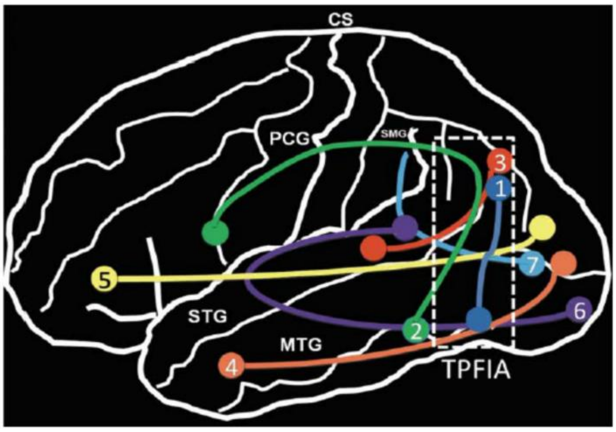

Figure 5.8: The schematic representation of the tracts recovered by diffusion tensor imaging tractography, showing the temporoparietal fibre intersection area (TPFIA) in a left hemisphere (rectangular box with dotted lines). 1 = Vertical* portion of the superior longitudinal fasciculus; 2 = arcuate fasciculus; 3 = middle longitudinal fasciculus; 4 = inferior longitudinal fasciculus; 5 = inferior fronto-occipital fasciculus; 6 = optic radiations; 7 = tapetum. Abbreviations: CS = central sulcus; MTG = middle temporal gyrus; PCG = precentral gyrus; SMG = supramarginal gyrus; STG = superior temporal gyrus. Adapted with permission from Lippincott Williams and Wilkins/Wolters Kluwer Health: Neurosurgery 72, ons87-ons97, Martino et al. (2013a), copyright (2013). *In the original paper, this tract (1) is labelled as "horizontal" portion, which appears to be an error given the orientation of this fibre bundle. Indeed, the same tract is labelled as "vertical segment of the SLF" in another paper (Martino et al., 2011).

A second path from vOT courses dorsally and curves forward over the occipital horn of the lateral ventricle forming an arch (yellow arrow heads in Figure 5.7a). At approximately at the level of anterior commissure, it passes ventrolateral to the putamen, within the extreme/external capsule, and proceeds anteriorly towards orbitofrontal cortex. In some cases, it then splits in the orbitofrontal cortex into anterior, posterior or lateral orbital gyri. The trajectory and terminations of this path are consistent with the inferior fronto-occipital fasciculus (IFOF), which is suggested to be connected to fusiform gyrus at the posterior end (Crosby et al., 1962) and the orbitofrontal cortex at the anterior end (Martino et al., 2010a; Sarubbo et al., 2011). This path was identified in 89% of the cases in LH and 93% in RH (Table 5.1). It is worth noting, that with the current resolution of the DTI images (2.5 x 2.5 x 2.5mm), it is impossible to differentiate the external and extreme capsules based on the DWI data, but studies in human brains using Klingler’s dissection technique (Klingler, 1935; Klingler and Gloor, 1960; Ludwig and Klinger, 1956), demonstrate that the IFOF is part of both extreme and external capsule (Ebeling and Cramon, 1992; Kier et al., 2004; Martino et al., 2011; Türe et al., 2000; Wang et al., 2011a). Within the frontal lobe however, in only a small number of cases did this path terminate in Broca’s area, a region implicated to be functionally connected to vOT. Specifically, only 13% of the paths terminated in POp and 14% terminated in PTr.

Consistent with a review of the white matter tracts from DTI, anatomical drawings and gross dissection photographs (Jellison et al., 2004), these two tracts merge at the posterior end. From there they branch out superiorly towards middle occipital regions and ventrally towards inferior occipital and temporal regions (Figure 5.7 & Table 5.1). In around half of the cases they also reached angular gyrus (56% in RH; 48% in LH), most of which was towards the posterior part of angular gyrus where it boarders with the superior part of middle occipital gyrus. However, hardly any paths reached supramarginal gyrus (0% in RH; 7% in LH), another IPL region implicated to be functionally connected to vOT.

Table 5.1: Single seed analyses. The number (and percentages) of cases in which IFOF and ILF were identified are displayed for each hemisphere and each rostro-caudal mask. Also shown are the termination points of the paths and their number and percent of cases in which these were found.

Seed masks	RH			LH		
	ant (N = 9)	mid (N = 9)	post (N= 9)	ant (N = 10)	mid (N = 9)	post (N = 10)
Identified fasciculus						
IFOF	9/9 (100%)	8/9 (89%)	7/9 (78%)	9/10 (90%)	7/9 (78%)	9/10 (90%)
	24/27 (89%)			25/29 (86%)		
ILF	9/9 (100%)	9/9 (100%)	8/9 (79%)	10/1 (100%)	8/9 (89%)	9/10 (90%)
	26/27 (96%)			27/29 (93%)		
Terminations						
Orbitofrontal cortex	7/9 (78%)	7/9 (78%)	7/9 (78%)	7/10 (70%)	7/9 (78%)	9/10 (90%)
	21/27 (78%)			23/29 (79%)		
Pars Opercularis	1/9 (11%)	2/9 (22%)	0/9 (0%)	0/10 (0%)	1/9 (11%)	3/10 (30%)
	3/27(11%)			4/29 (14%)		
Pars Triangularis	0/9 (0%)	1/9 (11%)	0/9 (0%)	0/10 (0%)	1/9 (11%)	6/10 (60%)
	1/27 (3%)			7/29 (24%)		
Angular gyrus	5/9 (56%)	5/9 (56%)	5/9 (56%)	4/10 (40%)	4/9 (44%)	6/10 (60%)
	15/27 (56%)			14/29 (48%)		
Supramarginal gyrus	0/9 (0%)	0/9 (0%)	0/9 (0%)	0/10 (0%)	0/9 (0%)	2/10 (20%)
	0/27 (0%)			2/29 (7%)		
Middle occipital gyrus	9/9 (100%)	8/9 (89%)	8/9 (89%)	10/1 (100%)	9/9 (100%)	10/1 (100%)
	25/27 (93%)			29/29 (100%)		
Temporal pole	9/9 (100%)	9/9 (100%)	9/9 (100%)	10/1 (100%)	8/9 (89%)	9/10 (90%)
	27/27 (100%)			27/29 (93%)		

In addition to the long-range association fibres, a set of local U-fibres were also identified from each vOT seed mask. These local U-shaped paths become identifiable at a higher threshold since it eliminates paths with lower connection probability and leaves U-fibres, which have higher connection probability to the vOT seeds (Figure 5.7b). More specifically, they linked adjacent vOT seed masks and were located lateral to the ILF, consistent with previous studies (Catani et al., 2003; Mandonnet et al., 2009). This local connection is part of the occipito-temporal projection system, a series of short U-fibres connecting temporal lobe with the visual association cortex (Catani et al., 2003; Tusa and Ungerleider, 1985), distinct from ILF (Catani et al., 2003).

In summary, the first set of analyses that seeded vOT robustly identified the ILF that terminated in the anterior temporal pole at the anterior end as well as in the middle occipital regions at the posterior end. It also demonstrated the IFOF in most cases, terminating in the orbitofrontal cortex. Moreover, local U-fibres linking the adjacent vOT sub-regions were also found. However, despite claims that vOT provides input to higher order association areas such as inferior parietal lobule or Broca's area during reading, evidence for direct anatomical connections between these two regions was not strong. There are two potential explanations: it is possible that the current, undirected tractography lacked sensitivity to detect these paths or, it may be that vOT is indirectly connected to these regions. For example, the functional link between vOT and SMG (Kawabata Duncan et al., 2013; van der Mark et al., 2011) may be due to a common driving input such as the middle occipital gyrus rather than a direct path between them. Therefore additional analyses were conducted to investigate these possibilities. These took advantage of the fact that there were strong prior hypotheses concerning the anatomical pathways to enhance detection sensitivity using a hypothesis-driven two-ROI approach.

Two-ROI analyses

This set of two ROI analyses investigated the more specific hypotheses that brain regions functionally connected to vOT such as the supramarginal gyrus (SMG), angular gyrus (ANG) and Broca's area (pars opercularis (POp) and pars triangularis (PTr)) are linked to vOT via direct cortico-cortical connections.

vOT and Broca's area

Many studies report functional connectivity between vOT and Broca's area (e.g., Booth et al., 2002; Mechelli et al., 2005) and thus it is possible that this is mediated by direct anatomical connections linking these regions via IFOF. Even though my first analysis robustly identified the IFOF, there was only weak evidence of a path linking vOT to Broca's area. Consequently, a second set of analyses were conducted to investigate this further. The two main sub-regions of Broca's area, POp and PTr, were investigated separately and the results provided additional evidence for a pathway linking vOT to both POp and PTr. On the whole, the trajectory of the pathways linking vOT to POp and PTr were very similar, both taking two separate courses from vOT. The first pathway had the same trajectory as that of the IFOF described earlier (Figure 5.9). It proceeded from vOT anteriorly towards the orbitofrontal cortex, where it appeared to merge with the second path (see below) at the level of either the anterior or middle short gyrus of the insula before reaching separate parts of POp. Specifically, the IFOF terminated in its ventral portion while the second pathway terminated in its dorsal portion (Figure 5.10). For PTr, they travelled slightly further on and merged at a slightly more anterior location, dorsomedial to the junction of the anterior and superior peri-insular sulci. In some cases, it then split into PTr and anterior, posterior or lateral orbital gyri. As with POp, the IFOF terminated in the ventral and the second path terminated in the dorsal portions of PTr. The second pathway proceeded from vOT anteriorly to join the TPIFA, from which it curved antero-superiorly forming an arch. It then continued anteriorly towards the frontal lobe at the level superior to the claustrum (consistent with the superior longitudinal fasciculus, SLF) until it merged with the IFOF before terminating in the dorsal portions of POp or PTr. There were no systematic differences across the rostro-caudal vOT seeds or hemispheres although there was a trend that these pathways to be found more in the posterior vOT (Table 5.2). Dissection studies have demonstrated that the superficial layer of IFOF terminates at inferior frontal gyrus (POp and PTr) at the anterior end and at posterior basal temporal regions including fusiform gyrus at the posterior end (Martino et al., 2010a; Sarubbo et al., 2011) and current data showing IFOF between vOT and Broca's area are consistent with this. In contrast, the second pathway seemed to be a combination of distinct fasciculi given its trajectory and is thus likely to be a false positive. This path was probably recovered, not as a true path between these regions, but as a result of the posterior part of IFOF merging with SLF to around the occipital horn of the lateral ventricle where numerous white matter bundles intersect (i.e., the TPIFA Martino et al., 2013a). This view was further strengthened by the finding that here were no cases where only the second pathway was recovered in the absence of the IFOF at the given threshold.

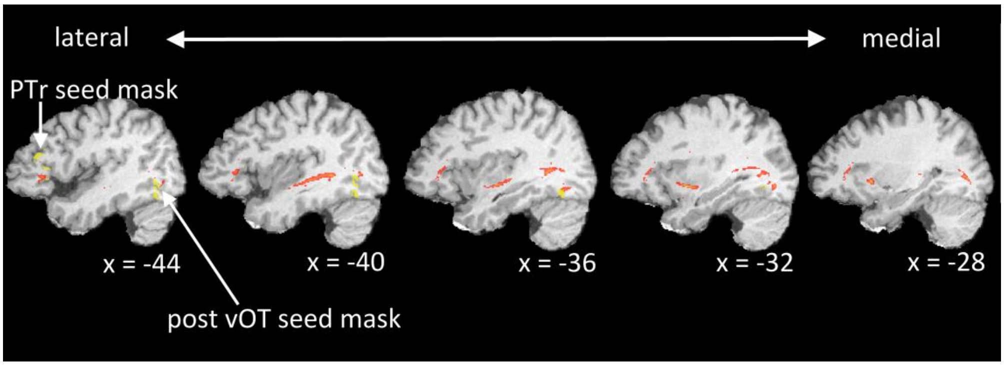

Figure 5.9: Tractography results (in red to yellow) for between vOT and posterior PTr in one subject (standard apace). The characteristic double-hook-like trajectory of the IFOF can be identified as it moves from lateral (x = –44) to medial slices (x = –28)

Table 5.2: The number and the percentages of cases showing paths connecting vOT and Broca's area via IFOF.

	RH			LH		
	ant	mid	post	ant	mid	post
Pars Opercularis (30%)	0/9	3/9	4/9	1/10	3/9	6/10
	7/27(26%)			10/29 (34%)		
Pars Triangularis (44.5%)	3/9	5/9	4/9	2/10	4/9	7/10
	12/27(44%)			13/29(45%)		

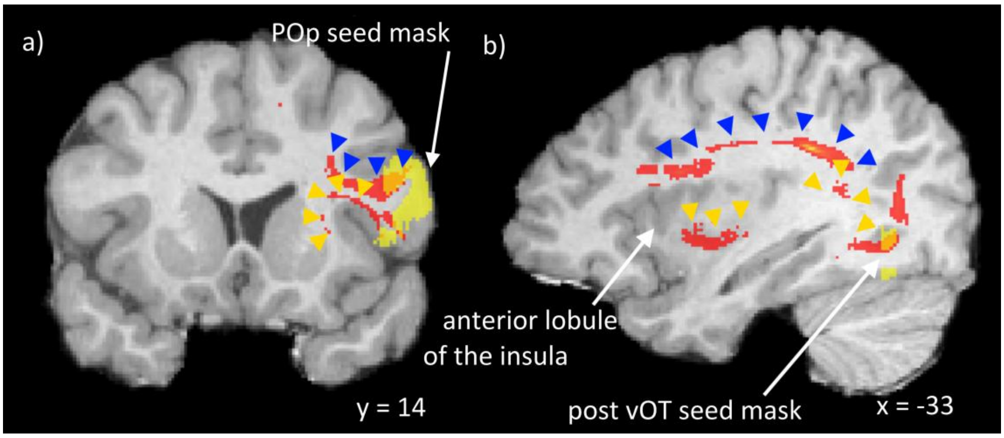

Figure 5.10: The IFOF (yellow arrow heads) and the second pathway (blue arrow heads), reaching POp separately, from the tractography results (in red to yellow) for posterior vOT and POp in one subject (standard apace). a) The paths entering or exiting the POp mask (yellow) separately; the IFOF in its ventral portion and the second pathway in its dorsal portion. Shown on a coronal slice, at y = 14. b) Trajectory of IFOF and SLF on the parasagittal slice at x = –33.

Given the strong possibility that the second pathway, which merged with the IFOF in the vicinity of Broca's area was a false positive, it was crucial to confirm that the IFOF independently reached Broca's area. Therefore, I investigated this by running the same analyses with an additional exclusion mask placed where the SLF passed through in order to block the second pathway (Figure 5.11). If the IFOF was indeed reaching Broca's area via the SLF, these analyses would eliminate the IFOF as well as the second pathway since any paths between vOT and Broca's area that reached the exclusion mask would be discarded. Consistent with my earlier visual inspection, these analyses still recovered the IFOF, but not the second pathway, confirming the IFOF between vOT and Broca's area, independent of the SLF.

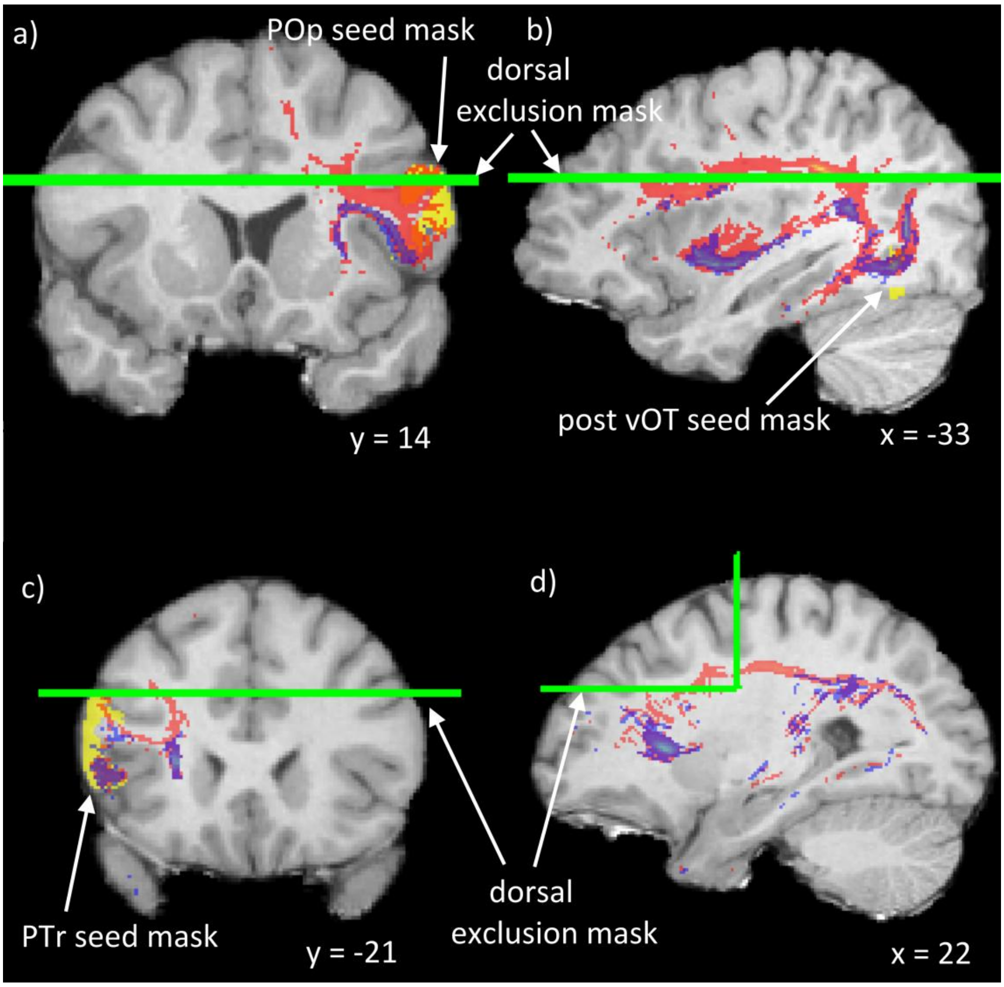

Figure 5.11: Comparison of the tractography results between vOT and Broca's area with (in blue to light blue) and without (in red to yellow) an additional dorsal exclusion mask. While the original tractography (in red to yellow) resulted in both the IFOF and the second pathway via the SLF, the new one with the additional dorsal exclusion mask recovered only the IFOF, confirming that this pathway independently reached Broca's area. These are shown with no thresholding in order to demonstrate that this difference is not due to the arbitrary threshold. a) & b): The tractography results for posterior vOT and POp in LH. Only the IFOF reaches the POp seed mask (yellow) with the dorsal exclusion mask placed at z = 24 to 26 (in green). c) & d): The tractography results for anterior vOT and PTr in RH. The dorsal exclusion mask was placed at z = 30 to 31 and y = –7 to 8, forming a 90-degree angle seen in a parasagittal slice.

vOT and angular gyrus (ANG)

Connections between vOT and ANG have been implicated since the time of Dejerine (1892). In the first set of analyses, paths were found between these two regions in approximately half of the cases. However, since these paths pass through the TPIFA (Martino et al., 2013a) where many white matter bundles intersect, it was not easy to visually delineate their exact course and determine whether the paths were genuine or a false positive. The two-ROI analyses between vOT and ANG increased evidence for paths connecting these two regions, particularly in the left hemisphere (Table 5.3).

Table 5.3: The number and the percentages of cases showing paths connecting vOT and angular gyrus.

	RH			LH		
	ant	mid	post	ant	mid	post
Angular gyrus (82%)	8/9	7/9	5/9	10/10	7/9	9/10
	20/27(74%)			26/29 (96%)		

There were two different courses to the recovered paths, the anterior and the posterior ones (Figure 5.12a). Both paths ran vertically between vOT and ANG. The anterior one, identified at around y = –48 to –55 travelled anteriorly from vOT into the ILF and superiorly into the IFOF and/or the arcuate fasciculus (AF). At around the deep white matter of the ascending posterior segment of the parallel sulcus, it then branched dorsolaterally into the ANG in the form of U-fibres linking the adjacent areas within ANG (Figure 5.12b). Given this trajectory, this path is likely to be a combination of these separate fasciculi, rather than an independent path. In contrast, the posterior path, identified at around y= –64 to –71, was continuous. From the fusiform and inferior temporal gyri in vOT it travelled superiorly towards the posterior portion of angular gyrus (Figure 5.12c). The location and the trajectory of this path is consistent with the vertical (or perpendicular) occipital fasciculus (VOF) of Wernicke, described by early anatomists (Crosby et al., 1962; Déjerine, 1895; Gray and Lewis, 1918; Greenblatt, 1973; Greenblatt, 1976; Herrick, 1915; Larsell, 1951; Wernicke, 1900, 1903) and reported in recent DTI studies (Wakana et al., 2004; Yeatman et al., 2013). This white matter bundle connects fusiform gyrus with posterior parietal region passing through the front part of the occipital lobe (Crosby et al., 1962; Gray and Lewis, 1918). However, the number of cases where this path was found in the current data was small (Table 5.4). It was found in the left hemisphere more (24%) than the right (7%).

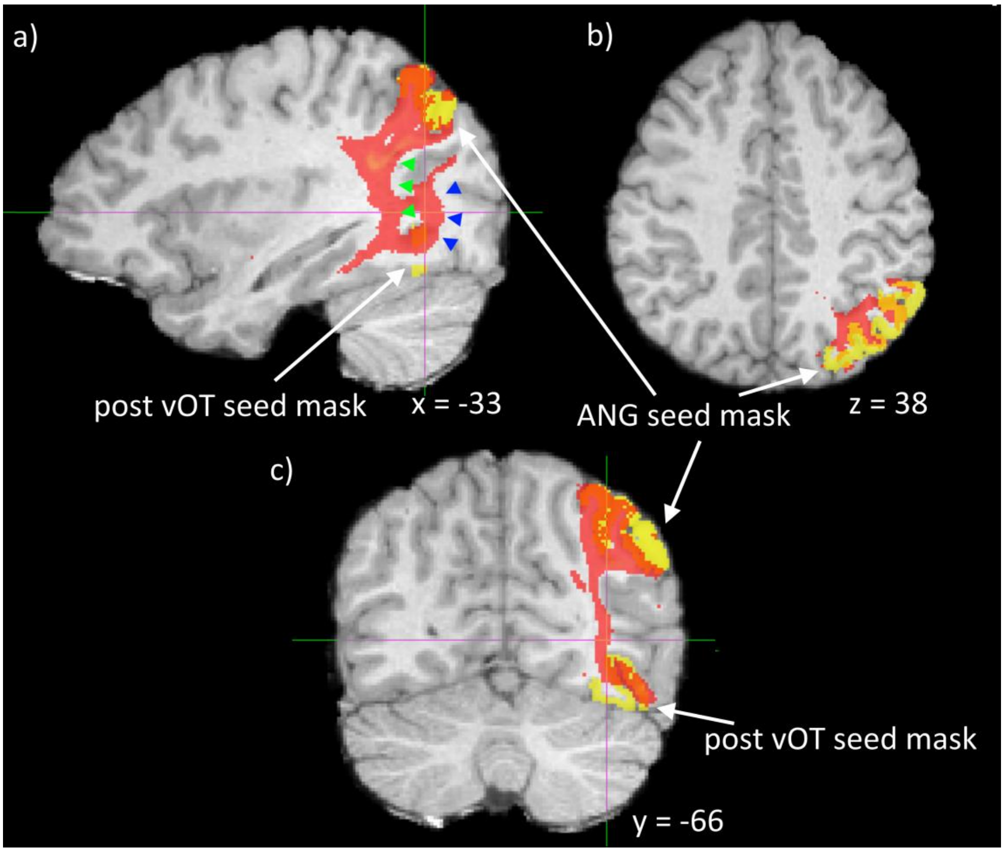

Figure 5.12: Paths (red to yellow) recovered between angular gyrus and posterior vOT in one subject (standard space) in left hemisphere. Also shown are the ANG seed mask and the posterior vOT seed masks in yellow. a) There were two different courses to the recovered paths: the anterior one at around y = –48 to –55 (light green arrow heads) and the posterior one at around y = –64 to –71 (blue arrow heads), shown on a parasagittal slice at x = –33. b) The anterior path reaching ANG in the form of U-fibres linking the adjacent areas in ANG at z = 38. c) The posterior path. This continuous vertical path corresponds to the posterior, rather than the anterior, path seen on the parasagittal slice (indicated by the green crosshair).

Table 5.4: The number and the percentages of cases showing VOF between vOT and ANG

RH			LH		
ant	mid	post	ant	mid	post
0/9	1/9	1/9	4/10	2/9	1/10
2/27(7%)			7/29(24%)		

In order to ensure that this is indeed an independent vertical path, these results were inspected on the colour-coded orientation map using the original DW-MRI images from the same subject (Figure 5.13). This type of display allows the diffusion vectors to be represented with the RGB coding: left-right (Red), anterior-posterior (Green) and superior-inferior (Blue) orientations and the intensity corresponds to the fractional anisotropy (FA, brighter = higher). These maps, however, normally show the images of the entire brain making it hard to identify the exact position of the paths recovered by the tractography. Therefore the tractography results were registered to the FA image and used to multiply the diffusion vector images, in order to make a “cut out” image of the tractography results on the colour-coded orientation map. This displayed the posterior path primarily in blue (the superior-inferior direction), separate from the IFOF and the AF, which are in green (the anterior-posterior orientation). This indicates that the posterior path is independent of adjacent IFOF and/or AF, consistent with Wakana and colleagues (2004), unlike the anterior path which corresponds to the green on this map.

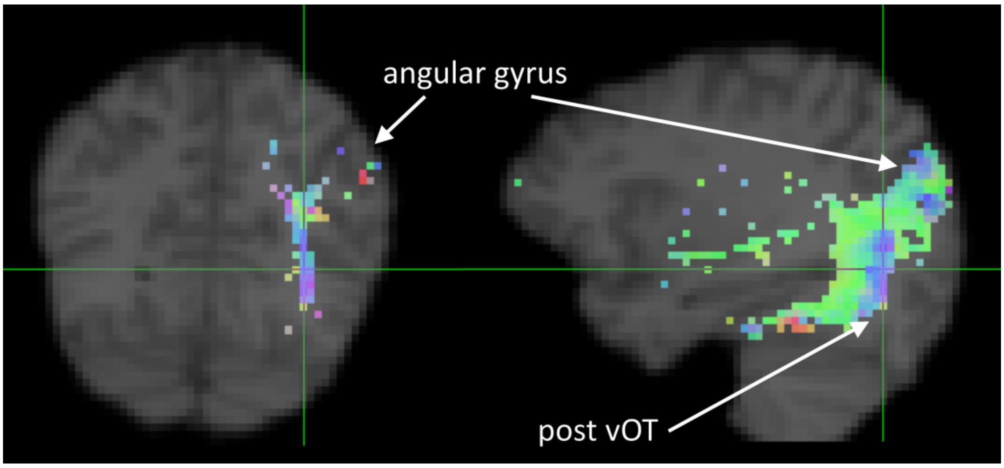

Figure 5.13: The colour-coded orientation map showing the results for the two-ROI analyses for posterior vOT and ANG in one subject overlaid on the subject's native structural image in diffusion space. The intensity corresponds to the fractional anisotropy (brighter = higher) and the colours indicate the direction of the diffusion: left-right (red), anterior-posterior (green) and superior-inferior (blue) orientations. The posterior path can be seen primarily in blue indicating that this runs vertically, separate from IFOF and AF, which are in anterior-posterior orientation (green).

vOT and supramarginal gyrus (SMG)

As with ANG, functional connectivity between vOT and SMG has also been reported (e.g., Kawabata Duncan et al., 2013; van der Mark et al., 2011). Direct pathways between these regions were hardly observed in the first analyses (0% in RH and 7% in LH). The evidence for such a pathway was still difficult to obtain even with two-ROI analyses (Figure 5.14). The recovered paths first travelled from vOT antero-superiorly to join the ILF. Then it proceeded postero-superiorly along the ILF until it reached the occipital horn of the lateral ventricle. From there it briefly travelled antero-inferiorly along the curve of the IFOF before joining the AF/SLF. Finally it took a sharp turn dorsolaterally into SMG. Although these paths were found in 52% in RH and 34% in LH (Table 5.5), they are most likely to be a false positive - probably the result of separate fasciculi combined by the tractography algorithm, given the trajectory.

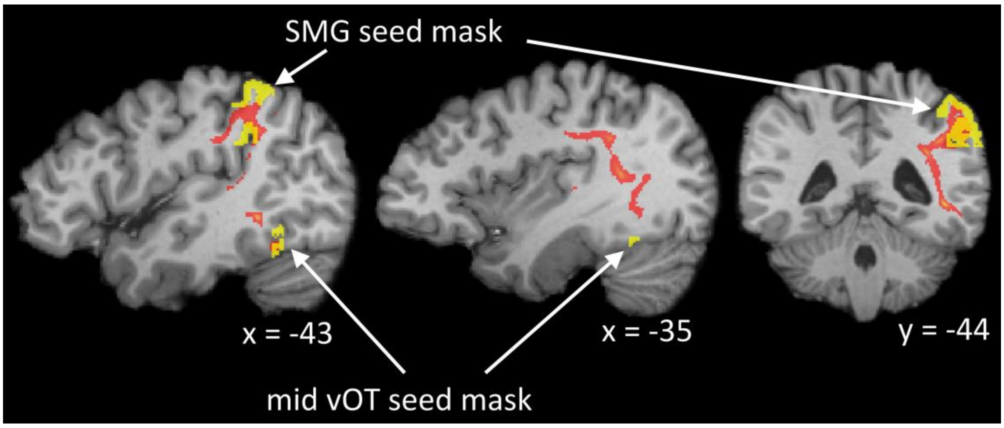

Figure 5.14: Tractography results for middle vOT and SMG (in red to yellow). Also shown are the SMG seed mask and the middle vOT seed masks in yellow.

Table 5.5: The number and the percentages of cases showing paths between vOT and SMG

RH			LH		
ant	Mid	post	ant	mid	post
5/9	6/9	3/9	2/10	2/9	6/10
14/27(52%)			10/29 (34%)		

In order to investigate this, these paths were also inspected in the colour-coded orientation map. As suspected, they demonstrated a mixture of colours indicating several different orientations (Figure 5.15). This suggests that, unlike the case above with vOT and ANG, this path is indeed a combination of various paths recovered by the tractography algorithm. This interpretation is also supported by the fact that this path is located in the TPIFA, where seven known white matter tracts intersect. Taken together, these observations are sufficient evidence that this path is indeed a false positive.

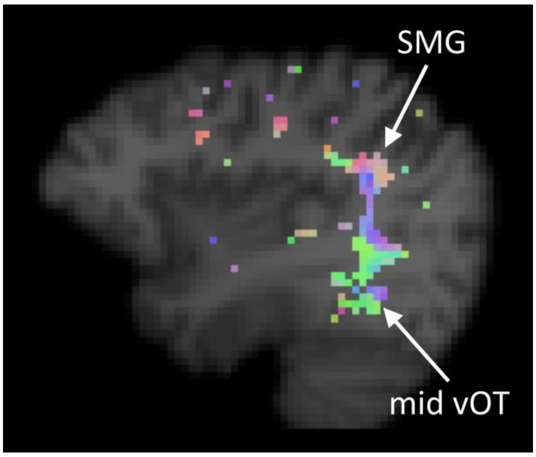

Figure 5.15: The colour-coded orientation map showing a typical results for the two-ROI analyses for vOT and SMG in one subject, overlaid on the subject's native structural image in diffusion space. The path between the two regions is in multiple of colours indicating that this is comprised of different white matter bundles orientated in different directions.

In summary, the second set of analyses produced reasonable evidence to suggest that there was a direct anatomical link between vOT and Broca's area via the IFOF. Moreover, it provided further support for the existence of the VOF between vOT and ANG. However, there was no evidence for the direct pathways between vOT and SMG.

An alternative hypothesis for the functional connectivity between SMG and vOT

Given the lack of paths between vOT and SMG, it was then hypothesised that middle occipital gyrus (MOG) may be the common input source for SMG and vOT, which could explain the presence of functional connectivity between vOT and SMG (Kawabata Duncan et al., 2013; van der Mark et al., 2011) in the absence of direct anatomical connection between them. Therefore, a third set of analyses investigated the link between MOG and SMG, as well as MOG and vOT.

MOG and SMG

First, two-ROI analyses with SMG and middle occipital gyrus (MOG) were run. The results showed robust paths between these regions in most cases (9 out of 10 subjects). These paths run between the lateral surface of SMG and that of MOG via the deep white matter where they join the long horizontal part of the superior longitudinal fasciculus (SLF) (Figure 5.16), which connects frontal, temporal, parietal and occipital lobes (Catani et al., 2002; Crosby et al., 1962; Türe et al., 2000). These paths are consistent with the short U-fibres that are part of the SLF, which connect parieto-occipital cortex (Catani et al., 2002; Crosby et al., 1962).

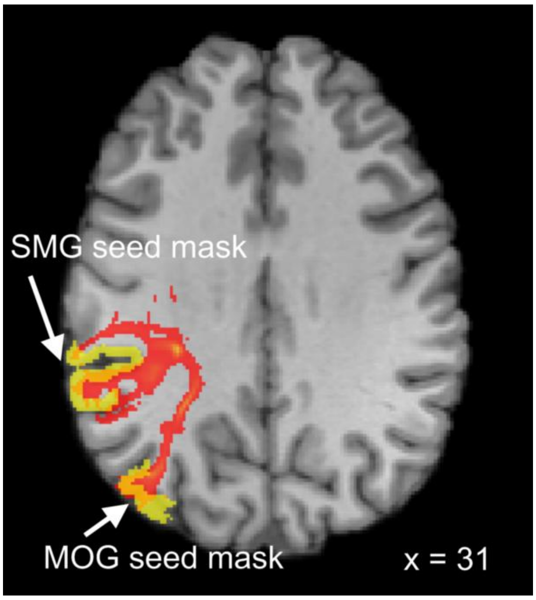

Figure 5.16: Results for the two-ROI analyses between supramarginal gyrus and middle occipital gyrus in RH from a single subject (standard space), shown in an axial slice at z = 31. The two seed masks are shown in yellow (SMG = supramarginal gyrus, MOG = middle occipital gyrus).

Indeed, this possibility became more probable when these paths were overlaid with an additional MOG single seed analysis I ran. Consistent with the literature (Catani et al., 2002; Crosby et al., 1962), these paths lie lateral to longer paths running from the occipital to the parietal and then to the frontal cortices (Figure 5.17).

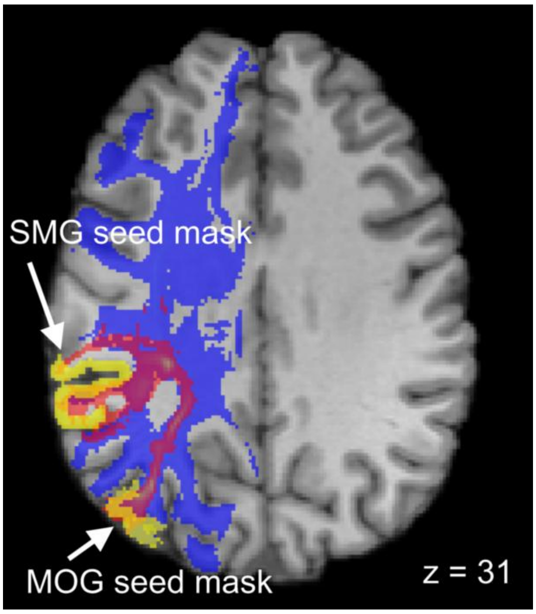

Figure 5.17: Overlay of the tractography results for the MOG – SMG (in red to yellow) and a single seed MOG (in blue to light blue) analyses. The MOG – SMG path is lateral to the long paths running from the occipital to the parietal and then to the frontal cortices, which corresponds to the SLF.

MOG and vOT

Even though strong pathways between vOT and MOG were found in the first analyses, in order to confirm that it is the same parts of MOG that send projections to both SMG and vOT hence the common source of input, additional two ROI analyses between vOT and MOG were run. Strong pathways between these regions were found in all cases. There were no systematic differences across the rostro-caudal vOT seeds or hemispheres. From each vOT seed mask, the paths travelled posterosuperiorly joining the ILF towards the middle occipital regions. Having passed the level of the posterior vOT seed mask (approximately y = –68 to –71), they then split into inferior and superior parts of MOG. Although these paths are consistent with the ILF, they also correspond to the posterior end of the IFOF. Since these fibres merge posteriorly (Jellison et al., 2004), it is not possible to determine whether these paths between MOG and vOT belong to the IFL or IFOF. Nevertheless, comparing these analyses and the SMG & MOG analyses, it can be clearly seen that these paths overlap in the MOG seed mask (Figure 5.18).

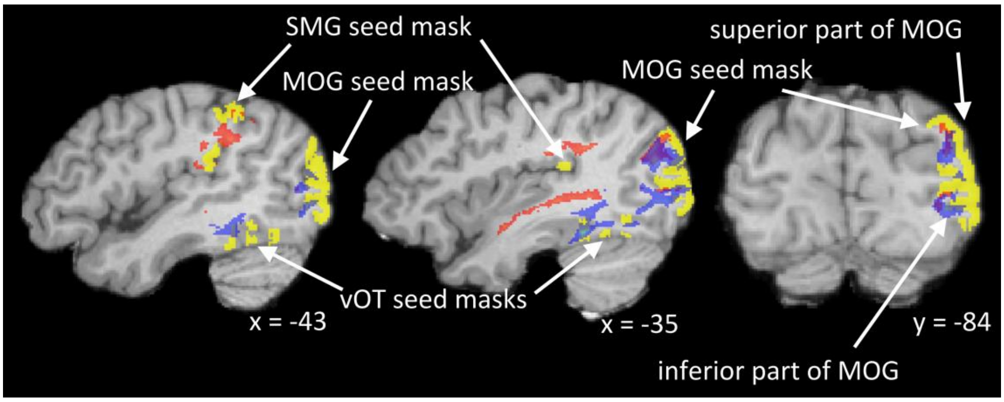

Figure 5.18: The tractography results for SMG and MOG (in red to yellow); and anterior vOT and MOG (in blue to light blue), showing the MOG terminations for these pathways overlap in both the superior and inferior parts.

To conclude, the third set of analyses revealed robust pathways between MOG and SMG as well as vOT. In addition, it was confirmed that these distinct paths were traced to the same parts of MOG, providing support for the hypothesis that MOG may be the common input source for SMG and vOT.

5.4. DISCUSSION

The aim of this study was to investigate the cortico-cortical pathways linking vOT with other regions within the language system and compare these to the functional connections seen in visual word recognition. My main findings were: i) a clear path along the inferior longitudinal fasciculus linking vOT with the temporal pole extensively and also with the middle occipital gyrus (MOG); ii) a set of local U-fibres linking adjacent vOT regions; iii) a path via the inferior fronto-occipital fasciculus (IFOF) linking vOT with Broca's area; iv) no direct path linking vOT and SMG but the same region of MOG linked to both vOT on the ventral surface and SMG more dorsally; and v) a path via the vertical occipital fibres linking vOT with the posterior angular gyrus. In addition, there were no systematic differences in anatomical connectivity across the rostro-caudal sub-regions of vOT. Finally, although there were numerical differences in the probability values of the recovered paths between the left and right hemispheres, no systematic hemispheric differences were found from the current analyses.

Consistent with previous reports (Catani et al., 2003; Crosby et al., 1962), the current study robustly identified the inferior longitudinal fasciculus (ILF) linking vOT with the temporal pole anteriorly and with middle occipital gyrus posteriorly. In addition, a set of local U-fibres linking the adjacent areas along the ILF was also consistently identified. These U-fibres, a part of the occipito-temporal projection system (Catani et al., 2003; Tusa and Ungerleider, 1985), are widely considered to be distinct from the ILF (Catani et al., 2003; Martino et al., 2011). However, the existence of the ILF as a long association fibre bundle, independent of the sets of local U-fibres, has also been questioned (Tusa and Ungerleider, 1985). Tusa and Ungerleider (1985) claimed that their dissection of human and monkey brains together with autoradiography experiments showed that only local short fibres connected the occipital and temporal lobes. The ILF (long or short) is a part of the visual object recognition pathway (Mishkin et al., 1983; Ungerleider and Mishkin, 1982) and involved in reading and other language functions (Catani and de Schotten, 2012). In any case, this ventral pathway also seems to provide an indirect anatomical link between vOT and Broca’s area (Gil-Robles et al., 2013; Ashtari, 2012; Mandonnet et al., 2007) in that it connects vOT with the anterior temporal lobes which are then linked to Broca’s area via the uncinate fasciculus (Ebeling and Cramon, 1992; Kier et al., 2004; Türe et al., 2000).

Another ventral pathway, the inferior fronto-occipital fasciculus (IFOF), was also identified in the majority of cases, primarily linking vOT with the orbitofrontal cortex. Subsequent two-ROI analyses that seeded vOT and Broca’s area confirmed that this pathway also terminated in Broca’s area. Both of these termination regions are consistent with previous dissection studies of this fasciculus (Martino et al., 2010a; Sarubbo et al., 2011). This white matter tract is considered to be the direct pathway connecting the occipital, posterior temporal and orbito-frontal cortex (Catani and Mesulam, 2008; Mandonnet et al., 2007) and its involvement in semantic processing has been demonstrated by intraoperative direct stimulation (Duffau et al., 2005; Mandonnet et al., 2007). Interestingly, a recent study with intraoperative subcortical stimulation demonstrated a double dissociation between reading/visual (symbol) recognition and picture naming in these two ventral pathways (Gil-Robles et al., 2013). Specifically, direct stimulation on the IFOF induced a transient impairment of picture naming (semantic paraphasia), sparing reading and symbol recognition whereas disturbances of reading (letter-by-letter reading) and symbol recognition (unable to describe the identity or the purpose) were elicited by stimulation on the ILF with no effect on picture naming. The results demonstrate that these two pathways are functionally dissociable, but the interpretation requires a degree of caution. Picture naming and word reading place slightly different demands on their cognitive components (i.e., visual pattern recognition, semantic processing, phonological processing, etc) and consequently engage both common and distinct neuroanatomical systems (Bookheimer et al., 1995; Mechelli et al., 2007; Price et al., 2006; Vandenberghe et al., 1996). Clearly, further studies are needed to investigate which cognitive processes are (at least partially) mediated by these fasciculi and also how these fasciculi contribute to specific cognitive processes.

Although a previous study demonstrated different functional connections between sub-regions of Broca's area and rostro-caudal sub-regions of vOT (Mechelli et al., 2005), the tractography results did not show these patterns. The pathways between the sub-regions of Broca's area (POp and PTr) and of vOT were in fact very similar. POp and PTr have been implicated in phonological and semantic processing respectively (Devlin et al., 2003; McDermott et al., 2003; Roskies et al., 2001). It is possible that the sub-regions of vOT and Broca's area share the same set of complex anatomical connectivity and such differences in functional connectivity are the consequence of transient neuronal couplings between them, which reflects a greater involvement of a particular processing required for certain types of word processing. Another possibility is that the differential functional connections reported might reflect indirect anatomical connections between the two regions mediated by a different third region, rather than direct connections that might not differ across the sub-regions. Functional connections do not necessarily entail direct anatomical connections and direct anatomical connections do not deny the existence of indirect ones. Anatomical connections are likely to be much more complex than functional connections that can be demonstrated with dynamic causal modelling (DCM) since DCM can only test the connections between the regions included in the model. Different functional connections for the sub-regions might reflect separate phonological and semantic routes between vOT and Broca via a different third region, associated with phonological and semantic processing respectively. Broca's area has been implicated in unification of phonology, semantics as well as syntax (Hagoort, 2005). It is therefore conceivable that a region that plays such a complex functional role has very rich anatomical connections with many cortical regions.

Despite the functional connections previously implicated between vOT and SMG, a region associated with phonological processing (Joubert et al., 2004; Moore and Price, 1999; Stoeckel et al., 2009), no direct pathway was found linking these regions even with the two ROI analyses. This result is in sharp contrast to Epelbaum and colleagues (2008), who reported that these regions were linked via the arcuate fasciculus (AF). As with the current study, they used individual's vOT regions as the seed masks and the trajectory of this pathway was in fact very similar (Epelbaum et al., 2008, Figure 3, p.967) to the one that I indentified from the two-ROI analyses that seeded vOT and SMG. This path is located in the TPIFA, where seven white matter tracts are known to intersect. As a result, the area is one where tractography is particularly prone to false positives and the interpretation of the results requires a caution. Indeed, the colour-coded orientation map (Figure 5.15) showed this "pathway" to be made up of various orientations, suggesting that there is not a single path linking vOT and SMG. In addition, previous studies agree that the AF does not cover the ventral part of the brain (e.g., Catani and de Schotten, 2012; Crosby et al., 1962; Dick and Tremblay, 2012; Glasser and Rilling, 2008; Martino et al., 2013b; Türe et al., 2000). Although Epelbaum and colleagues attribute this path to mediating top-down influence from the parietal lobe to vOT, the current analyses do not support their view.

While there was no evidence of a direct path, there was strong evidence for a common area of MOG projecting to both vOT and SMG in parallel. The analyses that seeded MOG and SMG found short U-fibres that are part of the SLF, which connect parieto-occipital cortex (Catani et al., 2002; Crosby et al., 1962). In addition, subsequent analyses that seeded MOG and vOT confirmed that the same parts of MOG send projections to both vOT and SMG. This can explain the functional connections between vOT and SMG in the absence of the direct anatomical connections because they are driven by a common input. The finding that both are linked to the same parts of MOG is consistent with the parallel visual pathways seen in non-human primates (Merigan and Maunsell, 1993; Mishkin et al., 1983; Ungerleider and Mishkin, 1982). This suggests that the visual information is directly sent to SMG from the occipital cortex via these dorsal pathways.

An important implication of this is that vOT is not the first bifurcation point where the visual input splits to the dorsal and ventral routes. This suggests that despite the consensus that vOT plays an important role on reading, it is not always necessary. If vOT is not the first and central node in the reading network that provides input to the rest of the reading system, it may be possible for reading to be sustained following damage to vOT because the dorsal route could still function independently albeit with some difficulty. This is consistent with the dual neural circuit model of reading, which is based on observations on patients with alexia in Japanese kana and/or kanji (Iwata, 1984, 2011; Sakurai, 2004). This hypothesis proposes that the early visual analysis occurs in the primary visual cortex and the visual association cortices including MOG, from which the visual information takes dual routes: a dorsal routes (including SMG) and a ventral routes (including vOT) during reading. Under normal circumstances, the two routes interact to aid fluent reading. Therefore damage to either route could cause some degree of impairment, which results from the damaged path (or nodes on the path) itself and also from restricted interaction with the other intact route, depending on the extent and the exact location of the damage. This can explain why temporo-parietal, instead of occipitotemporal, lesions can lead to alexia (Philipose et al., 2007; Warrington and Shallice, 1980) and why lesions in the left occipitotemporal region do not always result in reading impairment (Hillis et al., 2005). It is also consistent with recent DCM studies investigating multiple reading routes that found that reading does not always involve vOT (Richardson et al., 2011; Seghier et al., 2012).

In addition, it has an important implication for some cognitive models of reading that stipulate a single region corresponding to "orthography". If visual information splits at the level of MOG, then anatomically one possibility is that this early visual area is such a region. However, this is unlikely given that in non-human primates, the selectivity of the neurons in the secondary visual cortex (V2), in which middle occipital gyrus is located, includes contours, texture, size, orientation and spatial frequency (Anzai et al., 2007; Hegdé and Van Essen, 2000). In humans, this area is also implicated in processing of shape-edges in illusory contours (Seghier and Vuilleumier, 2006). The LCD model, a reading-specific account of vOT, also proposes that the neurons in V2 detect local contours such as letter fragments (Dehaene et al., 2005), based on neurophysiological studies of object recognition (Riesenhuber and Poggio, 1999; Rolls, 2000). In other words, information processed here is less complex than orthography. Collectively, it strongly suggests that there is not a single region specific to orthography that exclusively feeds written information to the rest of the reading network.

Finally, the current study identified a possible direct anatomical link between vOT and ANG, namely the vertical occipital fasciculus (VOF) of Wernicke. This pathway was also reported in a recent tractography study of vOT (Yeatman et al., 2013). However, there are a number of important methodological differences between the current results and those of Yeatman and colleagues that are worth discussing. First of all, Yeatman and colleagues used deterministic tractography. As discussed in Chapter 2, deterministic approach is less reliable than probabilistic algorithm and unsuitable for tracing through areas with low diffusion anisotropy such as in grey matter. Moreover, Yeatman and colleagues placed the seed mask in the white matter. Oddly, its exact location was not mentioned in the paper. As the authors themselves noted, with such methods it is simply not possible to determine whether the recovered paths terminated in the cortical grey matter. Instead, it demonstrates, at best, that the pathways "pass within close proximity" (Yeatman et al., 2013, p. 151) to the cortical region of interest. In contrast, the current study employed probabilistic tractography, which is able to trace through grey matter and the seed masks were placed in the cortical grey matter. In addition, the current study used 60 diffusion directions whereas Yeatman and colleagues used 12. Furthermore, the current study used the tracing step of 0.5mm, in comparison to Yeatman and colleagues' 1mm. All of these differences reflect in the uncertainty of tractography and the ability to determine the cortical terminations of the recovered paths. Therefore, the current study provides evidence that vOT and ANG were anatomically linked via the VOF, rather than suggesting that there is a pathway passing by "in close proximity" to vOT and ANG.

This fasciculus is relatively unknown, which is surprising given that its reported cortical terminations include vOT and ANG (Crosby et al., 1962; Déjerine, 1895; Gray and Lewis, 1918; Greenblatt, 1973; Larsell, 1951; Tilney and Riley, 1921). Greenblatt (1973) postulated that the VOF might be a pathway involved in reading because it was the only known white matter tract that connected the inferomedial occipital and parietal cortex. However, he later questioned the true identity of this fasciculus (Greenblatt, 1976), pointing out that the location and connections were the same as what Krieg (1966; 1973) depicted and described as the extreme sagittal stratum. Although this fasciculus has re-emerged in recent tractography studies (Oishi et al., 2008; Wakana et al., 2004; Yeatman et al., 2013), anatomical nomenclature is built on consensus and deeply related to anatomical consistency (Mori, personal communication). Clearly, further studies are required to investigate the existence, location and functional role of this fasciculus.

Conclusion

The current study investigated the cortico-cortical, intra-hemispheric connectivity of vOT that might mediate the functional connections implicated between vOT and other regions within the language system in previous studies. The results showed both direct and indirect pathways between vOT and Broca's area. Moreover, vOT was found to be anatomically linked to ANG via the VOF. While no pathway was found between vOT and SMG, the same parts of MOG appeared to send projections to both SMG and vOT, which accounts for the functional connections in the absence of direct anatomical pathways between these regions. Because tractography cannot reveal directionality, it is not possible to determine whether the pathways recovered by the tractography algorithm are feed-forward or feedback projections. However, most cortical pathways are reciprocal (Kravitz et al., 2013; Van Essen and Maunsell, 1983) and there are more backward afferents than forward efferent connections (Friston, 2002, 2003, 2005; Friston and Price, 2001). Therefore, it is a reasonable assumption that the pathways recovered in the current study are bi-directional. Altogether, these results are consistent with the parallel cortical visual pathways seen in other primates and support a modified neuroanatomical model of reading with multiple routes, not only from visual to higher-order language areas but also within the language areas.

By delineating these pathways anatomically, we gain a deeper understanding of how the nodes of the reading network may interact and can begin to map a basic circuitry diagram. Such a diagram is important since it enables us to understand the nature of the inter-regional interactions such as functional connectivity and how the system may degrade following a focal damage. Moreover, it provides a better understanding of degenerate systems (i.e., alternative systems that can perform the same functions, Price and Friston, 2002) and lays a foundation for constructing biologically-informed computational models of reading that would exceed the current cognitive models.

6. GENERAL DISCUSSION

The aim of this thesis was to empirically evaluate a central claim of the Interactive Account of vOT functioning that this area responds not only to bottom-up processing demands of the visual stimuli but is also influenced by automatic, top-down non-visual processing demands. I evaluated this claim both functionally and anatomically. The results provide strong support. Specifically:

- Activation in vOT was significantly modulated by stimulus changes during the same task when low level visual features were held constant (Chapters 3 & 4), indicating that vOT activation is influenced by top-down processing of non-visual aspects of the stimuli; and it interactively integrates bottom-up visual properties with top-down higher order information that is not present in the visual stimuli.

- Activation in vOT was modulated by task changes when stimuli were held constant (Chapter 3), indicating that vOT activation is influenced by top-down non-visual processing demands required by the task.

- Visual familiarity, as opposed to lexical frequency, had a strong effect on vOT activation (Chapter 4). In contrast, lexical frequency modulated activation in a region of the inferior temporal gyrus lateral to the visual familiarity effect in vOT.

- Activation in vOT was higher for Japanese Kanji than Hiragana words when visual familiarity was held constant and this was not due to their inherent differences in visual complexity (Chapter 4). This suggests that the activation difference in vOT reflects the different association between the surface form of the word and its non-visual properties.

- Anatomical pathways linking vOT to other regions of the neural reading circuit were identified, which may underlie these bottom-up and top-down interactions (Chapter 5). Critically, the data suggest that vOT is not the first bifurcation point in the reading network and provide the anatomical basis for biologically-informed computational models of reading.

With respect to function, I systematically tested the claim that top-down influences significantly affect vOT activation when bottom-up visual properties of the stimuli are carefully controlled. The data strongly demonstrated that vOT responses are not solely dependent on visual properties of the stimuli. Instead, non-visual properties such as semantics and phonology automatically interact with visual properties of the stimuli. This is a central claim of the Interactive Account and clearly distinguishes it from the LCD model and the orthographic input lexicon account, both of which emphasise bottom-up, feed-forward processing. In addition, this interaction of visual and non-visual properties need not be limited to alphabetic stimuli or even to orthographic visual information, and is thus compatible with vOT activation to non-alphabetic word stimuli and non-orthographic stimuli. In other words, the Interactive Account addresses two aspects of the reading-specific theories of vOT that are incompatible with the data presented in this thesis and elsewhere. First, activity in vOT is task-dependent even when the bottom-up visual properties of the stimuli are carefully controlled and second, it is applicable to both alphabetic and non-alphabetic orthography as well as non-orthographic stimuli.

With respect to anatomy, I investigated the anatomical basis of functional interaction and further demonstrated that contrary to the assumption common of most neurological models of reading, vOT is not the central source of written input to the rest of the reading system. Instead, the anatomical pathways that mediate interactions between vOT and other cortical regions in the reading network are consistent with the parallel cortical visual pathways seen in other primates. vOT is part of the ventral stream that works together with the dorsal stream in parallel. These pathways are functionally connected, but the nodes in the two pathways may not be anatomically directly connected. These findings speak to a revised neuroanatomical model of reading with multiple interactive routes, between visual and higher-order language areas as well as within the language areas themselves (Figure 6.1).

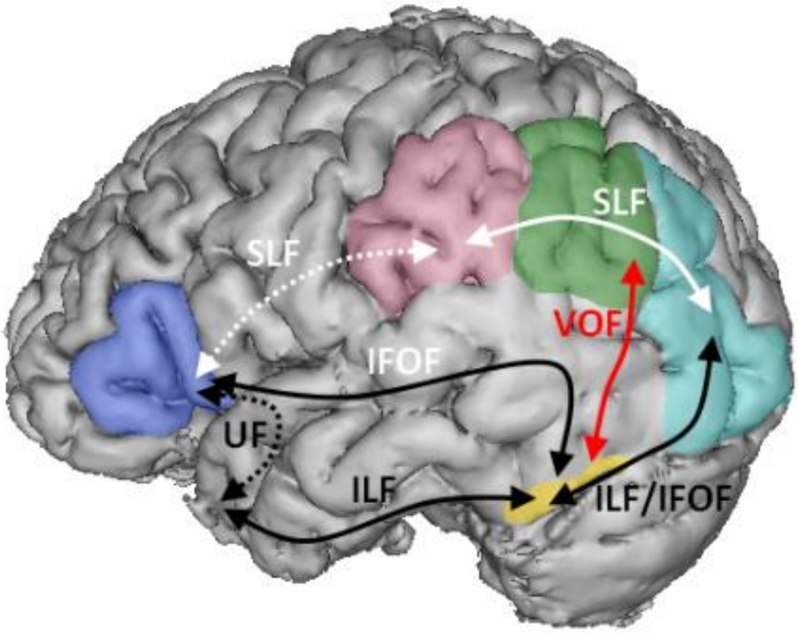

Figure 6.1: Modified neuroanatomical model of reading. Visual information splits from middle occipital gyrus (light blue) dorsally to supramarginal gyrus (pink) and ventrally to vOT (yellow). Angular gyrus (green) and vOT are linked directly. vOT is linked to Broca's area directly as well as indirectly via the temporal pole. Broca's area (dark blue) unifies the information from both routes for articulation. Anatomical pathways are shown with solid lines (demonstrated in this thesis) and dotted lines (not demonstrated in this thesis). The white lines represent the dorsal pathway and the black lines represent the ventral pathway (the red line belongs to neither). Double arrowheads indicate that information flow via these pathways is bi-directional. Abbreviations: SLF = superior longitudinal fasciculus; UF = uncinate fasciculus; IFOF = inferior fronto-occipital fasciculus; VOF = vertical occipital fasciculus; ILF = inferior longitudinal fasciculus.

This model is fundamentally interactive and the connections linking the nodes are bidirectional. In contrast to older neuroanatomical models, this model mirrors cognitive models, where interactivity is a fundamental property necessary to explain a wide range of behavioural data. Moreover, it is consistent with interactive nature demonstrated in the monkey neurophysiology and neuroanatomy literature (Kravitz et al., 2013). This model has at least two alternative routes from visual cortex to articulation, each of which is specialised for a different aspect of visual processing, consistent with parallel cortical pathways in the monkey (Ungerleider and Mishkin, 1982). These two pathways unite in Broca's area where integration of lexical information takes place (Hagoort, 2005). This is broadly consistent with cognitive models that emphasise a semantic route (ventral) and a single orthography-to-phonological route (dorsal) such as the distributed connectionist model by Seidenberg & McClelland (1989), rather than those which stipulate two routes for orthography-to-phonological conversion (lexical and sub-lexical) as in the dual route cascading model by Coltheart and colleagues (2001).

Furthermore, although vOT plays an important role in this model, it is not as central as it is in older models where it feeds the orthographic information to the whole reading system. This has an important implication as briefly noted in Chapter 5. Specifically, vOT is not the site of lexical or prelexical orthographic representations. Instead, vOT's contributions to reading are such that it helps recognise familiar orthographic patterns and link them with their higher order properties. The direct link to the temporal lobe areas is most likely to be involved in semantics and the indirect links to supramarginal gyrus is expected to concern phonology. Higher order properties from these regions are sent into vOT via predictive coding mechanisms by top-down backward connections.

Both functional and anatomical data presented in this thesis point to the same broad conclusion: neither vOT function nor its connectivity is specific to reading but they parallel visual object recognition and face processing (Bar et al., 2006; Cohen Kadosh et al., 2010; Gazzaley et al., 2007; Gilaie-Dotan et al., 2009; Kravitz et al., 2013; Kveraga et al., 2007; Schrader et al., 2009). Both the LCD model and the orthographic input lexicon account are built on the assumption that vOT is the first and essential node in the reading network and plays a reading-specific role. As such, they are not compatible with the majority of the data presented in this thesis. The Interactive Account, on the other hand, is not specialised for written words and proposes that the vOT response is the combination of bottom-up sensory information, top-down predictions from higher order processing areas and the mismatch between the two. All the data in this thesis are consistent with the Interactive Account.

Limitations & Future directions

Needless to say, this revised model is not yet complete. In order to fully understand the reading system during normal function, which will then aid the understanding of its break down due to neurological diseases and developmental dyslexia, other important regions need to be investigated both functionally as well as anatomically and integrated into the model. These would include other cortical areas such as left inferior/middle frontal regions where activation is increased during semantic word matching in dyslexic populations across alphabetic and non-alphabetic orthographies (Hu et al., 2010); the right hemisphere where activation is often associated with non-alphabetic orthography (Mei et al., 2013; Nakamura et al., 2005c; Tan et al., 2000); left superior temporal sulcus, which is involved in a reading pathway that does not include vOT (Richardson et al., 2011; Seghier et al., 2012); and sub-cortical regions such as the putamen, which is also suggested to play an important role in reading (Seghier and Price, 2010). In addition, further investigations of functional contributions of anatomical pathways such as the inferior fronto-occipital fasciculus and inferior longitudinal fasciculus (as seen in Chapter 5) will be valuable.

Investigations of top-down influences on vOT would ideally involve an investigative method with millisecond temporal resolution and sub-millimetre spatial resolution such as current source density (CSD) analysis (Schroeder et al., 1998). CSD samples neuronal electrical signals separately from each cortical layer, allowing one to separate feed-forward signals entering mostly layer IV from feedback signal entering outside layer IV (Van Essen and Maunsell, 1983). Moreover, the temporal resolution of the methodology allows the observation of the precise timing of these signals. Unfortunately, a non-invasive method for human investigation that offers this combination of temporal and spatial resolution is currently not available. Therefore, in this thesis, it was not possible to uncover fine-grained temporal and spatial characteristics of top-down influences.

With regard to anatomical investigations, DW-MRI and tractography are rapidly advancing. As noted in Chapter 2, model-free methods such as diffusion spectrum imaging (Wedeen et al., 2005) and persistent angular structure (PAS) MRI (Jansons and Alexander, 2003) will provide better accuracy of tractography than model-based approaches. However, currently the data acquisition requirements are higher with these techniques (e.g., a greater number of measurements, longer acquisition time, higher b-values) and the computation time is also longer (Seunarine and Alexander, 2009). Therefore, these are difficult or impractical methods to apply, given the resources available for research purposes at present. Similarly, fitting more complex, biologically plausible models that include additional parameters to the diffusion profile (e.g., NODDI, neurite orientation dispersion and density imaging, Zhang et al., 2012) would provide more detailed information about the microstructure of tissues. However, these techniques are still in their infancy and further assessment is required before they are in routine (clinical) use (Zhang et al., 2012). These options will undoubtedly help us better understand the structures of the reading pathways such as the vertical occipital fasciculus when they become more reliable and/or accessible.

Finally, as demonstrated in Chapter 4, the importance of using non-alphabetic orthographies in reading research needs to be emphasised. The neural information processing underlying reading requires a systematic investigation of a range of different languages and scripts since the evidence points to the common reading network regardless of orthographies. In particular, the uniqueness of the Japanese orthographic system is extremely useful in providing valuable information required for a deeper understanding of this complex system (e.g., Iwata, 2011; Sakurai, 2004).

REFERENCES

Amano, S., Kondo, T., 2003a. Nihongo-no Goi-tokusei [Lexical properties of Japanese] (Vol. 7), CD-ROM version (in Japanese). NTT database series. Sanseido, Tokyo

Amano, S., Kondo, T., 2003b. Nihongo-no Goi-tokusei [Lexical properties of Japanese] (Vols. 1–6), CD-ROM version (in Japanese). NTT database series. Sanseido, Tokyo.

Andersson, J.L., Hutton, C., Ashburner, J., Turner, R., Friston, K., 2001. Modeling geometric deformations in EPI time series. NeuroImage 13, 903-919.

Anwander, A., Tittgemeyer, M., Von Cramon, D.Y., Friederici, A.D., Knösche, T.R., 2007. Connectivity-based parcellation of Broca's area. Cerebral Cortex 17, 816-825.

Anzai, A., Peng, X., Van Essen, D.C., 2007. Neurons in monkey visual area V2 encode combinations of orientations. Nature Neuroscience 10, 1313-1321.

Ashtari, M., 2012. Anatomy and functional role of the inferior longitudinal fasciculus: a search has just begun. Developmental Medicine & Child Neurology 54, 6-7.

Axer, H., Klingner, C.M., Prescher, A., 2012. Fiber anatomy of dorsal and ventral language streams. Brain and Language.

Baayen, R.H., Pipenbrook, R., 1995. The Celex lexical database. Linguistuc Data Consortium, University of Pennsylvania, Philadelphia.

Bailey, P., von Bonin, G., McCulloch, W.S., 1950. The isocortex of the chimpanzee. University of Illinois Press, Urbana.

Bar, M., 2003. A cortical mechanism for triggering top-down facilitation in visual object recognition. Journal of Cognitive Neuroscience 15, 600-609.

Bar, M., Kassam, K.S., Ghuman, A.S., Boshyan, J., Schmidt, A.M., Dale, A.M., Hämäläinen, M.S., Marinkovic, K., Schacter, D.L., Rosen, B.R., Halgren, E., 2006. Top-down facilitation of visual recognition. Proceedings of the National Academy of Sciences of the United States of America 103, 449-454.

Beckmann, C.F., Jenkinson, M., Smith, S.M., 2003. General multilevel linear modeling for group analysis in FMRI. NeuroImage 20, 1052-1063.

Behrens, T., Jbabdi, S., 2009. MR diffusion tractography. In: Johansen-Berg, H., Behrens, T. (Eds.), Diffusion MRI: From quantitative measurement to in-vivo neuroanatomy. Academic Press, San Diego pp. 333-351.

Behrens, T.E.J., Berg, H.J., Jbabdi, S., Rushworth, M.F.S., Woolrich, M.W., 2007. Probabilistic diffusion tractography with multiple fibre orientations: What can we gain? NeuroImage 34, 144-155.

Behrens, T.E.J., Johansen-Berg, H., Woolrich, M.W., Smith, S.M., Wheeler-Kingshott, C.A.M., Boulby, P.A., Barker, G.J., Sillery, E.L., Sheehan, K., Ciccarelli, O., Thompson, A.J., Brady, J.M., Matthews, P.M., 2003a. Non-invasive mapping of connections between human thalamus and cortex using diffusion imaging. Nature Neuroscience 6, 750-757.

Behrens, T.E.J., Woolrich, M.W., Jenkinson, M., Johansen-Berg, H., Nunes, R.G., Clare, S., Matthews, P.M., Brady, J.M., Smith, S.M., 2003b. Characterization and Propagation of Uncertainty in Diffusion-Weighted MR Imaging. Magnetic Resonance in Medicine 50, 1077-1088.

Behrmann, M., Nelson, J., Sekuler, E.B., 1998. Visual complexity in letter-by-letter reading: 'Pure' alexia is not pure. Neuropsychologia 36, 1115-1132.

Ben-Shachar, M., Dougherty, R.F., Deutsch, G.K., Wandell, B.A., 2007a. Differential sensitivity to words and shapes in ventral occipito-temporal cortex. Cerebral Cortex 17, 1604-1611.

Ben-Shachar, M., Dougherty, R.F., Wandell, B.A., 2007b. White matter pathways in reading. Current Opinion in Neurobiology 17, 258-270.

Berens, P., Logothetis, N.K., Tolias, A.S., Ecker, A., Berens, P., Keliris, G., Bethge, M., Logothetis, N., Tolias, A., 2013. Local field potentials, BOLD and spiking activity - relationships and physiological mechanisms. Nature 2, 67.

Beversdorf, D.Q., Ratcliffe, N.R., Rhodes, C.H., Reeves, A.G., 1997. Pure alexia: Clinical-pathologic evidence for a lateralized visual language association cortex. Clinical Neuropathology 16, 328-331.

Binder, J., Mohr, J., 1992. The Topography of callosal reading pathways. A case-control analysis. Brain 115, 1807-1826.

Binder, J., Price, C.J., 2001. Functional neuroimaging of language. In: Canbeza, R., Kingston, A. (Eds.), Handbook of functional neuroimaging of cognition. MIT Press, pp. 187-251.

Binder, J.R., Frost, J.A., Hammeke, T.A., Bellgowan, P.S.F., Rao, S.M., Cox, R.W., 1999. Conceptual processing during the conscious resting state: A functional MRI study. Journal of Cognitive Neuroscience 11, 80-93.

Binder, J.R., Swanson, S.J., Hammeke, T.A., Morris, G.L., Mueller, W.M., Fischer, M., Benbadis, S., Frost, J.A., Rao, S.M., Haughton, V.M., 1996. Determination of language dominance using functional MRI: A comparison with the Wada test. Neurology 46, 978-984.

Biswal, B., Zerrin Yetkin, F., Haughton, V.M., Hyde, J.S., 1995. Functional connectivity in the motor cortex of resting human brain using echo-planar mri. Magnetic Resonance in Medicine 34, 537-541.

Bitan, T., Booth, J.R., Choy, J., Burman, D.D., Gitelman, D.R., Mesulam, M.M., 2005. Shifts of effective connectivity within a language network during rhyming and spelling. Journal of Neuroscience 25, 5397-5403.

Bitan, T., Cheon, J., Lu, D., Burman, D.D., Booth, J.R., 2009. Developmental increase in top-down and bottom-up processing in a phonological task: An effective connectivity, fMRI study. Journal of Cognitive Neuroscience 21, 1135-1145.

Bitan, T., Cheon, J., Lu, D., Burman, D.D., Gitelman, D.R., Mesulam, M.M., Booth, J.R., 2007. Developmental changes in activation and effective connectivity in phonological processing. NeuroImage 38, 564-575.

Bledowski, C., Kaiser, J., Rahm, B., 2010. Basic operations in working memory: Contributions from functional imaging studies. Behavioural Brain Research 214, 172-179.

Bokde, A.L.W., Tagamets, M.A., Friedman, R.B., Horwitz, B., 2001. Functional interactions of the inferior frontal cortex during the processing of words and word-like stimuli. Neuron 30, 609-617.

Bolger, D.J., Perfetti, C.A., Schneider, W., 2005. Cross-cultural effect on the brain revisited: Universal structures plus writing system variation. Human Brain Mapping 25, 92-104.

Bookheimer, S.Y., Zeffiro, T.A., Blaxton, T., Gaillard, W., Theodore, W., 1995. Regional cerebral blood flow during object naming and word reading. Human Brain Mapping 3, 93-106.

Booth, J.R., Burman, D.D., Meyer, J.R., Gitelman, D.R., Parrish, T.B., Mesulam, M.M., 2002. Functional anatomy of intra- and cross-modal lexical tasks. NeuroImage 16, 7-22.

Booth, J.R., Lu, D., Burman, D.D., Chou, T.L., Jin, Z., Peng, D.L., Zhang, L., Ding, G.S., Deng, Y., Liu, L., 2006. Specialization of phonological and semantic processing in Chinese word reading. Brain Research 1071, 197-207.

Boussaoud, D., Desimone, R., Ungerleider, L.G., 1991. Visual topography of area TEO in the macaque. Journal of Comparative Neurology 306, 554-575.

Brodmann, K., 1909. Vergleichende Lokalisationslehre der Grosshirnrinde in ihren Prinzipien dargestellt auf Grund des Zellenbaues. Barth, Leipzig.

Bruno, J.L., Zumberge, A., Manis, F.R., Lu, Z.L., Goldman, J.G., 2008. Sensitivity to orthographic familiarity in the occipito-temporal region. NeuroImage 39, 1988-2001.

Brunswick, N., McCrory, E., Price, C.J., Frith, C.D., Frith, U., 1999. Explicit and implicit processing of words and pseudowords by adult developmental dyslexics. A search for Wernicke's Wortschatz? Brain 122, 1901-1917.

Bullier, J., 2001. Integrated model of visual processing. Brain Research Reviews 36, 96-107.

Bullier, J., Nowak, L.G., 1995. Parallel versus serial processing: new vistas on the distributed organization of the visual system. Curr Opin Neurobiol 5, 497-503.

Cai, Q., Paulignan, Y., Brysbaert, M., Ibarrola, D., Nazir, T.A., 2010. The left ventral occipito-temporal response to words depends on language lateralization but not on visual familiarity. Cerebral Cortex 20, 1153-1163.

Carreiras, M., Mechelli, A., Estevez, A., Price, C.J., 2007. Brain activation for lexical decision and reading aloud: Two sides of the same coin? Journal of Cognitive Neuroscience 19, 433-444.

Carreiras, M., Mechelli, A., Price, C.J., 2006. Effect of word and syllable frequency on activation during lexical decision and reading aloud. Human Brain Mapping 27, 963-972.

Carreiras, M., Seghier, M.L., Baquero, S., Estévez, A., Lozano, A., Devlin, J.T., Price, C.J., 2009. An anatomical signature for literacy. Nature 461, 983-986.

Caspers, J., Zilles, K., Eickhoff, S.B., Schleicher, A., Mohlberg, H., Amunts, K., 2013. Cytoarchitectonical analysis and probabilistic mapping of two extrastriate areas of the human posterior fusiform gyrus. Brain Structure and Function 218, 511-526.

Castles, A., Coltheart, M., 1993. Varieties of developmental dyslexia. Cognition 47, 149-180.

Catani, M., de Schotten, M.T., 2012. Atlas of Human Brain Connections. OUP Oxford.

Catani, M., Howard, R.J., Pajevic, S., Jones, D.K., 2002. Virtual in Vivo interactive dissection of white matter fasciculi in the human brain. NeuroImage 17, 77-94.

Catani, M., Jones, D.K., Donato, R., Ffytche, D.H., 2003. Occipito-temporal connections in the human brain. Brain 126, 2093-2107.

Catani, M., Mesulam, M., 2008. The arcuate fasciculus and the disconnection theme in language and aphasia: History and current state. Cortex 44, 953-961.

Cattell, J.M., 1886. The time it takes to see and name objects. Mind os-XI, 63-65.

Chao, L.L., Haxby, J.V., Martin, A., 1999. Attribute-based neural substrates in temporal cortex for perceiving and knowing about objects. Nature Neuroscience 2, 913-919.

Chee, M.W.L., Soon, C.S., Lee, H.L., 2003a. Common and segregated neuronal networks for different languages revealed using functional magnetic resonance adaptation. Journal of Cognitive Neuroscience 15, 85-97.

Chee, M.W.L., Weekes, B., Lee, K.M., Soon, C.S., Schreiber, A., Hoon, J.J., Chee, M., 2000. Overlap and dissociation of semantic processing of Chinese characters, English words, and pictures: evidence from fMRI. NeuroImage 12, 392-403.

Chee, M.W.L., Westphal, C., Goh, J., Graham, S., Song, A.W., 2003b. Word frequency and subsequent memory effects studied using event-related fMRI. NeuroImage 20, 1042-1051.

Chen, Y., Fu, S., Iversen, S.D., Smith, S.M., Matthews, P.M., 2002. Testing for dual brain processing routes in reading: A direct contrast of Chinese character and Pinyin reading using fMRI. Journal of Cognitive Neuroscience 14, 1088-1098.

Ciccarelli, O., Behrens, T.E., Altmann, D.R., Orrell, R.W., Howard, R.S., Johansen-Berg, H., Miller, D.H., Matthews, P.M., Thompson, A.J., 2006. Probabilistic diffusion tractography: A potential tool to assess the rate of disease progression in amyotrophic lateral sclerosis. Brain 129, 1859-1871.

Clarke, S., 1994. Association and intrinsic connections of human extrastriate visual cortex. Proceedings of the Royal Society B: Biological Sciences 257, 87-92.

Cohen Kadosh, K., Henson, R.N.A., Cohen Kadosh, R., Johnson, M.H., Dick, F., 2010. Task-dependent activation of face-sensitive cortex: An fMRI adaptation study. Journal of Cognitive Neuroscience 22, 903-907.

Cohen, L., Dehaene, S., Naccache, L., Lehericy, S., Dehaene-Lambertz, G., Henaff, M.A., Michel, F., 2000. The visual word form area: spatial and temporal characterization of an initial stage of reading in normal subjects and posterior split-brain patients. Brain 123 (Pt 2), 291-307.

Cohen, L., Jobert, A., Le Bihan, D., Dehaene, S., 2004. Distinct unimodal and multimodal regions for word processing in the left temporal cortex. NeuroImage 23, 1256-1270.

Cohen, L., Lehericy, S., Chochon, F., Lemer, C., Rivaud, S., Dehaene, S., 2002. Language-specific tuning of visual cortex? Functional properties of the Visual Word Form Area. Brain 125, 1054-1069.

Coltheart, M., 1981. The MRC Psycholinguistic Database. Quarterly Journal of Experimental Psychology 33A, 497-505.

Coltheart, M., 2004. Are there lexicons? The Quarterly Journal of Experimental Psychology Section A 57, 1153-1171.

Coltheart, M., Rastle, K., Perry, C., Langdon, R., Ziegler, J., 2001. DRC: A dual route cascaded model of visual word recognition and reading aloud. Psychological Review 108, 204-256.

Cornelissen, P.L., Kringelbach, M.L., Ellis, A.W., Whitney, C., Holliday, I.E., Hansen, P.C., 2009. Activation of the left inferior frontal gyrus in the first 200 ms of reading: evidence from magnetoencephalography (MEG). PLoS One 4, e5359.

Crosby, E.C., Humphrey, T., Lauer, E.W., 1962. Correlative anatomy of the nervous system. The Macmillan Company, New York.

Damasio, A.R., Damasio, H., 1983. The anatomic basis of pure alexia. Neurology 33, 1573-1573.

Davis, C.J., 2005. N-watch: A program for deriving neighborhood size and other psycholinguistic statistics. Behavior Research Methods 37, 65-70.

Dehaene, S., Cohen, L., 2007. Cultural recycling of cortical maps. Neuron 56, 384-398.

Dehaene, S., Cohen, L., 2011. The unique role of the visual word form area in reading. Trends in Cognitive Sciences 15, 254-262.

Dehaene, S., Cohen, L., Sigman, M., Vinckier, F., 2005. The neural code for written words: a proposal. Trends in Cognitive Sciences 9, 335-341.

Dehaene, S., Jobert, A., Naccache, L., Ciuciu, P., Poline, J.B., Bihan, D.L., Cohen, L., 2004. Letter binding and invariant recognition of masked words: Behavioral and neuroimaging evidence. Psychological Science 15, 307-313.

Dehaene, S., Le Clec, H.G., Poline, J.B., Le Bihan, D., Cohen, L., 2002. The visual word form area: a prelexical representation of visual words in the fusiform gyrus. NeuroReport 13, 321-325.

Dehaene, S., Naccache, L., Cohen, L., Bihan, D.L., Mangin, J.F., Poline, J.B., Rivière, D., 2001. Cerebral mechanisms of word masking and unconscious repetition priming. Nature Neuroscience 4, 752-758.

Dejerine, J., 1891. Sur un cas de cecite verbale avec agraphie suivi d'autopsie. Memoires de la Societe de Biologie 3, 197–201.

Dejerine, J., 1892. Contribution a l’etude anatomoclinique et clinique des differentes varietes de cecite verbal. Memoires de la Societe de Biologie 4, 61-90.

Déjerine, J.J., 1895. Anatomie des centres nerveux. Rueff, Paris.

Démonet, J.F., Thierry, G., Cardebat, D., 2005. Renewal of the neurophysiology of language: Functional neuroimaging. Physiological Reviews 85, 49-95.

Desikan, R.S., Ségonne, F., Fischl, B., Quinn, B.T., Dickerson, B.C., Blacker, D., Buckner, R.L., Dale, A.M., Maguire, R.P., Hyman, B.T., Albert, M.S., Killiany, R.J.,

2006. An automated labeling system for subdividing the human cerebral cortex on MRI scans into gyral based regions of interest. NeuroImage 31, 968-980.

Devlin, J.T., 2009. Current Perspectives on Imaging Language. In: Kraft, E., Gulyás, B., Pöppel, E. (Eds.), Neural Correlates of Thinking. Springer Berlin Heidelberg, pp. 123-139.

Devlin, J.T., Jamison, H.L., Gonnerman, L.M., Matthews, P.M., 2006a. The role of the posterior fusiform gyrus in reading. Journal of Cognitive Neuroscience 18, 911-922.

Devlin, J.T., Matthews, P.M., Rushworth, M.F.S., 2003. Semantic processing in the left inferior prefrontal cortex: A combined functional magnetic resonance imaging and transcranial magnetic stimulation study. Journal of Cognitive Neuroscience 15, 71-84.

Devlin, J.T., Sillery, E.L., Hall, D.A., Hobden, P., Behrens, T.E.J., Nunes, R.G., Clare, S., Matthews, P.M., Moore, D.R., Johansen-Berg, H., 2006b. Reliable identification of the auditory thalamus using multi-modal structural analyses. NeuroImage 30, 1112-1120.

Dick, A.S., Tremblay, P., 2012. Beyond the arcuate fasciculus: Consensus and controversy in the connectional anatomy of language. Brain 135, 3529-3550.

Dietz, N.A.E., Jones, K.M., Gareau, L., Zeffiro, T.A., Eden, G.F., 2005. Phonological decoding involves left posterior fusiform gyrus. Human Brain Mapping 26, 81-93.

Duffau, H., Gatignol, P., Mandonnet, E., Peruzzi, P., Tzourio-Mazoyer, N., Capelle, L., 2005. New insights into the anatomo-functional connectivity of the semantic system: A study using cortico-subcortical electrostimulations. Brain 128, 797-810.

Duncan, K.J., Pattamadilok, C., Devlin, J.T., 2010. Investigating occipito-temporal contributions to reading with TMS. Journal of Cognitive Neuroscience 22, 739-750.

Duncan, K.J., Pattamadilok, C., Knierim, I., Devlin, J.T., 2009. Consistency and variability in functional localisers. NeuroImage 46, 1018-1026.

Ebeling, U., Cramon, D.v., 1992. Topography of the uncinate fascicle and adjacent temporal fiber tracts. Acta Neurochirurgica 115, 143-148.

Eger, E., Schweinberger, S.R., Dolan, R.J., Henson, R.N., 2005. Familiarity enhances invariance of face representations in human ventral visual cortex: fMRI evidence. NeuroImage 26, 1128-1139.

Eickhoff, S.B., Jbabdi, S., Caspers, S., Laird, A.R., Fox, P.T., Zilles, K., Behrens, T.E.J., 2010. Anatomical and functional connectivity of cytoarchitectonic areas within the human parietal operculum. Journal of Neuroscience 30, 6409-6421.

Ekstrom, A., 2010. How and when the fMRI BOLD signal relates to underlying neural activity: The danger in dissociation. Brain Research Reviews 62, 233-244.

Engel, S.A., Glover, G.H., Wandell, B.A., 1997. Retinotopic organization in human visual cortex and the spatial precision of functional MRI. Cerebral Cortex 7, 181-192.

Epelbaum, S., Pinel, P., Gaillard, R., Delmaire, C., Perrin, M., Dupont, S., Dehaene, S., Cohen, L., 2008. Pure alexia as a disconnection syndrome: new diffusion imaging evidence for an old concept. Cortex 44, 962-974.

Felleman, D.J., Van Essen, D.C., 1991. Distributed Hierarchical Processing in the Primate Cerebral Cortex. Cerebral Cortex 1, 1-47.

Fiebach, C.J., Friederici, A.D., Müller, K., Von Cramon, D.Y., 2002. fMRI evidence for dual routes to the mental lexicon in visual word recognition. Journal of Cognitive Neuroscience 14, 11-23.

Fiebach, C.J., Ricker, B., Friederici, A.D., Jacobs, A.M., 2007. Inhibition and facilitation in visual word recognition: Prefrontal contribution to the orthographic neighborhood size effect. NeuroImage 36, 901-911.

Fiez, J.A., Balota, D.A., Raichle, M.E., Petersent, S.E., 1999. Effects of lexicality, frequency, and spelling-to-sound consistency on the functional anatomy of reading. Neuron 24, 205-218.

Fiez, J.A., Petersen, S.E., 1998. Neuroimaging studies of word reading. Proc Natl Acad Sci U S A 95, 914-921.

Fischl, B., Van Der Kouwe, A., Destrieux, C., Halgren, E., Ségonne, F., Salat, D.H., Busa, E., Seidman, L.J., Goldstein, J., Kennedy, D., Caviness, V., Makris, N., Rosen, B., Dale, A.M., 2004. Automatically Parcellating the Human Cerebral Cortex. Cerebral Cortex 14, 11-22.

Fleming, S.M., Thomas, C.L., Dolan, R.J., 2010. Overcoming status quo bias in the human brain. Proceedings of the National Academy of Sciences of the United States of America 107, 6005-6009.

Ford, A., McGregor, K.M., Case, K., Crosson, B., White, K.D., 2010. Structural connectivity of Broca's area and medial frontal cortex. NeuroImage 52, 1230-1237.

Friston, K., 2002. Functional integration and inference in the brain. Progress in Neurobiology 68, 113-143.

Friston, K., 2003. Learning and inference in the brain. Neural Networks 16, 1325-1352.

Friston, K., 2005. A theory of cortical responses. Philosophical Transactions of the Royal Society B: Biological Sciences 360, 815-836.

Friston, K., 2010. The free-energy principle: A unified brain theory? Nature Reviews Neuroscience 11, 127-138.

Friston, K., Kiebel, S., 2009. Predictive coding under the free-energy principle. Philosophical Transactions of the Royal Society B: Biological Sciences 364, 1211-1221.

Friston, K.J., 1994. Functional and effective connectivity in neuroimaging: a synthesis. Human Brain Mapping 2, 56-78.

Friston, K.J., Harrison, L., Penny, W., 2003. Dynamic causal modelling. NeuroImage 19, 1273-1302.

Friston, K.J., Price, C.J., 2001. Dynamic representations and generative models of brain function. Brain Research Bulletin 54, 275-285.

Frost, R., 1998. Toward a strong phonological theory of visual word recognition: true issues and false trails. Psychol Bull 123, 71-99.

Frost, S.J., Mencl, W.E., Sandak, R., Moore, D.L., Rueckl, J.G., Katz, L., Fulbright, R.K., Pugh, K.R., 2005. A functional magnetic resonance imaging study of the tradeoff between semantics and phonology in reading aloud. NeuroReport 16, 621-624.

Frost, S.J., Sandak, R., Mencl, W.E., Landi, N., Moore, D., Della Porta, G., J.G., R., Katz, L., Pugh, K.R., 2008. Neurobiological studies of skilled and impaired word reading. In: Elena, L.G., Adam, J.N. (Eds.), Single-Word Reading: Behavioral and Biological Perspectives Taylor & Francis Group, New York, pp. 355-376.

Fu, S., Chen, Y., Smith, S., Iversen, S., Matthews, P.M., 2002. Effects of word form on brain processing of written Chinese. NeuroImage 17, 1538-1548.

Gaillard, R., Naccache, L., Pinel, P., Clémenceau, S., Volle, E., Hasboun, D., Dupont, S., Baulac, M., Dehaene, S., Adam, C., 2006. Direct intracranial, FMRI, and lesion evidence for the causal role of left inferotemporal cortex in reading. Neuron 50, 191-204.

Gazzaley, A., Rissman, J., Cooney, J., Rutman, A., Seibert, T., Clapp, W., D'Esposito, M., 2007. Functional interactions between prefrontal and visual association cortex contribute to top-down modulation of visual processing. Cereb Cortex 17 Suppl 1, i125-135.

Geschwind, N., 1962. The Anatomy of Acquired Disorders of Reading. In: Money, J. (Ed.), Reading disability. Johns Hopkins Press, Baltimore, pp. 115-129.

Geschwind, N., 1965. Disconnexion syndromes in animal and man Brain 88, 585.

Geschwind, N., Fusillo, M., 1966. Color-naming defects in association with alexia. Archives of Neurology 15, 137-146.

Gil-Robles, S., Carvallo, A., Jimenez, M.D.M., Gomez Caicoya, A., Martinez, R., Ruiz-Ocaña, C., Duffau, H., 2013. Double dissociation between visual recognition and picture naming: A study of the visual language connectivity using tractography and brain stimulation. Neurosurgery 72, 678-686.

Gilaie-Dotan, S., Perry, A., Bonneh, Y., Malach, R., Bentin, S., 2009. Seeing with profoundly deactivated mid-level visual areas: non-hierarchical functioning in the human visual cortex. Cereb Cortex 19, 1687-1703.

Glasser, M.F., Rilling, J.K., 2008. DTI tractography of the human brain's language pathways. Cerebral Cortex 18, 2471-2482.

Glezer, L.S., Jiang, X., Riesenhuber, M., 2009. Evidence for Highly Selective Neuronal Tuning to Whole Words in the "Visual Word Form Area". Neuron 62, 199-204.

Glover, G.H., 1999. Deconvolution of impulse response in event-related BOLD fMRI. NeuroImage 9, 416-429.

Gold, B.T., Balota, D.A., Jones, S.J., Powell, D.K., Smith, C.D., Andersen, A.H., 2006. Dissociation of automatic and strategic lexical-semantics: Functional magnetic resonance imaging evidence for differing roles of multiple frontotemporal regions. Journal of Neuroscience 26, 6523-6532.

Graves, W.W., Grabowski, T.J., Mehta, S., Gupta, P., 2008. The left posterior superior temporal gyrus participates specifically in accessing lexical phonology. Journal of Cognitive Neuroscience 20, 1698-1710.

Gray, H., Lewis, W.H., 1918. Anatomy of the human body. Lea & Febiger, Philadelphia.

Greenblatt, S.H., 1973. Alexia without agraphia or hemianopsia. Anatomical analysis of an autopsied case. Brain 96, 307-316.

Greenblatt, S.H., 1976. Subangular alexia without agraphia or hemianopsia. Brain and Language 3, 229-245.

Greicius, M.D., Krasnow, B., Reiss, A.L., Menon, V., 2003. Functional connectivity in the resting brain: A network analysis of the default mode hypothesis. Proceedings of the National Academy of Sciences of the United States of America 100, 253-258.

Ha Duy Thuy, D., Matsuo, K., Nakamura, K., Toma, K., Oga, T., Nakai, T., Shibasaki, H., Fukuyama, H., 2004. Implicit and explicit processing of kanji and kana words and non-words studied with fMRI. NeuroImage 23, 878-889.

Hagmann, P., Thiran, J.P., Jonasson, L., Vandergheynst, P., Clarke, S., Maeder, P., Meuli, R., 2003. DTI mapping of human brain connectivity: Statistical fibre tracking and virtual dissection. NeuroImage 19, 545-554.

Hagoort, P., 2005. On Broca, brain, and binding: A new framework. Trends in Cognitive Sciences 9, 416-423.

Harm, M.W., Seidenberg, M.S., 2004. Computing the meanings of words in reading: cooperative division of labor between visual and phonological processes. Psychol Rev 111, 662-720.

Hasson, U., Levy, I., Behrmann, M., Hendler, T., Malach, R., 2002. Eccentricity Bias as an Organizing Principle for Human High-Order Object Areas. Neuron 34, 479-490.

Hauk, O., Davis, M.H., Pulvermüller, F., 2008. Modulation of brain activity by multiple lexical and word form variables in visual word recognition: A parametric fMRI study. NeuroImage 42, 1185-1195.

Haxby, J.V., Hoffman, E.A., Gobbini, M.I., 2002. Human neural systems for face recognition and social communication. Biological psychiatry 51, 59-67.

Hegdé, J., Van Essen, D.C., 2000. Selectivity for complex shapes in primate visual area V2. The Journal of neuroscience : the official journal of the Society for Neuroscience 20.

Heiervang, E., Behrens, T.E.J., Mackay, C.E., Robson, M.D., Johansen-Berg, H., 2006. Between session reproducibility and between subject variability of diffusion MR and tractography measures. NeuroImage 33, 867-877.

Hellyer, P.J., Woodhead, Z.V.J., Leech, R., Wise, R.J.S., 2011. An investigation of Twenty/20 vision in reading. Journal of Neuroscience 31, 14631-14638.

Herbster, A.N., Mintun, M.A., Nebes, R.D., Becker, J.T., 1997. Regional cerebral blood flow during word and nonword reading. Hum Brain Mapp 5, 84-92.

Herrick, C.J., 1915. An Introduction to Neurology. W.B. Saunders Co., Philadelphia.

Hillis, A.E., Newhart, M., Heidler, J., Barker, P., Herskovits, E., Degaonkar, M., 2005. The roles of the "visual word form area" in reading. NeuroImage 24, 548-559.

Hinshelwood, J., 1900. Letter-, Word- and Mind-blindness. H.K. Lewis.

Holl, N., Noblet, V., Rodrigo, S., Dietemann, J.L., Mekhbi, M.B., Kehrli, P., Wolfram-Gabel, R., Braun, M., Kremer, S., 2011. Temporal lobe association fiber tractography as compared to histology and dissection. Surgical and Radiologic Anatomy 33, 713-722.

Horwitz, B., Rumsey, J.M., Donohue, B.C., 1998. Functional connectivity of the angular gyrus in normal reading and dyslexia. Proceedings of the National Academy of Sciences of the United States of America 95, 8939-8944.

Hsu, C.H., Lee, C.Y., Marantz, A., 2011. Effects of visual complexity and sublexical information in the occipitotemporal cortex in the reading of Chinese phonograms: A single-trial analysis with MEG. Brain and Language 117, 1-11.

Hu, W., Lee, H.L., Zhang, Q., Liu, T., Geng, L.B., Seghier, M.L., Shakeshaft, C., Twomey, T., Green, D.W., Yang, Y.M., Price, C.J., 2010. Developmental dyslexia in Chinese and English populations: Dissociating the effect of dyslexia from language differences. Brain 133, 1694-1706.

Ino, T., Nakai, R., Azuma, T., Kimura, T., Fukuyama, H., 2009. Recognition and reading aloud of kana and kanji word: An fMRI study. Brain Research Bulletin 78, 232-239.

Iwai, E., Mishkin, M., 1969. Further evidence on the locus of the visual area in the temporal lobe of the monkey. Experimental Neurology 25, 585-594.

Iwata, M., 1984. Kanji versus Kananeuropsychological correlates of the Japanese writing system. Trends in Neurosciences 7, 290-293.

Iwata, M., 2011. Dual neural circuit model of reading and writing. Brain and Nerve 63, 821-829.

Jacobs, A.M., Graf, R., Kinder, A., 2003. Receiver operating characteristics in the lexical decision task: evidence for a simple signal-detection process simulated by the multiple read-out model. Journal of Experimental Psychology: Learning, Memory, and Cognition 29, 481-488.

Jansons, K.M., Alexander, D.C., 2003. Persistent angular structure: New insights from diffusion MRI data. Dummy version. Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), pp. 672-683.

Jbabdi, S., Johansen-Berg, H., 2011. Tractography: Where Do We Go from Here? . Brain Connectivity 1, 169-183.

Jellison, B.J., Field, A.S., Medow, J., Lazar, M., Salamat, M.S., Alexander, A.L., 2004. Diffusion Tensor Imaging of Cerebral White Matter: A Pictorial Review of Physics, Fiber Tract Anatomy, and Tumor Imaging Patterns. American Journal of Neuroradiology 25, 356-369.

Jenkinson, M., Bannister, P., Brady, M., Smith, S., 2002. Improved optimization for the robust and accurate linear registration and motion correction of brain images. NeuroImage 17, 825-841.

Jenkinson, M., Smith, S., 2001. A global optimisation method for robust affine registration of brain images. Medical Image Analysis 5, 143-156.

Jiang, X., Rosen, E., Zeffiro, T., VanMeter, J., Blanz, V., Riesenhuber, M., 2006. Evaluation of a shape-based model of human face discrimination using fMRI and behavioral techniques. Neuron 50, 159.

Johnston, J.C., 1981. Understanding word perception: Clues from studying the word superiority effect. . In: Tzeng, O.J.L., Singer, H. (Eds.), Perception of print: Reading research in experimental psychology Lawrence Erlbaum Associates, Hillsdale, NJ, pp. 65-81.

Jones, D.K., Horsfield, M.A., Simmons, A., 1999. Optimal strategies for measuring diffusion in anisotropic systems by magnetic resonance imaging. Magnetic Resonance in Medicine 42, 515-525.

Joubert, S., Beauregard, M., Walter, N., Bourgouin, P., Beaudoin, G., Leroux, J.M., Karama, S., Lecours, A.R., 2004. Neural correlates of lexical and sublexical processes in reading. Brain and Language 89, 9-20.

Juch, H., Zimine, I., Seghier, M.L., Lazeyras, F., Fasel, J.H.D., 2005. Anatomical variability of the lateral frontal lobe surface: Implication for intersubject variability in language neuroimaging. NeuroImage 24, 504-514.

Kanwisher, N., McDermott, J., Chun, M.M., 1997. The fusiform face area: A module in human extrastriate cortex specialized for face perception. Journal of Neuroscience 17, 4302-4311.

Katanoda, K., Yoshikawa, K., Sugishita, M., 2001. A functional MRI study on the neural substrates for writing. Human Brain Mapping 13, 34-42.

Kawabata Duncan, K.J., Twomey, T., Parker Jones, Ō., Seghier, M.L., Haji, T., Sakai, K., Price, C.J., Devlin, J.T., 2013. Inter- and Intrahemispheric Connectivity Differences When Reading Japanese Kanji and Hiragana. Cerebral Cortex.

Keller, S.S., Highley, J.R., Garcia-Finana, M., Sluming, V., Rezaie, R., Roberts, N., 2007. Sulcal variability, stereological measurement and asymmetry of Broca's area on MR images. Journal of Anatomy 211, 534-555.

Keller, T.A., Carpenter, P.A., Just, M.A., 2001. The neural bases of sentence comprehension: a fMRI examination of syntactic and lexical processing. Cerebral Cortex 11, 223-237.

Kherif, F., Josse, G., Price, C.J., 2011. Automatic top-down processing explains common left occipito-temporal responses to visual words and objects. Cereb Cortex 21, 103-114.

Kiehl, K.A., Liddle, P.F., Smith, A.M., Mendrek, A., Forster, B.B., Hare, R.D., 1999. Neural pathways involved in the processing of concrete and abstract words. Hum Brain Mapp 7, 225-233.

Kier, E.L., Staib, L.H., Davis, L.M., Bronen, R.A., 2004. MR imaging of the temporal stem: Anatomic dissection tractography of the uncinate fasciculus, inferior occipitofrontal fasciculus, and Meyer's loop of the optic radiation. American Journal of Neuroradiology 25, 677-691.

Kiriyama, I., Miki, H., Kikuchi, K., Ohue, S., Matsuda, S., Mochizuki, T., 2009. Topographic analysis of the inferior parietal lobule in high-resolution 3D MR imaging. American Journal of Neuroradiology 30, 520-524.

Kiyosawa, M., Itoh, M., Nakagawa, Y., Kobayashi, N., Tamai, M., 1995. Effect of kanji and kana reading on cerebral blood flow patterns measured by PET. Japanese Journal of Ophthalmology 39, 198-205.

Klingler, J., 1935. Erleichterung der makroskopischen Praeparation des Gehirns durch den Gefrierprozess. Schweiz Arch Neurol Psychiatr 36, 247-256.

Klingler, J., Gloor, P., 1960. The connections of the amygdala and of the anterior temporal cortex in the human brain. The Journal of Comparative Neurology 115, 333-369.

Koyama, M.S., Kelly, C., Shehzad, Z., Penesetti, D., Castellanos, F.X., Milham, M.P., 2010. Reading networks at rest. Cerebral Cortex 20, 2549-2559.

Kravitz, D.J., Saleem, K.S., Baker, C.I., Mishkin, M., 2011. A new neural framework for visuospatial processing. Nature Reviews Neuroscience 12, 217-230.

Kravitz, D.J., Saleem, K.S., Baker, C.I., Ungerleider, L.G., Mishkin, M., 2013. The ventral visual pathway: an expanded neural framework for the processing of object quality. Trends in Cognitive Sciences 17, 26-49.

Krieg, W.J.S., 1966. Functional neuroanatomy. Brain Books, Evanston, Illinois.

Krieg, W.J.S., Mark, M., 1973. Architectonics of Human Cerebral Fiber Systems. Brain Books, Evanston, Illinois.

Kronbichler, M., Bergmann, J., Hutzler, F., Staffen, W., Mair, A., Ladurner, G., Wimmer, H., 2007. Taxi vs. taksi: On orthographic word recognition in the left ventral occipitotemporal cortex. Journal of Cognitive Neuroscience 19, 1584-1594.

Kronbichler, M., Hutzler, F., Wimmer, H., Mair, A., Staffen, W., Ladurner, G., 2004. The visual word form area and the frequency with which words are encountered: Evidence from a parametric fMRI study. NeuroImage 21, 946-953.

Kronbichler, M., Klackl, J., Richlan, F., Schurz, M., Staffen, W., Ladurner, G., Wimmer, H., 2009. On the functional neuroanatomy of visual word processing: Effects of case and letter deviance. Journal of Cognitive Neuroscience 21, 222-229.

Ku, S-P., Tolias, A.S., Logothetis, N.K., Goense, J. fMRI of the face-processing network in the ventral temporal lobe of awake and anesthetized macaques. Neuron 70, 352-362

Kuo, W.J., Yeh, T.C., Lee, C.Y., Wu, Y.T., Chou, C.C., Ho, L.T., Hung, D.L., Tzeng, O.J.L., Hsieh, J.C., 2003. Frequency effects of Chinese character processing in the brain: An event-related fMRI study. NeuroImage 18, 720-730.

Kveraga, K., Boshyan, J., Bar, M., 2007. Magnocellular projections as the trigger of top-down facilitation in recognition. Journal of Neuroscience 27, 13232-13240.

Larsell, O., 1951. Anatomy of the nervous system. Appleton-Century-Crofts, New York.

Law, I., Kannao, I., Fujita, H., Lassen, N.A., Miura, S., Uemura, K., 1991. Left supramarginal/angular gyri activation during reading of syllabograms in the Japanese language. Journal of Neurolinguistics 6, 243-251.

Lee, C.Y., Tsai, J.L., Kuo, W.J., Yeh, T.C., Wu, Y.T., Ho, L.T., Hung, D.L., Tzeng, O.J.L., Hsieh, J.C., 2004. Neuronal correlates of consistency and frequency effects on Chinese character naming: An event-related fMRI study. NeuroImage 23, 1235-1245.

Leff, A.P., Crewes, H., Plant, G.T., Scott, S.K., Kennard, C., Wise, R.J.S., 2001. The functional anatomy of single-word reading in patients with hemianopic and pure alexia. Brain 124, 510-521.

Lin, P., Hasson, U., Jovicich, J., Robinson, S., 2010. A Neuronal Basis for Task-Negative Responses in the Human Brain. Cereb Cortex doi:10.1093/cercor/bhq151.

Liu, Y., Dunlap, S., Fiez, J., Perfetti, C., 2007. Evidence for neural accommodation to a writing system following learning. Hum Brain Mapp 28, 1223-1234.

Logothetis, N.K., 2002. The neural basis of the blood-oxygen-level-dependent functional magnetic resonance imaging signal. Philosophical Transactions of the Royal Society B: Biological Sciences 357, 1003-1037.

Logothetis, N.K., 2003. The underpinnings of the BOLD functional magnetic resonance imaging signal. Journal of Neuroscience 23, 3963-3971.

Logothetis, N.K., 2008. What we can do and what we cannot do with fMRI. Nature 453, 869-878.

Logothetis, N.K., Pauls, J., Augath, M., Trinath, T., Oeltermann, A., 2001. Neurophysiological investigation of the basis of the fMRI signal. Nature 412, 150-157.

Ludwig, E., Klinger, J., 1956. Atlas Cerebri Humani. Little Brown & Co., Boston.

Lukatela, G., Turvey, M.T., 1994. Visual Lexical Access Is Initially Phonological: 2. Evidence From Phonological Priming by Homophones and Pseudohomophones. Journal of Experimental Psychology: General 123, 331-353.

Mandonnet, E., Gatignol, P., Duffau, H., 2009. Evidence for an occipito-temporal tract underlying visual recognition in picture naming. Clinical Neurology and Neurosurgery 111, 601-605.

Mandonnet, E., Nouet, A., Gatignol, P., Capelle, L., Duffau, H., 2007. Does the left inferior longitudinal fasciculus play a role in language? A brain stimulation study. Brain 130, 623-629.

Marshall, J.C., Newcombe, F., 1973. Patterns of paralexia: A psycholinguistic approach. Journal of psycholinguistic research 2, 175-199.

Martino, J., Brogna, C., Robles, S.G., Vergani, F., Duffau, H., 2010a. Anatomic dissection of the inferior fronto-occipital fasciculus revisited in the lights of brain stimulation data. Cortex 46, 691-699.

Martino, J., Da Silva-Freitas, R., Caballero, H., Marco De Lucas, E., García-Porrero, J.A., Vázquez-Barquero, A., 2013a. Fiber dissection and diffusion tensor imaging tractography study of the temporoparietal fiber intersection area. Neurosurgery 72, ons87-ons97.

Martino, J., De Witt Hamer, P.C., Berger, M.S., Lawton, M.T., Arnold, C.M., De Lucas, E.M., Duffau, H., 2013b. Analysis of the subcomponents and cortical terminations of the perisylvian superior longitudinal fasciculus: A fiber dissection and DTI tractography study. Brain Structure and Function 218, 105-121.

Martino, J., de Witt Hamer, P.C., Vergani, F., Brogna, C., de Lucas, E.M., Vázquez-Barquero, A., García-Porrero, J.A., Duffau, H., 2011. Cortex-sparing fiber dissection: An improved method for the study of white matter anatomy in the human brain. Journal of Anatomy 219, 531-541.

Martino, J., Vergani, F., Robles, S.G., Duffau, H., 2010b. New insights into the anatomic dissection of the temporal stem with special emphasis on the inferior fronto-occipital fasciculus: implications in surgical approach to left mesiotemporal and temporoinsular structures. Neurosurgery 66, 4-12.

Mayall, K., Humphreys, G.W., Mechelli, A., Olson, A., Price, C.J., 2001. The effects of case mixing on word recognition: Evidence from a PET study. Journal of Cognitive Neuroscience 13, 844-853.

Mazoyer, B., Zago, L., Mellet, E., Bricogne, S., Etard, O., Houdé, O., Crivello, F., Joliot, M., Petit, L., Tzourio-Mazoyer, N., 2001. Cortical networks for working memory and executive functions sustain the conscious resting state in man. Brain Research Bulletin 54, 287-298.

McCandliss, B.D., Cohen, L., Dehaene, S., 2003. The visual word form area: Expertise for reading in the fusiform gyrus. Trends in Cognitive Sciences 7, 293-299.

McCann, R.S., Besner, D., Davelaar, E., 1988. Word Recognition and Identification: Do Word-Frequency Effects Reflect Lexical Access? Journal of Experimental Psychology: Human Perception and Performance 14, 693-706.

McClelland, J.L., Rumelhart, D.E., 1981. An interactive activation model of context effects in letter perception: I. An account of basic findings. Psychological Review 88, 375-407.

McDermott, K.B., Petersen, S.E., Watson, J.M., Ojemann, J.G., 2003. A procedure for identifying regions preferentially activated by attention to semantic and phonological relations using functional magnetic resonance imaging. Neuropsychologia 41, 293-303.

Mechelli, A., Crinion, J.T., Long, S., Friston, K.J., Lambon Ralph, M.A., Patterson, K., McClelland, J.L., Price, C.J., 2005. Dissociating reading processes on the basis of neuronal interactions. Journal of Cognitive Neuroscience 17, 1753-1765.

Mechelli, A., Gorno-Tempini, M.L., Price, C.J., 2003. Neuroimaging studies of word and pseudoword reading: consistencies, inconsistencies, and limitations. Journal of Cognitive Neuroscience 15, 260-271.

Mechelli, A., Humphreys, G.W., Mayall, K., Olson, A., Price, C.J., 2000. Differential effects of word length and visual contrast in the fusiform and lingual gyri during reading. Proceedings of the Royal Society B: Biological Sciences 267, 1909-1913.

Mechelli, A., Josephs, O., Lambon Ralph, M.A., McClelland, J.L., Price, C.J., 2007. Dissociating stimulus-driven semantic and phonological effect during reading and naming. Human Brain Mapping 28, 205-217.

Mechelli, A., Price, C.J., Friston, K.J., Ishai, A., 2004. Where bottom-up meets top-down: Neuronal interactions during perception and imagery. Cerebral Cortex 14, 1256-1265.

Mei, L., Xue, G., Lu, Z.L., He, Q., Zhang, M., Xue, F., Chen, C., Dong, Q., 2013. Orthographic transparency modulates the functional asymmetry in the fusiform cortex: An artificial language training study. Brain and Language 125, 165-172.

Menon, V., Desmond, J., 2001. Left superior parietal cortex involvement in writing: integrating fMRI with lesion evidence. Cognitive brain research 12, 337-340.

Mercure, E., Dick, F., Johnson, M.H., 2008. Featural and configural face processing differentially modulate ERP components. Brain Research 1239, 162-170.

Merigan, W.H., Maunsell, J.H.R., 1993. How Parallel are the Primate Visual Pathways? Annual Review of Neuroscience 16, 369-402.

Miklossy, J., Clarke, S., Van der Loos, H., 1991. The long distance effects of brain lesions: Visualization of axonal pathways and their terminations in the human brain by the Nauta method. Journal of Neuropathology and Experimental Neurology 50, 595-614.

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

Mitzdorf, U., 1985. Current source-density method and application in cat cerebral cortex: Investigation of evoked potentials and EEG phenomena. Physiological Reviews 65, 37-100.

Mitzdorf, U., 1987. Properties of the evoked potential generators: current source-density analysis of visually evoked potentials in the cat cortex. International Journal of Neuroscience 33, 33-59.

Moore, C.J., Price, C.J., 1999. Three distinct ventral occipitotemporal regions for reading and object naming. NeuroImage 10, 181-192.

Morton, J., 1969. Interaction of information in word recognition. Psychological Review 76, 165-178.

Mummery, C., Patterson, K., Hodges, J., Price, C., 1998. Functional neuroanatomy of the semantic system: divisible by what? Journal of Cognitive Neuroscience 10, 766-777.

Mummery, C.J., Shallice, T., Price, C.J., 1999. Dual-process model in semantic priming: A functional imaging perspective. NeuroImage 9, 516-525.

Murphy, K., Garavan, H., 2004. Artifactual fMRI group and condition differences driven by performance confounds. NeuroImage 21, 219-228.

Murtha, S., Chertkow, H., Beauregard, M., Evans, A., 1999. The neural substrate of picture naming. Journal of Cognitive Neuroscience 11, 399-423.

Naidich, T.P., Valavanis, A.G., Kubik, S., Penn, R.D., Roberts, D.W., 1995. Anatomic relationships along the low-middle convexity: Part I - Normal specimens and magnetic resonance imaging. Neurosurgery 36, 517-532.

Nair, D.G., 2005. About being BOLD. Brain Research Reviews 50, 229-243.

Nakamura, K., Dehaene, S., Jobert, A., Le Bihan, D., Kouider, S., 2005a. Subliminal convergence of Kanji and Kana words: Further evidence for functional parcellation of the posterior temporal cortex in visual word perception. Journal of Cognitive Neuroscience 17, 954-968.

Nakamura, K., Dehaene, S., Jobert, A., Le Bihan, D., Kouider, S., 2005b. Subliminal convergence of Kanji and Kana words: further evidence for functional parcellation of the posterior temporal cortex in visual word perception. J Cogn Neurosci 17, 954-968.

Nakamura, K., Honda, M., Hirano, S., Oga, T., Sawamoto, N., Hanakawa, T., Inoue, H., Ito, J., Matsuda, T., Fukuyama, H., Shibasaki, H., 2002. Modulation of the visual word retrieval system in writing: A functional MRI study on the Japanese orthographies. Journal of Cognitive Neuroscience 14, 104-115.

Nakamura, K., Oga, T., Okada, T., Sadato, N., Takayama, Y., Wydell, T., Yonekura, Y., Fukuyama, H., 2005c. Hemispheric asymmetry emerges at distinct parts of the occipitotemporal cortex for objects, logograms and phonograms: A functional MRI study. NeuroImage 28, 521-528.

Nazir, T.A., Ben-Boutayab, N., Decoppet, N., Deutsch, A., Frost, R., 2004. Reading habits, perceptual learning, and recognition of printed words. Brain and Language 88, 294-311.

Nobre, A.C., Allison, T., McCarthy, G., 1994. Word recognition in the human inferior temporal lobe. Nature 372, 260-263.

Norris, D., McQueen, J.M., Cutler, A., 2000. Merging information in speech recognition: Feedback is never necessary. Behavioral and Brain Sciences 23.

Ogawa, S., Tank, D.W., Menon, R., Ellermann, J.M., Kim, S.G., Merkle, H., Ugurbil, K., 1992. Intrinsic signal changes accompanying sensory stimulation: Functional brain mapping with magnetic resonance imaging. Proceedings of the National Academy of Sciences of the United States of America 89, 5951-5955.

Oishi, K., Zilles, K., Amunts, K., Faria, A., Jiang, H., Li, X., Akhter, K., Hua, K., Woods, R., Toga, A.W., Pike, G.B., Rosa-Neto, P., Evans, A., Zhang, J., Huang, H., Miller, M.I., van Zijl, P.C.M., Mazziotta, J., Mori, S., 2008. Human brain white matter atlas: Identification and assignment of common anatomical structures in superficial white matter. NeuroImage 43, 447-457.

Oldfield, R.C., 1971. The assessment and analysis of handedness: The Edinburgh inventory. Neuropsychologia 9, 97-113.

Oztas, E., 2003. Neuronal tracing. Neuroanatomy 2, 2-5.

Paap, K.R., Newsome, S.L., McDonald, J.E., Schvaneveldt, R.W., 1982. An activation-verification model for letter and word recognition: The word-superiority effect. Psychological Review 89, 573.

Parker, G.J.M., Alexander, D.C., 2005. Probabilistic anatomical connectivity derived from the microscopic persistent angular structure of cerebral tissue. Philosophical Transactions of the Royal Society B: Biological Sciences 360, 893-902.

Passingham, R.E., Stephan, K.E., Kötter, R., 2002. The anatomical basis of functional localization in the cortex. Nature Reviews Neuroscience 3, 606-616.

Pauling, L., Coryell, C.D., 1936. The Magnetic Properties and Structure of Hemoglobin, Oxyhemoglobin and Carbonmonoxyhemoglobin. Proc Natl Acad Sci U S A 22, 210-216.

Perry, C., Ziegler, J.C., Zorzi, M., 2007. Nested incremental modeling in the development of computational theories: The CDP+ model of reading aloud. Psychological Review 114, 273-315.

Petersen, S.E., Fox, P.T., Posner, M.I., Mintun, M., Raichle, M.E., 1988. Positron emission tomographic studies of the cortical anatomy of single-word processing. Nature 331, 585-589.

Petr, R., Holden, L.B., Jirout, J., 1949. The Efferent Intercortical Connections of Teh Superficial Corten of the Temporal Lobe (Macaca Mulatta)*. Journal of Neuropathology & Experimental Neurology 8, 100.

Petrides, M., Pandya, D.N., 2002. Comparative cytoarchitectonic analysis of the human and the macaque ventrolateral prefrontal cortex and corticocortical connection patterns in the monkey. European Journal of Neuroscience 16, 291-310.

Philipose, L.E., Gottesman, R.F., Newhart, M., Kleinman, J.T., Herskovits, E.H., Pawlak, M.A., Marsh, E.B., Davis, C., Heidler-Gary, J., Hillis, A.E., 2007. Neural regions essential for reading and spelling of words and pseudowords. Annals of neurology 62, 481-492.

Phillips, J.A., Humphreys, G.W., Noppeney, U., Price, C.J., 2002. The neural substrates of action retrieval: An examination of semantic and visual routes to action. Visual Cognition 9, 662-684.

Pinsk, M.A., Arcaro, M., Weiner, K.S., Kalkus, J.F., Inati, S.J., Gross, C.G., Kastner, S., 2009. Neural representations of faces and body parts in macaque and human cortex: A comparative fMRI study. Journal of Neurophysiology 101, 2581-2600.

Plaut, D.C., McClelland, J.L., Seidenberg, M.S., Patterson, K., 1996. Understanding Normal and Impaired Word Reading: Computational Principles in Quasi-Regular Domains. Psychological Review 103, 56-115.

Price, C.J., 2000. The anatomy of language: contributions from functional neuroimaging. Journal of Anatomy 197, 335-359.

Price, C.J., Devlin, J.T., 2003. The myth of the visual word form area. NeuroImage 19, 473-481.

Price, C.J., Devlin, J.T., 2011. The Interactive Account of ventral occipitotemporal contributions to reading. Trends in Cognitive Sciences 15, 246-253.

Price, C.J., Friston, K.J., 1997. Cognitive Conjunction: A New Approach to Brain Activation Experiments. NeuroImage 5, 261-270.

Price, C.J., Friston, K.J., 2002. Degeneracy and cognitive anatomy. Trends in Cognitive Sciences 6, 416-421.

Price, C.J., Friston, K.J., 2005. Functional ontologies for cognition: The systematic definition of structure and function. Cognitive Neuropsychology 22, 262-275.

Price, C.J., McCrory, E., Noppeney, U., Mechelli, A., Moore, C.J., Biggio, N., Devlin, J.T., 2006. How reading differs from object naming at the neuronal level. Neuroimage 29, 643-648.

Price, C.J., Mechelli, A., 2005. Reading and reading disturbance. Current Opinion in Neurobiology 15, 231-238.

Price, C.J., Noppeney, U., Phillips, J., Devlin, J.T., 2003. How is the fusiform gyrus related to category-specificity? Cogn Neuropsychol 20, 561-574.

Price, C.J., Wise, R.J., Frackowiak, R.S., 1996. Demonstrating the implicit processing of visually presented words and pseudowords. Cereb Cortex 6, 62-70.

Price, C.J., Wise, R.J.S., Watson, J.D.G., Patterson, K., Howard, D., Frackowiak, R.S.J., 1994. Brain activity during reading. The effects of exposure duration and task. Brain 117, 1255-1269.

Raichle, M.E., Fiez, J.A., Videen, T.O., MacLeod, A.M.K., Pardo, J.V., Fox, P.T., Petersen, S.E., 1994. Practice-related changes in human brain functional anatomy during nonmotor learning. Cerebral Cortex 4, 8-26.

Raichle, M.E., MacLeod, A.M., Snyder, A.Z., Powers, W.J., Gusnard, D.A., Shulman, G.L., 2001. A default mode of brain function. Proceedings of the National Academy of Sciences of the United States of America 98, 676-682.

Raichle, M.E., Snyder, A.Z., 2007. A default mode of brain function: A brief history of an evolving idea. NeuroImage 37, 1083-1090.

Rao, R.P.N., Ballard, D.H., 1999. Predictive coding in the visual cortex: A functional interpretation of some extra-classical receptive-field effects. Nature Neuroscience 2, 79-87.

Reicher, G.M., 1969. Perceptual recognition as a function of meaningfulness of stimulus material. Journal of Experimental Psychology 81, 275-280.

Reimer, J.F., Lorsbach, T.C., Bleakney, D.M., 2008. Automatic semantic feedback during visual word recognition. Memory and Cognition 36, 641-658.

Richardson, F.M., Seghier, M.L., Leff, A.P., Thomas, M.S.C., Price, C.J., 2011. Multiple routes from occipital to temporal cortices during reading. Journal of Neuroscience 31, 8239-8247.

Riesenhuber, M., Poggio, T., 1999. Hierarchical models of object recognition in cortex. Nature Neuroscience 2, 1019-1025.

Rilling, J.K., 2006. Human and NonHuman primate brains: Are they allometrically scaled versions of the same design? Evolutionary Anthropology 15, 65-77.

Rilling, J.K., Glasser, M.F., Preuss, T.M., Ma, X., Zhao, T., Hu, X., Behrens, T.E.J., 2008. The evolution of the arcuate fasciculus revealed with comparative DTI. Nature Neuroscience 11, 426-428.

Rilling, J.K., Seligman, R.A., 2002. A quantitative morphometric comparative analysis of the primate temporal lobe. Journal of Human Evolution 42, 505-533.

Roberts, D.J., Woollams, A.M., Kim, E., Beeson, P.M., Rapcsak, S.Z., Lambon Ralph, M.A., 2012. Efficient Visual Object and Word Recognition Relies on High Spatial Frequency Coding in the Left Posterior Fusiform Gyrus: Evidence from a Case-Series of Patients with Ventral Occipito-Temporal Cortex Damage. Cerebral Cortex.

Rogers, H., 2005. Writing Systems: A Linguistic approach. Blackwell Publishing, Malden, MA.

Rolls, E.T., 2000. Functions of the primate temporal lobe cortical visual areas in invariant visual object and face recognition. Neuron 27, 205-218.

Roskies, A., Fiez, J., Balota, D., Raichle, M., Petersen, S., 2001. Task-dependent modulation of regions in the left inferior frontal cortex during semantic processing. Journal of Cognitive Neuroscience 13, 829-843.

Rosson, M.B., 1983. From SOFA to LOUCH: lexical contributions to pseudoword pronunciation. Memory and Cognition 11, 152-160.

Rumelhart, D.E., McClelland, J.L., 1982. An interactive activation model of context effects in letter perception: II. The contextual enhancement effect and some tests and extensions of the model. Psychological Review 89, 60-94.

Rumsey, J.M., Horwitz, B., Donohue, B.C., Nace, K., Maisog, J.M., Andreason, P., 1997. Phonological and orthographic components of word recognition. A PET-rCBF study. Brain 120, 739-759.

Rumsey, J.M., Horwitz, B., Donohue, B.C., Nace, K.L., Maisog, J.M., Andreason, P., 1999. A functional lesion in developmental dyslexia: Left angular gyral blood flow predicts severity. Brain and Language 70, 187-204.

Sakurai, Y., 2004. Varieties of alexia from fusiform, posterior inferior temporal and posterior occipital gyrus lesions. Behavioural Neurology 15, 35-50.

Sakurai, Y., Momose, T., Iwata, M., Sudo, Y., Kumakura, Y., Ohtomo, K., Kanazawa, I., 2001. Cortical activation in reading assessed by region of interest-based analysis and statistical parametric mapping. Brain Research Protocols 6, 167-171.

Sakurai, Y., Momose, T., Iwata, M., Sudo, Y., Ohtomo, K., Kanazawa, I., 2000. Different cortical activity in reading of Kanji words, Kana words and Kana nonwords. Cognitive Brain Research 9, 111-115.

Sakurai, Y., Momose, T., Iwata, M., Watanabe, T., Ishikawa, T., Takeda, K., Kanazawa, I., 1992. Kanji word reading process analysed by positron emission tomography. NeuroReport 3, 445-448.

Salmelin, R., Kiesilä, P., Uutela, K., Salonen, O., 1996. Impaired visual word processing in dyslexia revealed with magnetoencephalography. Annals of neurology 40, 157-162.

Sarubbo, S., De Benedictis, A., Maldonado, I.L., Basso, G., Duffau, H., 2011. Frontal terminations for the inferior fronto-occipital fascicle: anatomical dissection, DTI study and functional considerations on a multi-component bundle. Brain Structure and Function, 1-17.

Schmahmann, J.D., Pandya, D., 2009. Fiber Pathways of the Brain. Oxford University Press, USA.

Schrader, S., Gewaltig, M.O., Körner, U., Körner, E., 2009. Cortext: A columnar model of bottom-up and top-down processing in the neocortex. Neural Networks 22, 1055-1070.

Schroeder, C.E., Mehta, A.D., Givre, S.J., 1998. A spatiotemporal profile of visual system activation revealed by current source density analysis in the awake macaque. Cerebral Cortex 8, 575-592.

Seghier, M.L., Lee, H.L., Schofield, T., Ellis, C.L., Price, C.J., 2008. Inter-subject variability in the use of two different neuronal networks for reading aloud familiar words. NeuroImage 42, 1226-1236.

Seghier, M.L., Neufeld, N.H., Zeidman, P., Leff, A.P., Mechelli, A., Nagendran, A., Riddoch, J.M., Humphreys, G.W., Price, C.J., 2012. Reading without the left ventral occipito-temporal cortex. Neuropsychologia 50, 3621-3635.

Seghier, M.L., Price, C.J., 2010. Reading aloud boosts connectivity through the putamen. Cerebral Cortex 20, 570-582.

Seghier, M.L., Price, C.J., 2011. Explaining left lateralization for words in the ventral occipitotemporal cortex. Journal of Neuroscience 31, 14745-14753.

Seghier, M.L., Vuilleumier, P., 2006. Functional neuroimaging findings on the human perception of illusory contours. Neuroscience and Biobehavioral Reviews 30, 595-612.

Seidenberg, M.S., McClelland, J.L., 1989. A Distributed, Developmental Model of Word Recognition and Naming. Psychological Review 96, 523-568.

Seunarine, K., Alexander, D., 2009. Multiple fibers: beyond the diffusion tensor. In: Johansen-Berg, H., Behrens, T. (Eds.), Diffusion MRI: From quantitative measurement to in-vivo neuroanatomy. Academic Press, San Diego.

Shaywitz, B.A., Shaywitz, S.E., Blachman, B.A., Pugh, K.R., Fulbright, R.K., Skudlarski, P., Mencl, W.E., Constable, R.T., Holahan, J.M., Marchione, K.E., Fletcher, J.M., Lyon, G.R., Gore, J.C., 2004. Development of left occipitotemporal systems for skilled reading in children after a phonologically- based intervention. Biol Psychiatry 55, 926-933.

Shaywitz, S.E., Shaywitz, B.A., 2008. Paying attention to reading: The neurobiology of reading and dyslexia. Development and Psychopathology 20, 1329-1349.

Shulman, G.L., Fiez, J.A., Corbetta, M., Buckner, R.L., Miezin, F.M., Raichle, M.E., Petersen, S.E., 1997. Common blood flow changes across visual tasks: II. Decreases in cerebral cortex. Journal of Cognitive Neuroscience 9, 648-663.

Smith, M.C., Besner, D., 2001. Modulating semantic feedback in visual word recognition. Psychonomic Bulletin and Review 8, 111-117.

Smith, S.M., Fox, P.T., Miller, K.L., Glahn, D.C., Fox, P.M., Mackay, C.E., Filippini, N., Watkins, K.E., Toro, R., Laird, A.R., Beckmann, C.F., 2009. Correspondence of the brain's functional architecture during activation and rest. Proceedings of the National Academy of Sciences of the United States of America 106, 13040-13045.

Spitsyna, G., Warren, J.E., Scott, S.K., Turkheimer, F.E., Wise, R.J.S., 2006. Converging language streams in the human temporal lobe. Journal of Neuroscience 26, 7328-7336.

Springer, J.A., Binder, J.R., Hammeke, T.A., Swanson, S.J., Frost, J.A., Bellgowan, P.S., Brewer, C.C., Perry, H.M., Morris, G.L., Mueller, W.M., 1999. Language dominance in neurologically normal and epilepsy subjects A functional MRI study. Brain 122, 2033-2046.

Starrfelt, R., Gerlach, C., 2007. The Visual What For Area: Words and pictures in the left fusiform gyrus. NeuroImage 35, 334-342.

Starrfelt, R., Habekost, T., Leff, A.P., 2009. Too little, too late: Reduced visual span and speed characterize pure alexia. Cerebral Cortex 19, 2880-2890.

Stoeckel, C., Gough, P.M., Watkins, K.E., Devlin, J.T., 2009. Supramarginal gyrus involvement in visual word recognition. Cortex 45, 1091-1096.

Sugihara, G., Kaminaga, T., Sugishita, M., 2006. Interindividual uniformity and variety of the "writing center": a functional MRI study. NeuroImage 32, 1837-1849.

Suzuki, W., Amaral, D., 2003a. Where are the perirhinal and parahippocampal cortices? A historical overview of the nomenclature and boundaries applied to the primate medial temporal lobe. Neuroscience 120, 893-906.

Suzuki, W.A., Amaral, D.G., 2003b. Perirhinal and parahippocampal cortices of the Macaque monkey: Cytoarchitectonic and chemoarchitectonic organization. Journal of Comparative Neurology 463, 67-91.

Suzuki, W.L., Amaral, D.G., 1994. Perirhinal and parahippocampal cortices of the macaque monkey: Cortical afferents. The Journal of Comparative Neurology 350, 497-533.

Szwed, M., Cohen, L., Qiao, E., Dehaene, S., 2009. The role of invariant line junctions in object and visual word recognition. Vision Research 49, 718-725.

Szwed, M., Dehaene, S., Kleinschmidt, A., Eger, E., Valabrègue, R., Amadon, A., Cohen, L., 2011. Specialization for written words over objects in the visual cortex. NeuroImage 56, 330-344.

Tan, L.H., Laird, A.R., Li, K., Fox, P.T., 2005. Neuroanatomical correlates of phonological processing of Chinese characters and alphabetic words: A meta-analysis. Human Brain Mapping 25, 83-91.

Tan, L.H., Liu, H.L., Perfetti, C.A., Spinks, J.A., Fox, P.T., Gao, J.H., 2001. The neural system underlying Chinese logograph reading. NeuroImage 13, 836-846.

Tan, L.H., Spinks, J.A., Feng, C.M., Siok, W.T., Perfetti, C.A., Xiong, J., Fox, P.T., Gao, J.H., 2003. Neural systems of second language reading are shaped by native language. Human Brain Mapping 18, 158-166.

Tan, L.H., Spinks, J.A., Gao, J.H., Liu, H.L., Perfetti, C.A., Xiong, J., Stofer, K.A., Pu, Y., Liu, Y., Fox, P.T., 2000. Brain activation in the processing of Chinese characters and words: A functional MRI study. Human Brain Mapping 10, 16-27.

Tarkiainen, A., Helenius, P., Hansen, P.C., Cornelissen, P.L., Salmelin, R., 1999. Dynamics of letter string perception in the human occipitotemporal cortex. Brain 122, 2119-2131.

Tilney, F., Riley, H.A., 1921. The Form and Functions of the Central Nervous System. Hoeber, New York.

Tomassini, V., Jbabdi, S., Klein, J.C., Behrens, T.E.J., Pozzilli, C., Matthews, P.M., Rushworth, M.F.S., Johansen-Berg, H., 2007. Diffusion-weighted imaging tractography-based parcellation of the human lateral premotor cortex identifies dorsal and ventral subregions with anatomical and functional specializations. Journal of Neuroscience 27, 10259-10269.

Tournier, J.D., Calamante, F., Gadian, D.G., Connelly, A., 2004. Direct estimation of the fiber orientation density function from diffusion-weighted MRI data using spherical deconvolution. NeuroImage 23, 1176-1185.

Tsao, D.Y., Moeller, S., Freiwald, W.A., 2008. Comparing face patch systems in macaques and humans. Proceedings of the National Academy of Sciences of the United States of America 105, 19514-19519.

Tuch, D.S., 2004. Q-ball imaging. Magnetic Resonance in Medicine 52, 1358-1372.

Türe, U., Yaşargil, M.G., Friedman, A.H., Al-Mefty, O., 2000. Fiber dissection technique: Lateral aspect of the brain. Neurosurgery 47, 417-427.

Tusa, R.J., Ungerleider, L.G., 1985. The inferior longitudinal fasciculus: A reexamination in humans and monkeys. Annals of neurology 18, 583-591.

Ulrich, R., Miller, J., 1994. Effects of Truncation on Reaction Time Analysis. Journal of Experimental Psychology: General 123, 34-80.

Ungerleider, L.G., Mishkin, M., 1982. Two cortical visual systems. Analysis of Visual Behavior, 549-586.

van der Mark, S., Bucher, K., Maurer, U., Schulz, E., Brem, S., Buckelmüller, J., Kronbichler, M., Loenneker, T., Klaver, P., Martin, E., Brandeis, D., 2009. Children with dyslexia lack multiple specializations along the visual word-form (VWF) system. NeuroImage 47, 1940-1949.

van der Mark, S., Klaver, P., Bucher, K., Maurer, U., Schulz, E., Brem, S., Martin, E., Brandeis, D., 2011. The left occipitotemporal system in reading: Disruption of focal fMRI connectivity to left inferior frontal and inferior parietal language areas in children with dyslexia. NeuroImage 54, 2426-2436.

Van Essen, D.C., Maunsell, J.H., 1983. Hierarchical organization and functional streams in the visual cortex. Trends in Neurosciences 6, 370-375.

Van Hoesen, G.W., 1982. The parahippocampal gyrus: New observations regarding its cortical connections in the monkey. Trends in Neurosciences 5, 345-350.

van Turennout, M., Bielamowicz, L., Martin, A., 2003. Modulation of neural activity during object naming: effects of time and practice. Cerebral Cortex 13, 381-391.

van Turennout, M., Ellmore, T., Martin, A., 2000. Long-lasting cortical plasticity in the object naming system. Nature Neuroscience 3, 1329-1334.

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

von Bonin, G., Bailey, P., 1947. The Neocortex of Macaca Mulatta. University of Illinois Press.

von Economo, C.F., 1929. The Cytoarchitectonics of the Human Cerebral Cortex. Oxford University Press, London.

Wakana, S., Jiang, H., Nagae-Poetscher, L.M., Van Zijl, P.C.M., Mori, S., 2004. Fiber Tract-based Atlas of Human White Matter Anatomy. Radiology 230, 77-87.

Walton, M.E., Devlin, J.T., Rushworth, M.F., 2004. Interactions between decision making and performance monitoring within prefrontal cortex. Nat Neurosci 7, 1259-1265.

Wang, F., Sun, T., Li, X., Xia, H., Li, Z., 2011a. Microsurgical and tractographic anatomical study of insular and transsylvian transinsular approach. Neurological Sciences 32, 865-874.

Wang, X., Yang, J., Shu, H., Zevin, J.D., 2011b. Left fusiform BOLD responses are inversely related to word-likeness in a one-back task. NeuroImage 55, 1346-1356.

Warrington, E.K., Shallice, T., 1980. Word-form dyslexia. Brain: a journal of neurology 103, 99.

Wedeen, V.J., Hagmann, P., Tseng, W.Y.I., Reese, T.G., Weisskoff, R.M., 2005. Mapping complex tissue architecture with diffusion spectrum magnetic resonance imaging. Magnetic Resonance in Medicine 54, 1377-1386.

Weller, J.A., Levin, I.P., Shiv, B., Bechara, A., 2007. Neural correlates of adaptive decision making for risky gains and losses. Psychological Science 18, 958-964.

Wernicke, C., 1900. Atlas des Gehirns Vol.2. Schletter'sche Buchhandlung.

Wernicke, C., 1903. Atlas des Gehirns Vol.3. Schletter'sche Buchhandlung.

Wheeler, D.D., 1970. Processes in word recognition. Cognitive Psychology 1, 59-85.

Whitlock, D.G., Nauta, W.J.H., 1956. Subcortical projections from the temporal neocortex in Macaca mulatta. The Journal of Comparative Neurology 106, 183-212.

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

Woodhead, Z.V.J., Brownsett, S.L.E., Dhanjal, N.S., Beckmann, C., Wise, R.J.S., 2011a. The visual word form system in context. Journal of Neuroscience 31, 193-199.

Woodhead, Z.V.J., Wise, R.J.S., Sereno, M., Leech, R., 2011b. Dissociation of sensitivity to spatial frequency in word and face preferential areas of the fusiform gyrus. Cerebral Cortex 21, 2307-2312.

Woolrich, M.W., Behrens, T.E.J., Beckmann, C.F., Jenkinson, M., Smith, S.M., 2004. Multilevel linear modelling for FMRI group analysis using Bayesian inference. NeuroImage 21, 1732-1747.

Woolrich, M.W., Ripley, B.D., Brady, M., Smith, S.M., 2001. Temporal autocorrelation in univariate linear modeling of FMRI data. NeuroImage 14, 1370-1386.

Worsley, K.J., Marrett, S., Neelin, P., Vandal, A.C., Friston, K.J., Evans, A.C., 1996. A unified statistical approach for determining significant signals in images of cerebral activation. Human Brain Mapping 4, 58-73.

Wu, C.Y., Ho, M.H.R., Chen, S.H.A., 2012. A meta-analysis of fMRI studies on Chinese orthographic, phonological, and semantic processing. NeuroImage 63, 381-391.

Xu, B., Grafman, J., Gaillard, W.D., Ishii, K., Vega-Bermudez, F., Pietrini, P., Reeves-Tyer, P., DiCamillo, P., Theodore, W., 2001. Conjoint and extended neural networks for the computation of speech codes: the neural basis of selective impairment in reading words and pseudowords. Cerebral Cortex 11, 267-277.

Xue, G., Chen, C., Jin, Z., Dong, Q., 2006. Language experience shapes fusiform activation when processing a logographic artificial language: An fMRI training study. NeuroImage 31, 1315-1326.

Xue, G., Jiang, T., Chen, C., Dong, Q., 2008. Language experience shapes early electrophysiological responses to visual stimuli: The effects of writing system, stimulus length, and presentation duration. NeuroImage 39, 2025-2037.

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, 146-155.

Yoncheva, Y.N., Zevin, J.D., Maurer, U., McCandliss, B.D., 2010. Auditory selective attention to speech modulates activity in the visual word form area. Cerebral Cortex 20, 622-632.

Yoon, H.W., Cho, K.D., Chung, J.Y., Park, H.W., 2005. Neural mechanisms of Korean word reading: A functional magnetic resonance imaging study. Neuroscience Letters 373, 206-211.

Yousry, T.A., Schmid, U.D., Alkadhi, H., Schmidt, D., Peraud, A., Buettner, A., Winkler, P., 1997. Localization of the motor hand area to a knob on the precentral gyrus. A new landmark. Brain 120, 141-157.

Zhang, H., Schneider, T., Wheeler-Kingshott, C.A., Alexander, D.C., 2012. NODDI: Practical in vivo neurite orientation dispersion and density imaging of the human brain. NeuroImage 61, 1000-1016.

Zhang, Y., Brady, M., Smith, S., 2001. Segmentation of brain MR images through a hidden Markov random field model and the expectation-maximization algorithm. IEEE Transactions on Medical Imaging 20, 45-57.

Ziegler, J.C., Goswami, U., 2005. Reading acquisition, developmental dyslexia, and skilled reading across languages: A psycholinguistic grain size theory. Psychological Bulletin 131, 3-29.